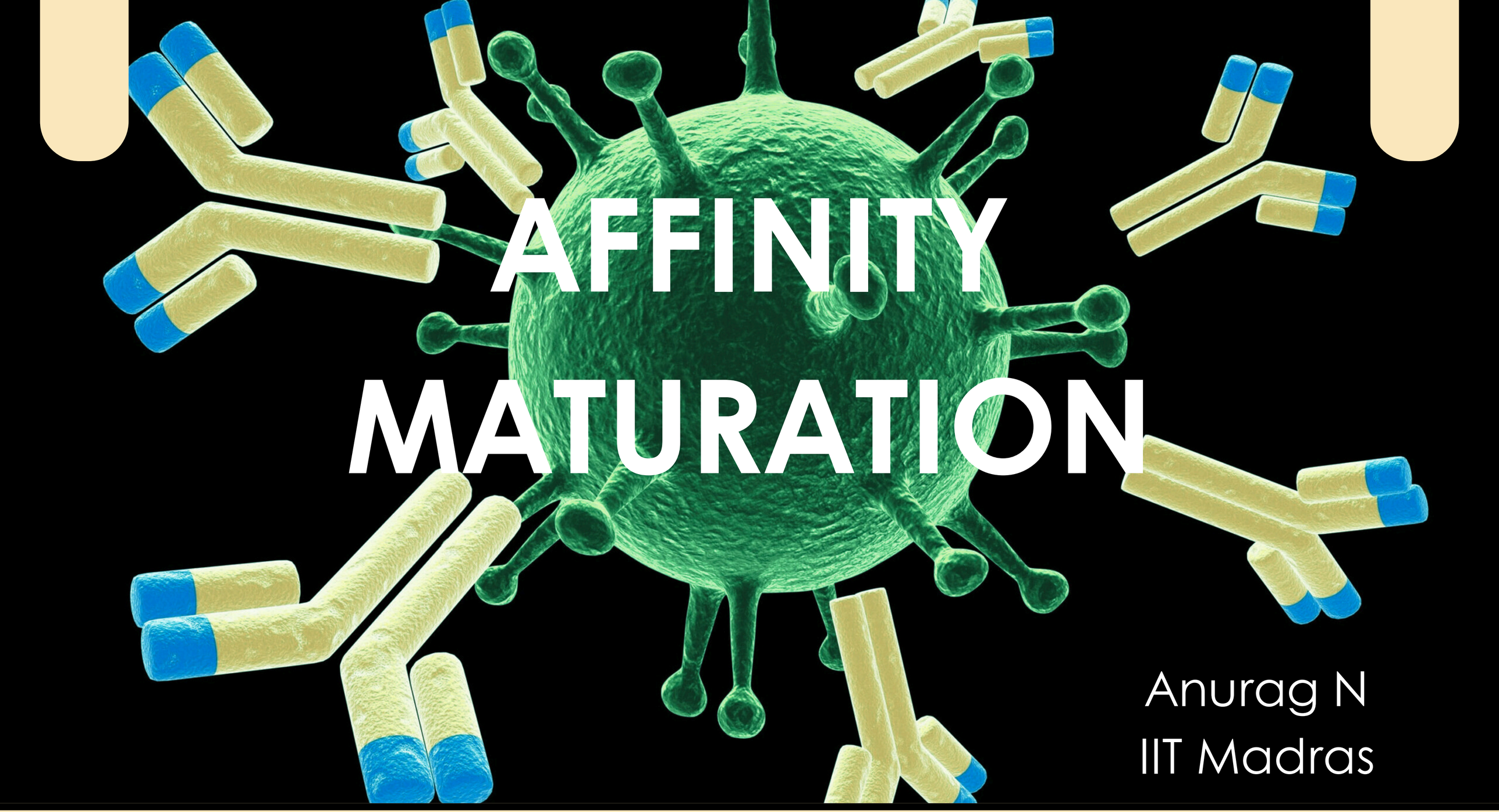


AFFINITY MATURATION

Anurag N
IIT Madras

Affinity maturation is the process by which activated B cells produce antibodies with increased affinity for antigen during the course of an immune response.

Affinity maturation aims to enhance binding affinity, specificity and stability.

BACKGROUND

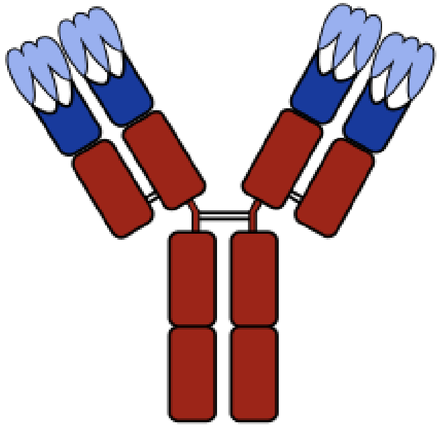
B cells which are part of the adaptive immune system get activated and proliferate during an immune response.

The B cell receptor(BCR) has:

- A Variable region which interacts with the antigen
- A Constant region which determines its class

Changes in the variable region affect how the B Cell interacts with the antigen.

