IntuitionLabs

INTUITIONLABS / RESEARCH & PRACTICAL GUIDANCE

Claude Biomolecular Modeling: Validation & Deployment

> Build practical AI for pharma and biotech with IntuitionLabs. We help life-science teams turn complex information and workflows into useful software, governed knowledge systems and AI tools.

Published by IntuitionLabs · 2026-09-19 · claude biomolecular modeling · claude science · drug discovery · protein structure prediction · flashpairformer · biomolecular modeling · pharma ai · model validation · gpu inference

Original article: https://intuitionlabs.ai/articles/claude-biomolecular-modeling-validation-deployment

In this article

Executive Summary

Anthropic released **36 open-source inference-optimization kits** on September 17, 2026, covering cofolding and structure prediction, structure generation, inverse folding, protein language models, and genomics. The repository is a reference release, not a new biomolecular foundation model, and is explicitly described as **not maintained** (github.com). Anthropic reports that Claude helped optimize more than 30 models in under four weeks (anthropic.com). Those development-effort and performance observations should be treated as vendor evidence to reproduce, not as an industry productivity benchmark.

The headline results separate two operating points. **Exact mode** reportedly averaged **1.6x faster** across 14 structure-prediction models on NVIDIA H100 GPUs, while **Fast mode** averaged **4.1x faster** across the 13 models that offered it (www-cdn.anthropic.com) (www-cdn.anthropic.com). Across **1,925 model-target pairs**, pooled acceptable-interface share changed by less than one percentage point in every tested mode (www-cdn.anthropic.com). This does not prove equivalence on a sponsor's sequences, workload mix, or hardware.

Big mode is a memory and sharding capability, not a claim of improved scientific generalization. It enabled accurate predictions above **10,000 tokens** on one eight-GPU node, but all seven capability runs between roughly 31,000 and **70,320 residues** completed without producing an accurate structure (www-cdn.anthropic.com) (www-cdn.anthropic.com). A pilot should therefore test Exact first, promote Fast only after endpoint-specific non-inferiority checks, and reserve Big for inputs that fail memory feasibility in other modes.

Deployment is feasible but not turnkey. Each kit pins its own stack and generally supplies a Dockerfile, Apptainer definition, and requirements lockfile. All kits are documented against an **H100 80 GB on Linux x86-64**, with driver floors varying from 525 to 580 (github.com). No independent reproduction of the new speed claims was located by the September 19 evidence cutoff. The go decision should require a pinned commit, stock-versus-optimized paired runs, blind holdout data, confidence intervals, scientific endpoint thresholds, resource telemetry, provenance, and rollback. That risk-based approach aligns with joint FDA and EMA principles calling for validation proportionate to context of use (fda.gov).

Introduction and Background

For teams evaluating **generative AI validation in pharma**, the orchestration layer and the scientific model require separate evidence.

“Claude biomolecular modeling” can refer to three different layers that should not be conflated. First, **Claude Science** is a beta workbench, not a separate model (claude.com). Second, Claude can assist scientists with literature synthesis, code, workflow orchestration, and tool use. Third, the new repository contains model-specific code paths that accelerate established specialist models. Official materials describe the workbench as the place where specialized tools work together (claude.com), not as a replacement for structure predictors or design engines.

This distinction matters for **Claude AI for drug discovery**. Claude may coordinate a Boltz-2 prediction, help inspect outputs, or generate analysis code, while the upstream biomolecular model still produces the molecular prediction. Anthropic says Claude Science can connect to more than 60 scientific databases and domain-specific open models (claude.com). Anthropic separately lists literature review, hypothesis generation, genomic analysis, and regulatory-drafting support among life-sciences workflows (anthropic.com). None of these statements turns an in-silico score into evidence of drug efficacy.

For an adjacent consultancy such as IntuitionLabs, the appropriate role is governance and implementation advice, not a competing product row. Its published operating approach emphasizes governed workflows and scaling from observed evidence (intuitionlabs.ai) and measuring quality, reliability, risk signals, and support burden before expansion (intuitionlabs.ai). That posture fits this report's central recommendation: pilot the code as a controlled infrastructure change, and validate its scientific impact separately from its runtime improvement.

