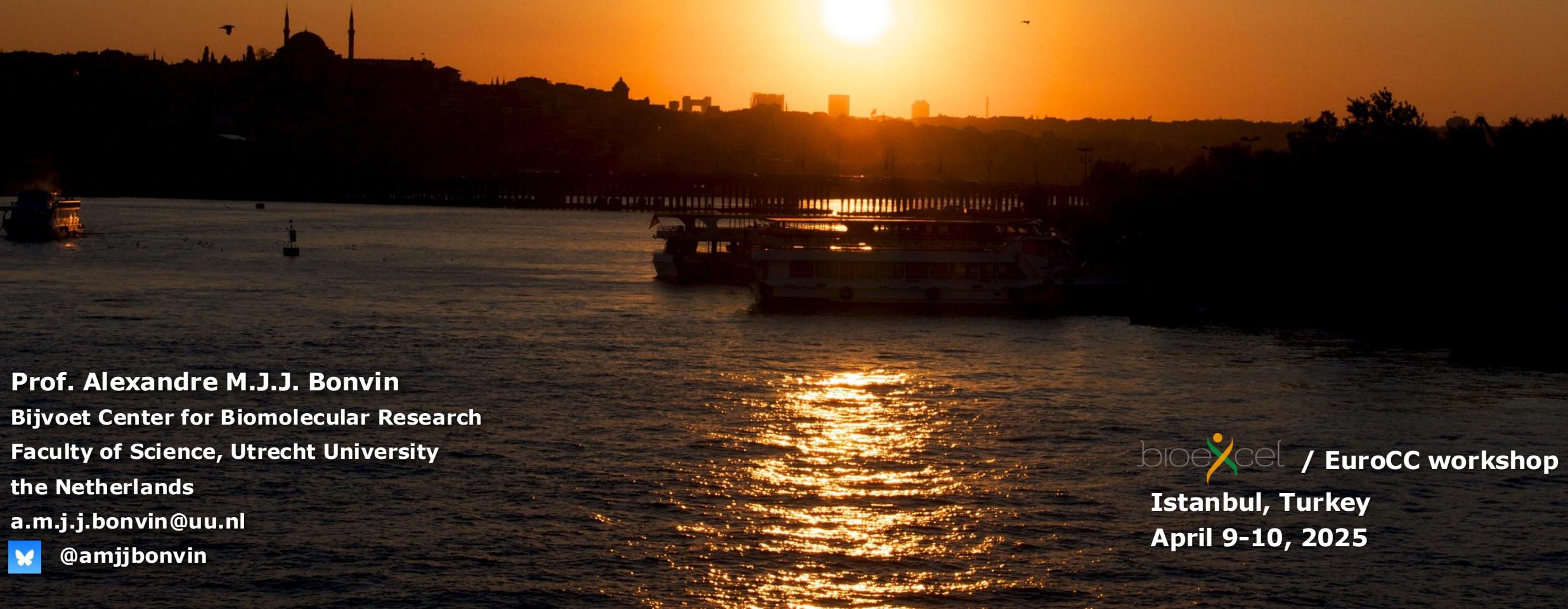


Integrative modeling of biomolecular complexes



Prof. Alexandre M.J.J. Bonvin

Bijvoet Center for Biomolecular Research

Faculty of Science, Utrecht University

the Netherlands

a.m.j.j.bonvin@uu.nl



[@amjjbonvin](https://twitter.com/amjjbonvin)

 / EuroCC workshop

Istanbul, Turkey

April 9-10, 2025

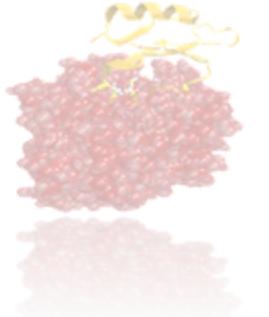
1st HADDOCK workshop
Boğaziçi University
June 2011



Overview

- **Introduction**
 - **Information-driven docking with HADDOCK**
 - **Introducing HADDOCK3**
 - **AI-based antibody-antigen modelling**
 - **Protein cyclic-peptides**
 - **Protein-glycans**
 - **Conclusions & perspectives**

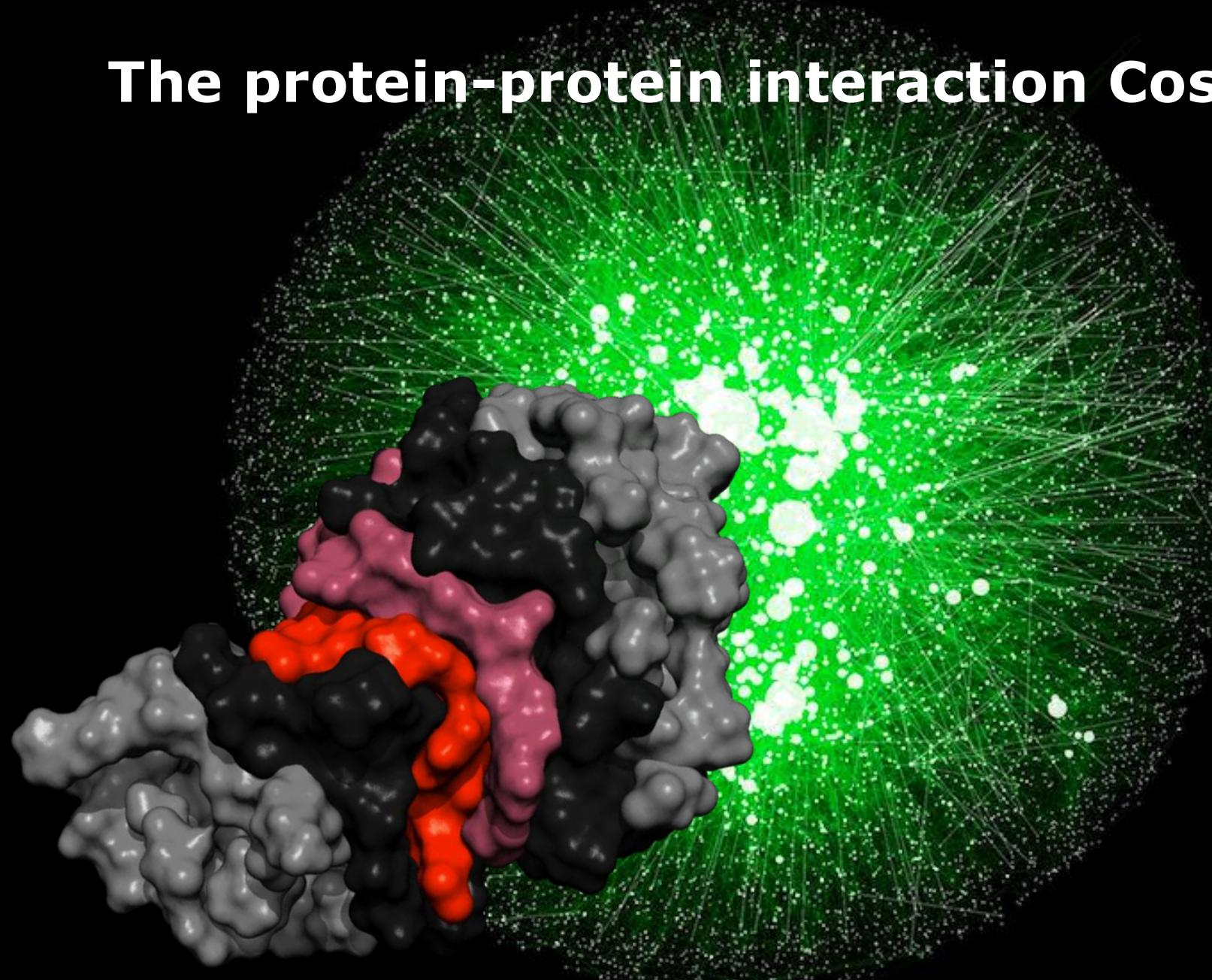
The social network of proteins



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



The protein-protein interaction Cosmos



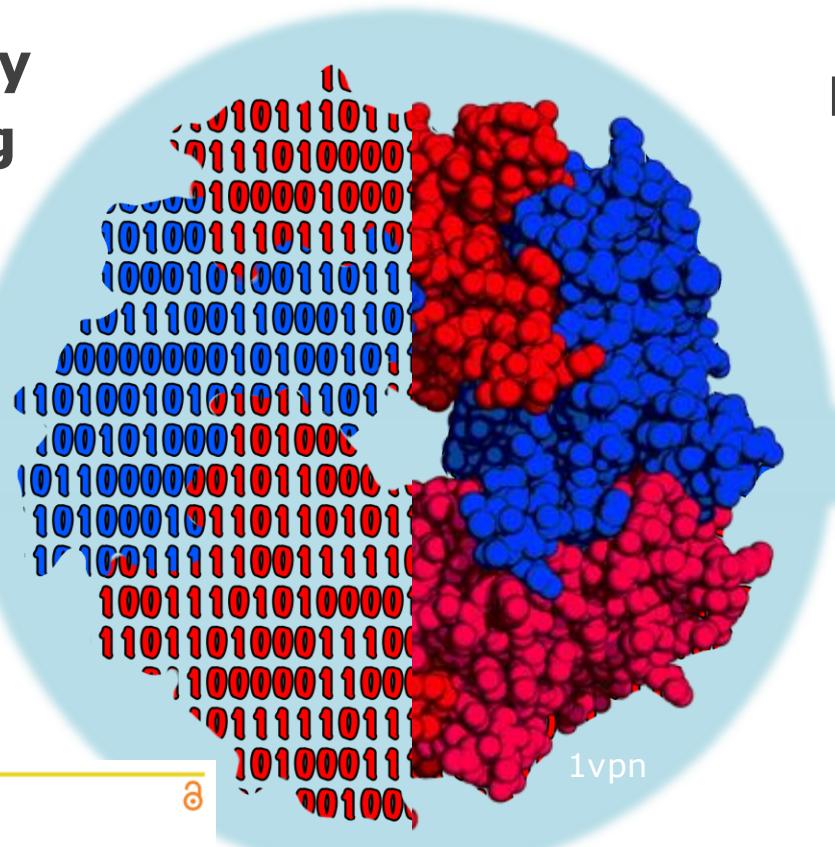
Structural biology of interactions

Docking
Molecular
Dynamics

Homology
Modeling

Threading

SAS
MS
X-Ray
EPR
NMR
FRET
Cryo-EM



Experiment

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 The copyright holder for this preprint (which was not certified by peer review) is the author/funder.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 Žídek¹,
Zielinski¹, Alex B
Clancy¹, Pushmeet
Kaur¹, DeepMind¹,
London¹

ARTICLE

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

OP

Improved prediction
using AlphaFold2

Patrick Bryant^{1,2,3,4,5}, Gabriele Pozzati^{1,2,3,4,5}

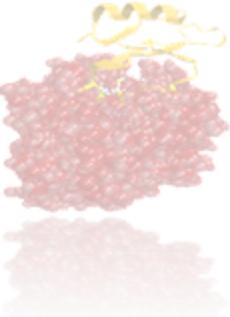
Article
Protein structure prediction
photo-crosslinking and
deep learning

Received: 26 September 2022
Accepted: 6 February 2023
Koija Stahl⁶
& Juri Rapps⁶

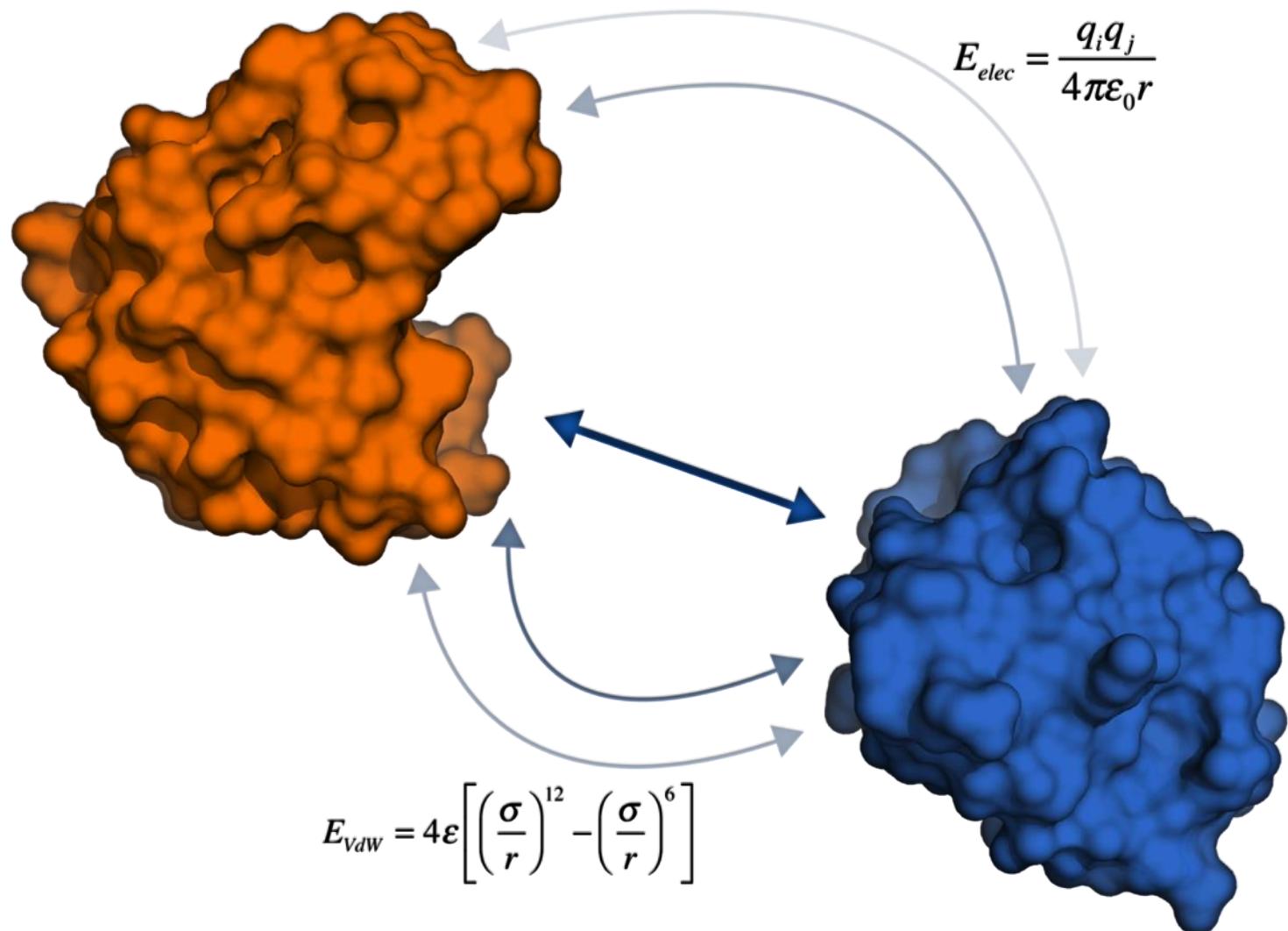
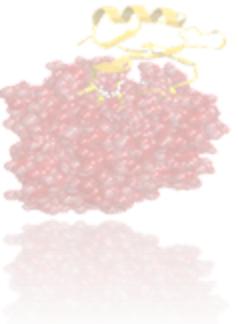
Josh Abramson, Jonas Adler, Jack Dunger, Richard Evans, Tim Green, Alexander Pritzel, Olaf Ronneberger, Lindsay Willmore, Andrew J. Ballard, Joshua Bambrick, Sebastian W. Bodenstein, David A. Evans, Chia-Chun Hung, Michael O'Neill, David Reiman, Kathryn Tunyasuvunakool, Zachary Wu, Akvilė Žemgulytė,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

Nature (2024) | [Cite this article](#)

