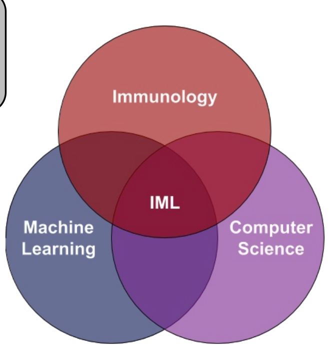
Improving TCR-pMHC Structure Prediction with TCR-Specific Spatial Restraint Features

Joakim Clifford, Postdoctoral Researcher at

Immunoinformatics and Machine Learning Group, Technical University of Denmark (DTU)

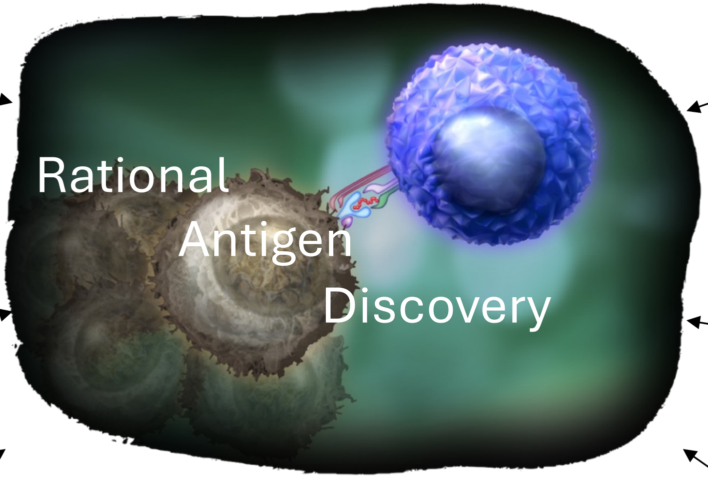
*“Development and application of mathematical models
to characterize important players of the human immune system”*



From sequence data to biomedical insights

B cell epitope prediction

Antigen processing



TCR structure prediction

MHC::peptide binding prediction

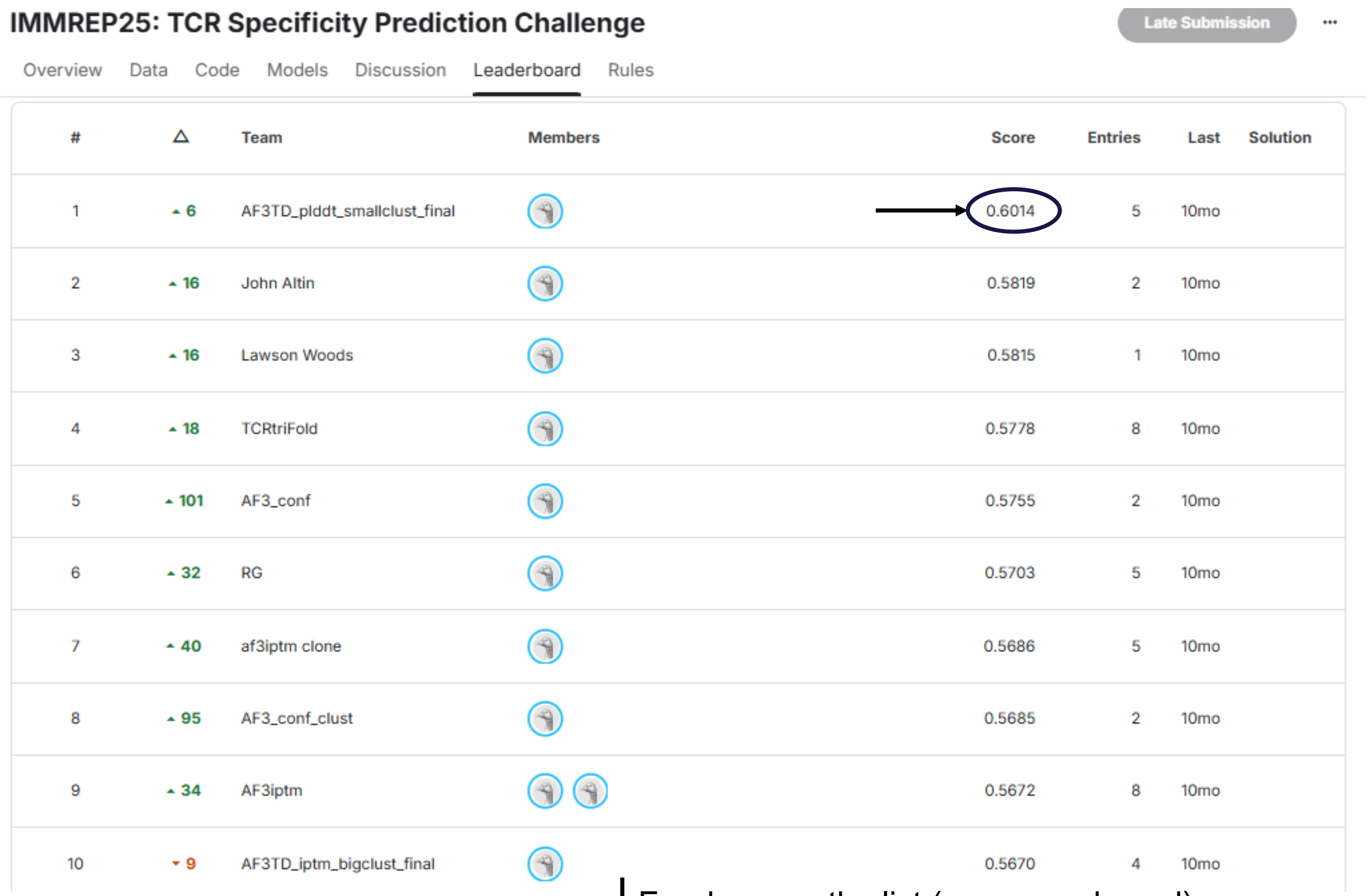
TCR::antigen binding prediction

Protein Immunogenicity Prediction

T-cell epitope prediction

TOOLS:
NetMHCpan
NetMHCIIpan
NetTCR
NetTCRstruct
NetChop
PopCover
GibbsCluster
Bepipred
Discotope
AbEpiTope
BepiPocket/
DiscoPocket

- **IMMREP2025:** Unseen peptides
- Structure based methods (**AF3, Chai1, Boltz1**) dominate leaderboard
- Sequence-based methods from (**winners from IMMREP2023**) were at <u>the bottom.</u>
- Although performance is still not that great
  Best method: AUC01 of 0.6

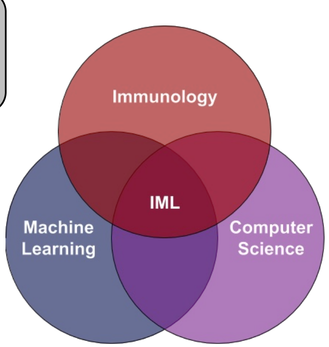


IMMREP25: TCR Specificity Prediction Challenge

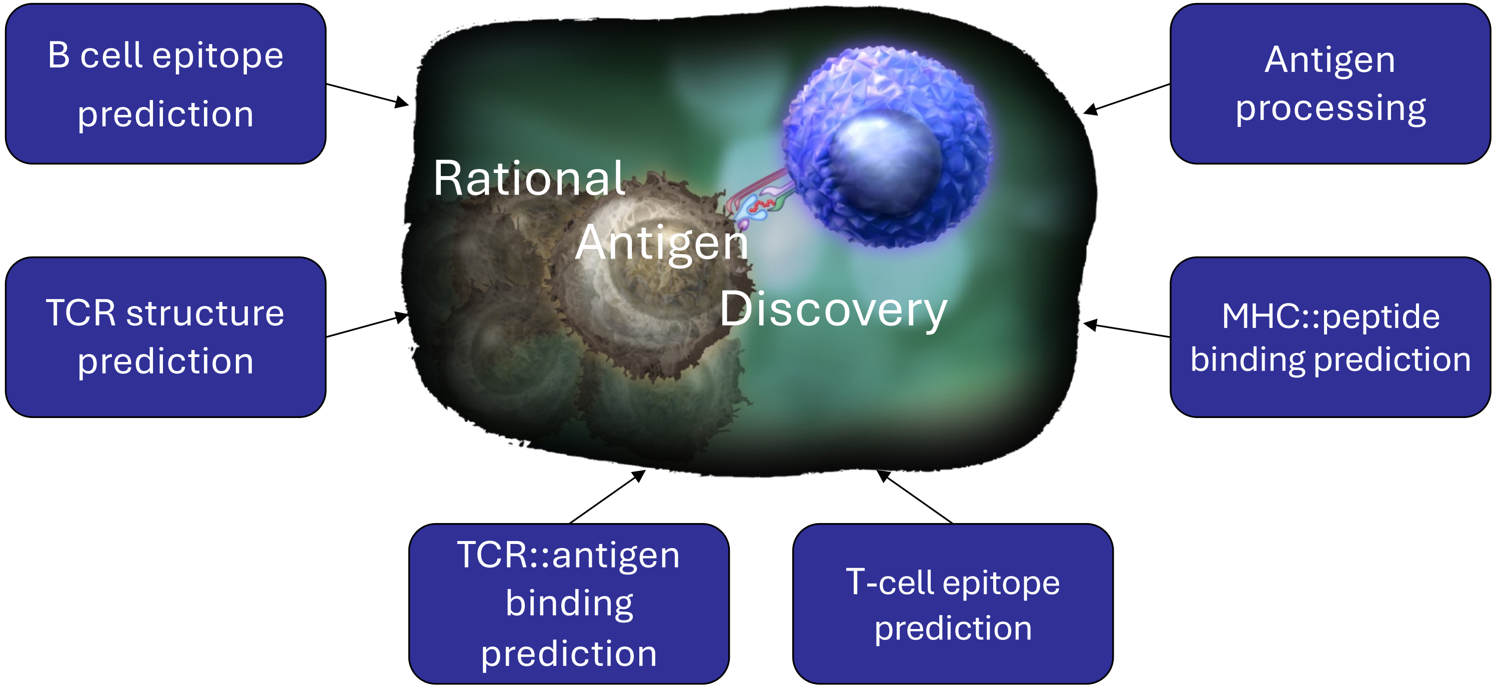
#	△	Team	Members	Score	Entries	Last	Solution
1	▲ 6	AF3TD_plddt_smallclust_final		0.6014	5	10mo	
2	▲ 16	John Altin		0.5819	2	10mo	
3	▲ 16	Lawson Woods		0.5815	1	10mo	
4	▲ 18	TCRtriFold		0.5778	8	10mo	
5	▲ 101	AF3_conf		0.5755	2	10mo	
6	▲ 32	RG		0.5703	5	10mo	
7	▲ 40	af3iptm clone		0.5686	5	10mo	
8	▲ 95	AF3_conf_clust		0.5685	2	10mo	
9	▲ 34	AF3iptm		0.5672	8	10mo	
10	▼ 9	AF3TD_iptm_bigclust_final		0.5670	4	10mo	