Key Changes

An optimization portfolio, not one molecular model

The release packages **36 drop-in kits** for inference paths in open protein and genomics machine-learning tools. The report groups them into six families: 14 cofolding or structure-prediction packages, three hallucination packages, six structure-generation packages, three inverse-folding packages, seven genomics packages, and three protein-language packages (www-cdn.anthropic.com). This breadth makes the repository more like a collection of adapters than a unified application.

The portfolio spans materially different scientific jobs:

- **Complex structure prediction:** Boltz-2, Chai-1, OpenFold3, Protenix and related kits predict biomolecular structures or interfaces.
- **Protein design:** Structure-generation tools create candidates, while inverse-folding tools propose sequences for candidate backbones.
- **Inverse folding:** ESM-IF1 and ProteinMPNN map structure to sequence candidates. The original ProteinMPNN study reported **52.4% native sequence recovery**, versus 32.9% for Rosetta, on its own benchmark (pmc.ncbi.nlm.nih.gov). That upstream result is not evidence about Anthropic's optimized kit.

- **Genomics:** The collection covers long-context nucleotide modeling and base-resolution chromatin-accessibility prediction.

Exact, Fast, and Big are distinct scientific choices

Exact mode aims to preserve stock outputs. That is the lowest-friction pilot path because it isolates implementation performance from output drift. **Fast mode** allows small documented numerical differences within a tool's own seed-to-seed variation (github.com). A team still must show that aggregate scientific decisions remain stable. “Within seed variation” is a vendor-defined engineering criterion, not a sponsor-specific acceptance criterion.

Big mode adds the memory machinery used by Fast and may shard pair representations across GPUs on one host (github.com). It can make a previously infeasible input runnable, but feasibility and accuracy are separate gates. The report itself says Big extends what can be computed, not what the models learned.

FlashPairformer changes the kernel layer

Pairformer is a neural-network block operating on single and pair representations. In AlphaFold 3, it replaced Evoformer and works without the multiple-sequence-alignment representation used by Evoformer (nature.com). **FlashPairformer** is Anthropic's name for its Pairformer kernel family, including CUDA-native triangle attention and triangle multiplication as well as portable implementations (github.com).

Vendor benchmarks report **2.7x** average acceleration for triangle-attention kernels at the common pair width (www-cdn.anthropic.com). At pair width 256, they report **2.9x** for triangle attention and **3.2x** for triangle multiplication (www-cdn.anthropic.com). The measurement was a geometric mean over seven input sizes from 256 to 2,048 tokens (www-cdn.anthropic.com), with medians from at least 20 CUDA-graph replays (www-cdn.anthropic.com). These are kernel microbenchmarks, not end-to-end prediction throughput, and should be reproduced independently on the exact software and GPU stack used by the pilot.

Repository, Models, and Mode Selection

Each kit carries a pinned stock upstream release, patches or wrappers, environment files, tests, and kit-specific instructions. Anthropic's new code is Apache License 2.0 (github.com), but upstream code and model weights retain their own terms. A legal and scientific intake must therefore review both the kit license and the upstream model or weight license.

Table 1 maps representative kits to their scientific context. It is not an exhaustive list of all 36 packages.

Kit and pin	Scientific task	Modes and reference hardware	License evidence and pilot note
Boltz-2 2.2.1	Complex structure and interaction prediction	Exact, Fast, Big; Big documented up to 3,000 tokens on one H100 80 GB (github.com)	Upstream model and code are MIT licensed (github.com); validate each required artifact.
Chai-1 0.6.1	Complex structure prediction with optional restraints	Exact, Fast, Big; upstream requires CUDA and bfloat16 support (github.com)	Apache 2.0 code and weights; compare all generated samples.

Kit and pin	Scientific task	Modes and reference hardware	License evidence and pilot note
OpenFold3 0.4.1	Biomolecular structure prediction	Exact, Fast, Big; Linux x86-64 and CUDA 12.8-capable driver (github.com)	Apache 2.0; the Python route also needs CUDA toolkit 12.8 and a compiler.
ProteinMPNN, commit 8907e667	Inverse folding and sequence design	Exact only; no Fast, Big, or multi-GPU route (github.com)	MIT upstream; use task-specific recovery and designability endpoints.
Evo 2 0.6.0	DNA scoring, inference, and generation	Exact and Fast; pinned 7B uses one device and 40B uses two (github.com)	Apache 2.0 components; 40B deployment should be capacity-tested separately.
ChromBPNet 1.0.1	Base-resolution chromatin accessibility	Exact and Fast; Fast is the default	MIT upstream; the kit covers only the `pred_bw` subcommand (github.com).

