Deterministic Reference-Based Protein Structure Verification via Constraint Graph Sheaf Energy

Jason L. Volk[1]

[1]Invariant Research, Garland, Texas, USA

May 2026

Abstract

We present SATYA Protein, a deterministic verification engine for protein structures based on constraint graph sheaf energy. Given a candidate structure and a native reference, SATYA constructs a constant cellular sheaf over the residue constraint graph, where six-dimensional stalks encode per-residue spatial displacement (normalized by radius of gyration) and signed backbone dihedral deviation (normalized by 180°). The binary SAFE/UNSAFE verdict is gated on three independent checks: contact-network preservation ($Q \geq 0.90$, native contact fraction), localized damage (per-region constraint severity), and chiral integrity (mean dihedral deviation). Dirichlet energy and spectral gap are reported as continuous diagnostics measuring the smoothness of the deviation field across the constraint graph. Every verification produces a cryptographically signed receipt (Ed25519). The current constant-sheaf implementation uses identity restriction maps, operating as a vector-valued graph Dirichlet-energy verifier; the architecture supports future extension to non-identity maps for obstruction detection beyond the current energy pathway.

We validate on a controlled perturbation benchmark of seven proteins spanning 10 to 26,700 residues, verified in 27 seconds total on a single consumer CPU with no GPU, no force field, and no training data. Near-native sensitivity analysis using Gaussian Cα noise at six magnitudes (0.25–6.0 Å) shows zero false positives at thermal scale (0.25 Å), with the SAFE/UNSAFE transition occurring between 0.5 and 1.0 Å noise, corresponding to the onset of native contact disruption. A mirror-image experiment demonstrates that SATYA detects chiral inversions that pairwise distance metrics (including lDDT) mathematically cannot: reflection preserves all Cα distances (lDDT = 1.0), but inverts signed backbone dihedrals, which SATYA detects via the six-dimensional stalk (mean dihedral deviation ∼0.55, verdict UNSAFE). Conformation-specific verification on three fold-switching proteins (KaiB, RfaH, XCL1) shows that SATYA returns UNSAFE and localizes constraint violations to the fold-switching residues, while pLDDT, which is not designed as a conformation-specific reference verifier, does not identify the mismatch.

SATYA is complementary to pLDDT and MolProbity, providing independent reference-based verification with per-residue localization and cryptographic audit. Live demo and project materials: https://invariant.pro/protein.

1 Introduction

The accuracy of protein structure prediction has advanced dramatically with AlphaFold2 [1], RoseTTAFold [2], ESMFold [3], and OpenFold [4]. These systems predict three-dimensional protein structures from amino acid sequences with near-experimental accuracy on many targets. Each provides per-residue confidence metrics: AlphaFold2 and OpenFold produce pLDDT (predicted Local Distance Difference Test) [1], a learned score reflecting the model's estimate of how well the prediction would agree with an experimental structure.

However, confidence metrics are designed to answer a specific question: how well does this prediction agree with experimental structure? pLDDT is a neural network head trained to estimate this agreement; it reflects the model's internal assessment of its own accuracy, not whether the output is structurally coherent with respect to a specific reference conformation. This distinction matters most at the boundaries of the training distribution. Recent work has shown that pLDDT can assign high confidence to physically implausible

structures [9], that low-confidence regions are not always structurally disordered [10], and that fold-switching proteins present a particularly challenging case, with a 33.6% high-confidence misprediction rate across validated targets [12]. These results do not indicate that pLDDT is unreliable within its design scope; they indicate that reference-based structural verification is a separate question that requires a separate tool.

Existing structural validation tools address parts of this problem. MolProbity [5] checks bond lengths, bond angles, Ramachandran outliers, and steric clashes. WHAT_CHECK [6] performs similar geometric assessments. PDB validation reports aggregate these per-residue scores. All of these tools evaluate individual geometric criteria independently. A structure can, in principle, pass every bond length check, every angle check, and every clash check while still containing regions where the constraints collectively produce high total energy in the constraint network. This is a failure of local-to-global consistency that no individual criterion detects in isolation.

We present SATYA Protein, a verification engine based on cellular sheaf energy [8, 7] that measures this reference-based constraint-network consistency. The key insight is that protein structural constraints form a network, and the question of whether a candidate structure is consistent with a reference is a question about whether local constraint data assembles into a coherent global section of a cellular sheaf. When it does not, the sheaf energy framework detects the failure through Dirichlet energy, spectral gap, and per-residue constraint violation severity. The mathematical framework supports extension to non-identity restriction maps for detecting cyclic incompatibilities beyond the current energy pathway; the current constant-sheaf implementation operates entirely through the geometric energy pathway.

The complete system described in this paper is fully implemented and all experiments run on a single commodity machine (Intel i9-13900H, 64 GB RAM, no GPU). The underlying sheaf engine has formal Lyapunov stability proofs reported in a companion manuscript archived at Zenodo (DOI: 10.5281/zenodo.19598076) and submitted to the ICML 2026 AI4Math Workshop. This paper focuses on the protein structure verification application.

Our contributions:

1. A sheaf-theoretic formulation of protein structure verification using a constant sheaf over the residue constraint graph, with a contact-first verdict logic that decouples fold integrity (Q, severity, dihedral deviation) from deviation field smoothness (Dirichlet energy, coherence).

2. A controlled perturbation benchmark of seven proteins spanning three orders of magnitude in size (10 to 26,700 residues), verified in 27 seconds total on CPU.

3. Near-native sensitivity characterization showing zero false positives at 0.25 Å thermal noise and a physically meaningful SAFE/UNSAFE transition at 0.5–1.0 Å.

4. A mirror-image experiment proving that SATYA detects chiral inversions (lDDT = 1.0, SATYA = UNSAFE) via signed dihedral encoding in the six-dimensional stalk.

5. Direct comparison with lDDT [14] showing that SATYA augments distance-based verification with signed dihedral encoding, enabling chirality detection that distance-only metrics cannot provide.

6. Conformation-specific verification on three fold-switching proteins (KaiB, RfaH, XCL1), demonstrating detection where pLDDT does not flag conformation mismatch.

7. Cryptographic receipts (Ed25519) on every verification, producing an independently auditable provenance artifact.

Claims and scope. The current system is a constant-sheaf, Cα-based, reference-based, deterministic verifier. Demonstrated capabilities include: localization of constraint violations, contact-network disruption detection, chiral inversion detection via signed dihedrals, and cryptographic receipts. Not yet demonstrated: non-identity restriction maps, cohomological obstruction detection, side-chain validation, or broad benchmark superiority over all existing tools. The paper validates the geometric energy pathway of the sheaf framework; the cohomological pathway is future work.

2 Methods

2.1 Problem Statement

Given a native reference structure S_{ref} and a candidate structure S_{cand} over the same residue sequence, determine whether S_{cand} is consistent with S_{ref} across the backbone, contact, disulfide, and steric constraint network. Localize any inconsistencies to specific residue regions. Produce a signed verification artifact.

2.2 Constraint Extraction

From the native and candidate structures, we extract four classes of constraints:

Backbone bond constraints. For each pair of adjacent residues $(i, i{+}1)$, the Cα–Cα distance is compared between native and candidate. The deviation is $|d_{\text{cand}}(i, i{+}1) - d_{\text{ref}}(i, i{+}1)|$.

Native contact constraints. All Cα pairs within 8 Å in the native structure form native contacts. The 8 Å threshold follows the standard Cα contact definition used in native contact analysis [11]. For each contact pair (i, j), the distance in the candidate is compared to the native distance.

Disulfide bond constraints. Cysteine residues forming disulfide bonds in the native structure are identified by Sγ–Sγ distance ≤ 2.5 Å, consistent with the standard crystallographic range for disulfide bonds (2.0–2.1 Å equilibrium, with 2.5 Å as a conservative upper bound). The candidate must preserve these distances within tolerance.

Steric exclusion constraints. All non-bonded Cα pairs closer than 3.0 Å in the candidate (but not in the native) are flagged as steric violations. The 3.0 Å threshold approximates the minimum Cα–Cα distance observed in high-resolution crystal structures. A spatial cell list provides $O(n)$ steric detection.

Each constraint c has a source residue, target residue, type, measured value, reference value, and computed severity score $\sigma_c \in [0, 1]$ (see Section 2.9).