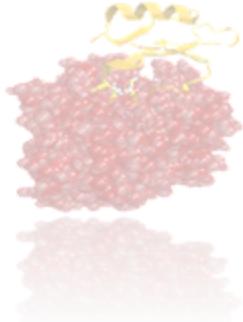
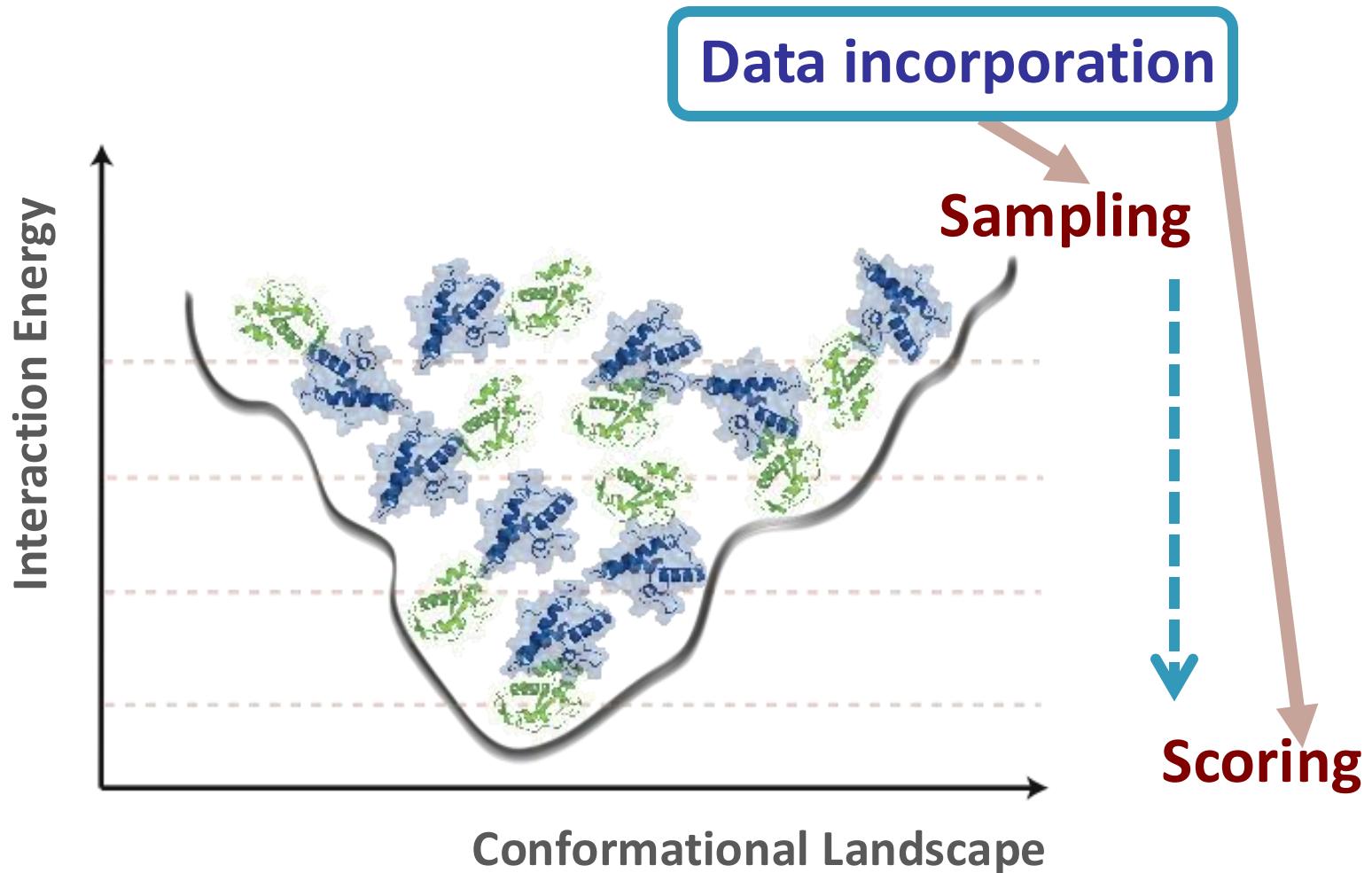
-resolution experiments
by the experimental community



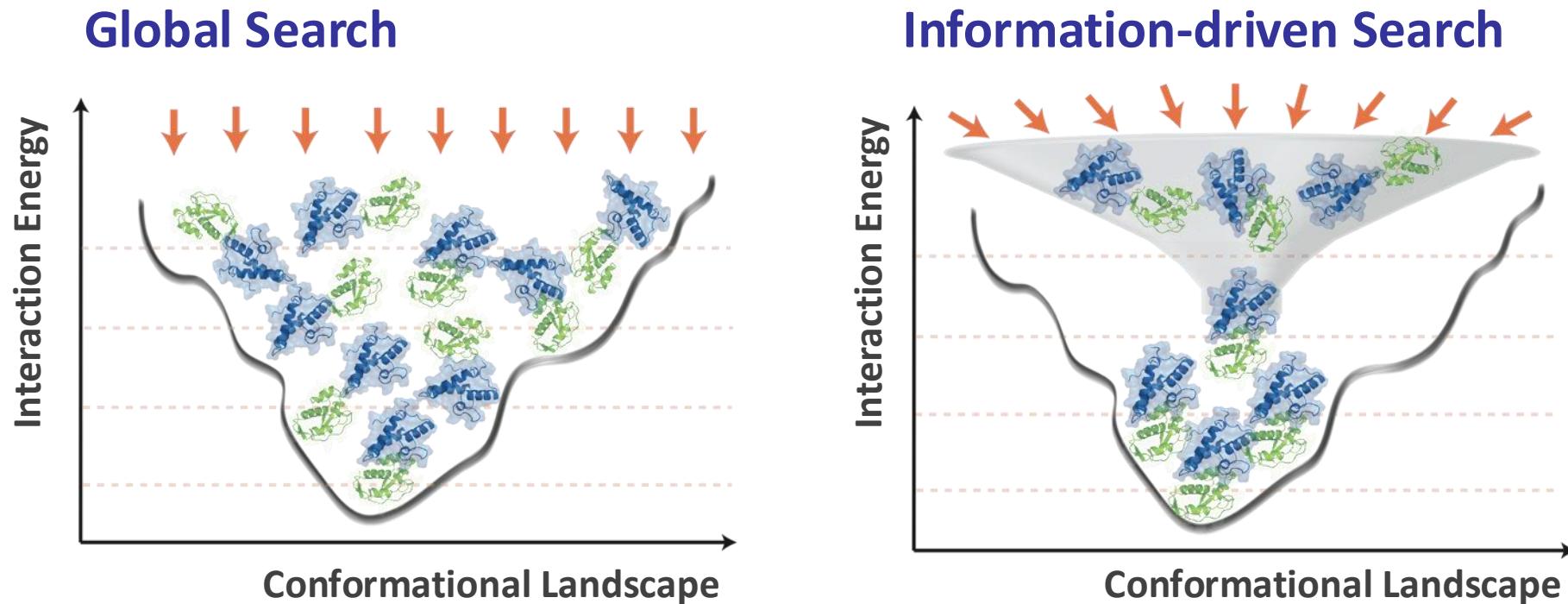
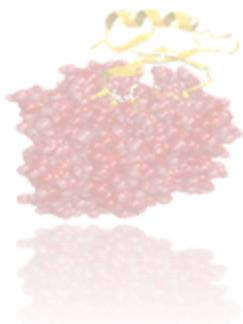
Molecular Docking



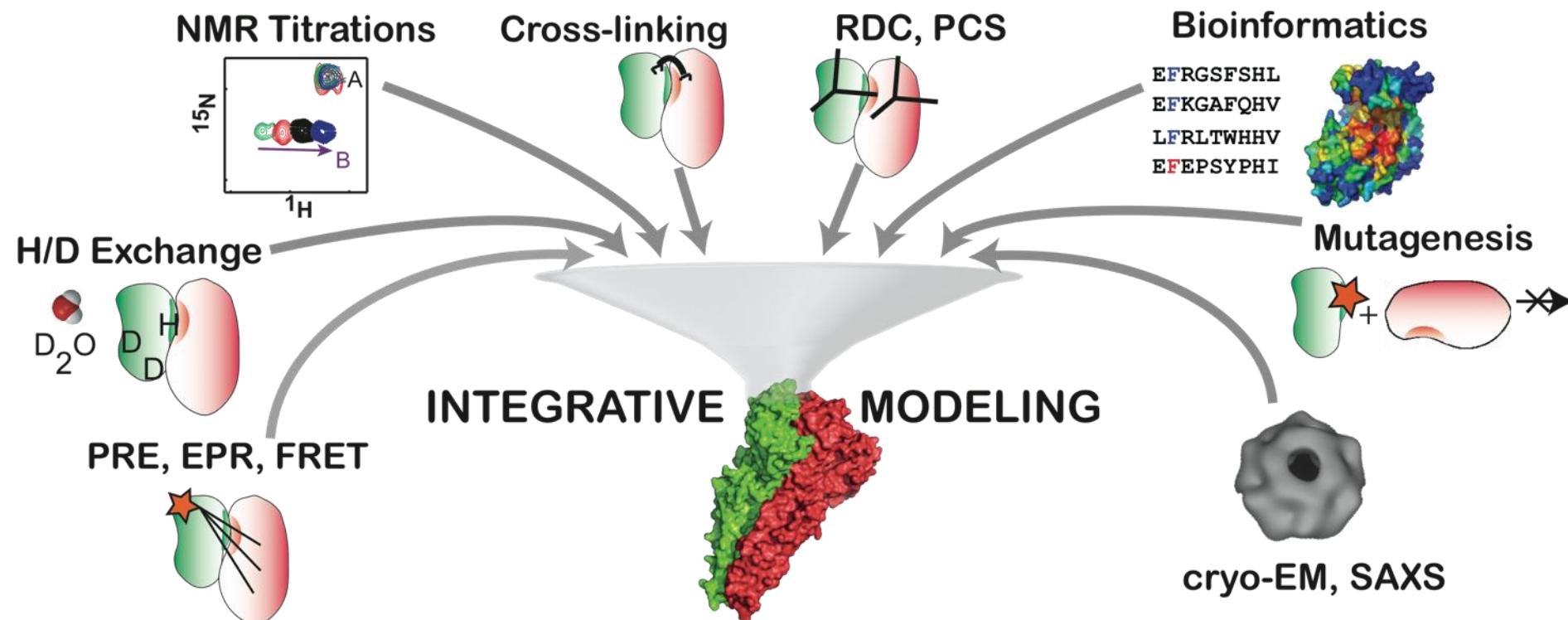
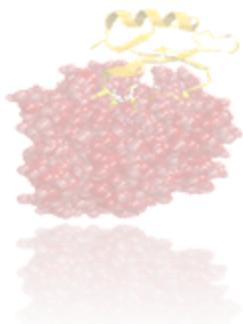
Methodology



Data Integration during Sampling



What is Integrative Modeling?



Overview

- **Introduction**
 - **Information-driven docking with HADDOCK**
 - **Introducing HADDOCK3**
 - **AI-based antibody-antigen modelling**
 - **Protein cyclic-peptides**
 - **Protein-glycans**
 - **Conclusions & perspectives**

HADDOCK: An integrative modeling platform

Incorporates ambiguous and low-resolution data to aid the docking

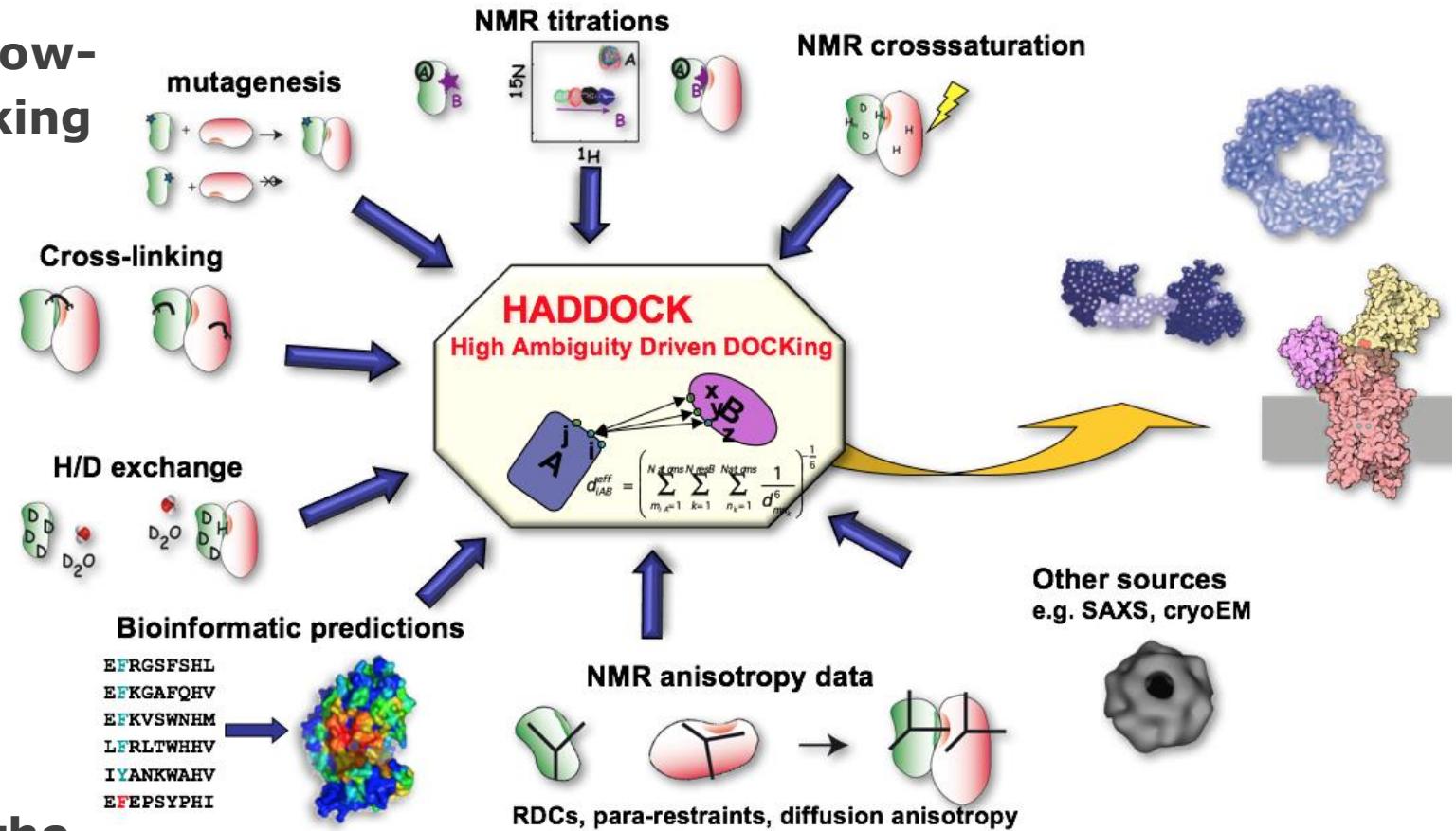
Capable of docking up to 20 molecules (2.4 version)

Symmetries can be leveraged

Allows for flexibility at the interface

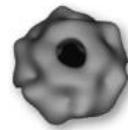
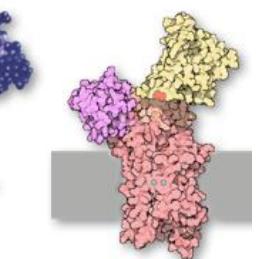
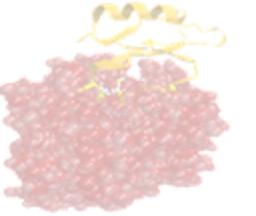
Final flexible refinement in explicit solvent

Consistent performance over the years in CAPRI



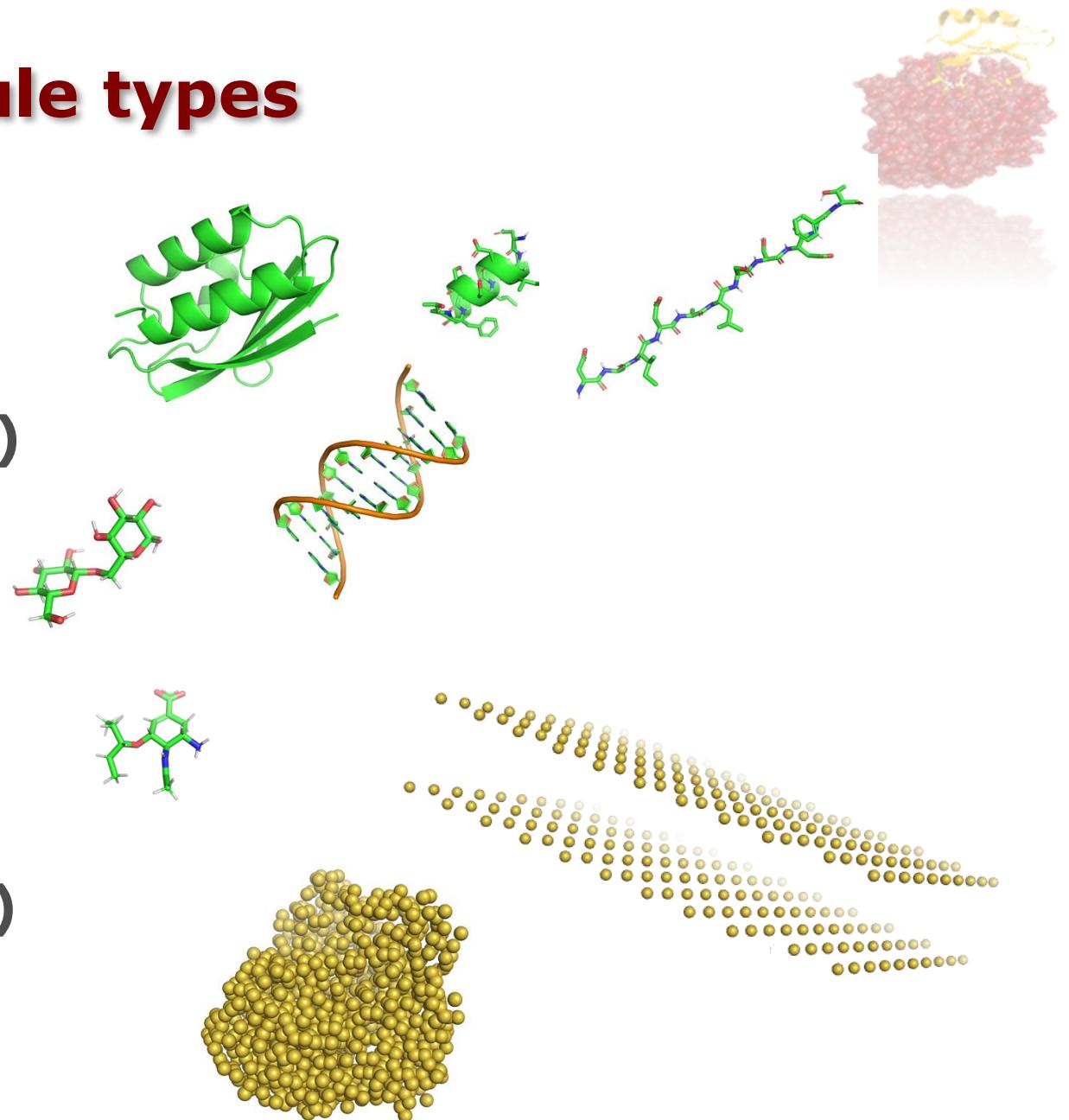
Dominguez, Boelens & Bonvin. JACS 125, 173 (2003).