The matrix shows why there is no single installation or validation answer. Even within one repository, supported modes, driver floors, memory ceilings, endpoint types, and licensing differ. A program should approve a small bill of materials per kit rather than approve the repository as one undifferentiated component.

Implementation Considerations and Process Changes

Deployment paths and environment pinning

The repository documents **Docker, Apptainer, and Python virtual-environment** routes. As of the evidence cutoff, users had to build containers from each kit's definition because prebuilt images were only planned (github.com). That raises the importance of recording the base-image digest, package locks, compiler, Torch, CUDA, CPython, driver, GPU SKU, firmware, and model-weight hashes.

Recommended deployment controls are:

- **Pin by immutable identifier.** Docker recommends selecting base images by digest to guarantee the same version on rebuild (docs.docker.com). Record the Anthropic commit and every upstream commit or tag.
- **Generate provenance and a software bill of materials.** Docker BuildKit supports software bill of materials attestations through a Supply-chain Levels for Software Artifacts-compliant process (docs.docker.com). SLSA provenance describes where, when, and how an artifact was produced (slsa.dev).
- **Verify container identity.** Apptainer can verify a signed SIF image as a bit-for-bit reproduction of the signed original (apptainer.org). Cosign also verifies that a signature payload's digest matches the container (docs.sigstore.dev).
- **Protect build secrets.** Docker advises secret mounts rather than build arguments because secret mounts are excluded from provenance (docs.docker.com).
- **Fail closed.** The kits say that an optimization that cannot engage fails by name rather than silently reverting to stock. Preserve this behavior in schedulers and alerting.

Claude Science and pharmaceutical data boundaries

Claude Science can run locally or on a remote machine over Secure Shell, and it can submit jobs to a laboratory high-performance computing cluster (anthropic.com). Its artifacts can include the exact code and environment used (anthropic.com), which is useful evidence but does not by itself validate a result.

Teams must review the separate data plane. Claude Science stores project history locally, but prompts can include file contents, code output, connector results, images, and memory. Anthropic says running beside the data keeps datasets on controlled infrastructure except for content Claude reads into the conversation (claude.com). Covered-model requests are retained for 30 days (claude.com). Remote jobs execute outside the sandbox with the user's host permissions (claude.com). These facts should drive network, credential, data-classification, and least-privilege controls.

Reproduction Protocol and Scientific Validation

The goal is to measure whether the optimized path preserves an agreed scientific decision while improving resources. The protocol should be preregistered before examining optimized outputs. NIST recommends evaluating capability claims through empirically validated methods (nvlpubs.nist.gov).

A paired reproduction design

1. **Freeze scope.** Name one kit, upstream pin, Anthropic commit, mode, command, task, and context of use. Do not mix kit results into one go decision.
2. **Create three data partitions.** Use a smoke-test set, a development set for engineering, and a blind holdout. Preserve biological separation by sequence identity, family, structure, or time as appropriate. Different partition strategies test distinct generalization regimes (frontiersin.org).
3. **Pin execution.** Record host, GPU, driver, CUDA, libraries, precision, batch size, random seeds, recycling and sampling settings, environment variables, hashes, and nondeterminism. Protein-ML reproducibility guidance specifically calls for hardware context, numerical precision, seeds, and runtime parameters (frontiersin.org).
4. **Run stock and optimized paths in pairs.** Randomize order where thermal or shared-cluster effects could bias latency. Use warm-up runs, then enough repetitions to estimate dispersion. End-to-end sample size should be based on observed variance.
5. **Measure the full workload.** Capture wall time, GPU time, peak allocated and reserved memory, energy if available, CPU and I/O time, preprocessing, model initialization, database search, retries, failures, and output count. A review found **72%** of surveyed benchmarking studies omitted computational resources and runtime, a useful warning against reporting acceleration without the denominator (nature.com).
6. **Evaluate scientific endpoints.** For interfaces, DockQ combines fraction of native contacts, ligand root-mean-square deviation, and interface root-mean-square deviation into one continuous measure (journals.plos.org). Anthropic uses DockQ above 0.23 as acceptable (anthropic.com), but teams should also inspect continuous shifts, confidence, calibration, ranking stability, and task-specific downstream metrics.

