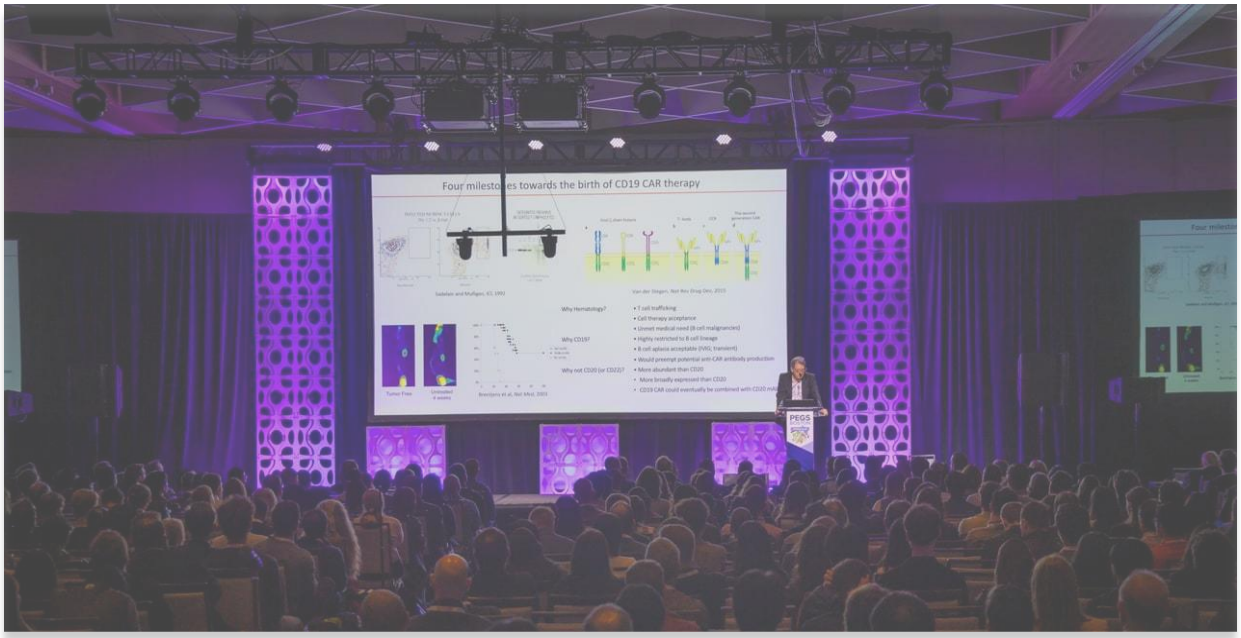

Conference Track Summary –

Machine Learning for Protein Engineering

Brought to you by Cambridge Healthtech Institute, a division of Cambridge Innovation Institute

250 1st Ave, Suite 300, Needham, MA 02494
T: 781.247.6262

Executive Overview

The **PEGS Boston 2026** Machine Learning for Protein Engineering track highlighted the rapid evolution of AI from a protein prediction tool into a practical platform for designing functional biologics. Across presentations, speakers emphasized that advances in generative models, protein language models, and active learning are accelerating binder discovery, antibody optimization, and protein engineering. Discussions focused on integrating computational design with experimental validation, multi-parameter optimization, and rigorous benchmarking to produce molecules that are not only functional, but also developable and therapeutically relevant.

Most Frequently Covered Issues

1. Generative protein design
AI is enabling de novo design of proteins with increasingly sophisticated structures and biochemical functions.

2. Multi-property optimization
Developability, stability, specificity, and manufacturability are becoming design objectives alongside affinity.

3. AI-guided discovery workflows
Integrated computational platforms are helping prioritize candidates earlier and reduce experimental cycles.

4. Active learning and experimental feedback
Iterative laboratory validation continues to improve model performance and guide successive rounds of design.

5. Benchmarking AI performance
Meaningful evaluation frameworks are essential for identifying where AI delivers reliable scientific value.

Recurring Takeaways

- AI is accelerating protein engineering by integrating design, prediction, and optimization.
- Early computational triage focuses laboratory resources on stronger candidates.
- Successful protein design requires balancing multiple therapeutic properties simultaneously.
- Experimental validation remains essential for confirming AI-generated predictions.
- Robust benchmarks and high-quality data will determine the next generation of protein design tools.

