

BIOMOL EVIDENCE ALPHA v1.4

Protected biomolecular evidence workflows for Structural Triage and MD Evidence

Controlled researcher-alpha methodology for protected, inspectable evidence packages.

Capability	Status
Pilot keys and first-use activation	Passed
Usage counters and quotas	Passed
Protected Structural Triage and MD Evidence results	Passed
External non-AdK NhaA MD Evidence gate	Passed

Biomol Evidence Alpha White Paper

Protected biomolecular evidence workflows for structural triage and molecular-dynamics review

Executive summary

Biomol Evidence Alpha is a controlled researcher-alpha platform for protected, inspectable biomolecular evidence workflows.

It currently supports two operational functions:

1. Structural Triage
2. MD Evidence

A third function, Integrated Evidence Dossier, is planned as the synthesis layer.

The platform is designed to help researchers move from raw protein-structure panels and compatible MD trajectories to structured evidence packages with provenance, computed metrics, artifact bundles, manifest files, and checksum verification.

Biomol Evidence Alpha is not positioned as a black-box claim engine. It is an evidence organization, review-prioritization, and artifact-generation platform for expert scientific review.

1. Problem

Protein researchers often face two related operational problems.

First, large structure panels can contain hundreds or thousands of protein targets. Manual inspection of every predicted or available structure is slow, inconsistent, and difficult to document.

Second, molecular-dynamics trajectories can be difficult to summarize and share in a consistent evidence format. Researchers may compute trajectory metrics locally, but collaborator-ready evidence packages often require provenance, file identity, reproducible metrics, manifests, and checksums.

The practical bottleneck is not only computation. It is the absence of a governed evidence workflow.

2. Function 1: Structural Triage

Structural Triage helps researchers prioritize protein-structure panels for manual review.

Inputs

- CSV with `uniprot_accession`
- optional `target_id`

- optional category
- optional notes

Outputs

- ranked target table
- review action
- triage score
- confidence/range evidence
- source provenance
- artifact bundle
- manifest
- checksum file

Current protected-result behavior

Registered Structural Triage results are owned by the researcher key that created them.

Result pages and artifact downloads require the correct researcher key.

Current pilot scope

- guided first review: 100-250 targets
- standard pilot: up to 1,000 targets
- larger runs: scheduled request

3. Function 2: MD Evidence

MD Evidence helps researchers convert compatible topology/trajectory packages into protected evidence summaries.

Inputs

- topology file
- trajectory file
- metadata JSON or CSV

Confirmed format paths

- PSF + DCD
- PDB + DCD
- GRO + XTC

Outputs

- file identity

- SHA-256 hashes
- parser status
- frame count
- atom count
- residue count
- RMSD summary
- RMSF summary
- radius-of-gyration summary
- evidence bundle
- manifest
- checksum file

Current protected-result behavior

Registered MD Evidence results are owned by the researcher key that created them.

Result pages and artifact downloads require the correct researcher key.

Current pilot scope

- compatible small-to-medium trajectories
- larger trajectory support by coordination

4. Pilot access model

Biomol Evidence Alpha uses controlled pilot keys.

A standard pilot key currently provides:

- 20 jobs
- 30 days from first registered job
- up to 1,000 Structural Triage targets
- up to 500 MB MD Evidence upload quota

The pilot access page shows:

- jobs remaining
- Structural Triage target quota remaining
- MD Evidence upload quota remaining
- days remaining
- access status

This creates a controlled, auditable pilot model before broader researcher exposure.

5. Protected result model

Every registered job belongs to the researcher key that created it.

Protected surfaces include:

- job result pages
- result JSON
- ranked tables
- metric outputs
- evidence cards
- manifest files
- checksum files
- artifact downloads

The system behavior is:

- correct key: allow
- missing key: deny
- wrong key: deny
- unused legacy/demo records: handled as public technical archive when explicitly framed

6. Evidence bundle model

Biomol Evidence Alpha is organized around evidence bundles rather than one-off scores.

A typical evidence bundle includes:

- result JSON
- CSV tables
- Markdown evidence card
- manifest
- SHA-256 checksum file

This supports collaborator review, reproducibility checks, and institutional auditability.

7. Function 3: Integrated Evidence Dossier

Integrated Evidence Dossier is the planned synthesis layer.

It will combine:

- an owned Structural Triage result
- an owned MD Evidence result
- target identity
- project metadata
- optional researcher notes

Planned outputs:

- PDF dossier
- JSON evidence package
- provenance summary
- computed metrics
- review checklist
- artifact links
- manifest
- checksums
- scope notes

The goal is to give researchers and reviewers one coherent evidence brief instead of separate technical outputs.

8. Validation status

Current operational validation includes:

- Structural Triage registered run
- MD Evidence registered run
- protected result access
- protected artifact downloads
- pilot access status
- usage counters
- admin export
- post-reboot service recovery
- public-language live page audit

Historical technical records are retained in the Technical Archive.

9. External validation plan

The next MD Evidence credibility step is one external non-AdK trajectory.

The proposed candidate is the NhaA equilibrium dataset from MDAnalysisData. It is a public membrane-protein trajectory with GRO topology and XTC trajectory format. For pilot testing, a small chunk should be created from the larger dataset and run through the protected MD Evidence workflow.

A successful or clearly characterized failure case should produce:

- protected MD Evidence result
- parser status
- frame/atom/residue counts
- RMSD/RMSF/radius-of-gyration summaries when parseable
- evidence bundle

- manifest
- checksum file
- validation note

10. Scope

Biomol Evidence Alpha organizes computational evidence for expert review.

Scientific interpretation, biological conclusions, therapeutic relevance, or regulated use require separate expert review and independent validation.

The platform is designed to support review, not replace scientific judgment.

External NhaA gate

MD Evidence passed an external non-AdK validation gate using a public NhaA membrane-protein trajectory package.

- Job ID: `stage2_20260502_090924_4r3swx6m`
- Format path: GRO + XTC
- Frames analyzed: 20
- Atoms: 60,702
- Residues: 1,314
- Protein selection atoms: 11,624
- RMSD mean Å: 2.759966
- RMSF mean Å: 2.064565
- Radius of gyration mean Å: 31.295446
- Parse status: parsed
- Errors: none
- Warnings: none

This gate validates protected job creation, parsing, descriptive metric computation, and evidence packaging for a non-AdK trajectory path.