



Solving 3D puzzles of biomolecular interactions by physics- and AI-based integrative modelling

Prof. Alexandre M.J.J. Bonvin
Bijvoet Centre for Biomolecular Research
Faculty of Science, Utrecht University
the Netherlands
a.m.j.j.bonvin@uu.nl
 @amjjbonvin.bsky.social

International Symposium on Grid and Clouds
Academia Sinica, Taipei, Taiwan
March 16-20, 2026





NMR: 1.2 GHz, 900, 750, 600, 600US, 2x500 MHz
800WB-DNP, 400WB-DNP, 700US, 500WB MHz

e-infrastructure: >1800 CPU cores + GPUs + EGI HTC resources (>100'000 CPU cores)



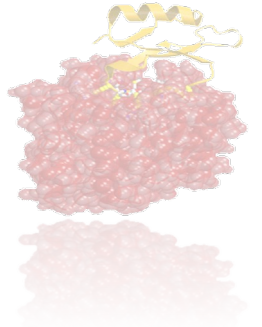
National



and European infrastructure



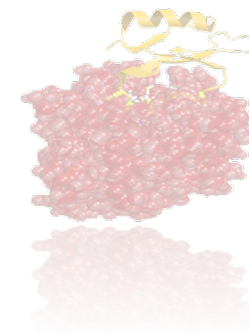
The social network of proteins



Majority of 'life' depends on interactions, particularly protein-protein

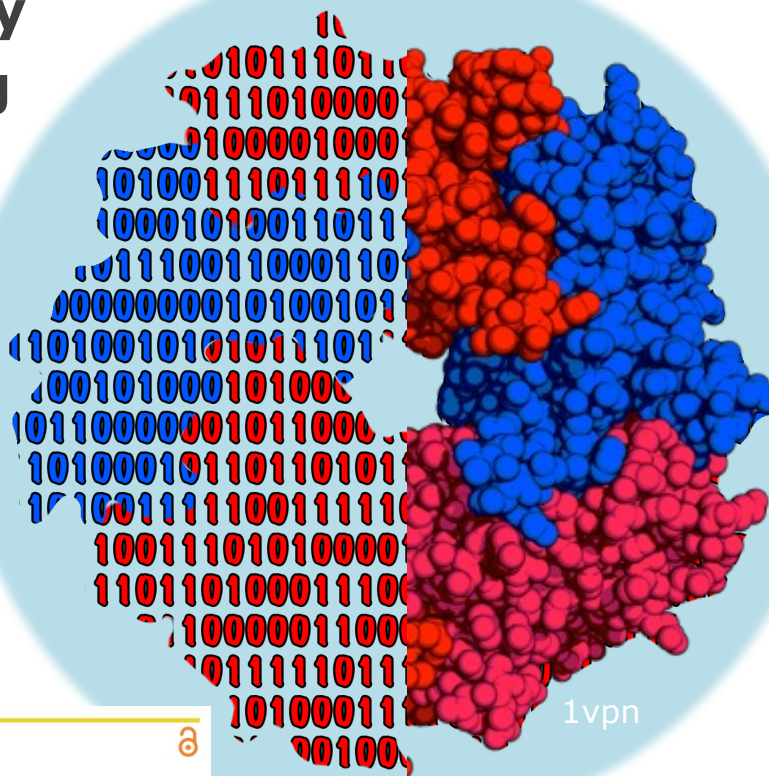


Structural biology of interactions



Docking
Molecular Dynamics
Homology Modeling
Threading

MS
SAS
X-Ray
EPR
FRET
NMR
Cryo-EM



Experiment

1vpn

-resolution experiments
by the experimental community

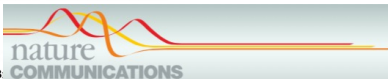
bioRxiv preprint doi: <https://doi.org/10.1101/2021.10.04.463034>; this version posted October 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



2021-10-04

Protein complex prediction with AlphaFold-Multimer

Richard Evans^{1*},
Augustin Židek¹,
Zielinski¹, Alex B.
Clancy¹, Pushme¹
¹DeepMind, London,



ARTICLE

<https://doi.org/10.1038/s41467-022-28865-w>

OP nature biotechnology

Improved prediction using AlphaFold2

Patrick Bryant^{1,2,3✉}, Gabriele Pozzati^{1,2}

Article

Protein structure prediction with AlphaFold2

Received: 26 September 2022

Accepted: 6 February 2023

Kolja Stahl[✉]

& Juli Rapps

Accurate structure prediction of biomolecular interactions with AlphaFold 3

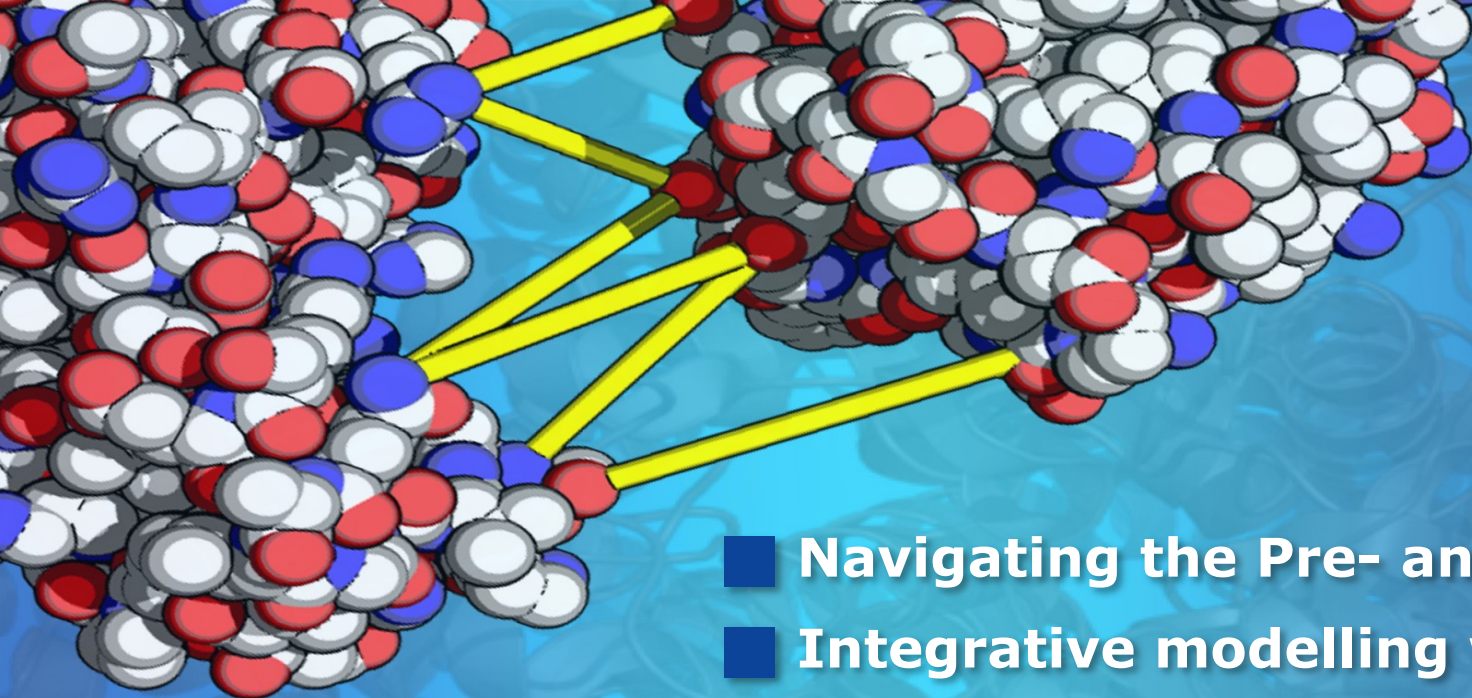
Josh Abramson, Jonas Adler, Jack Dunger, Richard Evans, Tim Green, Alexander Pritzel, Olaf Ronneberger, Lindsay Willmore, Andrew J. Ballard, Joshua Bambrick, Sebastian W. Bodenstein, David A. Evans, Chia-Chun Hung, Michael O'Neill, David Reiman, Kathryn Tunyasuvunakool, Zachary Wu, Akvilė Žemgulytė, Eirini Arvaniti, Charles Beattie, Ottavia Bertolli, Alex Bridgland, Alexey Cherepanov, Miles Congreve, ... John M. Jumper ✉ + Show authors

Nature (2024) | Cite this article



Universiteit Utrecht

[Faculty of Science
Chemistry]



Overview

- Navigating the Pre- and Post-AlphaFold Divide
- Integrative modelling with HADDOCK
- Introducing HADDOCK3
- Application examples
 - An NMR-based application – where AlphaFold fails
 - AI-based antibody-antigen modelling
 - Cyclic peptides
- Conclusions & perspectives

```
111000100101010110101010
100011010101000101111010110101
10100001000111101010101110000110101
0001010101011110010100110010101010100
0101011111110101010000101010100000101010100
001011010101110101000111101010101000101001010101111
0001111111010101011001100010111000101011010100010101000101
```

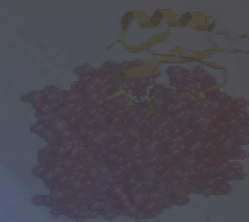
CRITICAL ASSESSMENT THE RISE OF AI / ALPHAFOLD



Marc Lensink



Ezgi Karaca



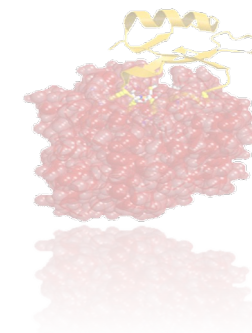
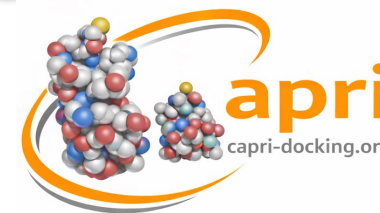
Critical Assessment

of Structure Predictions

- **CASP**
- Since 1994
- Prediction “season”
 - Intense 2 to 3 months
- Protein monomers
- (traditionally)

of PRedicted Interactions

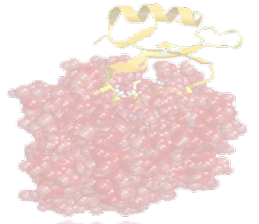
- **CAPRI**
- Since 2001
- Predictions on a “rolling” basis
 - Fits with publication schedule
 - 3 to 4 weeks per prediction round
- Protein complexes



Since 2014 joint prediction rounds for
complexes/assemblies



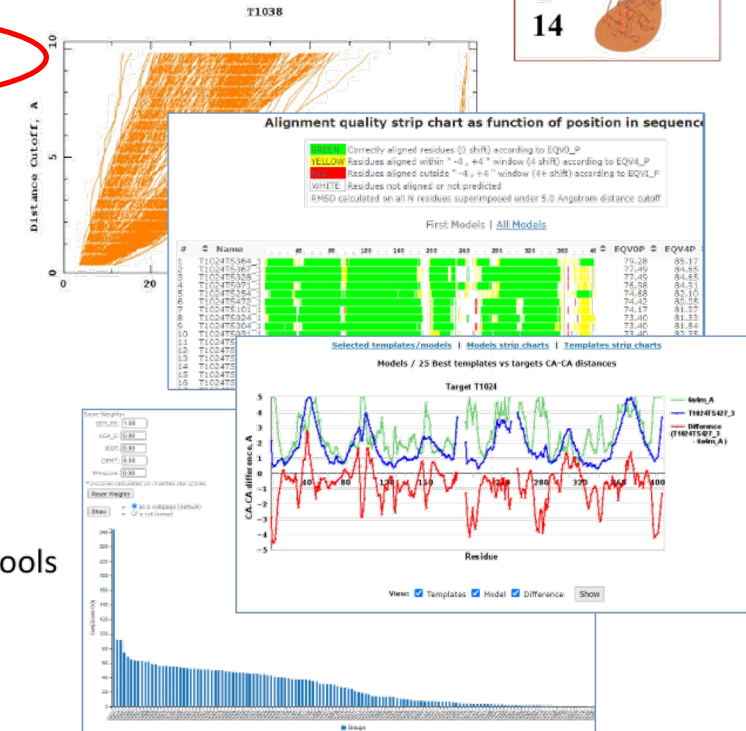
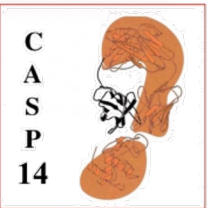
Critical Assessment is a Big Deal



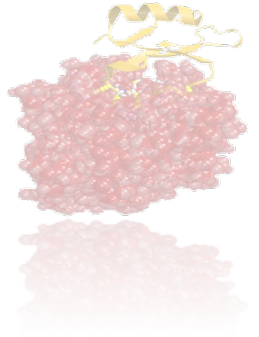
- CASP has been around for 30+ years
- Reference for structure prediction
- CAPRI - started 20+ years ago and specializes in the prediction of protein complexes
- Blind prediction experiments catalyze method development
- Critical Assessment defines the state-of-the-art in the field

CASP14 data at a glance

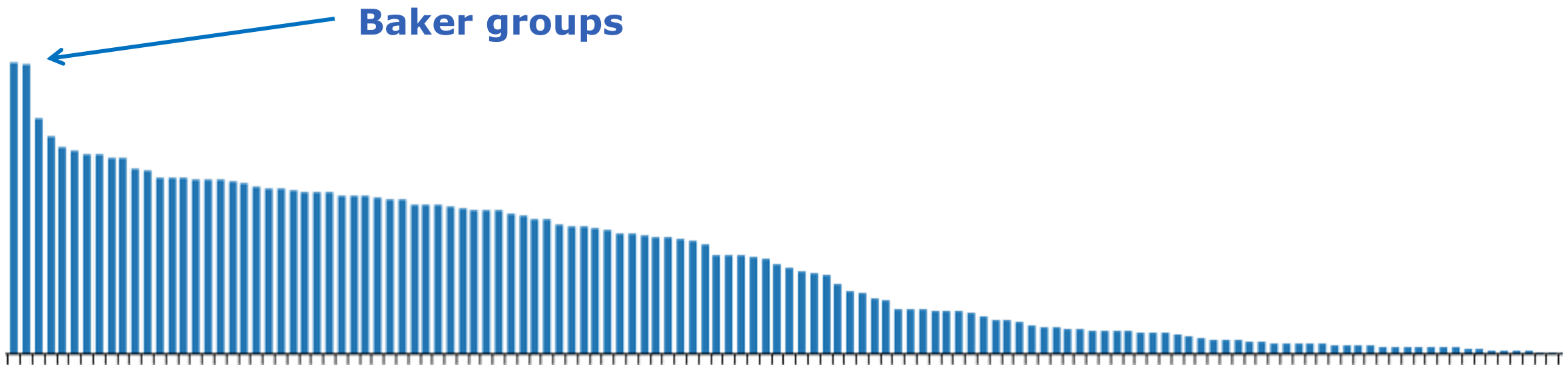
- >200 prediction methods
- ~100 research centers
- ~350 predictors
- >80 targets
- >67,000 models
- >5,000,000 scores
- ~430 GB of data
- >30 different software tools
- >20 visualization tools



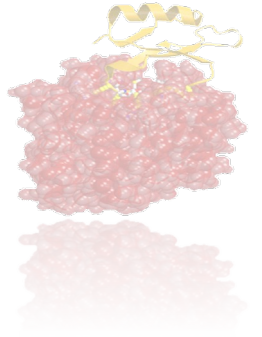
CASP14 - 2020



“Regular” targets (meaning single proteins)



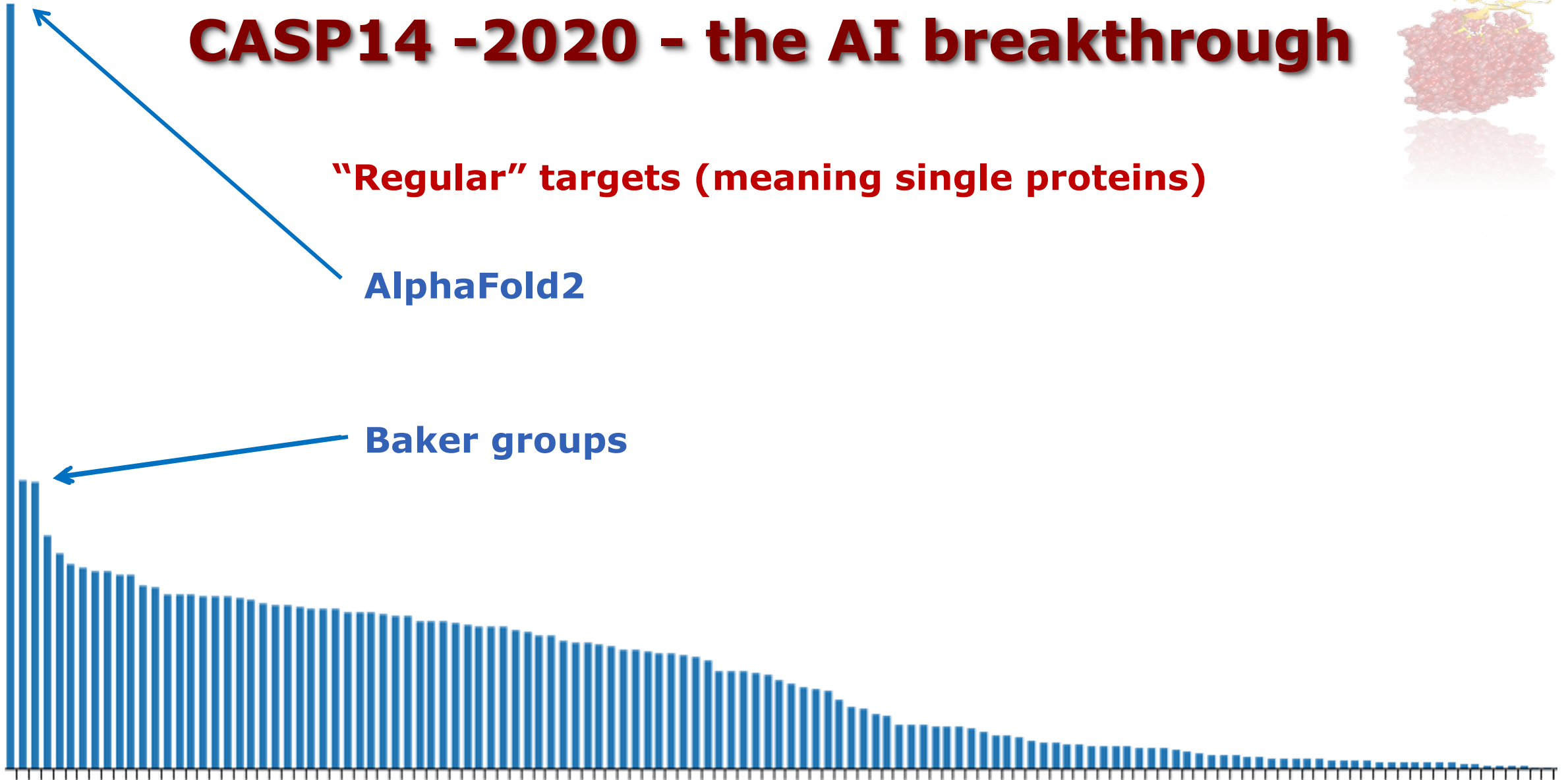
CASP14 -2020 - the AI breakthrough



“Regular” targets (meaning single proteins)

AlphaFold2

Baker groups



The 2024 Nobel prize in Chemistry



David Baker (1962)
U. Washington, Seattle;
Howard Hughes Medical
Institute, USA



Demis Hassabis (1976)
Google DeepMind, London,
United Kingdom

Ill. Niklas [Elmehed](#) © Nobel Prize Outreach



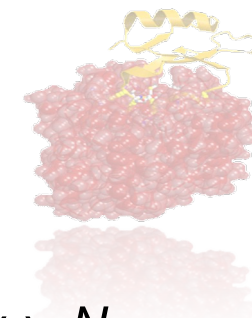
John Jumper (1985)
Google DeepMind, London,
United Kingdom

one half to David Baker
“for computational protein design”

and the other half jointly to Demis Hassabis and John Jumper
“for protein structure prediction”



CASP14/CAPRI

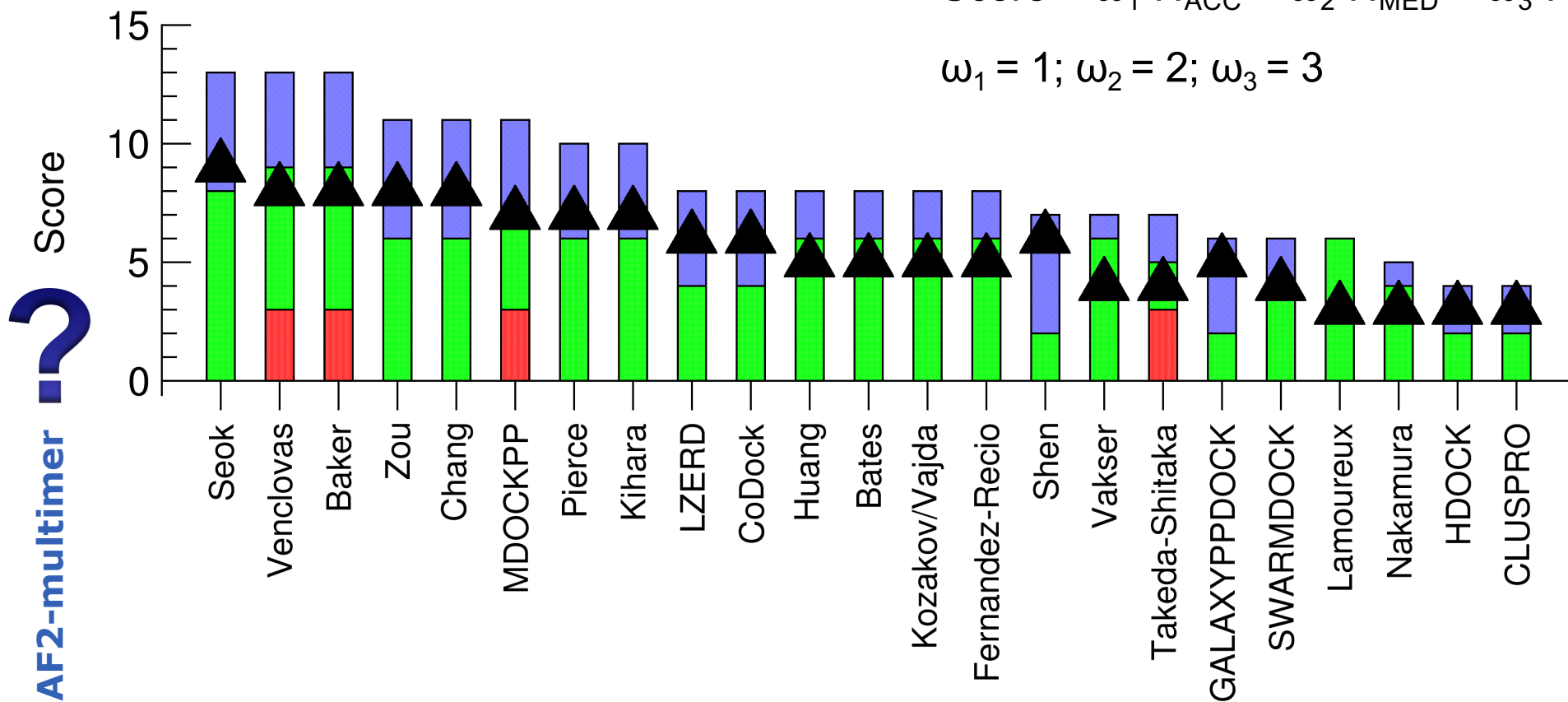


14 Complexes

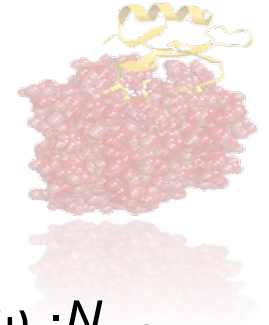
High – Medium – Acceptable

$$\text{Score} = \omega_1 \cdot N_{\text{ACC}} + \omega_2 \cdot N_{\text{MED}} + \omega_3 \cdot N_{\text{HIGH}}$$

$$\omega_1 = 1; \omega_2 = 2; \omega_3 = 3$$



CASP14/CAPRI

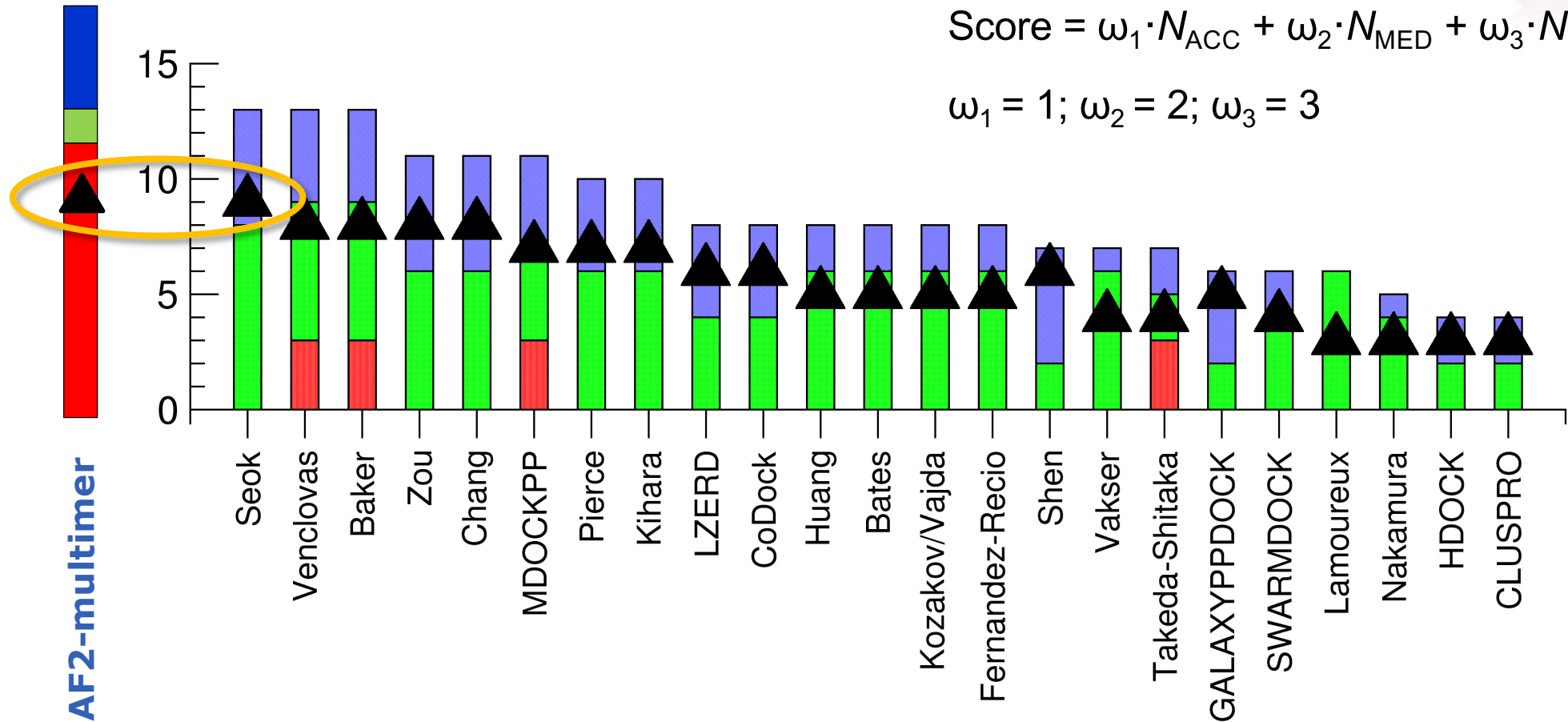


14 Complexes

High – Medium – Acceptable

$$\text{Score} = \omega_1 \cdot N_{\text{ACC}} + \omega_2 \cdot N_{\text{MED}} + \omega_3 \cdot N_{\text{HIGH}}$$

$$\omega_1 = 1; \omega_2 = 2; \omega_3 = 3$$



“Off-the-shelf” AF2-multimer: same number of targets (▲), but increased quality





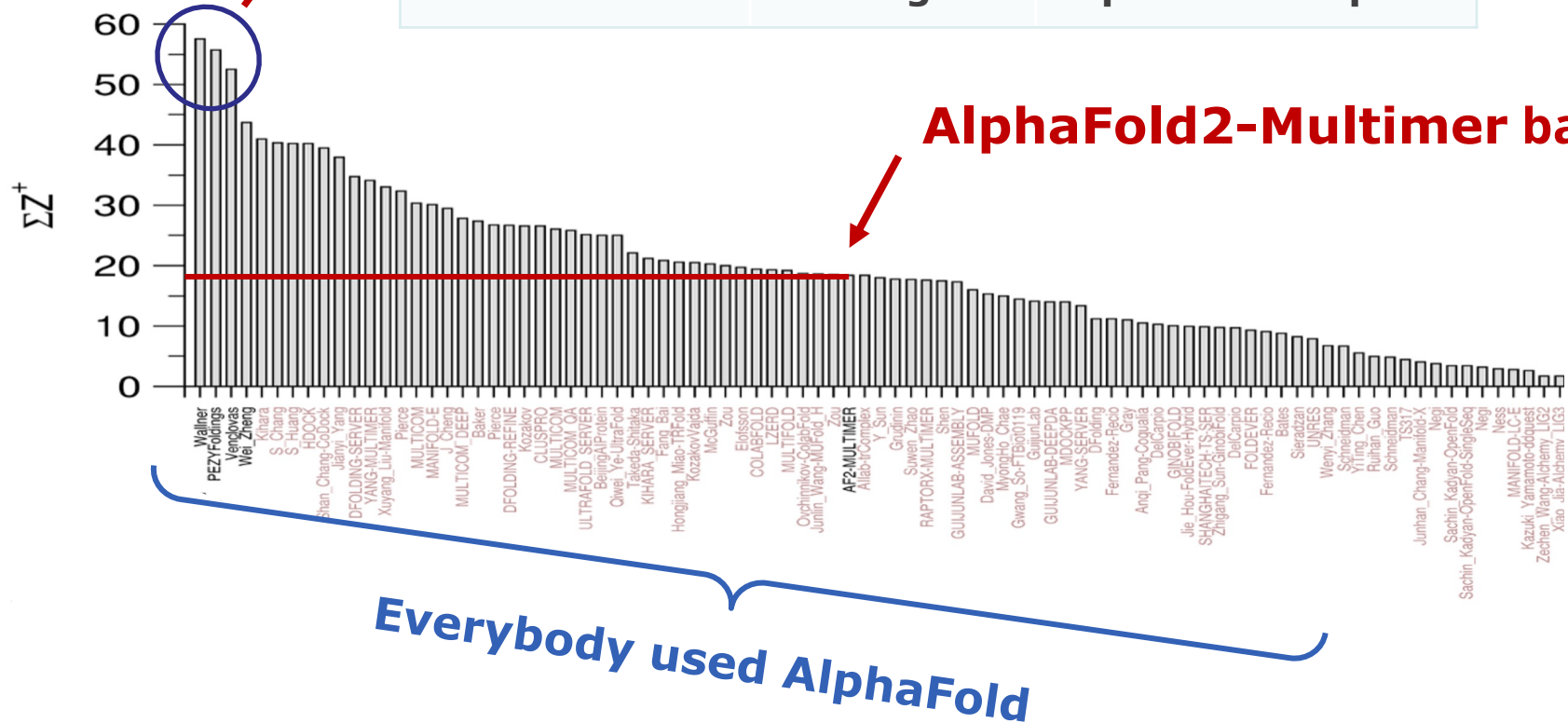
GOING BEYOND ALPHAFOLD SAMPLING AND SCORING



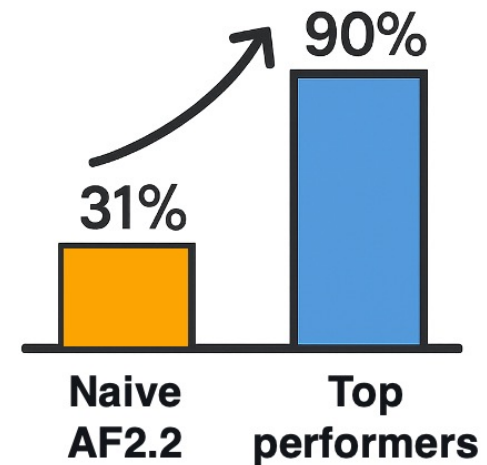
CASP15/CAPRI – 2022 -Assembly prediction