Executive Overview 2
Accurate Protein-Binder Design Using BindCraft 4
Designing Biochemical Function with Generative AI 5
De Novo Miniprotein Design: Developing Drug-Like Molecules from Scratch 6
LensAI: A Multi-Parametric AI Platform for Integrated Antibody Discovery, De-Risking, and Engineering 7
Application of AI to Developability Screening, a Skeptic's View 8
TherAbDesign: Bridging AI and Biophysics for Antibody Developability Optimization 9
Scalable, Robust and Easy to Use Biologics Discovery & Development with Generative AI 10
Evaluation of Digital Protein Design Tools in an Industry Setting 11

Accurate Protein-Binder Design Using BindCraft

Full Video Here

Lennart Nickel, Graduate Student, Biotechnology & Bioengineering, École Polytechnique Fédérale de Lausanne

De novo binder design can shorten the path from a biological question to a useful experimental molecule. BindCraft uses AlphaFold-based hallucination, sequence optimization, and final filtering to generate compact protein binders, allowing top candidates to move into wet-lab testing without the large library screens traditionally required for binder discovery. (00:03:46–00:07:05)

Across diverse protein families, testing roughly 10 to 20 designs often produced binders with nanomolar affinity, although some targets remained challenging. Co-folding lets both binder and target adjust during design, creating complementary interfaces, while sequence optimization produces molecules with strong secondary structure and high thermal and chemical stability. (00:07:05–00:10:46)

The practical value extends beyond binding measurements. Designed proteins blocked a bacterial toxin from engaging claudin and rescued cells in a concentration-dependent manner, turning a process that can take months into a rapid proof-of-concept workflow that can test biological hypotheses within weeks. (00:12:06–00:14:16)

High hit rates also make direct functional screening increasingly realistic. Binders incorporated into engineered viral capsids produced target-specific cell transduction without prior binding prescreens, illustrating how designed proteins can function as modular components within larger systems. Remaining challenges include conditional behavior, multi-target logic, post-translational modifications, and standardized benchmarking across design methods. (00:15:26–00:21:20)

Key Takeaways

- De novo design can generate useful binders from small computational candidate sets.
- Co-folding supports complementary interfaces and highly stable designed proteins.
- Rapid binder generation enables faster biological hypothesis testing.
- Standardized benchmarks are needed to define where computational binder design succeeds and fails.

Designing Biochemical Function with Generative AI

Full Video Here

Rohith Krishna, PhD, Postdoctoral Fellow, Computational Biology & Machine Learning, University of Washington

De novo protein design is moving beyond generating structures toward specifying precise biochemical functions. Diffusion models can start from noise and build proteins around defined requirements, while a combination of soft and hard constraints lets researchers control features such as hydrogen bonding, binding geometry, symmetry, active sites, nucleic acids, and small molecules. (00:02:51–00:07:54)

Combining constraints opens functions that are difficult to reach by modifying natural proteins. One example produced an enzyme around a complex active site containing metals, six functional residues, and a post-translational modification, while related approaches can pair sequence-specific binding with protease activity or impose symmetry on functional enzymes for multivalent and clustered architectures. (00:10:08–00:14:31)

Compact, sequence-specific DNA binders show another direction for generative design. Proteins of roughly 100 to 150 amino acids were designed against specific DNA sequences, achieving single-digit nanomolar affinities and sequence selectivity, suggesting a route toward smaller programmable DNA-binding systems that could eventually be coupled with other functional or conditional modules. (00:14:56–00:17:43)

Structural constraints can also shape post-translational outcomes rather than binding alone. By controlling how deeply glycan atoms were buried within designed proteins, researchers protected selected sugars from enzymatic processing and generated proteins carrying high-mannose glycans with controllable branching patterns. Such capabilities expand the biochemical behaviors accessible to design, although safety, immunogenicity, and pharmacology remain important translational challenges. (00:18:06–00:25:36)

Key Takeaways

- Generative models can design toward biochemical function, not simply protein structure or binding.
- Composable constraints enable control over active sites, symmetry, molecular interactions, and geometry.
- Compact designed proteins can achieve strong, sequence-specific DNA recognition.
- Structural design can control glycosylation patterns and other complex biochemical properties.

De Novo Miniprotein Design: Developing Drug-Like Molecules from Scratch

Full Video Here

James Bowman, PhD, CTO, AI Proteins

De novo miniproteins offer unusually direct control over properties that matter for therapeutic development. Compact 40 to 60 amino acid designs can combine stable tertiary structure with favorable tissue penetration, while deliberate placement of hydrophobic and soluble surfaces provides an opportunity to engineer drug-like behavior from the beginning rather than inherit liabilities from natural scaffolds. (00:01:37–00:04:14)

Binding is only an early step toward a viable therapeutic. High-throughput purification and characterization can generate the empirical data needed to predict expression, stability, specificity, and other developability properties from sequence, allowing poor candidates to be removed before costly experiments and reducing repeated design-make-test cycles. (00:04:14–00:10:22)

Specificity requires testing beyond clean biophysical systems because nonspecific interactions may emerge only in complex biological environments. Cell-based measurements can complement SPR, while sequence-space sampling helps preserve molecular diversity and identify candidates with different surface chemistries that may influence stability, tissue uptake, and other properties that cannot yet be reliably predicted. (00:10:22–00:16:48)

A selective TNFR1 antagonist illustrates how these principles can converge in a functional molecule. The optimized miniprotein combined low-picomolar TNFR1 affinity with undetectable TNFR2 binding, retained activity in primary-cell assays, and controlled TNF-dependent inflammation in vivo after cross-species reactivity was deliberately engineered into the design. (00:17:15–00:24:32)

Key Takeaways

- Small de novo proteins can be engineered for therapeutic properties from the outset.
- Property prediction can remove weak candidates before expensive experimental cycles.
- Biological specificity requires empirical testing in increasingly complex systems.
- Multi-property optimization can advance designed binders toward functional therapeutic candidates.