BACKGROUND



The variable regions with heavy and light chains have polypeptide segments called **Complementarity-determining regions (CDRs)** which bind to their specific antigen and this segment determines the binding activity of the respective antibody.

IN VIVO

The process of affinity maturation occurs naturally in the body.

Somatic Hypermutation:

- **Activation-induced (cytidine) deaminase (AID)** produced in the activated germinal center B cells deaminates cytidine to uridine in the DNA sequence of the antibody.
- This change is corrected by **mismatch repair/ base excision** where DNA Polymerase inserts a new base into the DNA strand.
- Mutations like this occur constantly in the B cells which can result in better antibody-antigen binding.

IN VITRO

- Random mutations inside the CDRs are introduced using radiation, chemical mutagens or error-prone PCR.
- Diversity can be increased by **chain shuffling** where a fixed heavy chain is combined with a library of light chains, or vice versa, to create new antibody variants.
- Variants are selected based on their performace in simulated environments.

IN SILICO

- In vivo and in vitro methods are time consuming as they are governed by the biological rate.
- With the advancements in protein structure prediction, molecular dynamics, and algorithms to identify and introduce mutations, we can enhance binding without relying solely on traditional, time-consuming experimental methods.
- These are faster, cheaper, safer than lab-based methods.
- These work in complementary to wet-lab optimization efforts.

IN SILICO

There are multiple strategies that can be used to perform insilico analysis.

- Sequence-based optimization
- Structure-based optimization

These strategies optimize the CDR regions of the proteins of interest and use a metric to track the performance of the mutations.

Structure-based Optimization

What it uses:

- 3D structure of the antibody (often with the antigen)
- Derived from: X-ray, Cryo-EM, AlphaFold, homology modeling

Goal:

- Analyze physical interactions between antibody and antigen
- Predict the impact of mutations on binding energy, stability, and structure

Structure-based Optimization

How it works:

This approach analyzes the atomic interactions at the binding site.

It directly simulates how the mutation:

- Changes hydrogen bonding, salt bridges or hydrophobic contacts
- Introduces/removes steric clashes

It calculates $\Delta\Delta G$ of binding and folding.

Structure-based Optimization

Tools:

- FoldX, Rosetta, PyRosetta
- Molecular Dynamics (GROMACS, AMBER)
- SwissSidechain, DynaMut, HADDOCK

Outputs:

- Binding ΔΔG (change in binding energy upon mutation)
- Interface maps, H-bonds, salt bridges
- Structural stability (RMSD, flexibility, packing)

Structure-based Optimization

Strengths:

- Physically interpretable (can see why a mutation works)
- Captures steric clashes, H-bonds, buried hydrophobic packing
- Useful for rational design of CDR loops and hotspot residues

Limitations:

- Requires accurate or modeled 3D structure
- Computationally expensive
- May not capture long-range epistasis or coevolution

Sequence-based Optimization

What it uses:

- Primary amino acid sequence (FASTA)
- Antibody repertoire data, multiple sequence alignments (MSA), deep mutational scanning

Goal:

- Predict effects of mutations using learned patterns, co-evolution, and statistical or ML models

Sequence-based Optimization

How it works:
It looks for statistical patterns in antibody sequences known to bind well. Some models are trained on millions of natural or synthetic antibody sequences.

"Antibodies with a glycine at position 99 in HCDR3 tend to bind better, so try that."

These methods don't need a structure. They are based on:

- Evolutionary conservation (e.g., consensus sequence design)
- Machine learning (deep learning models trained on datasets of binders)
- Language models (e.g., AntiBERTy, ESM) that treat sequences like sentences

Sequence-based Optimization

Tools:

- DeepAb, ProAffiMuSeq, AbLSTM, AntiBERTy
- Consensus design, PSSMs, LSTM/transformers
- Language models (e.g., ESM, ProtBERT)

Outputs:

- Predicted ΔΔG or binding score
- Likelihood of mutation being beneficial

Sequence-based Optimization

Strengths:

- Fast and scalable (no need for structure)
- Can learn from massive datasets (e.g., phage display)
- Effective for library design, screening, or low-resource settings

Limitations:

- It has limited interpretability (a ‘black box’)
- No physical insight (doesn’t show clashes or H-bond formation)
- Less accurate for rare or novel CDR structure

What to use when?

Scenario	Best Approach
You have a solved or modeled 3D structure	Structure-based
You only have antibody sequences	Sequence-based
You want quick mutation scanning	Sequence-based
You want high-resolution energy modeling	Structure-based
You’re training an ML model on screening data	Sequence-based
You want to rationally design loop mutations	Structure-based

CASE STUDY

Paper:

Towards an optimal monoclonal antibody with higher binding affinity to the receptor-binding domain of SARS-CoV-2 spike proteins from different variants

Methodology:

They ran 3 sets of mutations designed using RAbD(RosettaAntibodyDesign):

- all CDRs considered for full design
- all CDR's L1 regions were considered for full design
- all CDRs except L1 and H3 regions were considered for full design

They generated 91,800 candidates for (A), 72,000 candidates for (B), and 72,000 candidates for (C) scenarios.