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

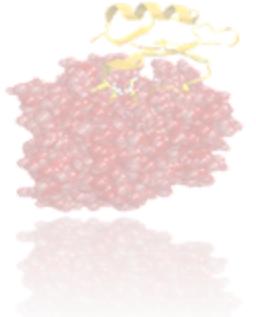


Molecule types

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

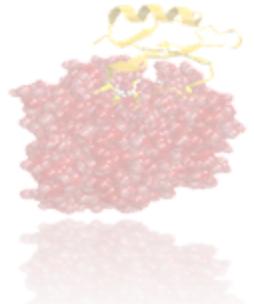


Restraint types

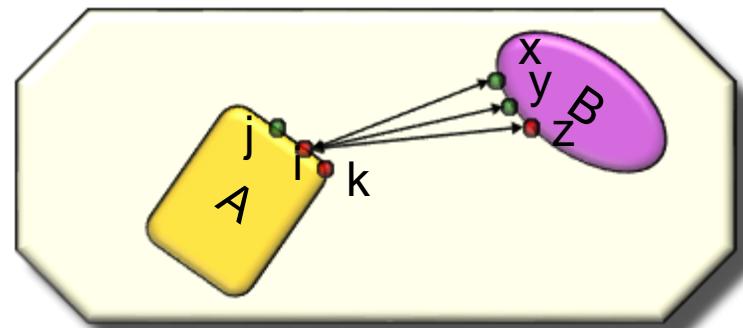


- **Distances (ambiguous / unambiguous)**
- **Dihedral angles**
- **NMR residual dipolar couplings**
- **NMR pseudo contact shifts**
- **NMR relaxation anisotropy**
- **Cryo-EM**
- **Shapes**
- **Radius of gyration**
- **Symmetry (C₂ – C₆)**
- **Non-crystallographic symmetry**

Data-driven docking with HADDOCK



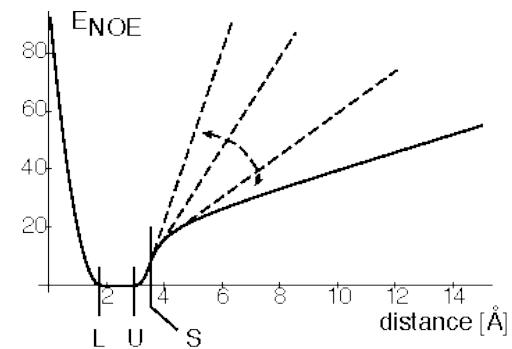
List of interface residues
for protein A



List of interface residues
for protein B

Effective distance d_{iAB}^{eff}
calculated as

$$d_{iAB}^{\text{eff}} = \left(\sum_{m_{iA}=1}^{N_{\text{atoms}}} \sum_{k=1}^{N_{\text{resB}}} \sum_{n_k=1}^{N_{\text{atoms}}} \frac{1}{d_{mn_k}^6} \right)^{-\frac{1}{6}}$$



Ambiguous Interaction Restraint:
a residue must make contact with any residue from
the other list

Different fraction of restraints (typically 50%)
randomly deleted for each docking trial to deal with
inaccuracies and errors in the information used

$$E_{\text{NOE}} = \begin{cases} (r - L)^2 & \text{if } r < L \\ 0 & \text{if } L < r < U \\ (U - r)^2 & \text{if } U < r < S \\ A(r - U)^{-1} + B(r - U) + C & \text{if } r > S \end{cases}$$

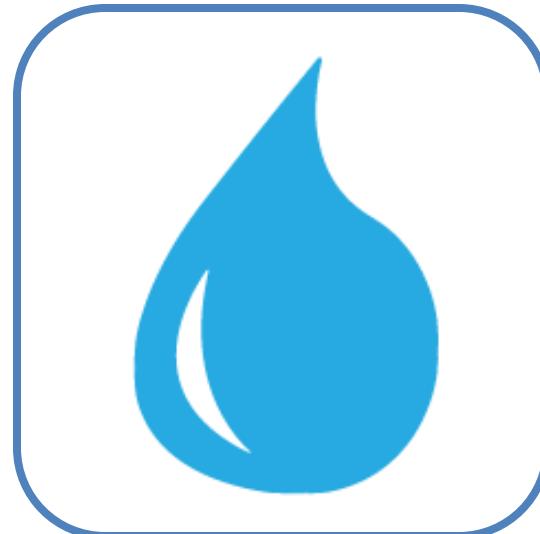
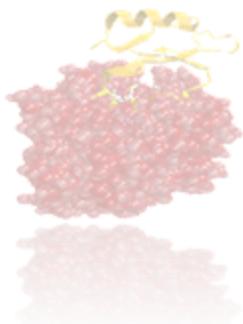
(Nilges & Brunger 1991)



Universiteit Utrecht

Faculty of Science
Chemistry

HADDOCK docking protocol



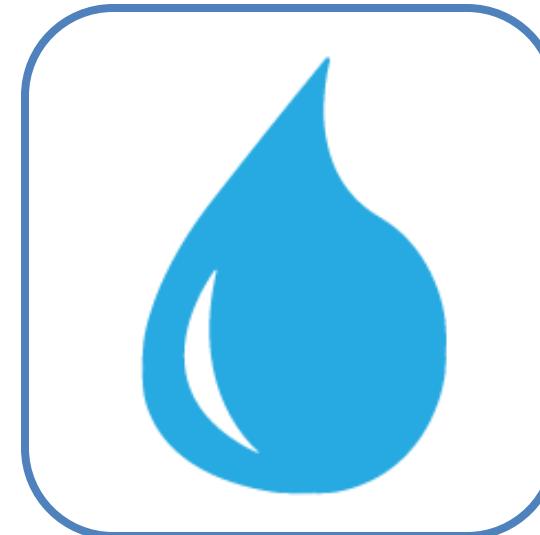
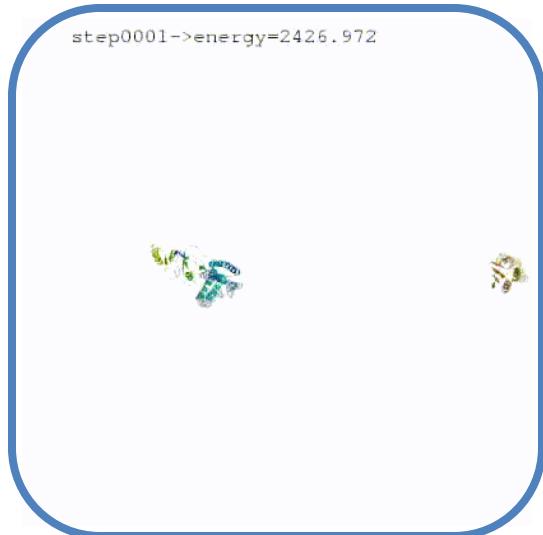
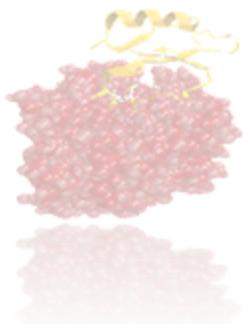
it0

it1

itw

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

HADDOCK docking protocol



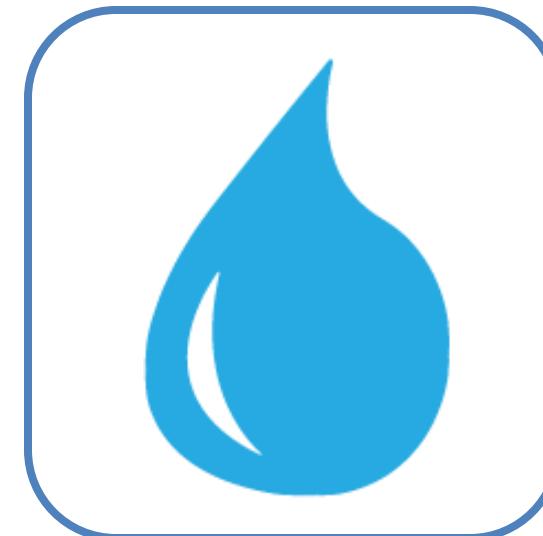
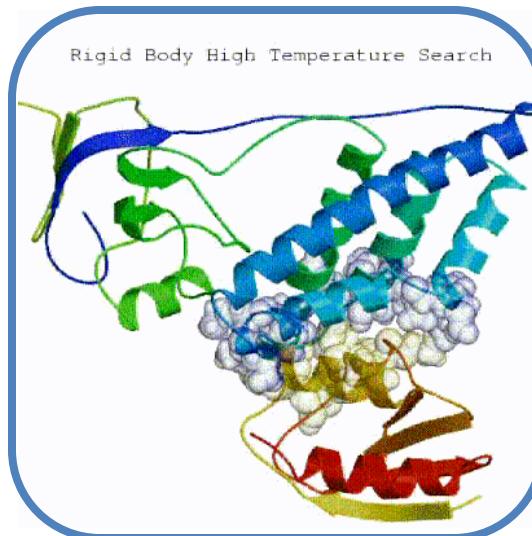
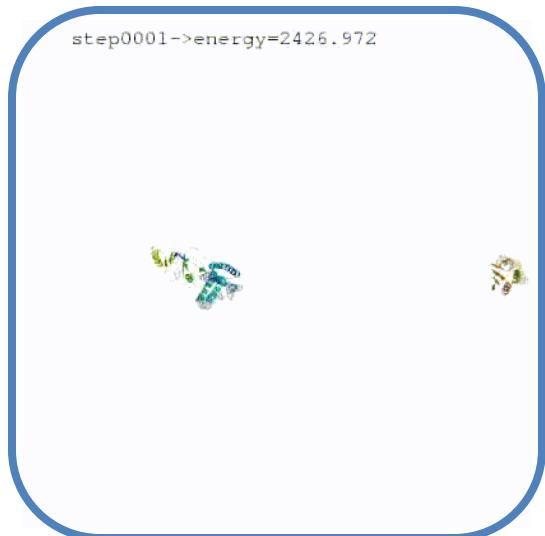
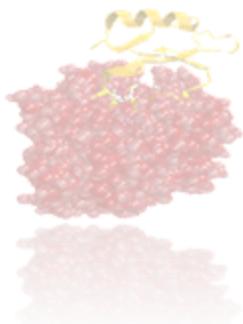
it0

it1

itw

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

HADDOCK docking protocol



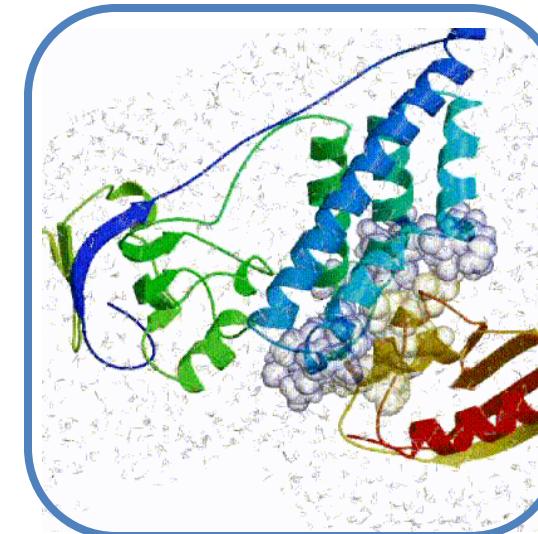
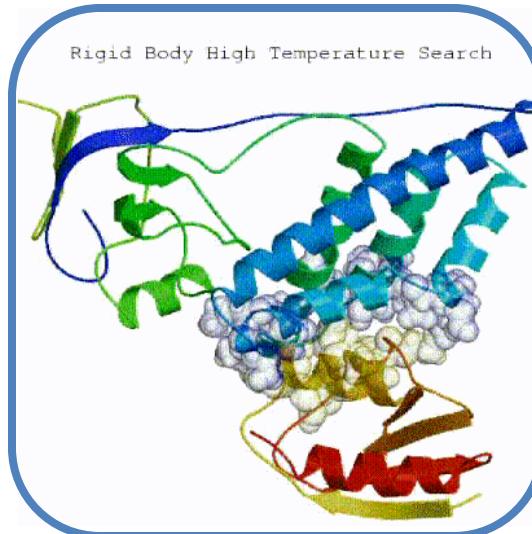
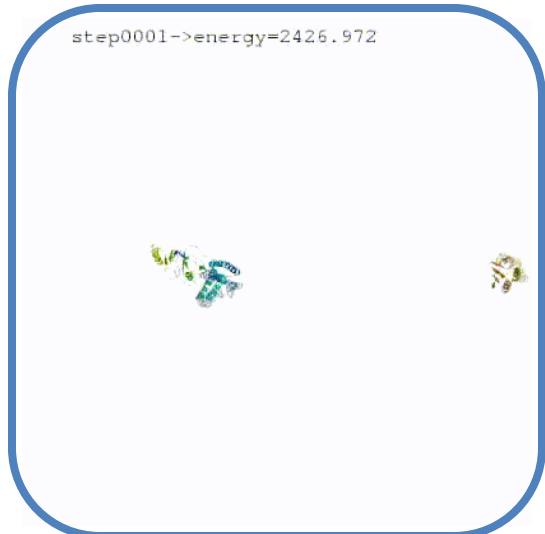
it0

it1

itw

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

HADDOCK docking protocol

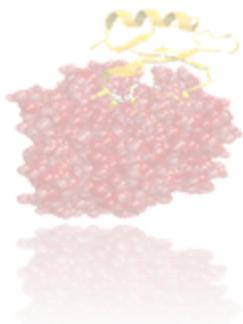


it0

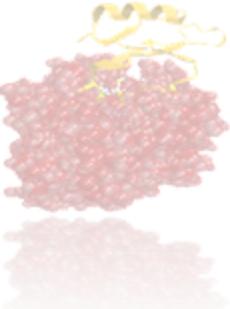
it1

itw

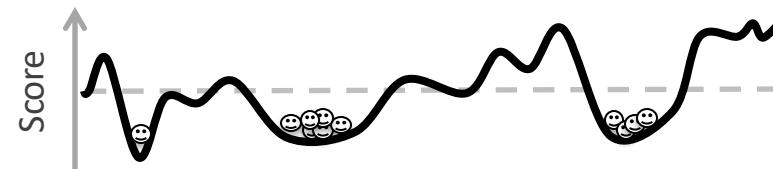
**Succession of energy minimization and molecular dynamics protocols
reminiscent of NMR structure calculations**



Energetics & Scoring



- **OPLS non-bonded parameters** (Jorgensen, *JACS* 110, 1657 (1988))
- **8.5 Å non-bonded cutoff, switching function, $\Sigma=10$**
- **Clustering of solutions**
- **Ranking based on cluster-based HADDOCK score:**