7. **Estimate uncertainty.** Report paired differences with confidence intervals, distributions, and per-stratum results. FDA's draft framework says all performance estimates should include confidence intervals (fda.gov). Do not infer equivalence merely because a difference is not statistically significant.
8. **Lock and rerun.** Rebuild from provenance, rerun a subset, archive outputs, and demonstrate a switch back to stock. Test data should remain independent of development data (fda.gov).

Context-of-use gates

Table 2 scales the evidence burden to how a result will be used.

Context	Permissible role	Minimum evidence	Release gate
Engineering exploration	Test installation, throughput, and feasibility on public or non-sensitive data	Exact-mode smoke tests, environment hash, stock comparison, resource telemetry	Reproducible build and no silent fallback
Exploratory research	Prioritize hypotheses or candidates for further computational and laboratory work	Blind holdout, multiple seeds, paired endpoint distributions, expert review, documented uncertainty	No material degradation on predefined metrics; all results labeled in silico
Research decision support	Influence experimental allocation or portfolio choices	Independent validation owner, representative data, stratified thresholds, provenance, monitoring, rollback	Signed protocol and acceptance report; deviation handling
Regulatory evidence support	Contribute to evidence entering a drug-development decision	Formal context of use, fit-for-use data, risk assessment, traceable processing, version control, lifecycle plan	Engage relevant quality and regulatory functions; meet applicable requirements

The regulatory boundary is important. FDA's January 2025 draft guidance is non-binding and specifically excludes AI used in drug discovery (fda.gov). EMA likewise notes that discovery use can have low regulatory impact when suboptimal performance affects only the developer (ema.europa.eu). Once an output informs regulated evidence or a high-impact decision, the expectations become proportionate to that role.

Data Analysis and Evidence

Anthropic's principal structure benchmark used **FoldBench-Lite**, which combines part of FoldBench with more recent Protein Data Bank entries (www-cdn.anthropic.com). Across 13 model configurations and 1,925 target pairs, the vendor reports less than a one-percentage-point change in acceptable-interface share and no pooled change statistically distinguishable from zero (www-cdn.anthropic.com). The confidence intervals used paired percentile bootstrap resampling with **20,000 resamples** within each model (www-cdn.anthropic.com). That is a serious evaluation design, but it remains one vendor's selection of models, targets, modes, baselines, and hardware.

Table 3 separates the reported observation from the decision a pilot must reproduce.

Vendor-reported observation	Evidence boundary	Required reproduction result
Exact averaged **1.6x** across 14 structure models	H100 80 GB, specified upstream pins, vendor baseline	Paired wall time and GPU time on the sponsor's pinned stack, with bitwise output comparison
Fast averaged **4.1x** across 13 eligible structure models	Small numeric differences permitted; pooled accuracy analysis	Endpoint non-inferiority by model, target class, size, and seed, with confidence intervals
FlashPairformer triangle attention averaged **2.7x**	Kernel forward-pass microbenchmark, batch size one	Kernel and end-to-end profiles, including preprocessing and compilation
Accurate Big runs exceeded **10,000 tokens**	Single eight-GPU node and selected examples	Predefined accuracy and confidence on representative large complexes
Capability run reached **70,320 residues**	Eight B300 GPUs; all seven very-large predictions inaccurate	Treat as memory feasibility only, not validated scientific output

The table prevents two category errors. First, throughput is not accuracy. Second, successful completion is not scientific validity. The largest-system result belongs in capacity planning, not in a validated-accuracy claim, which makes a clean temporal holdout essential.

Transparent cost and throughput worksheet

Use local prices and observed timings rather than a universal cloud price. Anthropic's report used **$5 per H200 GPU-hour** for one specific cost model, but explicitly tied that value to a Modal list price and host allowance (www-cdn.anthropic.com). It should not be generalized across providers, regions, commitments, or dates.

For each workload, calculate:

- **Predictions per GPU-hour** = successful predictions / total GPU-hours.
- **GPU cost per accepted prediction** = hourly GPU rate × GPU count × wall hours / accepted predictions.
- **Total pilot run cost** = GPU cost + CPU and storage cost + Claude token cost + engineer time + validation-review time.
- **Retry-adjusted cost** = total run cost / final accepted outputs, including failed and repeated runs in the numerator.
- **Latency reduction percent** = (stock time − optimized time) / stock time × 100.
- **Peak-memory reduction percent** = (stock peak memory − optimized peak memory) / stock peak memory × 100.
- **Decision-stability rate** = unchanged pass/fail or ranking decisions / paired cases × 100.