Wallner	Sampling	AlphaFold
PEZYFoldings	Pruning	AlphaFold input
Venclovas	Scoring	AlphaFold output



Increased success by customized AF2 pipelines





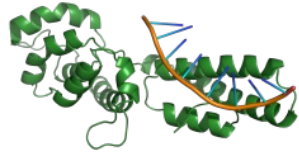
**WHAT STILL
DOESN'T WORK
WELL...**



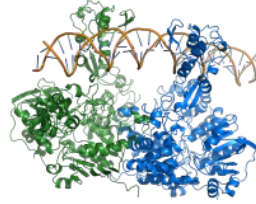
DNA and RNA binding



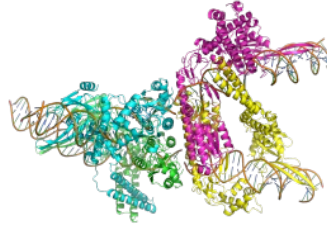
T304
M1282



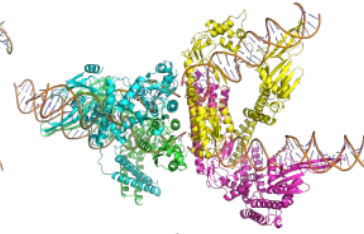
T306
M1276



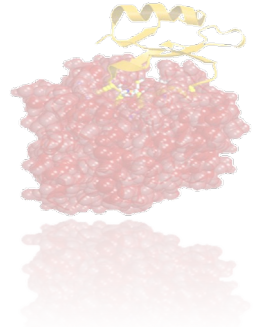
T308
M1287



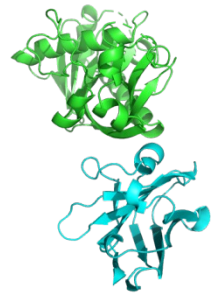
T254/T264
M1228v1
M1239v1



T255/T265
M1228v2
M1239v2

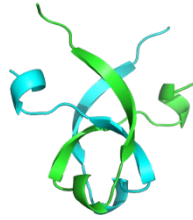


CASP16
CAPRI Rounds 57/58



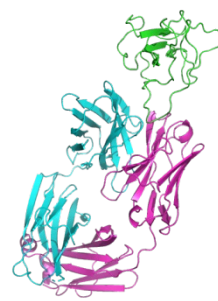
T207
H1142

**Nanobody
binding**



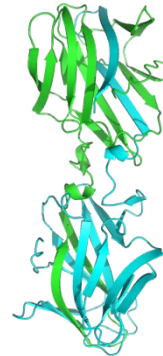
T213
T1160

**Ancient
protein
reconstruction**



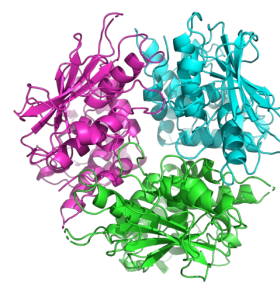
T216
H1166

**Antibody
binding**



T224
T1176

**Domain
swap**



T250/T252
T1249v1/T1249v2

**Multiple
conformations**

CASP15
CAPRI Round 54

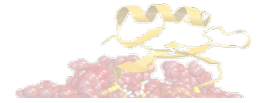




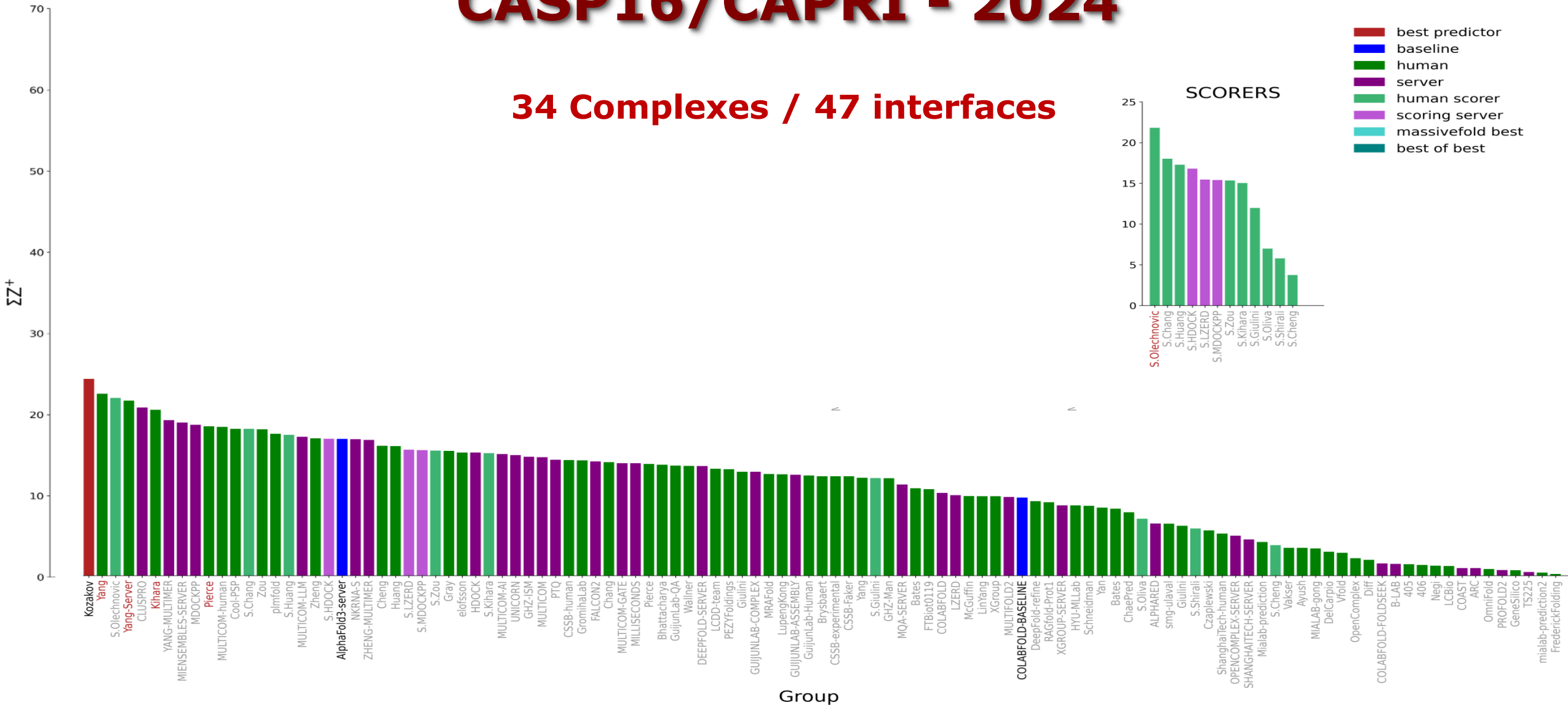
**AND WE CAN STILL
DO MUCH BETTER!**



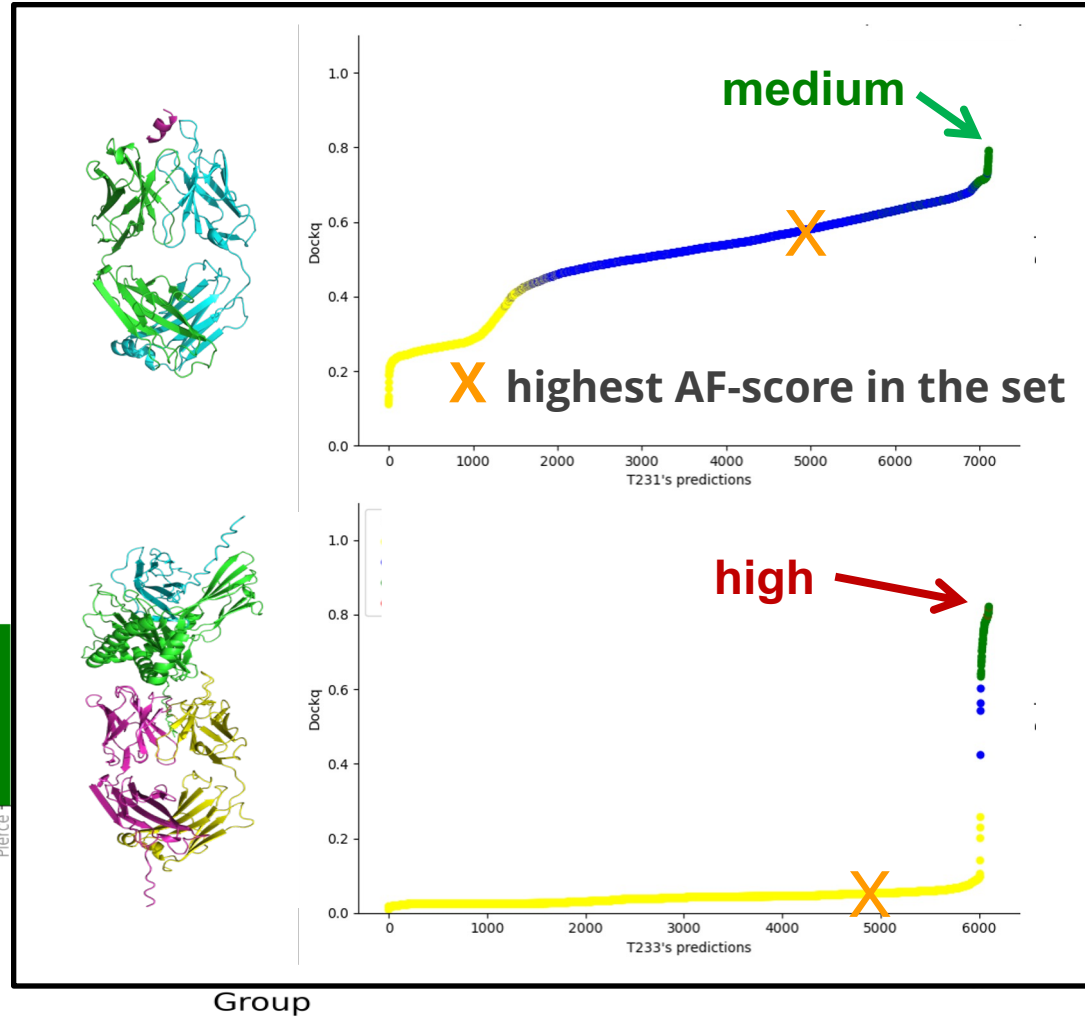
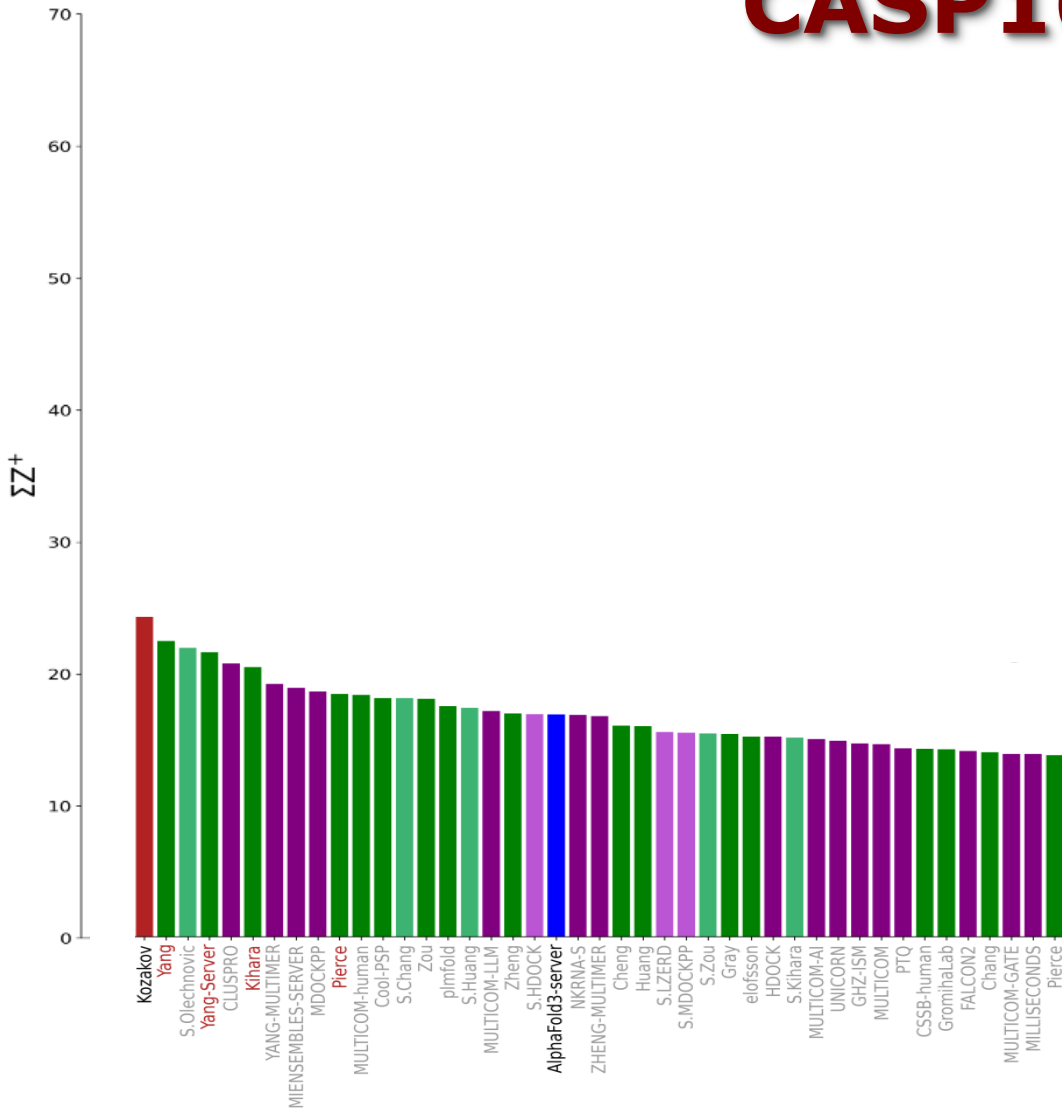
CASP16/CAPRI - 2024



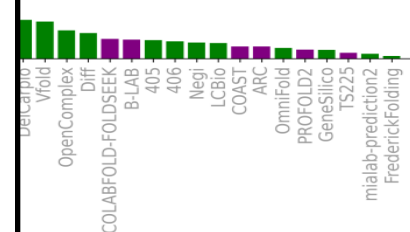
34 Complexes / 47 interfaces



CASP16/CAPRI - 2024

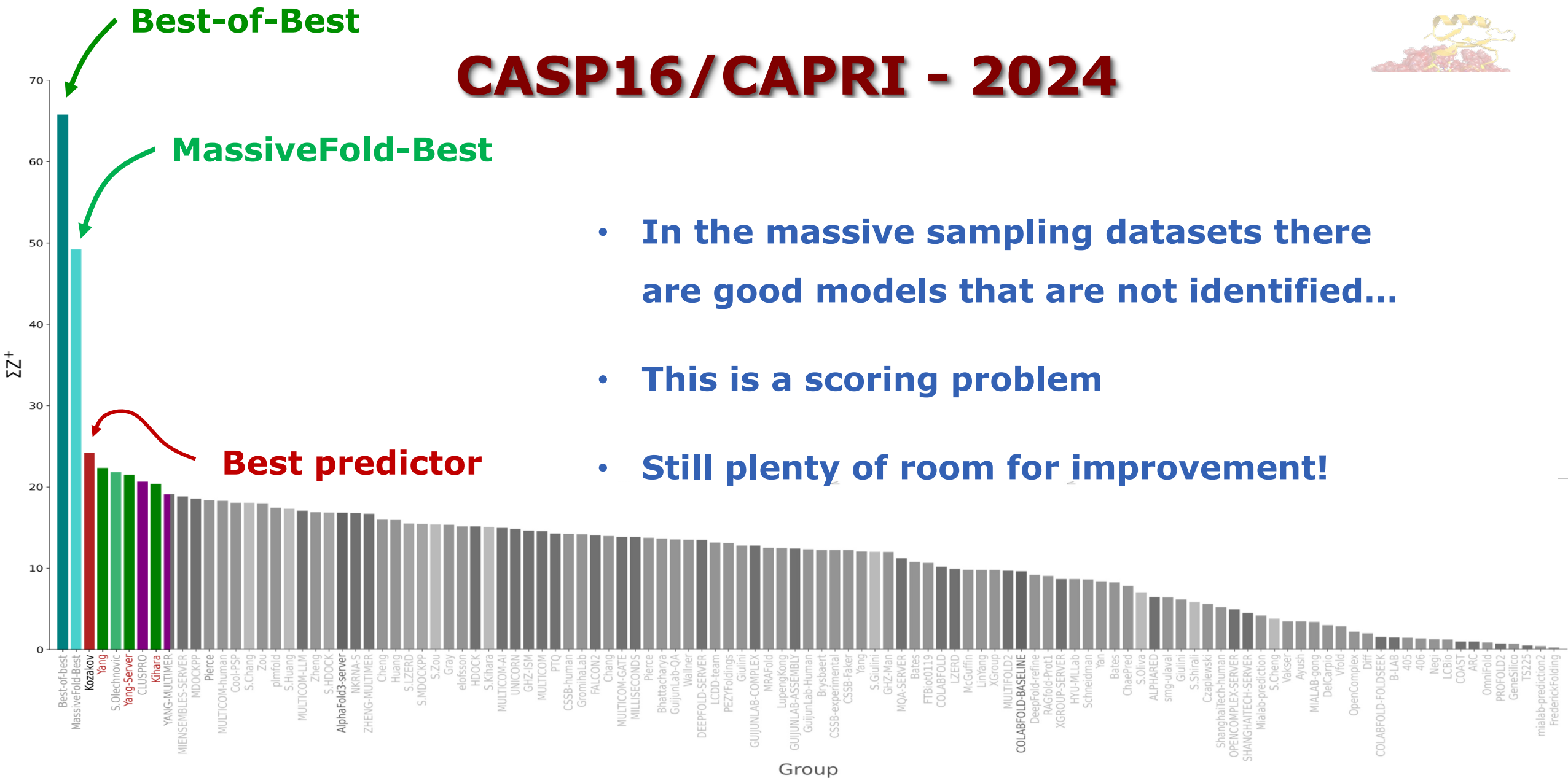


- best predictor
- baseline
- human
- server
- human scorer
- scoring server
- massivefold best
- best of best





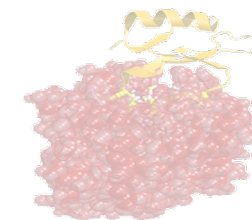
CASP16/CAPRI - 2024



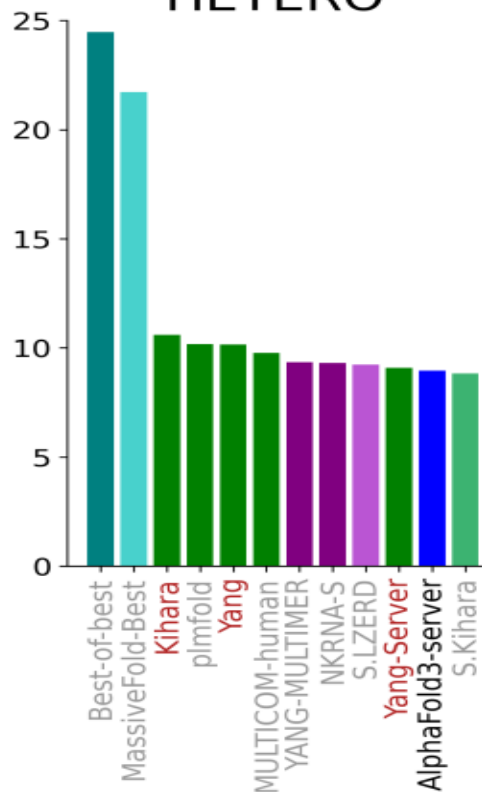
- In the massive sampling datasets there are good models that are not identified...
- This is a scoring problem
- Still plenty of room for improvement!



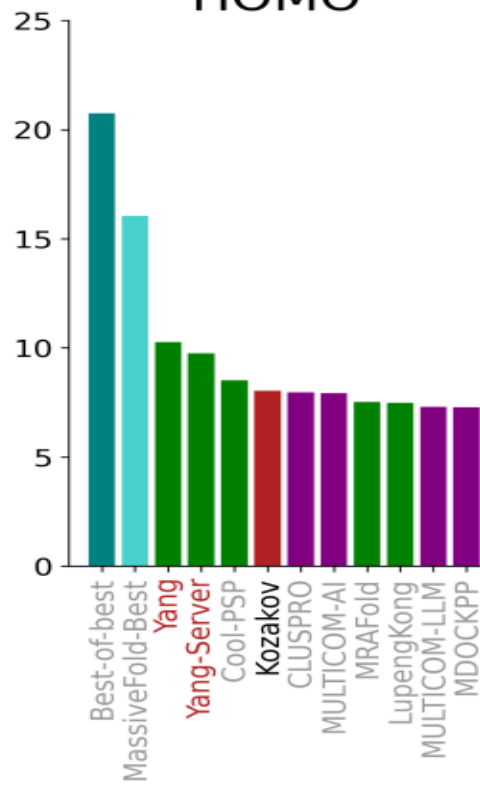
CASP16/CAPRI 2024



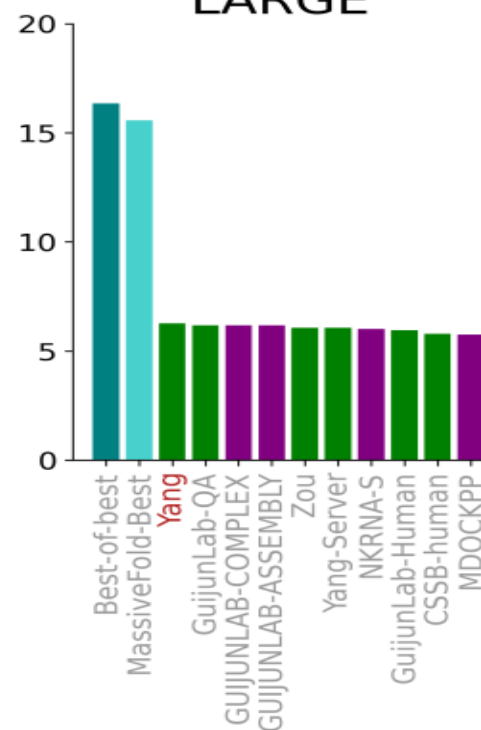
HETERO



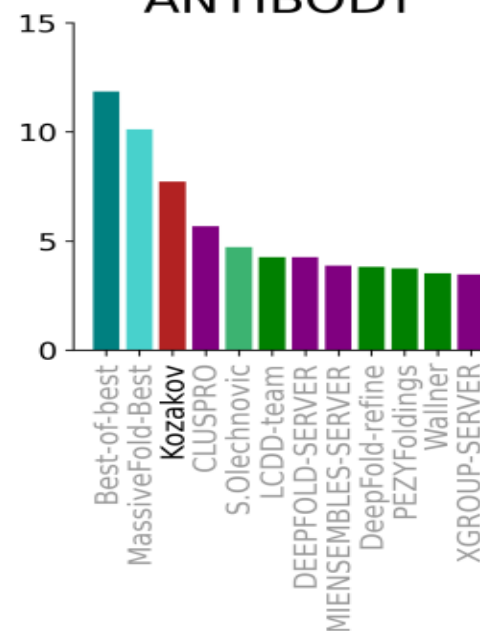
HOMO



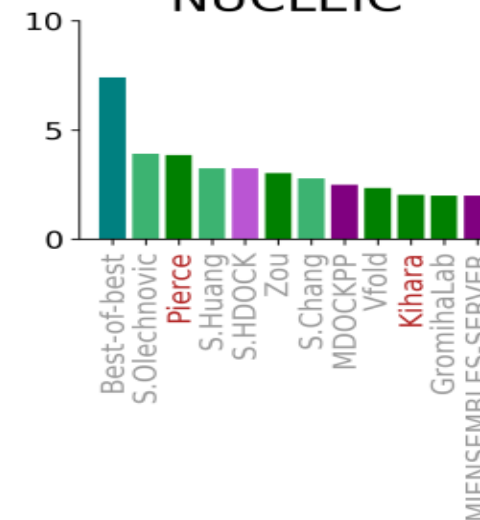
LARGE



ANTIBODY



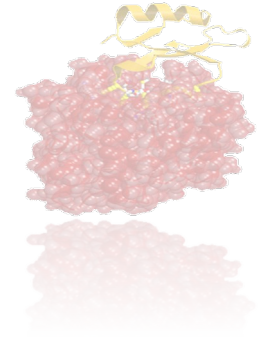
NUCLEIC



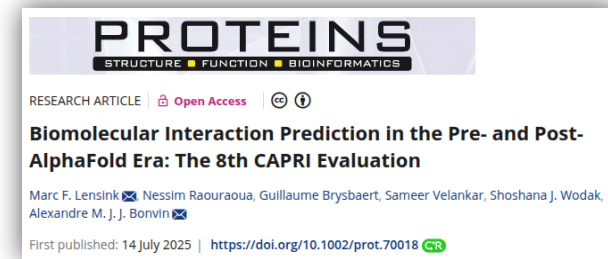
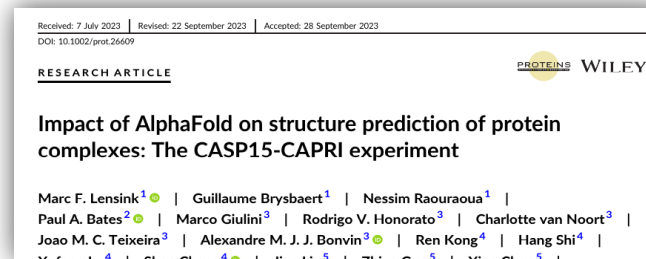
This hold for various types of complexes



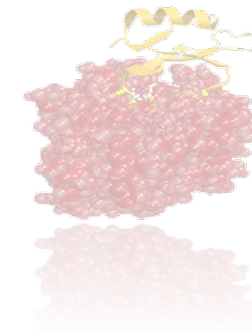
Conclusions



- There is clearly a *pre* and *post* AlphaFold world
- **CASP & CAPRI**
 - Dynamic experiments, catalysts for method development
 - Define the state-of-the-art
- **Challenges:**
 - Large domain motions, weak binding
 - Anti/nanobodies, intertwining, nucleic acids (especially RNA)
 - Detailed interface descriptions: side-chain positioning and water molecules
 - **Scoring**
 - **Getting targets!**