CASE STUDY

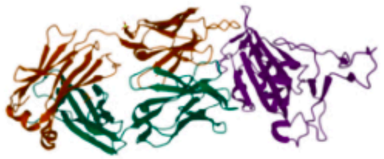
Methodology:

From the entire pool of candidates, they have selected the best-improved mAbs using two criteria:

- the complex must have an Ab-Ag interface score below − 150 REU (ROSETTA Energy Units) (the native interface score is − 65 REU)
- an Ab-Ag interface surface area larger than 1900 Å^2 (the native complex interface surface area is 2060 Å^2).

The top 10 candidates where then further analysed for their binding energies with the receptor-binding domain of SARS-CoV-2 spike proteins.

CASE STUDY



CR3022 + RBD_{wt}

Step 1 — *Rosetta Ab Design*

235000 Designed Abs

Step 2 — *Rosetta Score*

Top 10 Abs

Step 3 — *CG MD with SIRAH FF*

Step 4 — *Fast sampling CG MC with FORTE*

Top 10 Abs + RBD_{wt}

Binding Free Energy

Binding Free Energy

Step 5 — *Comparison of the results*

FORTE Validation + Selection of best candidates

Step 6 — ***FORTE*** *Theoretical Alanine Scanning on* ***Top 2 Abs***

Step 7 — *Fast sampling CG MC with* ***FORTE*** ***Top 10 Abs + $RBD_{VOC/VOI}$***

Step 8 — *Fast sampling CG MC with* ***FORTE*** **Top 2 improved Abs + $RBD_{WT,VOC/VOI}$**

Identification of improved Abs

Evaluation of Best designed Abs active against known VOC and VOI

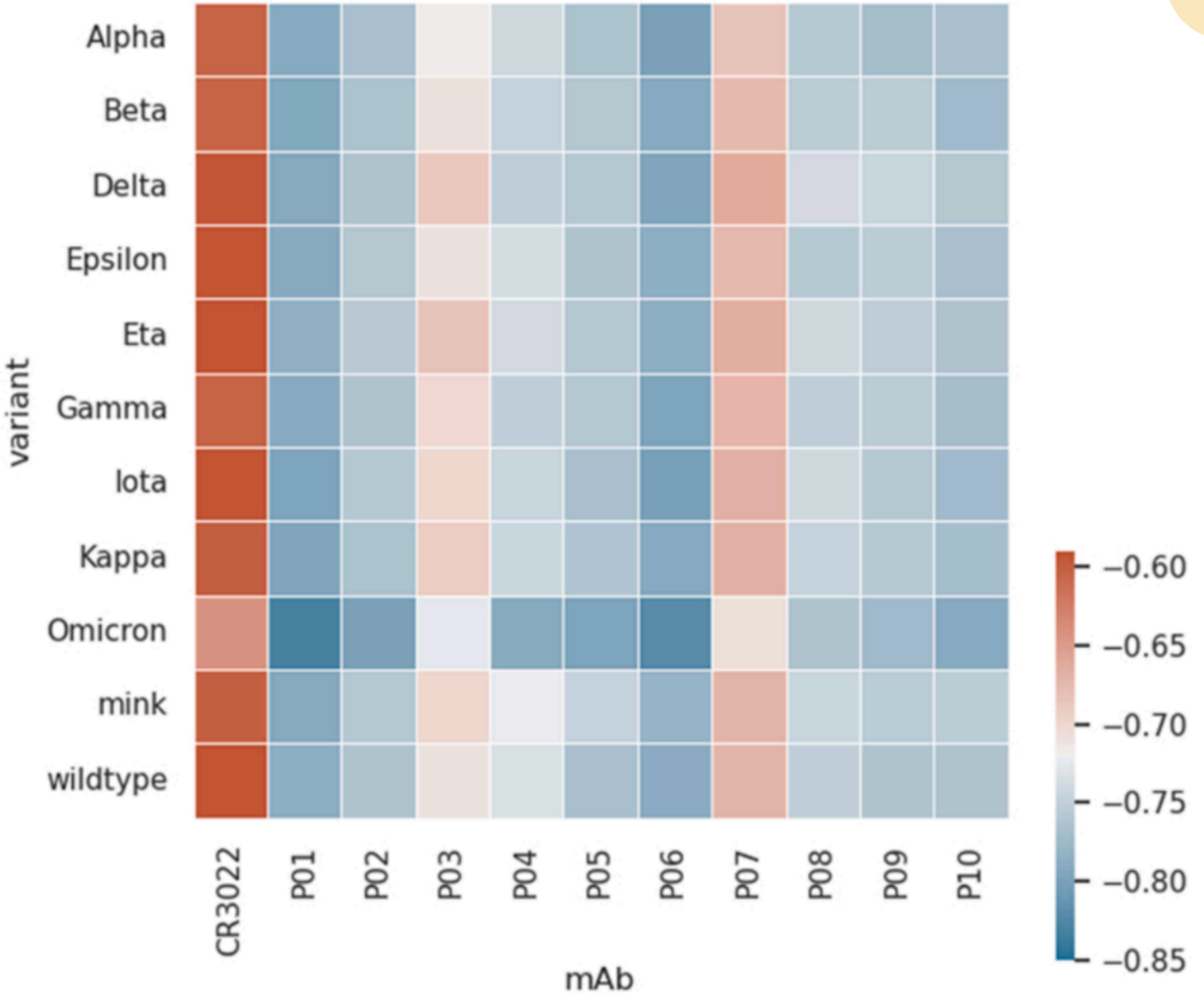
Step 9 — *AA MD of P01-RBD_{wt}*

Identification of structural determinants leading to the high affinity of the Top 1 Ab against RBD_{wt}

