THE ARISING

How AI Helps Scientists Design Proteins

PROTEIN DESIGN



NOVEMBER 2026

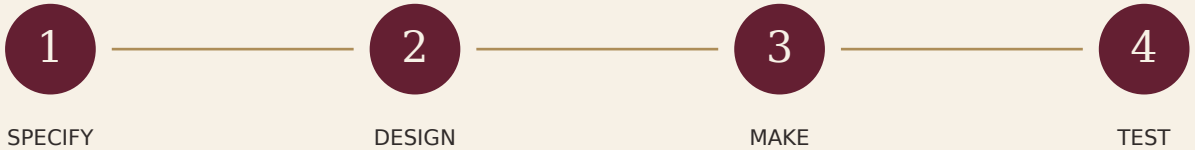
A molecule designed for a job

Proteins are working molecules. Their chains of amino acids fold into structures whose surfaces and movements help determine what they can do. Some catalyze chemical reactions; others carry signals, recognize targets or form larger assemblies. Designing a protein means proposing a sequence and structure that could perform a chosen task, then discovering whether the actual molecule behaves that way.

The specification might be simple to state: bind a particular surface on another protein, hold a useful chemical group in position, or assemble into a symmetrical object. Meeting it is harder. The molecule must make the intended contacts while remaining stable enough to produce and study. A convincing picture addresses only part of that challenge.

Structure prediction and protein design therefore start from different questions. Prediction asks what structure a given sequence is likely to adopt. Design asks which new sequence and structure could satisfy a goal. Prediction tools can help screen designs, but agreement between computer models is still computational evidence. Making the molecule adds a different kind of test. [1]

The opportunity is substantial: researchers can explore molecular arrangements beyond the examples they started with in nature. The useful outcome is a working molecule, not merely a novel shape. That distinction gives this field both its excitement and its discipline.



Design specifies a job. Prediction screens a proposal. Experiments test the physical molecule. Results inform the next design.

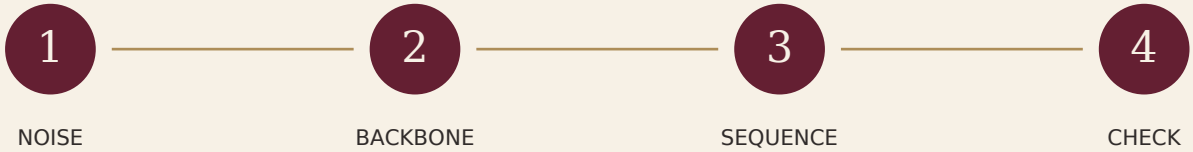
How diffusion becomes a design tool

RFdiffusion adapts a structure-prediction network to a generative task. During training, known protein structures are deliberately disrupted with noise. The model learns to recover plausible structure. When used for design, it starts from noise and iteratively proposes an organized protein backbone, guided where needed by constraints such as a binding target or a functional motif. [1]

A backbone describes the main chain's geometry. A functional motif is a local arrangement that matters for a task, such as a binding site. The surrounding protein must support that arrangement. This allows the design problem to begin with a desired interaction rather than requiring an already complete protein as the starting template.

In the 2023 workflow, a separate method, ProteinMPNN, proposed amino acid sequences for the generated backbones. Structure prediction was then used to assess whether a proposed sequence was likely to recover the intended structure. This sequence of tools matters: the backbone generator alone did not produce all the evidence needed to call a candidate successful. [1]

The denoising steps are mathematical operations in a computer, not a movie of a real protein folding. A successful computational screen means the proposal meets specified model-based criteria. Its experimental prospects still depend on properties those criteria may miss. The attraction of the method is its ability to generate diverse candidates under constraints, giving researchers more possibilities to test.



Conceptual workflow: disordered coordinates become a candidate backbone; a sequence is proposed and checked. This is not a physical folding trajectory.

What the laboratory established

Watson and colleagues' 2023 Nature study moved beyond displaying generated structures. The researchers experimentally characterized structures and functions of hundreds of designed symmetric assemblies, metal-binding proteins and protein binders. These are distinct design challenges: forming an organized assembly is different from recognizing a target surface. [1]

One particularly informative result was a cryo-electron microscopy structure of a designed binder in complex with influenza haemagglutinin. The observed structure was nearly identical to the design model, according to the paper. That comparison supports the accuracy of this molecular design in the tested complex. It does not show that the binder prevents influenza in people. [1]

Different experiments answer different questions. Producing a soluble protein shows that material can be obtained under the tested conditions. A binding assay tests interaction with a target. Structural measurement asks whether the molecule and its partner adopt the intended arrangement. Each is valuable; none can stand in for all the others.

The broad result is a more versatile approach to protein engineering. The paper's scope does not justify assuming that every generated candidate will work, that all targets are equally tractable, or that computational confidence measures clinical value. When reading a success rate, the denominator and the selection process are part of the result: candidates generated, candidates filtered, and candidates actually tested can be very different groups.



Three complementary laboratory questions. A positive answer in one column does not automatically answer the other two.