Far down on the list (sequence-based)

“Development and application of mathematical models
to characterize important players of the human immune system”



From sequence data to biomedical insights



B cell epitope prediction

TCR structure prediction

TCR::antigen binding prediction

T-cell epitope prediction

Antigen processing

MHC::peptide binding prediction

TOOLS:

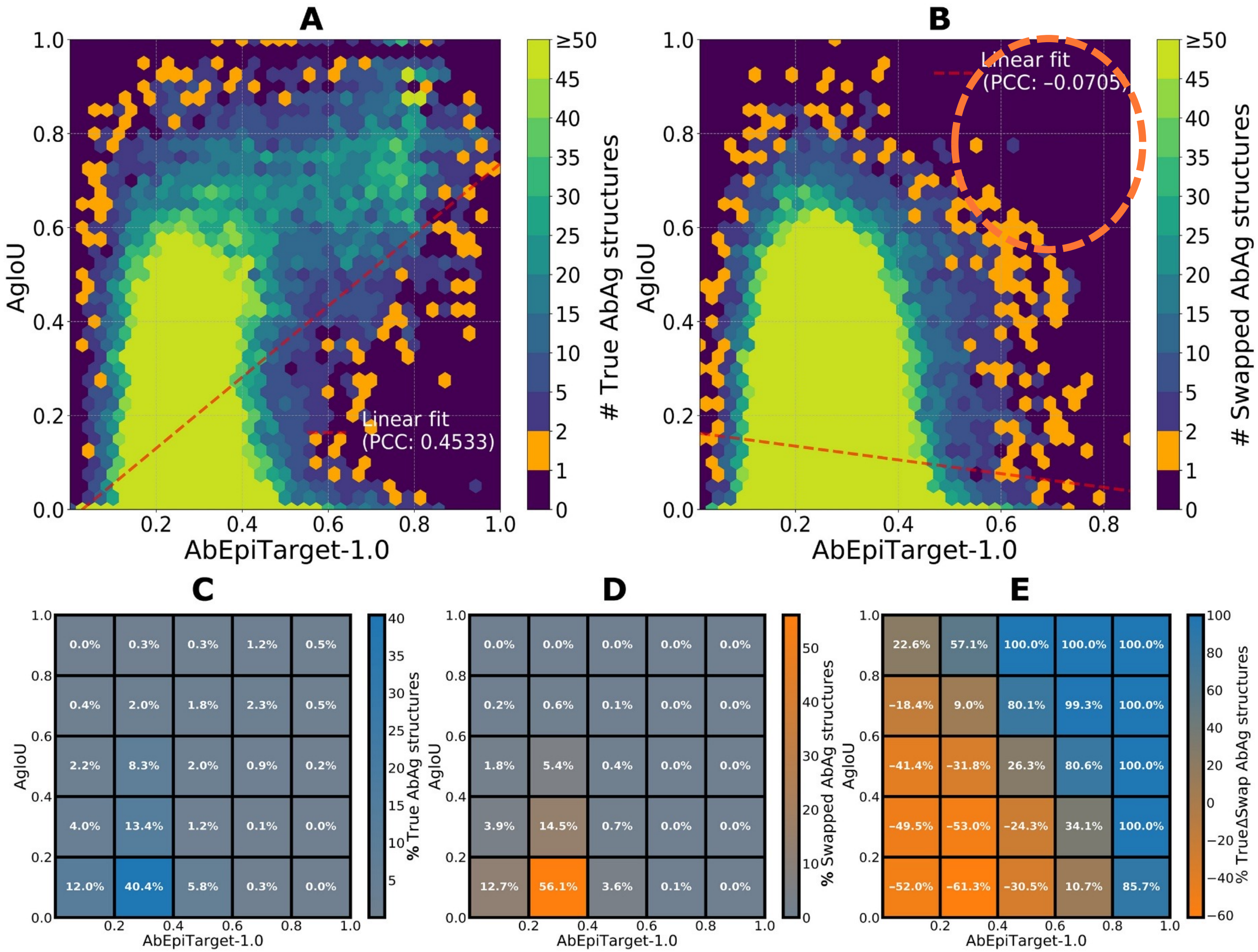
NetMHCpan
NetMHCIIpan
NetTCR
NetTCRstruct
NetChop
PopCover
GibbsCluster
Bepipred
Discotope
AbEpiTope
BepiPocket/
DiscoPocket

- **Plots (A-E):** Distribution of predicted antibody-antigen structures.

- **X-axis:** Confidence Score

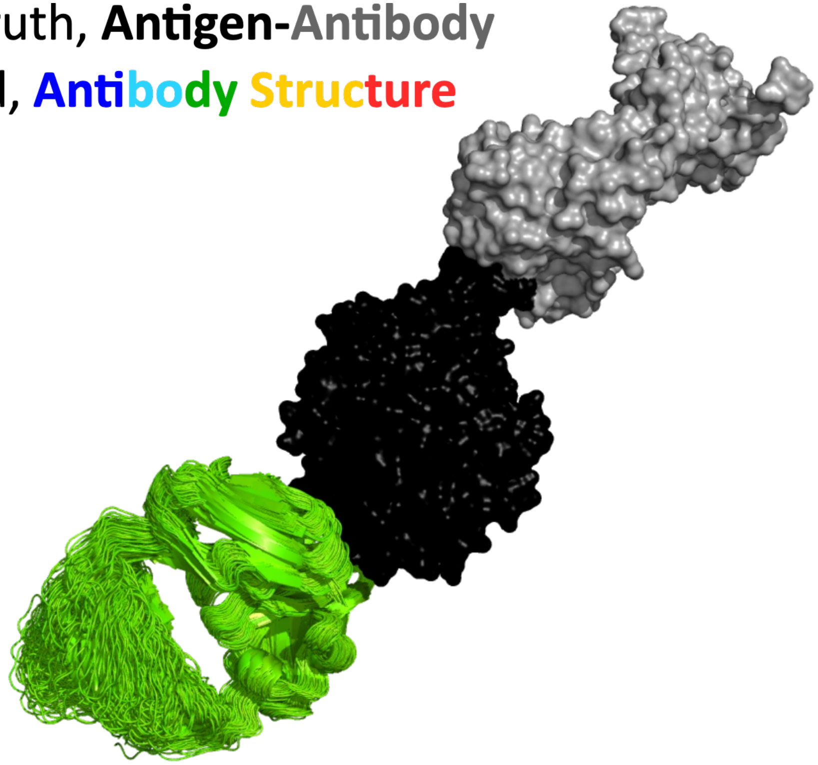
- **Y-axis:** AgIoU (epitope accuracy)

- **A) Correcly paired AbAgs can obtain high accuracy and score.**
- **B) Incorrecly paired AbAgs rarely receive high accuracy and score.**



***Clifford et al.**, Science Advances, June 2025*

Chai and AlphaFold can have very limited diversity in their predictions.