CASP/CAPRI Acknowledgements



CAPRI Management Committee

Alexandre Bonvin

Marc Lensink

Michael Sternberg

Sandor Vajda

Ilya Vakser

Sameer Velankar

Zhiping Weng

**All CAPRI participants
and contributing
experimentalists**

CASP

John Moult

Andriy Krystafovych

Krzysztof Fidelis

CAPRI Scientific Advisory Board

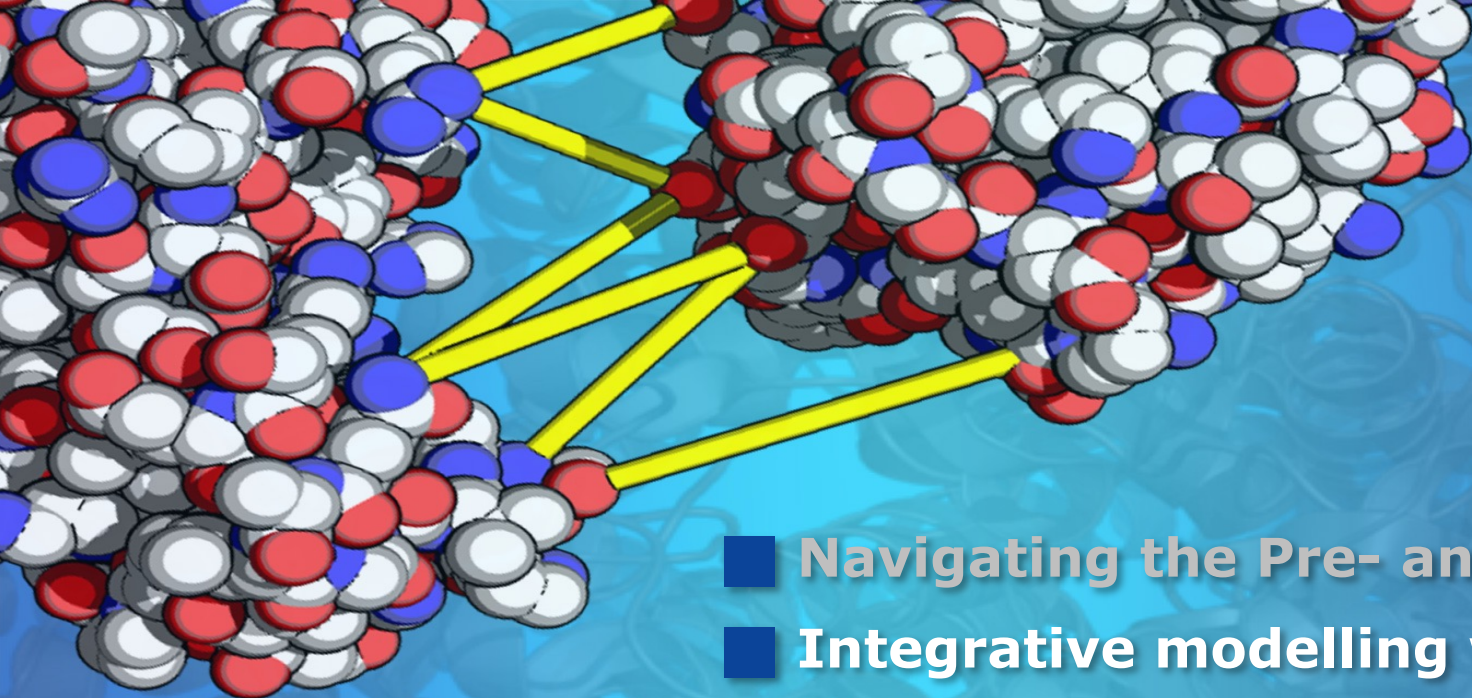
Rommie Amaro

Stephen Burley

Wah Chiu

Kristina Djinovic-Carugo





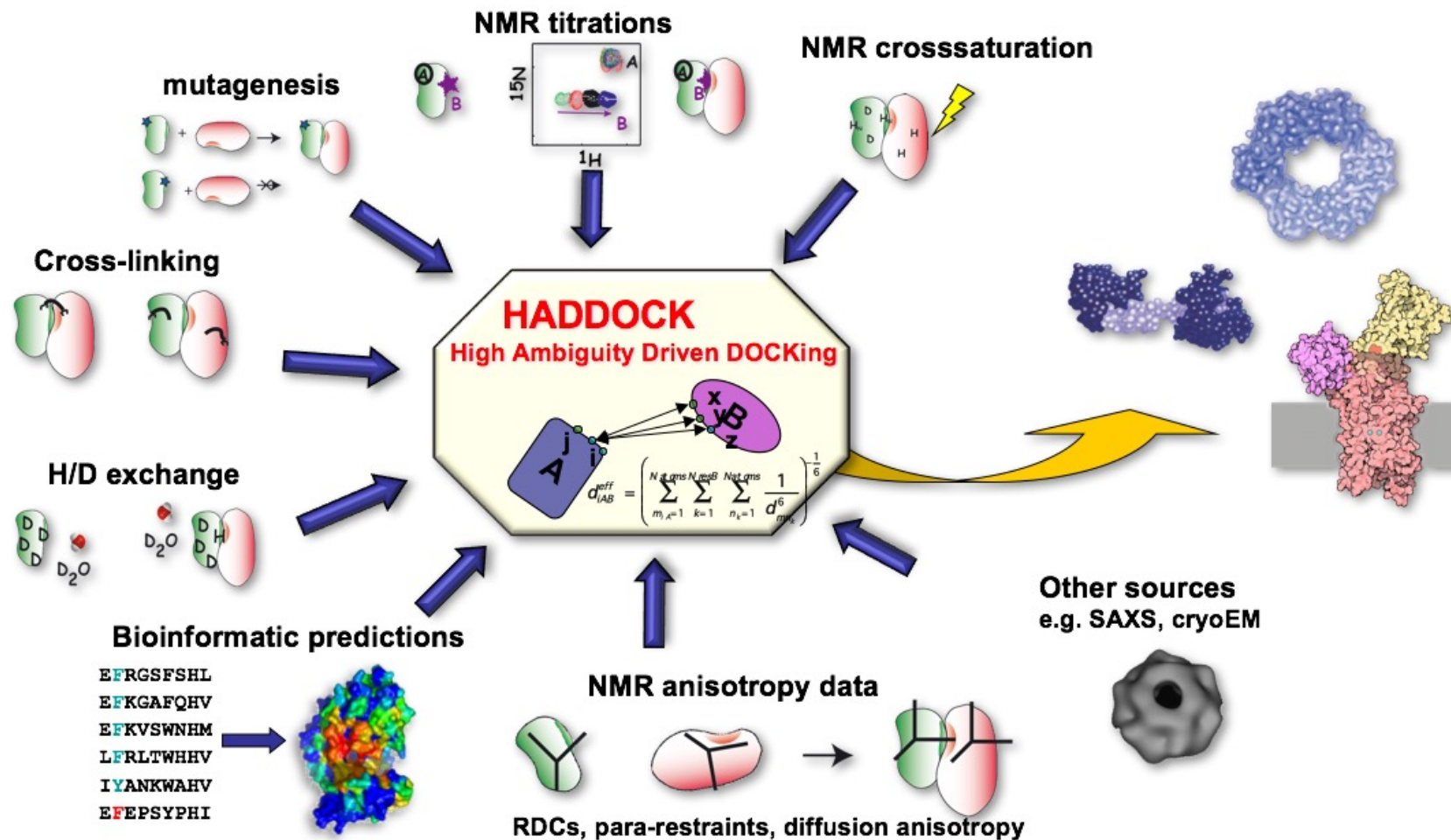
Overview

- Navigating the Pre- and Post-AlphaFold Divide
- Integrative modelling with HADDOCK
- Introducing HADDOCK3
- Application examples
 - An NMR-based application – where AlphaFold fails
 - AI-based antibody-antigen modelling
 - Cyclic peptides
- Conclusions & perspectives

```
111000100101010110101010
100011010101000101111010110101
10100001000111101010101110000110101
0001010101011110010100110010101010100
010101111110101010000101010100000101010100
0010110101011101010001111010101010001010010101011111
000111111010101011001100010111000101011010100010101000101
```

HADDOCK: An integrative modeling platform

<http://www.bonvinlab.org/software>

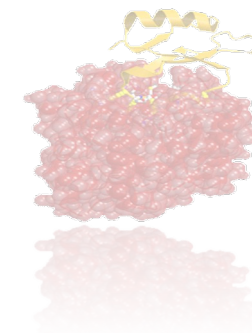


Journal of the American Chemical Society > Vol 125/Issue 7 > Article

ARTICLE | January 21, 2003

HADDOCK: A Protein-Protein Docking Approach Based on Biochemical or Biophysical Information

Cyril Dominguez, Rolf Boelens, and Alexandre M. J. J. Bonvin

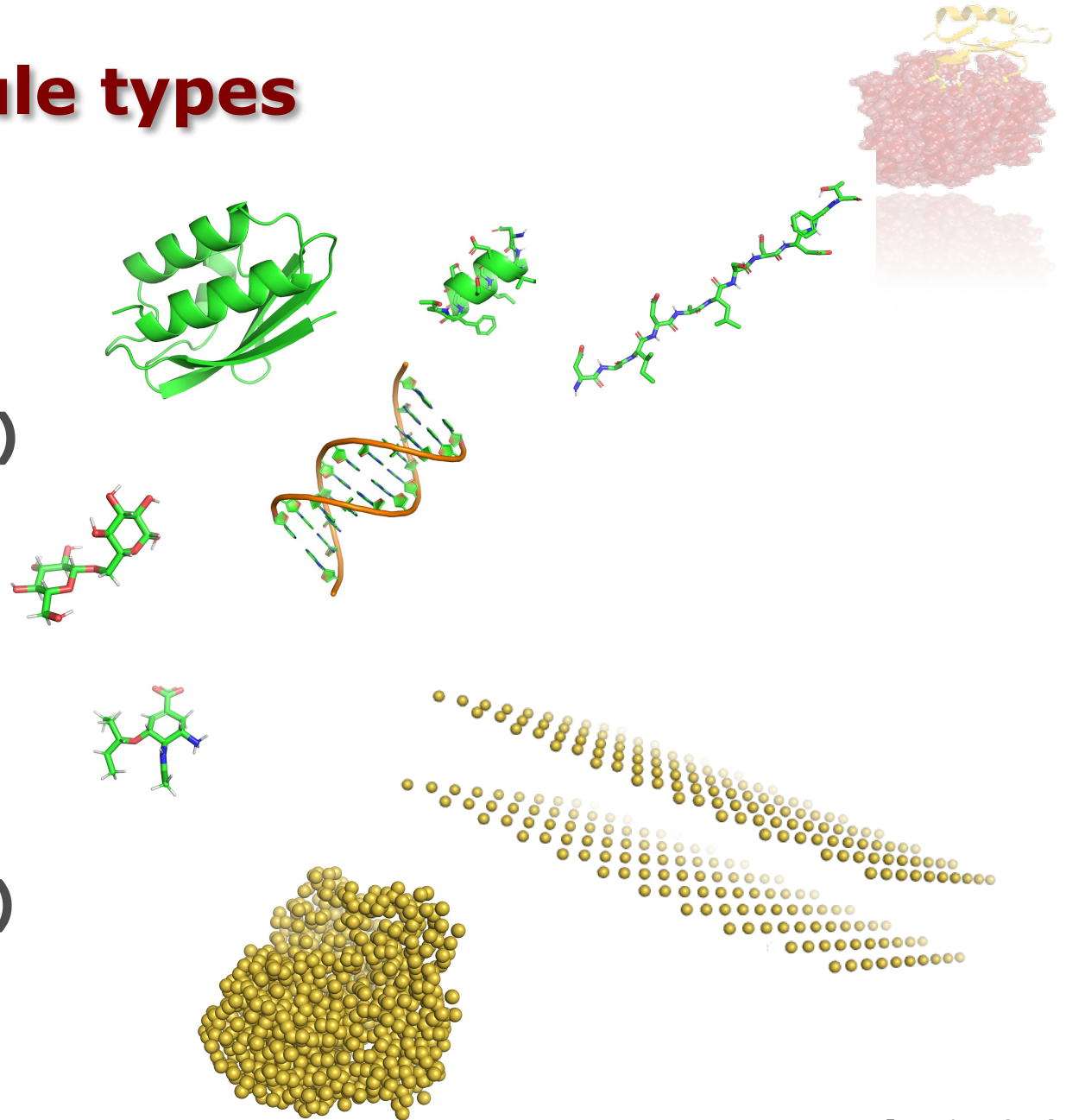


Universiteit Utrecht

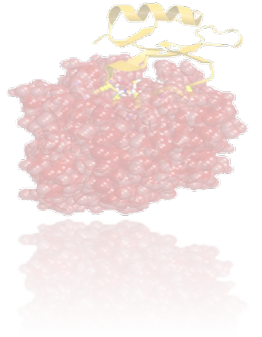
[Faculty of Science
Chemistry]

Molecule types

- **Various mixtures of:**
 - **Proteins / peptides**
 - **Nucleic acids (DNA/RNA)**
 - **Glycans**
 - **Glycosylated proteins**
 - **Small molecules**
 - **Shapes**
 - **Coarse graining (Martini)**



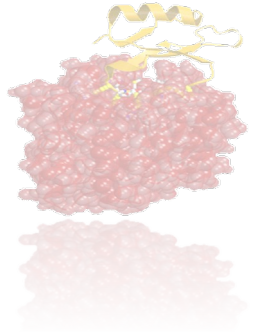
Restraint types



- Distances (ambiguous / unambiguous)
- Dihedral angles
- NMR residual dipolar couplings
- NMR pseudo contact shifts
- NMR relaxation anisotropy
- Cryo-EM
- Shapes
- Radius of gyration
- Symmetry (C2 – C6)
- Non-crystallographic symmetry



HADDOCK docking protocol



it0



it1

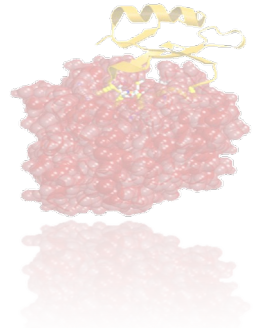


itw

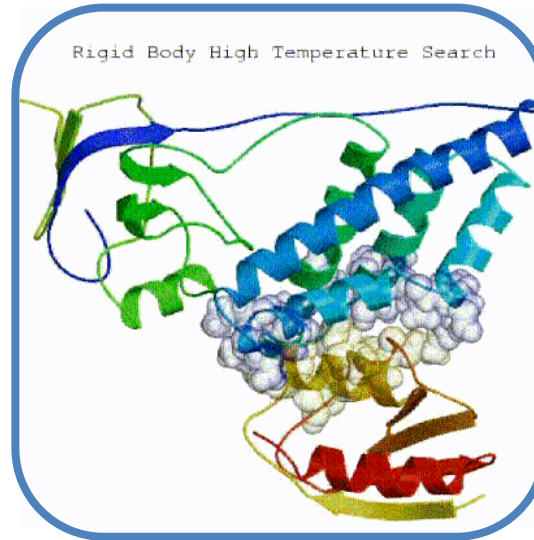
**Succession of energy minimization and molecular dynamics protocols
reminiscent of NMR structure calculations**



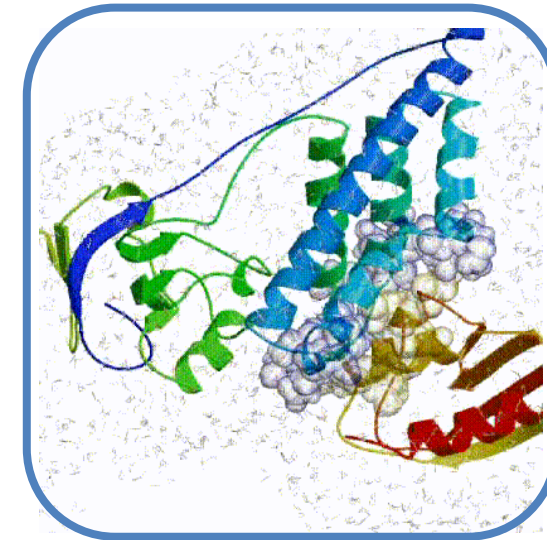
HADDOCK docking protocol



it0



it1

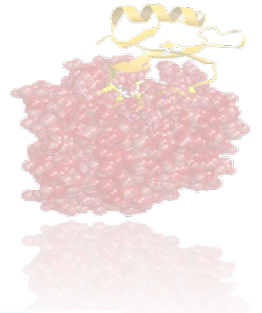


itw

**Succession of energy minimization and molecular dynamics protocols
reminiscent of NMR structure calculations**

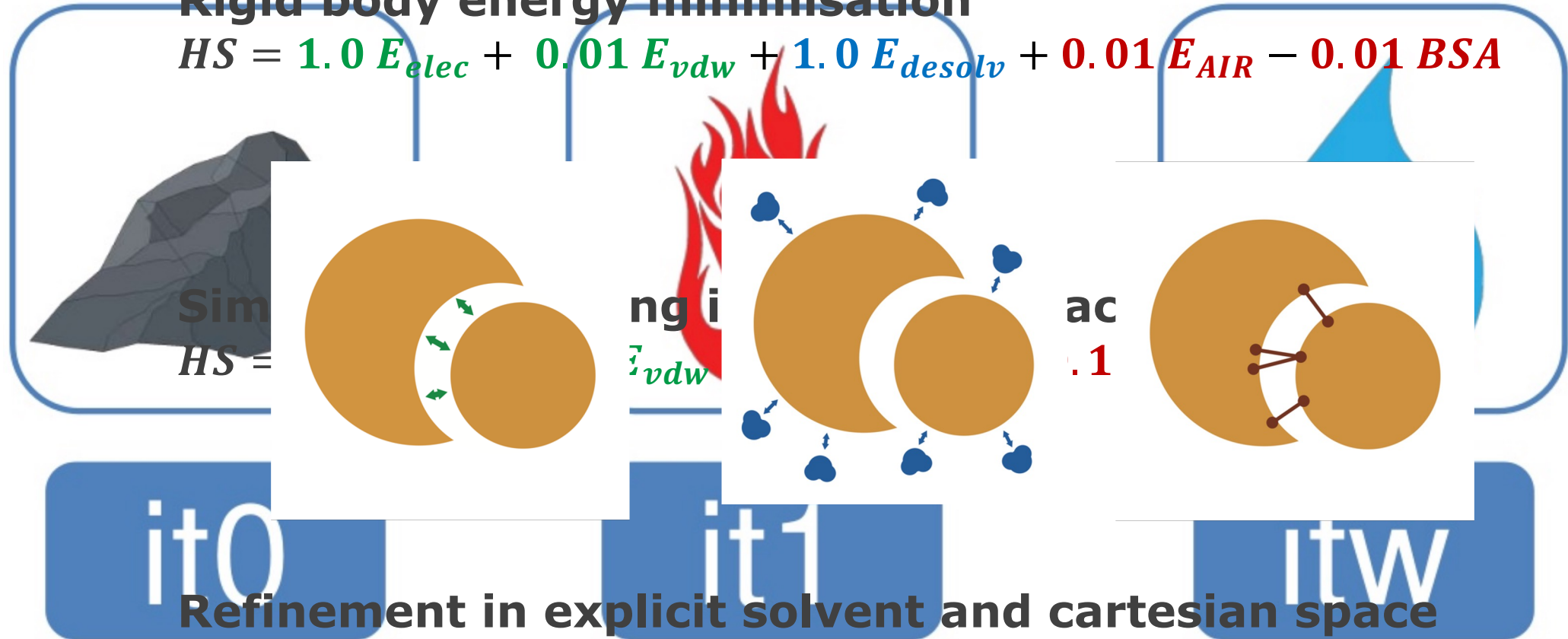


HADDOCK protocol



Rigid body energy minimisation

$$HS = 1.0 E_{elec} + 0.01 E_{vdw} + 1.0 E_{desolv} + 0.01 E_{AIR} - 0.01 BSA$$



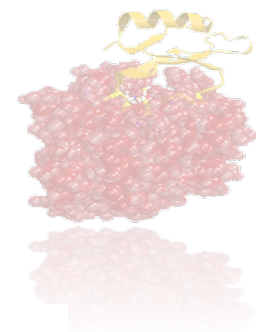
Refinement in explicit solvent and cartesian space

$$HS = 0.2 E_{elec} + 1.0 E_{vdw} + 1.0 E_{desolv} + 0.1 E_{AIR}$$



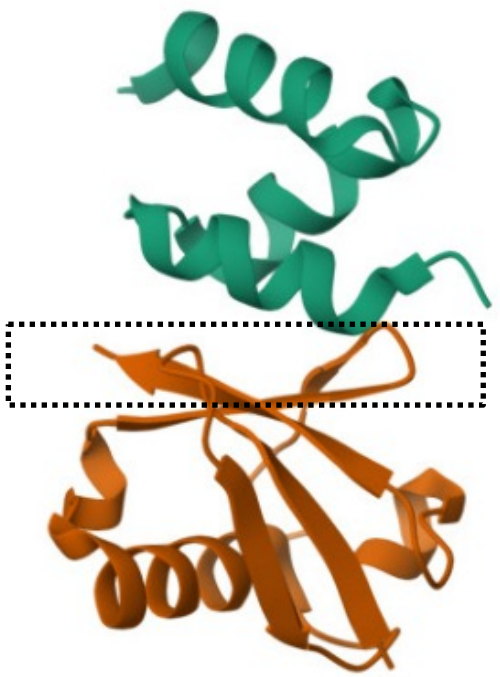
Manon Réau

DeepRank-GNN: Graph Neural Network

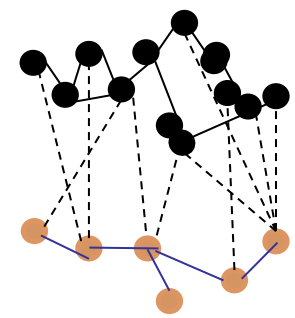


Node feature: PSSMs, Type, Polarity, BSA, Charge

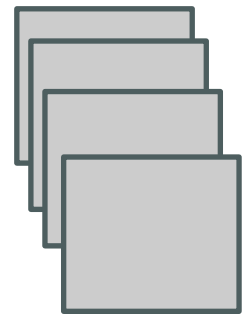
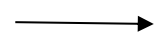
Edge feature: Distance



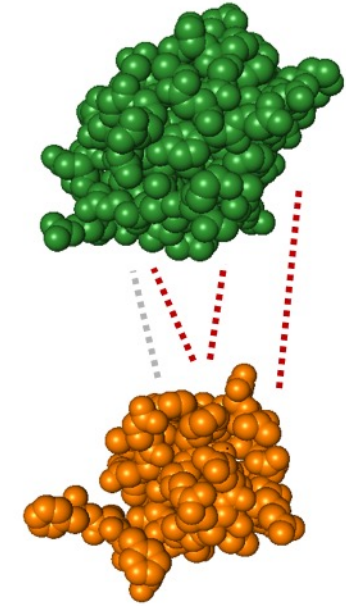
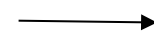
PPI structure



Residue-level interface graph



GNN training
(Model architecture Adapted from DeepRank-GNN)



fnat prediction

JOURNAL ARTICLE

DeepRank-GNN: a graph neural network framework to learn patterns in protein-protein interfaces

Manon Réau, Nicolas Renaud, Li C Xue, Alexandre M J J Bonvin ✉ Author Notes

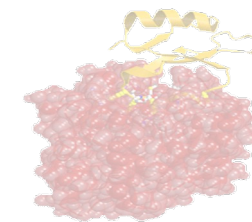
Bioinformatics, Volume 39, Issue 1, January 2023, btac759, <https://doi.org/10.1093/bioinformatics/btac759>





Xiaotong Xu

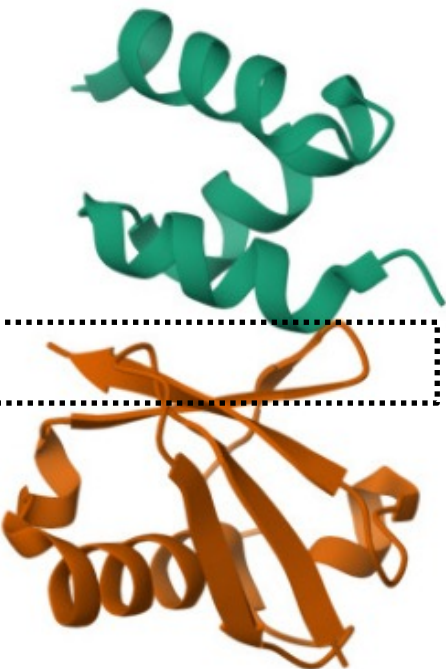
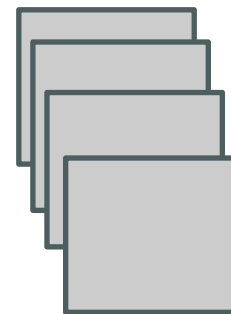
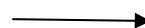
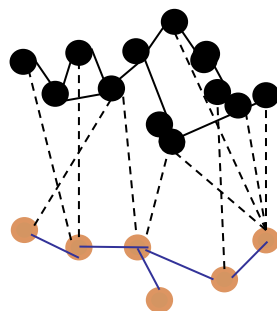
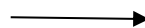
Graph Neural Network with protein language model embeddings



Node feature: ~~PSSMs~~, Type, Polarity, BSA, Charge

↑
ESM-2 embeddings

Edge feature: Distance



PPI structure

Residue-level interface graph

GNN training
(Model architecture Adapted from DeepRank-GNN)

Compared with DeepRank-GNN with PSSM features,

- **Faster!** (no need to generate the PSSM)
- **Equal or better performance on two PPI tasks**
 - Scoring
 - Classification of biological vs crystallographic interfaces

<https://github.com/haddock/DeepRank-GNN-esm>

JOURNAL ARTICLE

DeepRank-GNN-esm: a graph neural network for scoring protein-protein models using protein language model

Xiaotong Xu, Alexandre M J J Bonvin

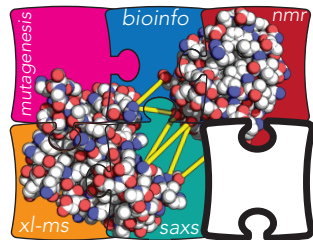
Bioinformatics Advances, Volume 4, Issue 1, 2024, vbad191,

<https://doi.org/10.1093/bioadv/vbad191>

Published: 05 January 2024 Article history



HADDOCK web portal (WeNMR)



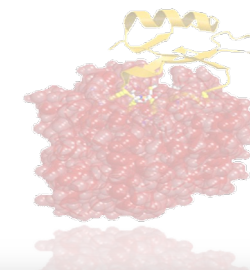
HADDOCK
High-Ambiguity Driven Docking

- > 79000 registered users
- > 700000 served runs since June 2008
- 65% on the EOSC HTC resources (>85% for 2.4)

Vargas Honorato *et al.* Nature Prot. 2024

Van Zundert *et al.* J.Mol.Biol. 2016

De Vries *et al.* Nature Prot. 2010



HADDOCK 2.4
@Bonvinlab

WELCOME TO THE UTRECHT BIOMOLECULAR INTERACTION WEB PORTAL >>

Welcome! **HADDOCK** (High Ambiguity Driven protein-protein **DOCK**ing) is an information-driven flexible docking approach for the modeling of biomolecular complexes.

HADDOCK distinguishes itself from ab-initio docking methods in the fact that it encodes information from identified or predicted protein interfaces in ambiguous interaction restraints (AIRs) to drive the docking process. It also allows to define specific unambiguous distance restraints (e.g. from MS cross-links) and supports a variety of other experimental data including NMR residual dipolar couplings, pseudo contact shifts and cryo-EM maps.

HADDOCK can deal with a large class of modeling problems including protein-protein, protein-nucleic acids and protein-ligand complexes, including multi-bodies (N>2) assemblies.

HADDOCK is one of the **flagship software** in the EU H2020 BioExcel Center of Excellence for Biomolecular Research.

Register | **Submit a new job** | **See our tutorials** | **Get Help**

<https://wenmr.science.uu.nl>

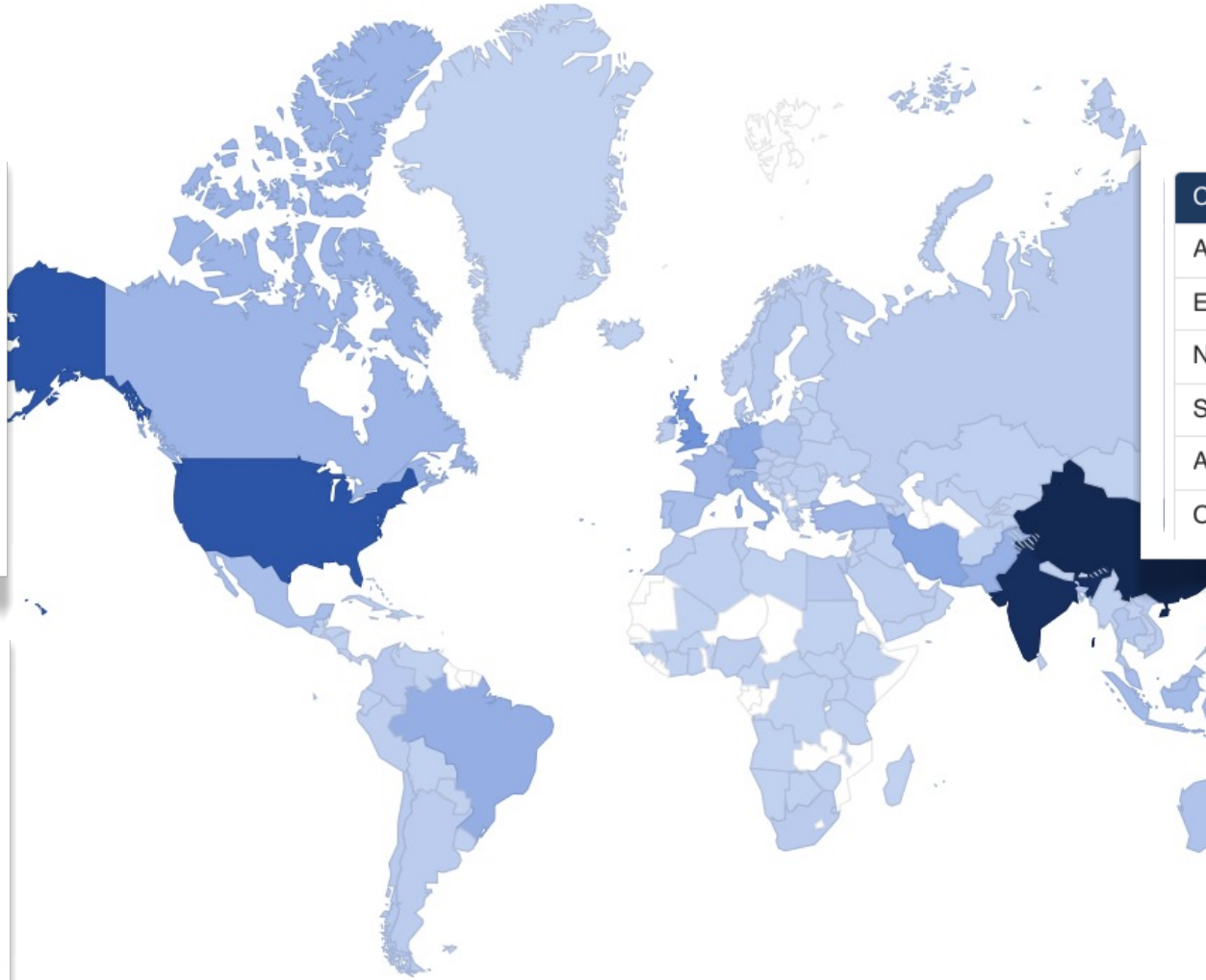
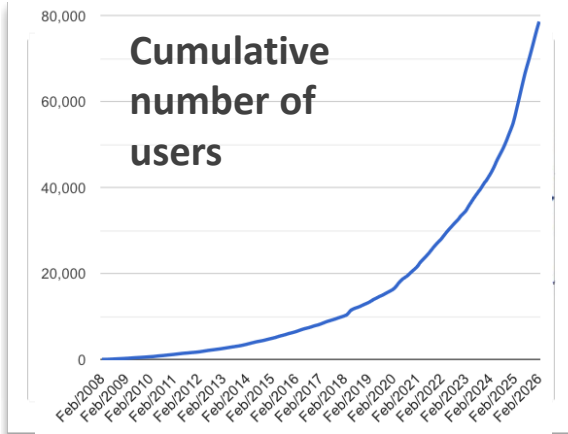
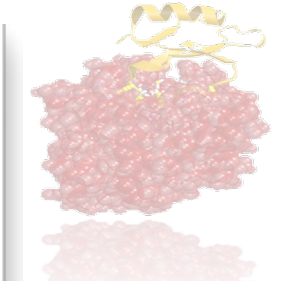


Universiteit Utrecht

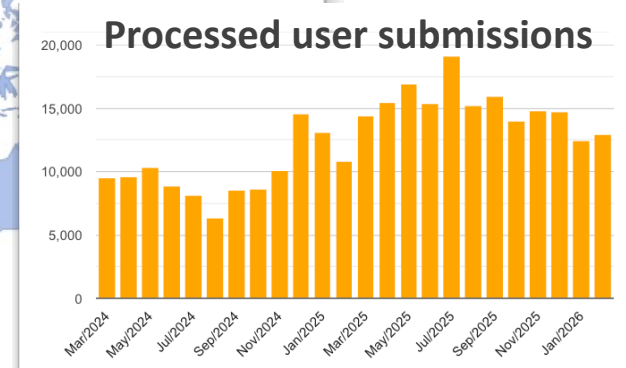
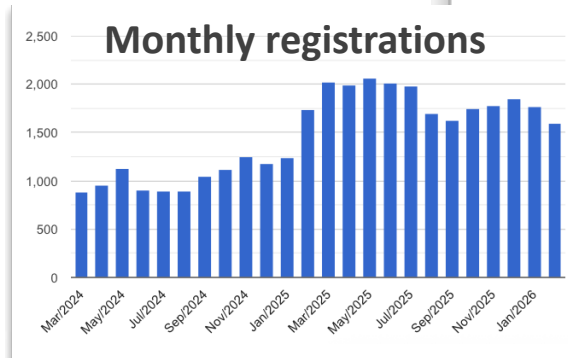
[Faculty of Science
Chemistry]

