AP-018

The Hidden State Isomorphism

Cell Signaling and Variational Inference as Instances of the Same Formal System

Laboratorios Alexandria

Alexandria Paper Series | June 2026

contact@laboratoriosalexandria.com

Origin: THESIS-20260626-1002 | Epistemic Grade A | Confidence 0.85

Domains: Life Sciences × Computation & AI

Abstract

A multi-agent deliberation system designed to discover cross-domain isomorphisms identified a formal structural correspondence between cell signaling dynamics and variational inference. We argue that these two domains are not merely analogous — they are instances of the same mathematical object: a dynamical system with hidden states, noisy observations, and a well-defined inference problem over the posterior distribution of those states. This isomorphism explains an otherwise puzzling empirical observation: why variational autoencoders (VAEs) work remarkably well for modeling biological systems. We show that the correspondence is not an engineering choice but a structural necessity arising from the shared formal architecture. We derive testable predictions — biological analogues of well-characterized VAE failure modes including mode collapse, posterior collapse, and KL vanishing — and locate the isomorphism within the broader context of the Free Energy Principle. The thesis was produced by constitutional multi-agent deliberation without human intervention, graded A with confidence 0.85, and subsequently verified through literature review.

1. Introduction

Variational autoencoders work remarkably well for biological data. VAEs whose latent spaces map to gene regulatory pathways achieve interpretable representations that outperform architectures with unconstrained latent spaces (Seninge et al., 2021). Physics-informed VAEs can solve stochastic differential equations governing biological kinetics (Gao et al., 2022). VAEs trained on single-cell perturbation data predict gene expression responses to unseen transcription factor knockouts (Yang et al., 2024). And VAEs whose decoders mirror the hierarchical structure of cellular ontologies generalize across perturbation types (Lotfollahi et al., 2023).

The standard interpretation is that VAEs are a powerful tool applied to a complex domain. But this framing leaves an important question unanswered: *why* do VAEs work so well for biology, specifically? Why does a generative architecture originally designed for image synthesis achieve its most interpretable results when its latent structure mirrors biological organization?

We propose that the answer is structural. VAEs do not merely model biological systems — they are, in a precise formal sense, the same kind of system. Both cell signaling and variational inference are instances of a single mathematical object: a dynamical system with hidden internal states that must be inferred from noisy, partial observations.

This thesis emerged from the deliberation infrastructure of Laboratorios Alexandria, a multi-agent epistemic system governed by constitutional constraints (Laboratorios Alexandria, 2026a). The system pairs adversarial analysts, dialecticians, and epistemic judges in structured deliberation sessions to identify cross-domain structural correspondences. From approximately 5,000 deliberated sessions producing 134 grade-A theses, the system identified this isomorphism by crossing domains — life sciences and computational intelligence — that human researchers rarely traverse simultaneously.

The contribution of this paper is not the application of VAEs to biology (a well-established practice) nor the philosophical claim that biological systems perform inference (articulated by the Free Energy Principle). It is the explicit, formal characterization of the structural isomorphism between the two systems at the level of their mathematical components — and the derivation of testable predictions that follow from it.

2. The Two Systems

2.1 Cell Signaling as a Hidden State System

A cell in its microenvironment processes extracellular signals (ligands, cytokines, mechanical forces) through intracellular cascades involving kinase phosphorylation, second messengers, and transcription factor activation. The cell's internal state — the vector of concentrations of signaling proteins, their phosphorylation states, the configuration of chromatin — is not directly observable from outside. What is measurable are downstream readouts: gene expression profiles (RNA-seq), protein abundances (proteomics), or morphological features (imaging).

This system is naturally described by a stochastic differential equation:

$$dX_m = \mu(X_m)dt + \sigma(X_m)dW_m$$

where X_m is the vector of internal signaling states, $\mu(X_m)$ encodes the deterministic biochemical kinetics (Michaelis-Menten, Hill functions, mass action), $\sigma(X_m)$ captures the stochasticity intrinsic to molecular interactions (thermal noise, copy number fluctuations), and W_m is a standard Wiener process.

Observations are made through a measurement function $Y = g(X_m) + \eta$, where g maps internal states to observables and η represents measurement noise (Kim et al., 2022; Gupta et al., 2022).

2.2 Variational Inference as a Hidden State System

A variational autoencoder posits the existence of latent variables z that generate observed data x through a stochastic process: $z \sim p(z)$, $x \sim p\theta(x|z)$. Because the true posterior $p(z|x)$ is intractable, the VAE introduces an approximate posterior $q\varphi(z|x)$ and optimizes the evidence lower bound (ELBO):

$$\mathcal{L}(\theta, \varphi; x) = E[\log p\theta(x|z)] - D_{kl}(q\varphi(z|x) \,||\, p(z))$$

The first term is the reconstruction fidelity — how well the decoder recovers observations from inferred latent states. The second term is the complexity cost — how far the approximate posterior departs from the prior (Kingma & Welling, 2014).

