2024/09/19

Generative AI is starting to disrupt drug development

Stef van Grieken (stef@cradle.bio)

Stef van Grieken

Now and before

Cradle CEO and Co-Founder	All the time
Operator Exchange Angel investing together with cool people	Sometimes
Google (AI, X, Flights, Android Auto) Group product manager	2014 - 2021
OpenState Foundation Founder and Chairman	Before that
Google Software Engineering Intern	Long time ago
European Parliament Policy Advisor	Very long time ago

Machine learning for drug development

Engineering proteins with Cradle

Examples from antibodies & vaccines

Teaching computers about proteins

Gen AI can be applied across the drug development process

causaly

"Summarize academic literature on disease"

SPRING

"Figure out what cells died on your plates"

?

"Select the right target groups for my trial"

?

"Write the regulatory report for launch"

TARGET IDENTIFICATION

TARGET TO HIT

LEAD OPTIMIZATION

PRE-CLINICAL

PHASE I

PHASE II

PHASE III

BenchSci

"Select right re-agents for experimental setup"

Cradle

"Select and optimize my hits for efficacy, safety and manufacturability"

?

"Make sure I recruit the right patients into trial"

Small molecules.
Big success rates in phase 1.

90%

90% of small molecules designed with Gen AI succeed in phase I clinical trials. That’s way above industry standard. We don't have enough data for larger molecules.

Outline

Machine learning for drug development

Engineering proteins with Cradle

Examples from antibodies & vaccines

Teaching computers about proteins

Cradle

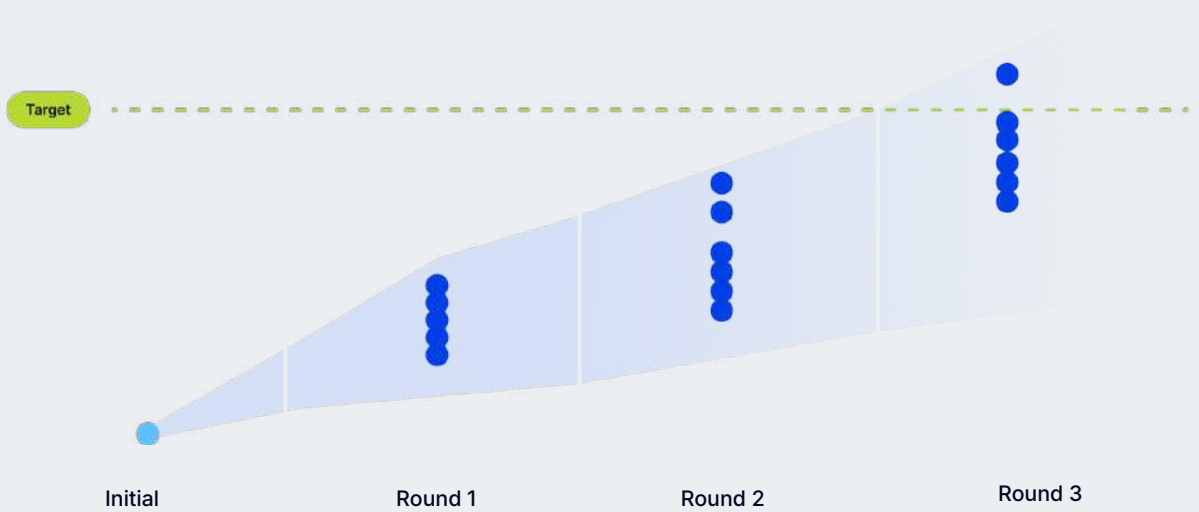
40+ people across Biology, Design, Software and Machine Learning.

Offices (and lab) in Amsterdam and Zürich.

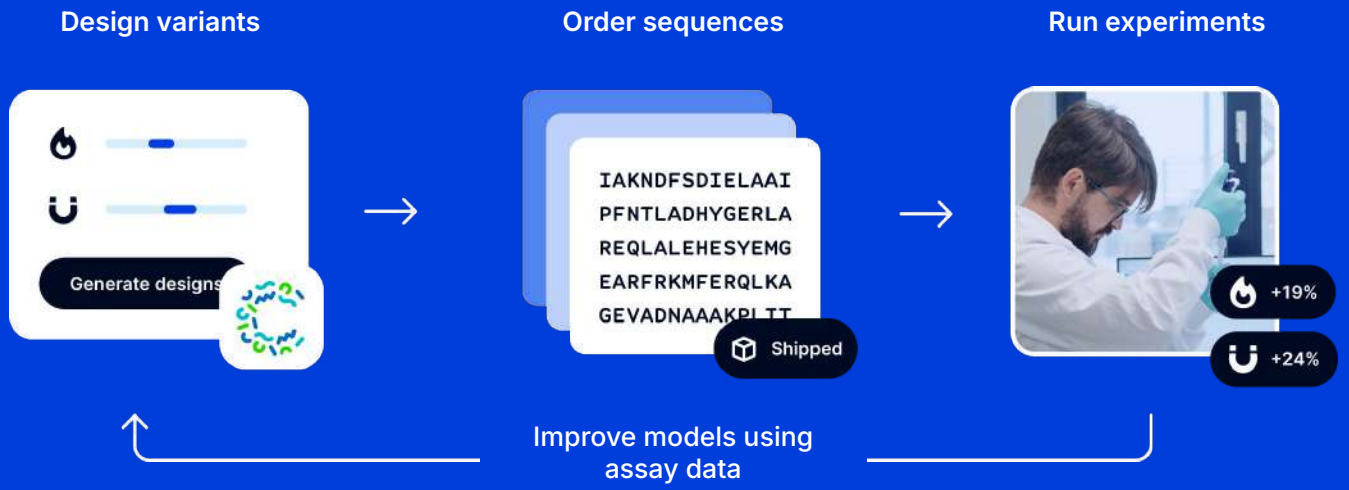
Founded in 2021.

We help make every experimental round better than the last

Protein quality over rounds

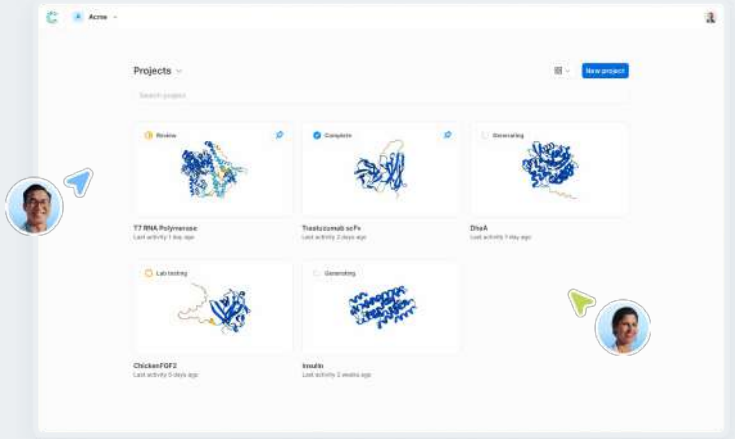


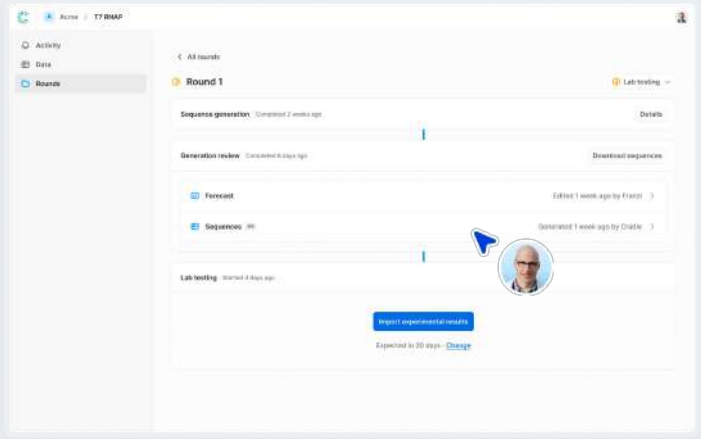
Cradle fits into and enhances the existing experimental workflows



