Critical Review of "Data-driven Discovery of Biophysical T Cell Receptor Co-specificity Rules"

Edoardo Livolsi — Independent Researcher

Manuscript under review

A.G.T. Pyo et al.,
"Data-driven Discovery of Biophysical T Cell Receptor Co-specificity Rules",
arXiv:2412.13722.

Summary

The manuscript proposes a data-driven framework, based on supervised contrastive learning, to infer so-called "biophysical rules" governing T cell receptor (TCR) co-specificity to peptide-MHC ligands. The authors claim that learned similarity metrics capture generalizable, ligand-independent principles underlying receptor-ligand binding.

While the work is technically competent from a machine-learning standpoint, it suffers from severe conceptual flaws, category errors between statistical inference and physical law, and systematic overinterpretation of fitted correlations as biophysical principles.

Major Conceptual Issues

1. Category Error: Statistics Mistaken for Physics

The central conceptual flaw is the identification of fitted statistical regularities with "biophysical rules".

The manuscript does not derive any rule from first principles, nor from an underlying Hamiltonian, energy functional, or variational principle. Instead, it infers a distance metric optimized to separate labeled and unlabeled sequence pairs. This is a statistical classifier, not a physical law.

Calling such outcomes "biophysical rules" is misleading and incorrect.

2. Absence of a Physical Model

No microscopic or mesoscopic physical model of TCR-pMHC interaction is presented. There is no binding energy, no force law, no conformational dynamics, and no thermodynamic description.

The probability expressions introduced are postulated, not derived. The exponential ansatz linking distance to co-specificity is assumed, then fitted. This reverses the logic of theoretical physics.

3. Pseudolikelihood Framed as Explanation

The manuscript relies heavily on pseudo-likelihood arguments and contrastive loss functions. While these are valid statistical tools, they do not explain why certain sequence features matter, only that they correlate with labels in the dataset.

The distinction between explanation and interpolation is repeatedly blurred.

Interpretational Overreach

4. Overstated Generalization Claims

Claims of generalization to unseen ligands are not supported by a structural argument. Generalization is empirically demonstrated within a restricted dataset and does not imply universality.

No invariance principle, conservation law, or scaling argument is provided to justify why the learned metric should apply beyond the sampled regime.

5. Misuse of Statistical Physics Language

References to spin glasses, landscapes, and order parameters are metaphorical. No explicit mapping to a known statistical mechanical model is constructed.

Such language creates an appearance of theoretical depth without delivering one.

Technical Limitations

6. Dependence on Data Curation

The inferred rules are sensitive to dataset composition, labeling quality, and sampling bias. This dependence is acknowledged only marginally and undermines claims of robustness.

7. Lack of Falsifiability

The framework lacks clear falsifiable predictions. Failure modes of the learned metric are not defined. As a consequence, the approach cannot be meaningfully tested against independent physical constraints.

Recommendation

This manuscript does not meet the standards for publication as a contribution to theoretical or biophysical understanding.

While the computational methodology is sound, the work:

- Introduces no new physical principle,
- Confuses statistical correlation with causation,

- Overstates the significance of machine-learned metrics,
- Relies on metaphor rather than theory.

I therefore recommend **rejection** in its current form. At most, the work could be considered as a methodological machine-learning study, provided all claims of discovering "biophysical rules" are removed.

Final Assessment

This is not a theory of TCR specificity, but a data-driven classifier dressed in the language of physics. The distinction is essential and currently violated throughout the manuscript.

Reviewer:
Edoardo Livolsi
Independent Researcher

Appendix A: Mathematical Non-Equivalence with the Universal Psi Equation

A.1 Definition of the Competing Constructions

The reviewed manuscript defines a sequence-dependent distance $d(\sigma, \sigma')$ as the solution of a data-dependent optimization problem. Explicitly, a loss functional $L[d; \mathcal{D}]$ is minimized over a finite empirical dataset $\mathcal{D}$:

$$\frac{\delta L}{\delta d} = 0.$$

All properties of the resulting metric depend explicitly on the choice of $\mathcal{D}$.