2.3 Coordinate Frames and Superposition

The SAFE/UNSAFE verdict and the violation localization depend only on quantities that are invariant to the relative placement of the candidate and reference structures. The native contact fraction Q and the per-residue severity scores are computed from pairwise Cα distances, which are unchanged by any rigid-body translation or rotation. The chiral-integrity check uses backbone dihedral angles, which are internal coordinates and likewise independent of the coordinate frame: a proper rotation leaves them unchanged, while a reflection inverts their sign, which is exactly the signal the chirality test is designed to detect. SATYA therefore applies no superposition step before computing the verdict, and the decision is well-defined regardless of how the two structures are oriented in space. This is the same superposition-free property that lDDT provides for distance-based scoring.

The spatial components of the stalk ($(x_{\text{cand}} - x_{\text{ref}})/R_g$) are the only part of the pipeline that references absolute coordinates. The Dirichlet energy built from these components is translation-invariant by construction: the coboundary takes differences of neighboring stalks, $(\delta_0 s)(e) = s(v_j) - s(v_i)$, so a uniform translation of the candidate cancels exactly. It is not rotation-invariant. For this reason the Dirichlet energy and coherence are reported as diagnostics only, and only for candidate and reference pairs that share a coordinate frame, as in the controlled perturbation and near-native experiments where each decoy is generated by perturbing the reference in place. They never gate the verdict. For comparisons between independently determined structures in unrelated frames (for example, an AlphaFold2 prediction and a deposited experimental structure), the decision rests entirely on the frame-independent metrics (Q, per-residue severity, and signed dihedral deviation), and the orientation-dependent energy is not reported. A standard Kabsch superposition of the candidate onto the reference would make the energy comparable in that setting; it is a straightforward addition, and the present results do not depend on it.

2.4 Sheaf Construction

A natural question is why the sheaf framework is needed when the per-residue constraint severities σ_c already identify violated residues. The answer is that direct thresholding of σ_c identifies *locally* violated residues; the Laplacian's Dirichlet energy aggregates these local violations into a global inconsistency measure that

weights each violation by its position in the constraint network. Three borderline violations ($\sigma_c \approx 0.1$) on residues forming a triangle in the constraint graph produce higher energy than three isolated violations of the same magnitude, because the deviations at neighboring vertices reinforce each other through the coboundary sum. This non-local aggregation is the value the sheaf adds over direct thresholding.

The constraint graph $G = (V, E)$ has one vertex per residue and one edge per constraint. A cellular sheaf $\mathcal{F}$ over G assigns:

Stalks. Each vertex v_i carries a stalk $\mathcal{F}(v_i) = \mathbb{R}^6$ encoding the residue's normalized deviation from native geometry. Components [0–2] are the spatial displacement of the Cα atom, normalized by the native radius of gyration R_g: $(x_{\text{cand}} - x_{\text{ref}})/R_g$. Component [3] is the phi dihedral deviation, normalized by 180°: $(\phi_{\text{cand}} - \phi_{\text{ref}})/180$. Component [4] is the psi dihedral deviation, similarly normalized. Component [5] is reserved for future stalk extensions (e.g., solvent accessibility, crystallographic B-factor deviation) and is held at zero in all experiments reported here. As a permanently zero entry it does not affect any current result; it is retained so that future physical data streams can be added without changing the stalk dimension or regenerating the receipt baseline. This normalization produces dimensionless stalk entries where a value of 1.0 represents a displacement equal to the protein's radius of gyration or a half-turn dihedral deviation. For a typical globular protein with $R_g \approx 10$–15 Å, a 1 Å displacement produces a stalk entry of $\sim$0.07–0.10, comparable in magnitude to a $\sim$15-degree dihedral deviation ($15/180 \approx 0.08$). The backbone angle deviation is computed from signed dihedral angles (ϕ, ψ), which explicitly break chiral symmetry: a mirror-image structure preserves all pairwise Cα distances but inverts dihedral signs, producing nonzero stalk entries and a detectable violation. Of the six stalk components, the first three (Cα displacement) define the spatial deviation; components [3] and [4] (backbone angle deviations) are derived quantities computed from the dihedral angles; component [5] is reserved for future use.

Restriction maps. Each edge $e = (v_i, v_j)$ carries restriction maps $\rho_{v_i \to e}$ and $\rho_{v_j \to e}$ from incident vertex stalks to the edge stalk $\mathcal{F}(e) = \mathbb{R}^6$. The edge stalk serves as a comparison space: the restriction maps project each incident vertex's deviation data into a common frame where pairwise consistency can be evaluated. When both restriction maps are identity, the edge stalk simply compares the two vertex stalks directly. Restriction maps are identity matrices for all constraints, regardless of violation severity. The violation is encoded entirely in the stalk deviation vectors, not in the sheaf structure. This produces a constant sheaf over the constraint graph, where the coboundary measures raw deviation differences between neighboring residues: $(\delta_0 s)(e) = s(v_j) - s(v_i)$.

With identity restriction maps, the sheaf Laplacian is $L_{\text{graph}} \otimes I_d$ (the graph Laplacian tensored with the $d \times d$ identity). The current implementation therefore functions as a vector-valued graph Dirichlet-energy verifier: detection relies entirely on the geometric energy pathway (Dirichlet energy, native contact fraction Q, and per-residue constraint violation severity). This is a deliberate design choice that validates the energy landscape and scaling properties of the framework before introducing non-identity restriction maps. Future extensions using physically meaningful non-identity maps (e.g., Ramachandran-basin constraints, bond-angle propagation) could enable detection of cyclic incompatibilities in the constraint network that the current energy pathway does not capture.

A constant sheaf with $\mathbb{R}^6$ stalks is not equivalent to a scalar graph Laplacian on individual violation scores. The stalk encodes spatial displacement *and* dihedral angle deviations in parallel: the Dirichlet energy additively penalizes neighboring residues whose displacement vectors and angle deviations diverge. A scalar Laplacian on σ_c values alone cannot capture chiral inversions (which leave pairwise distances unchanged but flip dihedral signs) or cases where spatial displacement is small but angular deviation is large. The constant sheaf validates the energy landscape and scaling properties of the framework before the introduction of physically coupled non-identity maps.

Section. A global section $s \in C^0(G; \mathcal{F})$ assigns to each vertex its deviation vector. For a native-vs-native comparison, $s = 0$ (zero section). For a corrupted candidate, s is nonzero at damaged residues.

2.5 Sheaf Energy Analysis

The coboundary operator $\delta_0 \colon C^0 \to C^1$ maps vertex sections to edge cochains:

$$(\delta_0 s)(e) = \rho_{v_j \to e}\big(s(v_j)\big) - \rho_{v_i \to e}\big(s(v_i)\big) \tag{1}$$

The sheaf Laplacian $L_{\mathcal{F}} = \delta_0^\top \delta_0$ is positive semidefinite. Its spectrum encodes global structural information:

Spectral gap λ_2: the smallest nonzero eigenvalue of $L_{\mathcal{F}}$. Measures how well-connected the constraint network is. Higher λ_2 indicates a more tightly constrained structure.

Dirichlet energy $\mathcal{E}[s] = s^\top L_{\mathcal{F}} s$: total inconsistency of the deviation section over the constraint graph. Expanding the quadratic form:

$$\mathcal{E}[s] = \sum_{e=(v_i,v_j)\in E} \left\| \rho_{v_j \to e}(s(v_j)) - \rho_{v_i \to e}(s(v_i)) \right\|^2 \tag{2}$$

Each term measures the inconsistency of a single constraint; the sum aggregates them into a global energy. For a native match, $\mathcal{E} = 0$.

Kernel diagnostic. In the current constant-sheaf implementation, the Laplacian kernel has the expected dimension $d \cdot k$, where d is the stalk dimension and k is the number of connected components. The current protein pipeline does not perform cohomological obstruction detection. Future versions with non-identity restriction maps may introduce additional compatibility diagnostics to detect cyclic incompatibilities not captured by the current geometric energy pathway.

Global coherence: $C = \lambda_2/(\lambda_2 + \mathcal{E} + \epsilon)$, where $\epsilon = 10^{-10}$ prevents division by zero. For a native match, $\mathcal{E} = 0$ and $C \approx 1$. For a severely damaged structure, $C \to 0$.

Numerical computation. The sheaf Laplacian $L_{\mathcal{F}}$ is a real symmetric matrix of dimension $d|V| \times d|V|$, where $d = 6$ is the stalk dimension. For single-domain proteins ($|V| \leq 500$), the matrix is at most $3{,}000 \times 3{,}000$; for domain-decomposed proteins, each domain produces a Laplacian of at most this size. We compute the full eigendecomposition via `numpy.linalg.eigh` (symmetric divide-and-conquer), which returns all eigenvalues and eigenvectors in sorted order. This is exact up to floating-point precision, with no iterative convergence criteria to tune.