Software

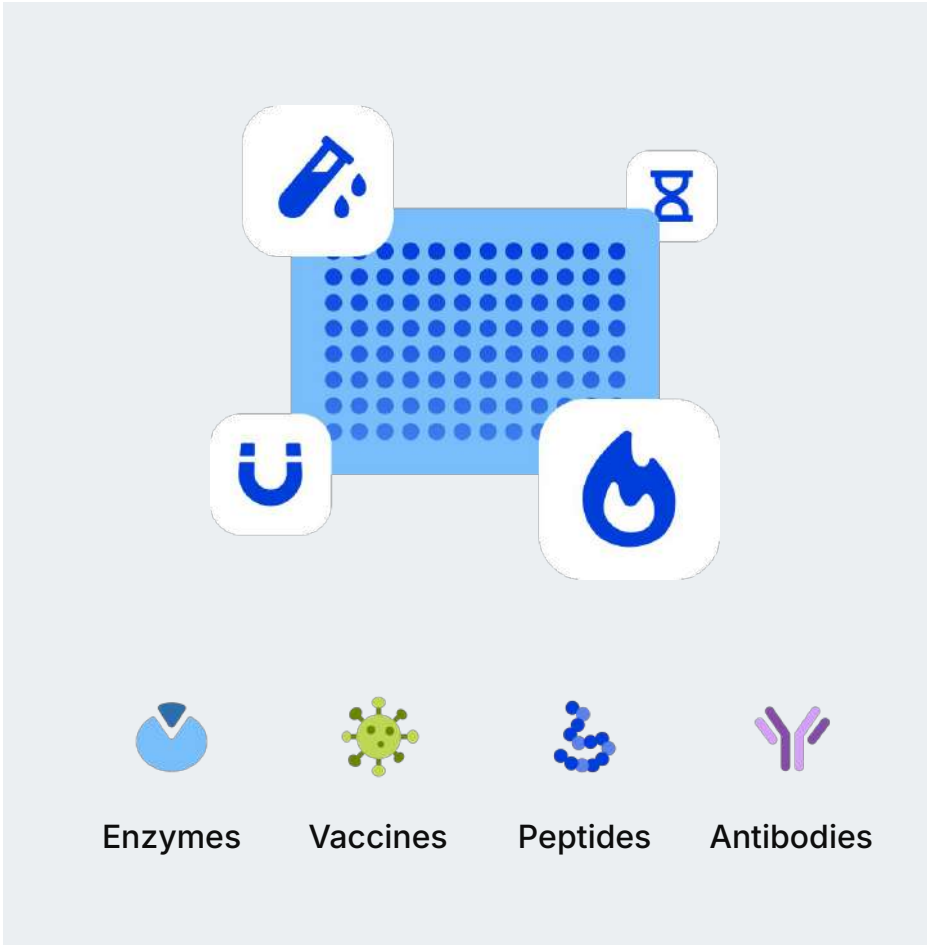
Cradle is easy to use by anyone in the lab – no ML expertise needed.





We support multi-property microtiter workflows for many modalities and +40 assays

A 96-well plate of clean data is often enough to fine-tune Cradle's core machine models to your specific assays and start seeing great results.



Our Lab

We improve our models using data from our Amsterdam lab.

Developing machine learning models & tools is a continuous process of improvement. We A/B test every new version of our models on a wide range of proteins.

We also develop ‘Foundational Datasets’ aimed specifically at understanding protein function. For example, we build a 10^7 library of antibody expression in CHO to help models understand what a well expressing antibody looks like.

19	Customers
30	Programs
6	Modalities
6+	Properties

Satisfaction

Our customers report high level of satisfaction

“Cradle's AI-based approaches were invaluable in enhancing our enzyme's activity. We were particularly impressed by its ability to effectively leverage our historical data, as well as the collaborative spirit the Cradle team exhibited throughout the project.”

Staff Scientist

“Cradle is great in terms of delivering their design solutions to customers within the set time frame.”

Director Computational Design

Customer satisfaction:

Outline

Machine learning for drug development
Engineering proteins with Cradle
Examples from antibodies & vaccines
Teaching computers about proteins

Antibodies

120+

120+ antibody-based therapeutics are now approved for treating serious diseases.

Humira and Herceptin are notable medicines for rheumatoid arthritis and breast cancer.

300B

Market value expected to reach 300B by 2025.

1: Lyu X, Zhao Q, Hui J, Wang T, Lin M, Wang K, Zhang J, Shentu J, Dalby PA, Zhang H, Liu B. The global landscape of approved antibody therapies. Antib Ther. 2022 Sep 6;5(4):233-257. doi: 10.1093/abt/tbac021. PMID: 36213257; PMCID: PMC9535261.
2: Lu, RM., Hwang, YC., Liu, IJ. et al. Development of therapeutic antibodies for the treatment of diseases. J Biomed Sci 27, 1 (2020). https://doi.org/10.1186/s12929-019-0592-z
image: Bo Wang, Sachith Gallolu Kankanamalage, Jianbo Dong, Yue Liu, Optimization of therapeutic antibodies, Antibody Therapeutics, Volume 4, Issue 1, January 2021, Pages 45–54, https://doi.org/10.1093/abt/tbab003

The process for developing a new antibody drug

Target identification and validation		Antibody generation and screening		Hit-to-lead		Lead optimization	
		Timeline	3-6 mo	Timeline	6-12 mo	Timeline	6-12 mo
Timeline	3-6 mo	Screening	10^6 - 10^{10} variants	Rounds	2-10's	Leads	2-5
Target success rate	~50%	Success rate	0.001-0.1%	Lead success rate	10-20%	Success rate	1-2

Cradle

Common engineering strategies

Hit Identification

- Natural diversity screens
- Site saturation libraries
- Immunization campaigns
- **De novo machine learning methods**