CASE STUDY

Results:

Heatmap with the minima binding free energy of interactions values for the SARS-CoV-2 RBD-mAbs complexation at pH 7 and 150 mM of NaCl by the CpH MC simulations (FORTE) for the RBD of the main critical variants and different Rosetta-designer binder candidates (P01 to P10).



CASE STUDY

Results:

Some of the best mutations identified were:

- I34M in CDRH1
- T63S in CDRH2
- I107L in CDRH3
- K36M in CDRL1
- S58T in CDRL2
- Y97F in CDRL3

CASE STUDY

Antibody	Target	Type of antibody	Structure of antibody	Affinity Maturation Software	Increased affinity	Year	Ref
cAb-CA05	HEWL	VHH	Model by MOE	GROMACS, MolFeat-EC	20x	2013	[173]
Bevacizumab	VEGF-A	Fab fragment	Crystal structure, Model by Modeller	PyMOL v1.3, GROMACS 4.0.7	18x	2013	[174]
11K2	MCP-1	ScFv	Crystal structure	GROMACS	4.6x	2014	[175]
bH1	VEGF-A HER2	Fab fragment	Crystal structure	SIE-SCWRL v4.0, FOLD X v3.0, Rosettav3.5	104x 46x	2017	[165]
D8	HBV	ScFv	Model by Modeller	PatchDock, FOLD X v3.0	10x	2018	[176]
A26.8	TcdA	VHH	Crystal structure	SIE-SCWRL, FOLD X, Rosetta	9x	2018	[167]
NbC18	PIGF	VHH	Model by Modeller	Cluspro, HADDOCK, GROMACS v 5.1.2	2.1x	2018	[177]
V_HNAC1	Human α-synuclein	VHH	Model by Swiss-Modeller	GROMACS, ClusPro, DelPhi software	48x	2018	[79]
AB1	muCCL20	ScFv	Model by SabPred	Discovery Studio 2016, Schrödinger Biologics Suite2016-3, Rosetta	4x	2019	[178]
Nb02	CD47	VHH	Model by Modeller v9.19	SIE-SCWRL, FOLD X, Rosetta, HADDOCK, Schrodinger	87.4x	2019	[140]
D2-L29	HEL	VHH	Crystal structure	MOE	12x	2020	[179]
VHH212	HIF-1α	VHH	Model by Modeller	GROMACS v4.5.5, HADDOCK, InterProSurf , mCSM-AB, OSPREY , FoldX	17.5x	2020	[150]
Adalimumab	TNF-α	Fab fragment	Model by PIGSPro v2 and ABodyBuilder	PatchDock, FireDock, ClusPro 2.0, HADDOCK 2.2, PRODIGY, YASARA, mCSM-AB etc.	82x	2021	[180]
E53	DENV	scFv	Crystal structure	Discovery Studio 4.0	100x	2021	[181]
AP2M21	hPCSK9	ScFv	Model by ABodyBuilder server	HawkDock, PyMol v2.3.0	24.2x	2021	[182]
KD035	VEGFR2(D2-3)	Fab fragment	Crystal structure	MOE, NAMD	8.2x	2021	[183]
CR3022	SARS-CoV-2 RBD	Fab fragment	Crystal structure	Discovery Studio 4.5, GROMACS 5.1.2	15x	2022	[184]
VHH2	TNF-α	VHH	Model by Alphafold2	GROMACS 2019, HADDOCK, FOLD X	3.9x	2023	[151]

Source: Affinity maturation of antibody fragments: A review encompassing the development from random approaches to computational rational optimization

Novel Method - Inverse Folding

Inverse folding involves designing optimal amino acid sequences for a fixed 3D structure.

Key innovation:
Generative modeling of sequences that fold into the desired CDR loop architecture.

Tools:

- Antifold: Language model that performs inverse folding using masked tokens.
- ProteinMPNN: Autoregressive model that proposes probable sequences for fixed backbones.

Advantages:
The predictied mutations are more biologically relevant due to structure preservation.

Antifold Model

AntiFold accepts a solved or predicted antibody variable domain structure as input.

It has been trained by fine-tuning parameters of ESM-IF1 which had more than 12 million protein structures.

The user may specify regions to sample new sequences for and a temperature parameter to control sequence diversity.