3. The Isomorphism

These two systems share more than superficial resemblance. They are both instances of a single formal object, which we define as a hidden state inference system (HSIS):

Definition. A hidden state inference system is a tuple $(S, O, f, g, \varepsilon, \eta)$ where S is a space of hidden states, O is a space of observations, $f: S \times \varepsilon \to S$ is a state transition function with noise ε, $g: S \times \eta \to O$ is an observation function with noise η, and the inference problem is to recover $P(S|O)$.

The following table establishes the component-level correspondence:

HSIS Component	Cell Signaling	Variational Autoencoder
Hidden state S	Protein concentrations, phosphorylation states	Latent variable z
Observation O	RNA-seq, proteomics, imaging	Observed data x
Transition f	Biochemical kinetics (μ, σ)	Prior $p(z)$ / dynamics
Observation map g	Measurement pipeline	Decoder $p\theta(x|z)$
Process noise ε	Molecular stochasticity	Sampling noise in z
Measurement noise η	Technical variability, dropout	Reconstruction noise
Inference target	P(signaling state | readouts)	$q\varphi(z|x) \approx p(z|x)$

The isomorphism holds at the level of mathematical structure. Both systems define the same inference problem: given noisy, incomplete observations, recover a posterior distribution over hidden states. This is not a metaphor. A VAE whose latent nodes correspond to biological pathways (as in VEGA; Seninge et al., 2021) is not "borrowing" structure from biology — it is re-implementing the same inference architecture that the biological system instantiates in biochemistry.

3.1 The Observation Function as the Structural Bridge

The key to understanding why VAEs work so well for biological data lies in the observation function g. In both systems, g is a lossy, noisy mapping from a lower-dimensional hidden state to a higher-dimensional observation space. When a biologically-informed VAE constrains its decoder to mirror the structure of known biological pathways, it is not imposing an external constraint — it is aligning the computational observation function with the biological one. The alignment works because both functions serve the same formal role in the same formal system.

4. Why VAEs Work for Biology

The isomorphism provides a principled explanation for a body of empirical results. VEGA (Seninge et al., 2021) achieves interpretable latent spaces by constraining VAE latent nodes to correspond to gene regulatory pathways. Under the isomorphism, this works because the VAE's latent structure is being aligned with the biological system's actual hidden states. PI-VAE (Gao et al., 2022) solves biological SDEs by incorporating physical constraints into the VAE, succeeding because the SDE constraints describe the transition function f in the biological HSIS. Biologically-informed VAEs (Lotfollahi et al., 2023) predict responses to genetic and pharmacological perturbations by structuring the decoder to reflect cellular ontologies. In each case, the success is not coincidental but consequential: the better the VAE's architecture mirrors the HSIS structure of the biological system, the better it performs.

5. Predictions: Biological Analogues of VAE Failure Modes

If the isomorphism holds, then pathological behaviors in one domain should have identifiable analogues in the other. VAEs exhibit three well-characterized failure modes: mode collapse, posterior collapse, and KL vanishing. We predict that each has a biological counterpart.

5.1 Mode Collapse → Terminal Differentiation and Senescence

In VAEs, mode collapse occurs when the decoder maps all latent codes to a restricted subset of the output space, losing the ability to generate diverse outputs. The biological analogue is the irreversible restriction of cellular state space during terminal differentiation or senescence. Bistable switches in cell signaling — well-documented in MAPK cascades (Ferrell & Xiong, 2001), cell cycle control (Sha et al., 2003), and adipocyte differentiation (Park et al., 2012) — drive cells into stable attractors from which they cannot return. A senescent cell has, through constitutive MYC degradation, irreversibly exited the cell cycle (Afifi et al., 2023). The formal correspondence is precise: in both systems, a loss of representational diversity in the observation function g reduces the effective dimensionality of the state-to-output mapping.

Testable prediction: The statistical signatures of mode collapse (reduced variance, clustering of representations) should be quantitatively detectable in the gene expression profiles of terminally differentiated cells compared to progenitor populations, measured by decreased participation ratio of principal components during differentiation.

5.2 Posterior Collapse → Receptor Desensitization and Immune Tolerance

In VAEs, posterior collapse occurs when the encoder ignores the input data, producing an approximate posterior identical to the prior regardless of the observation (He et al., 2019). The biological analogue is receptor desensitization: sustained or repeated ligand exposure causes cells to become unresponsive via

receptor phosphorylation, uncoupling from downstream effectors, and internalization (Kelly et al., 2008). The cell's inference system stops updating its internal state in response to external signals. Immune tolerance provides a system-level example: T cells exposed to persistent antigen become anergic.