Lead optimization

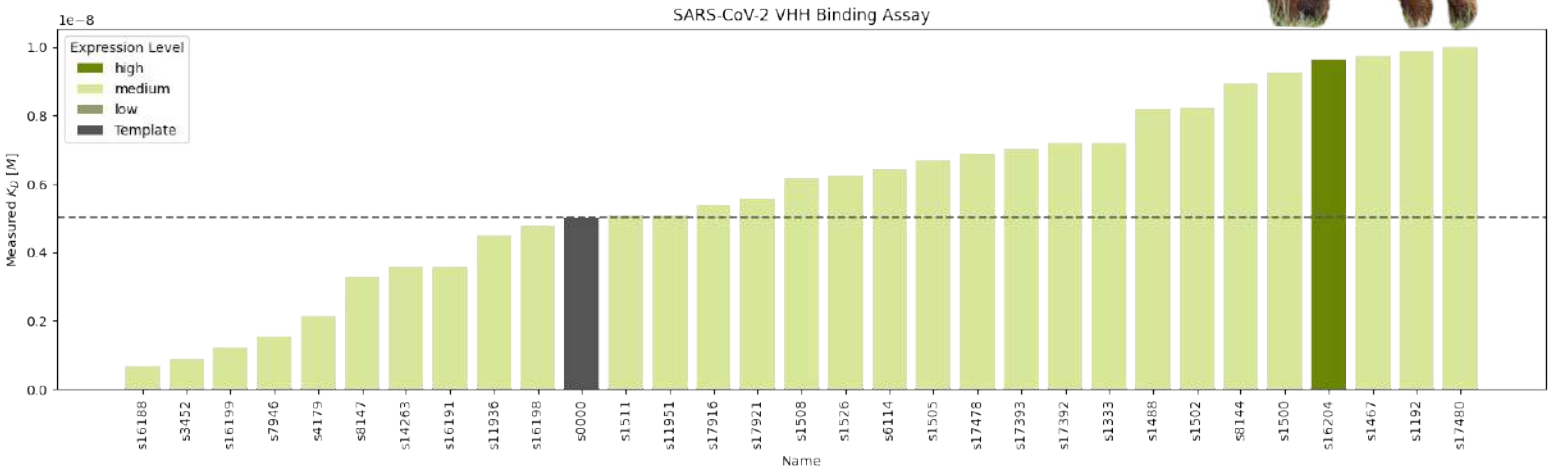
- Iterative mutation stacking / combinatorial libraries
- Directed evolution
- Rational design
- **AI-guided design ‘lab - in - the - loop’**

Hit Identification

Experiment 1: finding novel binders from an immunization campaign

Binding Affinity on Sars-CoV-2 with modified CDRs

- Conditioned generative models on phage display data from immunization campaign
- Generated and selected 96 sequences with mutations in framework and CDR1 & CDR2
- Binding affinity increased, creating 2 sub-nM binders
- High hit rate of increased binding

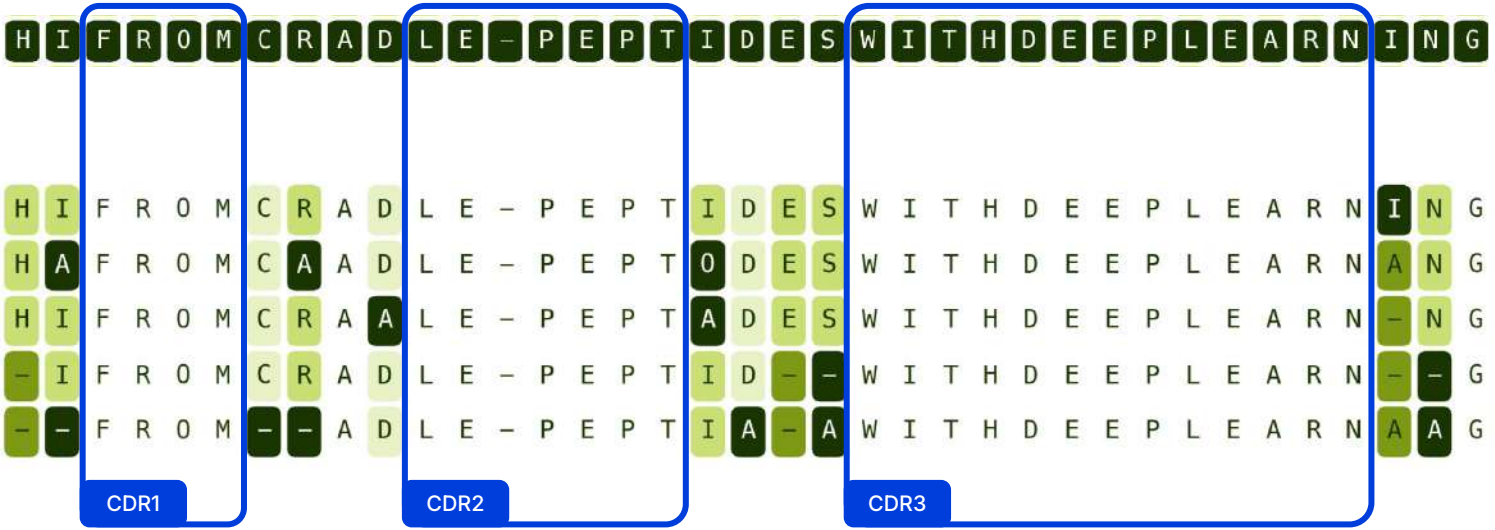


Tyson the alpaca takes heavyweight role in search for coronavirus vaccine | Reuters

Binding Maturation

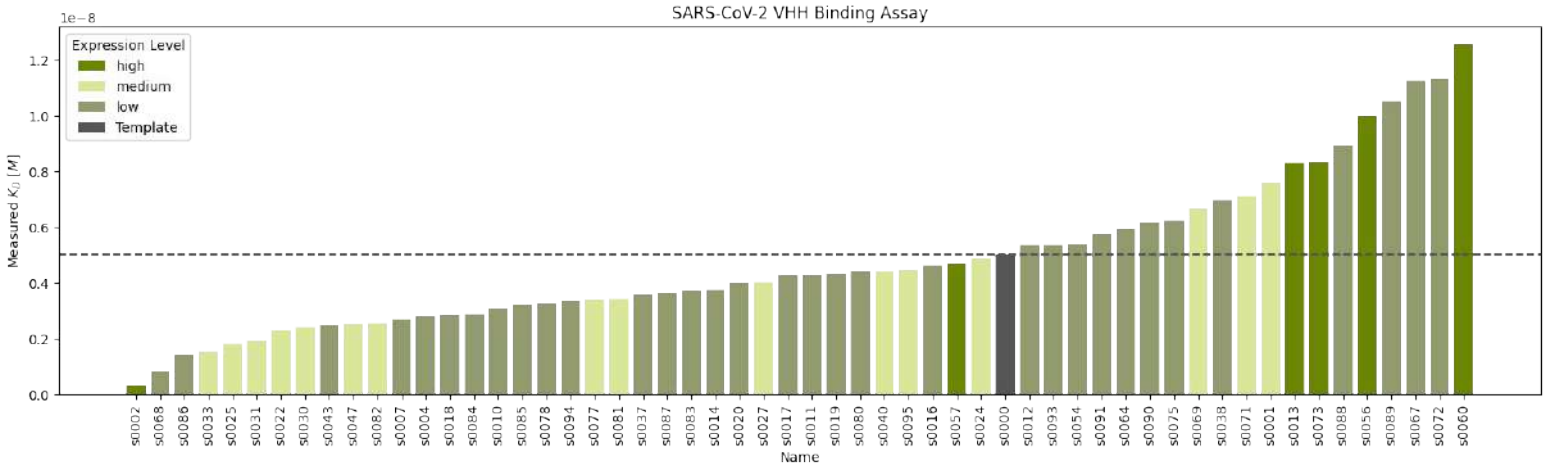
Experiment 2: Improving affinity against target