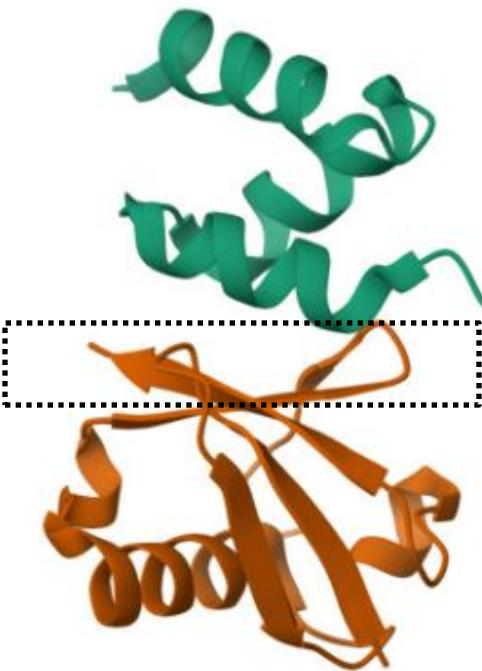
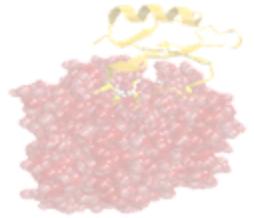
Rigid: $\text{Score} = 0.01 E_{\text{air}} + 0.01 E_{\text{vdW}} + 1.0 E_{\text{elec}} + 1.0 E_{\text{desolv}} - 0.01 \text{BSA}$

Flexible: $\text{Score} = 0.1 E_{\text{air}} + 1.0 E_{\text{vdW}} + 1.0 E_{\text{elec}} + 1.0 E_{\text{desolv}} - 0.01 \text{BSA}$

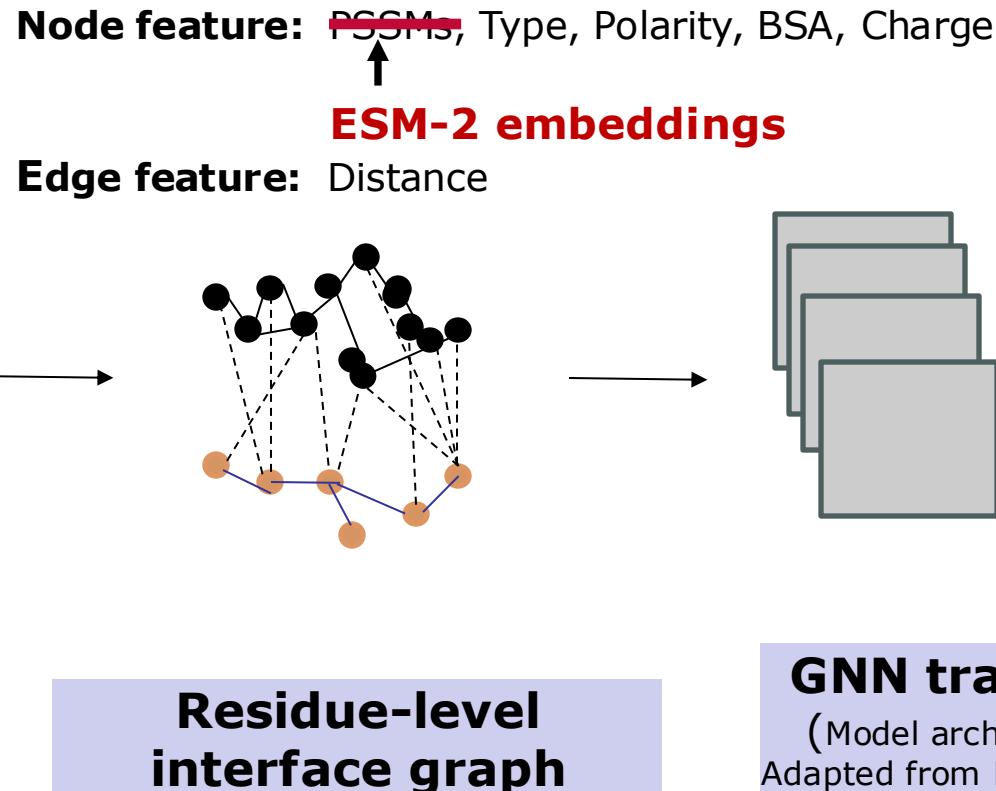
Water: $\text{Score} = 0.1 E_{\text{air}} + 1.0 E_{\text{vdW}} + 0.2 E_{\text{elec}} + 1.0 E_{\text{desolv}}$

- **E_{air} :** ambiguous interaction restraint energy
- **E_{desolv} :** desolvation energy using Atomic Solvation Parameters (Fernandez-Recio et al *JMB* 335, 843 (2004))
- **BSA:** buried surface area

Graph Neural Network with protein language model embeddings



PPI structure



Residue-level interface graph

GNN training
(Model architecture Adapted from DeepRank-GNN)

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>

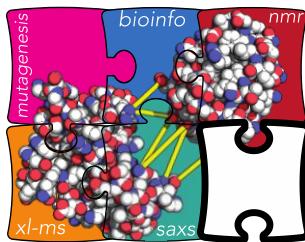
Published: 05 January 2024 Article history 

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



Universiteit Utrecht

[Faculty of Science Chemistry]



ADDOCK
High-Ambiguity Driven Docking

> 58000 registered users

> 690000 served runs since June 2008

➤ 65% on the EOSC HTC resources
(>85% for 2.4)

Vargas Honorato *et al.* Nature Prot. 2024

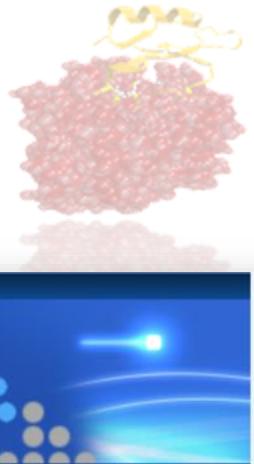
Van Zundert *et al.* J.Mol.Biol. 2016

De Vries *et al.* Nature Prot. 2010



SURF SARA

HADDOCK web portal (WeNMR)



HADDOCK 2.4
@Bonvinlab

WELCOME TO THE Utrecht BIOMOLECULAR INTERACTION WEB PORTAL >>

Welcome! HADDOCK (High Ambiguity Driven protein-protein DOCKing) 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.

ADDOCK
High Ambiguity Driven Docking

New to HADDOCK? To use the HADDOCK docking server you must have registered for an account.

Our server is easier than ever to use. Try our new submission interface!

HADDOCK is used for excellent science and so far it has been cited more than 5000 times!

Looking for support or questions about HADDOCK's usage? Check our BioExcel forum!

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

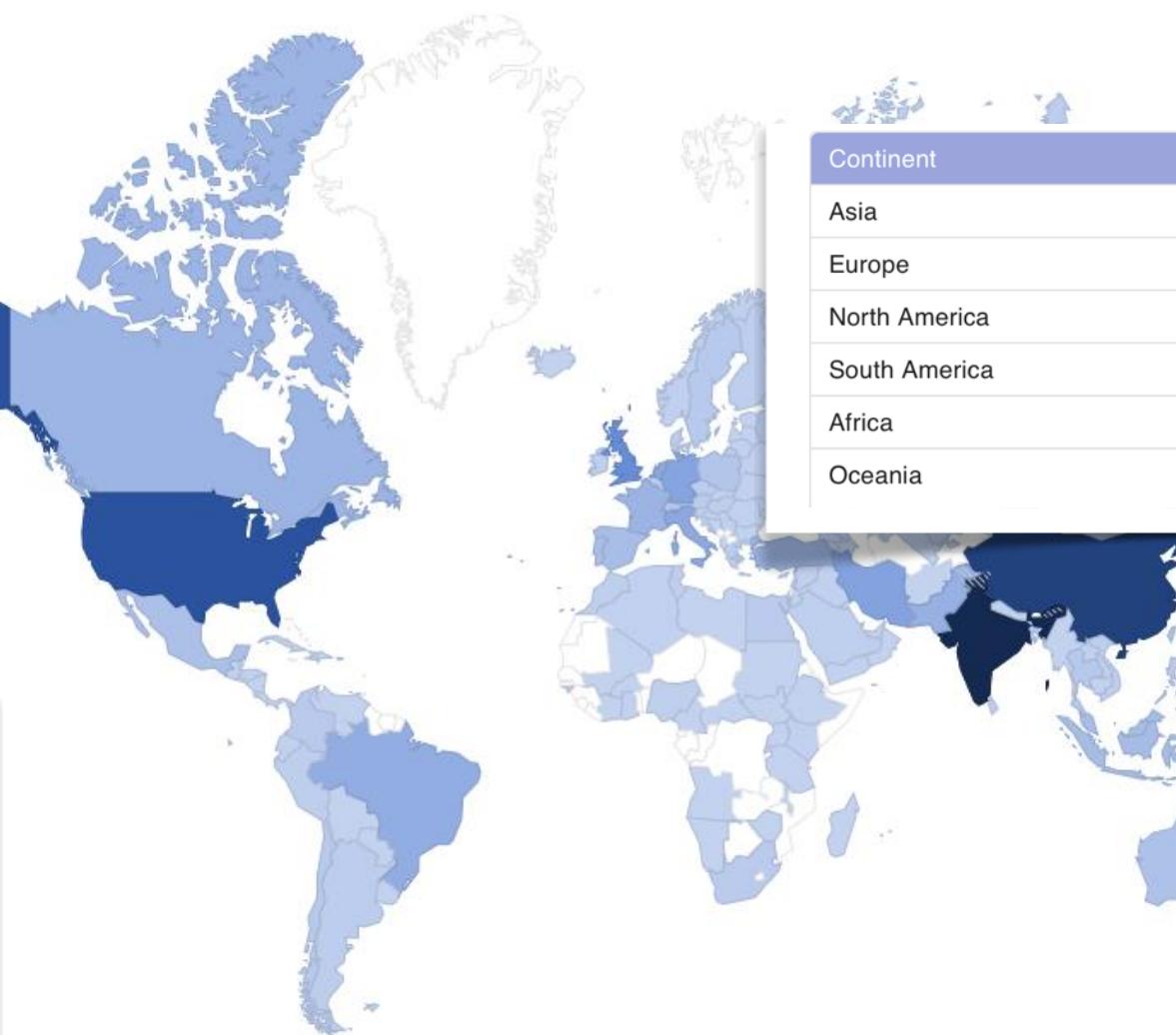
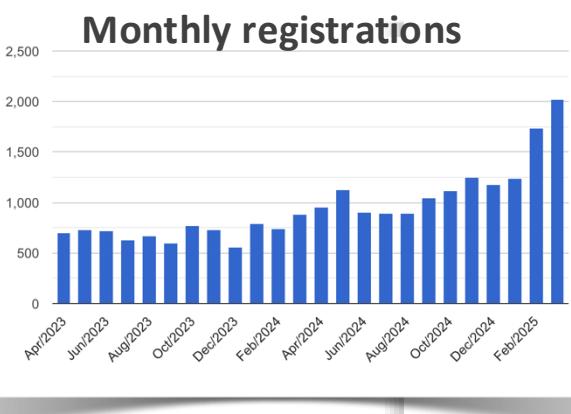
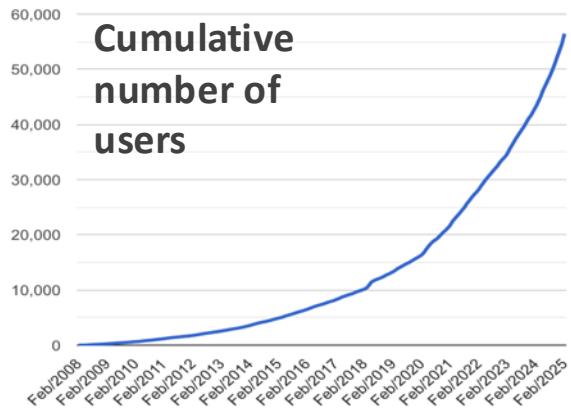
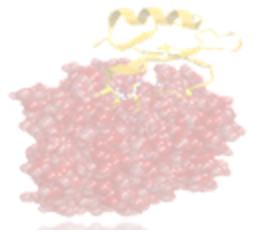


Universiteit Utrecht

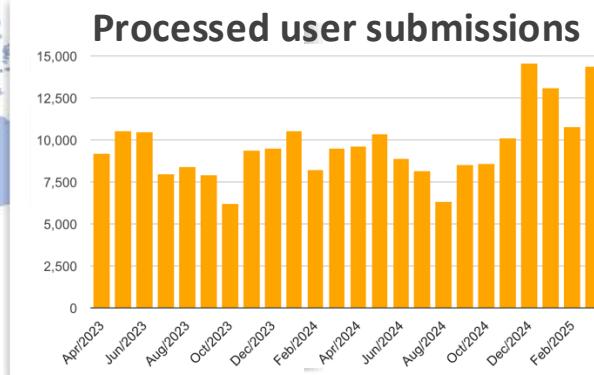
[Faculty of Science
Chemistry]

Haddock web portal

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

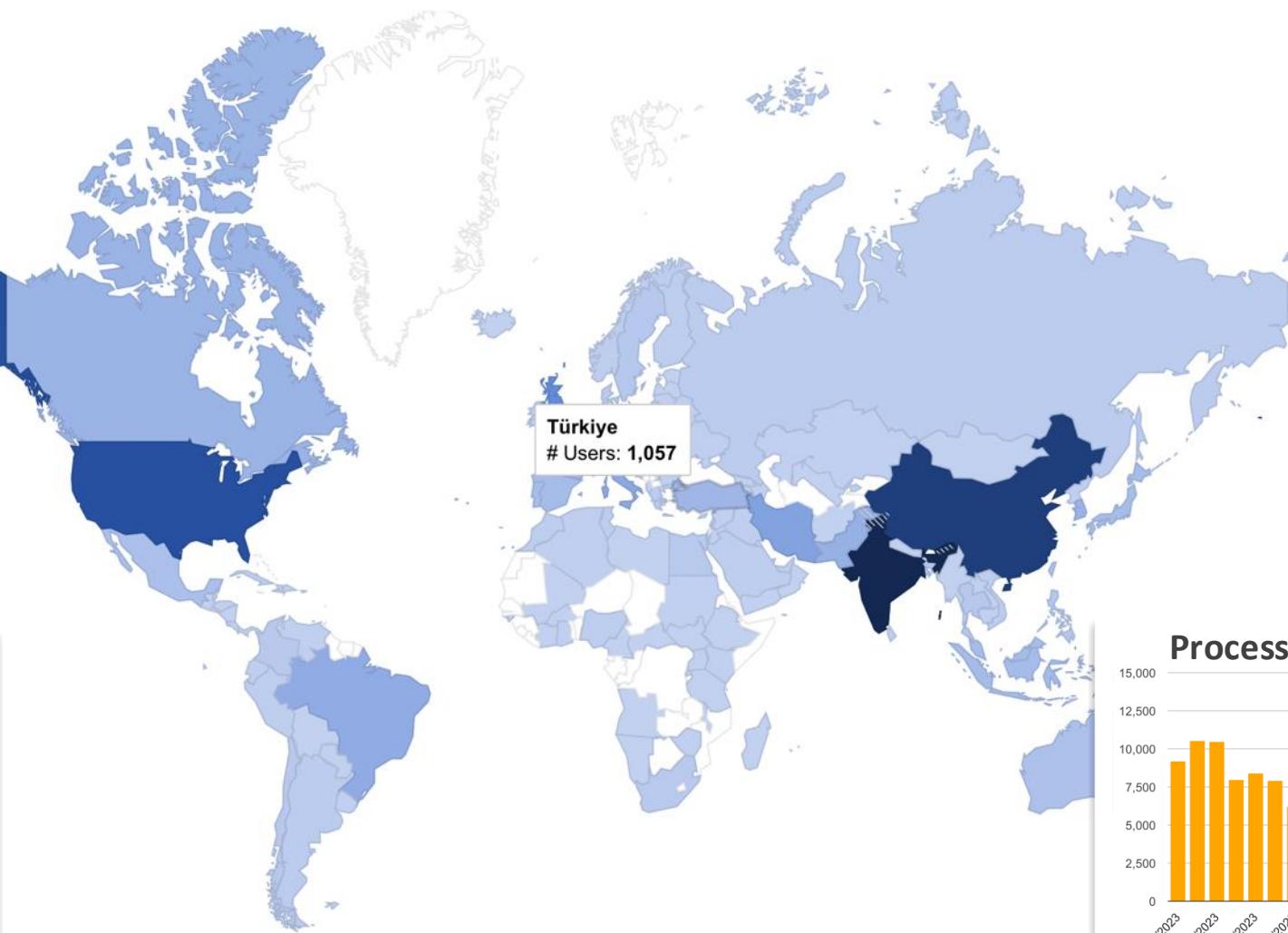
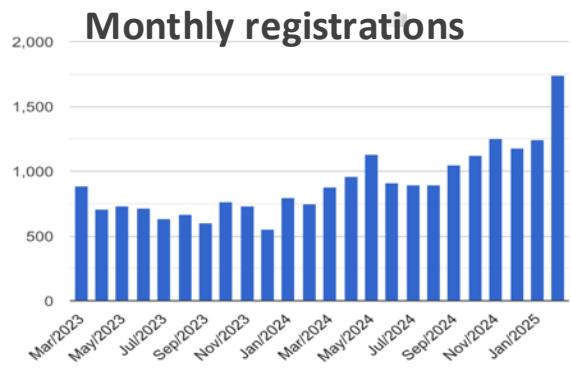
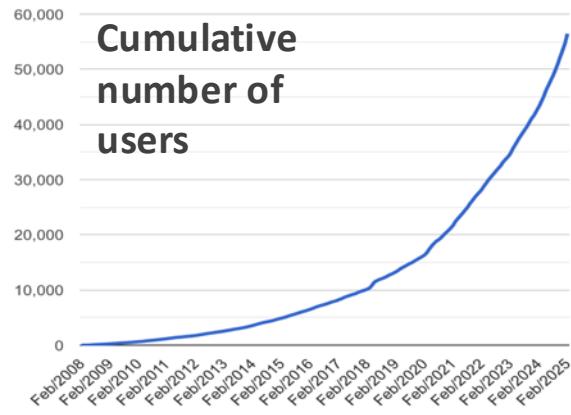