Haddock web portal

<https://wenmr.science.uu.nl/new/stats>

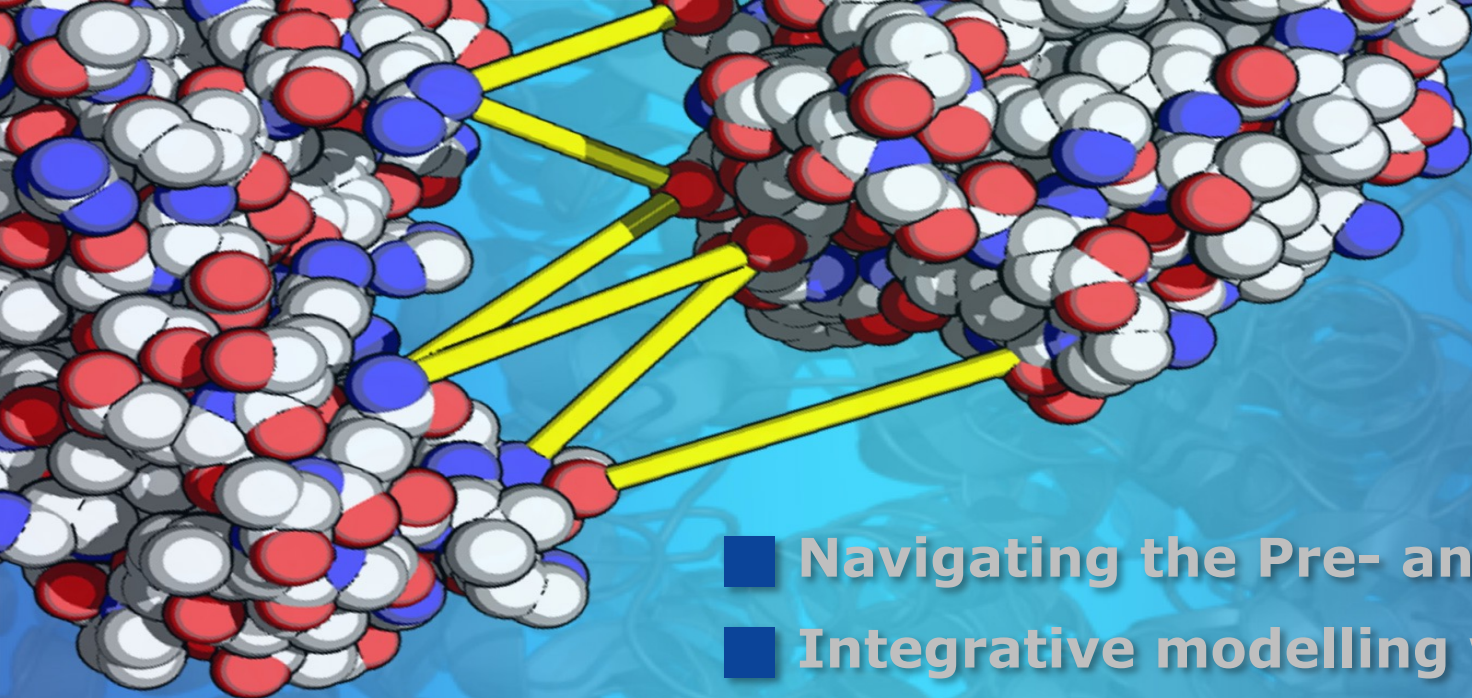


Continent	# of Users
Asia	44565
Europe	17088
North America	12185
South America	3228
Africa	1589
Oceania	893



The HADDOCK web portal is being used by **79554 users** across **184 countries!**



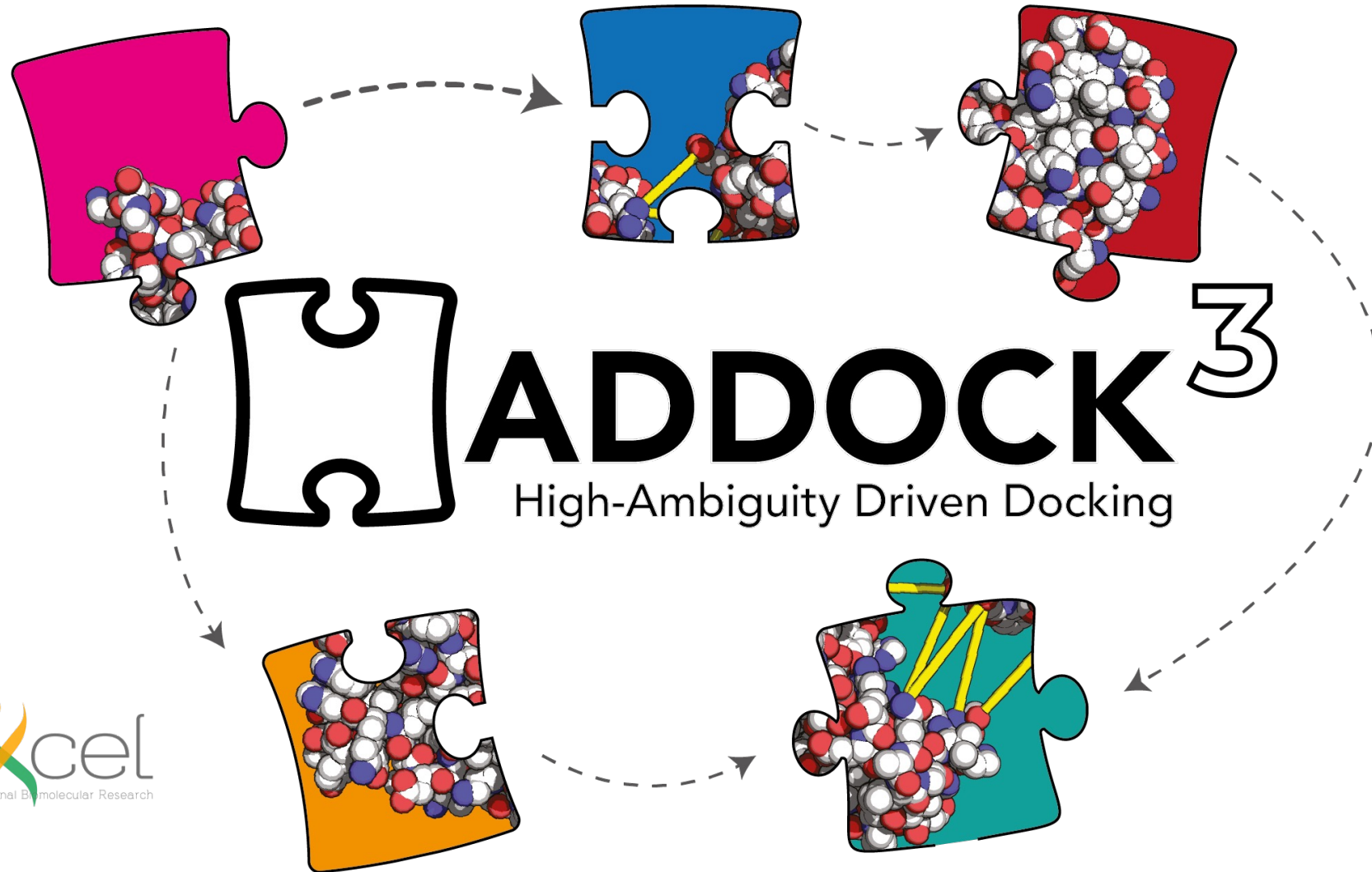


Overview

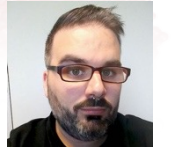
- Navigating the Pre- and Post-AlphaFold Divide
- Integrative modelling with HADDOCK
- Introducing HADDOCK3
- Application examples
 - An NMR-based application – where AlphaFold fails
 - AI-based antibody-antigen modelling
 - Cyclic peptides
- Conclusions & perspectives

```
111000100101010110101010
100011010101000101111010110101
10100001000111101010101110000110101
0001010101011110010100110010101010100
0101011111110101010000101010100000101010100
0010110101011101010001111010101010001010010101011111
0001111111010101011001100010111000101011010100010101000101
```

Introducing HADDOCK v3



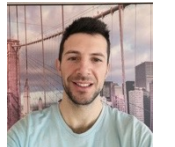
Rodrigo
Vargas Honorato



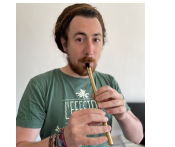
Brian Jimenez



João MC Teixeira

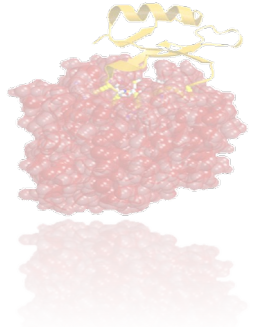


Marco Giulini



Victor Reys

From static workflow ...



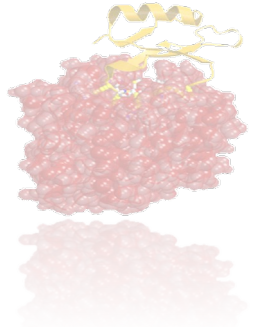
HADDOCK 2.X

Parameterizable fixed pipeline

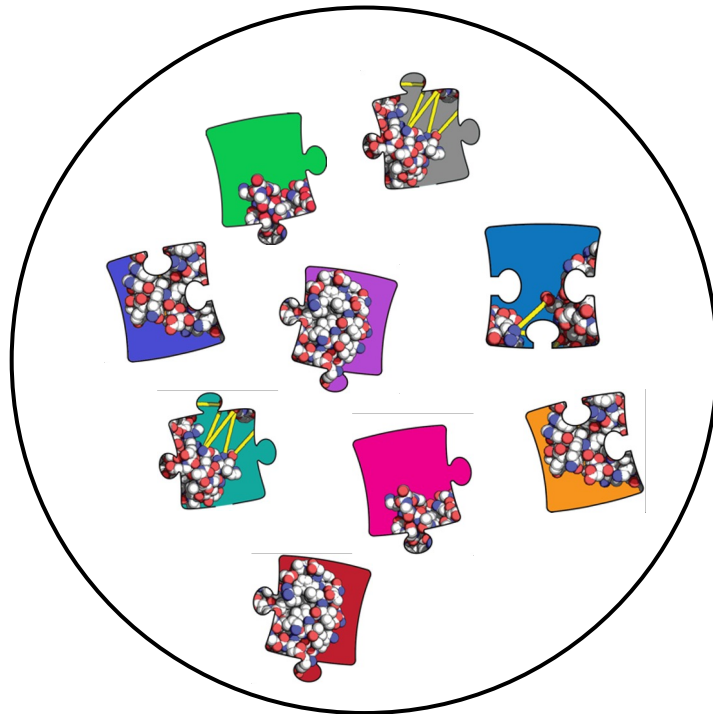




... to a modular workflow architecture



Catalogue of independent **modules**



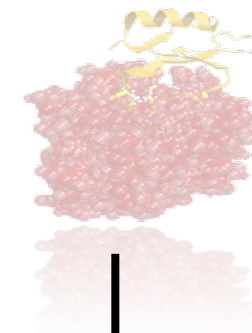
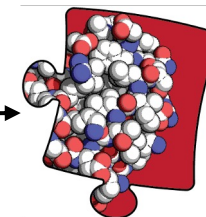
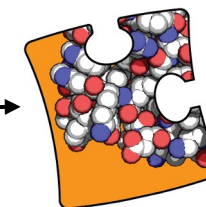
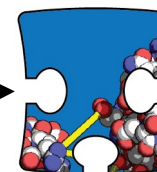
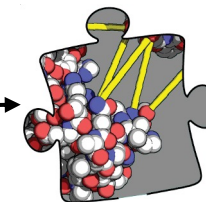
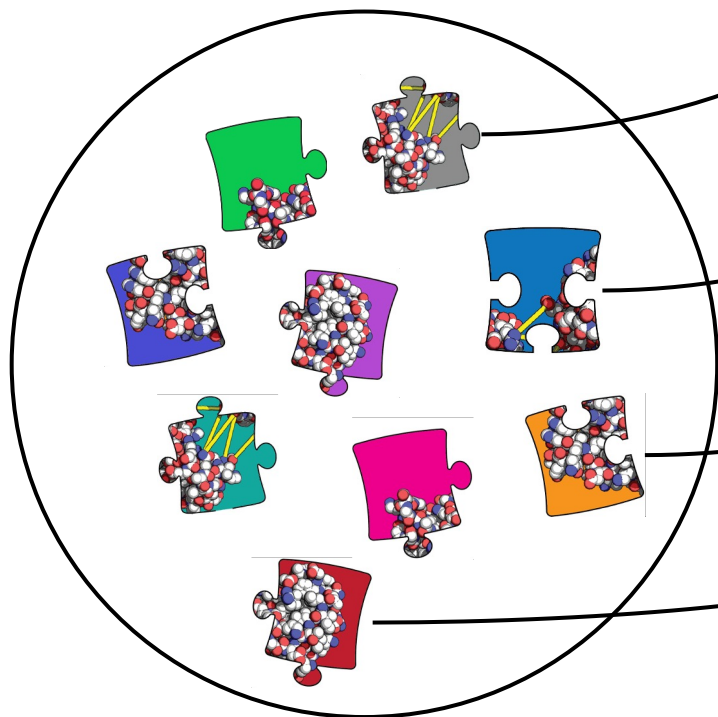
HADDOCK 2





Combining the different pieces

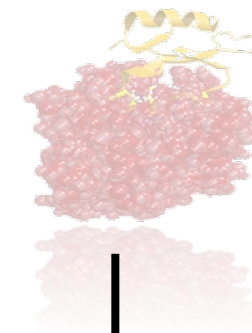
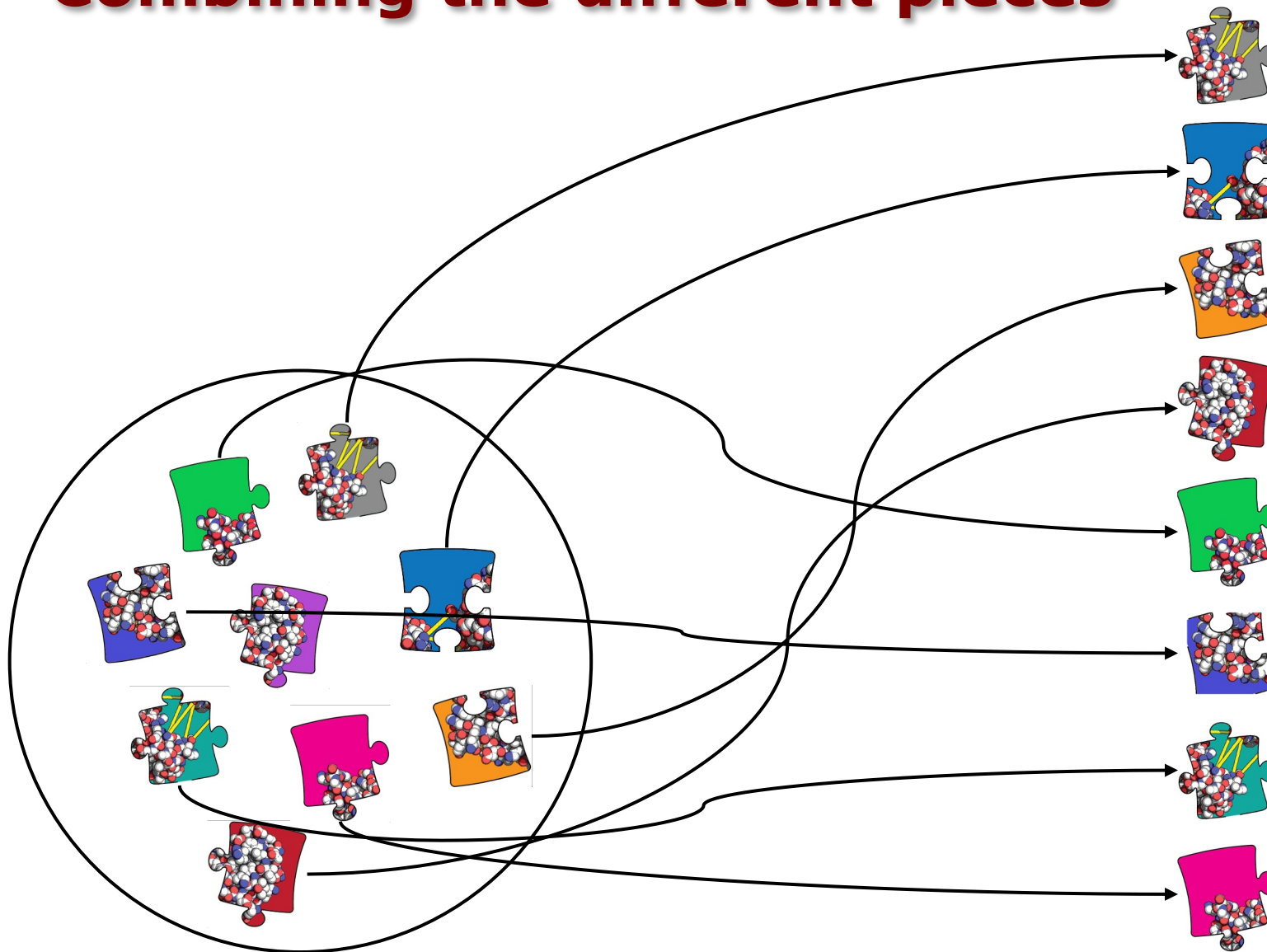
Catalogue of independent modules





Combining the different pieces

Catalogue of independent modules





Current status – v2025.11.0



topology

 All-atom topology





analysis

-  CAPRI eval
-  RMSD matrix
-  iRMSD matrix
-  Clust by FCC
-  Clust by RMSD
-  Select top model
-  Select top cluster
-  Contactmap
-  Alascan
-  DeepRank



sampling

-  Rigid Body
-  lightdock
-  gdock

refinement

-  Semi-Flexible ref.
-  Energy minimization
-  Short MD w/ explicit water
-  openMM

scoring

-  All-atom scoring
-  All-atom scoring w/ MDexp. water

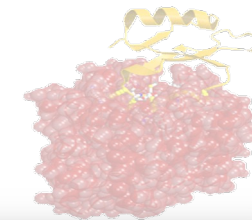
CLI

- Analysis
- Restraints generation
- Solvent accessibility

<https://github.com/haddock/haddock3>

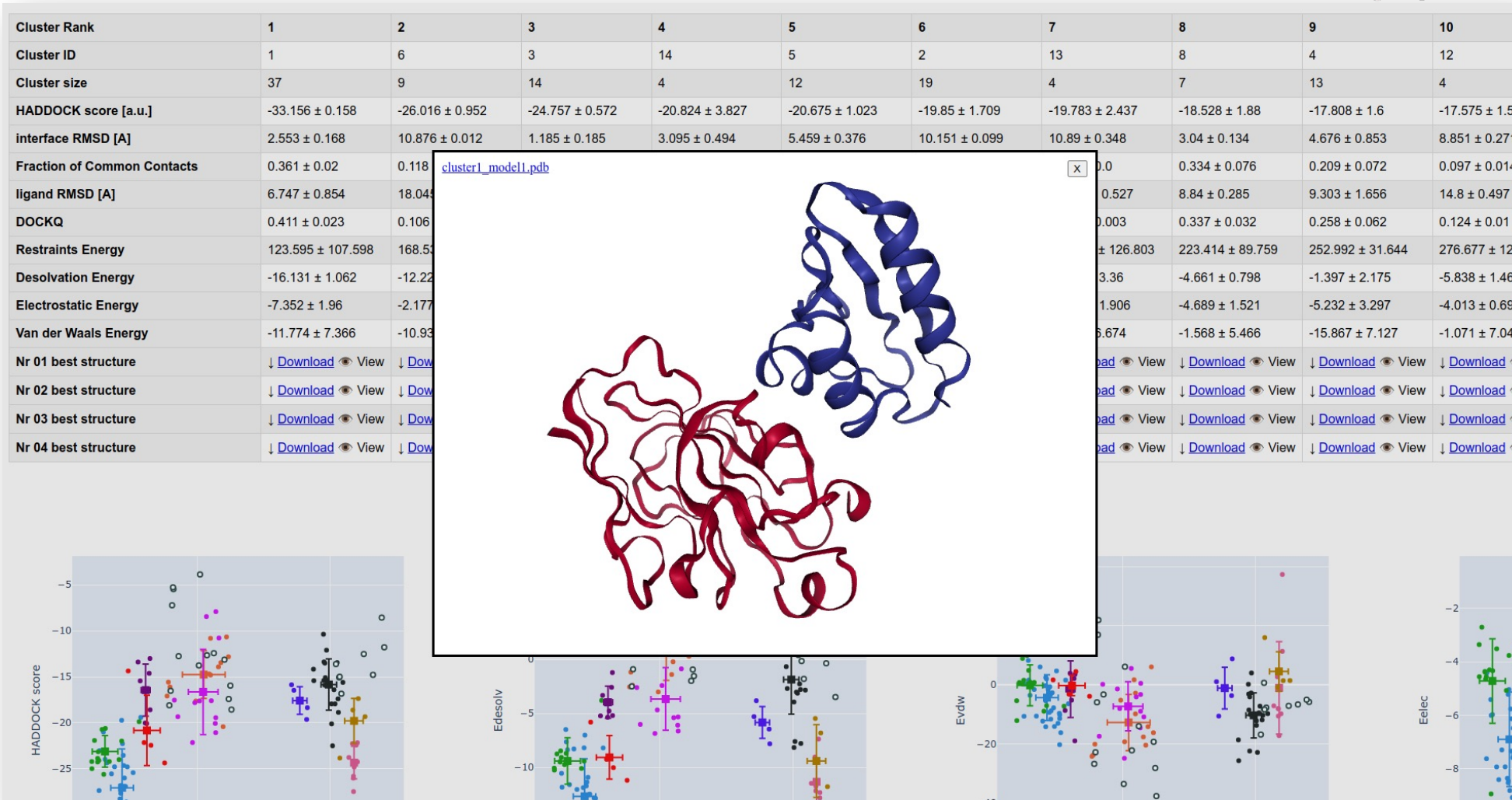
```
# =====  
# Protein-protein docking example with  
# NMR-derived ambiguous interaction restraints  
  
# directory in which the scoring will be done  
run_dir = "run1"  
  
# molecules to be docked  
molecules = [  
    "data/e2aP_1F3G.pdb",  
    "data/hpr_ensemble.pdb"  
]  
  
# =====  
[topoaa]  
  
[rigidbody]  
ambig_fname = "data/e2a-hpr_air.tbl"  
sampling = 1000  
  
[seletop]  
select = 200  
  
[flexref]  
ambig_fname = "data/e2a-hpr_air.tbl"  
  
[emref]  
ambig_fname = "data/e2a-hpr_air.tbl"  
  
[caprieval]  
  
[clustfcc]  
  
[seletopclusts]  
  
[caprieval]  
# =====
```

HADDOCK3 analyse report

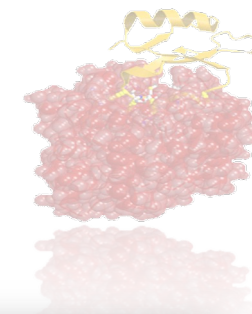


HADDOCK3 can automatically generate interactive analysis plots (similar to what the 2.4 server gives)

Will form the basis for the analysis interface



HADDOCK3 now available on ICE-CLOUD!



- As stand alone CLI
- As Jupyter notebook
- As webapp

Haddock3 Scenario Builder Upload Manage About Help Login

Scenarios
Explore and run predefined workflows.

Build
Use the workflow builder to create and submit a job.

Upload
Upload a workflow and submit as job.

Manage
Explore and analyse the results of completed jobs.

HADDOCK3 container
High-Ambiguity Driven Docking

Haddock3 Notebook notebook
HADDOCK, standing for High Ambiguity Driven protein-protein DOCKing, is a widely used computational tool for the...

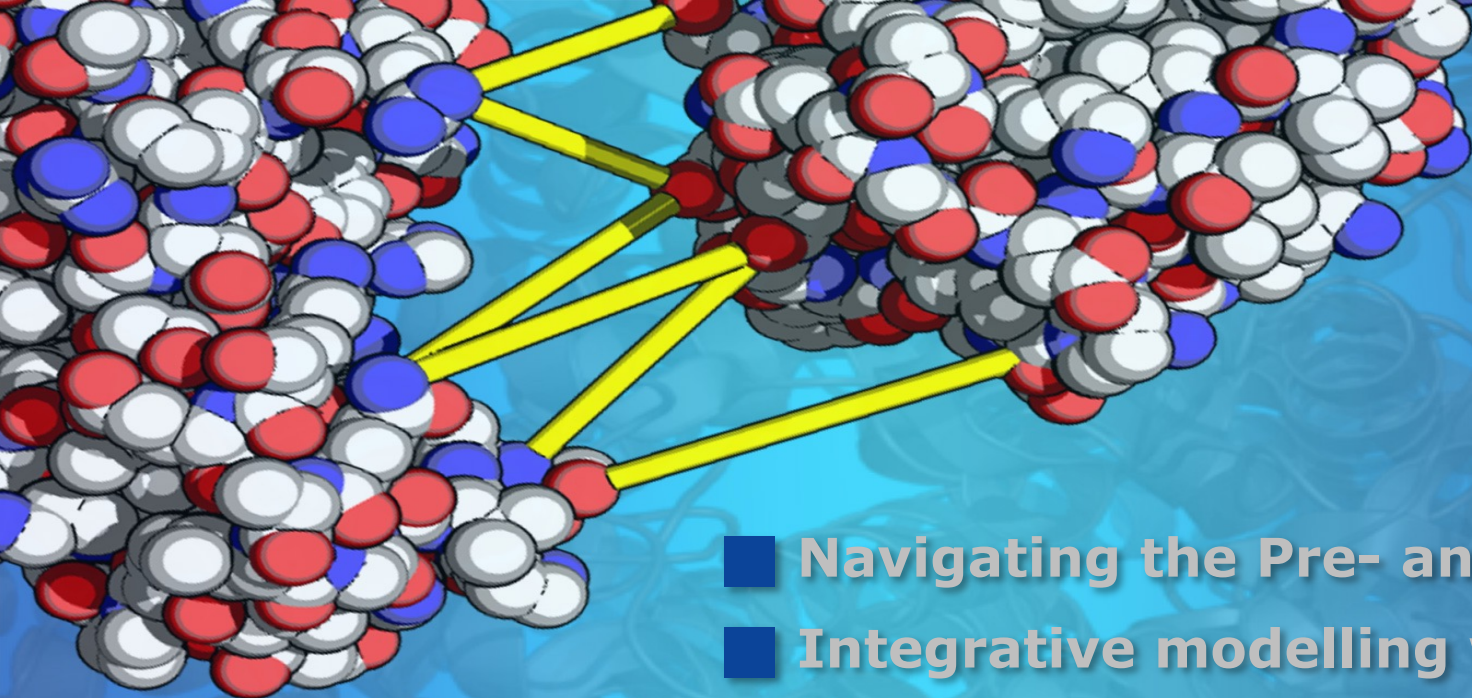
icecloud.in



[Faculty of Science
Chemistry]



Universiteit Utrecht

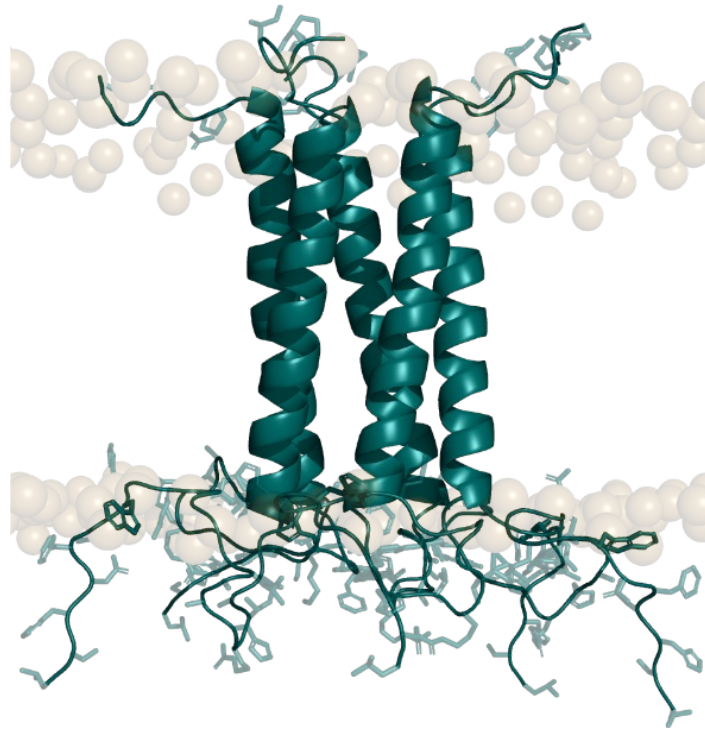


Overview

- Navigating the Pre- and Post-AlphaFold Divide
- Integrative modelling with HADDOCK
- Introducing HADDOCK3
- Application examples
 - An NMR-based application – where AlphaFold fails
 - AI-based antibody-antigen modelling
 - Cyclic peptides
- Conclusions & perspectives

```
111000100101010110101010
100011010101000101111010110101
10100001000111101010101110000110101
0001010101011110010100110010101010100
010101111110101010000101010100000101010100
0010110101011101010001111010101010001010010101011111
000111111010101011001100010111000101011010100010101000101
```

Mumps Virus Small Hydrophobic (SH) Protein (when alphafold fails...)