Testable prediction: The mutual information between extracellular ligand concentration and downstream pathway activation should approach zero during desensitization, following a trajectory quantitatively similar to the mutual information decay between VAE input and latent code during posterior collapse.

5.3 KL Vanishing → Metabolic Quiescence

KL vanishing occurs when the KL divergence term in the ELBO drives the approximate posterior to match the prior too closely, sacrificing information about individual data points for global simplicity. The biological analogue may be metabolic quiescence — the state in which cells minimize their departure from a baseline metabolic state, sacrificing responsiveness for energetic efficiency. Parr, Da Costa, and Friston (2020) demonstrated that for any Markov-blanketed system, the KL divergence has a direct thermodynamic interpretation as non-equilibrium free energy dissipation. Cellular dissipation constraints indeed limit metabolic versatility (Cossetto et al., 2025).

Testable prediction: The energetic cost of maintaining non-equilibrium signaling states (measurable as ATP consumption or entropy production rate) should correlate with the KL divergence between the cell's signaling profile and its quiescent baseline, and this correlation should break down in quiescent cells analogously to KL vanishing in VAE training.

6. Connection to the Free Energy Principle

The isomorphism we identify is deeply connected to, but distinct from, Karl Friston's Free Energy Principle (FEP). The FEP proposes that any self-organizing system maintaining itself far from thermodynamic equilibrium can be described as minimizing variational free energy (Friston, 2010). Friston's "Life as we know it" (2013) applied this framework to single cells, showing that the cell membrane functions as a Markov blanket separating internal from external states. Kuchling, Friston, Georgiev, and Levin (2020) extended this to morphogenesis, treating cells as Bayesian inference agents.

Our contribution is complementary. The FEP provides the *why*: living systems must minimize variational free energy to persist. We provide the *what*: the explicit structural mapping between the formal components of cell signaling and variational inference, at a level of specificity that generates testable quantitative predictions. Where the FEP tells us that cells should behave as if performing variational inference, our isomorphism tells us which biochemical components implement which inference operations, and predicts that specific computational pathologies should manifest as specific biological phenomena.

7. Implications

7.1 For Machine Learning

If VAEs and biological systems are instances of the same formal object, then biological evolution has had billions of years to solve engineering problems that ML researchers are encountering for the first time. Biological systems avoid mode collapse through multiple parallel pathways with overlapping but distinct activation profiles; they manage posterior collapse through receptor trafficking, dynamically

adjusting encoder capacity based on signal history. These architectural solutions may suggest innovations for VAE design.

7.2 For Cell Biology

The extensive mathematical theory of VAEs provides a formal toolkit for cell biology. The conditions under which posterior collapse occurs are well-characterized (He et al., 2019): it occurs when the decoder is powerful enough to reconstruct observations without the latent variable. Translated to biology: receptor desensitization should be more likely when cells have alternative pathways for maintaining homeostasis without processing the desensitized signal. This is consistent with observations that agonists which do not cause receptor internalization have higher propensities to develop tolerance (Kelly et al., 2008).

7.3 For Discovery Methodology

This isomorphism was identified by a constitutional multi-agent deliberation system processing documents from both domains simultaneously — not by a human researcher working at the intersection. The connection was invisible from within either discipline because cell biologists and ML researchers use different notation for the same mathematical objects. A system designed to cross those boundaries identified the connection by treating formal structure, rather than domain vocabulary, as the primary unit of analysis (Laboratorios Alexandria, 2026b).

8. Limitations

The isomorphism preserves some structure while discarding other. Several disanalogies must be acknowledged. Biological signaling operates across multiple spatial and temporal scales (millisecond ion channel dynamics to day-timescale epigenetic modification); VAEs typically operate at a single scale. Biological systems are not passive inference engines — they act on their environment through active inference (Friston, 2010). The “training” of biological inference occurs through evolution and development, not gradient descent. And biological systems are maintained far from thermodynamic equilibrium by continuous energy dissipation, a cost with no direct computational analogue. The isomorphism is structural, not complete: it holds at the level of the inference problem and its solution structure, not at the level of all implementation details.

9. Falsifiability

The isomorphism would be refuted if: (1) a cell signaling system were shown to be fundamentally incompatible with a hidden state model — if some aspect of signaling has no stochastic dynamical representation; (2) the specific quantitative predictions in Section 5 are empirically falsified; or (3) the operators of inference in both domains are shown to not preserve the same algebraic structure.

10. Conclusion

We have identified and formalized a structural isomorphism between cell signaling and variational inference. Both are instances of a hidden state inference system: a dynamical system with unobservable internal states that must be inferred from noisy observations. This isomorphism explains why VAEs achieve their most interpretable and effective results when applied to biological data — not because VAEs are a good tool for biology, but because both are the same formal object implemented in different substrates.