The worksheet should show median, interquartile range, tail latency, confidence interval, and failure count, not only the arithmetic mean. The H100 PCIe reference specification lists a maximum thermal design power of **350 W** (nvidia.com), so teams measuring energy can add electricity and cooling assumptions transparently. Do not invent inputs when the facility or cloud contract supplies them.

Pilot Architecture, Change Control, and Go or No-Go

A defensible pilot has separate orchestration, execution, evidence, and approval layers:

- **Orchestration layer:** Claude Science or an existing workflow manager can prepare commands and coordinate tools. Anthropic cautions that its reviewer checks against the execution record but does not rerun analyses (claude.com). Human and automated verification remain necessary.
- **Execution layer:** Run stock and optimized containers in isolated queues with read-only weights, least-privilege credentials, pinned digests, and bounded network access. The repository assumes trusted inputs, trusted weights, and a single-user machine or container.
- **Evidence layer:** Store manifests, hashes, logs, commands, seeds, inputs, outputs, metrics, and paired comparisons in an immutable run record. Joint FDA and EMA principles call for documentation of data provenance, processing steps, and analytical decisions (fda.gov).
- **Approval layer:** Separate the engineer who optimizes the run from the scientist or validation owner who reviews acceptance. NIST recommends distinct verification and validation roles from test and evaluation roles (nvlpubs.nist.gov).

Use this go or no-go scorecard:

- **Scope PASS:** one versioned kit, context of use, data boundary, and scientific decision are explicit.
- **Reproducibility PASS:** another operator can rebuild and rerun from stored provenance.
- **Performance PASS:** latency, throughput, and memory meet predefined thresholds with uncertainty reported.
- **Science PASS:** output and downstream decision metrics meet per-stratum acceptance thresholds on blind data.
- **Operations PASS:** monitoring, alarms, rollback, and stock fallback are exercised.
- **Security PASS:** images, weights, and binaries match approved digests; secrets and remote permissions are bounded.
- **Governance PASS:** changes have owners, review records, and retesting triggers. EMA says nontrivial stack changes in high-impact uses require re-evaluation (ema.europa.eu).
- **Rollback PASS:** stock execution can be restored without data migration or loss of traceability. NIST calls for mechanisms to supersede, disengage, or deactivate systems outside intended use (airc.nist.gov).

A no-go result can still be valuable. It can identify an unsupported driver, memory regression, unstable endpoint, licensing constraint, or workload where stock remains preferable. The point is to make that decision before optimized output enters a consequential research workflow.

Implications and Future Directions

The release suggests a practical pattern for **Claude AI pharma R&D**: use a general-purpose agent to help engineers understand and optimize specialist scientific software, while retaining specialist models, datasets, metrics, and accountable reviewers. This is more credible than treating Claude as a molecular simulator. It also means the evidence burden sits at several interfaces: Claude-generated code, the optimized kernel, the upstream model, the container, and the downstream scientific decision.

Several uncertainties should shape adoption. No independent execution of the September 17 benchmark was found. Reported speed and memory can vary with batch size, input size, and upstream version. Some Exact and Fast paths can use more memory than stock. Model-weight terms may differ from code licenses. These are normal reasons for a narrow pilot, not reasons to assume the code is unusable.

Future evidence should include third-party reproduction across A100, H100, H200, B200, and B300 systems; model-by-model accuracy distributions; energy and compilation costs; longer temporal holdouts; and maintenance commitments. OpenFold's earlier paper provides useful context: its authors reported three to four times faster inference than AlphaFold2 on a single 40 GB A100 (labsyspharm.org). It also shows why every acceleration claim needs a named baseline and environment.

Lifecycle controls should mature with context. The joint principles call for scheduled monitoring and periodic re-evaluation (fda.gov). FDA's draft framework also recommends specifying monitoring frequency and retesting triggers (fda.gov). A repository update, upstream model update, driver change, GPU migration, compiler change, new target population, or revised decision threshold can each trigger targeted revalidation.

Frequently Asked Questions (FAQs)

Is Claude itself predicting protein structures?

Not in the repository's core architecture. Claude helped generate inference optimizations for specialist open models. Claude Science can orchestrate those tools, and its official product page names Evo 2, Boltz-2, and OpenFold3 among connected models (claude.com). The specialist model produces the biological prediction.