Outputs:

- The overall tolerance to mutations without altering the backbone structure (perplexity)
- Individual amino acid probabilities

ProteinMPNN Model

How It Works:

- Input: Protein backbone (no side chains).
- Output: High-likelihood sequences that support that structure.
- Uses a graph-based neural network with message-passing layers.

Advantages:

- Extremely fast (milliseconds per design).

Antifold Model

Ranking performances of different Inverse Folding Models

Evaluation	Dataset	Model			
		ProteinMPNN	***ESM-IF1***	***AbMPNN***	***AntiFold***
Amino acid recovery, CDRH3 (%, ↑)	AbMPNN test-set	35%	43%	56%	**60%**
Sequence design, sampled CDR loops (RMSD, ↓)	AbMPNN test-set	1.03	1.01	0.98	**0.95**
AbAg binding affinity (S_r, ↑)	Warszawski, anti-lysozyme DMS	0.30	0.32	0.33	**0.42**
AbAg improved variants (Rank %, ↑)	Hie et al., 7x Ab affinity-maturation	73%	57%	55%	**80%**

Pipeline Planned

Input:
PDB file (If only FASTA available, predict structure using Alphafold)

Models to be Used:
AntiFold, ProteinMPNN and Rosetta Antibody Design (RAbD) to predict n number of sequences.

Predict mutant Ab Structures using Alphafold for the generated sequences.

Run FoldX/Rosetta Interaction analysis for the new structures to measure change in binding affinity.

Work done so far

- Script to generate n number of sequences using Antifold is working.
- Script to generate n number of sequences using ProteinMPNN is working.
- Script to run Alphafold locally to predict protein structures is ready.

Work done so far

Outputs generated:

- Light Chain of Herceptin

	27	28	29	36	37	38	56	57	65	105	106	107	108	109	114	115
Original	Q	D	V	N	T	A	S	A	S	Q	Q	H	Y	T	T	P
Seq 1	Q	D	V	G	T	A	S	A	S	Q	Q	Y	Y	S	T	P
Seq 2	Q	D	V	G	T	A	S	A	S	Q	Q	Y	Y	S	T	P
Seq 3	Q	G	V	G	T	A	S	A	S	Q	Q	Y	Y	S	T	P
Seq 4	Q	D	V	G	T	A	S	A	S	Q	Q	Y	Y	S	T	P
Seq 5	Q	G	V	G	T	A	S	A	S	Q	Q	Y	Y	T	T	P
Seq 6	Q	D	V	G	T	A	S	A	S	Q	Q	Y	Y	S	T	P
Seq 7	Q	D	V	G	T	A	S	A	S	Q	Q	Y	Y	S	T	P
Seq 8	Q	D	V	G	T	A	S	A	S	Q	Q	Y	Y	S	T	P
Seq 9	Q	D	V	G	T	A	S	A	S	Q	Q	Y	Y	S	T	P
Seq 10	Q	D	V	G	T	A	S	A	S	Q	Q	Y	Y	S	T	P
Seq 11	Q	D	V	G	T	A	S	A	S	Q	Q	Y	Y	S	T	P
Seq 12	Q	D	V	G	T	A	S	A	S	Q	Q	Y	Y	S	T	P
Seq 13	Q	D	V	G	T	A	S	A	S	Q	Q	Y	Y	S	T	P
Seq 14	Q	D	V	G	T	A	S	A	S	Q	Q	Y	Y	S	T	P
Seq 15	Q	D	V	G	T	A	S	A	S	Q	Q	Y	Y	S	T	P

Work done so far

Outputs generated:

- Heavy Chain of Herceptin

	27	28	29	30	35	36	37	38	56	57	58	59	62	63	64	65	105	106	107	108	109	110	111	112	113	114	115	116
Original	G	F	N	I	K	D	T	Y	I	Y	P	T	N	G	Y	T	S	R	W	G	G	D	G	F	Y	A	M	D
Seq 1	G	F	N	I	S	D	T	Y	I	Y	P	S	N	G	Y	T	A	N	E	Y	G	Y	Y	Y	Y	A	M	D
Seq 2	G	F	N	I	S	D	T	Y	I	Y	P	T	N	G	Y	T	A	S	E	Y	G	W	Y	Y	Y	A	M	D
Seq 3	G	F	N	I	S	D	T	Y	I	Y	P	T	N	G	Y	T	A	N	E	L	G	Y	Y	Y	Y	A	M	D
Seq 4	G	F	N	I	S	D	T	Y	I	Y	P	T	N	G	Y	T	A	T	E	Y	G	Y	Y	Y	Y	A	M	D
Seq 5	G	F	N	I	S	D	T	Y	I	Y	P	T	N	G	Y	T	A	N	E	L	G	Y	Y	Y	Y	A	M	D
Seq 6	G	F	N	I	S	D	T	Y	I	Y	P	T	N	G	Y	T	A	T	E	Y	G	Y	Y	Y	Y	A	L	D
Seq 7	G	F	N	I	S	D	T	Y	I	Y	P	T	N	G	Y	T	A	T	E	F	G	Y	Y	Y	Y	Y	L	D
Seq 8	G	F	N	I	S	D	T	Y	I	Y	P	A	N	G	Y	T	A	T	D	Y	G	W	Y	Y	Y	A	L	D
Seq 9	G	F	N	I	S	D	T	Y	I	Y	P	T	N	G	Y	T	A	T	E	Y	G	W	Y	Y	Y	A	M	D
Seq 10	G	F	N	I	S	D	T	Y	I	Y	P	A	N	G	Y	T	A	N	E	Y	G	Y	Y	Y	Y	A	M	D
Seq 11	G	F	N	I	S	D	T	Y	I	Y	P	A	N	G	Y	T	A	S	E	Y	G	Y	Y	Y	Y	Y	M	D
Seq 12	G	F	N	I	S	D	T	Y	I	Y	P	A	N	G	Y	T	A	N	E	Y	G	Y	Y	Y	Y	A	M	D
Seq 13	G	F	N	I	S	D	T	Y	I	Y	P	A	N	G	Y	T	A	T	E	Y	G	Y	Y	Y	Y	A	L	D
Seq 14	G	F	N	I	S	D	T	Y	I	Y	P	A	Y	G	Y	T	A	T	E	Y	G	Y	Y	Y	Y	A	M	D
Seq 15	G	F	N	I	S	D	T	Y	I	Y	P	A	N	G	Y	T	A	N	E	Y	G	Y	Y	Y	Y	A	M	D