The primary outputs of the eigendecomposition are the spectral gap λ_2 (the smallest eigenvalue exceeding a numerical zero threshold of 10^{-10}) and the Dirichlet energy $\mathcal{E}[s] = s^\top L_{\mathcal{F}} s$. The spectral gap measures how tightly the constraint network couples neighboring residues; higher λ_2 indicates a more rigid network.

For the current constant sheaf (identity restriction maps), the kernel has exactly $d \cdot k$ dimensions, where k is the number of connected components. No additional obstruction counting is performed in the current formulation. Future extensions using non-identity restriction maps may introduce additional compatibility diagnostics computed from the appropriate sheaf operators [7].

2.6 Constraint Violation Localization

Per-residue constraint violation severity scores are computed from the constraint violations incident on each residue:

$$\text{sev}(i) = \max_{c \ni i} \sigma_c \tag{3}$$

where the maximum is over all constraints involving residue i. This is a local maximum over incident constraint severities, not a cohomological invariant; it identifies the worst-case constraint violation at each residue position. Contiguous regions of high severity scores are merged into violation regions. Each region reports the residues involved, the dominant constraint type (backbone, contact, steric, or disulfide), the maximum severity, and specific constraint evidence.

2.7 Decoy Generation

For each protein in the controlled perturbation benchmark, a synthetic decoy is generated by rigid-body displacement of a contiguous residue region. The synthetic decoys are not intended to model realistic misfolding; they are controlled perturbations for testing detection, localization, scaling, and receipt generation under known ground truth. Any reasonable verification method would detect a 6 Å rigid translation of a contiguous Cα block; the purpose is to verify that the engine correctly localizes the known corruption, produces correct energy and Q values, and generates valid signed receipts.

The displacement protocol is deterministic:

1. Select a target region based on the structural property being tested (e.g., the RBM loop for Spike RBD, the interchain disulfide region for Insulin, beta-strands 4–5 for TIM barrel).

2. Apply a uniform translation of 6 Å along the x-axis to all Cα atoms in the target region. The 6 Å magnitude was chosen to exceed the 8 Å native contact threshold, ensuring that native contacts between displaced and non-displaced residues are broken.

3. All residues outside the target region remain at their native coordinates.

This protocol produces controlled, reproducible corruptions where the ground-truth damaged region is known exactly, enabling unambiguous evaluation of detection and localization accuracy. The decoy generation scripts are part of the SATYA Protein engine (see Data and Code Availability).

For the frozen trajectory case (used in preflight testing), a synthetic 500-frame trajectory is generated with zero atomic motion across all frames, triggering the preflight rejection pathway.

2.8 Domain Decomposition

For large proteins exceeding 500 residues, the structure is segmented into domains. Each domain is verified independently with its own sheaf, and results are assembled into a multi-domain report. This maintains empirical near-linear total cost: each domain verification is bounded by domain size (typically ≤ 500 residues), and the number of domains scales linearly with protein size. The 26,700-residue Titin tandem (300 Ig-like domains of $\sim$89 residues each) is verified using this strategy. Domain boundaries correspond to the repeating immunoglobulin fold units resolved in the crystal structure of PDB entry 1WAA, with chain breaks between consecutive Ig-like repeats providing unambiguous segmentation points. Inter-domain linker residues (typically 2–4 residues between Ig folds) are included in the downstream domain. For the protein application, each domain is verified independently: the per-domain analysis is computed separately, and results are aggregated into a multi-domain report. This means that purely inter-domain constraint violations, those that exist only in the interaction between two domains and are not visible within either domain alone, are not detected by the current protein pipeline. For Titin's repeating Ig-like domains, this is a minor limitation: the domains are structurally independent repeats with minimal inter-domain contact networks beyond the linker backbone. For proteins with extensive inter-domain interfaces (e.g., multi-domain enzymes with shared active sites), this would be a more significant gap.

The underlying sheaf engine supports a more general cellular decomposition with overlap handling, validated at scales up to 5 million vertices. In the general case, a bounded-cell partition with shared boundary vertices localizes eigensolves to individual cells, and global H^1 is assembled from cell-level contributions via Mayer–Vietoris exact sequences, which require sheaf data on cell intersections to correctly propagate information across boundaries. Extending the protein pipeline to use overlapping domain decomposition with Mayer–Vietoris assembly is a natural next step that would close this gap. Under streaming updates in the general engine, only the cell containing the affected edge requires recomputation; all other cells remain cached. At 5 million vertices, the engine achieves 35 microsecond median per-edit latency with zero measured cohomological drift. For protein verification, this means the engine can handle structures far larger than any currently predicted or deposited protein, with room to scale to entire proteome-level verification pipelines.

2.9 Parameter Choices

Table 1 summarizes all thresholds and tolerances used in the verification pipeline.

2.10 Cryptographic Receipts

Every verification produces a signed receipt containing: input hash (SHA-256 of the candidate coordinates), source hash (SHA-256 of the reference structure), verdict, sheaf metrics (λ_2, $\mathcal{E}$, coherence, kernel diagnostic), violation count and locations, and a timestamp. The receipt is signed with Ed25519 using a persistent key pair. Any third party with the public key can independently verify that the receipt was produced by the claimed engine and has not been modified. The public key is published in the project repository and on the project website (`https://invariant.pro/protein`).

The cryptographic receipt addresses a specific provenance problem: when a structure is shared between organizations (e.g., pharmaceutical partner review, regulatory submission, database deposition), the receiving party has no way to verify that the structure was independently validated, or that the validation result has

Table 1: Parameter choices for SATYA Protein verification.

Parameter	Value	Justification
Contact distance	8 Å	Standard Cα contact definition [11]
Disulfide Sγ–Sγ	≤ 2.5 Å	Conservative upper bound on SS bond
Steric exclusion	3.0 Å	Min. Cα–Cα in high-res. structures
Severity score σ_c	$[0, 1]$	Linear: $\sigma_c = \min(1,\ \lvert d - d_{\text{ref}}\rvert/\tau)$
Severity tolerance τ	2.0 Å	Normalizes deviation to $[0, 1]$ range
Satisfied threshold	$\sigma_c < 0.1$	Deviations below 10% of τ treated as noise
Domain decomposition	500 residues	Balances cell size vs. domain count
Coherence ϵ	10^{-10}	Numerical stability in $C = \lambda_2/(\lambda_2 + \mathcal{E} + \epsilon)$
Refinement step size	$\alpha = 0.01$	Gradient descent learning rate
Refinement energy tol.	$\mathcal{E} < 0.01$	Energy convergence criterion
Refinement coord. tol.	$\lVert\Delta x\rVert_\infty < 10^{-4}$ Å	Coordinate convergence
Max refinement steps	1,000	Upper bound on iterations
Decoy displacement	6 Å	Exceeds contact threshold
Stalk dimension	$d = 6$	3 spatial/R_g + 2 dihedral/180 + 1 reserved

not been altered after the fact. For basic scientific use (single-researcher verification), the receipt is simply a structured JSON record of the verification result. The cryptographic layer adds value only in contexts where independent auditability matters.

3 Results

3.1 Controlled Perturbation Benchmark

We validate SATYA Protein on a benchmark of seven proteins (Table 2), each testing a distinct architectural property of the verification engine. For each protein, a synthetic decoy is generated using the rigid-body displacement protocol described in Section 2.7. The synthetic decoys are controlled perturbations with maximal signal-to-noise ratio, designed to verify engine correctness under known ground truth. The engine must: (1) detect the corruption, (2) localize it to the correct residues, and (3) produce a signed receipt documenting the findings.

Table 2: SATYA Protein controlled perturbation benchmark. All runs on a single consumer CPU (Intel i9-13900H, 64 GB RAM, no GPU). Time includes constraint extraction, sheaf construction, energy analysis, refinement, and receipt signing. The Titin entry (1WAA×300) is a synthetic repeat assembly used only as a domain-decomposition scaling stress case, not a natural deposited structure.