LensAI: A Multi-Parametric AI Platform for Integrated Antibody Discovery, De-Risking, and Engineering

Full Video Here

Shuji Sato, VP of Innovative Solutions, Business Development, MindWalk

Antibody discovery generates abundant sequence and characterization data, but functional testing remains a major throughput constraint. Bringing computational analysis closer to the beginning of discovery can help teams prioritize candidates earlier, reducing iterative laboratory cycles before the most resource-intensive functional characterization begins. (00:00:00–00:03:03)

A unified data layer can connect sequence, structure, function, literature, and program-specific assay data throughout discovery. Pattern-based analysis can then identify relationships across these sources, supporting early decisions around target profiling, antigen design, epitopes, cross-reactivity, and screening tools while preserving traceability back to the underlying evidence. (00:03:03–00:09:09)

Once antibody sequences emerge, in silico analysis can rapidly organize and prioritize candidates before additional experiments. Paratyping groups antibodies by predicted binding behavior, while developability assessment, immunogenicity screening, and epitope mapping help distinguish candidates based on liabilities and likely function rather than relying solely on sequence similarity or affinity. (00:10:40–00:17:38)

Integrating these analyses with in vivo discovery creates a function-first path from immunization to optimized leads. High-throughput B-cell screening can recover diverse and rare antibodies, after which sequence-based humanization and multi-parameter optimization can be completed in days before producing selected clones for final functional testing, concentrating laboratory resources on stronger candidates. (00:12:21–00:18:36)

Key Takeaways

- Early computational analysis can reduce costly iterative discovery cycles.
- Connecting sequence, structure, function, and literature strengthens candidate prioritization.
- Developability and functional properties should inform selection alongside affinity.
- In silico optimization can focus laboratory testing on higher-quality antibody leads.

Application of AI to Developability Screening, a Skeptic's View

Full Video Here

Andrew C.R. Martin, DPhil, Emeritus Professor of Bioinformatics and Computational Biology, University College London

Generative antibody design warrants careful scrutiny because experimental validation remains unavoidable and established discovery methods already perform well. Some published approaches described as ab initio design still depend heavily on display libraries or affinity maturation, while stronger recent results leave open questions about unsuccessful targets, specificity, immunogenicity, and broader reproducibility. (00:00:00–00:04:57)

AI may offer more immediate value in antibody triage, including epitope prediction, affinity ranking, and developability assessment. Combining protein language models with graph neural networks improved epitope prediction over existing approaches, although performance remained imperfect, reinforcing the need to judge AI methods against practical baselines rather than assuming computational sophistication guarantees useful predictions. (00:04:57–00:08:14)

Developability may benefit from evaluating antibody sequences globally rather than scoring individual liabilities independently. Protein language model embeddings paired with nonlinear dimensionality reduction placed clinical-stage antibodies within a concentrated region of sequence space, and a separate held-back set fell within the same region, suggesting a practical way to triage large repertoires before deeper characterization. (00:08:14–00:17:26)

Predicting clinical success is harder because discontinuation can reflect efficacy, safety, target biology, or commercial decisions beyond intrinsic molecular properties. Even so, language-model features combined with a support vector machine separated approved and discontinued antibodies with encouraging performance, supporting layered AI-based triage while underscoring the limitations created by small training and independent test datasets. (00:17:26–00:25:57)

Key Takeaways

- Generative antibody design claims require rigorous experimental and cross-target validation.
- AI currently offers practical value in antibody triage and developability assessment.
- Global sequence representations may reveal developability signals missed by individual metrics.
- Limited datasets remain a major constraint on predicting clinical success.