Designing around a moving target

Short peptides can be difficult binding targets because their shapes may change. Some peptide hormones form a helix when they bind their receptors, yet are less structured on their own. A designed partner must accommodate the target's shape and the energetic cost of organizing it. Matching a single static picture can miss that problem. [2]

Vázquez Torres and colleagues addressed helical peptide binding through several approaches, including an extension of RFdiffusion for flexible targets. They also used partial diffusion: perturbing and rebuilding an existing design rather than beginning every search from scratch. The study reported picomolar-affinity binders to helical peptide targets, including designs made without subsequent experimental optimization. [2]

Affinity describes how tightly molecules bind under the measurement conditions. Picomolar is a concentration scale, not a measure of a patient's health. Tight binding can be useful, but specificity matters too: a reagent that captures similar unwanted molecules may give an ambiguous signal. The sample, assay and intended use determine which properties matter most.

The researchers demonstrated enrichment and subsequent mass-spectrometry detection of parathyroid hormone and glucagon, and constructed bioluminescence-based protein biosensors. PTH capture used supplemented human plasma; glucagon capture used spiked buffer because recovery in extract was low. The authors noted that glucagon-binder specificity would probably need improvement for practical applications. The work supports a route toward useful measurement tools, while clinical diagnostic accuracy, reliability across patient samples and improved treatment decisions remain separate questions. [2]



A binder can capture a target for measurement. Capture, detection and clinical interpretation are separate steps, each needing validation.

From a useful reagent to better care

A protein can become valuable without becoming a medicine. A selective laboratory reagent might help measure a hormone or probe a biological mechanism. A therapeutic candidate faces additional demands because it must reach the right place, act for an appropriate duration and avoid unacceptable harm. Evidence requirements should follow the proposed use.

For a diagnostic application, the question expands from whether a target can be detected to whether it can be measured reliably in the samples and settings where decisions will be made. Interference from other molecules, variation between samples and reproducibility across runs can change interpretation. A better analytical signal is promising; better clinical decisions require their own evidence.

For treatment, binding is only one property. Stability, manufacture, distribution, clearance, off-target interactions and unwanted immune responses can all matter. These are questions to investigate, not a prediction that a particular protein will fail. Preclinical activity can justify further research while leaving human safety and benefit unsettled.

The most credible optimism is about a better experimental cycle. More informative candidates could make each round of laboratory work more productive. Whether a particular workflow saves time or resources should be measured against an appropriate comparison. The two studies here establish engineering advances, not a general promise of faster cures. [1] [2]



Requirements depend on use. A laboratory reagent, a clinical test and a treatment need different evidence; none is an automatic next step.

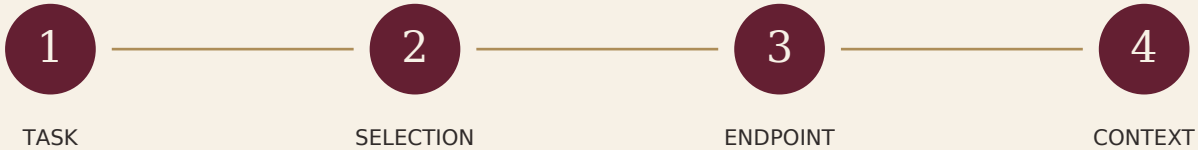
How to read the next protein headline

Start with the task. Was the goal to create a stable shape, bind a target, detect a molecule or change disease? Then identify what was actually measured. A structure-prediction score, a laboratory binding result and a patient outcome belong to different categories of evidence.

Ask how candidates were chosen. How many were generated, filtered and tested? What counted as success, and were failed designs reported? A spectacular example can establish possibility without establishing the reliability of the entire process. A fair comparison should use the same task, endpoint and experimental budget.

Look for specificity and context. Which competing targets were tested? Did the result hold in the relevant sample or biological setting? For a clinical claim, ask about people, comparators, harms and outcomes that matter to patients. A plausible mechanism should guide the next experiment rather than replace it.

Protein design is opening a wider space of molecular possibilities. Its durable achievement will be the ability to turn clear specifications into reproducible functions, and to learn from what fails. Computation helps propose the next molecule. Experiments determine what that molecule can actually do.



Follow the claim from the intended task to the measured endpoint. Keep uncertainty attached to the step where it remains.

Primary sources

[1] Watson JL, Juergens D, Bennett NR, et al. De novo design of protein structure and function with RFdiffusion. Nature. 2023;620:1089-1100. DOI: 10.1038/s41586-023-06415-8.
https://pubmed.ncbi.nlm.nih.gov/37433327/

[2] Vázquez Torres S, Leung PJY, Venkatesh P, et al. De novo design of high-affinity binders of bioactive helical peptides. Nature. 2024;626:435-442. DOI: 10.1038/s41586-023-06953-1.
https://pubmed.ncbi.nlm.nih.gov/38109936/

What these studies support

The first demonstrates a versatile protein-design method with experimental structural and functional validation. The second extends design to flexible helical peptide targets and demonstrates laboratory detection and biosensing applications. Neither study establishes patient benefit for the designed proteins discussed here.

This educational feature explains two foundational studies. It is not a treatment recommendation or a comprehensive survey of every protein-design platform.