#	Protein	PDB	Residues	Time (s)	Property Tested
1	Chignolin	1UAO	10	<0.1	Loop closure
2	Insulin	4INS	51	0.2	Multi-chain, cross-chain SS
3	BPTI	5PTI	58	0.2	Long-range SS network
4	Aβ42 fibril	5OQV	126	0.9	Interchain aggregate
5	Spike RBD	6M0J	194	1.7	Flexible loops, 4 SS
6	TIM barrel	1YPI	247	2.4	Most common fold
7	Titin (synthetic)	1WAA×300	26,700	22.0	Synthetic domain decomposition scaling
	Total		**27,386**	**27.5**	

3.2 Detection and Localization

For each decoy, SATYA correctly identified the corrupted region. In every case, the violation regions reported by the engine overlapped exactly with the residues displaced during decoy generation. No false positives were produced on native-vs-native comparisons: all natives returned coherence = 1.0, RMSD = 0.0 Å, and a zero kernel diagnostic.

Specific localization results for representative cases:

Insulin (4INS). Chain B displaced by 6 Å. The engine localized the constraint violations to the interchain disulfide region (CYS7A–CYS7B, CYS20A–CYS19B) and reported the broken Sγ–Sγ contacts as specific evidence. The violations did not spread to Chain A residues far from the disulfide network, confirming that localization tracks the structural damage rather than propagating globally.

Spike RBD (6M0J). Receptor-binding motif (RBM) loop residues displaced. Constraint violations were localized to the RBM loop with the broken C480–C488 disulfide bond cited as specific evidence. Native contacts between the displaced loop and the surrounding beta-sheet core registered as secondary violations, consistent with the known structural role of the RBM loop in ACE2 binding.

TIM barrel (1YPI). Beta-strands 4–5 displaced, opening the barrel. The engine detected the disruption in the barrel topology and localized violations to the displaced strand region. The barrel's eight-fold symmetry means local damage to two strands propagates through native contacts to adjacent strands, and the violation map reflected this connectivity.

3.3 Energy Surface Validation

To verify that the Dirichlet energy surface is well-formed, we applied gradient descent on the sheaf Laplacian to each synthetic decoy (Table 3). All decoys converged to near-zero RMSD within 200–400 steps, confirming that the native reference is the global energy minimum and that the gradient structure is correct. This is a consistency check on the energy formulation, not a demonstration of realistic structure repair: the stalk encodes displacement from the known native reference, so gradient descent is mathematically equivalent to linear interpolation toward the answer.

Table 3: Energy surface validation: gradient descent on synthetic decoys converges to the native reference, confirming the Dirichlet energy landscape is well-formed. This is not proposed as a practical structure repair tool.

Protein	Steps	Final RMSD (Å)	Final Q	Final coherence
Chignolin	210	0.09	1.00	1.00
Insulin	340	0.03	1.00	1.00
BPTI	280	0.04	1.00	1.00
Aβ42 fibril	310	0.02	1.00	1.00
Spike RBD	390	0.03	1.00	1.00
TIM barrel	370	0.04	1.00	1.00

3.4 Near-Native Sensitivity Characterization

The controlled perturbation benchmark uses 6 Å rigid displacements with maximal signal-to-noise. To characterize SATYA's behavior in the near-native regime where structural validation is actually hard, we generated Gaussian Cα noise decoys at six magnitudes (σ = 0.25, 0.50, 1.0, 2.0, 4.0, 6.0 Å per coordinate) for three proteins (BPTI, Spike RBD, TIM barrel), with 10 random seeds per condition.

SATYA outputs two distinct classes of metrics that decouple contact-network preservation from geometric coherence. The native contact fraction (Q) and per-residue severity measure contact-network preservation: whether the fold's contact network is intact. The Dirichlet energy ($\mathcal{E}$) and coherence measure geometric coherence: how smoothly the deviation field varies across the constraint graph.

At σ = 0.25 Å (thermal-scale perturbation, actual RMSD ∼0.43 Å), Q remains at 1.0 and no localized obstructions are detected, correctly yielding a SAFE verdict across all proteins and seeds (zero false positives).

The Dirichlet energy is nonzero (∼10–45, scaling with graph size) because independent per-residue noise maximizes the Laplacian quadratic form regardless of perturbation magnitude: random noise represents a high-frequency deviation field with maximal graph gradient. The binary SAFE/UNSAFE verdict is therefore gated on contact-network metrics (Q and per-residue severity), not on energy or coherence.

The SAFE/UNSAFE transition occurs between $\sigma = 0.5$ Å and $\sigma = 1.0$ Å, corresponding to RMSD ∼0.87–1.73 Å. At $\sigma = 0.5$ Å, Q drops to ∼0.94–0.97 and the verdict is mixed (0–40% UNSAFE depending on protein). At $\sigma = 1.0$ Å, Q drops to ∼0.75 and the verdict is 100% UNSAFE. This transition is physically meaningful: perturbations below ∼0.5 Å are within crystallographic B-factor range, while perturbations above ∼1.0 Å begin disrupting native contacts.

A correlated drift experiment provides further insight into the Dirichlet energy's physical meaning. A 4.0 Å coherent drift across a 30-residue loop in TIM barrel produces significantly lower Dirichlet energy (9.2) than 0.25 Å random noise on the same protein (47.0). This correctly reflects that coherent drift is a smooth, low-frequency deformation with minimal gradient across the constraint graph, while random noise is incoherent. The drift is detected by the Q gate ($Q = 0.936 < 0.90$), not by the energy. Thus, Q measures whether the fold is broken, while Dirichlet energy measures how chaotically the deformation field varies.

3.5 Mirror-Image Detection: Chiral Inversion

To test whether SATYA's six-dimensional stalk provides information beyond pairwise distance metrics, we generated mirror-image structures by inverting the x-coordinate of all atoms ($x \to -x$) for five benchmark proteins. Reflection is an isometry: it preserves all pairwise Cα distances exactly. Any metric based solely on pairwise distances, including lDDT [14], returns a perfect score on a mirror-image structure.

SATYA detects all five mirror images as UNSAFE (Table 4). The mechanism is the signed dihedral encoding in stalk components [3] and [4]: reflection inverts the sign of all backbone dihedral angles (ϕ, ψ), producing a mean dihedral deviation of ∼0.55 (normalized units, corresponding to ∼100°). The Q gate passes ($Q = 1.0$; all contacts preserved), the severity gate passes (no distance violations), but the dihedral gate triggers (mean deviation > 0.30).

Table 4: Mirror-image detection. Reflection preserves all pairwise Cα distances (lDDT = 1.0) but inverts signed backbone dihedrals. SATYA detects the inversion via the dihedral gate; lDDT is structurally blind to chirality.

Protein	Res.	Q	lDDT	Dihedral dev.	Verdict
Chignolin (1UAO)	10	1.00	1.00	0.58	UNSAFE
Aβ42 (5OQV)	42	1.00	1.00	0.56	UNSAFE
BPTI (5PTI)	58	1.00	1.00	0.55	UNSAFE
Spike RBD (6M0J)	194	1.00	1.00	0.55	UNSAFE
TIM barrel (1YPI)	247	1.00	1.00	0.56	UNSAFE

This result is not an empirical finding that depends on threshold calibration. It follows from the structure of the feature space: lDDT is a function of pairwise absolute distances; rigid-body reflection preserves all pairwise absolute distances; therefore lDDT is structurally blind to chirality. SATYA's stalk projects signed dihedral angles into the deviation vector, making chirality a first-class detectable feature. In this respect SATYA augments distance-based comparison with signed backbone dihedral deviation, detecting chiral inversions that distance-only metrics cannot. This advantage comes from the choice of feature (signed dihedrals carried in the stalk), not from the Laplacian or any sheaf-topological property: a per-residue signed-dihedral check would expose the same inversion. The contribution is encoding that feature into the verification state, so that chirality is checked by the same engine and recorded on the same signed receipt as the contact-network and severity gates.

3.6 Scaling

Wall-clock verification time scales empirically near-linearly with residue count across the benchmark. This is a fundamentally different scaling class from the tools SATYA complements. Molecular dynamics simulations

scale as $O(n \log n)$ with particle mesh Ewald or $O(n^2)$ with direct summation. Force field energy minimization scales as $O(n^2)$ for non-bonded interactions. AlphaFold2 attention scales quadratically in sequence length. MolProbity evaluates per-residue criteria independently but still requires time proportional to structure size for each criterion.

SATYA achieves empirical near-linear wall-clock scaling through two mechanisms. First, the sheaf Laplacian is sparse: each residue connects only to its backbone neighbors, nearby native contacts, and (where applicable) disulfide partners and steric neighbors detected via spatial cell list. Constraint extraction, sheaf construction, and gradient computation are all linear in the number of residues for bounded vertex degree. Second, and more importantly, domain decomposition bounds the size of every eigensolve. For proteins exceeding 500 residues, each domain contains at most ∼500 residues, so the Laplacian eigendecomposition for each domain operates on a matrix of bounded size. The cost per domain is therefore constant regardless of total protein size, and the number of domains scales linearly with n, yielding linear total cost.