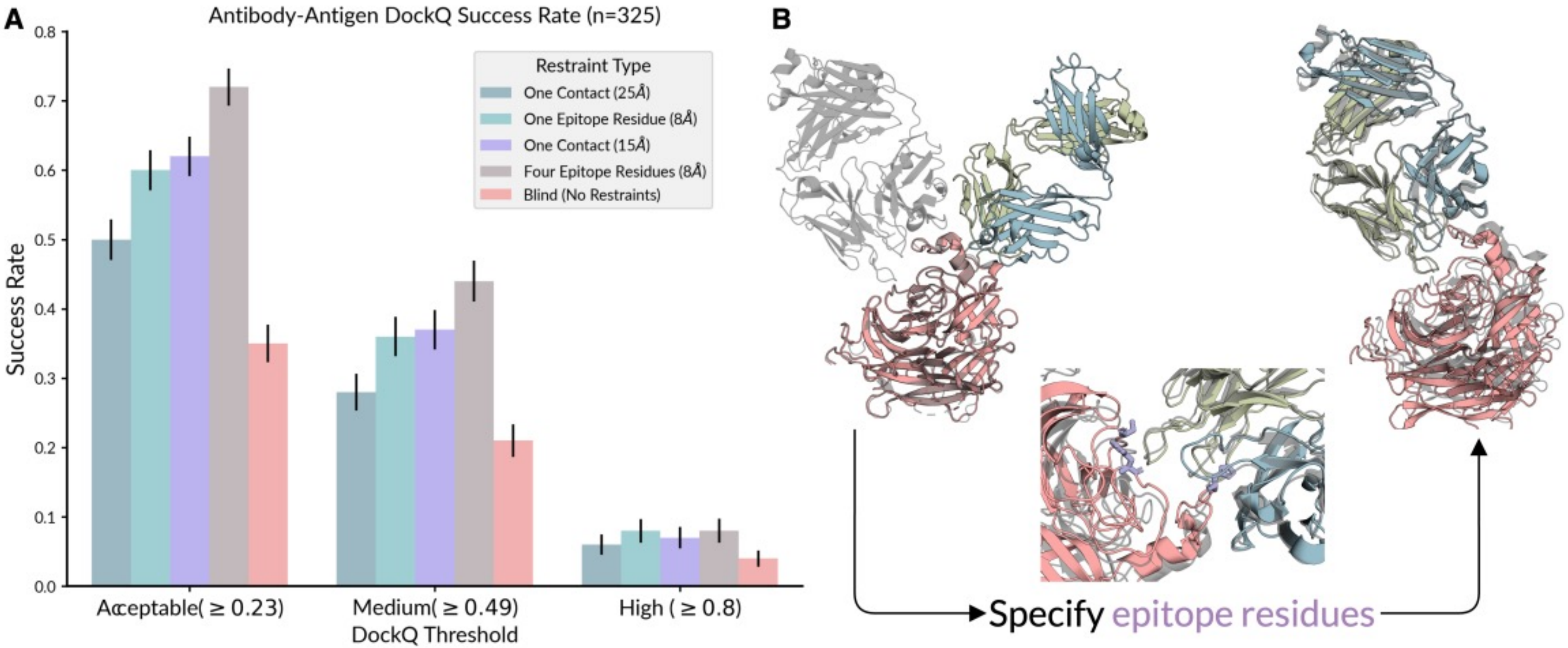
Groundtruth, **Antigen-Antibody**

Predicted, **Antibody Structure**

DockQ ≈ 0.0

Chai-1
Confidence Score: **0.684** - **0.699**

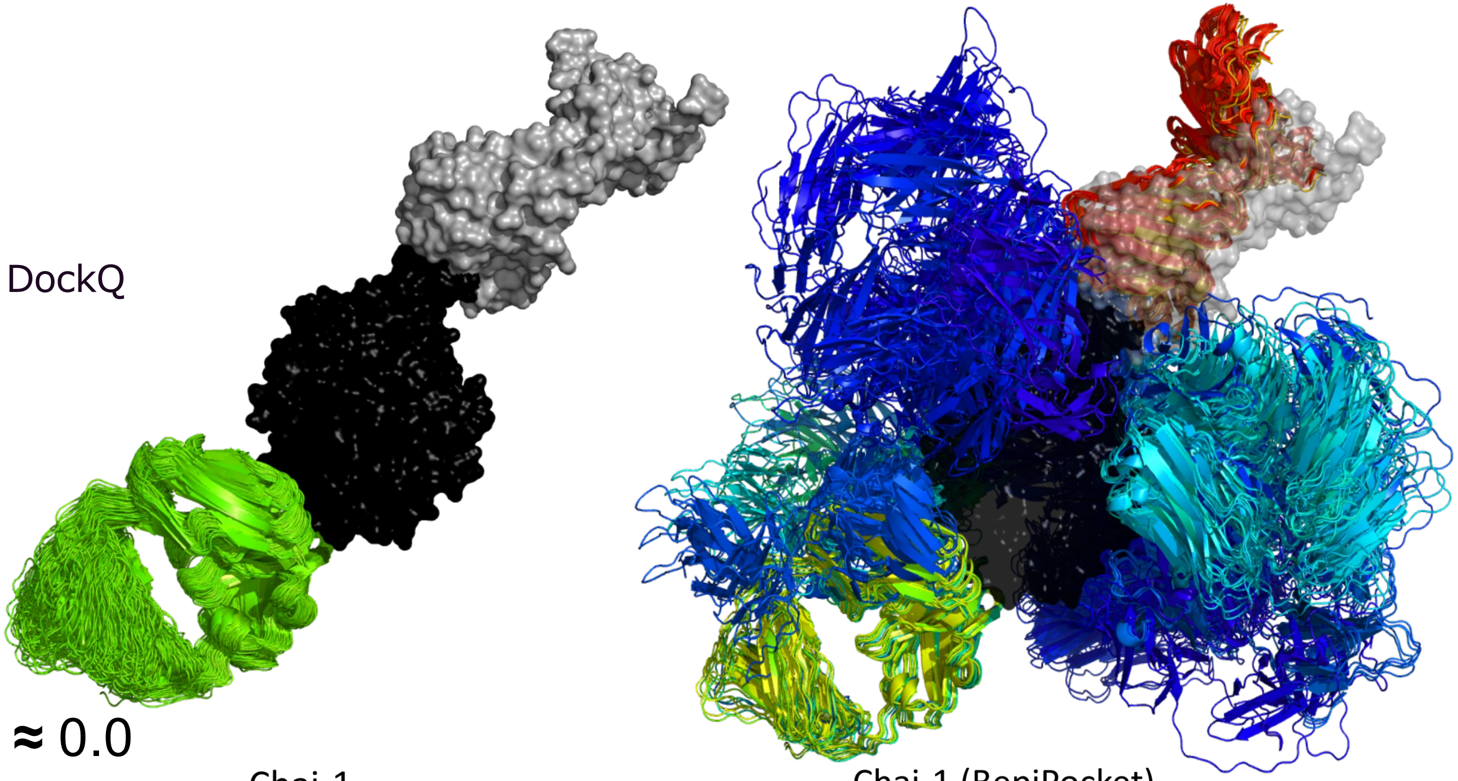
Chai-1 allows epitope restraints, improving accuracy



Translated from Chai Discovery et al., bioRxiv, October, 2024)

Antibodies-Antigen:

No DockQ → Acceptable/Medium/High DockQ



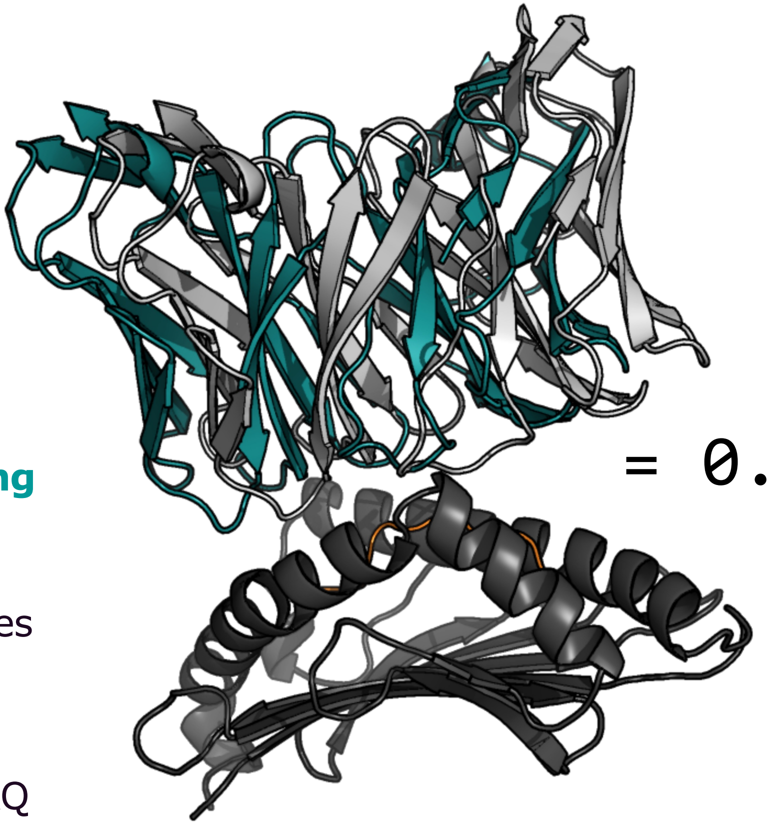
DockQ ≈ 0.0

Chai-1
Confidence Score: 0.684 - 0.699

DockQ ≈ 0.5

Chai-1 (BepiPocket)
Confidence Score: 0.492 - 0.808

Predicted TCR binding mode



= 0.376 DockQ

Crystallized TCRpeptideMHC

(PDB: 7N2R)

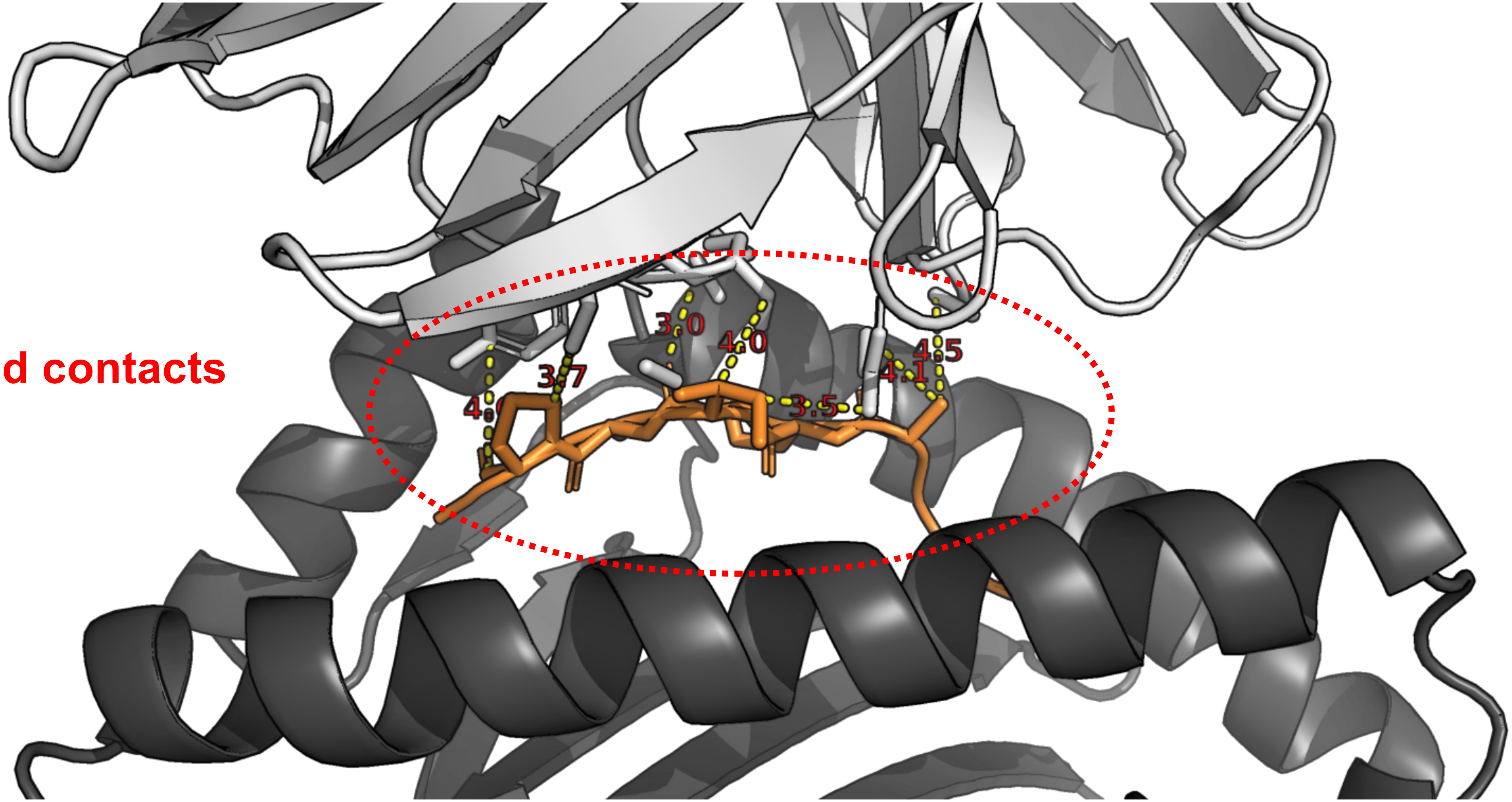
Accuracy for predicted TCRpMHC structures is generally a lot higher than AbAgs