Only making framework mutations

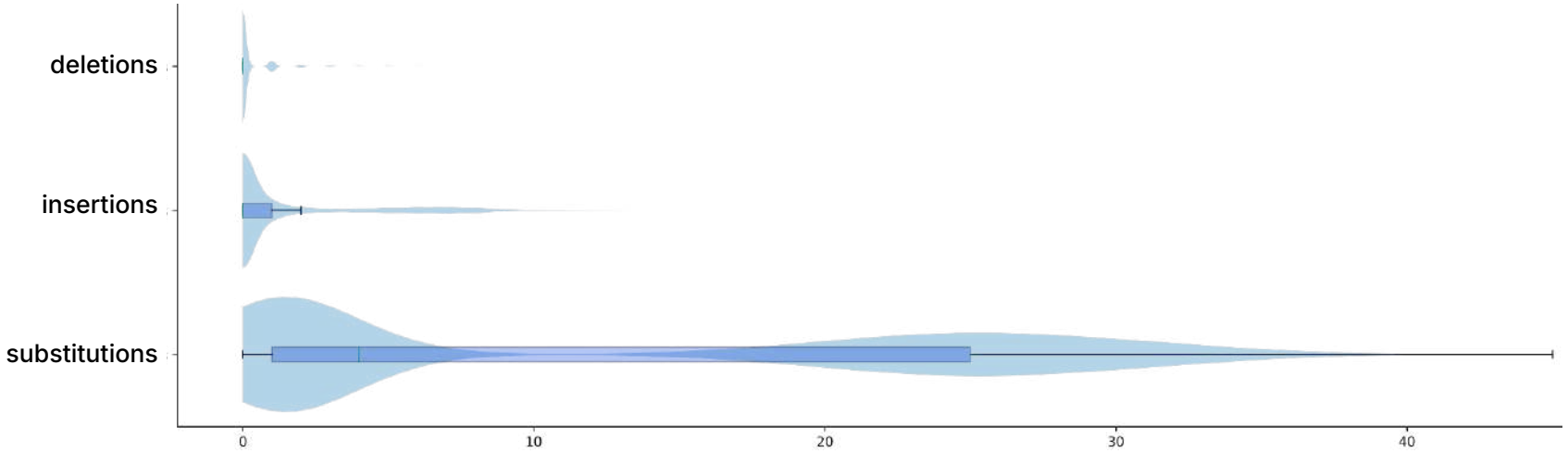


Improving affinity by 9x with framework mutations

- Conditioned generative models on alignment of similar sequences
- Generated and selected one round of 96 sequences with conserved CDRs, substitutions only
- Binding affinity increased by **9x**, achieving sub-nanomolar binding for 2 variants
- High hit rate of retained binding

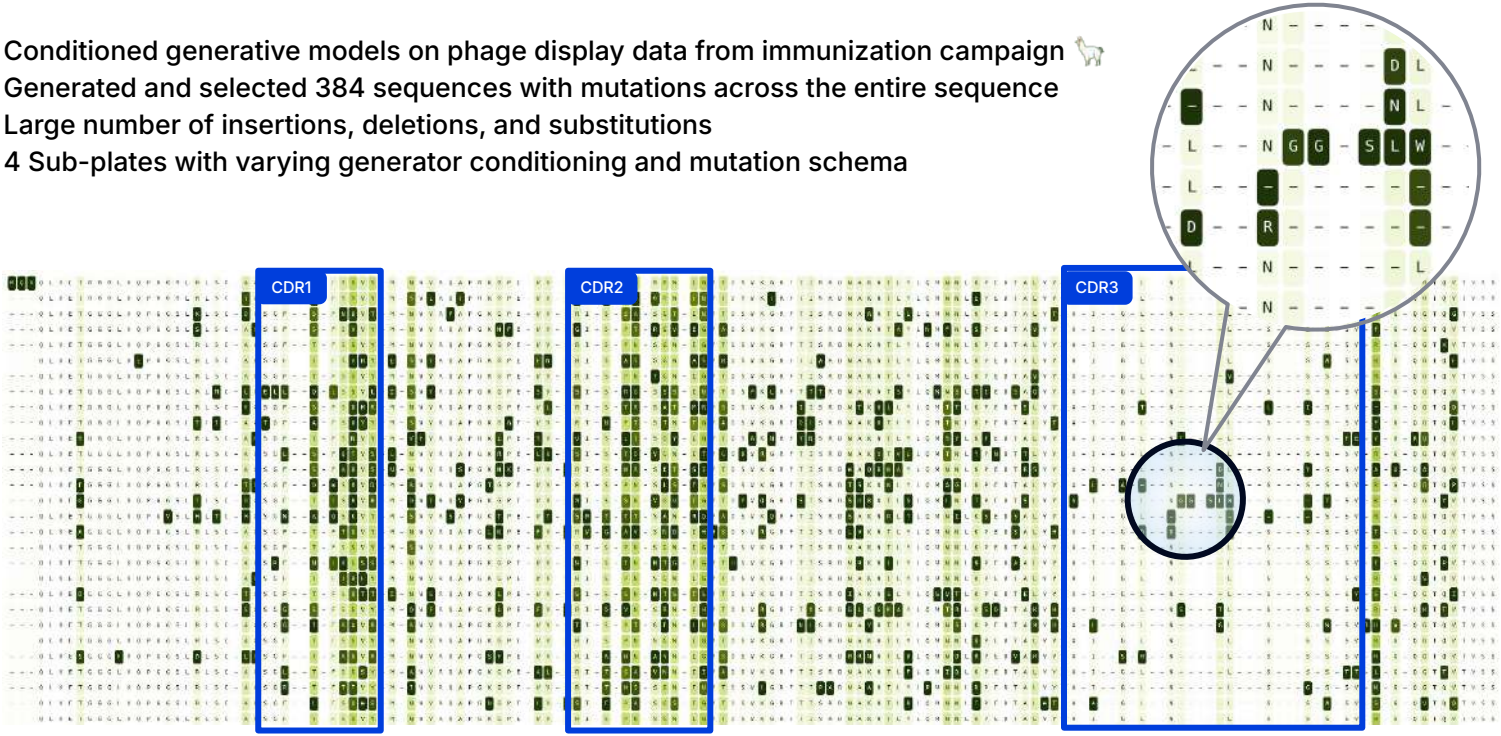


What if we introduced significant modification across the entire sequence?