Continent	# of Users
Asia	29870
Europe	13865
North America	9163
South America	2486
Africa	1089
Oceania	682

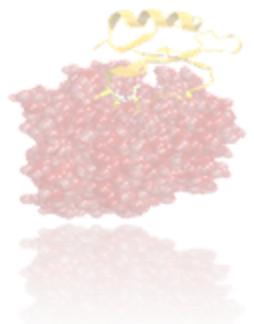
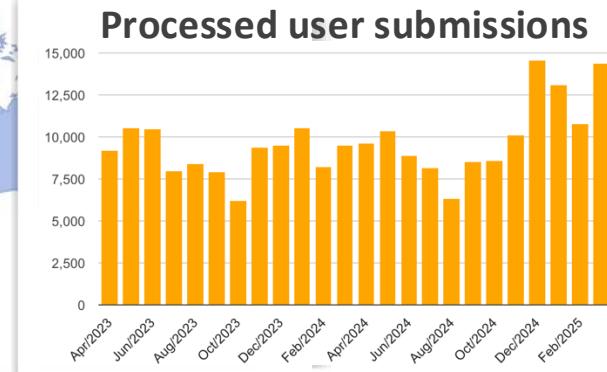


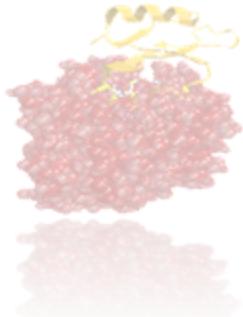
Haddock web portal

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



The HADDOCK web portal is being used by **58977 users** across **161 countries**!





HADDOCK support forum



The HADDOCK category is meant for discussing any topic related to the use of the HADDOCK software, either as a local installation or via the [HADDOCK web portal](#).

For details about HADDOCK please refer to

<https://www.bonvinlab.org/software/haddock2.4>

Feel free to create new topics related to your questions!

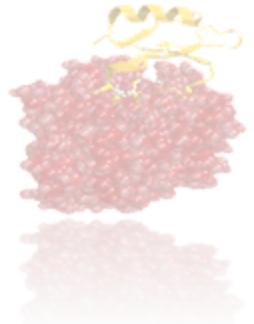
■ HADDOCK ▶ tags ▶ [Latest](#) New (1) Unread (93) Top

Topic	Replies	Views	Activity
CNS errors before/after recompilation	62	14.2k	Jul 11
Protein-ligand docking	47	8.4k	Jun 2021
Dealing with dimers	16	8.3k	May 2020
Interpretation of the Haddock results HADDOCK	7	6.5k	Sep 2021
Convert small-molecule SMILES string to pdb file for docking	1	5.8k	Jun 2018
What is, in general, a "good" score?	4	5.5k	Nov 2020

Bonvin Lab



Computational Structural Biology@Utrecht University



WHISCY

Interface prediction

CPORT

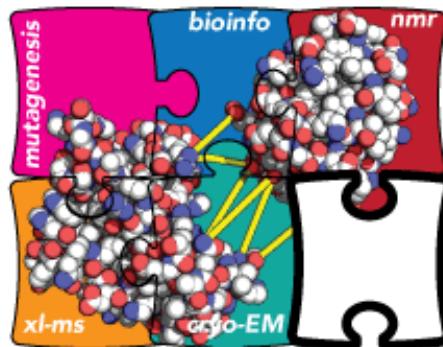
Interface prediction

Arctic3D

Known interfaces analysis

SpotON

HotSpot prediction



DisVis
Restraints analysis & visualization

Prodigy
*Affinity prediction
Crystal vs Biological*

PDB-tools
PDB files manipulation

ADDOCK
High-Ambiguity Driven Docking

PowerFit
Cryo-EM map fitting

ProABC2
Antibody paratope prediction

wenmr.science.uu.nl



Universiteit Utrecht

[Faculty of Science
Chemistry]

Overview

- **Introduction**
 - **Information-driven docking with HADDOCK**
 - **Introducing HADDOCK3**
 - **AI-based antibody-antigen modelling**
 - **Protein cyclic-peptides**
 - **Protein-glycans**
 - **Conclusions & perspectives**

Introducing HADDOCK v3



Rodrigo
Vargas Honorato



Brian Jimenez



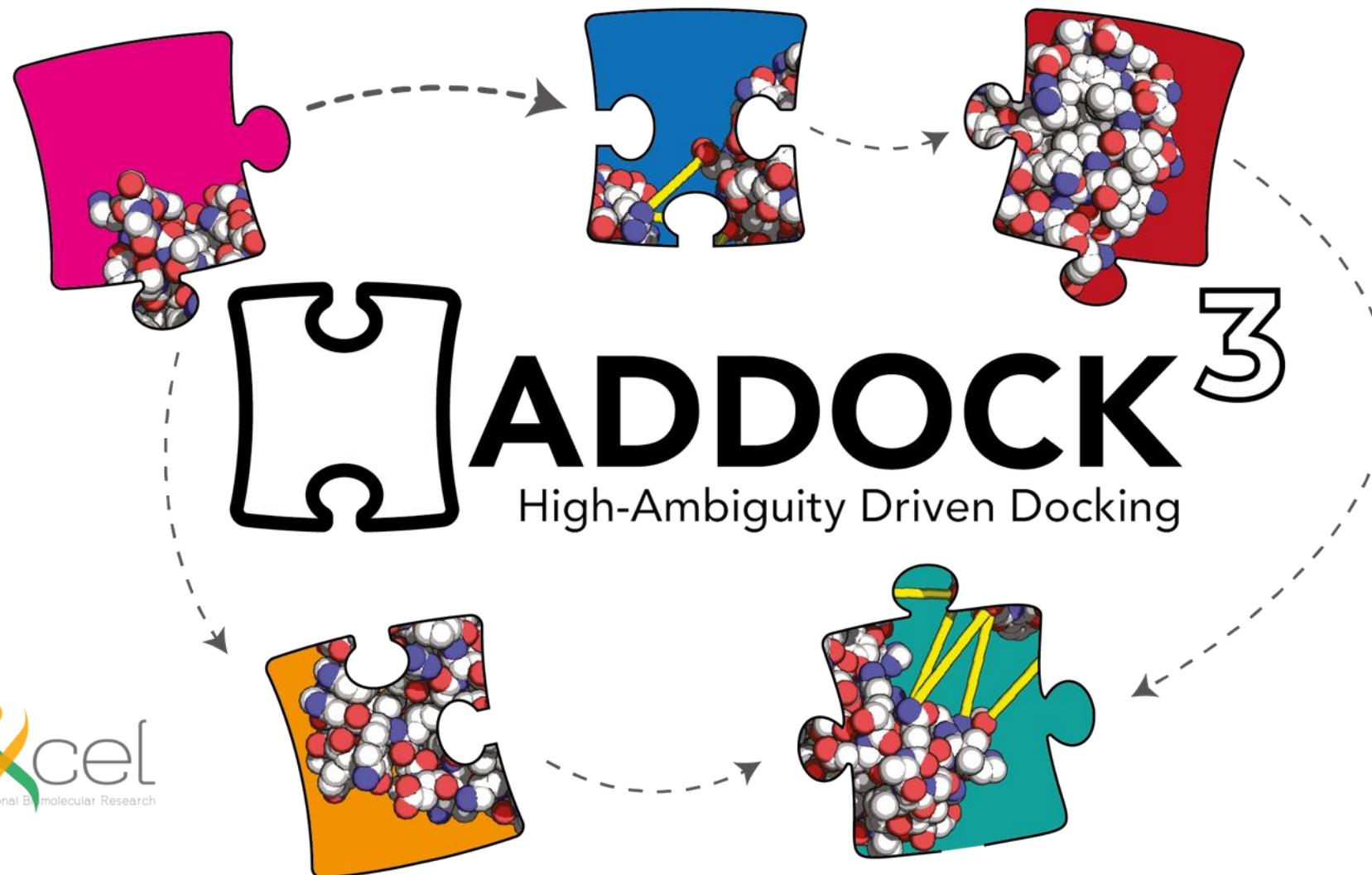
João MC Teixeira



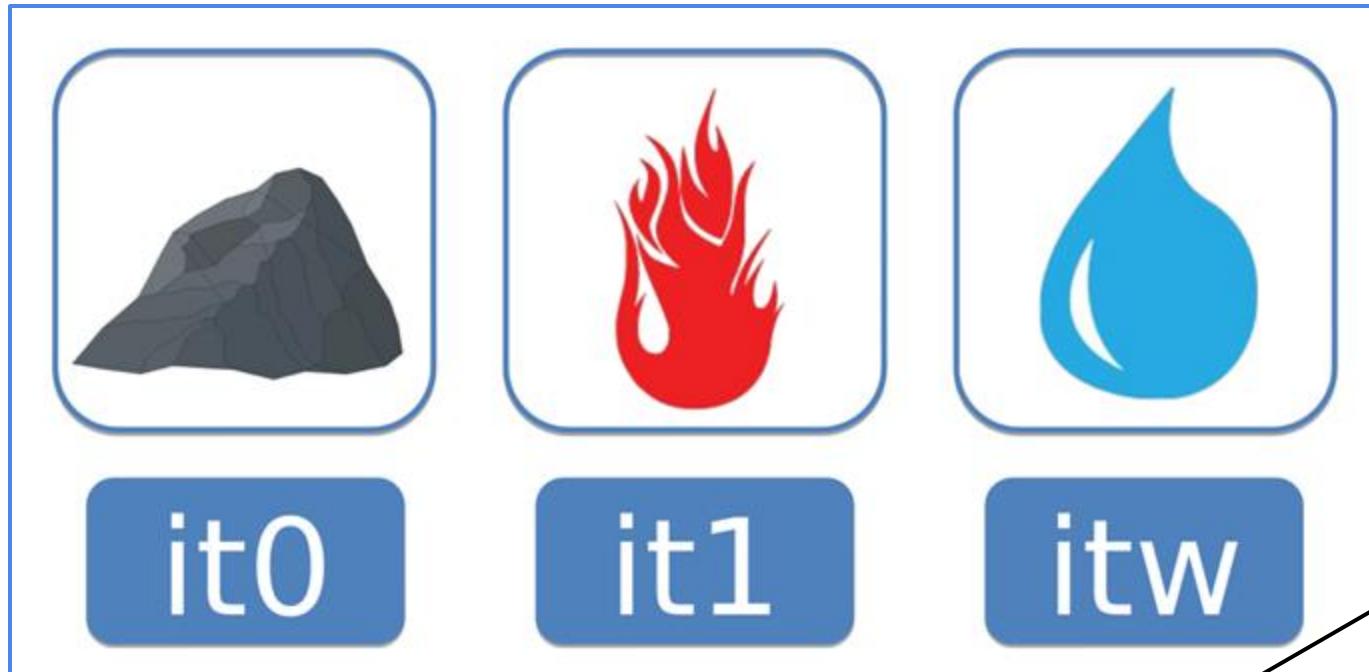
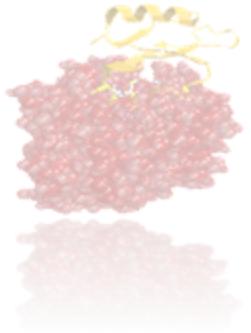
Marco Giulini



Victor Reys

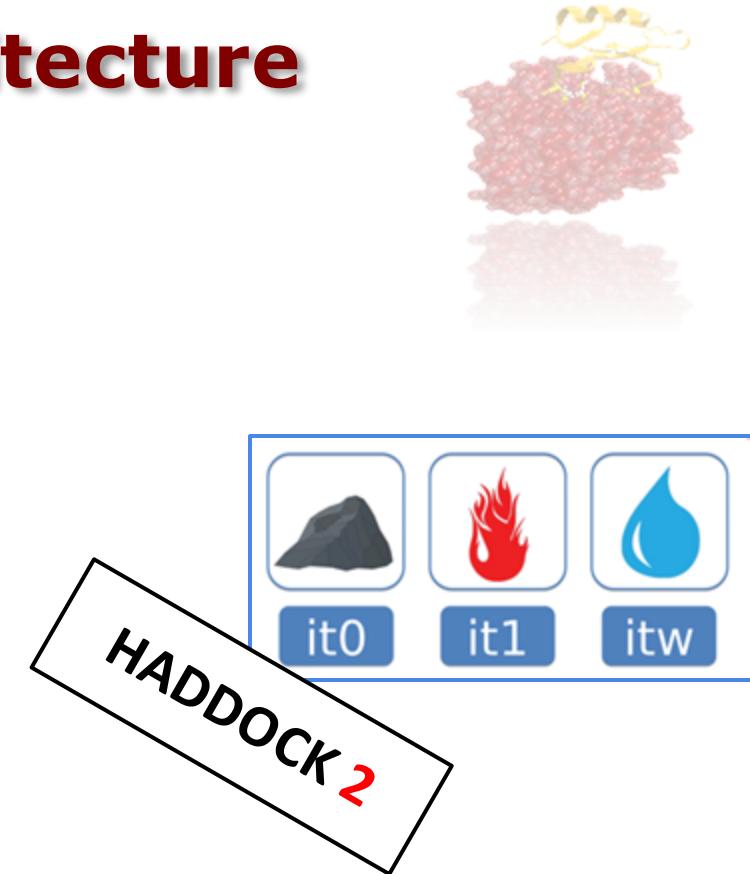
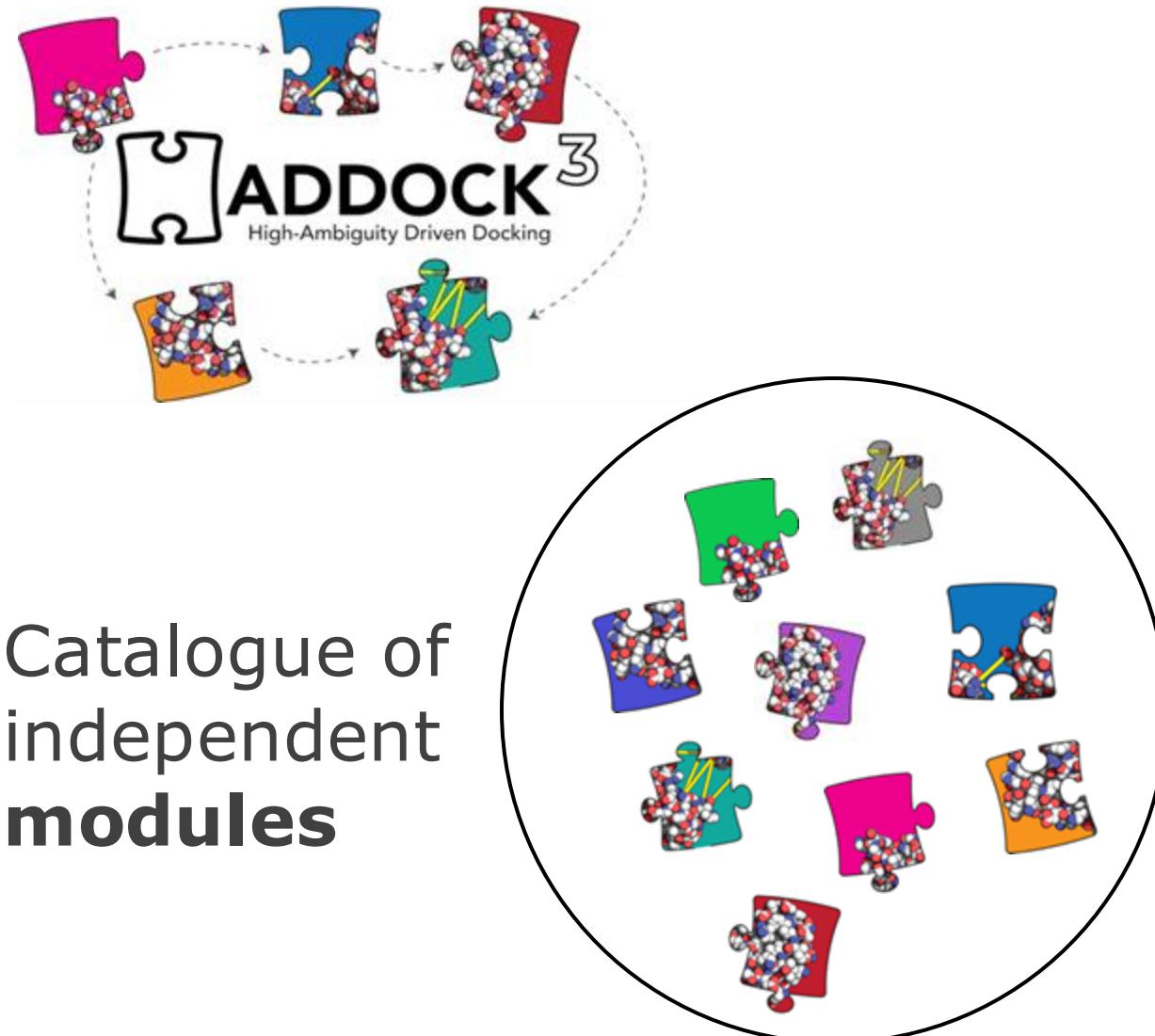


From static workflow ...