For TCRpMHC, we aim to:

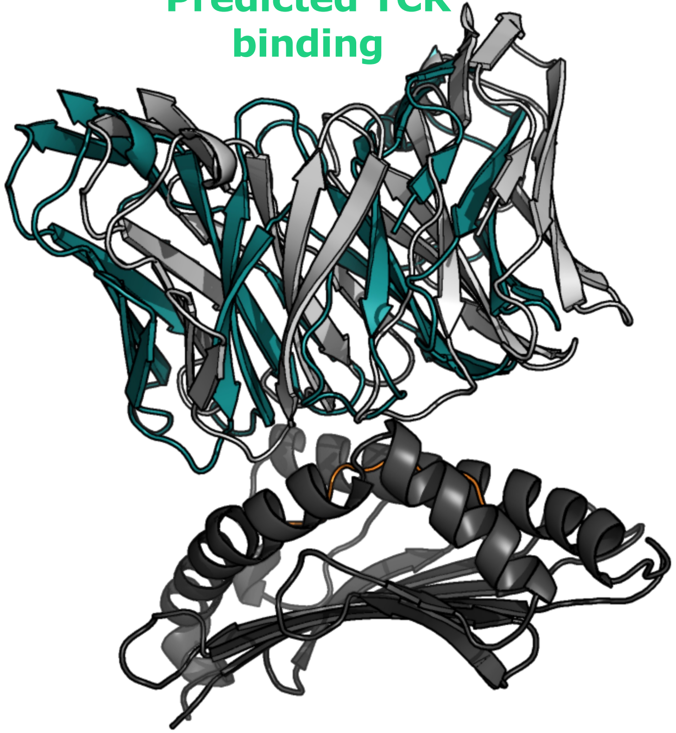
Acceptable/Medium → Medium/High DockQ

Hypothesis: Incorrect TCRs cannot achieve Medium/High accuracy

Can we cheat our way to **medium/high DockQ** accuracy?



Chai-1



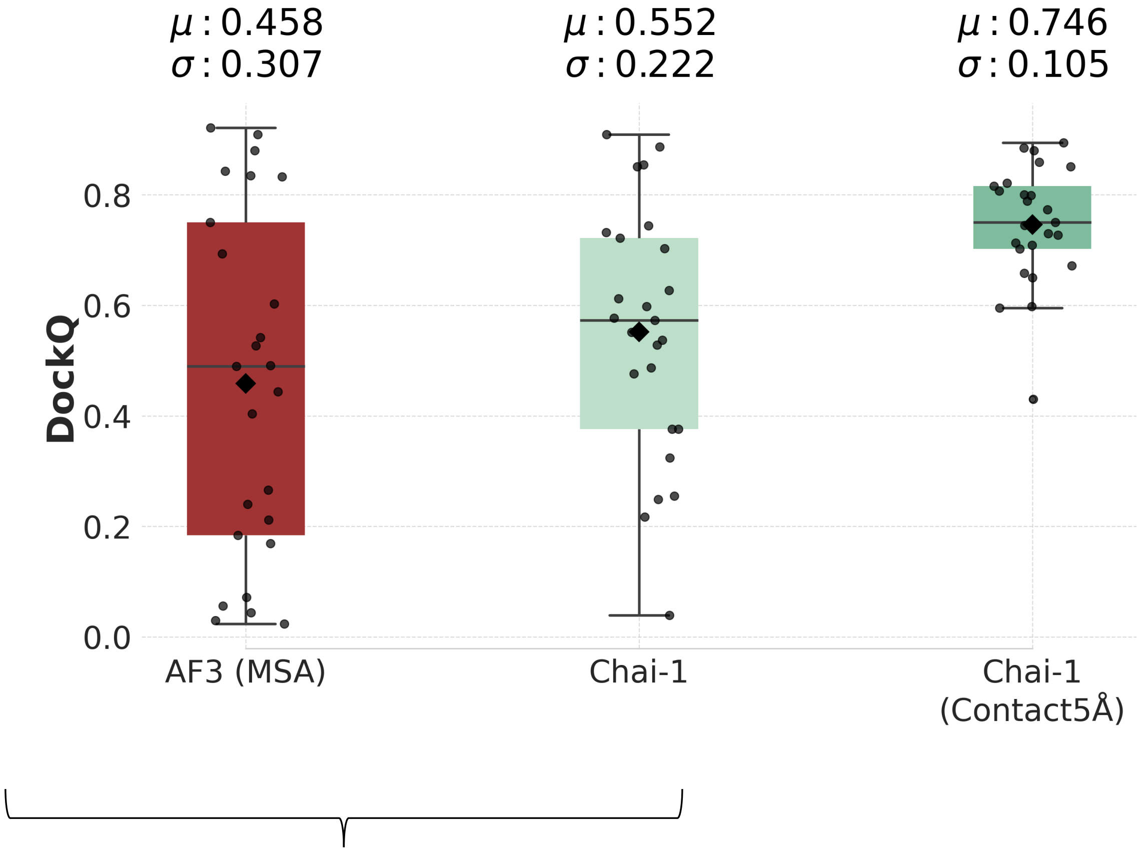
0.376 DockQ

Chai-1 with restraints

0.702 DockQ

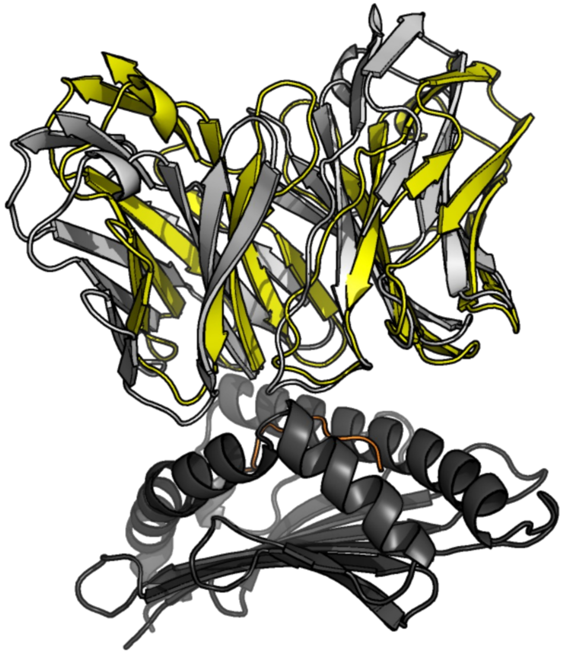
Solved contacts majorly improve structural accuracy

- Each dot is TCRpMHC structure.
- 25 TCRpMHCI structures
- **Chai-1 Contact5Å** Solved TCRpMHC contacts < 5Å



Without restraint methods

Chai-1
PepContact(5Å)

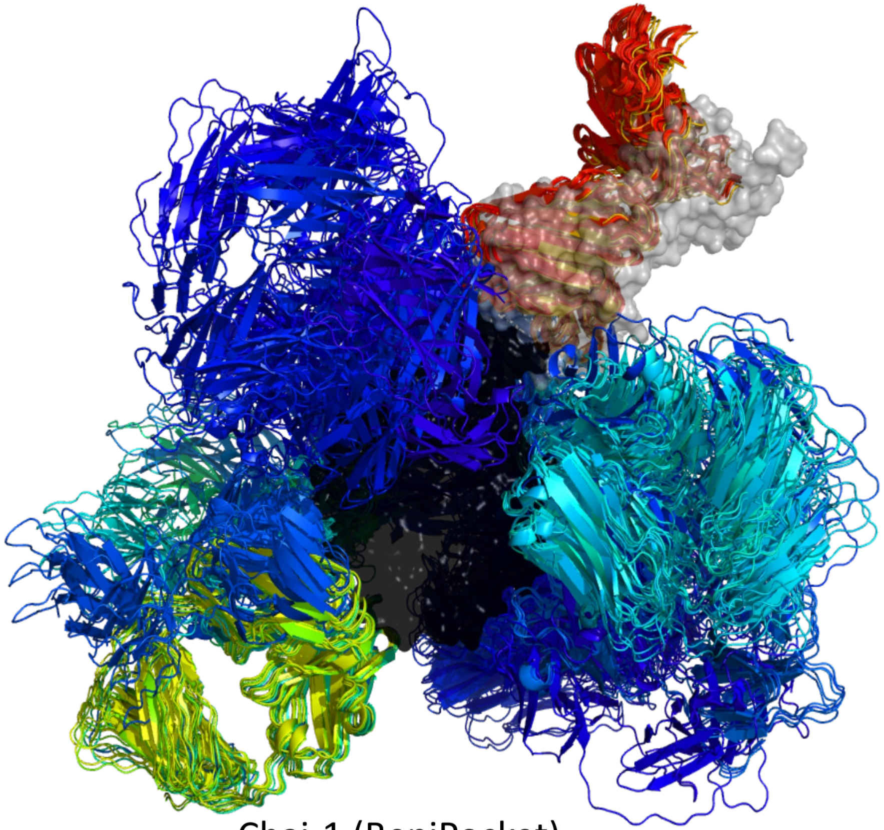


0.565 DockQ

We want to use a simple approach that worked well antibody-antigens.