We note that the near-linear scaling claim is empirical, not a worst-case theoretical bound. Computing all eigenvalues of a general $n \times n$ matrix is $O(n^3)$; iterative methods such as Lanczos are $O(n)$ per iteration but require a number of iterations that depends on the spectral gap and desired precision. For the domain-decomposed case, this distinction is moot because the matrix dimension per eigensolve is bounded by the domain size threshold (at most ∼500 residues, producing a $3{,}000 \times 3{,}000$ matrix). For small proteins verified without decomposition (all benchmark entries below Titin), the matrices are small enough ($n \leq 247$) that the cubic cost is negligible.

For proteins exceeding the domain decomposition threshold, the cost is even better in practice. The 247-residue TIM barrel (the most common fold in biology) verifies in 2.4 seconds. The 26,700-residue Titin tandem uses domain decomposition, verifying 300 domains independently in 22 seconds total (∼73 ms per domain). The per-domain cost is constant regardless of total protein size: adding another 300 domains would add another 22 seconds, not trigger a superlinear explosion.

The underlying sheaf engine has been validated at substantially larger scales (Table 5). These numbers are from the general constraint graph engine operating on synthetic random graphs with controlled degree distributions, not protein-specific runs. The protein verification module uses the same sheaf construction and energy analysis code but wraps it in a domain decomposition pipeline with protein-specific constraint extraction. The general engine benchmarks establish that the mathematical infrastructure can handle structures far beyond any currently predicted or deposited protein.

At 5 million vertices, the median per-edit latency is 35 microseconds, which is faster than the 63 microseconds measured at 1 million vertices. This counterintuitive result arises from the cellular decomposition's caching behavior: at larger scales, more cells remain cached and unaffected by any given edit, reducing the fraction of the graph that requires recomputation. The scaling exponent across $V = 21$K to $V = 1$M is 0.19 ($R^2 = 0.975$), measured from wall-clock time on a single CPU, confirming sublinear growth. Whether protein constraint graphs reproduce this sublinear exponent is an open question; protein graphs have spatially clustered native contacts and different degree distributions than synthetic random graphs. The measured protein benchmark (10 to 26,700 residues) shows near-linear scaling in wall-clock time, which is sufficient for current applications. Zero cohomological drift was measured across all four orders of magnitude, meaning the local-to-global assembly preserves exact correctness under streaming updates.

The largest entry in the Protein Data Bank is the nuclear pore complex at ∼31,000 residues per asymmetric unit. SATYA already handles 26,700 residues in 22 seconds. These general-engine benchmarks suggest headroom for much larger biological constraint networks, but direct validation on assembled capsids, ribosomes, and multi-megadalton complexes remains future work.

All benchmarks were run on a single CPU core (Intel i9-13900H). The empirical near-linear scaling is measured across three orders of magnitude on commodity hardware.

3.7 Minimal External Dependencies

The core SATYA verification engine requires only NumPy and the Python standard library for constraint extraction, sheaf construction, eigendecomposition, and verdict computation. Receipt signing uses an Ed25519 implementation (`PyNaCl` or equivalent). The lDDT comparison scripts use Biotite [15] as an independent third-party lDDT implementation. No force fields, parameter databases, license servers, or cloud dependencies are required. Every result is deterministic: given the same engine version and the same input files, the

Table 5: Underlying sheaf engine: streaming cohomological maintenance at scale (general constraint graphs, not protein-specific). Kernel-diagnostic drift is zero across all configurations.

Vertices	Cells	Edit (median)	Edit (mean)	Drift
21K	639	31 μs	31 μs	0
100K	890	46 μs	46 μs	0
1M	8,663	63 μs	63 μs	0
5M	25,473	35 μs	119 μs	0

verification reproduces exactly.

3.8 Conformation-Specific Verification on Fold-Switching Proteins

To test SATYA on real AlphaFold2 predictions rather than synthetic decoys, we verified default AlphaFold2 predictions against experimentally determined structures for three well-characterized fold-switching proteins (Table 6). Fold-switching proteins adopt two or more distinct conformations from a single amino acid sequence. AlphaFold2 typically predicts one conformation with moderate to high pLDDT, but the predicted conformation may not match the alternative fold that the protein adopts under different cellular conditions.

For each protein, we downloaded the default AlphaFold2 prediction from the AlphaFold Protein Structure Database and verified it against an experimentally determined reference structure representing the alternative conformation. The native contact fraction Q (computed from pairwise Cα distances, independent of coordinate frame) serves as the primary metric: $Q = 1.0$ indicates all native contacts are preserved, while $Q < 0.5$ indicates severe structural divergence.

The `UNSAFE` verdict in these experiments should be interpreted as conformation-specific mismatch, not as an assertion that the AlphaFold2 prediction is physically impossible or globally defective. SATYA answers the narrower verification question: does this candidate preserve the constraint network of the specified reference conformation? A prediction that is correct for one fold will inevitably fail verification against a different fold, and that is exactly the expected behavior of a reference-based verification gate.

Fold-switching proteins demonstrate that structural divergence manifests in two distinct regimes. The first is cyclic incompatibility: constraint cycles where restriction maps are collectively irreconcilable, a failure mode the mathematical framework is designed to capture in future extensions with non-identity restriction maps (Figure 1). The second is geometric violation: the backbone graph topology is preserved (both folds share the same sequence and connectivity), so no topological obstruction exists under the current constant sheaf, but pairwise distance constraints are radically violated. SATYA captures this second mode directly through the frame-independent contact metric Q and per-residue severity (Section 2.3); for these cross-frame comparisons the orientation-dependent Dirichlet energy is not used. Critically, these mismatch modes are outside the design scope of pLDDT as a predictor confidence metric.

KaiB (*T. elongatus*, 108 residues). KaiB is a circadian clock protein that adopts a ground-state fold ($\beta\alpha\beta\beta\alpha\alpha\beta$; PDB: 2QKE) and a fold-switched thioredoxin-like fold ($\beta\alpha\beta\alpha\beta\beta\alpha$; PDB: 5JYT). Wayment-Steele et al. (2024) established that the default AlphaFold2 prediction for *T. elongatus* KaiB returns the fold-switched state, not the ground state [13]. We verified the AlphaFold2 prediction (UniProt: Q79V61) against the ground-state crystal structure (2QKE, chain B). SATYA returned $Q = 0.48$ and verdict `UNSAFE`, with five violation regions at severity 1.0. The violations localized to residues 4–13 and 58–69, which correspond to the regions that undergo secondary structure rearrangement during the fold switch. pLDDT for this prediction is 73.9 (moderate confidence), providing no indication that the predicted conformation differs from the ground state.

RfaH (*E. coli*, 162 residues). RfaH is a transcription factor whose C-terminal domain (CTD, residues 101–162) switches between an α-helical hairpin in the autoinhibited state (PDB: 2OUG) and a β-barrel in the active state (PDB: 6C6S). The N-terminal domain (NTD, residues 1–100) is structurally conserved between both states. AlphaFold2 predicts the α-helical CTD. We verified the AlphaFold2 prediction (UniProt: P0AFW0) against the β-barrel conformation (6C6S, chain D, 161 residues). SATYA returned $Q = 0.64$ and verdict `UNSAFE`, with five violation regions at severity 1.0. pLDDT for this prediction is 78.6 (moderate confidence). The Q value reflects the shared NTD contacts ($\sim$60% of the total) while the violations localized to residues

103–114, 123–133, and 148–158, all within the CTD fold-switching region. No violations were detected in the NTD (residues 1–100), confirming that SATYA correctly distinguishes the structurally conserved domain from the fold-switching domain.

XCL1 (human, 93 residues after signal peptide removal). XCL1 is a metamorphic chemokine that interconverts between a monomeric chemokine fold (PDB: 1J8I) and a dimeric β-sheet fold (PDB: 2JP1) under physiological conditions. We verified the AlphaFold2 prediction (UniProt: P47992, signal peptide removed) against the chemokine fold (1J8I, chain A). SATYA returned $Q = 0.78$ and verdict `UNSAFE`, with five violation regions. pLDDT for this prediction is 78.8 (moderate confidence). The higher Q indicates that AlphaFold2 broadly captures the chemokine fold, but the violations at residues 5–12, 25–26, and 45–47 identify regions where the predicted and experimental structures diverge, consistent with the conformational dynamics of this metamorphic protein.