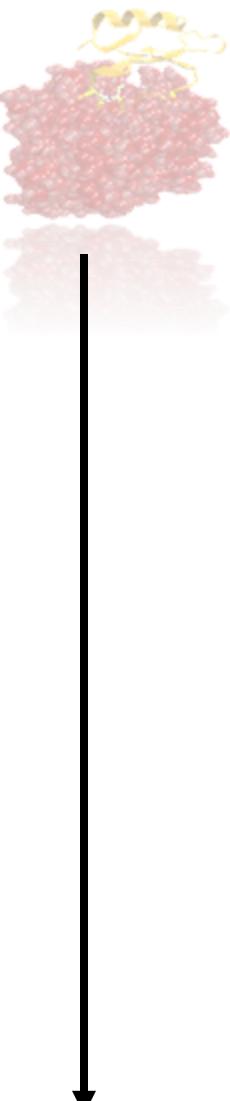
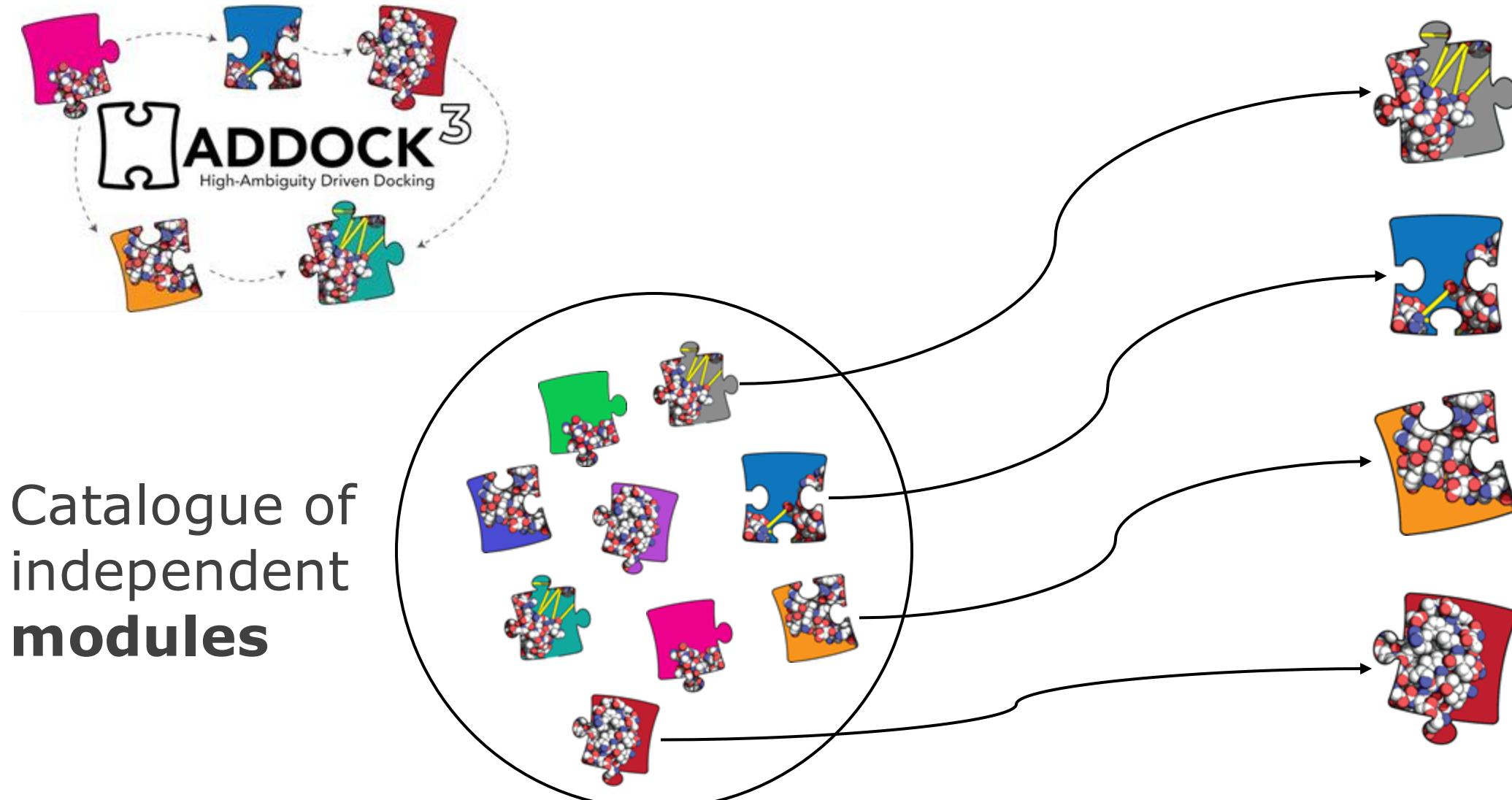


Parameterizable fixed pipeline

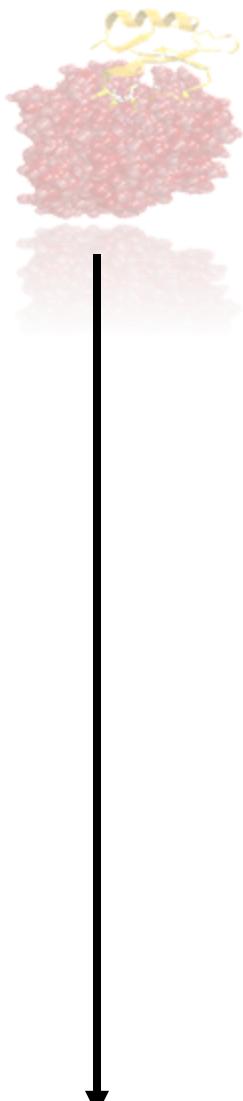
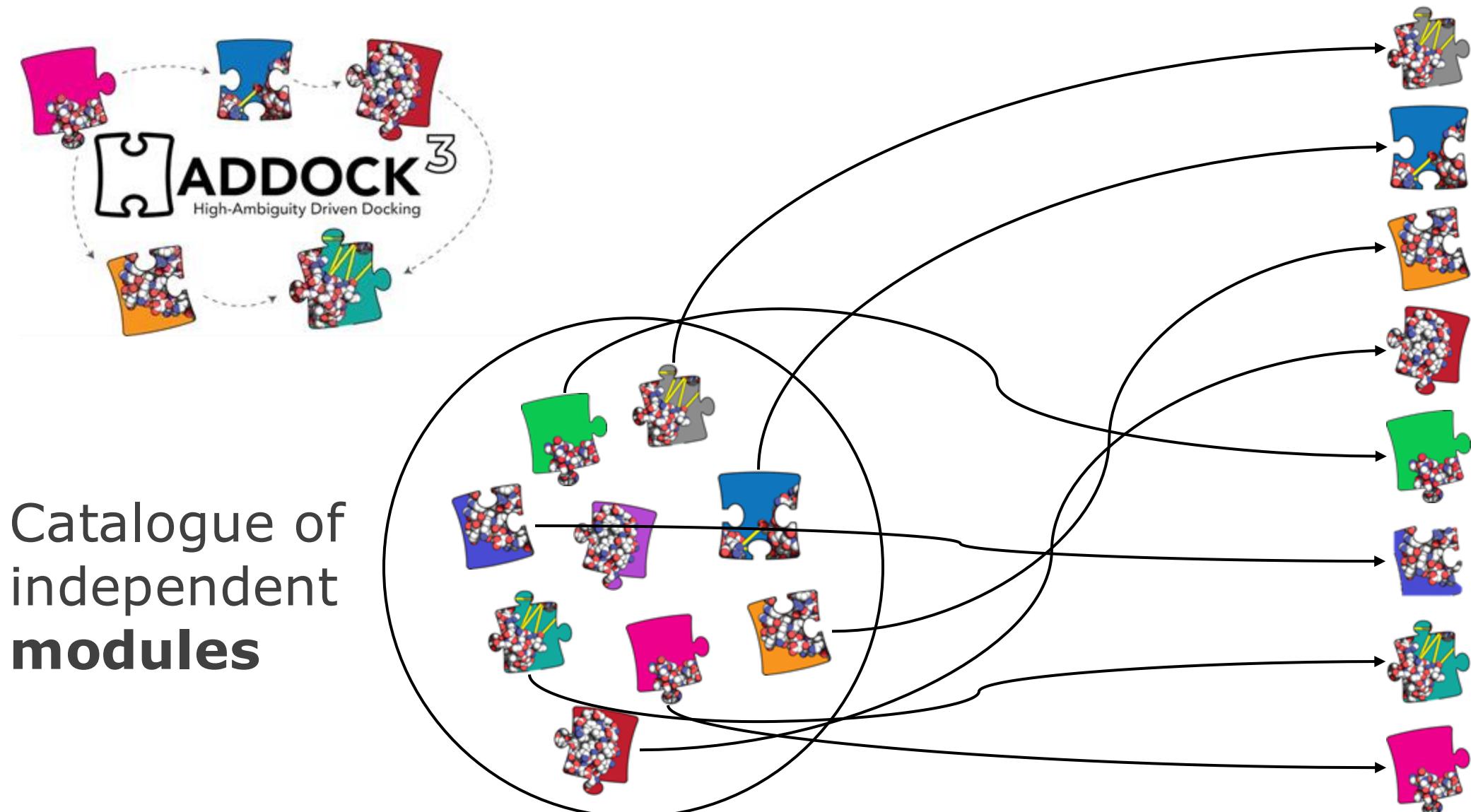
... to a modular workflow architecture

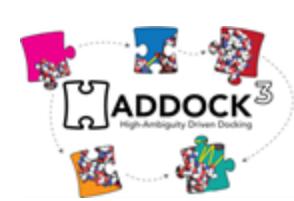


Combining the different pieces



Combining the different pieces





Current status

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]

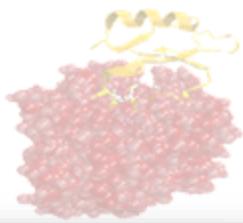
[seletopclusts]

[caprieval]