Slides adapted from
Birthe Kragelund

University of Copenhagen, DK



Kira Devantier

ScienceAdvances













[Current Issue](#) [First release papers](#) [Archive](#) [About](#)

HOME > SCIENCE ADVANCES > VOL. 11, NO. 23 > THE SH PROTEIN OF MUMPS VIRUS IS A DRUGGABLE PENTAMERIC VIROPORIN

RESEARCH ARTICLE | STRUCTURAL BIOLOGY



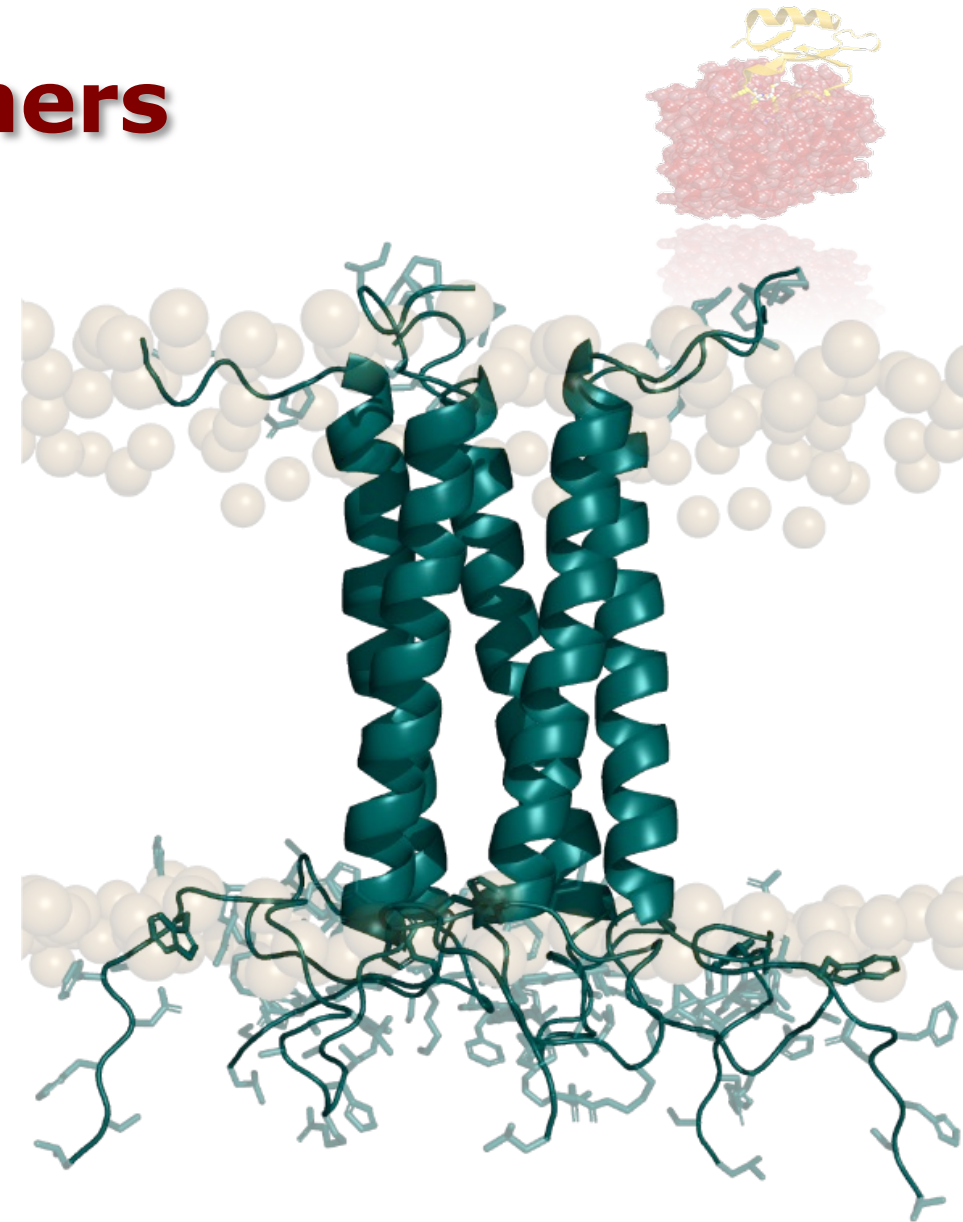
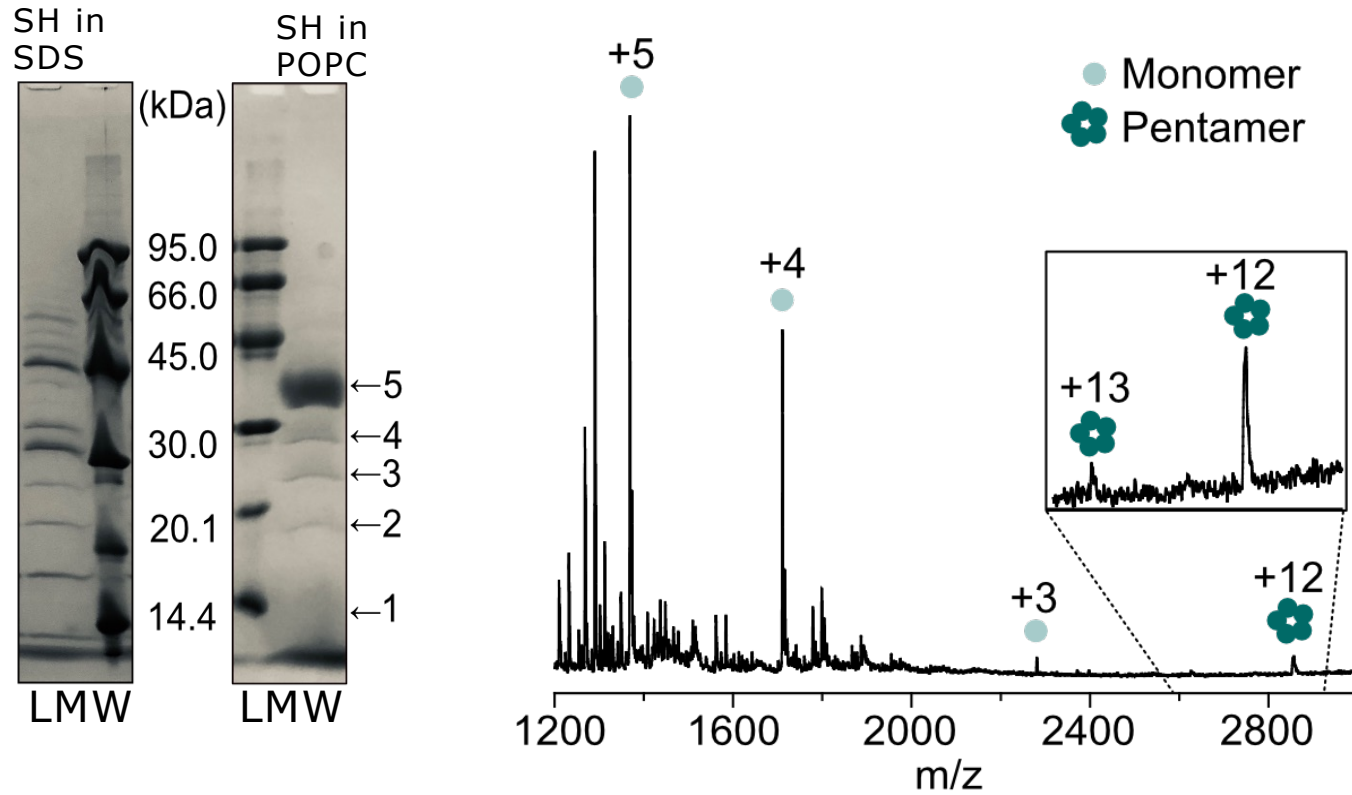
The SH protein of mumps virus is a druggable pentameric viroporin

KIRA DEVANTIER , TRINE L. TOFT-BERTELSEN, ANDREAS PRESTEL , VIKTORIA M. S. KJÆR , CAGLA SAHIN , MARCO GIULINI , STAVROULA LOUKA, KATJA SPIESS , ASMITA MANANDHAR, KATRINE QVORTRUP , TROND ULVEN , BO H. BENTZEN, ALEXANDRE MJJ BONVIN , NANNA MACAULAY , BIRTHE B. KRAGELUND , AND METTE M. ROSENKILDE  [fewer](#) [Authors Info & Affiliations](#)

SCIENCE ADVANCES • 6 Jun 2025 • Vol 11, Issue 23 • DOI: 10.1126/sciadv.ads3071

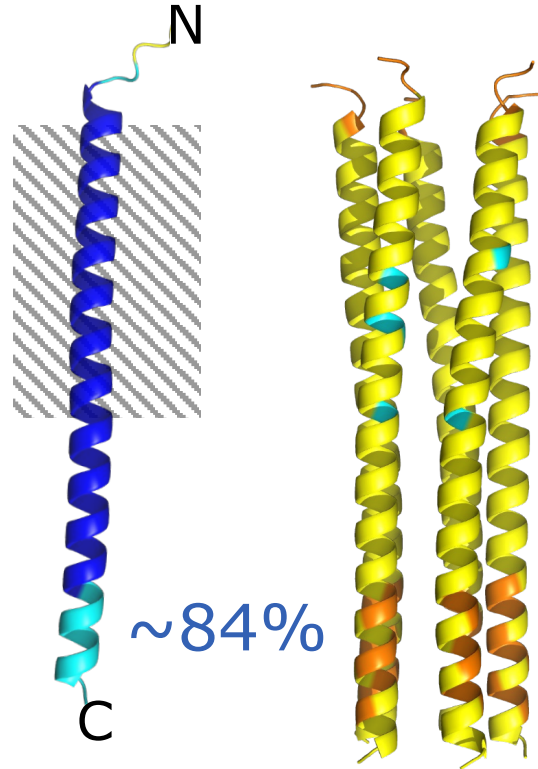
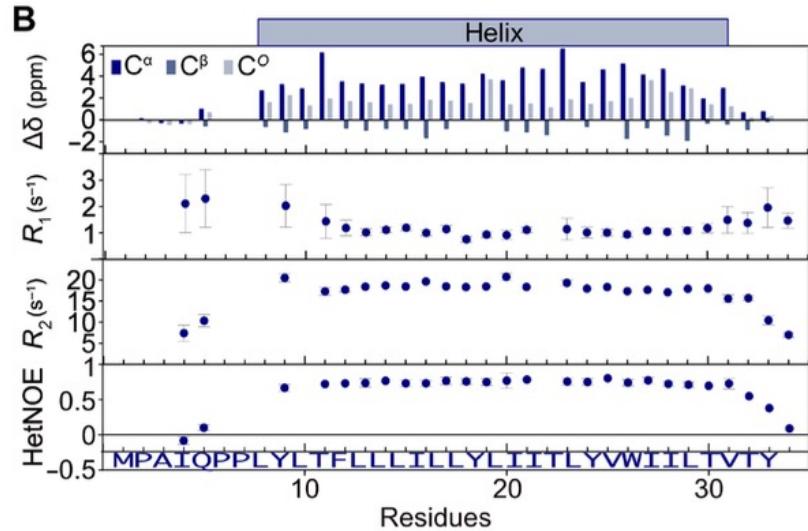
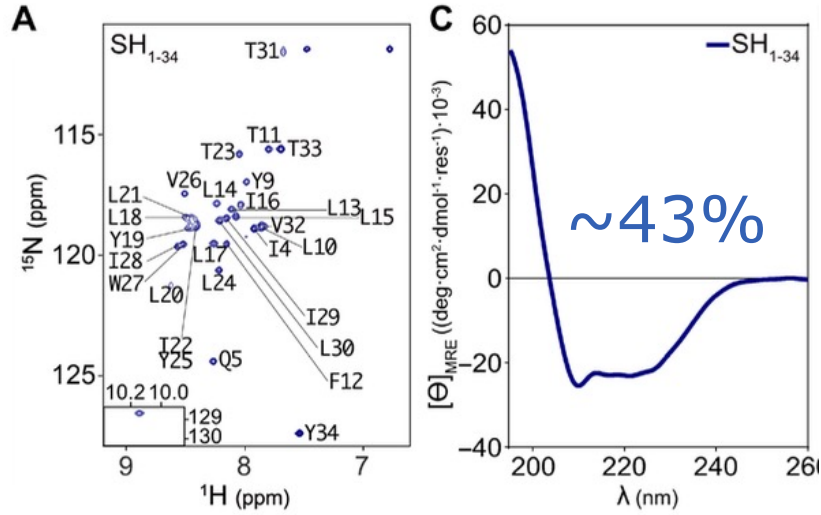


SH Forms Pentamers

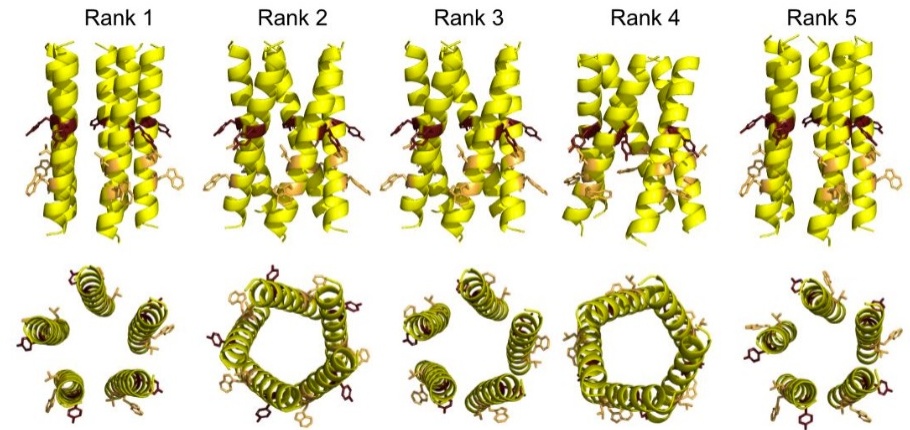


SH Forms Helical Structure

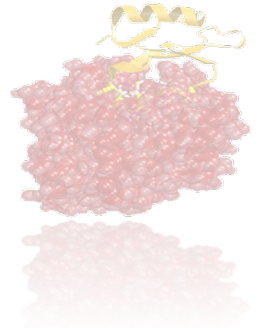
in DHPC micelles



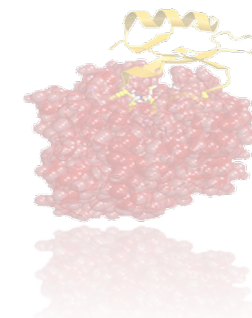
AlphaFold3 model confidence (pLDDT)	
very high (>90)	low (70-50)
confident (90-70)	very low (<50)



- **AF2/3 modelling overestimates the helix content and places polar residues facing the membrane** (adding lipids in AF3 does not help)

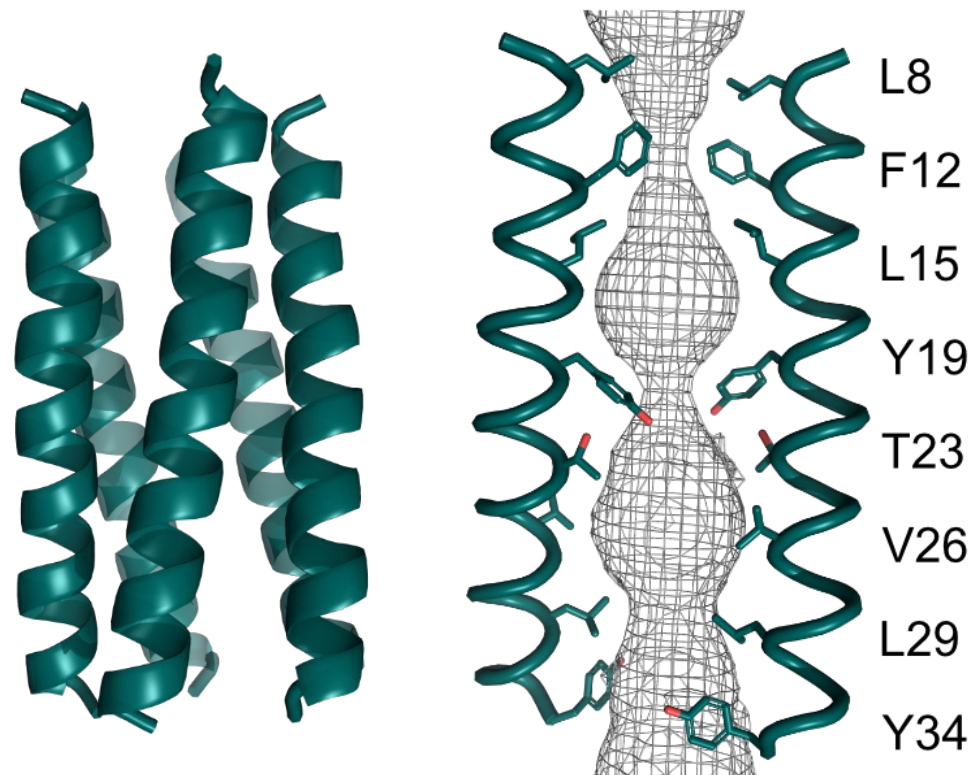
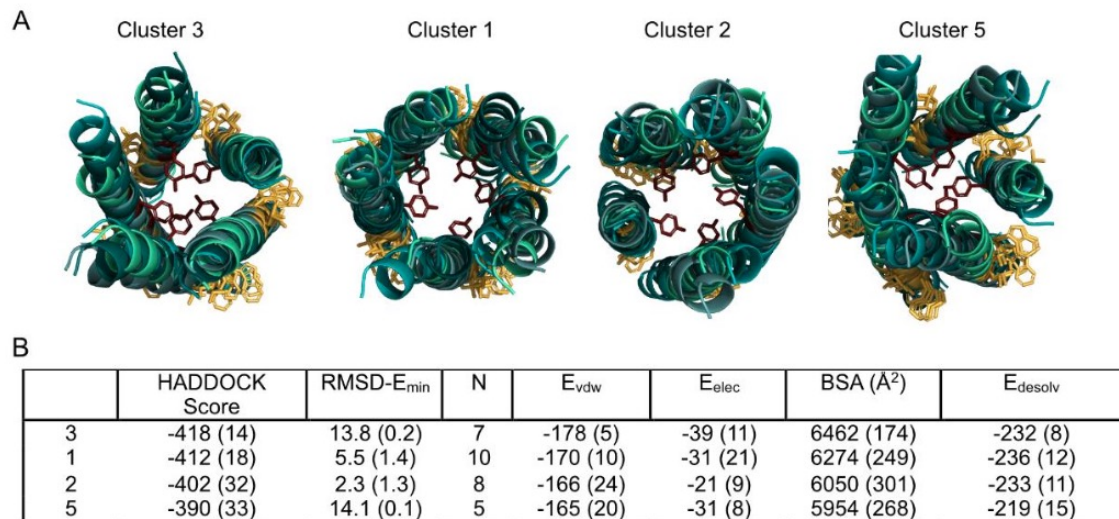


HADDOCK modelling

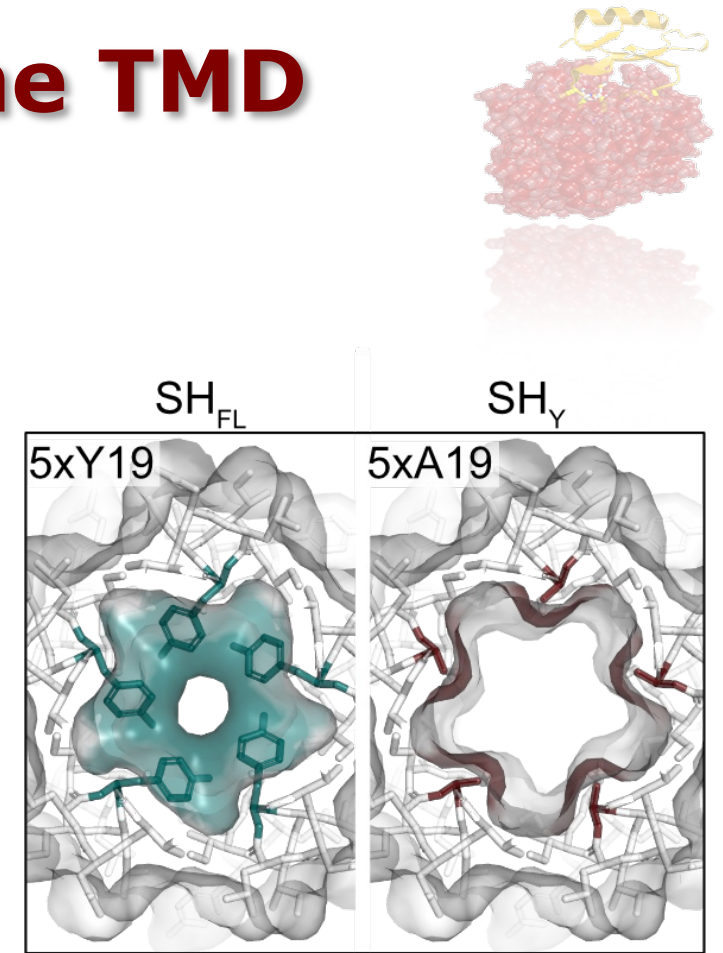
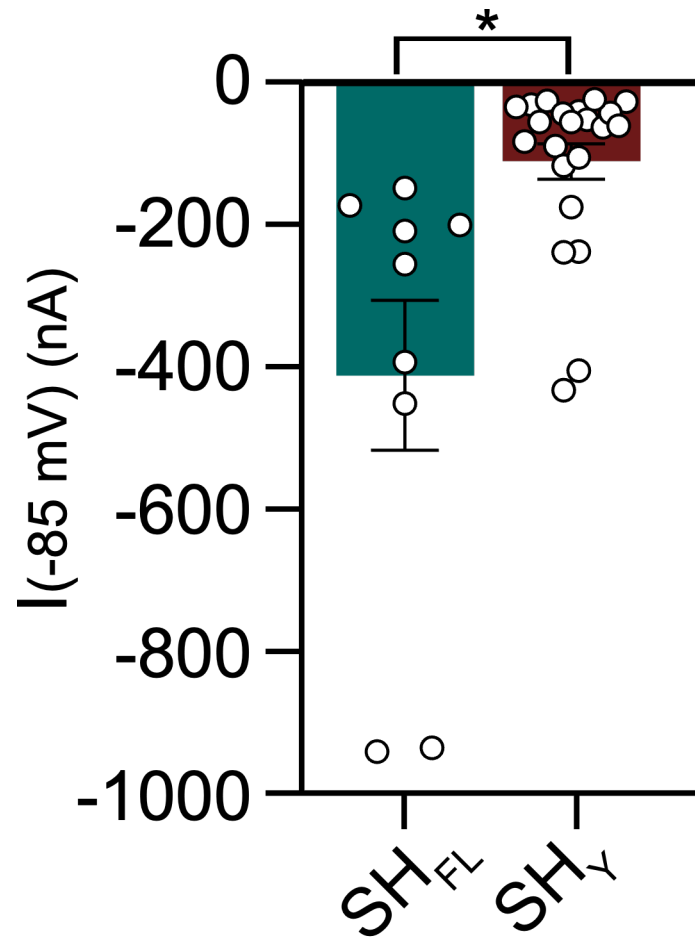
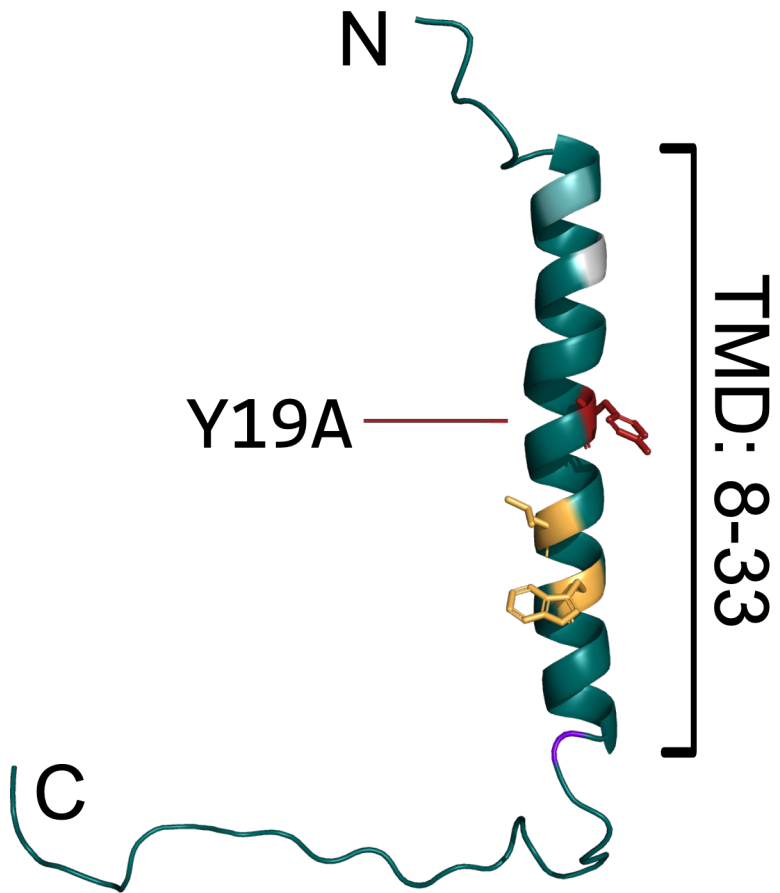


- Ab-initio modelling with C5 symmetry restraints
- Increased sampling

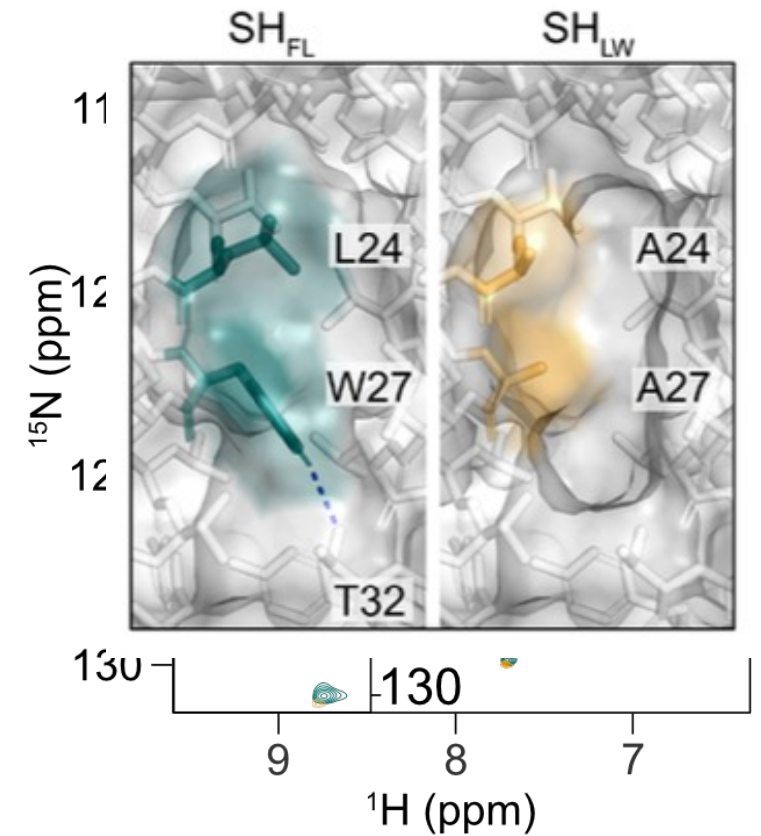
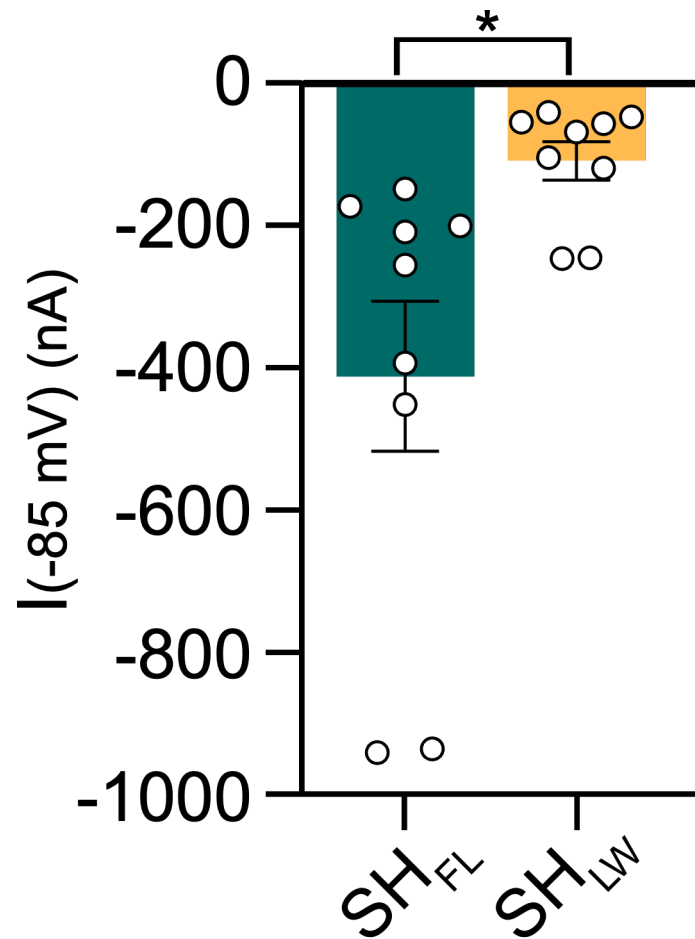
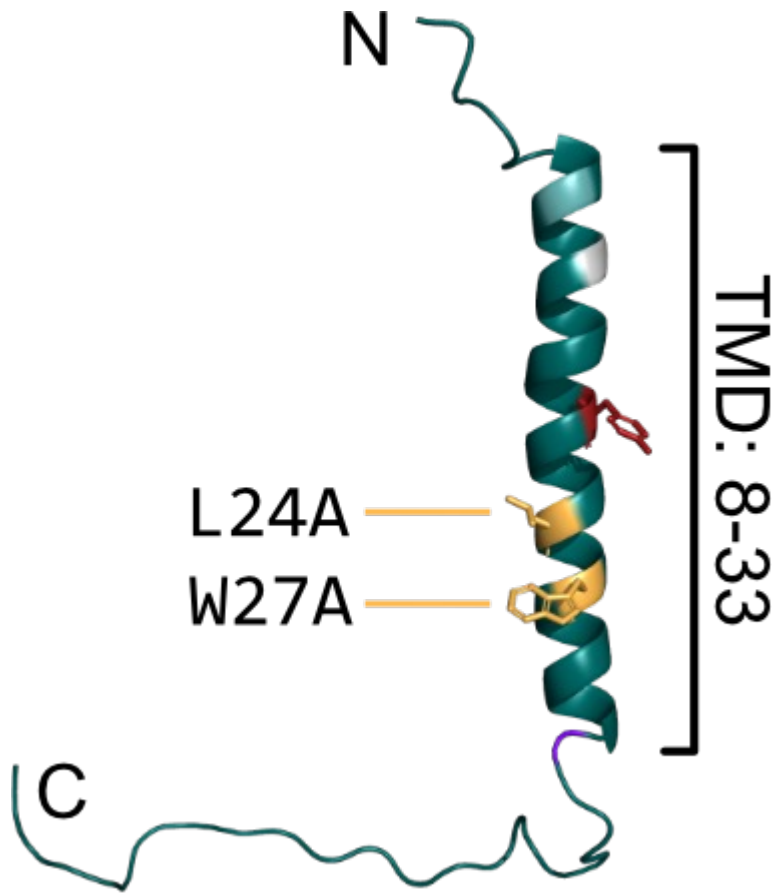
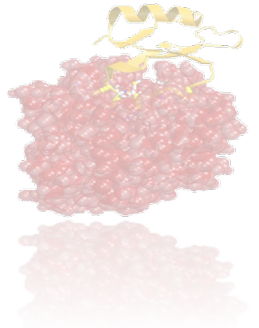
SH forms highly hydrophobic pores which conduct chloride