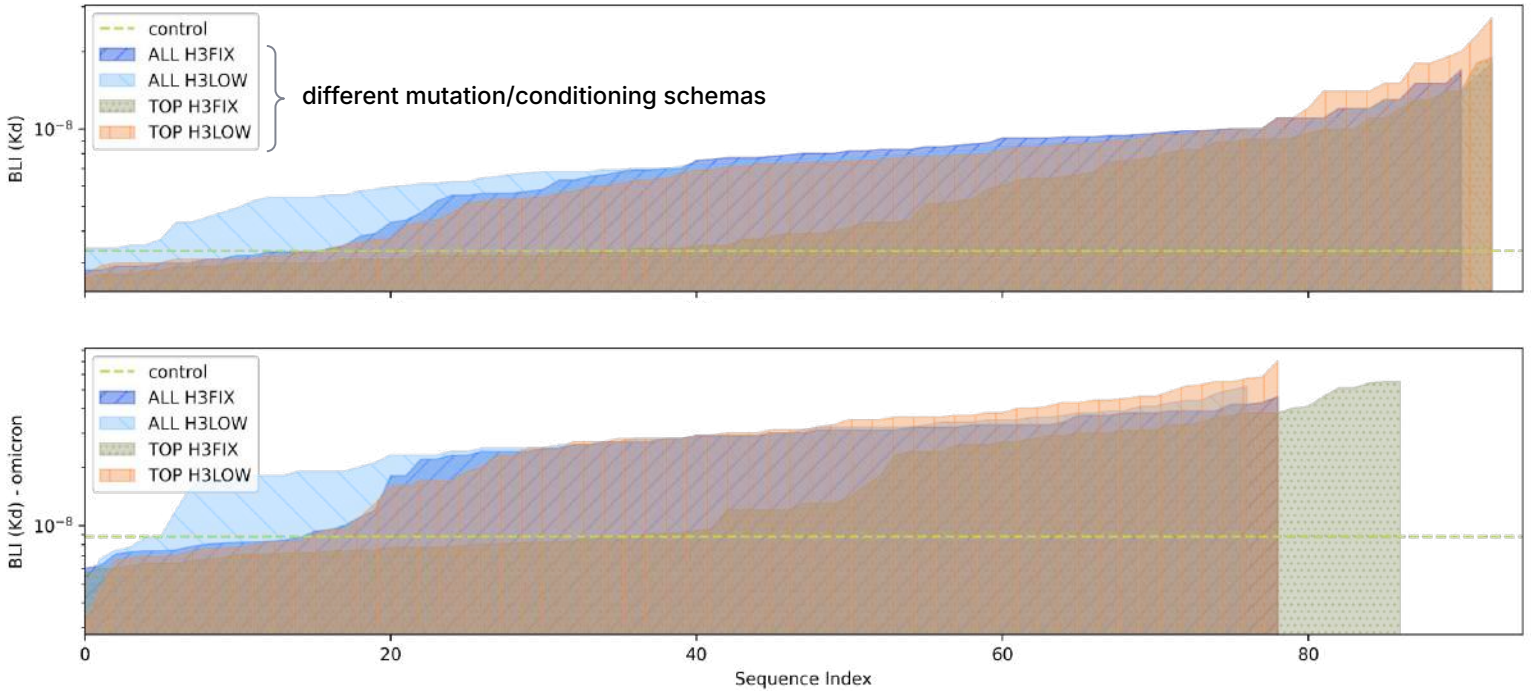


Allowing for insertions, deletions and substitutions across the sequence

- Conditioned generative models on phage display data from immunization campaign
- Generated and selected 384 sequences with mutations across the entire sequence
- Large number of insertions, deletions, and substitutions
- 4 Sub-plates with varying generator conditioning and mutation schema



Over 98% of the tested sequences retained binding



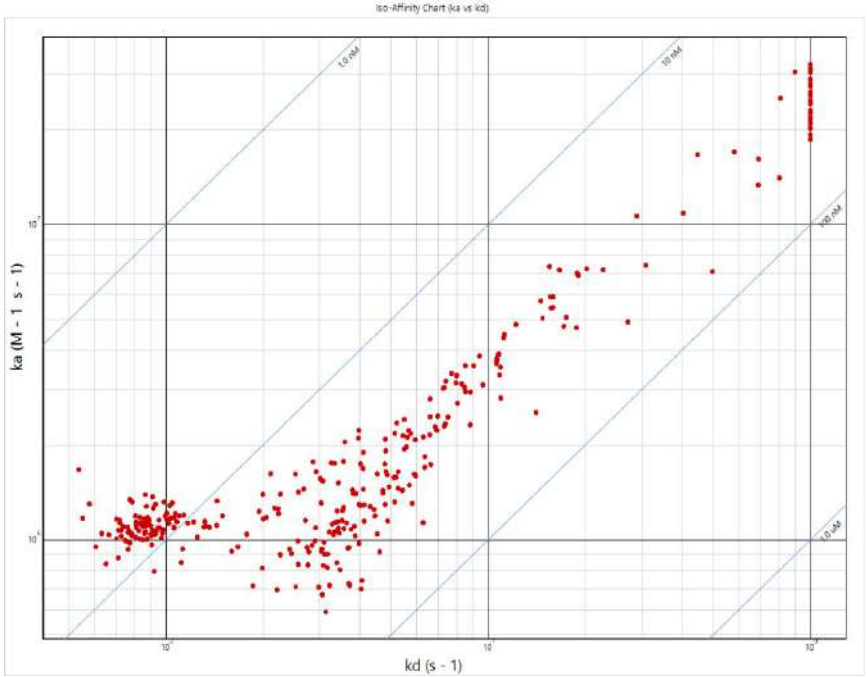
Improved cross-reactivity

Outcome: variants improved cross-reactivity to **Omicron** by almost 3-fold

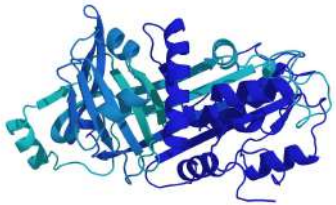
54/384 variants lost affinity to Omicron**

Ty1 VHH control has 2.5-fold weaker affinity against Omicron (8.75nM)

Our variants improved cross-reactivity to Omicron by **almost 3-fold** (3.2nM)



No polyspecific binding was observed



Ovalbumin
polyspecificity reagent

No Velcro
We tested non-specific binding of the variants using ovalbumin, a polyspecificity re-agent.

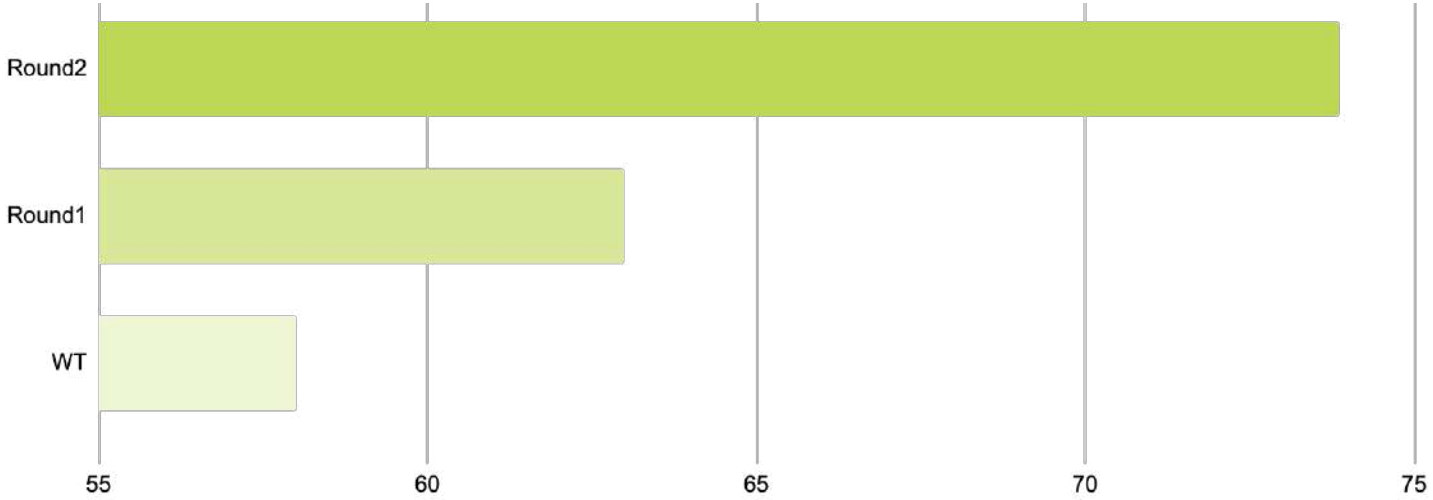
0/384 variants
showed binding to Ovalbumin