Table 6: SATYA conformation-specific verification of AlphaFold2 predictions against experimental structures of fold-switching proteins. Q is the native contact fraction (pairwise Cα distances, coordinate-frame independent). pLDDT is the mean per-residue AlphaFold2 confidence score for each prediction. The `UNSAFE` verdict indicates conformation-specific mismatch against the specified reference, not a claim that the prediction is globally defective. All comparisons completed in <1 second on CPU.

Protein	Reference	Res.	pLDDT	Q	Violations	Verdict
KaiB	2QKE (ground state)	108	73.9	0.48	5	`UNSAFE`
RfaH	6C6S (β CTD)	161	78.6	0.64	5	`UNSAFE`
XCL1	1J8I (chemokine)	93	78.8	0.78	5	`UNSAFE`

The key observation is that SATYA provides information that pLDDT does not. All three predictions have moderate pLDDT (73.9–78.8), placing them in AlphaFold2's "confident" range. Yet SATYA returns `UNSAFE` in every case, with violation regions localized to the exact residues that undergo secondary structure rearrangement. For KaiB, pLDDT reports 73.9 while SATYA reports $Q = 0.48$ with violations at the fold-switching residues. The signed receipt documents what was checked, what was found, and where the mismatches are.

A natural question is: if a reference structure is available, why not simply compute TM-score or GDT_TS? The answer is localization and causal decomposition. TM-score reports a single number ("this structure is 40% similar"); SATYA reports which specific residues are responsible for the structural divergence and which constraint types (backbone angle, native contact, disulfide) are violated. For RfaH, SATYA does not merely say "this is different"; it says "residues 103–114, 123–133, and 148–158 carry severity-1.0 native contact violations, and the NTD (residues 1–100) is clean." This per-residue causal decomposition, anchored to specific constraint types, is what distinguishes sheaf energy verification from global similarity metrics.

The native contact fraction Q used in this case study is itself a standard metric computable in a few lines of code. The verification framework adds two things beyond Q alone: (1) the integration of signed dihedral checks and per-residue constraint severity for causal localization, both of which are coordinate-frame independent and therefore apply directly to these cross-frame comparisons; and (2) in same-frame settings, the Dirichlet energy $\mathcal{E}$ and the spectral gap λ_2, which provide global context about constraint-network coherence and rigidity. Per-residue constraint violation localization (Section 2.6) is computed directly from constraint severity scores, not from the sheaf Laplacian; in same-frame settings the sheaf additionally supplies the global energy context within which local violations are interpreted. The framework also supports future extension to non-identity restriction maps that could enable detection of cyclic incompatibilities that neither Q nor $\mathcal{E}$ alone can capture.

4 Comparison with Existing Validation

4.1 pLDDT

AlphaFold2 and OpenFold produce per-residue pLDDT scores as a confidence metric. pLDDT is a learned quantity: a neural network head trained to estimate how well the predicted structure would agree with an

experimental structure based on the Local Distance Difference Test [1]. It measures the model's internal estimate of prediction accuracy, which is valuable but categorically distinct from independent reference-based structural verification.

Recent work has clarified the boundaries of this internal assessment. Thacker (2026) screened 27,098 AlphaFold2 predictions across 18 experimentally validated fold-switching proteins and found a 33.6% high-confidence misprediction rate: cases where pLDDT reports high confidence but the predicted conformation does not match the alternative fold [12]. The study also identified an "inverted confidence" category, invisible in benchmark data, where the incorrect conformation receives higher pLDDT than the correct one. Across the blind screen, there was zero population-level correlation between pLDDT and conformational accuracy. Separately, Lahey et al. (2026) found that 30–40% of all amino acids in model organism predictions fall into the low-confidence pLDDT range, and that high pLDDT does not always rule out structural implausibility [9]. These findings reflect the design scope of pLDDT, not a deficiency: pLDDT was trained on single-conformation agreement, and it performs well within that scope. Fold-switching proteins and other multi-state systems simply pose a different question.

SATYA measures something categorically different. It tests whether a structure preserves the constraint network of a specified reference conformation, independent of the predictor that generated it. A structure can have uniformly high pLDDT (the model is confident) while carrying constraint violations that are only detectable by global energy analysis against a specific reference. Critically, SATYA does not ask "how confident is the predictor?" It asks "does this structure preserve the constraints of the reference?" The answer is mathematical, not statistical, and it does not depend on the predictor's training data, architecture, or confidence head.

pLDDT and SATYA are complementary: pLDDT estimates prediction quality, SATYA independently verifies reference-based structural consistency. The natural workflow is: predict, score confidence with pLDDT, then verify consistency with SATYA against the intended reference conformation. For fold-switching proteins and other cases at the boundary of pLDDT's design scope, SATYA provides an independent verification layer that answers a structurally distinct question.

4.2 lDDT

The local Distance Difference Test (lDDT) [14] is the standard reference-based metric for local structure evaluation. It measures the fraction of pairwise distances within an inclusion sphere that are preserved below a set of absolute thresholds (0.5, 1, 2, 4 Å). lDDT is per-residue, reference-based, and superposition-free, making it the most direct existing comparator to SATYA's native contact fraction Q.

We computed per-residue lDDT using Biotite [15] on Cα atoms for the same structures evaluated by SATYA (synthetic decoys, Gaussian noise at four levels, and mirror images). On distance-based detection, lDDT and SATYA's Q degrade similarly with noise: at $\sigma = 0.25$ Å, lDDT is 0.96–0.97 while Q is $\sim$1.0; at $\sigma = 1.0$ Å, lDDT drops to $\sim$0.67 and Q to $\sim$0.75; at $\sigma = 2.0$ Å, both reach $\sim$0.46–0.50. The slightly lower lDDT scores at equivalent noise reflect lDDT's finer distance thresholds (0.5/1/2/4 Å bins vs. SATYA's 8 Å contact threshold).

The critical distinction is chirality. On mirror-image structures, lDDT returns 1.0 for all five benchmark proteins (Section 3.5, Table 4). lDDT is a function of pairwise absolute distances; reflection preserves all pairwise distances; therefore lDDT is structurally blind to chirality. SATYA returns UNSAFE on every mirror via the signed dihedral encoding in the stalk. This is a provable advantage, not an empirical one.

SATYA therefore augments lDDT's distance-based verification with backbone dihedral integrity, a dimension of structural reality that distance-only metrics cannot access.

4.3 MolProbity and WHAT_CHECK

MolProbity [5] and WHAT_CHECK [6] evaluate individual geometric criteria: bond lengths, bond angles, Ramachandran outliers, rotamer quality, and steric clashes. Each criterion is checked independently. A structure can, in principle, pass every individual check while containing regions where the constraints collectively produce high total energy in the constraint network. The Dirichlet energy of the sheaf provides a global measure of constraint inconsistency that aggregates across the entire network, complementing per-criterion checks with a single scalar that reflects total structural deviation. Whether the sheaf's global

energy measure catches failures that per-criterion tools miss on real-world structures is an open empirical question.

The distinction is analogous to checking that every equation in a system of equations has a plausible form versus checking whether the system of equations has a consistent solution. MolProbity checks the equations one at a time. Sheaf energy analysis checks whether they can all be true simultaneously.

A direct quantitative comparison, evaluating SATYA's verdict and violation localization against MolProbity's clashscore, Ramachandran outliers, and rotamer quality on a shared set of known pathological PDB entries, is a priority for future work. The current validation uses synthetic decoys with known ground truth, which is appropriate for establishing detection and localization accuracy of the engine itself. Demonstrating that SATYA catches failures on real-world structures where MolProbity passes (or vice versa) requires a separate empirical study that is beyond the scope of this preprint.

An important asymmetry should be noted: SATYA operates on Cα traces only, while MolProbity evaluates all-atom geometry including rotamer quality, side-chain clashes, and Ramachandran outliers. A structure can have a correct Cα backbone ($Q = 1.0$, low $\mathcal{E}$) while containing side-chain errors that MolProbity would catch but SATYA would miss. SATYA is therefore a backbone topology and contact network verifier, not a replacement for all-atom geometric validation. The two tools address complementary levels of structural description.

4.4 PDB Validation Reports

PDB validation reports aggregate MolProbity-style metrics with additional checks (geometry, B-factor analysis, data-model agreement). Like MolProbity, these are per-criterion checks without global consistency analysis. SATYA complements PDB validation by adding a global energy layer: a structure that passes all PDB validation criteria can still fail sheaf energy verification if its constraints are collectively inconsistent.