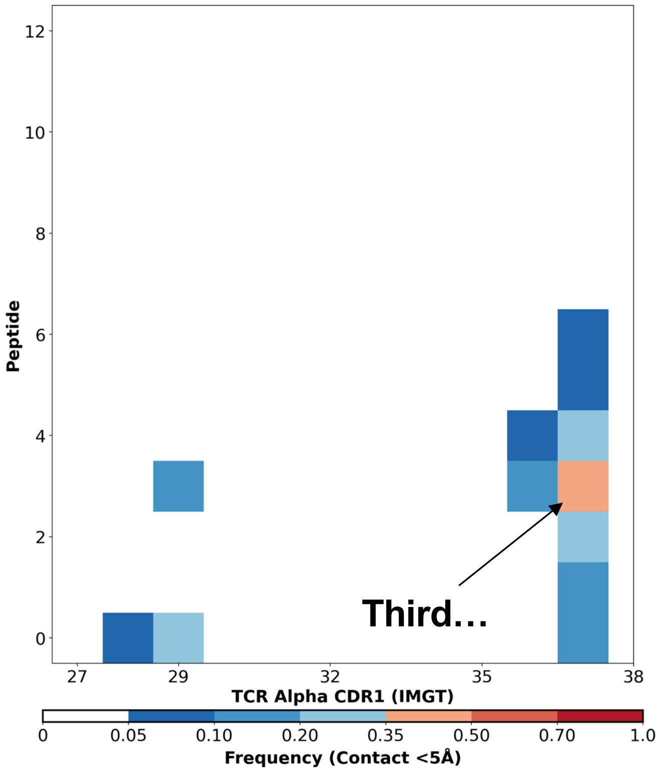
For **BepiPocket,** we specified only a single epitope-antibody restraint at a time.

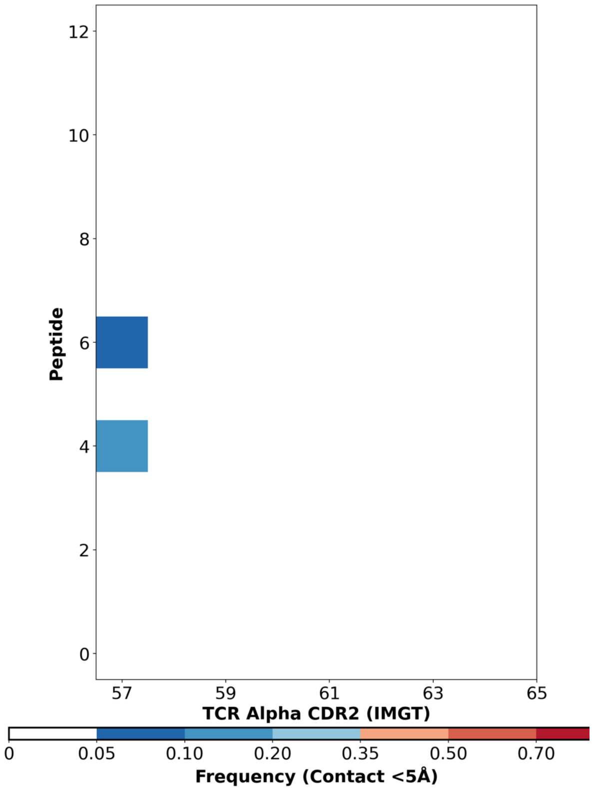
We will specify a single TCR-peptide restraint at a time

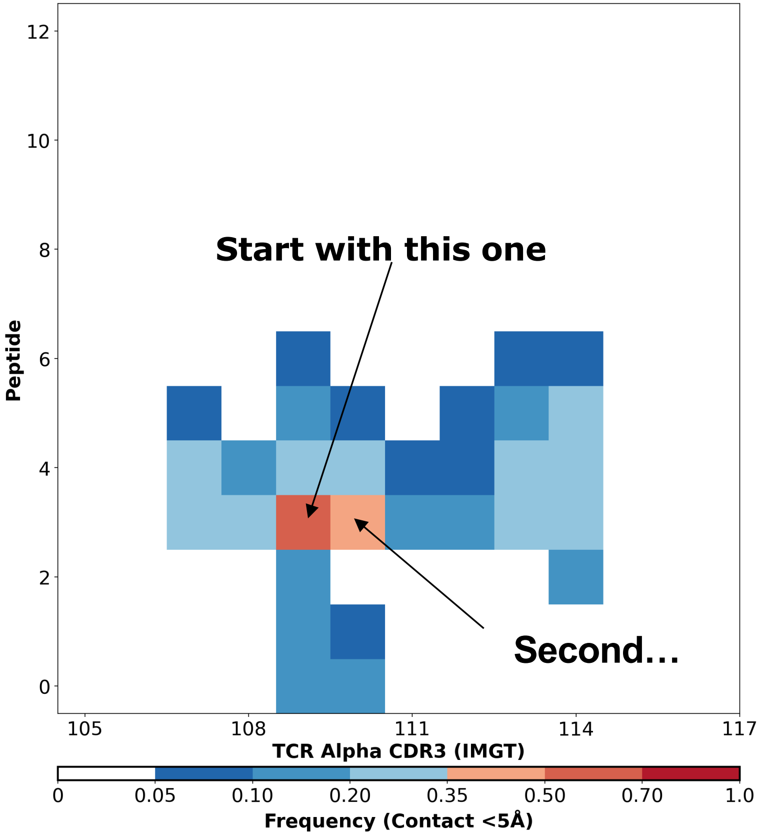


Chai-1 (BepiPocket)
Confidence Score: **0.492** - **0.808**

TCRPepCon-α

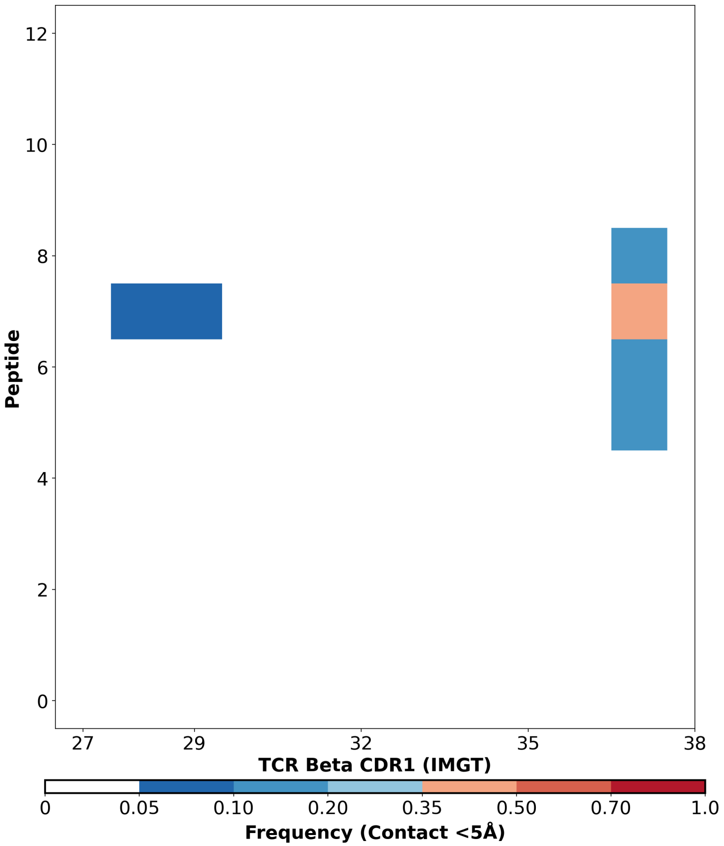

Third…



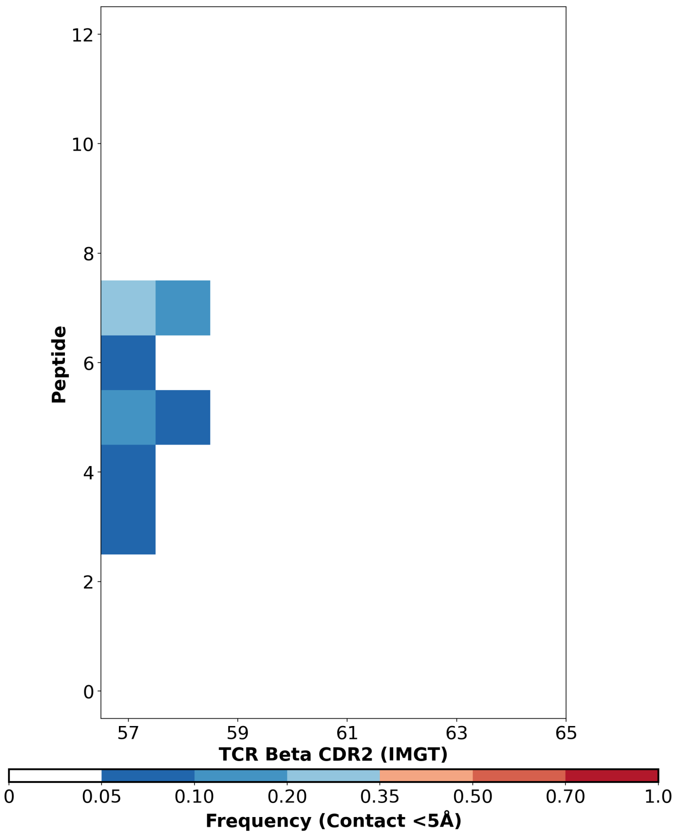

Start with this one

Second…

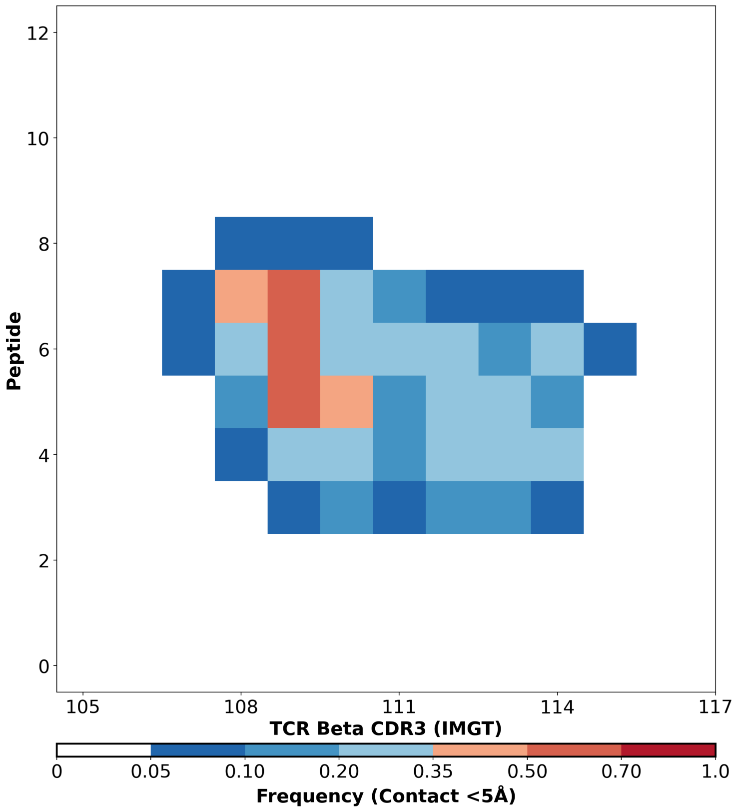
TCRPepCon-β



12
10
8
6
4
2
0
Peptide
27
29
32
35
38
TCR Beta CDR1 (IMGT)
0
0.05
0.10
0.20
0.35
0.50
0.70
1.0
Frequency (Contact <5Å)