Improving developability

Stable, non-aggregating and well expressing antibodies

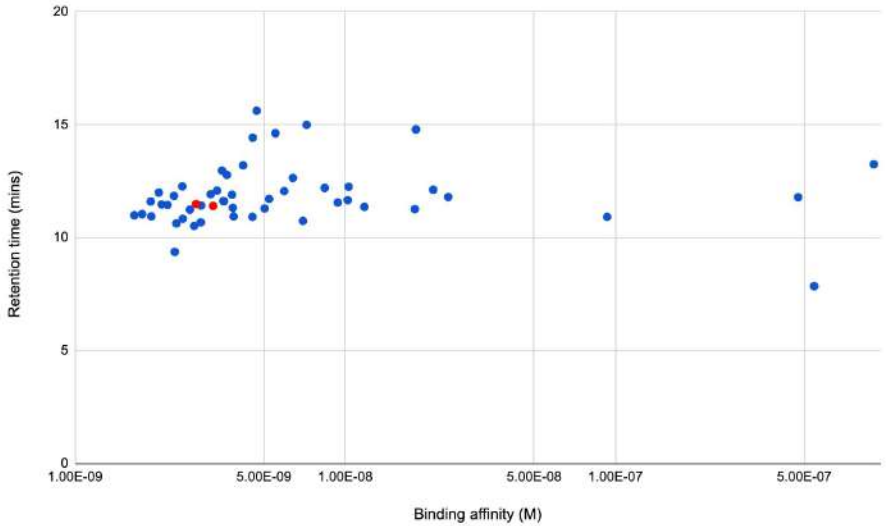
Sars-CoV-2 in vitro experiments: multi-round thermostability

- Conditioned generative model on alignment of similar sequences
- Ran two rounds of 96 sequences
- Substitutions only
- Improved Tm from 58 to 78 °C (28% lift)



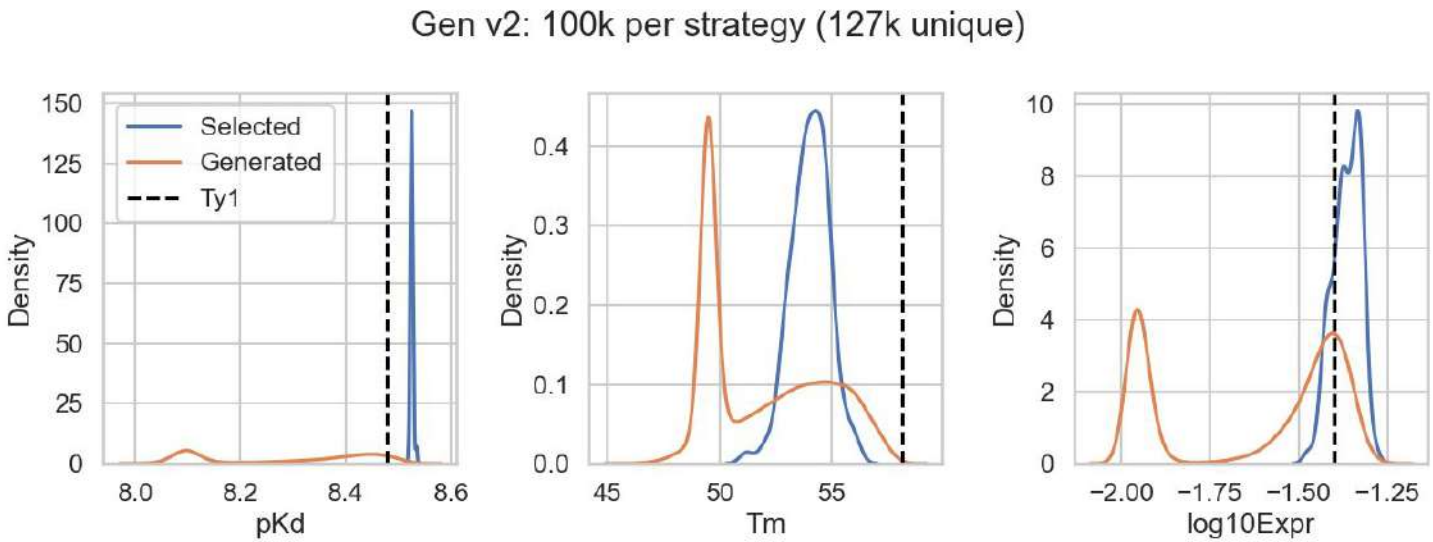
Maintaining favorable hydrophobicity

We observed with a diverse range of hydrophobicity values with similar or improved binding affinity.



What may happen in the future

Our models allow us to estimate what may happen in the future. This is an example of models expected distribution of assay values in the future for thermostability, expression and SARS-CoV-2 binding affinity



Developing a vaccine

Improving stability while maintaining therapeutic efficacy with Johnson & Johnson in Leiden

Case study

Johnson&Johnson
Innovative Medicine

A vaccine against *Staphylococcus aureus*

The pathogenic bacterium *S. aureus* causes upper respiratory and gut infections and was associated with over 1 million deaths in 2019.

Challenge: 'Antigen 1' is a chimeric protein that triggers an effective immune response to *S. aureus*. However, it required further stabilization while maintaining therapeutic efficacy.

Result: The recent collaboration between Cradle and J&J Innovative Medicine demonstrated the ability to stabilise and develop 'Antigen 1' in just 1 experimental round. In a similar vaccine development project, identification of a stable toxoid required 7 rounds of designs and 2 years to complete.

Feedback from team was that some of the most non-intuitive mutations were also among the most successful.

S. aureus

Time to target product profile

24mo

Antigen 1
7 rounds

2mo

Antigen 2
1 round

S. aureus project results and proof points

Rounds to completion

Team satisfaction (N=2 avg)

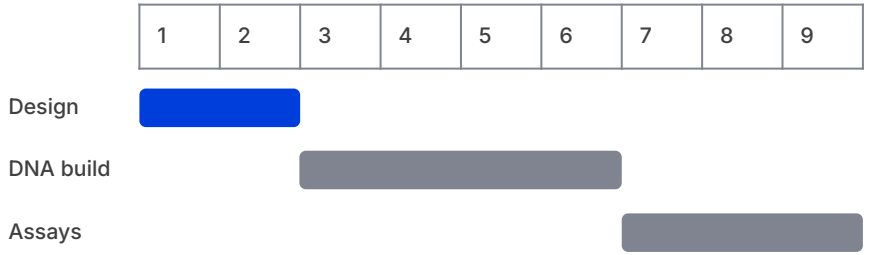
Case study

From 2 weeks to 3 days.