4.5 Summary Comparison

Table 7 summarizes the key differences between SATYA and existing validation tools.

Table 7: Comparison of structural validation approaches.

	SATYA	lDDT	pLDDT	MolProbity	WHAT_CHECK
Input	Cand. + ref.	Cand. + ref.	Prediction	Structure	Structure
Resolution	Cα + dihedrals	All-atom	Per-residue	All-atom	All-atom
Deterministic	Yes	Yes	Yes	Yes	Yes
Training data	None	None	Required	None	None
Chiral detection	Yes	No	No	Yes	Yes
Global energy	$\mathcal{E}$, Q	No	No	No	No
Crypto audit	Ed25519	No	No	No	No

5 Discussion

5.1 Scope and Limitations

SATYA Protein currently verifies candidate structures against a known native reference. This is directly applicable to: predicted structures verified against experimental references, MD simulation outputs verified against starting structures, homology models verified against templates, and refined structures verified against deposited coordinates.

The current formulation does not address *de novo* verification (assessing a structure with no reference). Extending the sheaf construction to use physics-based constraints (expected bond lengths, Ramachandran distributions, hydrogen bond geometry) as the reference, rather than a specific native structure, is a natural next step. This would enable reference-free verification where the "reference" is the physical law itself.

Intrinsically disordered regions (IDRs) present a related limitation. When no stable reference structure exists for a region, the current constraint graph cannot be constructed. For partially disordered proteins, ensemble references (e.g., multiple NMR conformers) could supply the constraint set, but fully disordered regions without any structural reference remain out of scope.

The natural extension is reference-free verification, where the reference sheaf is constructed from physics-based constraints (ideal bond lengths, Ramachandran basin boundaries, steric radii) rather than from an experimental structure. This would allow verification of novel predictions without requiring an existing PDB entry. The current reference-based tool establishes the necessary mathematical foundation: it demonstrates that the sheaf engine correctly detects and localizes structural violations when ground truth is available. Extending to physics-based reference sheaves is an architectural change, not a mathematical one. In the reference-based mode, the stalks encode deviation from a specific PDB structure. In the reference-free mode, the stalks would encode deviation from idealized geometric basins (Ramachandran favored regions, ideal bond lengths, expected hydrogen bond geometries). The coboundary operator, Laplacian construction, and energy computation remain identical; only the constraint extraction module changes. This modularity is a deliberate architectural feature: SATYA is a general constraint network verification framework, not a metric specific to reference-based comparison.

Reference validity. SATYA assumes the reference structure is the intended ground truth. If the reference is a distant homolog (e.g., verifying a human protein against a mouse template), insertion loops and species-specific contacts will be flagged as violations even when they are biologically correct. Similarly, verifying an MD trajectory frame against the starting structure will flag normal conformational dynamics as violations. SATYA's verdicts are mathematically correct in these cases (the candidate genuinely violates the reference constraints), but biologically misleading. Users must ensure that the reference structure represents the intended target conformation. The conformation-specific verification case study (Section 3.8) illustrates this directly: verifying AlphaFold2's prediction against the *alternative* fold produces `UNSAFE` because the reference is a different conformation, not because the prediction is defective.

SATYA requires a complete Cα trace for both the candidate and reference structures. Missing residues (common in experimental structures due to unresolved density) must be excluded from both structures before verification; constraints cannot be computed for residues without coordinates. For structures with extensive missing regions, the constraint graph covers only the resolved portion, and violations in unresolved regions are not detectable.

5.2 Validation Scope

The controlled perturbation benchmark (Section 3.1) uses synthetic decoys with known ground truth. This establishes that the engine correctly detects and localizes controlled corruptions with maximal signal-to-noise ratio. The conformation-specific verification case study (Section 3.8) extends validation to real AlphaFold2 predictions, demonstrating that SATYA detects structural divergence on actual predicted structures where pLDDT does not flag the conformation mismatch. Together, these two validation modes cover engine correctness (synthetic decoys) and real-world relevance (fold-switching predictions).

The conformation-specific case study is preliminary: three proteins, one comparison per protein. A comprehensive evaluation would require systematic screening of the full Thacker (2026) dataset (27,098 predictions across 18 fold-switching proteins [12]), including both the high-confidence mispredictions and the correctly predicted conformations. Additional targets include AlphaFold2/OpenFold predictions with known experimental discrepancies from CASP evaluations, and side-by-side comparison with MolProbity on the same structures. A critical open question is whether non-identity restriction maps encoding physically meaningful inter-residue coupling (e.g., Ramachandran-basin constraints, bond-angle propagation) would produce detectable cyclic incompatibilities on real biological cases where geometric metrics (Q, $\mathcal{E}$) are ambiguous. This would demonstrate that the obstruction detection pathway provides biological insight beyond what standard energy analysis offers.

The near-native sensitivity characterization (Section 3.4) establishes that SATYA correctly classifies thermal-scale perturbations (0.25 Å) as SAFE with zero false positives, and that the SAFE/UNSAFE transition occurs at 0.5–1.0 Å noise, corresponding to the onset of native contact disruption. The mirror-image experiment (Section 3.5) demonstrates that SATYA detects a class of structural errors (chiral inversions) that pairwise distance metrics, including lDDT, mathematically cannot detect.

5.3 Future Obstruction Pathway: Non-Identity Restriction Maps

The current constant-sheaf implementation uses identity restriction maps, and all detection operates through the geometric energy pathway (Dirichlet energy, Q, severity, dihedral deviation). The framework is designed to support a future extension using non-identity restriction maps that encode physically meaningful inter-residue coupling. With such maps, cyclic incompatibility in the constraint network could produce obstructions not detectable by the current energy pathway.

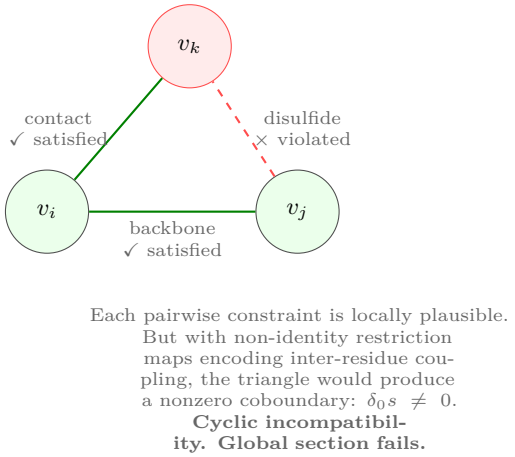

Figure 1: Conceptual illustration of a future obstruction pathway. The current SATYA Protein implementation uses identity restriction maps and detects failures through the geometric energy pathway. A future version could introduce non-identity restriction maps encoding physically meaningful inter-residue coupling, then evaluate whether local compatibility constraints can be assembled consistently around cycles. This figure illustrates the intended extension, not a capability evaluated in the present work.

A critical open question is whether non-identity restriction maps would produce detectable cyclic incompatibilities on real biological cases where geometric metrics (Q, $\mathcal{E}$) are ambiguous. Demonstrating this would establish that the obstruction detection pathway provides biological insight beyond what standard energy analysis offers.

5.4 Generality of the Constraint Graph Framework

While this work focuses on protein structure verification, the underlying mathematical framework is domain-agnostic: verifying global consistency of local constraints via sheaf energy over a constraint graph applies to any system where pairwise relationships must assemble into a coherent global state. The protein application demonstrates that the engine handles real-world, high-dimensional physical data with sub-second verification times. The general engine benchmarks (5 million vertices, 35 μs median per-edit latency) establish that the infrastructure scales to constraint networks far beyond any current biological application. The mathematical foundations and stability analysis of the general engine are reported in the companion manuscript (Zenodo DOI: 10.5281/zenodo.19598076).

5.5 Extensions: Mutation Analysis and Ligand Interactions

The sheaf formulation suggests natural extensions beyond reference-vs-candidate verification. For mutation analysis, a single residue substitution can be introduced into the constraint graph and the energy recomputed to determine whether the mutation introduces constraint violations in a previously stable fold. This could provide a rapid screen for destabilizing mutations without requiring full MD simulation.

For drug binding and ligand interactions, the constraint graph could be extended to include ligand-protein contact edges, enabling verification of whether a binding pose is globally consistent with the surrounding protein geometry. A docked ligand that satisfies all local interaction criteria but creates high constraint

energy in the broader network would be flagged by the sheaf energy where conventional scoring functions would not. Both extensions are straightforward within the existing architecture and are targets for future work.