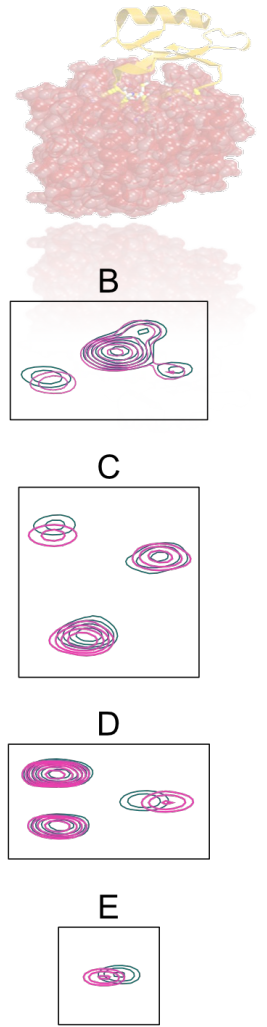
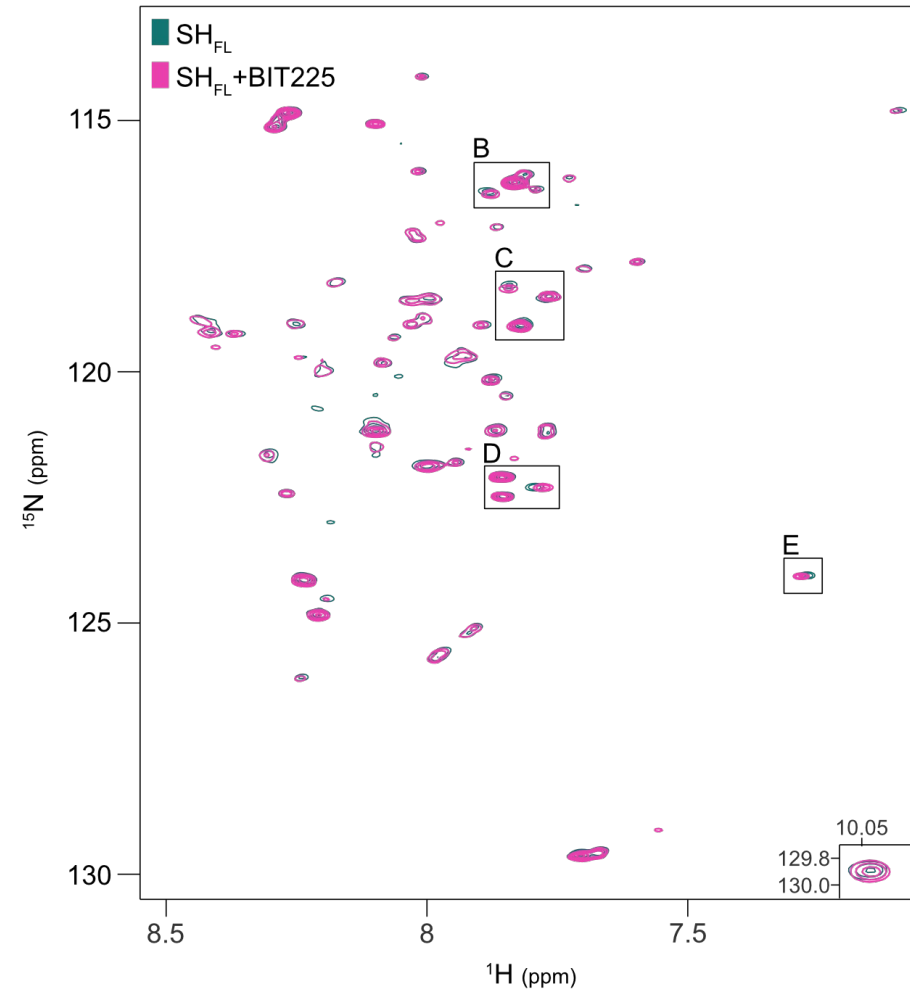
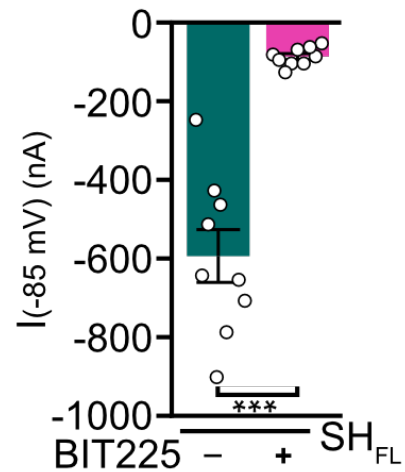
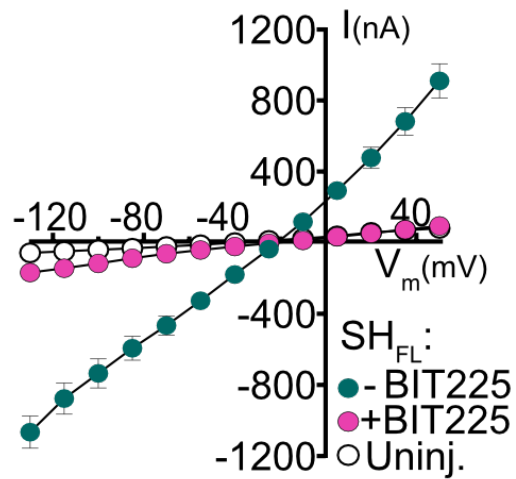
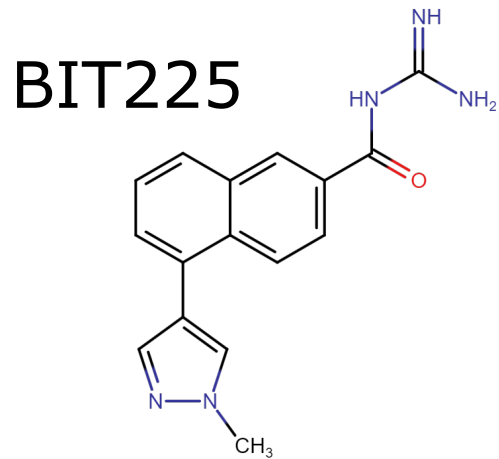
Important Residues in the TMD



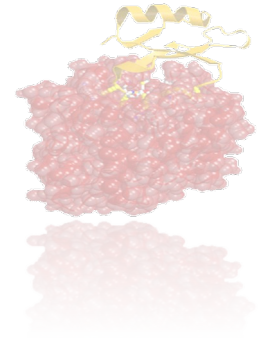
Important Residues in the TMD



SH is druggable



Conclusions



- SH modelling with Alphafold fails to produce reasonable models
- NMR identifies the transmembrane helical region and disordered termini
- HADDOCK modelling of the pentamer points to key residues for both chloride transport (Y19) and structural integrity of the pore (L24, W27) (confirmed by NMR)
- SH qualifies as druggable viroporin



ScienceAdvances

Current Issue First release papers Archive About ▾

HOME > SCIENCE ADVANCES > VOL. 11, NO. 23 > THE SH PROTEIN OF MUMPS VIRUS IS A DRUGGABLE PENTAMERIC VIROPORIN

RESEARCH ARTICLE | STRUCTURAL BIOLOGY

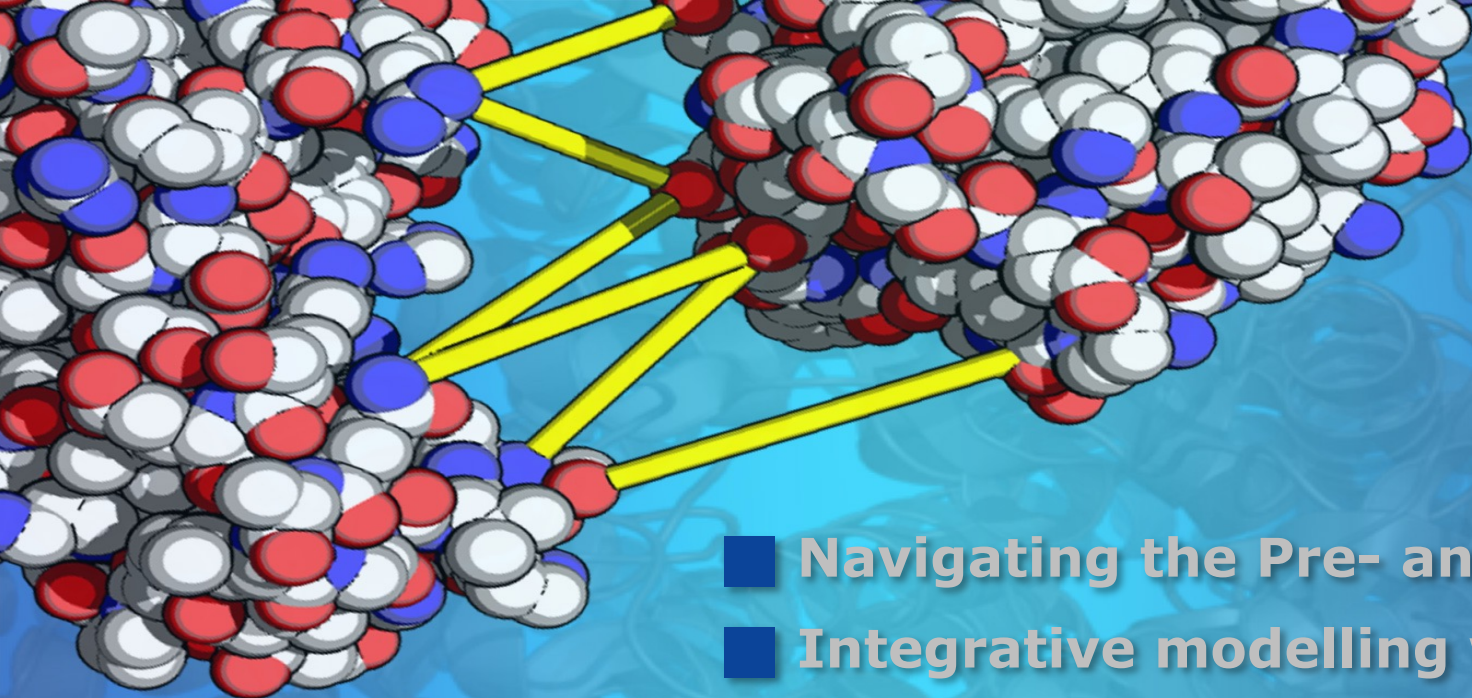
f X t in g s e

The SH protein of mumps virus is a druggable pentameric viroporin

KIRA DEVANTIER , TRINE L. TOFT-BERTELSEN, ANDREAS PRESTEL , VIKTORIA M. S. KJÆR , ÇAGLA SAHIN , MARCO GIULINI , STAVROULA LOUKA, KATJA SPIESS , ASMITA MANANDHAR, KATRINE QVORTRUP , TROND ULVEN , BO H. BENTZEN, ALEXANDRE MJJ BONVIN , NANNA MACAULAY , BIRTHE B. KRAGELUND , AND METTE M. ROSENKILDE  fewer [Authors Info & Affiliations](#)

SCIENCE ADVANCES • 6 Jun 2025 • Vol 11, Issue 23 • DOI: 10.1126/sciadv.ads3071



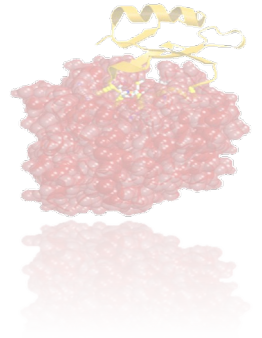


Overview

- Navigating the Pre- and Post-AlphaFold Divide
- Integrative modelling with HADDOCK
- Introducing HADDOCK3
- Application examples
 - An NMR-based application – where AlphaFold fails
 - AI-based antibody-antigen modelling
 - Cyclic peptides
- Conclusions & perspectives

```
111000100101010110101010
100011010101000101111010110101
10100001000111101010101110000110101
0001010101011110010100110010101010100
0101011111110101010000101010100000101010100
001011010101110101000111101010101000101001010101111
0001111111010101011001100010111000101011010100010101000101
```

An AI revolution in antibody modelling

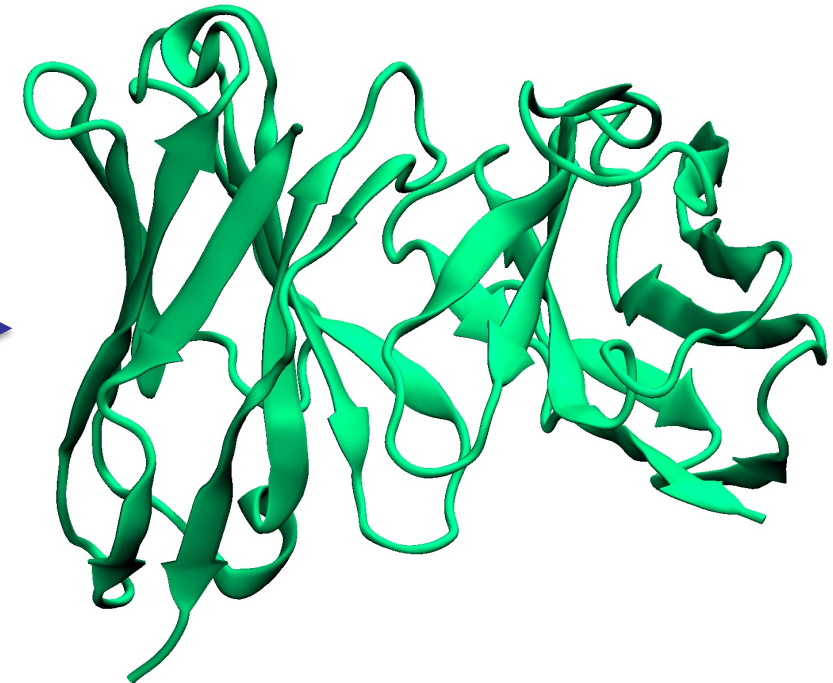
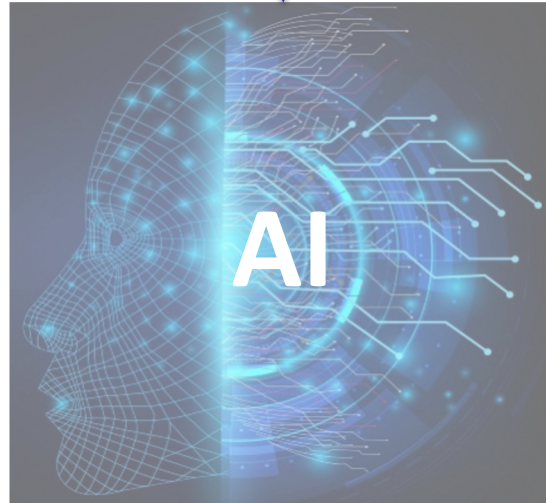


> Heavy chain

```
EVKLVESGGGLVQPGGSLRLSCAASGF  
AFSSYDMSWVRQAPGKRLEWVATISGG  
GRYTYYPDTVKGRFTISRDNKNSHYL  
QMNSLRAEDTAVYFCASPYGGYFDVWG  
QGTLVTVSS
```

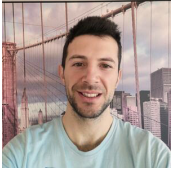
> Light chain

```
EIVLTQSPATLSLSPGERATLSCRASQ  
SISNFLHWYQQKPGQAPRLLIKYASQS  
ISGIPARFSGSGSGTDFTLTISSLEPE  
DFAVYFCQQSNSWPHTFGQGTKVEIK
```



Fast (few seconds per antibody in some cases) generation of antibody structures from sequence





Marco Giulini



Constantin Schneider

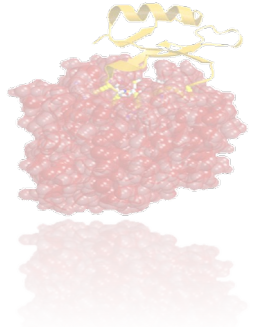


Nikita Desai



Daniel Cutting

Putting the predictors to the test



- **No structures present in the training set**
- **No Immunoglobulin-binding proteins**
- **No redundancy**
- **79 unique antibody-antigen complexes**

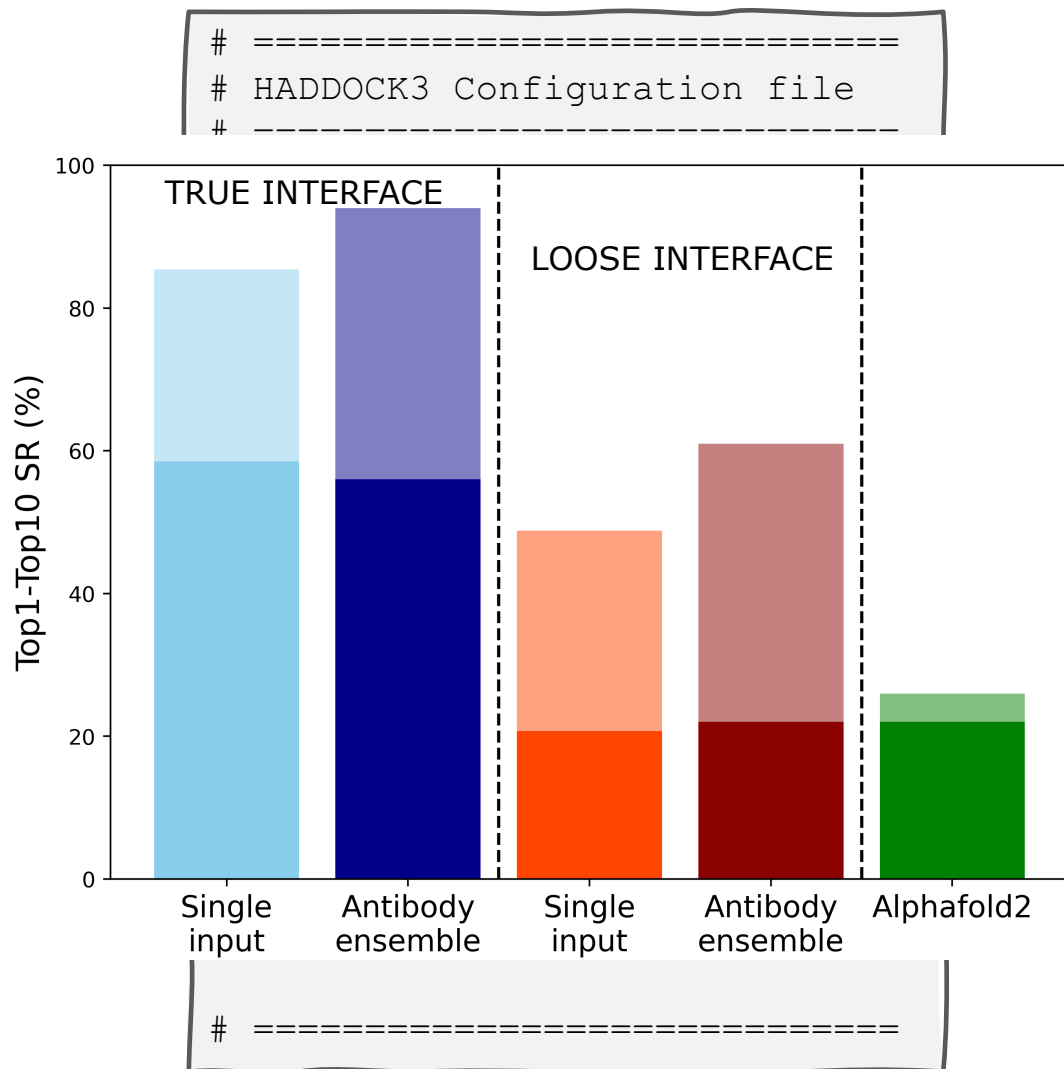
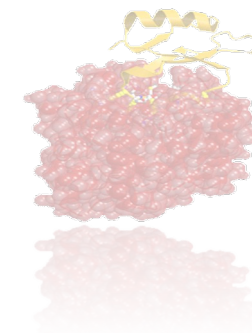
- **AlphaFold, Immune Builder & IgFold put to the test**

Schneider *et al.* (2022) *Nucleic Acids Res.* 50(D1):D1368-D1372

Dunbar *et al.* (2014) *Nucleic Acids Res.* 42(D1):D1140-D1146

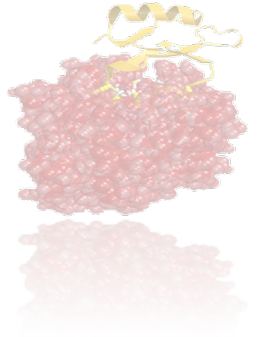


Antibody-antigen complex prediction

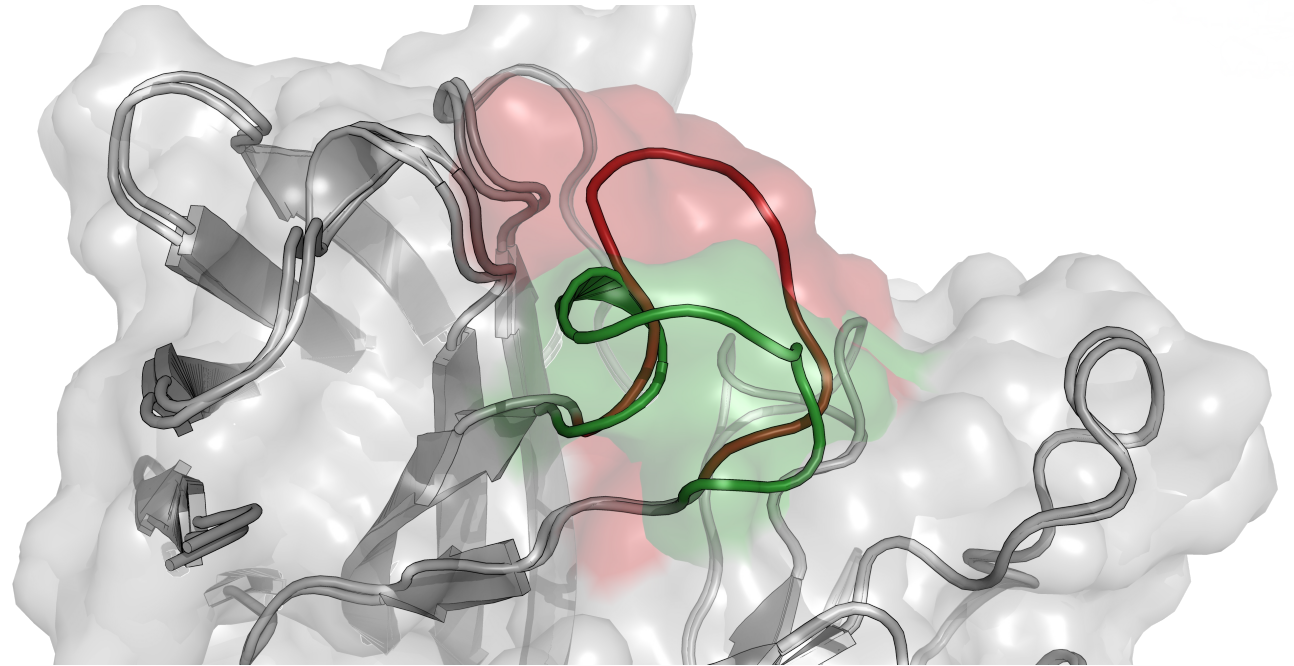


- **Fast HADDOCK3 protocol with limited sampling**
- **No coevolution** between antibody and antigen: **low Alphafold2 performance!**

H3 loop prediction



- **H3 is crucial** for the antigen recognition
- **Huge sequence and structure variability** due to its unique location at the intersection of the VDJ gene segments
- Absence of experimentally solved structural relatives



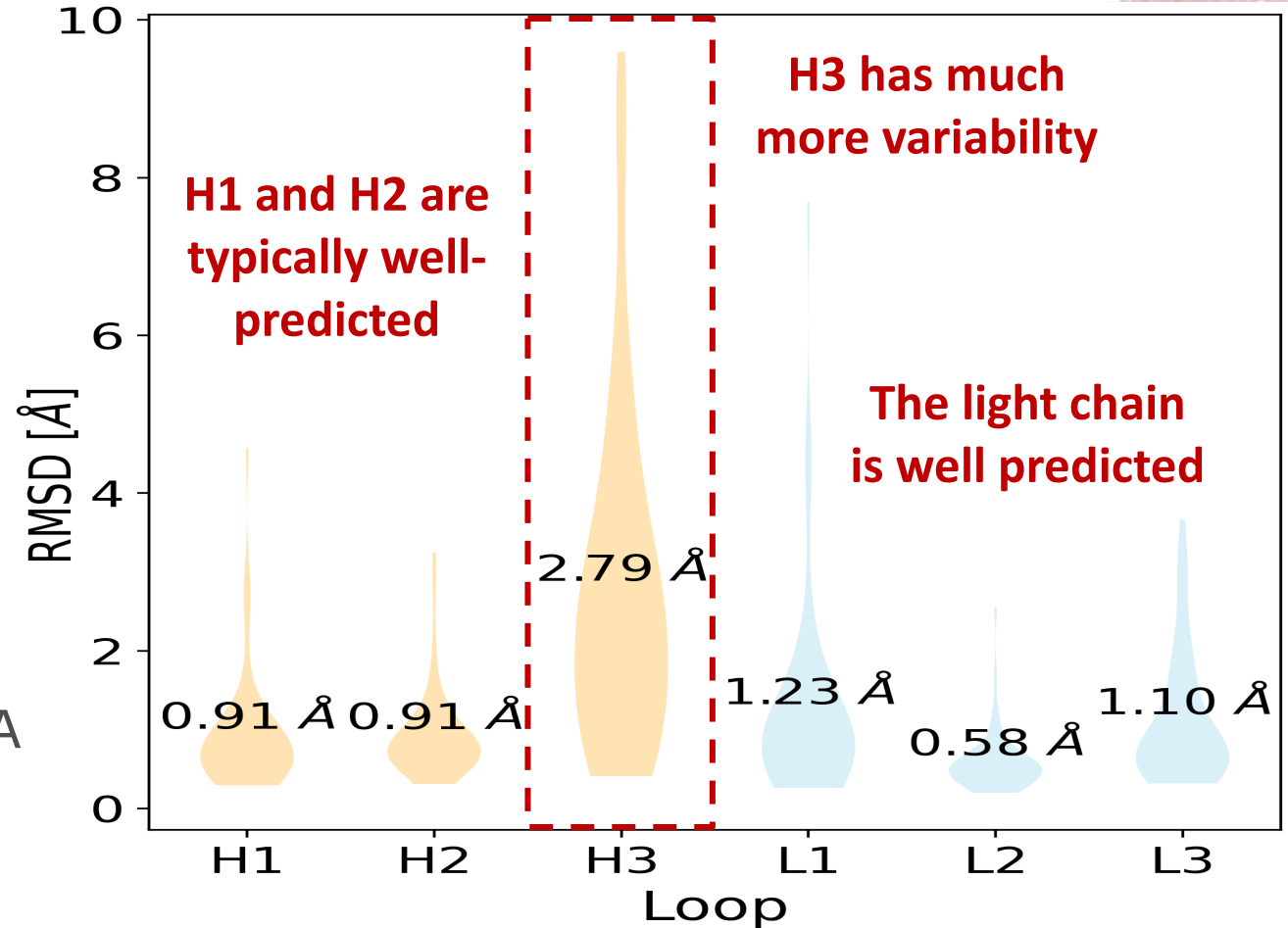
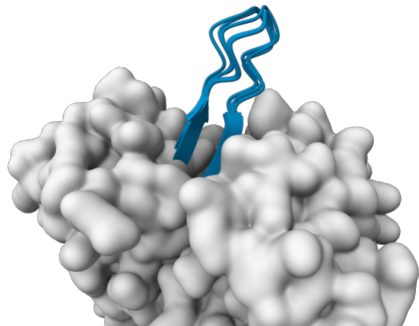
Weitzner, Dunbrack, Gray The origin of CDR H3 structural diversity, Structure 2015
Regep, ..., Deane The H3 loop of antibodies shows unique structural characteristics, Proteins 2017



Performance of AF2 on antibodies

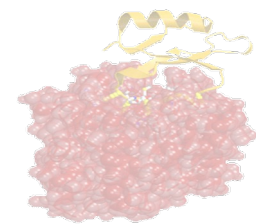
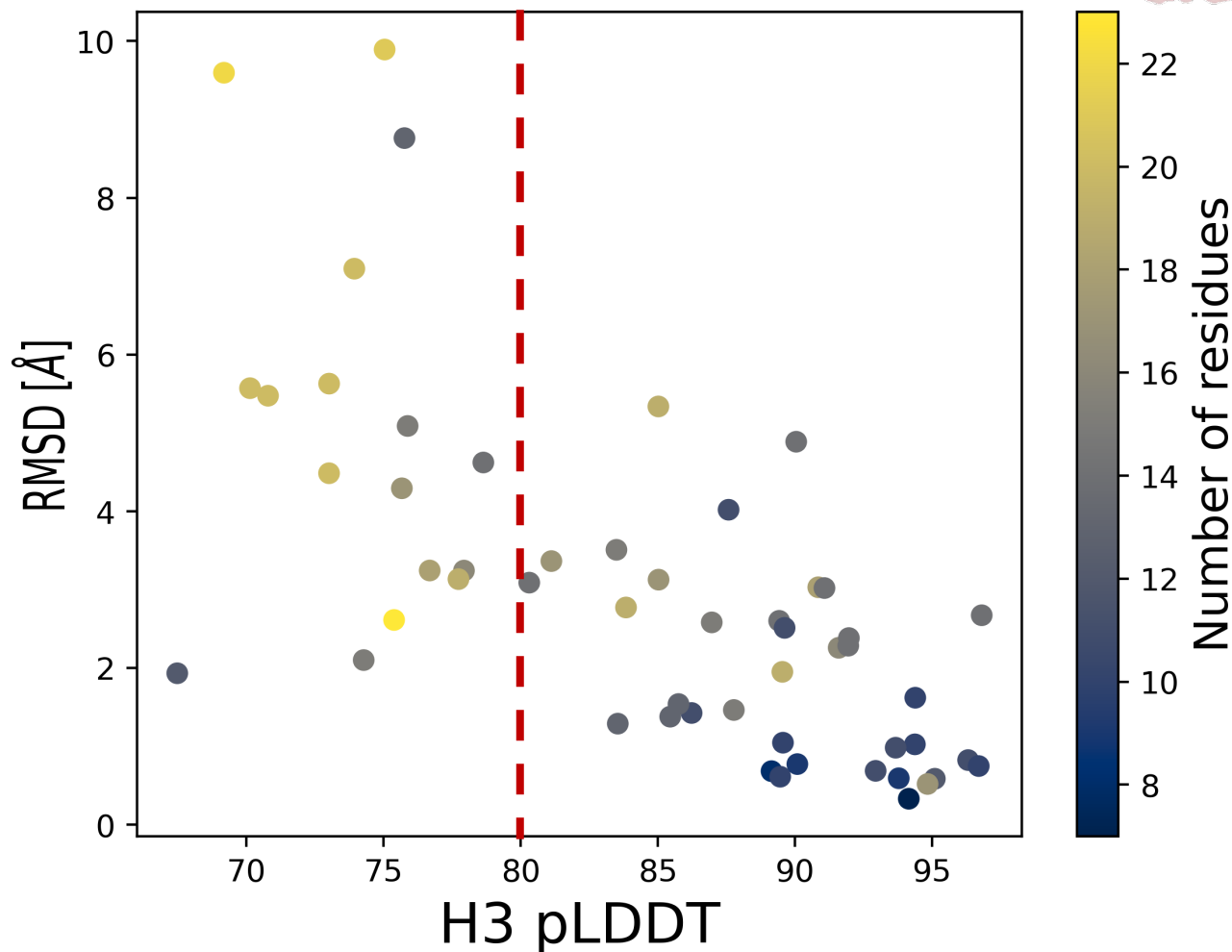