Work done so far

Outputs generated:

- log probabilities of mutation

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A
1	pdb_posins	pdb_chain	pdb_res	top_res	pdb_pos	perplexity	A	C	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W	Y	
2	1	B	E	M	1	1.4246	-5.4437	-7.5128	-5.8774	-5.8908	-7.0791	-4.5803	-6.737	-6.3564	-6.1898	-6.1864	-0.0556	-6.314	-5.2712	-6.2177	-6.1007	-5.0936	-5.7518	-5.4498	-8.1377	-7.6712		
3	2	B	V	V	2	1.0576	-9.4436	-12.1739	-12.8297	-8.9264	-11.8523	-11.1808	-12.6589	-4.7857	-11.2595	-8.2079	-9.3734	-11.5178	-11.7242	-10.3786	-11.5629	-13.9007	-9.6186	-0.0091	-11.7512	-12.069	
4	3	B	Q	Q	3	1.3579	-7.4898	-10.1733	-8.7166	-5.3761	-10.0227	-9.2407	-5.6702	-10.8549	-3.6848	-5.4162	-6.6697	-6.6712	-14.9773	-0.056	-4.5883	-8.3528	-7.2578	-6.127	-11.5208	-10.2835	
5	4	B	L	L	4	1.0018	-14.8226	-13.0936	-18.165	-13.3251	-12.0759	-16.1596	-15.8764	-11.5162	-15.4581	-0.0002	-9.1842	-17.002	-16.6845	-13.2437	-15.2965	-16.5375	-17.2368	-10.0609	-14.2967	-14.5446	
6	5	B	V	V	5	1.4596	-8.0742	-12.5306	-9.9095	-7.096	-11.0805	-11.777	-10.7384	-7.1538	-9.3452	-2.2728	-7.0053	-10.3985	-17.4751	-5.9478	-9.235	-10.416	-8.4645	-0.1154	-11.6766	-12.8229	
7	6	B	E	E	6	1.037	-9.9352	-11.8518	-9.1194	-0.0056	-18.0635	-10.1509	-12.9984	-14.3581	-11.6321	-12.8893	-12.4682	-11.0066	-12.0507	-5.245	-12.6855	-12.4557	-11.1796	-10.1175	-16.9983	-17.425	
8	7	B	S	S	7	1.0112	-8.5304	-12.323	-11.6044	-12.1497	-9.4227	-9.1618	-11.6864	-14.578	-11.6641	-11.4465	-12.8496	-11.4021	-15.511	-11.894	-8.9413	-0.0012	-7.5123	-13.505	-12.2436	-10.2063	
9	8	B	G	G	8	1.0032	-11.1719	-14.2084	-10.6695	-8.6287	-14.9323	-0.0003	-14.3702	-14.4391	-11.6173	-12.8262	-15.4072	-13.1575	-16.9041	-11.8195	-9.7296	-12.8203	-14.3756	-12.1946	-14.634	-14.823	
10	9	B	G	G	9	1	-15.0124	-18.6258	-15.9849	-15.416	-19.7803	0	-21.1393	-20.4298	-18.4566	-17.3566	-21.9295	-20.1057	-19.1834	-18.9667	-14.3371	-16.0793	-18.6935	-17.2207	-18.8946	-20.6485	
11	11	B	G	G	11	1.0005	-14.0254	-16.2285	-11.2322	-12.7179	-15.583	0	-17.2611	-15.8515	-14.7871	-12.4239	-16.5867	-15.7953	-17.8111	-15.0853	-11.3672	-13.1814	-14.9871	-13.0274	-15.9735	-17.0993	
12	12	B	L	L	12	1.0083	-11.7213	-14.2436	-14.3295	-11.6333	-12.1142	-13.9393	-13.2219	-11.2776	-10.4375	-0.0008	-9.6892	-13.8178	-9.1902	-9.8265	-8.4539	-12.4222	-12.2996	-8.0011	-14.0312	-14.9945	
13	13	B	V	V	13	1.0271	-7.9193	-9.7535	-13.3894	-9.4947	-13.6983	-12.6306	-13.5319	-6.5868	-9.1153	-8.4065	-7.7155	-11.2793	-14.4915	-9.2019	-9.4045	-9.7007	-8.6267	-0.0031	-16.0401	-14.9836	
14	14	B	Q	Q	14	2.1058	-6.506	-12.0573	-11.264	-5.4045	-11.5025	-12.6843	-6.904	-8.4547	-1.6089	-7.6119	-7.3077	-7.7585	-7.6309	-0.2854	-3.4273	-7.5091	-5.5143	-6.3229	-11.4652	-10.5183	
15	15	B	P	P	15	1.0092	-8.2771	-13.9097	-13.5085	-10.9235	-14.6275	-12.86	-14.1583	-11.9643	-12.69	-9.282	-12.5626	-14.0516	-0.0009	-10.9315	-11.0189	-8.2286	-8.9567	-9.0523	-15.6313	-15.575	
16	16	B	G	G	16	1.012	-10.6304	-15.8529	-11.9482	-9.011	-16.1844	-0.0013	-13.442	-15.9967	-7.2528	-12.099	-13.686	-13.3912	-15.6089	-11.3457	-7.8157	-10.5994	-10.6661	-12.5925	-12.7136	-15.9582	
17	17	B	G	G	17	1.0003	-14.3418	-19.2345	-18.6762	-16.1278	-21.9462	0	-18.4002	-19.3984	-13.8076	-17.7676	-19.1823	-17.0728	-19.2035	-16.2522	-10.8086	-15.6294	-17.2303	-17.4025	-20.2011	-22.181	