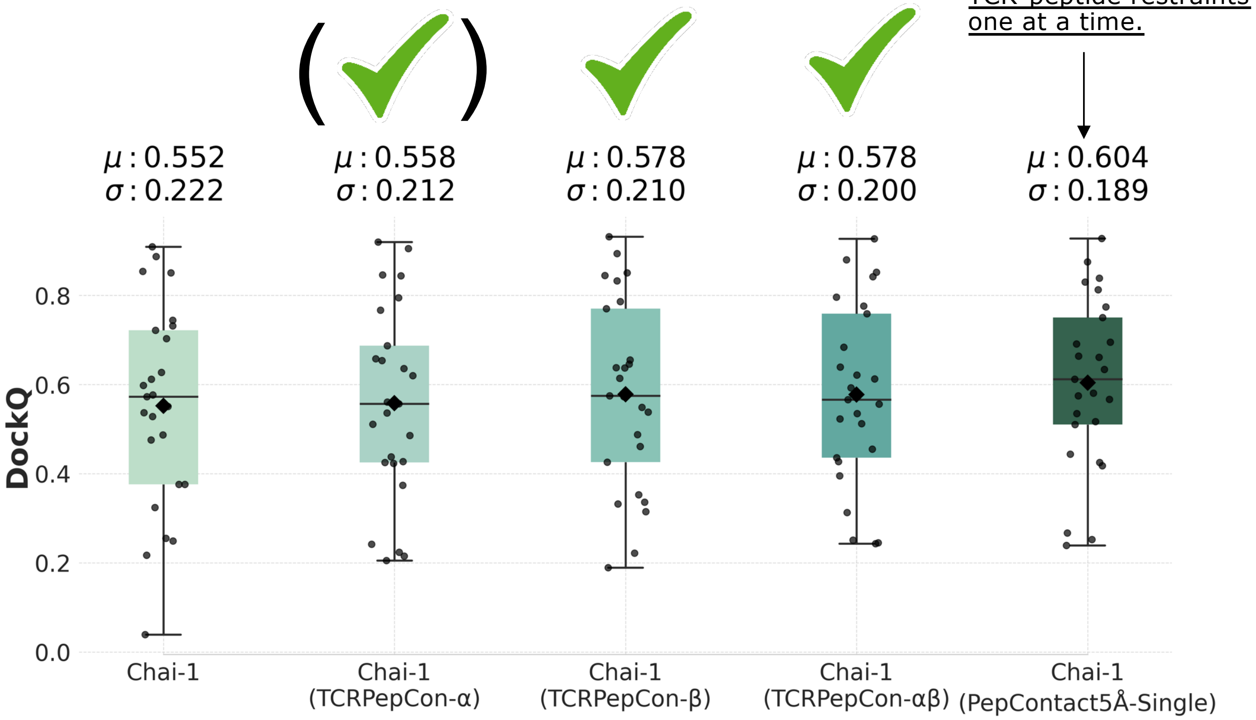


12
10
8
6
4
2
0
Peptide
57
59
61
63
65
TCR Beta CDR2 (IMGT)
0
0.05
0.10
0.20
0.35
0.50
0.70
1.0
Frequency (Contact <5Å)



12
10
8
6
4
2
0
Peptide
105
108
111
114
117
TCR Beta CDR3 (IMGT)
0
0.05
0.10
0.20
0.35
0.50
0.70
1.0
Frequency (Contact <5Å)

- Small increase **TCRPepCon-α**
- Increase for **TCRPepCon-β**
- Increase **TCRPepCon-αβ**
- Some increase in predicted structural accuracy



Upperbound (cheat method):
Running solved TCR-peptide restraints one at a time.

Conclusion and Perspectives

- TCRpMHC structure prediction guided with restraint features does improve predicted structural accuracy.
- But we still have not figured out to obtain the performance gains, we get for predicting antibody-antigen structures, for predicting TCRpMHC structure prediction.
- Canonical set of TCRpMHC restraint features.

Acknowledgements

Immunoinformatics and Machine Learning group:

Professor Morten Nielsen

Sebastian Deleuran

Jonas Birkelund

Yat-Tsai Richie Wan

Eric Bautista

Martiela Vaz de Freitas

Siqi Zhang

IEDB

Bjoern Peters

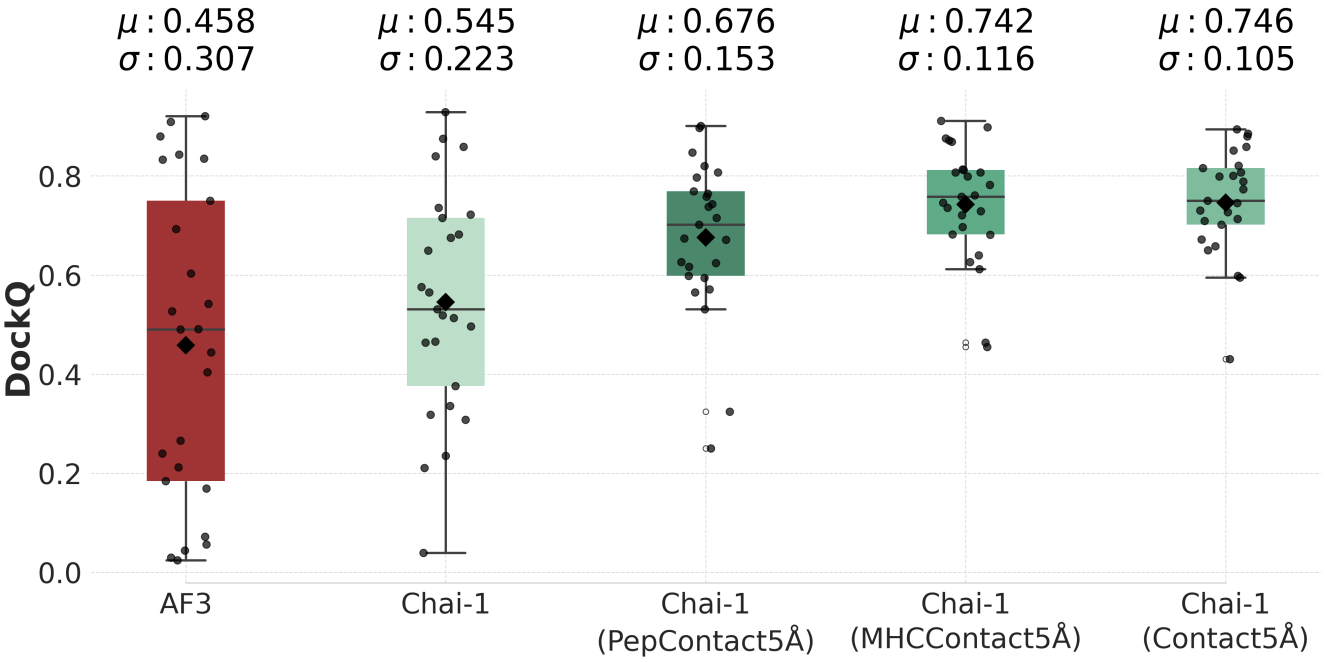
Eve Richardson

Lonneke Scheffer

Thank you for listening!

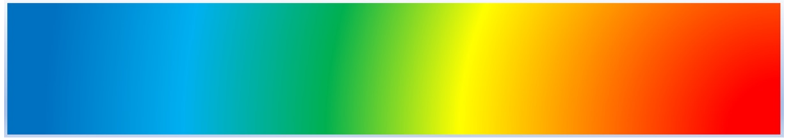
Peptide, MHC, and PeptideMHC contacts

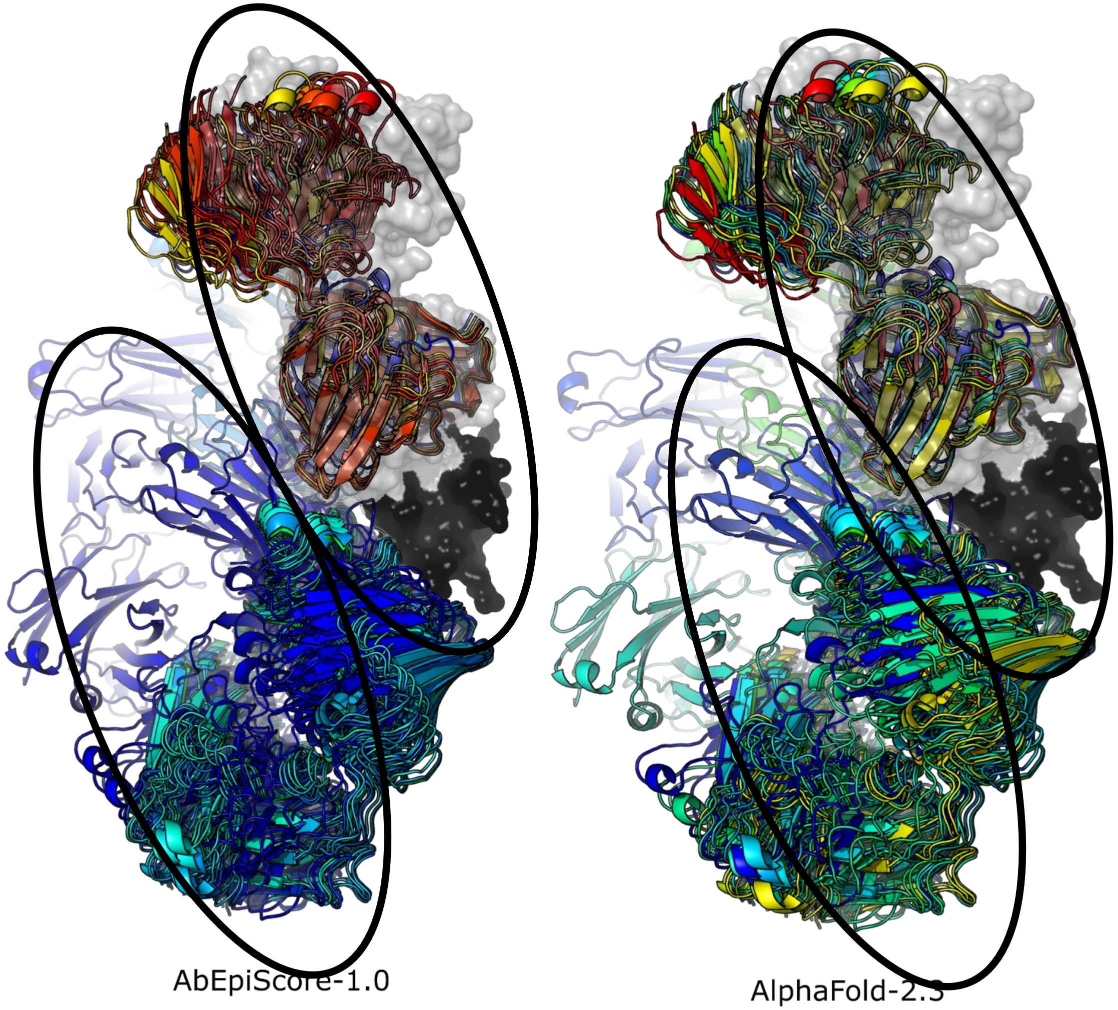
- **Chai-1 PepContact5Å:** Solved TCRp contacts < 5Å
- **Chai-1 MHCContact5Å:** Solved TCRMHC contacts < 5Å
- **Chai-1 Contact5Å:** Solved TCRpMHC contacts < 5Å