- **Dataset: 54 unique antibodies from SAbDab**
 - No sequence redundancy
 - No structures present in AF training set
- **Generate AF2 25 models each**
- **AF sampling** (multiple seeds, MSA subsampling/clustering) **is not stellar**

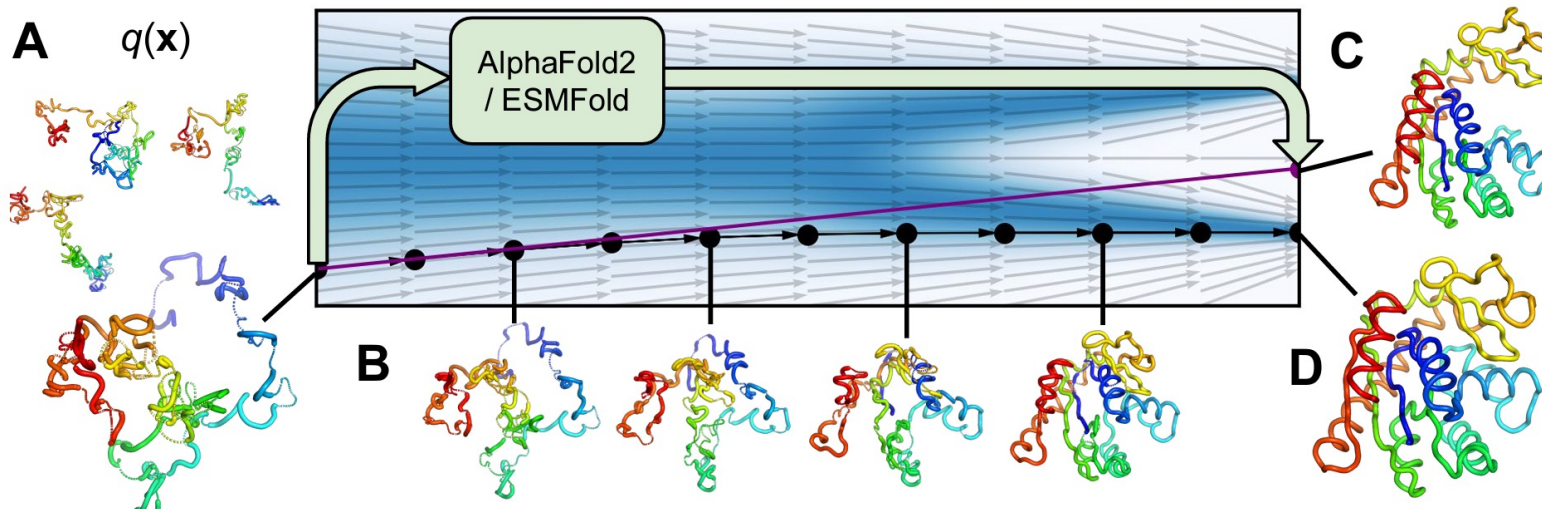
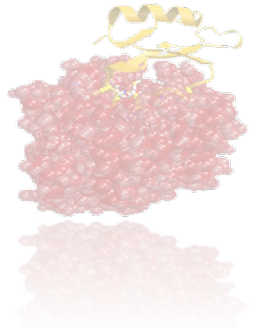


H3 pLDDT as an indicator for RMSD

- Pearson correlation = **-0.73**
- **80** is a good threshold to obtain reasonable models with AF2
- With this indicator, we can focus on antibodies with low AF2 confidence

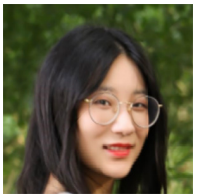


Alphaflow: Generative modelling on top of Alphafold



Take AlphaFold and repurpose it into a **generative modelling framework** to obtain a **highly diverse output**

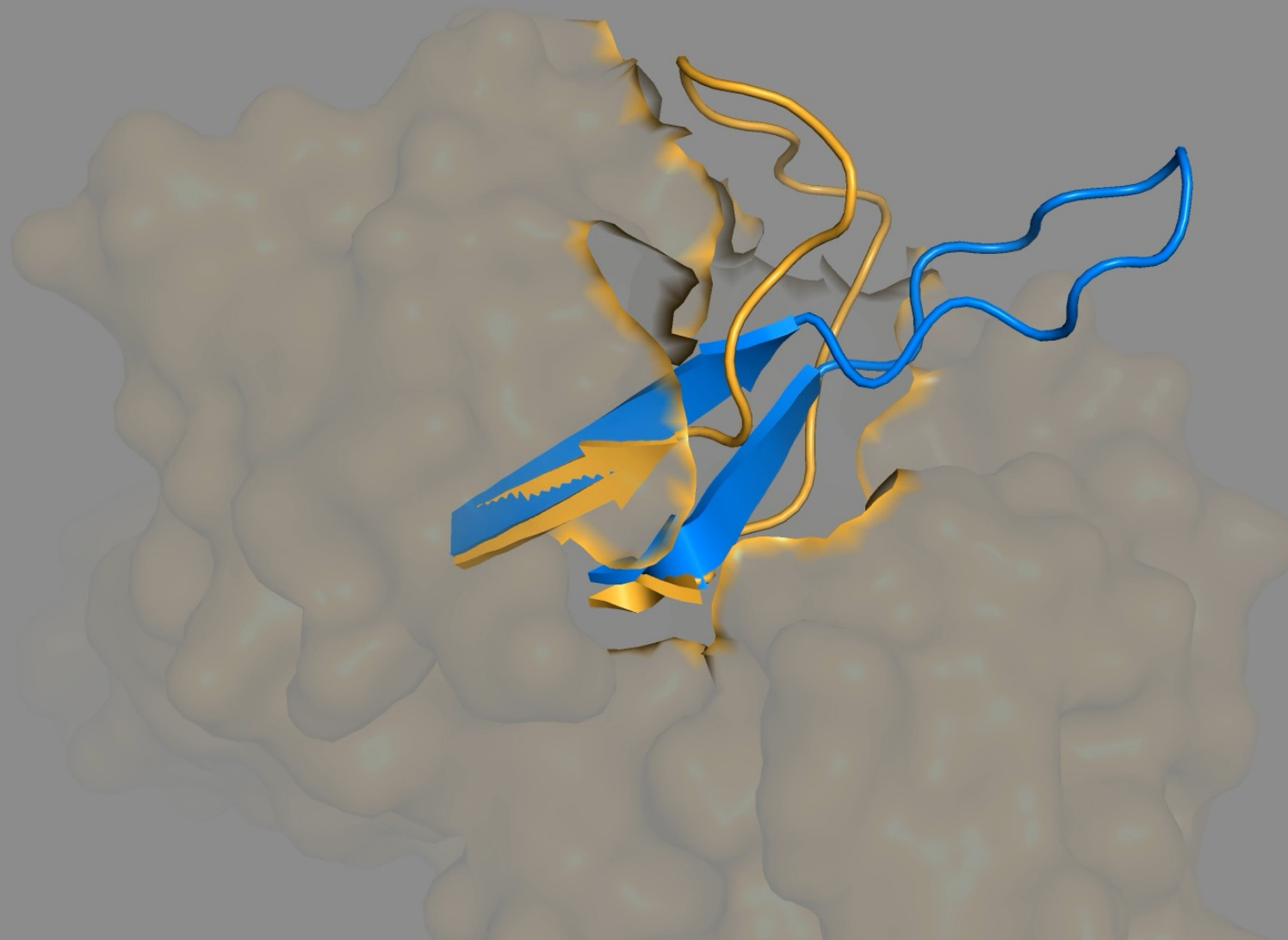
Flow matching framework preserving AlphaFold architecture



Xiaotong Xu

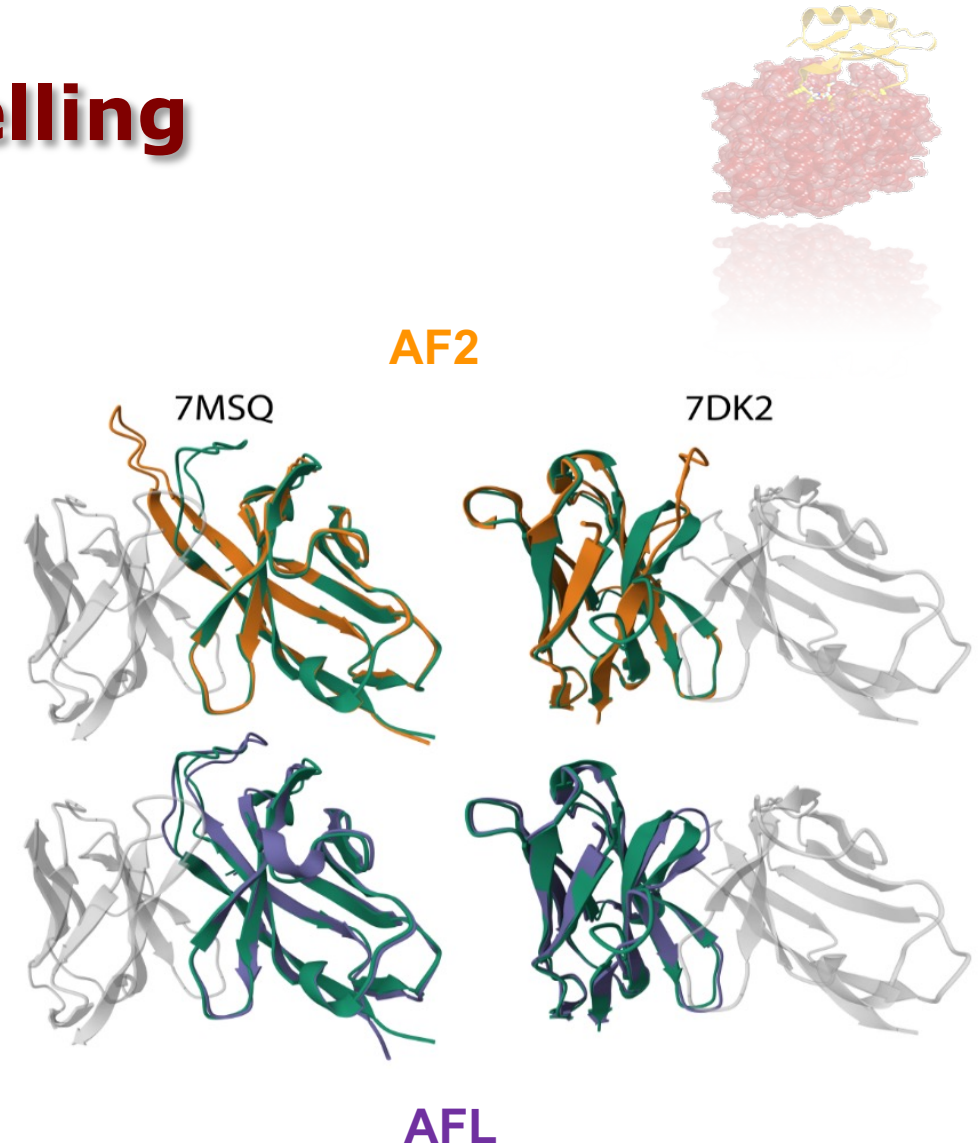
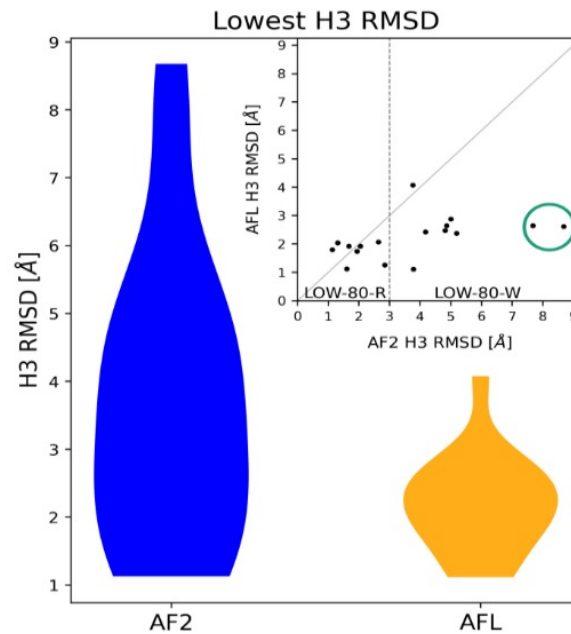
AlphaFold meets flow matching for generating protein ensembles
B Jing, B Berger, T Jaakkola - 2024



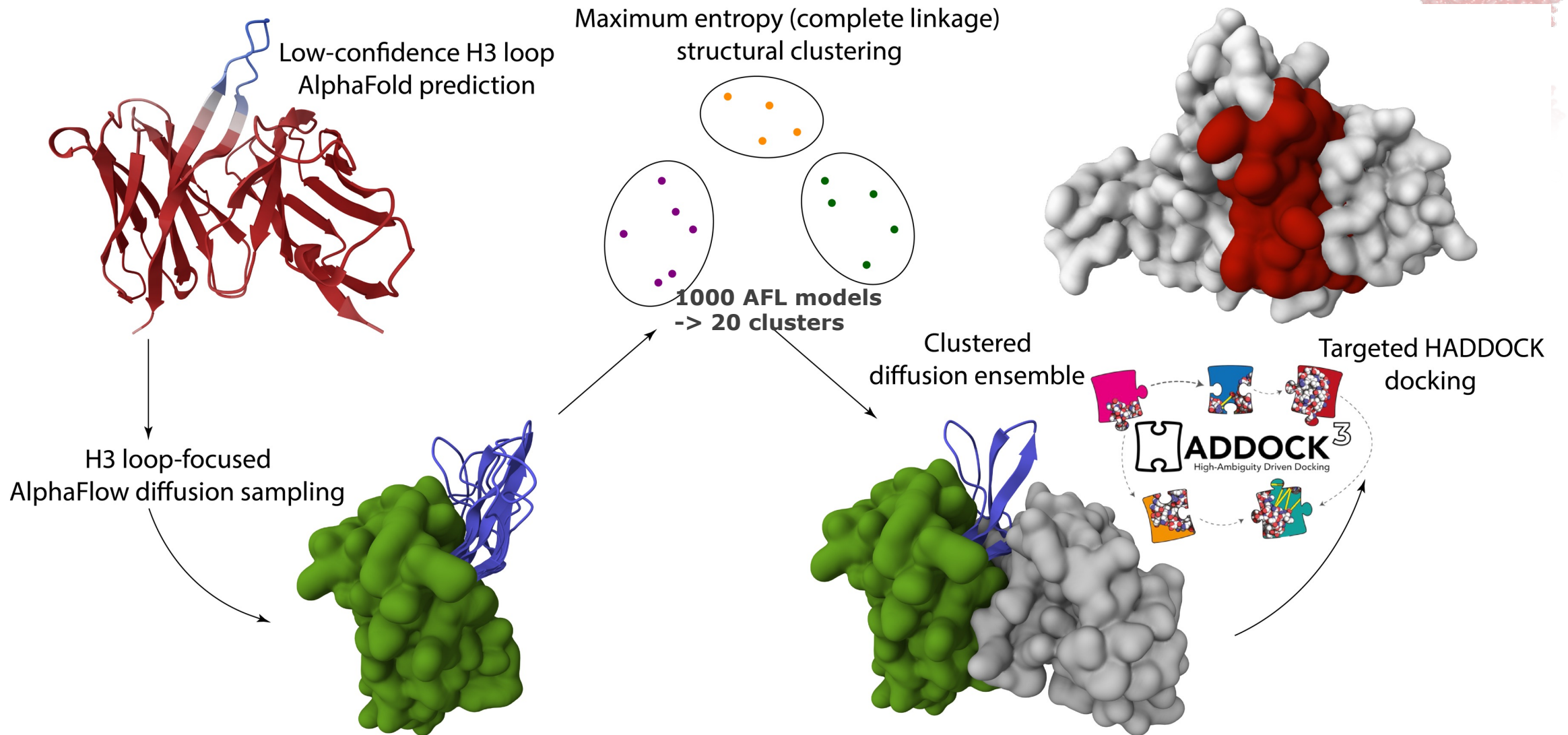


AlphaFlow modelling

- 1000 models for each antibody heavy chain
- AlphaFlow always recovers some models with $H3\text{-RMSD} < 3\text{\AA}$
- Fishing the good models out of the sample is difficult (pLDDT is not predictive here)



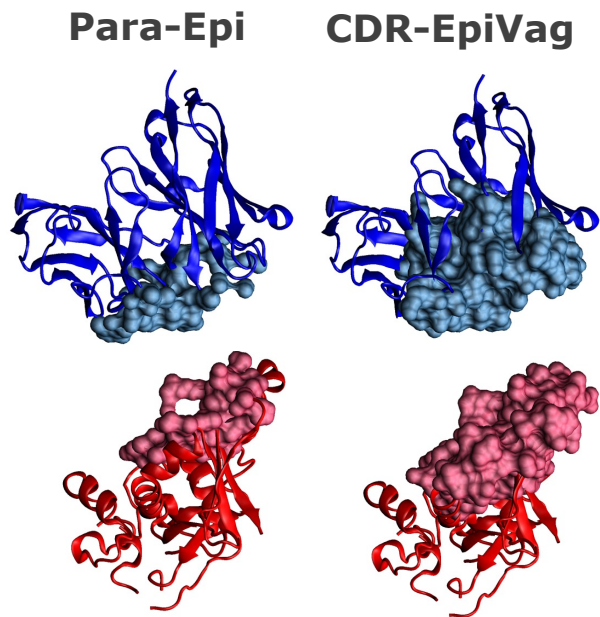
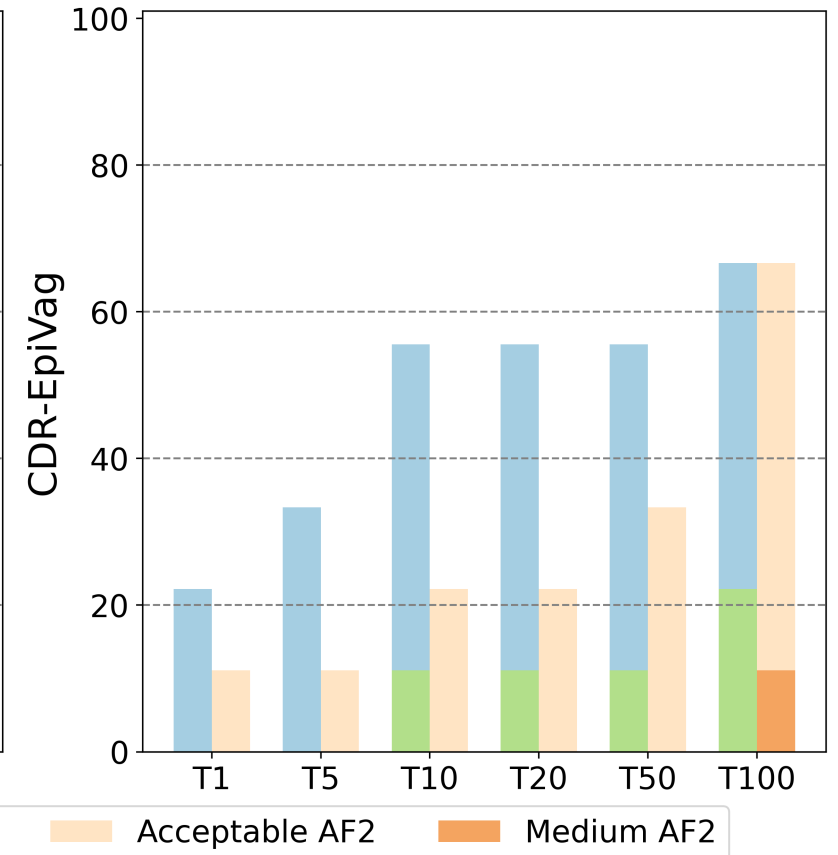
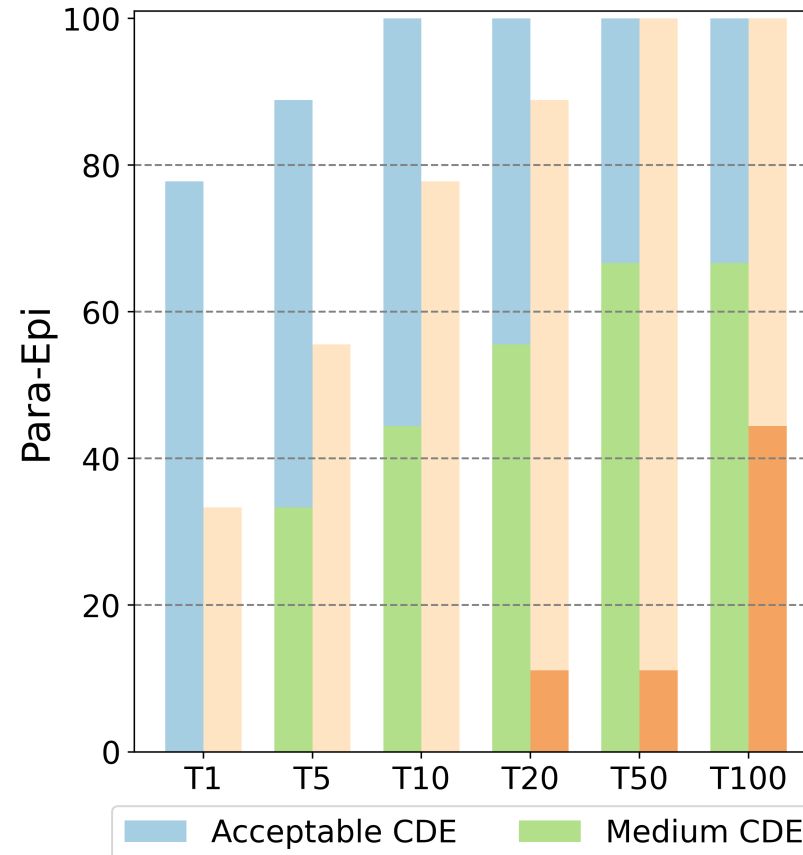
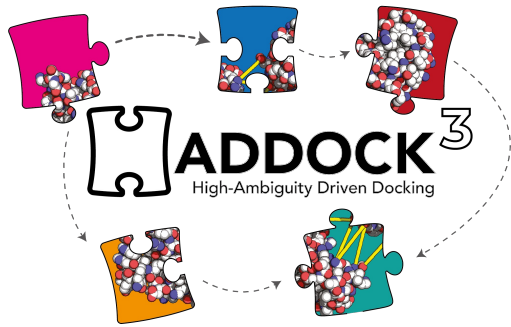
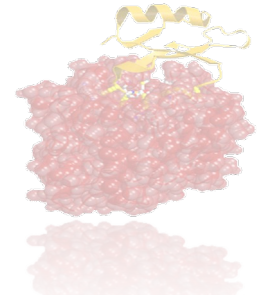
Modelling workflow



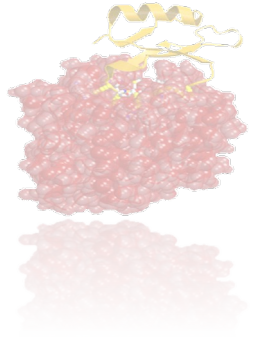
Xu, Giulini & Bonvin *Bioinfo. Adv.* 2025



Improved docking performance



AlphaFold3: did it solve everything?



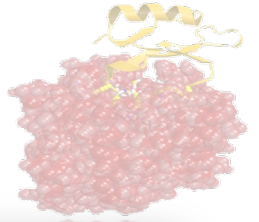
- **AF3 improved the success of antibody-antigen complex modelling** (at the cost of massive sampling)
- **MSA's importance is de-emphasized + diffusion module guarantees more diversity in the sampled structures**
- **On our small dataset of difficult cases, AF3 helped only in 1 out of 9 cases (reaching an H3-RMSD = 2.1Å)**
- **Neither AF2 or AF3 were able to provide any acceptable solutions within the top10**



J Abramson et al Accurate structure prediction of biomolecular interactions with AlphaFold 3, Nature 2024




Conclusions



- AlphaFold revolutionized the field of antibody/nanobody modelling, but there are still challenges in the complex and H3 loop prediction
- Exploring more widely the conformational space of H3 loops is beneficial for difficult cases
- Similar observation for nanobodies (with framework also playing a role)
- Despite the AI era, there is still (for now?) a need for more “classical” integrative modelling approaches

JOURNAL ARTICLE

Towards the accurate modelling of antibody-antigen complexes from sequence using machine learning and information-driven docking

Marco Giulini, Constantin Schneider, Daniel Cutting, Nikita Desai, Charlotte M Deane, Alexandre M J J Bonvin  [Author Notes](#)

Bioinformatics, Volume 40, Issue 10, October 2024, btae583,
<https://doi.org/10.1093/bioinformatics/btae583>

JOURNAL ARTICLE

Improved prediction of antibody and their complexes with clustered generative modelling ensembles

Xiaotong Xu, Marco Giulini, Alexandre M J J Bonvin  [Author Notes](#)

Bioinformatics Advances, Volume 5, Issue 1, 2025, vbaf161,
<https://doi.org/10.1093/bioadv/vbaf161>



bioRxiv
THE PREPRINT SERVER FOR BIOLOGY

Combining AI structure prediction and integrative modelling for nanobody-antigen complexes

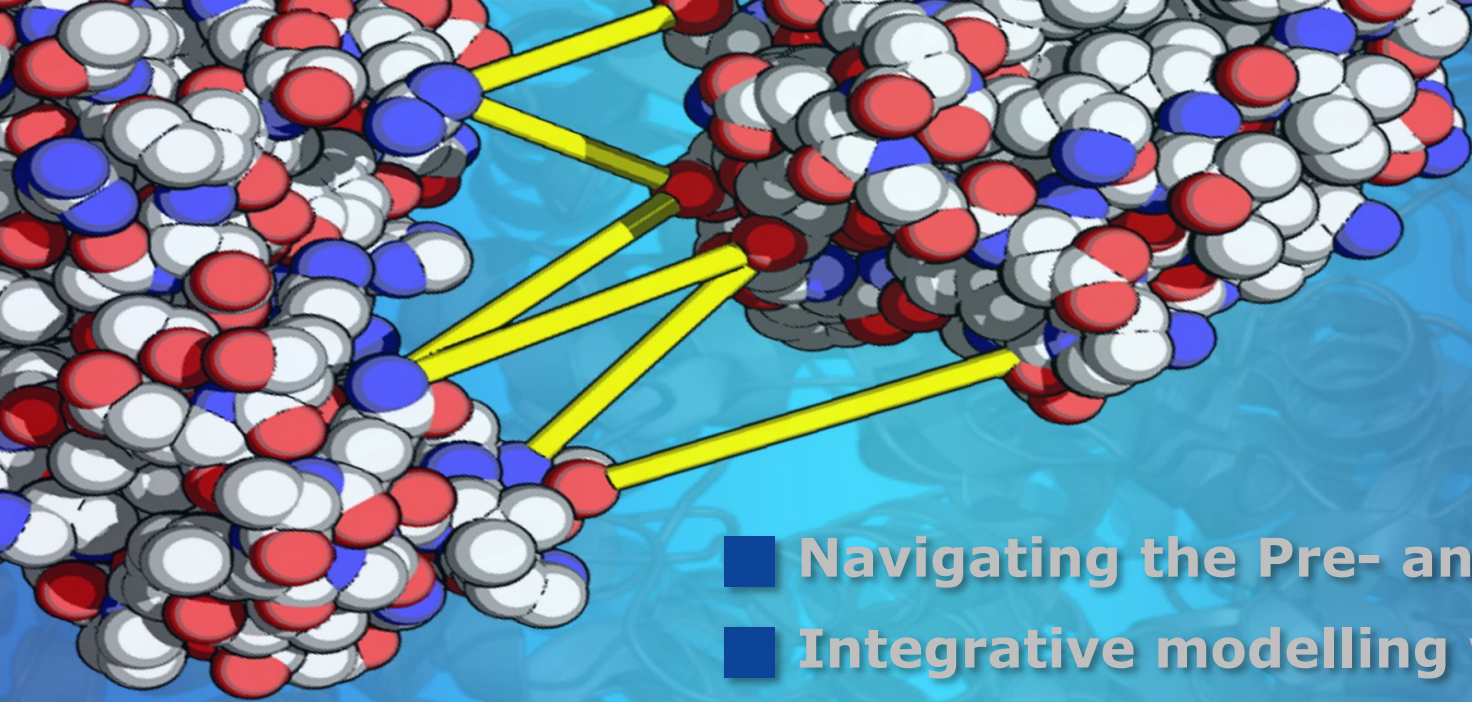
 Miguel Sánchez-Marín,  Marco Giulini,  Alexandre M.J.J. Bonvin

doi: <https://doi.org/10.1101/2025.07.01.662355>





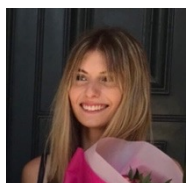
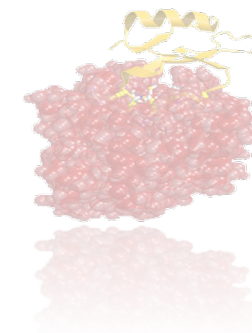
Overview



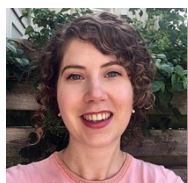
- Navigating the Pre- and Post-AlphaFold Divide
- Integrative modelling with HADDOCK
- Introducing HADDOCK3
- Application examples
 - An NMR-based application – where AlphaFold fails
 - AI-based antibody-antigen modelling
 - Cyclic peptides
- Conclusions & perspectives

```
111000100101010110101010
100011010101000101111010110101
10100001000111101010101110000110101
0001010101011110010100110010101010100
0101011111110101010000101010100000101010100
001011010101110101000111101010101000101001010101111
0001111111010101011001100010111000101011010100010101000101
```