18	18	B	S	S	18	1.0169	-7.4153	-11.8782	-11.0825	-10.9178	-9.0855	-8.5821	-11.8734	-14.8401	-10.3657	-9.9394	-11.4662	-11.6862	-10.8282	-10.6766	-7.6799	-0.0018	-8.3459	-12.9608	-11.3863	-11.7049	
19	19	B	L	L	19	1.0225	-11.8888	-14.4955	-15.7659	-11.2815	-14.0551	-15.1369	-13.1174	-12.0108	-9.2624	-0.003	-8.8489	-13.5165	-14.137	-8.7932	-5.9811	-11.5133	-13.5403	-11.2797	-15.5878	-15.8937	
20	20	B	R	R	20	1.0512	-10.2601	-12.1846	-11.1634	-9.2464	-13.1469	-11.3764	-11.5745	-11.7146	-5.2437	-9.3433	-10.9315	-11.0922	-13.3953	-7.9976	-0.0072	-7.9147	-7.1802	-9.1928	-12.7175	-13.6552	
21	21	B	L	L	21	1.0019	-14.3334	-15.9134	-20.9312	-17.6795	-11.1016	-17.8277	-15.9544	-10.0499	-17.617	-0.0002	-9.9461	-17.7771	-16.6679	-14.9359	-16.5462	-14.948	-18.3725	-9.7326	-14.2014	-16.5424	
22	22	B	S	S	22	1.0562	-7.3022	-10.256	-9.9587	-9.206	-10.7639	-9.8886	-10.9543	-10.3376	-10.1882	-10.6949	-11.4318	-8.7082	-13.0386	-10.3679	-9.0372	-0.0082	-5.0131	-8.7579	-12.9793	-10.9175	
23	23	B	C	C	23	1.0035	-7.964	-0.0004	-16.313	-15.2335	-16.8142	-11.6045	-17.8235	-15.2486	-19.955	-13.935	-15.1134	-16.5862	-18.9208	-17.0242	-17.8061	-11.9818	-13.6625	-11.4606	-17.2049	-20.2313	
24	24	B	A	A	24	1.4783	-0.0871	-11.0269	-8.7256	-5.8591	-11.3194	-6.1484	-9.9744	-8.7237	-7.927	-7.0432	-10.9745	-10.4103	-8.8484	-8.9011	-6.9645	-5.6306	-3.4628	-3.204	-11.2162	-12.687	
25	25	B	A	A	25	1.017	-0.0021	-12.7891	-14.4591	-14.2954	-14.7737	-6.4045	-17.7125	-15.3351	-17.1313	-14.0083	-13.3154	-14.9328	-11.0233	-17.0314	-15.4714	-9.0964	-8.49	-9.0576	-16.2507	-17.7408	
26	26	B	S	S	26	1.0646	-6.3597	-12.8402	-8.1313	-10.5097	-6.643	-9.0569	-9.2158	-9.9486	-9.2764	-8.8299	-11.0626	-7.3995	-14.2096	-9.5652	-7.2647	-0.0078	-6.558	-9.7321	-9.8496	-6.8984	
27	27	B	G	G	27	1.0047	-10.9354	-14.2238	-9.5254	-8.7746	-14.8894	-0.0004	-12.6311	-16.6389	-10.8055	-13.6538	-15.1823	-11.109	-18.0312	-11.1453	-9.0092	-10.7838	-13.8104	-13.9677	-14.0806	-13.6494	
28	28	B	F	F	28	1.0052	-15.3086	-14.8452	-14.5895	-16.5892	-0.0006	-16.6679	-12.0544	-12.141	-16.3581	-9.4819	-13.1148	-14.4225	-16.4845	-17.1412	-16.7042	-13.6609	-14.9084	-12.8811	-12.8736	-7.6634	
29	29	B	N	N	29	3.4656	-5.293	-10.2079	-3.2153	-6.0761	-7.3529	-6.7122	-6.6213	-5.9978	-4.3168	-6.0769	-6.3044	-0.5072	-7.0031	-7.4112	-4.7704	-1.8003	-1.8985	-8.7937	-8.2972	-6.3382	
30	30	B	I	I	30	1.2128	-11.8215	-12.9214	-13.0024	-13.3484	-11.2319	-13.5337	-14.7077	-0.0398	-13.9948	-4.8489	-7.2364	-11.3174	-7.7924	-10.5979	-12.5602	-10.4287	-7.7585	-3.5242	-16.15	-16.3599	
31	35	B	K	S	35	7.6158	-3.5693	-6.8444	-4.5372	-4.6691	-6.0655	-3.677	-5.1112	-3.2965	-1.7603	-6.5436	-5.9277	-2.1843	-11.0414	-5.5685	-2.6546	-0.989	-2.4351	-3.8327	-6.987	-3.3498	
32	36	B	D	D	36	1.4615	-6.8022	-11.3831	-0.0724	-4.2327	-8.4434	-6.3673	-6.5739	-8.9209	-7.2694	-9.0598	-8.9945	-3.5604	-12.6292	-7.6369	-6.3096	-4.5772	-5.1212	-9.0169	-7.5879	-6.1028	
33	37	B	T	T	37	4.2542	-5.4705	-5.8845	-1.8862	-5.6873	-5.2913	-5.8286	-4.8599	-5.2458	-6.5094	-8.6778	-9.0512	-3.3429	-10.7873	-6.7734	-7.3581	-2.1873	-0.5818	-4.4385	-7.238	-2.3601	
34	38	B	Y	Y	38	1.8028	-6.2834	-7.0129	-6.2996	-7.895	-2.6837	-8.5504	-5.3366	-7.4354	-8.1677	-5.7386	-9.1724	-5.2781	-10.7326	-8.789	-7.4524	-4.4464	-6.5482	-6.8612	-3.6097	-0.1388	