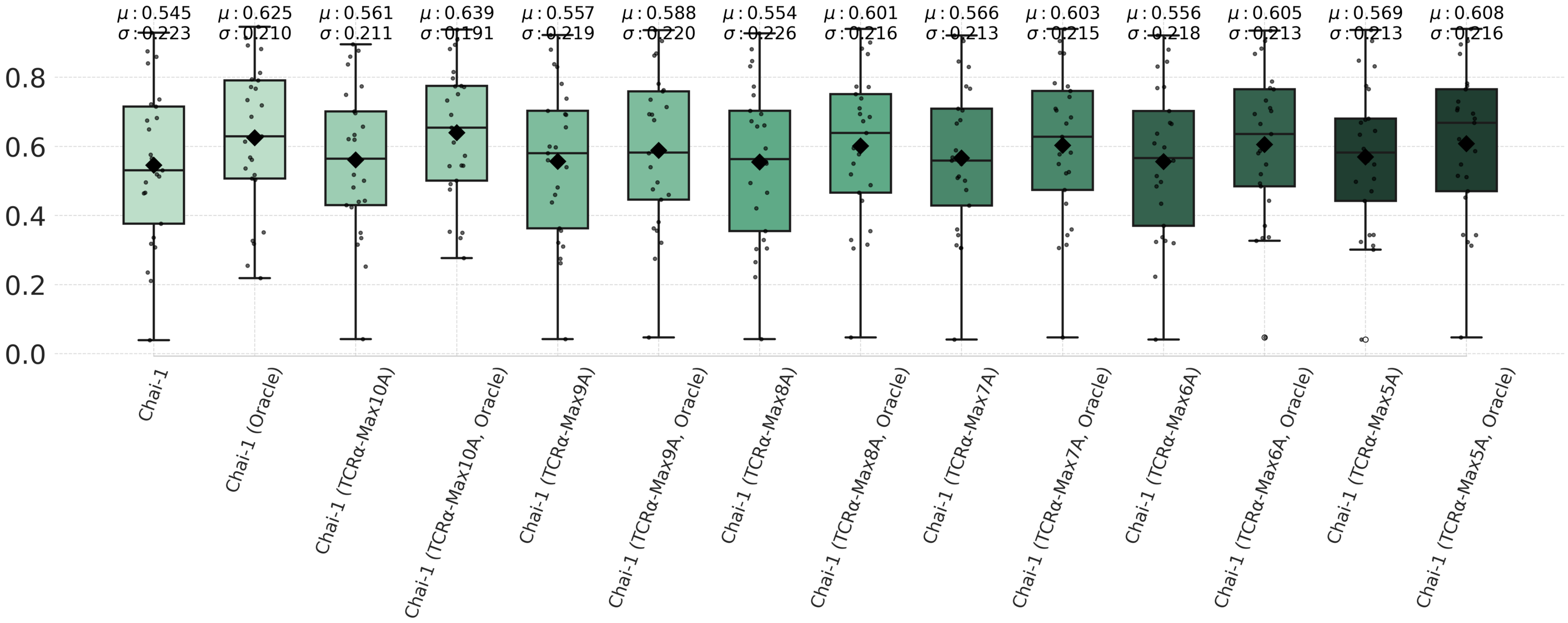
AbEpiTope: Predicting Antibody /BCR specificity

- Crystallize antibody (grey) to insulin growth factor II (**black**)
- Colored predicted antibody docking poses (blue, green, yellow, red)
- Score: AbEpiScore-1.0 and AlphaFold-2.3.



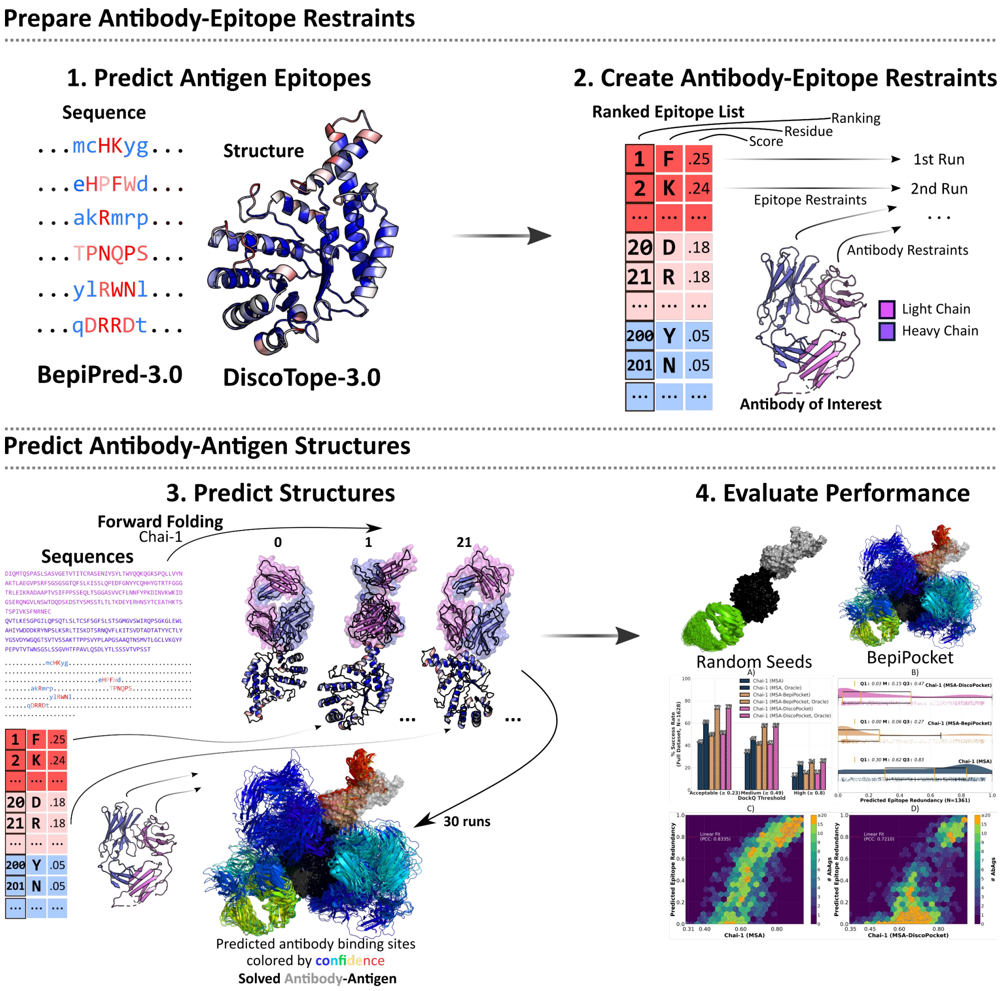


***Clifford et al.**, Science Advances, June 2025*


μ : 0.545
σ : 0.223
μ : 0.625
σ : 0.210
μ : 0.561
σ : 0.211
μ : 0.639
σ : 0.191
μ : 0.557
σ : 0.219
μ : 0.588
σ : 0.220
μ : 0.554
σ : 0.226
μ : 0.601
σ : 0.216
μ : 0.566
σ : 0.213
μ : 0.603
σ : 0.215
μ : 0.556
σ : 0.218
μ : 0.605
σ : 0.213
μ : 0.569
σ : 0.213
μ : 0.608
σ : 0.216
0.8
0.6
0.4
0.2
0.0
Chai-1
Chai-1 (Oracle)
Chai-1 (TCRα-Max10A)
Chai-1 (TCRα-Max10A, Oracle)
Chai-1 (TCRα-Max9A)
Chai-1 (TCRα-Max9A, Oracle)
Chai-1 (TCRα-Max8A)
Chai-1 (TCRα-Max8A, Oracle)
Chai-1 (TCRα-Max7A)
Chai-1 (TCRα-Max7A, Oracle)
Chai-1 (TCRα-Max6A)
Chai-1 (TCRα-Max6A, Oracle)
Chai-1 (TCRα-Max5A)
Chai-1 (TCRα-Max5A, Oracle)

- Predict restraint features with B cell epitope prediction tools (BepiPred-3.0 and Discotope-3.0)

- Use these features to guide structure prediction
- Improves structural accuracy
- Vastly improves diversity antibody-antigen predicted structures.