Which mode should a pharmaceutical research team test first?

Start with **Exact** whenever available because it tests runtime improvement without intended output change. Add **Fast** only with endpoint-specific equivalence or non-inferiority criteria. Use **Big** only for workloads that exceed other modes' memory feasibility, and never infer accuracy from successful completion.

Does a DockQ score above 0.23 prove a useful drug candidate?

No. DockQ is an interface-structure quality measure. The original DockQ work optimized class boundaries at **0.23, 0.49, and 0.80** (journals.plos.org). It does not establish binding in a laboratory, safety, exposure, efficacy, or clinical utility.

Can the kits be used in a regulated pharmaceutical workflow?

Potentially, but open-source availability is not validation. The organization must define context of use, qualify infrastructure, verify data and model provenance, establish acceptance criteria, control versions, document deviations, and maintain monitoring and rollback. Part 11 applies to electronic records created or maintained under FDA record requirements, not automatically to every exploratory artifact (fda.gov).

Are the reported results independently reproduced?

No independent reproduction of the September 17, 2026 kit benchmarks was located by the September 19 cutoff. The available quantitative evidence is Anthropic's announcement, report, and repository. Upstream papers validate their own models and baselines, not the new optimizations. NIST advises against extrapolating performance from narrow or anecdotal assessments (nvlpubs.nist.gov).

Conclusion

Anthropic's Claude-assisted biomolecular modeling release is deployable enough to justify a controlled technical pilot, but not mature enough to skip reproduction. It provides a broad, model-specific optimization portfolio with concrete environment pins, Exact, Fast, and Big operating modes, and promising vendor benchmarks. It is also an unmaintained reference release without prebuilt images or independent benchmark confirmation at the evidence cutoff.

The best near-term decision is selective. Choose one or two kits tied to a high-volume workload, reproduce stock and Exact on a pinned H100 environment, then test Fast against predefined scientific endpoints. Evaluate Big as a memory-feasibility tool and keep the accurate above-10,000-token result separate from the inaccurate very-large capability runs. Report throughput per GPU-hour, retry-adjusted cost, peak memory, endpoint stability, uncertainty, and failure rates.

For exploratory work, proportionate controls may be lightweight. For outputs that influence experiments, portfolios, or regulated evidence, require independent review, traceable provenance, representative holdouts, lifecycle monitoring, and tested rollback. The central principle is simple: Claude's contribution to optimization can be valuable, but a faster biomolecular calculation becomes decision-grade only after the organization demonstrates that the calculation remains scientifically fit for its declared use.

Source links

Links cited in this article, in order of first appearance. Full addresses are provided for readers using extracted text.

1. https://github.com/anthropics/uplifting-biomolecular-modeling
2. https://www.anthropic.com/research/claude-uplifts-biomolecular-modeling
3. https://intuitionlabs.ai/articles/pharma-ai-vendor-due-diligence-checklist
4. https://intuitionlabs.ai/articles/nvidia-gpus-in-pharma-industry
5. https://www-cdn.anthropic.com/c03643714397d9d396fa1ce1794f5f9f7863a82c.pdf
6. https://intuitionlabs.ai/articles/evaluating-biology-foundation-models
7. https://www.fda.gov/media/189581/download