By contrast, the Universal Psi framework is defined by a single variational equation, fixed a priori and independent of any dataset.

A.2 The Universal Psi Equation (Original Form)

The Universal Psi Equation is given by the stationarity condition

$$\frac{\delta}{\delta \Psi^\dagger(x)} \left\{ \left| \sum_{i,j} \Psi_i^\dagger \Gamma \Psi_j \right|^2 - \lambda_1 \sum_i \left(\Psi_i^\dagger \Psi_i - 1 \right)^2 - \lambda_2 \sum_{i,j,k} \mathrm{Tr} \left[\Gamma \left(\Psi_i \Psi_j^\dagger - \Psi_k \Psi_i^\dagger \right) \right]^2 \right\} = 0.$$

This equation is taken as exact and complete. No term may be removed, reduced, or reparameterized.

A.3 Global Constraint Structure

Let Ψ^* be any admissible configuration satisfying the Universal Psi Equation.

By definition, Ψ^* is not obtained by minimizing an external loss function, but by satisfying simultaneously all constraints encoded in the functional.

Hence, for any admissible variation $\delta\Psi$, one has

$$\delta \left\{ \left| \sum_{i,j} \Psi_i^\dagger \Gamma \Psi_j \right|^2 - \lambda_1 \sum_i \left(\Psi_i^\dagger \Psi_i - 1 \right)^2 - \lambda_2 \sum_{i,j,k} \mathrm{Tr} \left[\Gamma \left(\Psi_i \Psi_j^\dagger - \Psi_k \Psi_i^\dagger \right) \right]^2 \right\} = 0.$$

This condition enforces correlations between all components Ψ_i through a single global constraint.

A.4 Constraint-Induced Correlations

Let $O_a[\Psi]$ and $O_b[\Psi]$ be two observables defined on the space of fields.

Since both are evaluated on the same stationary configuration Ψ^*, their variations satisfy

$$\delta O_a = 0 \qquad \delta O_b = 0$$

under all admissible variations.

Therefore, any nontrivial correlation between O_a and O_b arises necessarily from the shared variational constraint and not from statistical fitting or empirical optimization.

A.5 Dataset Dependence of the Data-Driven Metric

The learned metric d satisfies only

$$\frac{\delta L}{\delta d} = 0,$$

where L depends explicitly on the dataset $\mathcal{D}$.

Given two datasets $\mathcal{D}_1$ and $\mathcal{D}_2$, one generally obtains

$$d(\mathcal{D}_1) \neq d(\mathcal{D}_2).$$

Hence the data-driven construction lacks closure: its defining equation changes with the data.

A.6 Proof of Non-Existence of a Mapping

Assume that a mapping exists between the two constructions, such that

$$d(\sigma, \sigma') = \Phi(\Psi(\sigma), \Psi(\sigma'))\,.$$

Then stationarity of the Universal Psi Equation would imply

$$\frac{\delta d}{\delta \Psi^\dagger} = 0$$

as a necessary condition.

No such condition follows from the data-driven optimization $\delta L/\delta d = 0$. Therefore such a mapping cannot exist.

A.7 Falsifiability

The Universal Psi framework is falsified by the existence of a field configuration Ψ for which

$$\frac{\delta}{\delta \Psi^\dagger(x)} \{\cdots\} \neq 0.$$

The data-driven framework admits no analogous falsification criterion, since any failure can be absorbed by redefining L or retraining d.

A.8 Mathematical Conclusion

The two constructions are governed by fundamentally different equations:

$$\frac{\delta}{\delta \Psi^\dagger(x)} \{\cdots\} = 0 \qquad \text{versus} \qquad \frac{\delta L}{\delta d} = 0.$$

These equations are not equivalent and cannot be transformed into one another.

The first defines a closed variational law. The second defines a dataset-dependent statistical estimator.

This completes the proof.