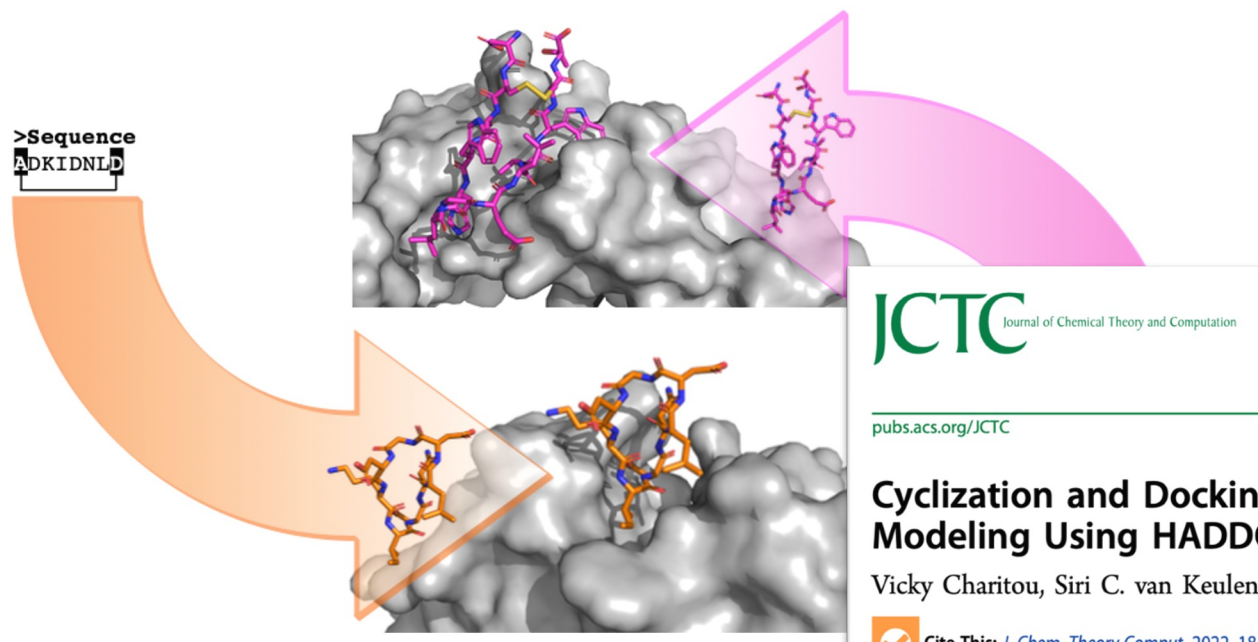
Modelling protein-cyclic peptide complexes with HADDOCK



Vicky Charitou



Siri van Keulen



JCTC
Journal of Chemical Theory and Computation

pubs.acs.org/JCTC



Article

Cyclization and Docking Protocol for Cyclic Peptide–Protein Modeling Using HADDOCK2.4

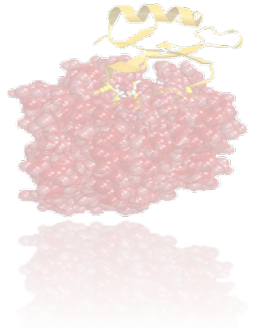
Vicky Charitou, Siri C. van Keulen,* and Alexandre M. J. J. Bonvin*

Cite This: *J. Chem. Theory Comput.* 2022, 18, 4027–4040

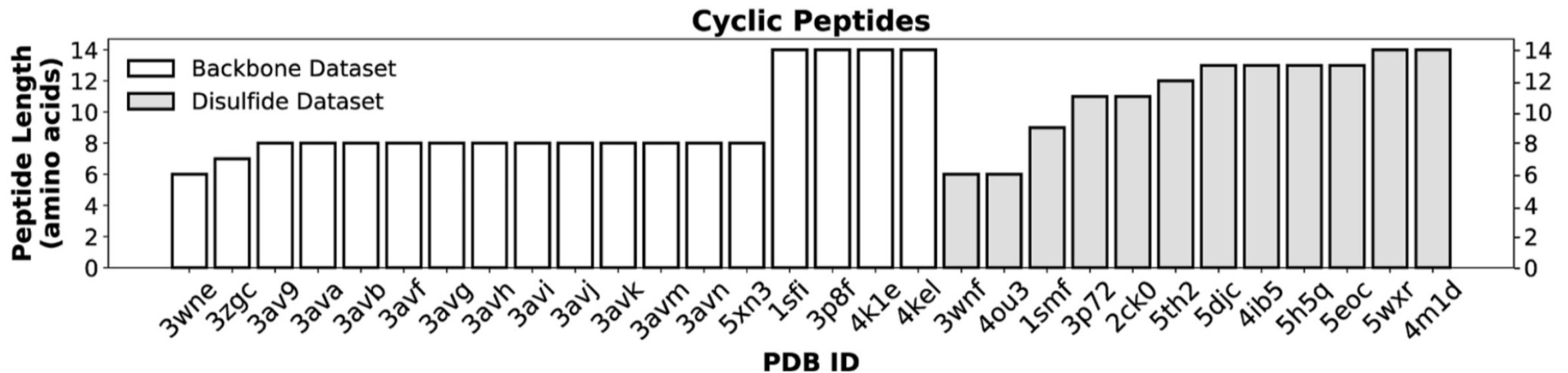
Read Online



Benchmark dataset



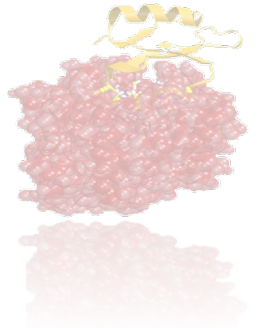
- 18 cyclic peptides (14 short ≤ 10 a.a.; 4 long 14 a.a)
- 12 disulfide peptides (3 short, 9 long)



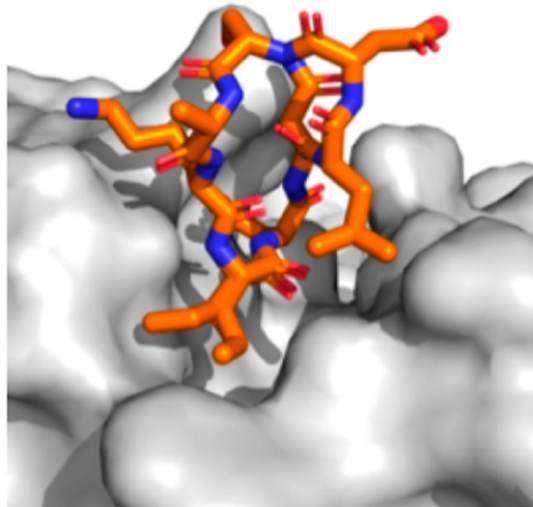
<https://github.com/haddock/cyclic-peptides>



Benchmark dataset

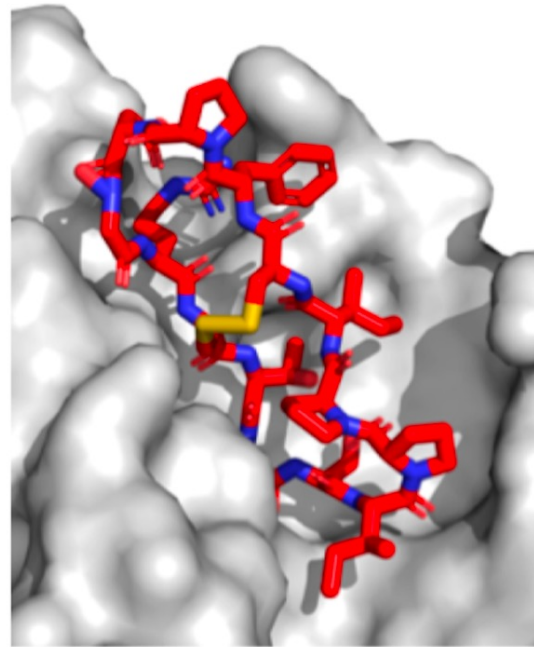


Short - backbone peptides
6-10 residues



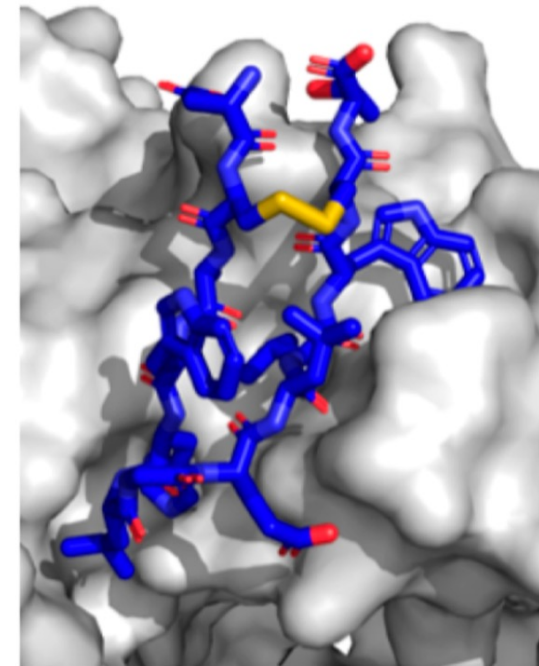
PDB ID: 3av9 (8 aa)

Long - backbone peptides
14 residues



PDB ID: 3p8f (14 aa)

Disulfide peptides
6-14 residues



PDB ID: 5djc (13 aa)



Peptide cyclisation protocol

Peptide Cyclization

Step 1 - PyMOL

polyproline



β -sheet

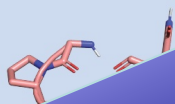


Generation of peptide starting conformations

2 structures



Step 2 - HADDOCK

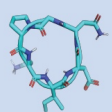


and distance

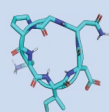
10 s



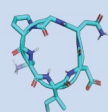
Step 3 - HADDOCK



Bond cyclization seed 1



Bond cyclization seed 2



Bond cyclization seed 3

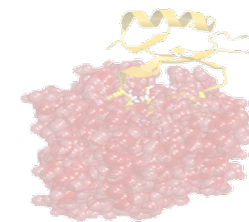
10 structures

+

10 structures

+

10 structures



- Based on the refinement stage (related angle space, full increased sampling)

- Step2: Distance restraints to bring the "termini" together (400 models, no electrostatics, RMSD clustering)

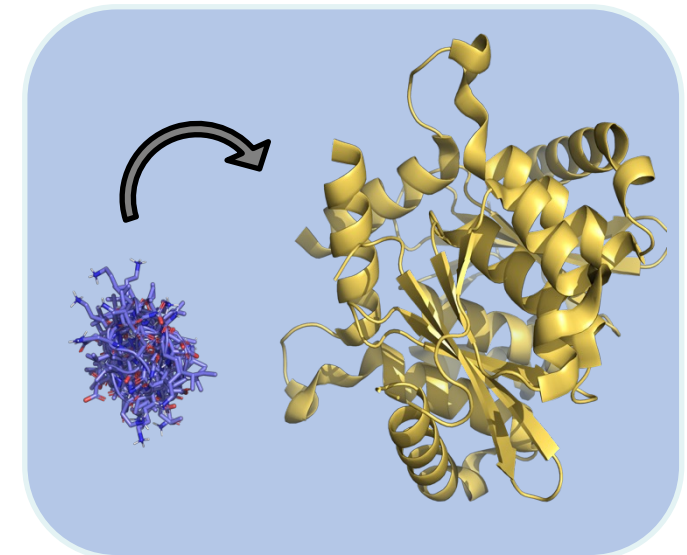
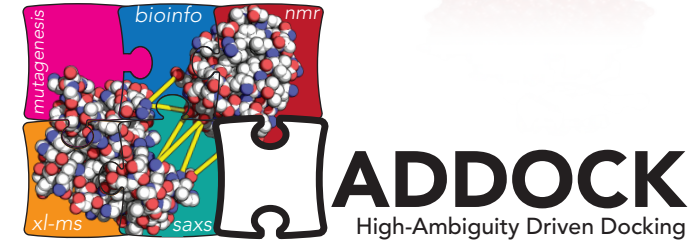
- Step3: Covalent bonds created automatically by HADDOCK (protocol repeated 3 time with different random seeds, RMSD clustering)

Can now be done within one single workflow in HADDOCK3

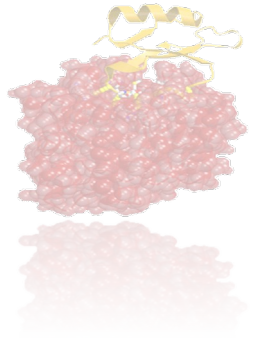


Ensemble docking protocol

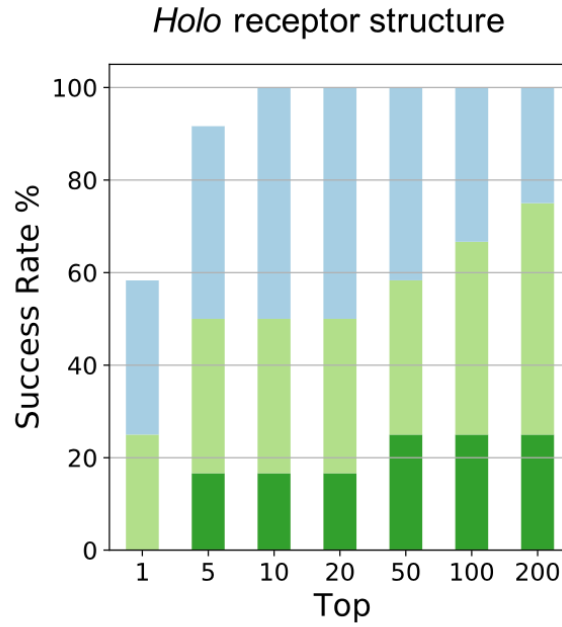
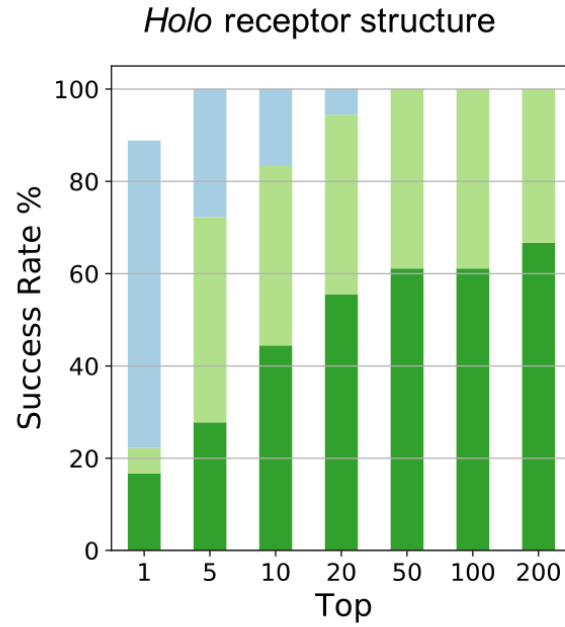
- Optimized HADDOCKing protocol:
 - Ensemble of 50 peptide conformations
 - 5000/400/400 sampling
 - #steps in flexible refinement increased by a factor 4
 - Peptides > 10 a.a. fully flexible and subjected to explicit solvent refinement
 - Restrained docking: Binding site on protein as active, peptide as passive



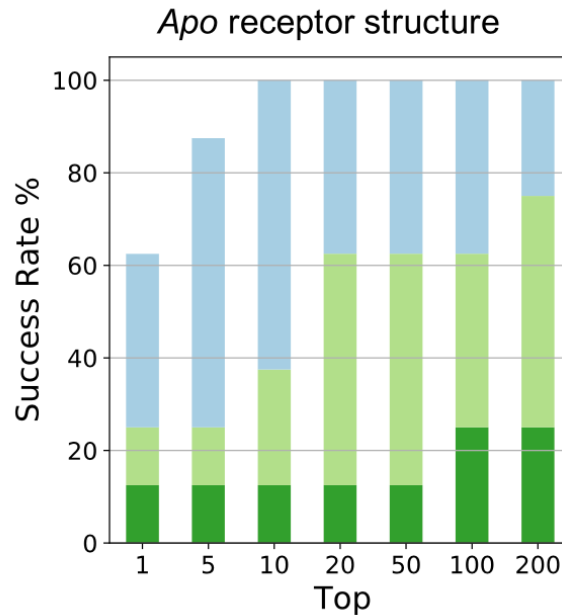
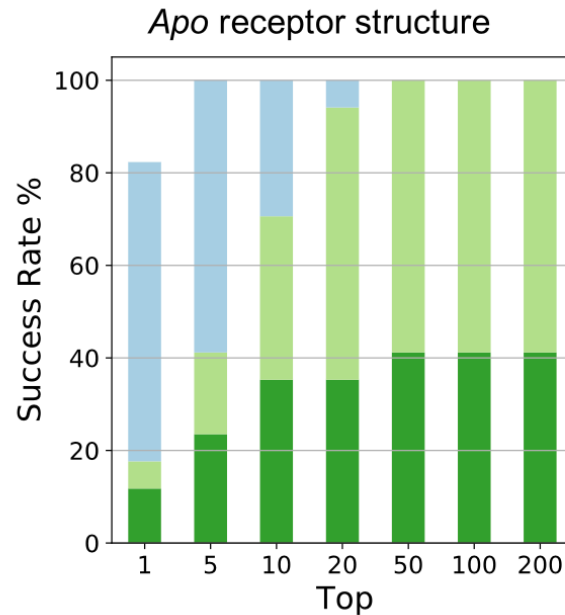
Overall performance



**Backbone
dataset
(#18)**



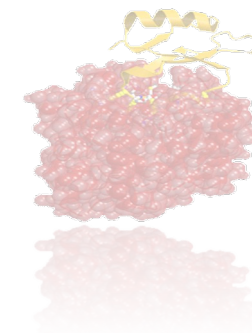
**Disulphide
dataset
(#12)**



■ fnat ≥ 0.8
■ fnat ≥ 0.5
■ fnat ≥ 0.2



Conclusions



- **Cyclisation protocol with distance restraints** (electrostatic off) **samples near native cyclic peptide conformations**
- **Ensemble docking protocol with increased flexibility for long (>10 a.a) peptides**
 - **100% acceptable top 10 models, 60% medium quality**
- **Long cyclic peptides remains challenging both in terms of sampling and scoring**

JCTC Journal of Chemical Theory and Computation



pubs.acs.org/JCTC

Article

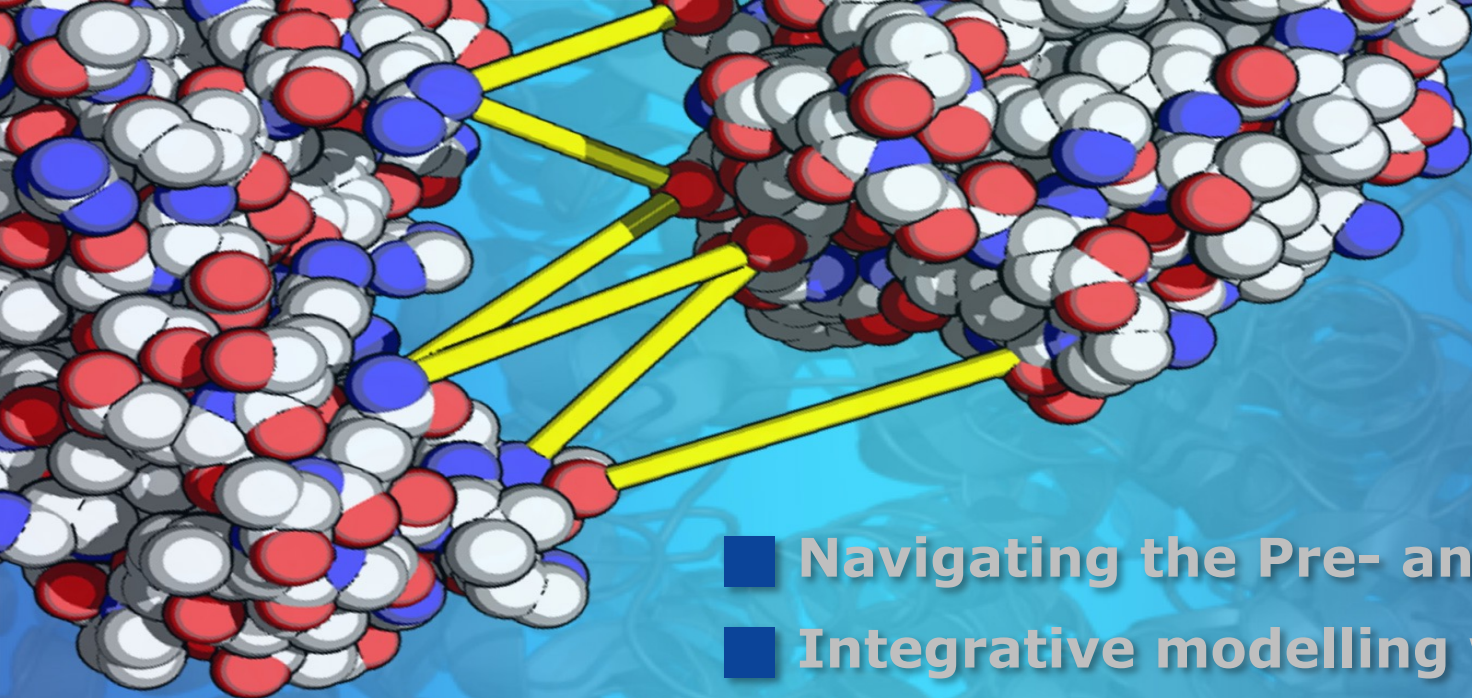
Cyclization and Docking Protocol for Cyclic Peptide–Protein Modeling Using HADDOCK2.4

Vicky Charitou, Siri C. van Keulen,* and Alexandre M. J. J. Bonvin*

 Cite This: *J. Chem. Theory Comput.* 2022, 18, 4027–4040

 [Read Online](#)



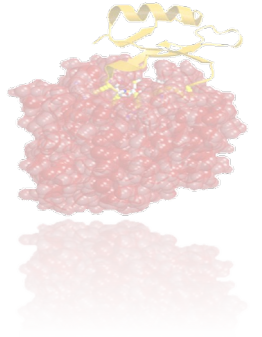


Overview

- Navigating the Pre- and Post-AlphaFold Divide
- Integrative modelling with HADDOCK
- Introducing HADDOCK3
- Application examples
 - An NMR-based application – where AlphaFold fails
 - AI-based antibody-antigen modelling
 - Cyclic peptides
- Conclusions & perspectives

```
111000100101010110101010
100011010101000101111010110101
10100001000111101010101110000110101
0001010101011110010100110010101010100
0101011111110101010000101010100000101010100
0010110101011101010001111010101010001010010101011111
0001111111010101011001100010111000101011010100010101000101
```

Conclusions & Perspective



- Celebrating 20+ years of HADDOCK developments!
- Despite the AI era, there is still (for now?) a need for more “classical” integrative modelling approaches (e.g. to deal with complexes consisting of a variety of molecule types, antibodies...)
- Increased flexibility with customizable workflows with **HADDOCK3** (which also allows incorporation of 3rd party software)



Acknowledgments: former and current CSB group@UU

€€



VICI
TOP-PUNT



WeNMR

West-Life

EGI-Engage

INDIGO-Datacloud

BioExcel CoE

EOSC-Hub

EGI-Ace

EOSC-Data Commons



bonvinlab.org/people



netherlands
Science center

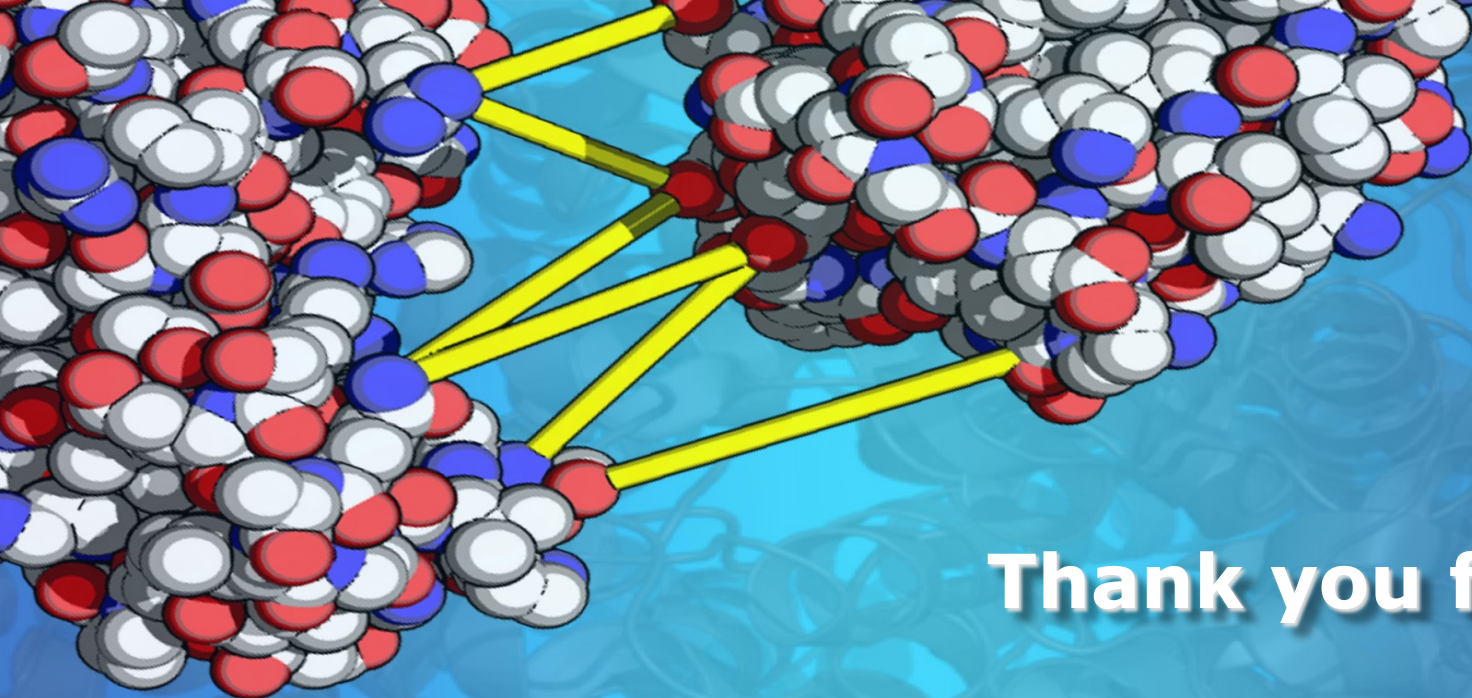
SURF SARA

Acknowledgments: the HADDOCK developers over the years

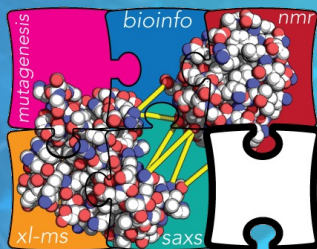


HADDOCK 20 years - Huizen NL, November 8-9 2023

110101010
010110101
000110101
101010100
101010100
101011111
000111111010101011001100010111000101011010100010101000101



Thank you for your attention!



ADDOCK
High-Ambiguity Driven Docking

HADDOCK online:

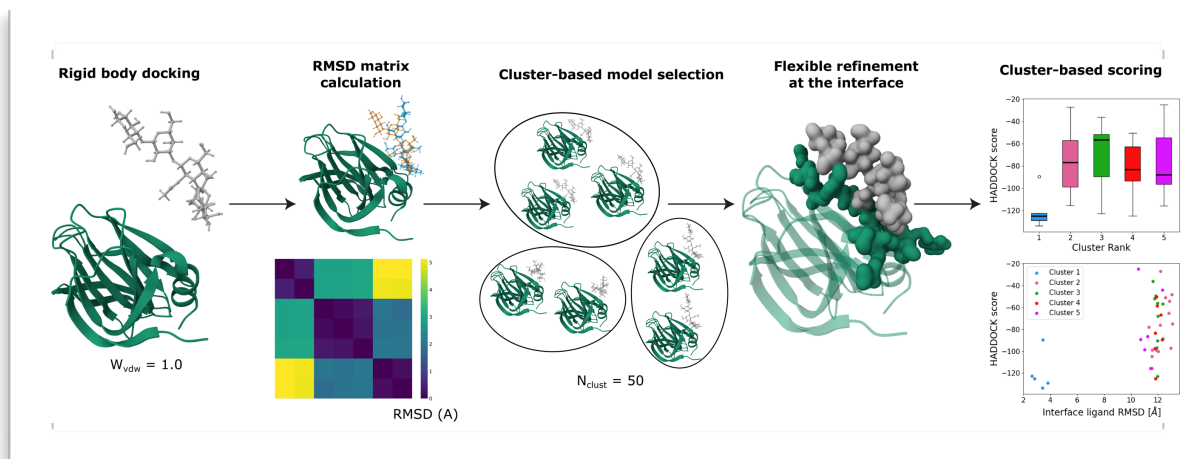
- <https://wenmr.science.uu.nl>
- <https://bonvinlab.org/software>
- <https://bonvinlab.org/haddock3-user-manual>
- <https://ask.bioexcel.eu>



Utrecht
Bioinformatics
Center

111000100101010110101010
100011010101000101111010110101
10100001000111101010101110000110101
0001010101011110010100110010101010100
0101011111101010100001010101000001010100
0010110101011101010001111010101000101001010101111
00011111110101010110011000101110001010110101000101010001010

Protein-glycans



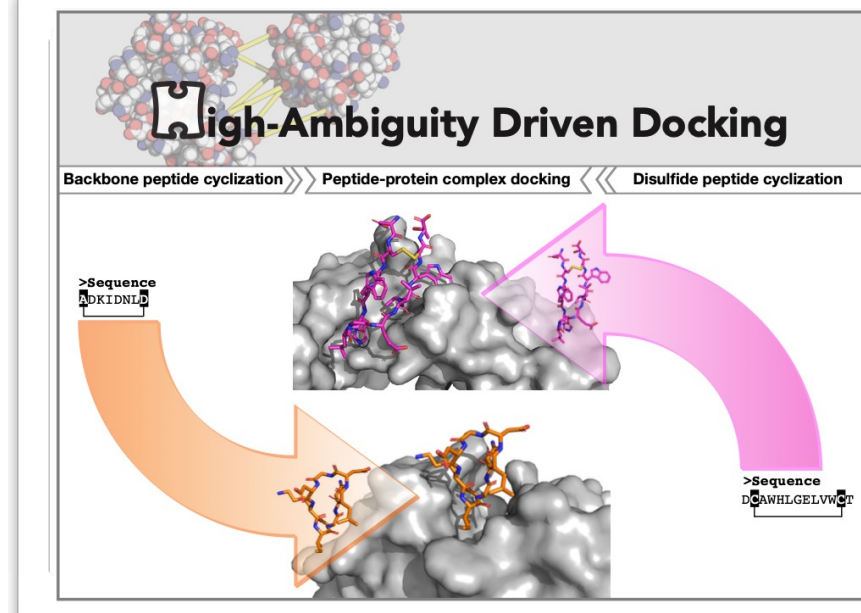
Modeling Protein–Glycan Interactions with HADDOCK

Anna Ranaudo, Marco Giulini, Angela Pelissou Ayuso, and Alexandre M. J. J. Bonvin*

Cite This: *J. Chem. Inf. Model.* 2024, 64, 7816–7825

Read Online

Protein-cyclic peptides



Cyclization and Docking Protocol for Cyclic Peptide–Protein Modeling Using HADDOCK2.4

Vicky Charitou, Siri C. van Keulen,* and Alexandre M. J. J. Bonvin*

Cite This: *J. Chem. Theory Comput.* 2022, 18, 4027–4040

Read Online