# =====
```



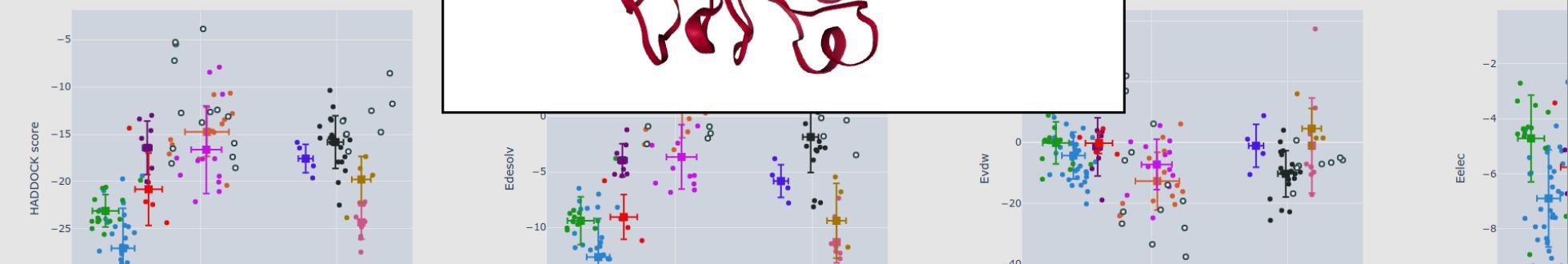
HADDOCK3 analyse report



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

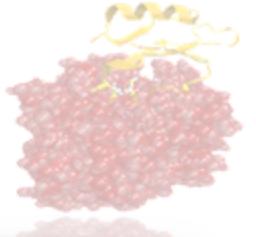
Will form the basis
for the analysis
interface

Cluster Rank	1	2	3	4	5	6	7	8	9	10
Cluster ID	1	6	3	14	5	2	13	8	4	12
Cluster size	37	9	14	4	12	19	4	7	13	4
HADDOCK score [a.u.]	-33.156 ± 0.158	-26.016 ± 0.952	-24.757 ± 0.572	-20.824 ± 3.827	-20.675 ± 1.023	-19.85 ± 1.709	-19.783 ± 2.437	-18.528 ± 1.88	-17.808 ± 1.6	-17.575 ± 1.5
interface RMSD [Å]	2.553 ± 0.168	10.876 ± 0.012	1.185 ± 0.185	3.095 ± 0.494	5.459 ± 0.376	10.151 ± 0.099	10.89 ± 0.348	3.04 ± 0.134	4.676 ± 0.853	8.851 ± 0.271
Fraction of Common Contacts	0.361 ± 0.02	0.118	cluster1_model1.pdb		Download View		Download View		Download View	
ligand RMSD [Å]	6.747 ± 0.854	18.045	Download View		Download View		Download View		Download View	
DOCKQ	0.411 ± 0.023	0.106	Download View		Download View		Download View		Download View	
Restraints Energy	123.595 ± 107.598	168.53	Download View		Download View		Download View		Download View	
Desolvation Energy	-16.131 ± 1.062	-12.22	Download View		Download View		Download View		Download View	
Electrostatic Energy	-7.352 ± 1.96	-2.177	Download View		Download View		Download View		Download View	
Van der Waals Energy	-11.774 ± 7.366	-10.93	Download View		Download View		Download View		Download View	
Nr 01 best structure	Download View	Download View	Download View		Download View		Download View		Download View	
Nr 02 best structure	Download View	Download View	Download View		Download View		Download View		Download View	
Nr 03 best structure	Download View	Download View	Download View		Download View		Download View		Download View	
Nr 04 best structure	Download View	Download View	Download View		Download View		Download View		Download View	

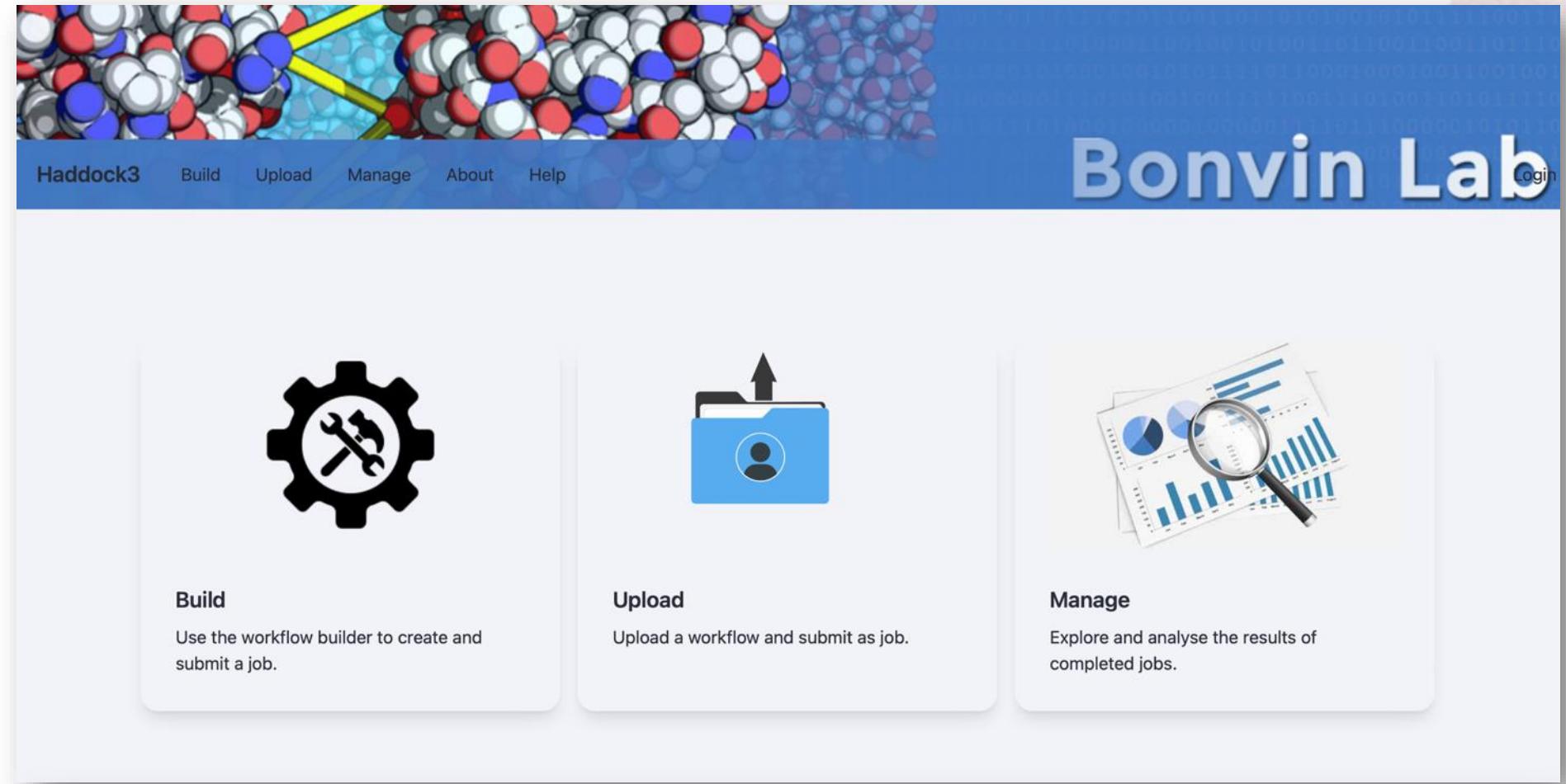


HADDOCK3 webapp

netherlands
eScience center



Will provide data management, analysis and visualization tools



Haddock3 Build Upload Manage About Help

Bonvin Lab Login

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

Upload
Upload a workflow and submit as job.

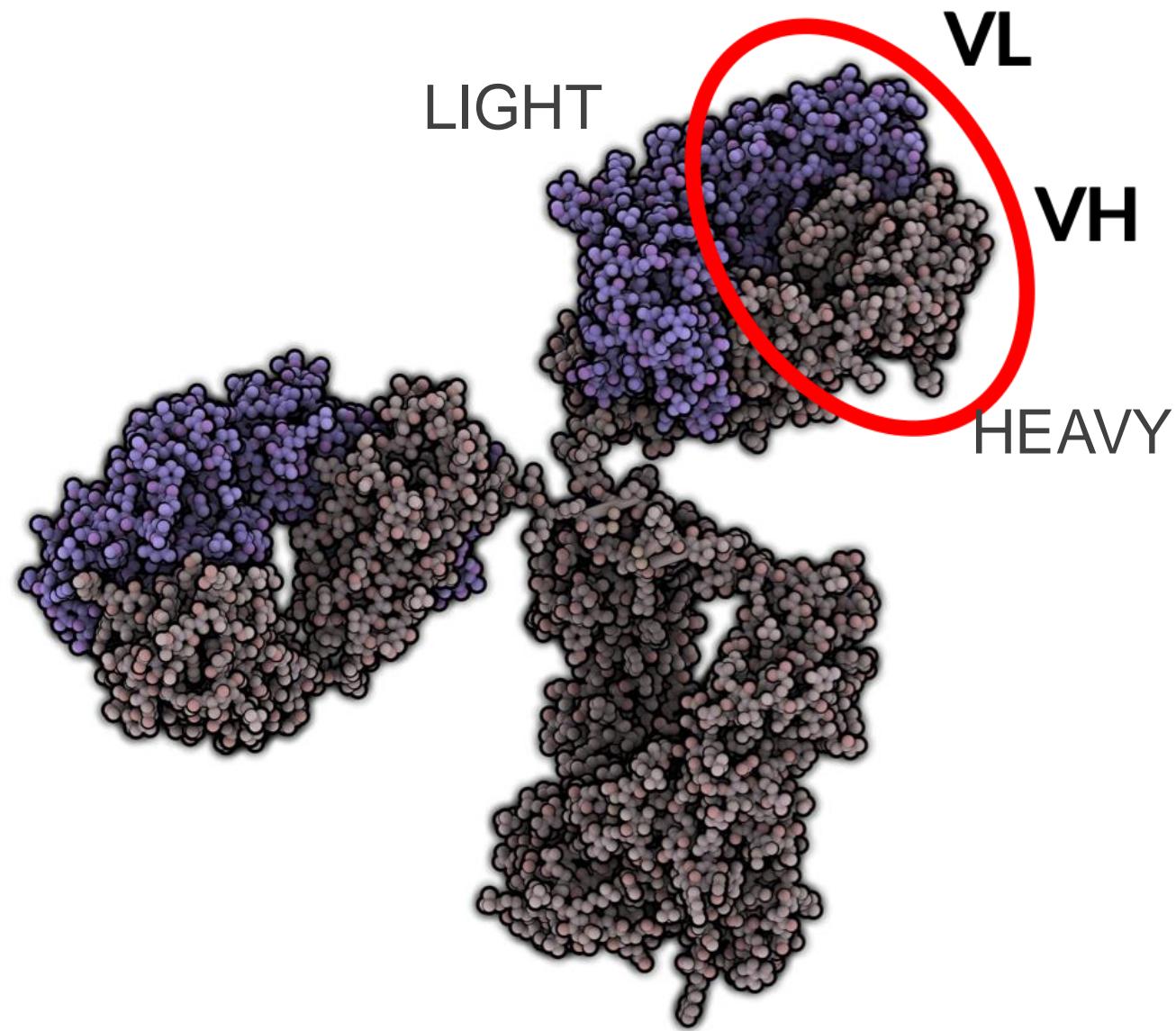
Manage
Explore and analyse the results of completed jobs.

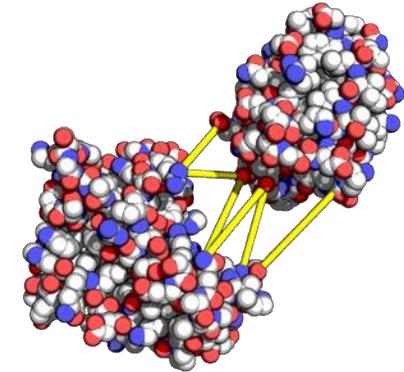
[github.com/i-VRESSE/...](https://github.com/i-VRESSE/)

Overview

- **Introduction**
 - **Information-driven docking with HADDOCK**
 - **Introducing HADDOCK3**
 - **AI-based antibody-antigen modelling**
 - **Protein cyclic-peptides**
 - **Protein-glycans**
 - **Conclusions & perspectives**

Antibody structure



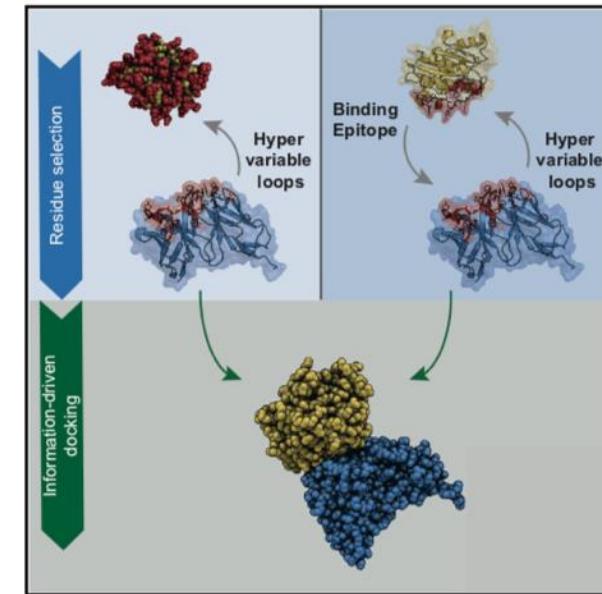


- **Using information directly to drive the modelling process improves the modelling of antibody-antigen complexes**

Structure

Modeling Antibody-Antigen Complexes by Information-Driven Docking

Graphical Abstract



Authors

Francesco Ambrosetti,
Brian Jiménez-García,
Jorge Roel-Touris,
Alexandre M.J.J. Bonvin

Correspondence
a.m.j.bonvin@uu.nl

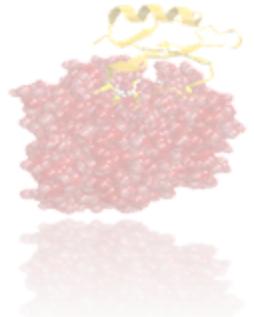
In Brief

Ambrosetti et al. demonstrate that, for the modeling of antibody-antigen complexes, using information about hypervariable loops and, when available, a loose definition of the epitope, improves the docking results. In this context, HADDOCK, which directly uses this information to guide docking, performs better than the other three software applications tested.

Highlights

- Accurate prediction of antibody-antigen structure is still a challenge
- Hypervariable loops of antibodies can be used to bias the modeling process
- Antigen epitope information, even loosely defined, is valuable for docking
- HADDOCK shows better performance and generates higher-accuracy models