The isomorphism generates quantitative, testable predictions. Mode collapse should manifest as terminal differentiation. Posterior collapse should manifest as receptor desensitization. KL vanishing should manifest as metabolic quiescence. These predictions bridge two well-developed mathematical literatures, creating opportunities for cross-pollination in both directions.

The discovery itself is a case study in cross-domain epistemic infrastructure. The isomorphism was identified by a multi-agent deliberation system designed to find structural correspondences across disciplinary boundaries — the kind of connection that is obvious in retrospect but invisible from within either discipline alone.

References

Afifi, M. M., Crncec, A., Cornwell, J. A., Cataisson, C., Paul, D., Ghorab, L. M., Hernandez, M. O., Wong, M., Kedei, N., & Cappell, S. D. (2023). Irreversible cell cycle exit associated with senescence is mediated by constitutive MYC degradation. Cell Reports, 42(9), 113079.

Cossetto, T., Rodenfels, J., & Sartori, P. (2025). Thermodynamic dissipation constrains metabolic versatility of unicellular growth. Nature Communications, 16, 8543.

Ferrell, J. E., & Xiong, W. (2001). Bistability in cell signaling: How to make continuous processes discontinuous, and reversible processes irreversible. Chaos, 11(1), 227–236.

Friston, K. (2010). The free-energy principle: a unified brain theory? Nature Reviews Neuroscience, 11(2), 127–138.

Friston, K. (2013). Life as we know it. Journal of the Royal Society Interface, 10(86), 20130475.

Gao, H., et al. (2022). Physics-Informed Variational Auto-Encoder for stochastic differential equations. Computer Methods in Applied Mechanics and Engineering, 398, 115236.

Kim, Hong & Kim (2022). Systematic inference identifies a major source of heterogeneity in cell signaling dynamics. Science Advances, 8(11), eabl4598.

Hasan, A., et al. (2021). Identifying latent stochastic differential equations. IEEE Transactions on Signal Processing, 70, 89–104.

He, J., Spokoyny, D., Neubig, G., & Berg-Kirkpatrick, T. (2019). Lagging inference networks and posterior collapse in variational autoencoders. Proceedings of ICLR 2019.

Kelly, E., Bailey, C. P., & Henderson, G. (2008). Agonist-selective mechanisms of GPCR desensitization. British Journal of Pharmacology, 153(S1), S379–S388.

Kim, E.-J. (2022). Information geometry and non-equilibrium thermodynamic relations. Journal of Statistical Mechanics, 2021(9), 093406.

Kingma, D. P., & Welling, M. (2014). Auto-encoding variational Bayes. Proceedings of ICLR 2014.

Kirchhoff, M., Parr, T., Palacios, E., Friston, K., & Kiverstein, J. (2018). The Markov blankets of life. Journal of the Royal Society Interface, 15(138), 20170792.

Kuchling, F., Friston, K., Georgiev, G., & Levin, M. (2020). Morphogenesis as Bayesian inference. Physics of Life Reviews, 33, 88–108.

Laboratorios Alexandria (2026a). AP-009: Constitutional governance as engineering strategy for autonomous epistemic systems. Zenodo.

Laboratorios Alexandria (2026b). AP-014: Near-knowledge — Measuring epistemic distance in cross-domain discovery. Zenodo.

Lotfollahi, M., et al. (2023). Biologically informed deep learning to query gene programs in single-cell atlases. Bioinformatics, 39(6), btad387.

Menon, S., Blei, D., & Vondrick, C. (2022). Forget-me-not! Contrastive critics for mitigating posterior collapse. Proceedings of ICML 2022.

Park, B. O., Ahrends, R., & Teruel, M. N. (2012). Consecutive positive feedback loops create a bistable switch that controls preadipocyte-to-adipocyte conversion. Cell Reports, 2(4), 976–990.

Parr, T., Da Costa, L., & Friston, K. (2020). Markov blankets, information geometry and stochastic thermodynamics. Phil. Trans. R. Soc. A, 378(2164), 20190159.

Seninge, L., et al. (2021). VEGA: VAE Enhanced by Gene Annotations. Nature Communications, 12, 5684.

Sha, W., et al. (2003). Hysteresis drives cell-cycle transitions in Xenopus laevis egg extracts. PNAS, 100(3), 975–980.

Wang, L., Walker, B. L., Iannaccone, S., Bhatt, D., Kennedy, P. J., & Tse, W. T. (2009). Bistable switches control memory and plasticity in cellular differentiation. PNAS, 106(16), 6638–6643.

Yang, M., et al. (2024). Manatee: predicting transcript-specific effects of genetic perturbations using a variational autoencoder. Scientific Reports, 14, 12345.

Laboratorios Alexandria
laboratoriosalexandria.com