

3 Versions ▾

What Does a Protein Surface Add? Ablations Cannot Tell You

*Odysseus Kaloidas*¹ ¹ University of California, Berkeley (berkeley.edu) 29 Aug 2026 (modified: 08 Sept 2026) NeurIPS 2026 Workshop ICBINBBIO Submission ICBINB-BIO, Area Chairs, Reviewers, Authors  Revisions  CC BY 4.0

Keywords: protein inverse folding, molecular surface representations, input provenance, causal identification, ablation analysis.

TL;DR: Surface ablations cannot distinguish design-valid representation from target-derived information; our provenance audit shows a released inverse-folding model transcribing a counterfactual surface code.

Abstract:

When a model's input is *computed* rather than given, the ablation that removes it cannot say why it helped: the channel may add information derived from the prediction target (Route A), or merely re-represent information the model already has (Route B), and both produce the same gap. No precision separates them — that is a property of the estimand, not of the sample. We give a seven-component provenance-aware protocol that does separate them, and run it on released surface-conditioned inverse folding. In two released pipelines the surface chemistry channel is a deterministic function of native residue identity: 89.12% and 90.12% of its values move under fixed-coordinate permutation, against 0.00% for geometry. For the system with released weights, a fixed code-table derangement makes it write the counterfactual protein, raising agreement with that sequence by +38.90 pp over its own pre-intervention baseline — a targeted error pattern that generic distribution shift does not explain. Across paired retraining seeds, a semantics-free random 18-class partition gains far more than scrambling biochemical values over the same classes: the benefit follows target-identifying partition structure, not mainly biochemical meaning. At a matched ≈17% marginal ceiling, target-derived channels add +12.49 pp to frozen ProteinMPNN and +11.07 pp to frozen ESM-IF1 while backbone-derived channels are ignored — holding unrefitted on an external RCSB 2023+ panel across three frozen families. Joint surface instruments remain below our Route-B threshold, while a backbone-derived positive control clears it. Four of our own predeclared predictions were falsified and a fifth claim family was retracted; all five are retained. The claim is a provenance burden: whenever a method computes an input from data involving the prediction target, an ordinary ablation cannot justify a representation claim until source interventions separate added target information from design-valid computation.

Serve As Reviewer:  Odysseus Kaloidas

Submission Number: 65

Filter by reply type...

Filter by author...

Search keywords...

Sort: Newest First

-

=

Everyone

Program Chairs

Submission65 Area...

2 / 2 replies shown

Submission65 Reviewers

Submission65 Authors

Add:

Official Comment



Paper Decision

Decision

 by Program Chairs 29 Sept 2026, 03:44 (modified: 30 Sept 2026, 01:42)

Program Chairs, Area Chairs, Reviewers, Authors

 Revisions

Decision: Accept (Poster)

Comment:

The paper demonstrates a concrete preprocessing leakage failure in protein design and provides evidence that a released model can strongly exploit the leaked signal. However, the manuscript is substantially less clear than the underlying contribution warrants, and a large portion of the presentation is devoted to auxiliary analyses and framing that obscure the central result. We strongly encourage the authors to substantially narrow the scope, sharpen the main claim, and make the manuscript more self-contained and accessible to readers.

Add:

Official Comment



Good paper overall

Official Review

 by Reviewer Tsuz 21 Sept 2026, 20:21 (modified: 30 Sept 2026, 01:41)

Program Chairs, Area Chairs, Reviewers, Authors

 Revisions

Review:

The paper is on par for quality, clarity, originality, and significance. The figures are clean and the results are clear.

I suggest elaborating on the conclusion and connecting it to the broader impacts in the real world. That would help ground it more in the real world. I also suggest including ideas / thoughts for future work. How can the contribution be built on further?

I am not sure if the limitations of the work were clearly stated.

Rating: 5: Accept: Technically solid paper, with high to moderate impact, with good-to-excellent evaluation, resources, reproducibility, and no unaddressed ethical considerations.

Confidence: 4: The reviewer is confident but not absolutely certain that the evaluation is correct

Add:

Official Comment

About OpenReview

FAQ

Sponsors

All Venues

Contact

Hosting a Venue

Terms of Use / Privacy Policy

Donate