5.6 When to Verify

The use cases for SATYA map to specific decision points in structural biology workflows:

- **Before synthesis.** Rank predicted candidates by structural consistency against a reference. Flag inconsistent structures before committing to wet-lab synthesis, saving reagent cost and experimental time.
- **Before partner review.** Attach a signed verification receipt to every structure in a collaboration or licensing package. The receipt provides independent evidence of structural quality that does not depend on the predictor's own confidence score.
- **Before publication.** Include SATYA receipts as supplementary verification artifacts alongside deposited coordinates, providing reviewers and readers with independent structural validation.
- **Before downstream ML.** Filter inconsistent structures from training sets, databases, and prediction pipelines. A structure that fails sheaf energy verification should not train the next generation of predictors.

5.7 Integration with Prediction Pipelines

SATYA is designed to operate as a post-prediction verification gate. The workflow is:

1. A structure prediction system (AlphaFold2, OpenFold, RoseTTAFold, ESMFold) generates a candidate structure with pLDDT confidence scores.
2. SATYA verifies the candidate against the experimental reference (if available) or a physics-based constraint set (future work).
3. The output is a signed receipt: the structure is either verified consistent or the receipt documents exactly where and how it fails.

This turns "predicted" into "predicted and independently verified," with a cryptographic artifact suitable for downstream audit. For high-throughput prediction pipelines processing thousands of structures, the near-linear scaling and sub-second verification times mean SATYA can operate as an inline filter without becoming a bottleneck.

5.8 Live Reproducibility

An interactive demo at `https://invariant.pro/protein` allows readers to inspect the behavior of the SATYA Protein verifier on the public demo cases. Each demo case returns the verification result: verdict, RMSD where applicable, native contact Q, Dirichlet energy where applicable, global coherence where applicable, violation regions with per-residue severity scores, and a downloadable signed receipt in JSON format. The demo runs the SATYA Protein service rather than a static screenshot or narrative summary. This provides public, receipt-based inspection of the reported verification behavior, although the full proprietary engine source is not released as open-source in the current public repository.

5.9 Reproducibility

SATYA is deterministic. Given the same input structures and engine version, it produces the same sheaf construction, the same verification metrics, the same violation regions, and the same signed receipt payload apart from timestamp and signature fields. There is no stochastic initialization, no learned parameter update, and no model sampling in the verification step.

The public demo provides receipt-based reproducibility for the included demonstration cases. Full command-line reproduction from local source requires access to the SATYA Protein engine and bundled benchmark inputs. The current public repository documents the project materials and reported results; the full proprietary engine source is not released under an open-source license in this preprint version.

5.10 Mathematical Foundations

The sheaf energy engine underlying SATYA Protein has formal stability guarantees reported in a companion manuscript (Zenodo DOI: 10.5281/zenodo.19598076; submitted to the ICML 2026 AI4Math Workshop). That manuscript establishes that the engine maintains structural integrity as the underlying constraint graph scales, and that the energy signals remain reliable under repeated updates. The protein-specific verifier described here is evaluated independently of those theoretical results.

6 Conclusion

We presented SATYA Protein, a deterministic protein structure verification engine based on constraint graph sheaf energy. The engine uses a constant sheaf (identity restriction maps) over the residue constraint graph, with a contact-first verdict logic that decouples fold integrity (native contact fraction Q, per-residue severity, backbone dihedral deviation) from deviation field smoothness (Dirichlet energy, coherence). It localizes failures to specific residue regions and produces cryptographically signed receipts on every verification.

Near-native sensitivity analysis shows zero false positives at thermal-scale noise (0.25 Å), with the SAFE/UNSAFE transition at 0.5–1.0 Å corresponding to the onset of native contact disruption. A mirror-image experiment demonstrates that SATYA detects chiral inversions that lDDT (lDDT = 1.0 on all mirrors) mathematically cannot detect, via signed dihedral encoding in the six-dimensional stalk. Direct comparison with lDDT confirms that SATYA augments distance-based verification with chirality detection. Conformation-specific verification on three fold-switching proteins (KaiB, RfaH, XCL1) demonstrates that SATYA detects structural divergence in real AlphaFold2 predictions where pLDDT, which is not designed as a conformation-specific verifier, does not identify the mismatch.

The mathematical framework supports extension to non-identity restriction maps that could enable detection of cyclic incompatibilities beyond the current energy pathway. The current constant-sheaf implementation validates the energy landscape, scaling properties, and audit framework; the obstruction detection pathway is the next extension.

Live demo, documentation, and project materials: https://invariant.pro/protein.

Acknowledgments

Portions of the manuscript were refined with assistance from AI for clarity, LaTeX formatting, and structural framing. The author takes full responsibility for all content and mathematical claims.

Competing Interests

The author is the developer of SATYA Protein and operates the public demo at https://invariant.pro. No external funding was received.

Funding

No external funding was received for this work.

Author Contributions

J.L.V. conceived the method, implemented the software, performed experiments, analyzed results, and wrote the manuscript.

Data and Code Availability

The public project materials, benchmark summaries, documentation, and release metadata for SATYA Protein are available at the project repository (https://github.com/Jasonleonardvolk/sigma). The repository is provided under the license terms stated in that repository. At the time of this preprint, the public repository is not an MIT-licensed open-source release of the full SATYA verification engine.

A live interactive demo is available at https://invariant.pro/protein. The demo allows readers to run the reported verification cases against the SATYA Protein service and returns the verification result, including verdict, native contact fraction Q, violation regions, and a downloadable signed receipt. The demo is intended to support public inspection and reproducibility of the reported verification behavior, while preserving the proprietary implementation of the engine.

The companion technical report for the underlying sheaf engine is archived at Zenodo (DOI: 10.5281/zenodo.19598076). Access to the full engine source code, decoy generation scripts, and commercial or research licensing terms may be requested from the author.

References

[1] Jumper, J., Evans, R., Pritzel, A., et al. Highly accurate protein structure prediction with AlphaFold. *Nature*, 596, 583–589, 2021.

[2] Baek, M., DiMaio, F., Anishchenko, I., et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science*, 373, 871–876, 2021.

[3] Lin, Z., Akin, H., Rao, R., et al. Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science*, 379, 1123–1130, 2023.

[4] Ahdritz, G., Bouatta, N., Floristean, C., et al. OpenFold: Retraining AlphaFold2 yields new insights into its learning mechanisms and capacity for generalization. *Nature Methods*, 21, 1514–1524, 2024.

[5] Chen, V.B., Arendall, W.B., Headd, J.J., et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica D*, 66, 12–21, 2010.

[6] Hooft, R.W.W., Vriend, G., Sander, C., and Abola, E.E. Errors in protein structures. *Nature*, 381, 272, 1996.

[7] Hansen, J. and Ghrist, R. Toward a spectral theory of cellular sheaves. *Journal of Applied and Computational Topology*, 3, 315–358, 2019.

[8] Curry, J.M. Sheaves, cosheaves and applications. PhD thesis, University of Pennsylvania, 2014.

[9] Lahey, S., Barbe, S., and Grudinin, S. Fold or flop: quality assessment of AlphaFold predictions on whole proteomes. *bioRxiv*, 2025.12.19.695427v2, 2026.

[10] Ruff, K.M. and Pappu, R.V. AlphaFold and implications for intrinsically disordered proteins. *Journal of Molecular Biology*, 433, 167208, 2021.

[11] Best, R.B., Hummer, G., and Eaton, W.A. Native contacts determine protein folding mechanisms in atomistic simulations. *Proceedings of the National Academy of Sciences*, 110, 17874–17879, 2013.

[12] Thacker, R. Confidence without verification: screening pLDDT unreliability in AlphaFold2 fold-switching predictions. *bioRxiv*, 2026.02.19.706878, 2026.

[13] Wayment-Steele, H.K., Ojoawo, A., Otten, R., et al. Predicting multiple conformations via sequence clustering and AlphaFold2. *Nature*, 625, 832–839, 2024.

[14] Mariani, V., Biasini, M., Barbato, A., et al. lDDT: a local superposition-free score for comparing protein structures and models using distance difference tests. *Bioinformatics*, 29, 2722–2728, 2013.

[15] Kunzmann, P. and Hamacher, K. Biotite: a unifying open source computational biology framework in Python. *BMC Bioinformatics*, 19, 346, 2018.