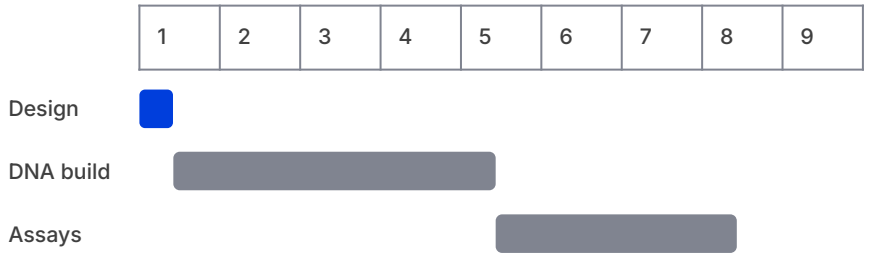
Using Cradle, Johnson & Johnson scientists uploaded lab data on a Friday and received newly generated sequences by Monday.

Before Cradle, the team typically spent 2 weeks designing a new library.

Before Cradle (Weeks per round)

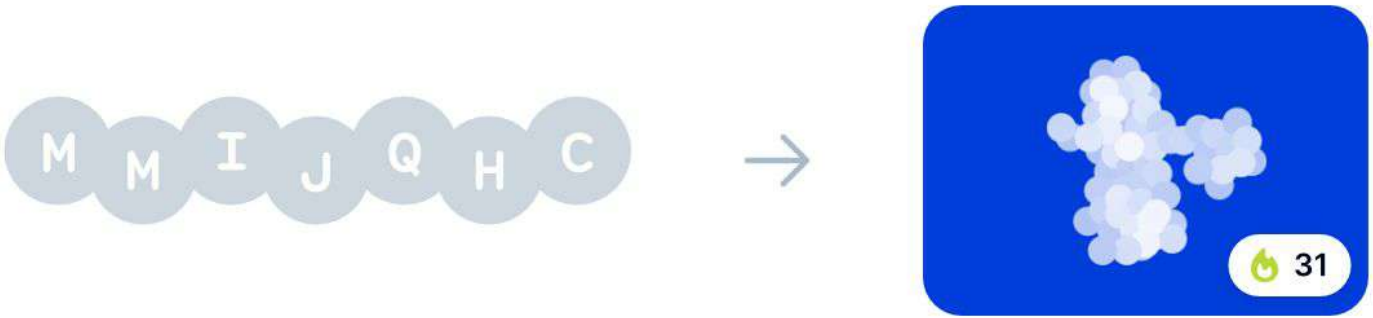


With Cradle (Weeks per round)



Machine learning for drug development
Engineering proteins with Cradle
Examples from antibodies & vaccines
Teaching computers about proteins

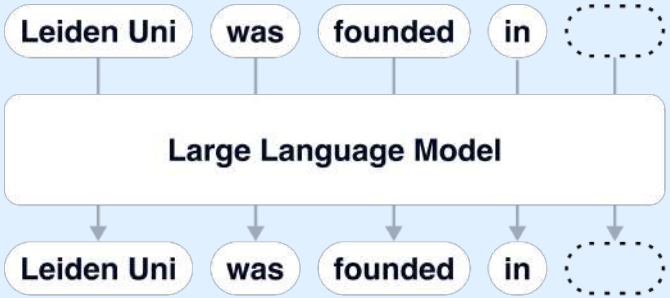
We can teach computers the relationship between a sequence and its function



Protein Language Models

Teaching computers about proteins

English

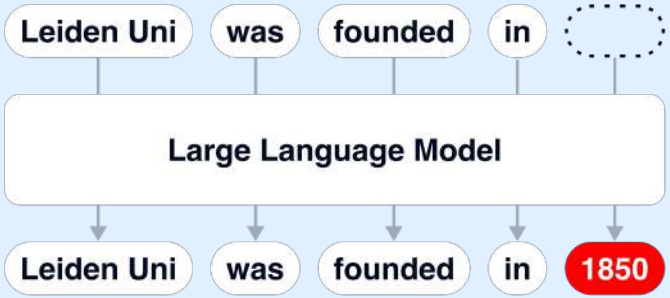


“Proteinish”

Protein Language Models

Teaching computers about proteins

English

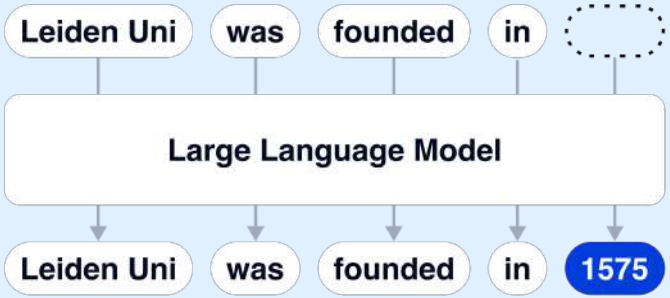


“Proteinish”

Protein Language Models

Teaching computers about proteins

English

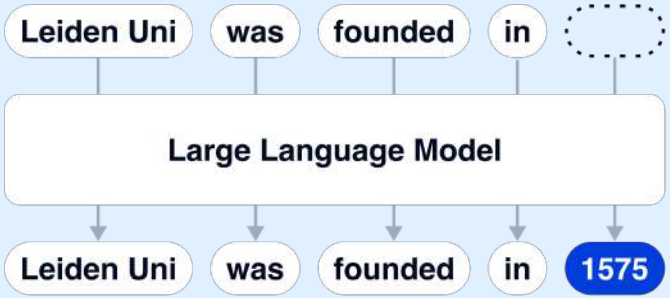


“Proteinish”

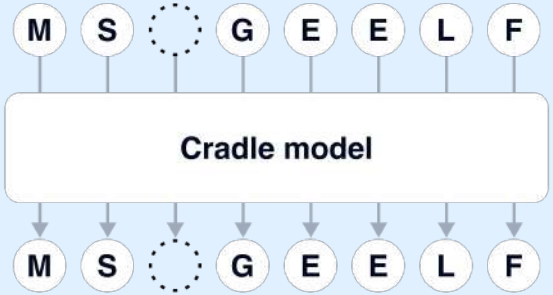
Protein Language Models

Teaching computers about proteins

English



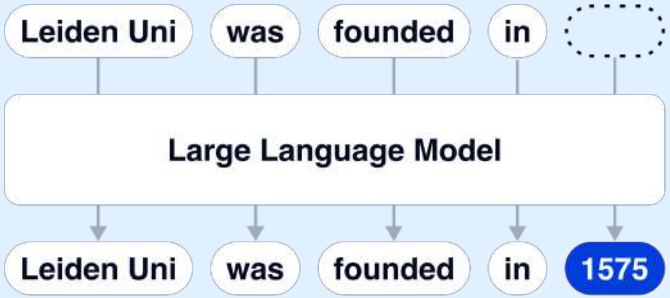
“Proteinish”



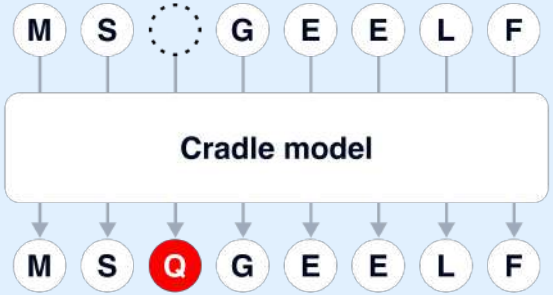
Protein Language Models

Teaching computers about proteins

English



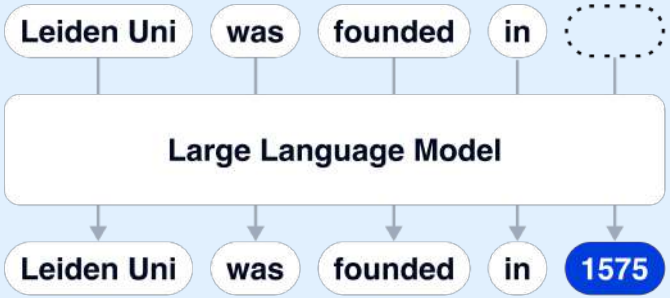
“Proteinish”