8. https://claude.com/product/claude-science?_bhlid=ebdfd5658eeeba69a7b8efa414dfe4d6d10e462f
9. https://intuitionlabs.ai/articles/claude-healthcare-life-sciences-ai-capabilities-2026
10. https://www.anthropic.com/news/claude-for-life-sciences
11. https://intuitionlabs.ai/
12. https://pmc.ncbi.nlm.nih.gov/articles/PMC9997061/
13. https://github.com/anthropics/uplifting-biomolecular-modeling/blob/main/common/opt_core/README.md
14. https://www.nature.com/articles/s41586-024-07487-w
15. https://github.com/anthropics/uplifting-biomolecular-modeling/blob/main/LICENSE
16. https://github.com/anthropics/uplifting-biomolecular-modeling/blob/main/boltz2/README.md
17. https://github.com/jwohlwend/boltz/blob/main/README.md?plain=1
18. https://github.com/chaidiscovery/chai-lab?_hsenc=p2ANqtz-_EE4qvi563UAs9LbooBDrRT3zSNEr_3LiFAgfGXBjjF4VRVlDNfk_YceKX-EkE7Bw4hPeA
19. https://github.com/anthropics/uplifting-biomolecular-modeling/blob/main/openfold3/README.md
20. https://github.com/anthropics/uplifting-biomolecular-modeling/blob/main/proteinmpnn/README.md
21. https://github.com/anthropics/uplifting-biomolecular-modeling/blob/main/evo2/STOCK.md
22. https://github.com/anthropics/uplifting-biomolecular-modeling/blob/main/chrombpnet/STOCK.md
23. https://docs.docker.com/build/building/best-practices/
24. https://docs.docker.com/build/metadata/attestations/sbom/
25. https://slsa.dev/spec/v1.2/build-provenance
26. https://apptainer.org/docs/user/1.2/signNverify.html
27. https://docs.sigstore.dev/cosign/verifying/verify/
28. https://docs.docker.com/build/metadata/attestations/slsa-provenance/
29. https://intuitionlabs.ai/articles/life-sciences-hpc-specialists
30. https://www.anthropic.com/news/claude-science-ai-workbench
31. https://claude.com/docs/claude-science/how-claude-science-works-with-your-data
32. https://claude.com/docs/claude-science/remote-compute-clusters
33. https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.600-1.pdf
34. https://www.frontiersin.org/journals/bioinformatics/articles/10.3389/fbinf.2026.1925740/full
35. https://www.nature.com/articles/s41467-019-09406-4
36. https://journals.plosone/article?id=10.1371/journal.pone.0161879
37. https://www.fda.gov/media/184830/download?_bhlid=4d54a7e74a14ce20df123608285aebfd01738f67
38. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/considerations-use-artificial-intelligence-support-regulatory-decision-making-drug-and-biological
39. https://www.ema.europa.eu/system/files/documents/scientific-guideline/reflection-paper-use-artificial-intelligence-ai-medicinal-product-lifecycle-en.pdf
40. https://www.nvidia.com/content/dam/en-zz/Solutions/gtcs22/data-center/h100/PB-11133-001_v01.pdf
41. https://claude.com/docs/claude-science/overview
42. https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf

43. https://airc.nist.gov/airmf-resources/airmf/5-sec-core/
44. https://labsyspharm.org/wp-content/uploads/2024/06/Ahdritz_et_al_Nat_Methods_2024.pdf
45. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0161879
46. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/part-11-electronic-records-electronic-signatures-scope-and-application
47. https://intuitionlabs.ai/articles/genai-proof-of-concept-in-pharma

THE PUBLISHER

About IntuitionLabs

IntuitionLabs is an AI consulting, custom software development and data engineering firm serving pharmaceutical, biotechnology, medical-device and other life-science organizations. We work with clinical, regulatory, medical-affairs, commercial, quality and IT teams to connect technology decisions with the work people need to accomplish.

AI consulting and adoption

Our AI enablement services cover readiness assessments, use-case selection, governance and policies, team workshops, adoption measurement and ongoing advisory support. We help organizations structure the information layer behind AI: source material, context, permissions and maintained knowledge that make generated answers useful and reviewable. Private LLM inference and hosted AI options support teams evaluating how to operate AI with appropriate control over their data and infrastructure.

Software, data and life-science workflows

IntuitionLabs develops custom software for pharma and biotech, integrates enterprise systems, and builds data engineering and business intelligence solutions. Areas of focus include AI agents, regulatory research, medical writing, medical affairs, CMC information, competitive intelligence and clinical-document workflows. Our eTMF intelligence work includes cross-system reconciliation and inspection-readiness support.

Enterprise platforms and regulated delivery

We provide Veeva services, application support, managed services, integrations and custom applications, alongside enterprise content work involving platforms such as Egnyte. For regulated workflows, our services include GxP enablement, computer-system validation and software development addressing 21 CFR Part 11 requirements. The applicable controls, validation responsibilities and acceptance criteria are defined for each engagement.

Work with IntuitionLabs

Explore AI enablement, pharma and biotech software development, data engineering and BI, and Veeva services. Contact IntuitionLabs to discuss your workflow, information sources and implementation needs.

IntuitionLabs publishes educational research to help life-science teams make informed technology decisions. Coverage of a product or organization does not imply a client relationship, endorsement or partnership.

Website: https://intuitionlabs.ai