Challenges

- Alphafold is very resource intensive(computationally). Online servers have limits on number of requests that can be made.
- Running Alphafold locally will require machines with more computational power (GPUs).
- FoldX is a licensed software. It would require us to purchase the liscense.

DISCUSSION

- Have to analyse the differences in effects of making mutations on just Abs and Ab-Ag complexes.

- All predictions done so far have used crystall structures. I will have to run examples of modelled structures to test the pipeline.

- We can run MD simulations on high score sequences to study them better.

- All these predicted mutant sequences will not guarantee the same results in real-world. They would require us to run in-vitro analysis to measure the actual values.

CONCLUSION

- Affinity maturation is central to therapeutic antibody development.
- In silico techniques accelerate design, reduce cost, and increase precision.
- Sequence-based and structure-based approaches complement each other.
- Inverse folding models (like ProteinMPNN, AntiFold) redefine rational antibody engineering.
- Integrated AI and physics-based pipelines are the future of rapid therapeutic discovery.

REFERENCES

- Høie, M. H., Hummer, A. M., Olsen, T. H., Aguilar-Sanjuan, B., Nielsen, M., & Deane, C. M. (2025). AntiFold: improved structure-based antibody design using inverse folding. Bioinformatics advances, 5(1), vbae202. https://doi.org/10.1093/bioadv/vbae202
- Neamtu, A., Mocci, F., Laaksonen, A., & Barroso da Silva, F. L. (2023). Towards an optimal monoclonal antibody with higher binding affinity to the receptor-binding domain of SARS-CoV-2 spike proteins from different variants. Colloids and surfaces. B, Biointerfaces, 221, 112986. https://doi.org/10.1016/j.colsurfb.2022.112986
- Li, J., Kang, G., Wang, J., Yuan, H., Wu, Y., Meng, S., Wang, P., Zhang, M., Wang, Y., Feng, Y., Huang, H., & de Marco, A. (2023). Affinity maturation of antibody fragments: A review encompassing the development from random approaches to computational rational optimization. International journal of biological macromolecules, 247, 125733. https://doi.org/10.1016/j.ijbiomac.2023.125733
- Shanker, V. R., Bruun, T. U. J., Hie, B. L., & Kim, P. S. (2023). Inverse folding of protein complexes with a structure-informed language model enables unsupervised antibody evolution. bioRxiv : the preprint server for biology, 2023.12.19.572475. https://doi.org/10.1101/2023.12.19.572475

THANK YOU