TherAbDesign: Bridging AI and Biophysics for Antibody Developability Optimization

Full Video Here

Amy Wang, PhD, Senior ML Scientist, Prescient Design, Genentech

Antibody optimization must balance affinity with properties such as stability, aggregation, and viscosity that determine whether a molecule can become a viable drug. Because gold-standard developability assays are expensive and low throughput, liabilities may emerge only after substantial engineering, creating lengthy cycles in which solving one problem can introduce another. (00:00:00–00:03:11)

Scalable biophysical proxies can provide earlier signals, but their usefulness depends on how those properties are defined and benchmarked. Comparisons of surface-based descriptors found little additional information from costly molecular dynamics simulations, while finer-grained hydrophobicity and electrostatic descriptors performed better than coarser alternatives at identifying experimental viscosity liabilities. (00:03:11–00:13:35)

Structure-based calculations can still become a bottleneck when screening large libraries. A sequence-based model trained to reproduce useful biophysical descriptors achieved strong correlation with structure-derived values while reducing scoring time dramatically, making it practical to evaluate large candidate sets earlier and preserve enough signal to identify molecules carrying elevated viscosity risk. (00:13:35–00:20:52)

Once developability properties can be predicted directly from sequence, the same framework can guide molecular design rather than simply filter candidates. In a test case with a known viscosity liability, proposed sequence changes largely recovered experimentally validated improvements, illustrating how scalable models can steer molecules toward more favorable biophysical space without being explicitly trained on viscosity itself. (00:20:52–00:26:34)

Key Takeaways

- Developability signals need to enter antibody optimization earlier.
- Useful proxies should be benchmarked before becoming design objectives.
- Sequence-based models can preserve valuable biophysical signals at much greater scale.
- Scalable developability prediction can support both candidate triage and molecular redesign.

Scalable, Robust and Easy to Use Biologics Discovery & Development with Generative AI

Full Video Here

Eli Bixby, Co Founder & Head of ML, ML Research & Engineering, Cradle

Scaling AI-enabled protein engineering requires more than a collection of models. An integrated platform can combine institutional, project, and public data with scientific constraints and experimental feedback, building fit-for-purpose predictive and generative models each round while preserving scientists' ability to define objectives, incorporate domain knowledge, and interpret expected outcomes. (00:01:03–00:07:38)

De novo binder generation shows why execution matters alongside model architecture. Across tested approaches, improvements in structural ranking had greater impact than switching generators, while performance varied substantially between users. Target difficulty also correlated with the quantity and quality of existing structural data, highlighting unresolved questions around generalization, confidence calibration, and robust library design. (00:07:38–00:11:17)

Zero-shot models can support early candidate selection, but their usefulness declines as optimization becomes more fine-grained. Developability and binding predictors showed uneven generalization across properties and formats, with performance deteriorating when asked to distinguish closely related variants. Active learning addresses this limitation by incorporating new experimental data across successive design rounds. (00:11:17–00:15:44)

Iterative learning becomes especially valuable for complex multi-property molecules. In bispecific optimization, prior monovalent data improved first-round affinity, while subsequent experimental results allowed binding between two arms to be balanced deliberately. Explicit constraints also preserved binding while shifting optimization toward thermostability, producing a substantial increase in melting temperature without sacrificing the targeted affinity window. (00:15:44–00:20:52)

Key Takeaways

- Scaling generative AI requires integrated workflows, data, and scientific expertise.
- Ranking strategy and execution can matter as much as the underlying generative model.
- Zero-shot predictions lose power during fine-grained lead optimization.
- Active learning enables deliberate multi-property optimization across experimental rounds.

Evaluation of Digital Protein Design Tools in an Industry Setting

Full Video Here

Koji Tsuda, PhD, Professor, Computational Biology & Medical Sciences, University of Tokyo

Language models offer a way to reuse learned protein representations across many antibody and nanobody tasks rather than building a separate model for each problem. Because performance can vary widely by application, meaningful comparison requires benchmarks spanning structural understanding, binding, and biophysical properties rather than relying on a single headline metric. (00:00:50–00:07:58)

Across eight encoder-model tasks, specialization helped selectively rather than universally. Antibody-focused language models performed better on antigen-binding and paratope prediction, while general protein models remained competitive elsewhere. Thermostability and affinity prediction were challenging across models, reinforcing that no single architecture consistently performs best across the full range of antibody-related problems. (00:08:02–00:19:17)

Decoder-based large language models showed similarly mixed results when antibody problems were framed as natural-language tasks. Fine-tuning improved some binding-scoring performance, while multimodal approaches combining protein representations with language models performed better on ranking. Directly asking LLMs to identify binders from large candidate pools, however, produced essentially random results. (00:19:17–00:25:38)

Benchmark design therefore shapes not only model evaluation but also where researchers invest effort. Poorly chosen tasks can reward progress that has little scientific relevance, while standardized datasets and evaluation protocols make limitations easier to identify. Current results also suggest smaller, specialized models may sometimes provide comparable performance more efficiently than very large general-purpose models. (00:25:38–00:29:44)

Key Takeaways

- No single language model performs best across all antibody and nanobody tasks.
- Antibody-specific models show advantages for several antigen-related problems.
- Current LLMs struggle with direct antibody library screening.
- Scientifically meaningful benchmarks are essential for directing useful model development.