Joakim Clifford and Morten Nielsen et al., *biorXiv preprint, September 2025*

Immunoinformatics and Machine Learning Group


Morten Nielsen, Professor,

Group Leader


Eric, PhD


Ziki, PhD

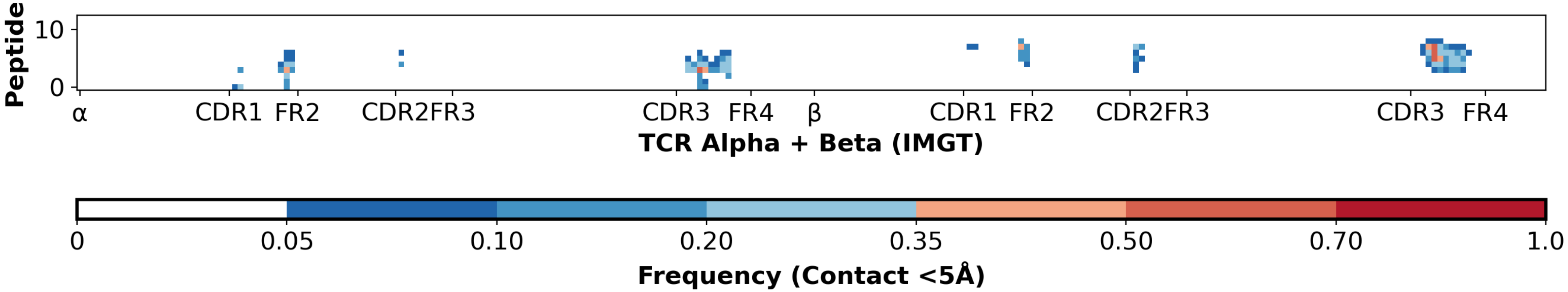

Sebastian, PhD


Pilar, PhD

Joakim, Postdoc

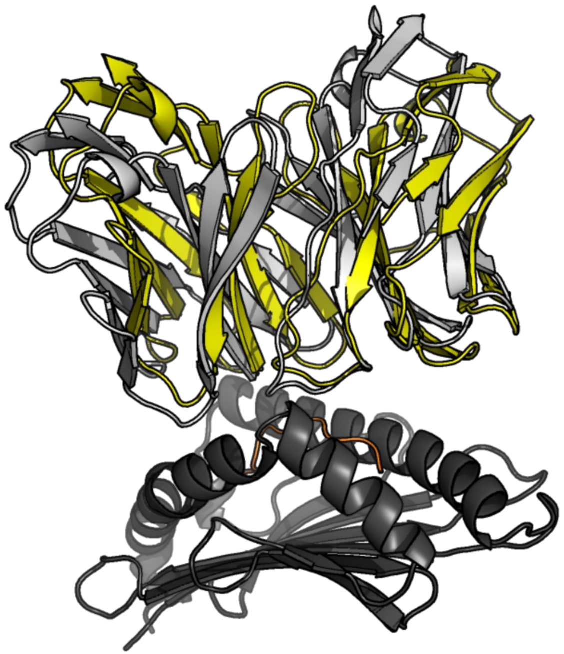
Martiela, Postdoc

TCRPepCon-α, TCRPepCon-β and TCRPepCon-αβ methods



1. Get frequency of contacts from 272 solved TCRpMHCI complexes.
2. Use these frequencies to create restraints for TCRpMHC structure prediction.
3. **TCRPepCon-α:** TCRAlpha-Peptide Restraints
4. **TCRPepCon-β:** TCRBeta-Peptide Restraints
5. **TCRPepCon-αβ:** TCRAlphaBeta-Peptide Restraints

Chai-1 PepContact(5Å)



0.565 DockQ

≈ 30% restraints

Chai-1 MHCContact(5Å)

0.697 DockQ

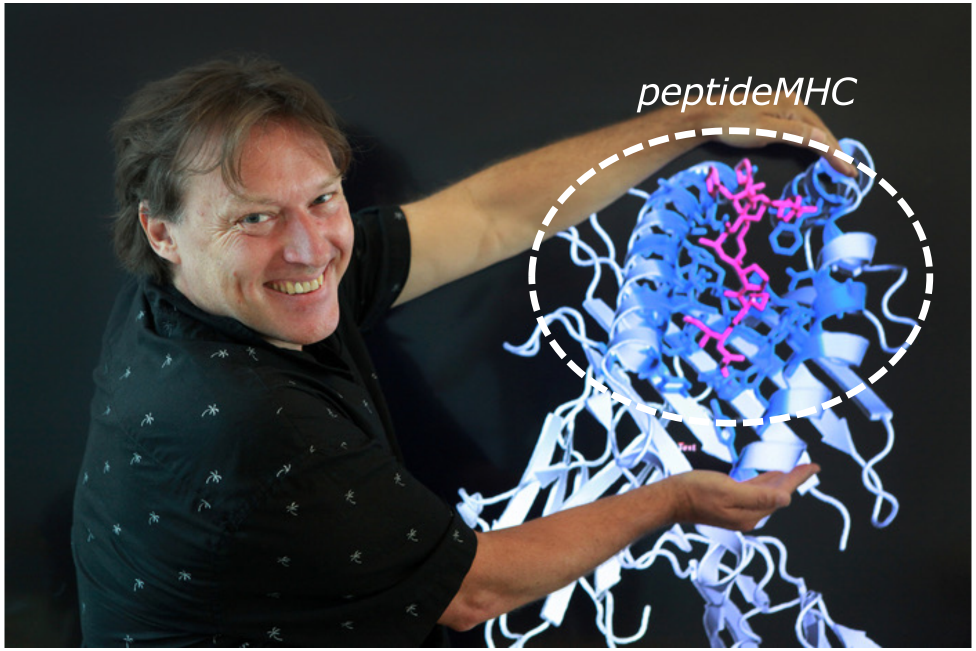
≈ 70% restraints

Chai-1 Contact (5Å)

0.702 DockQ

Immunoinformatics and Machine-Learning Group at DTU, Denmark

- *Understanding how T and B cells interact with their target molecules.*
- **Target molecules:** Antigens
- **Epitope:** Antigen interaction site.
- **T cell epitopes** are peptideMHCs.
- Our group creates mathematical models to predict B and T cell epitopes.
- Predicting epitopes helps us develop immunotherapies.



Morten Nielsen PhD, Professor

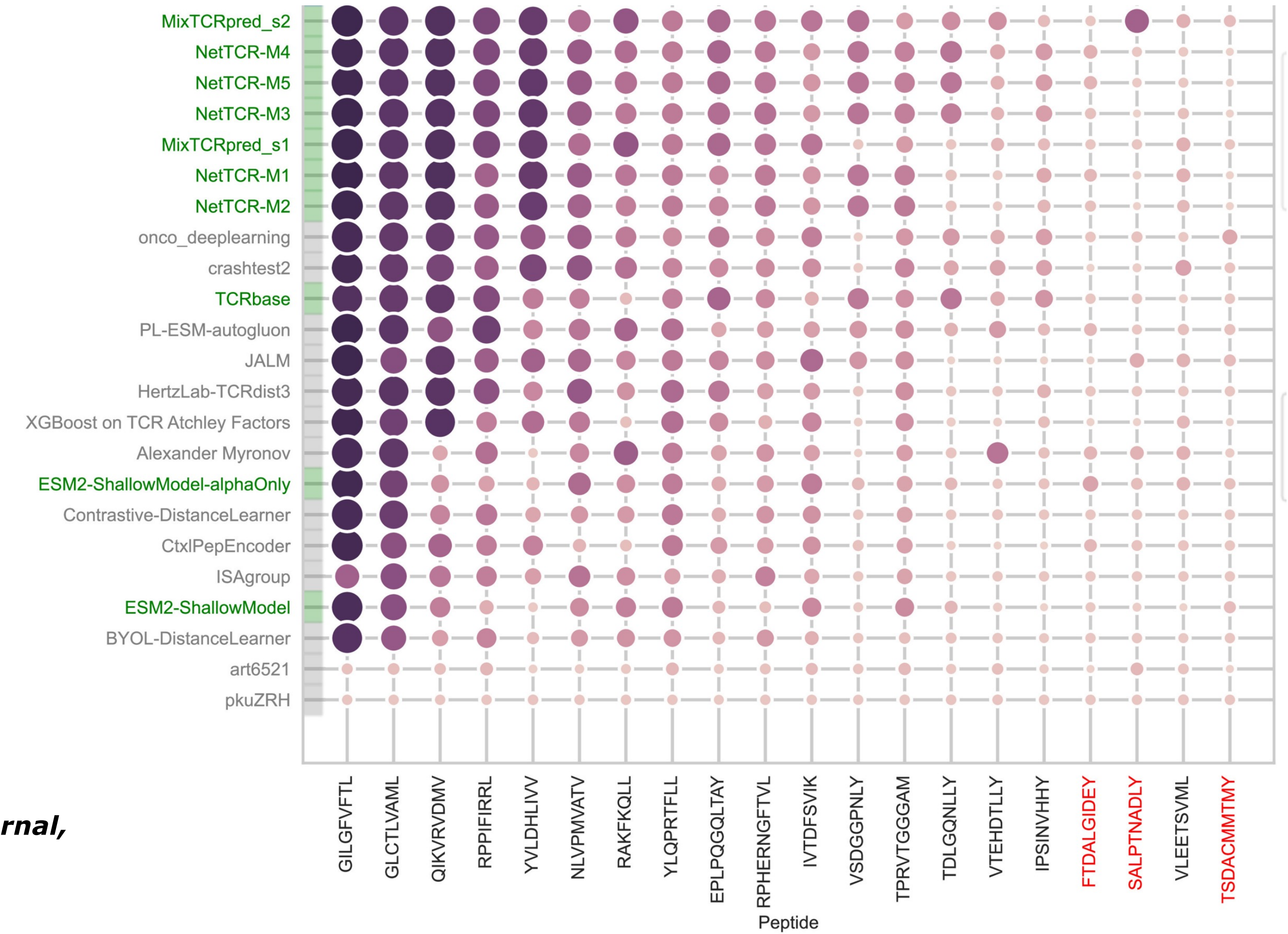
IMMREP23

Seen peptides is more or less a solved case.

Unseen peptides (peptides we have no data for), is very difficult.

Nielsen, Morten et al., Immunoinformatics Journal, 2026, *Lessons learned from the IMMREP23 TCR-epitope prediction challenge*

Bigger dots = Better performance



Seen peptides **Unseen peptides**