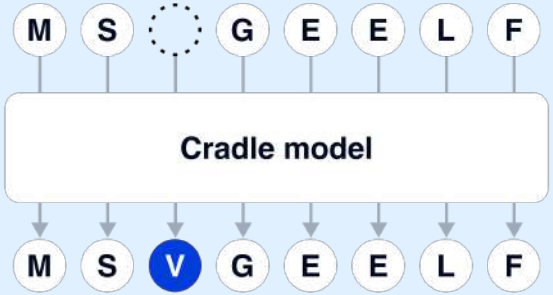
Protein Language Models

Teaching computers about proteins

English



“Proteinish”



Protein Language Models

Teaching computers about proteins

English

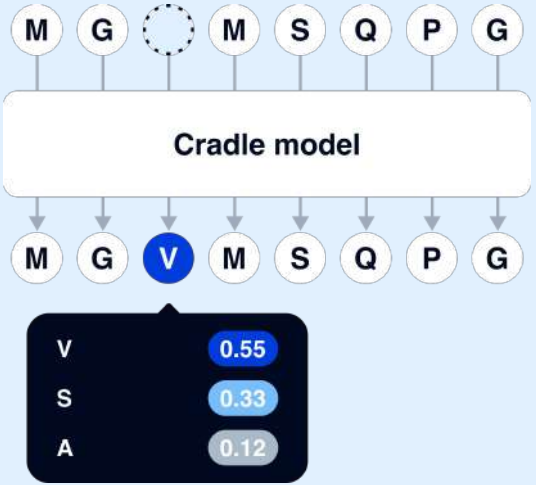
“Proteinish”

Protein Language Models

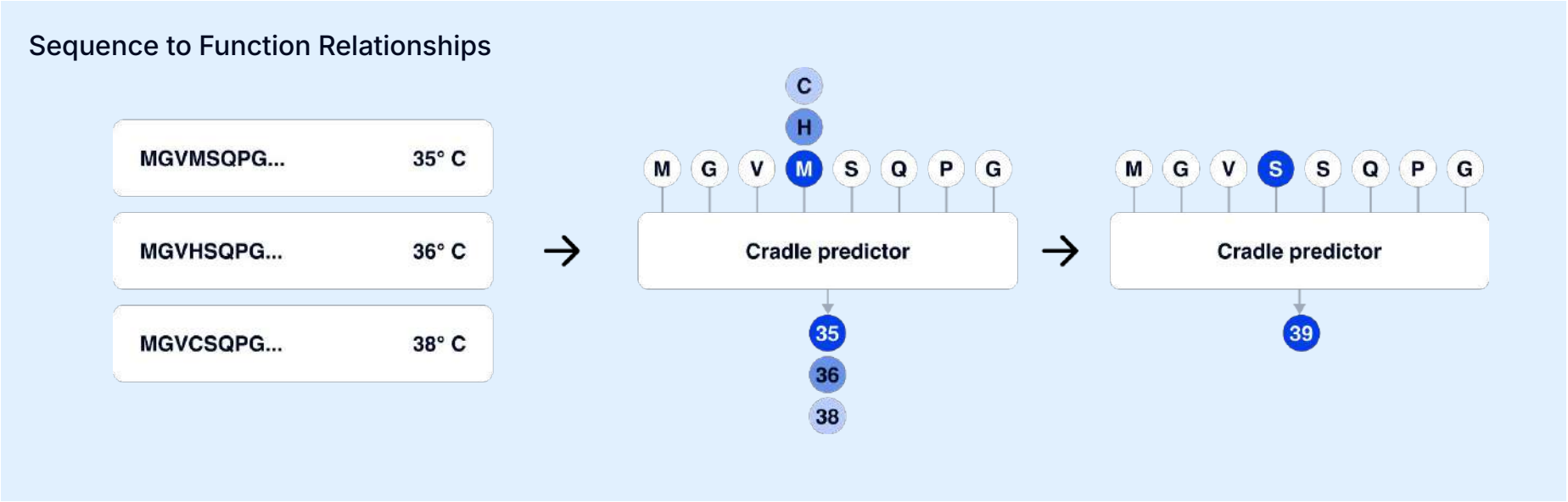
Teaching computers about proteins

English

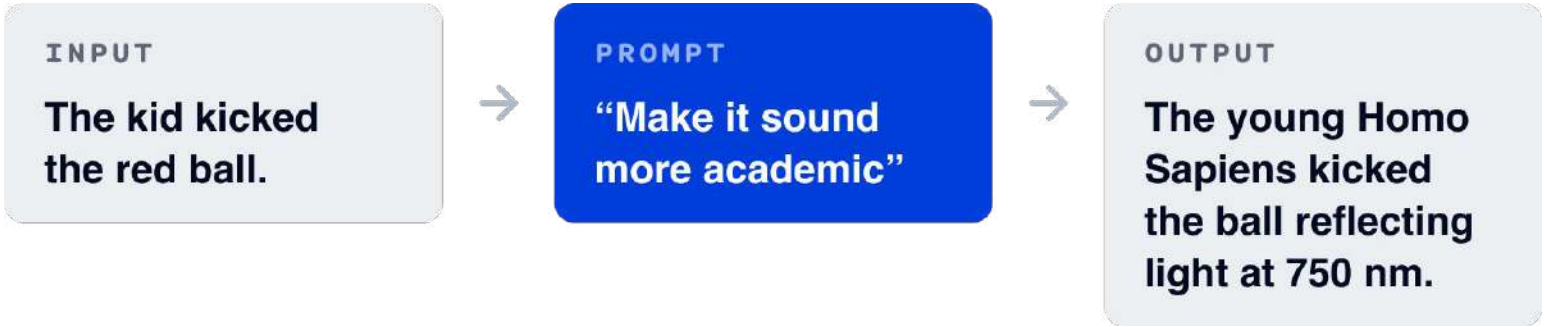
“Proteinish”



Learning functional relationships with lab data

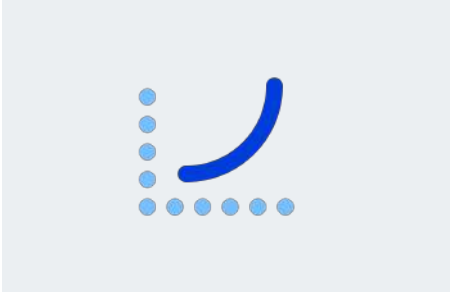


From prediction to generation



From prediction to generation

Cradle is ready to help you with your next project!



Accelerate your existing project

For ongoing or new projects we can help you reach your target profile in less rounds and reduce your R&D spend.

Start a new project

For new projects in antibodies, vaccines, peptides or proteins you can get started tomorrow. We only need an immunization campaign or a starting sequence.

THANKS

THANKS

THANKS

Website cradle.bio

Email stef@cradle.bio