Applications for anti/nanobody modelling



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>



bioRxiv

THE PREPRINT SERVER FOR BIOLOGY

Improved structural modelling of antibodies and their complexes with clustered diffusion ensembles

Marco Giulini, Xiaotong Xu, Alexandre MJJ Bonvin

doi: <https://doi.org/10.1101/2025.02.24.639865>

Information-driven docking of machine learning structure ensembles of nanobody-antigen complexes

Miguel Sánchez Marin¹, Marco Giulini¹, Alexandre M.J.J. Bonvin¹

¹ Bijvoet Centre for Biomolecular Research, Faculty of Science - Chemistry,
Utrecht University, Padualaan 8, 3584 CH, The Netherlands

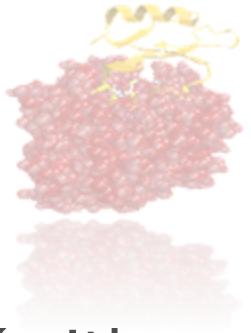
Soon on a preprint
server near you...



Universiteit Utrecht

[Faculty of Science
Chemistry]

Conclusions Antibodies



- **AI-based antibody modelling tools are accurate in generating antibodies quite close to their bound form** (with AF2 as best single method, but Immunebuilder much faster)
- **Fast HADDOCK3 protocol with very limited sampling**
- **Model diversity in the input models** (+quality of the information to guide the modelling) **is the key for (higher) HADDOCK3 docking success**
- **Scoring can still be improved**

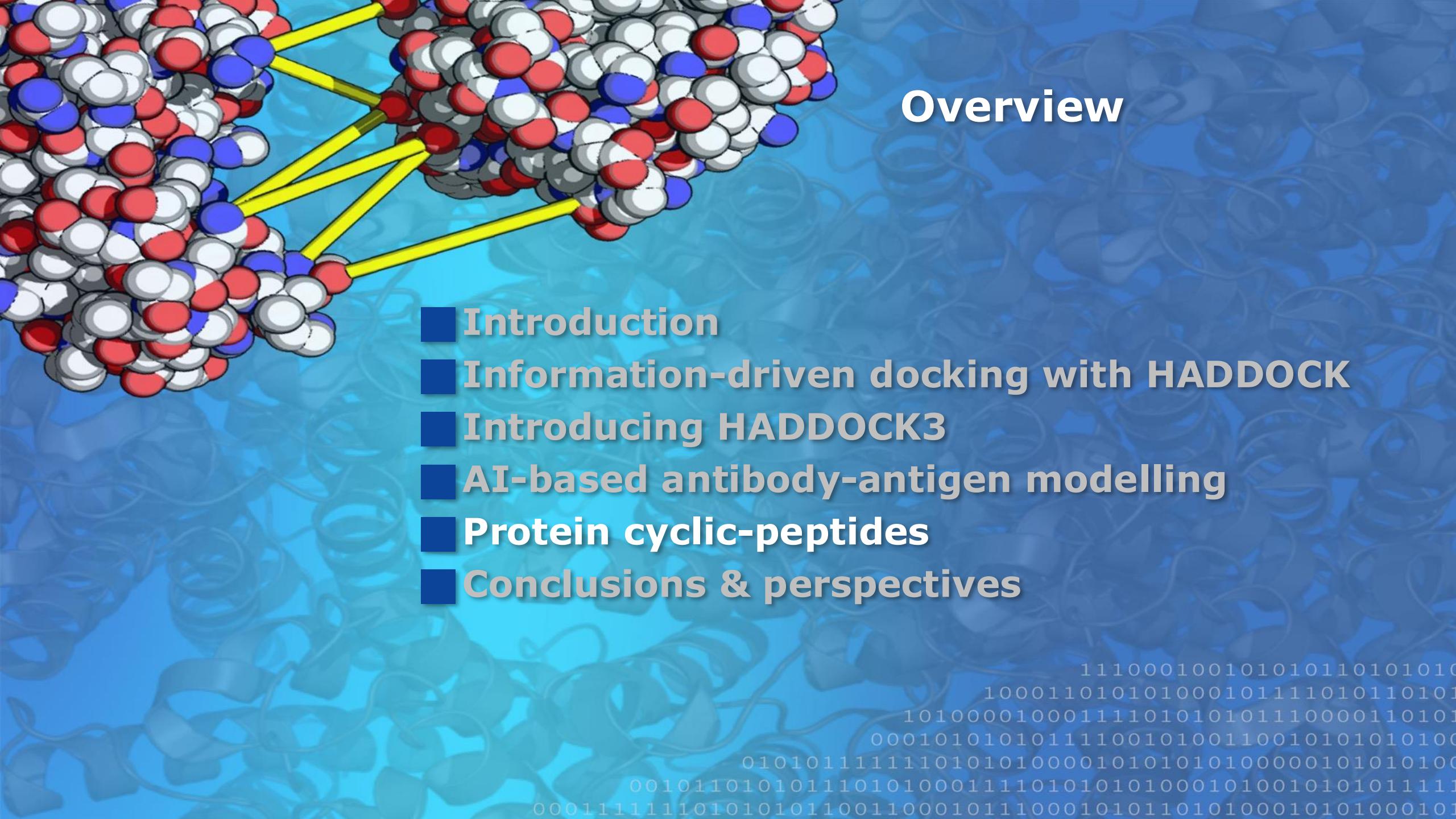
JOURNAL ARTICLE

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

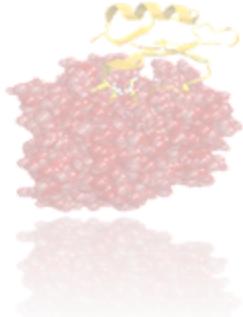
Bioinformatics, Volume 40, Issue 10, October 2024, btae583,
<https://doi.org/10.1093/bioinformatics/btae583>



Overview

- **Introduction**
- **Information-driven docking with HADDOCK**
- **Introducing HADDOCK3**
- **AI-based antibody-antigen modelling**
- **Protein cyclic-peptides**
- **Conclusions & perspectives**

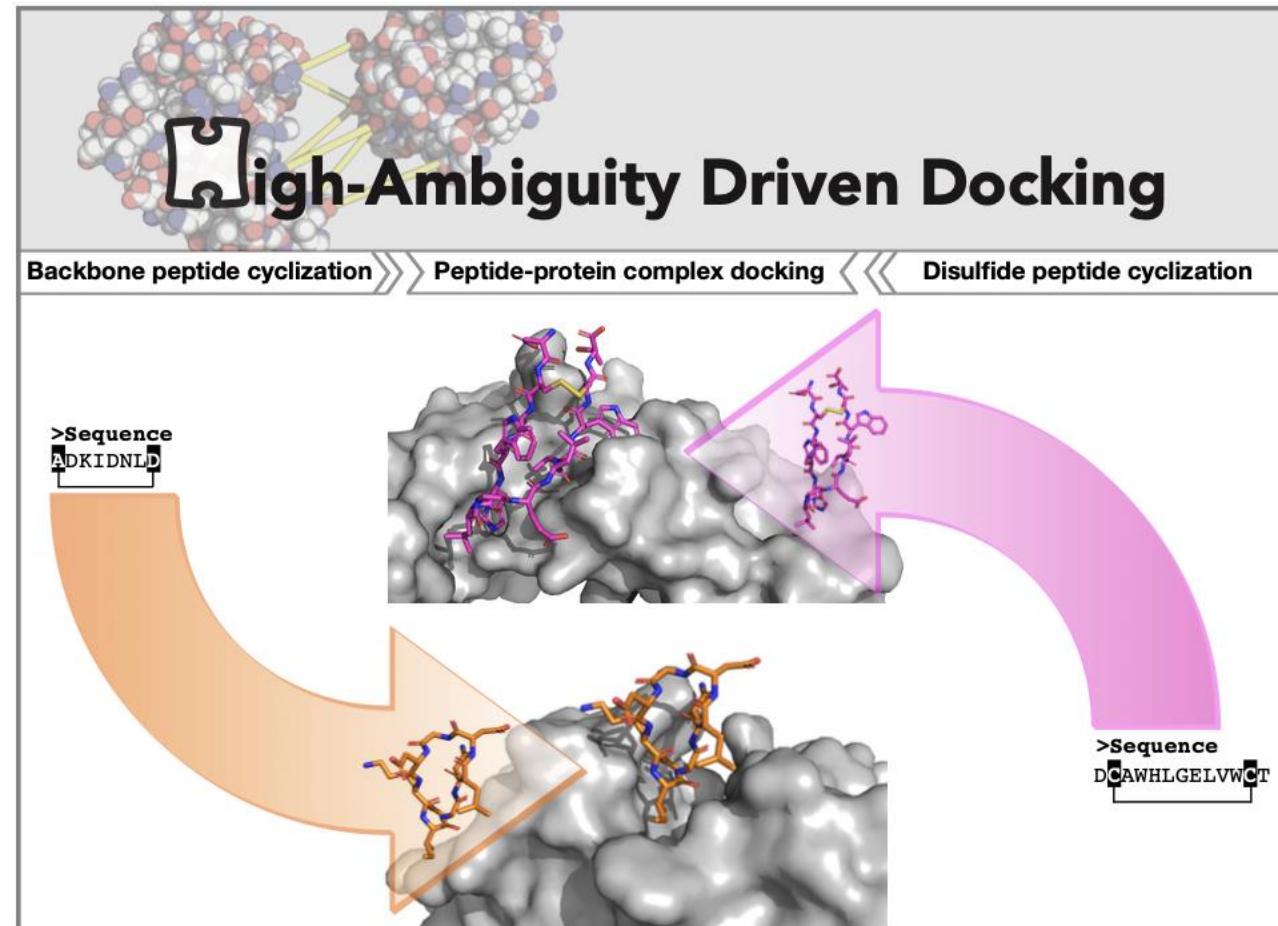
Modelling protein-cyclic peptide complexes with HADDOCK



Vicky Charitou



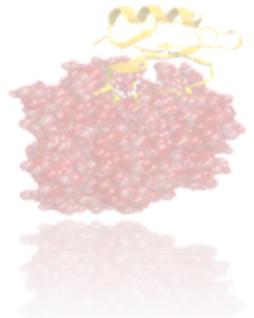
Siri van Keulen



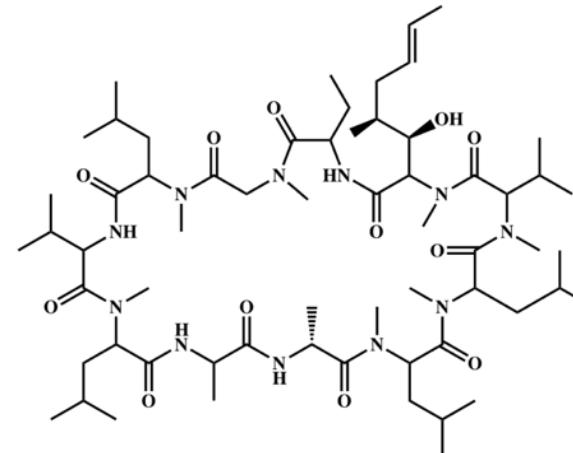
Universiteit Utrecht

[Faculty of Science
Chemistry]

Cyclic Peptides : Why?



- Promising therapeutic molecules
- Enhanced metabolic stability
- Better oral bioavailability
- E.g. cyclosporine A
- Cyclisation and conformational sampling challenge
- Only limited computational studies so far, e.g.



JCTC
Journal of Chemical Theory and Computation

 Cite This: *J. Chem. Theory Comput.* 2019, 15, 5161–5168

Letter

pubs.acs.org/JCTC

Docking Flexible Cyclic Peptides with AutoDock CrankPep

Yuqi Zhang[◎] and Michel F. Sanner^{*◎}

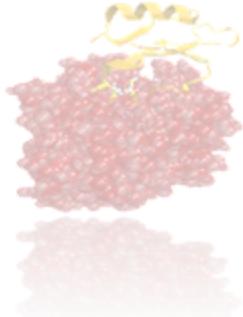
Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, California 92037, United States



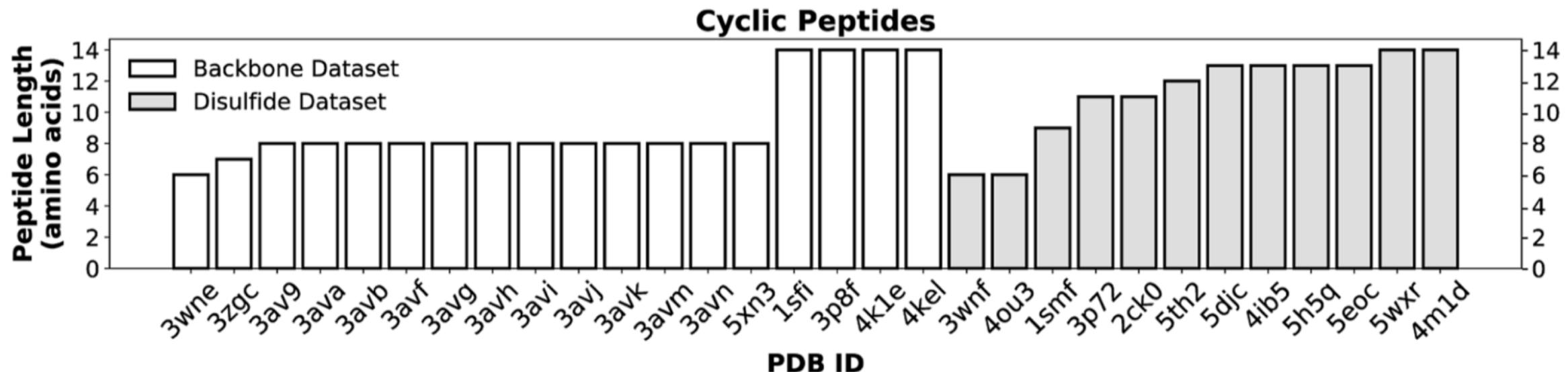
Universiteit Utrecht

[Faculty of Science
Chemistry]

Benchmark dataset

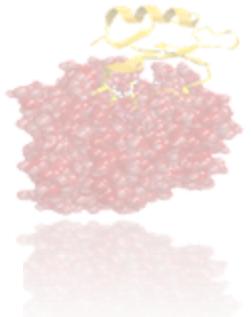


- **18 cyclic peptides (14 short \leq 10 a.a.; 4 long 14 a.a)**
- **12 disulfide peptides (3 short, 9 long)**

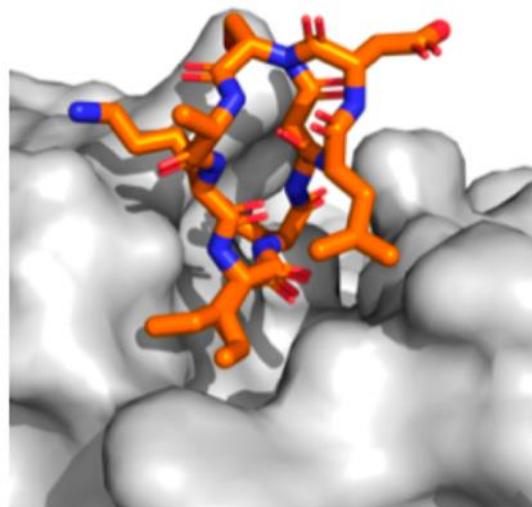


<https://github.com/haddocking/cyclic-peptides>

Benchmark dataset

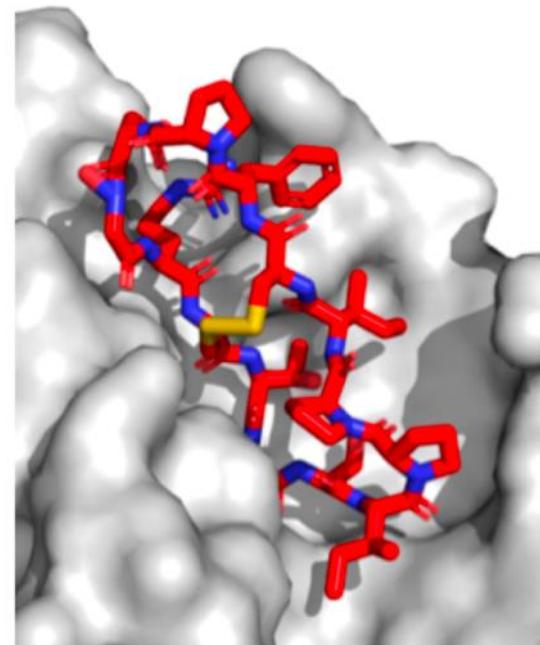


Short-backbone peptides
6-10 residues



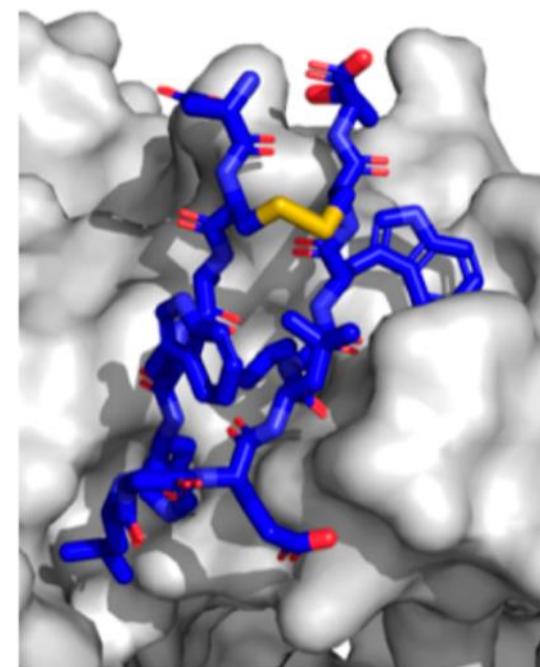
PDB ID: 3av9 (8 aa)

Long - backbone peptides
14 residues



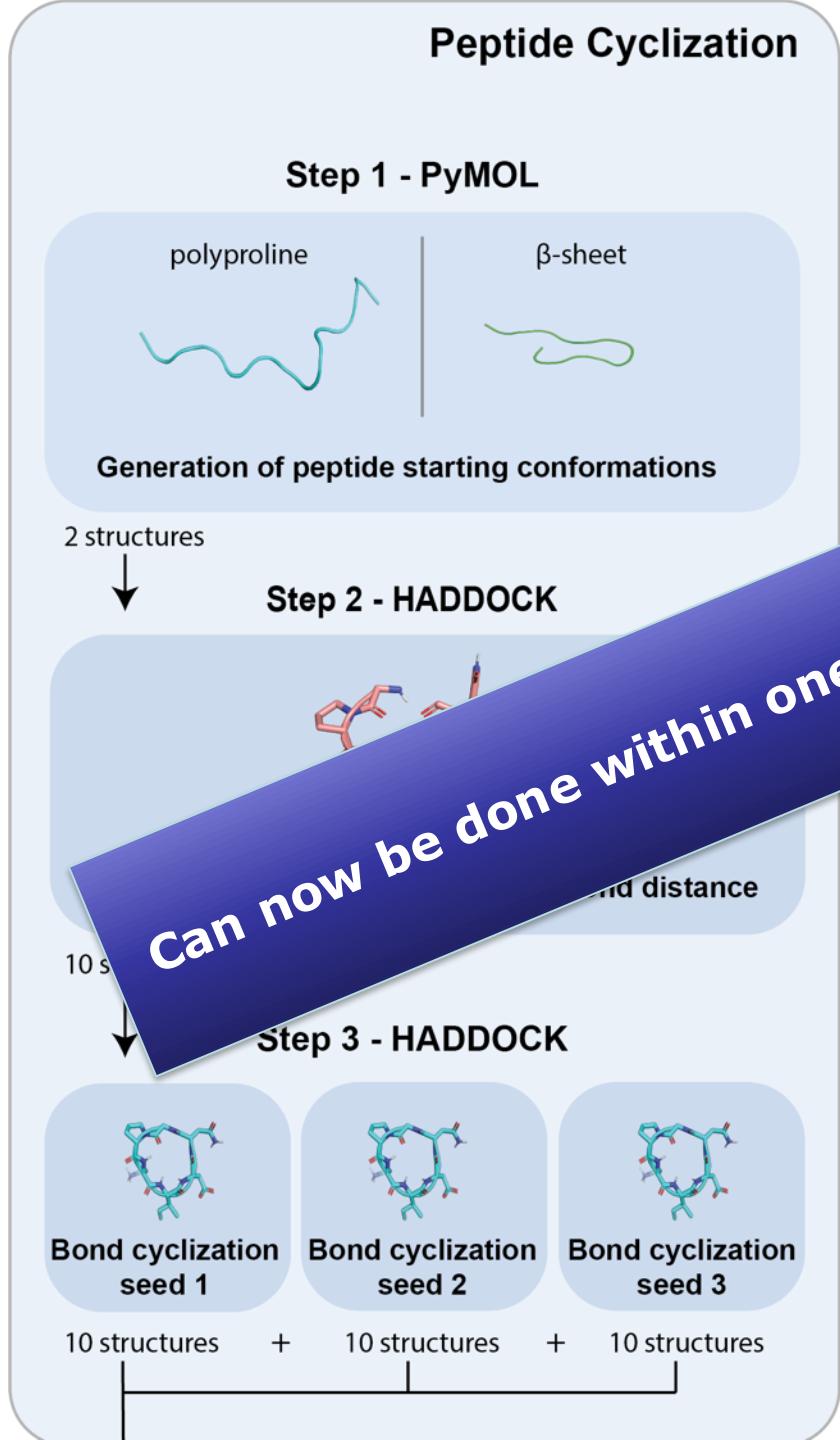
PDB ID: 3p8f (14 aa)

Disulfide peptides
6-14 residues

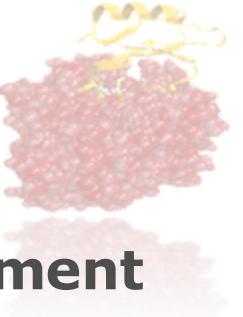


PDB ID: 5djc (13 aa)

Peptide cyclisation protocol

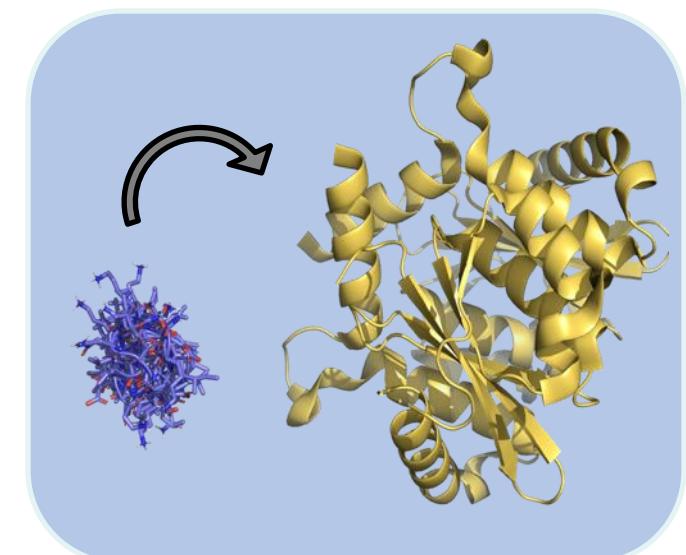


- **Based on the β -sheet stage** (fully automated, in angle space, full increased sampling)
- **Step2: Distance restraints to bring the “termini” together** (400 models, no electrostatics, RMSD clustering)
- **Step3: Covalent bonds created automatically by HADDOCK** (protocol repeated 3 time with different random seeds, RMSD clustering)



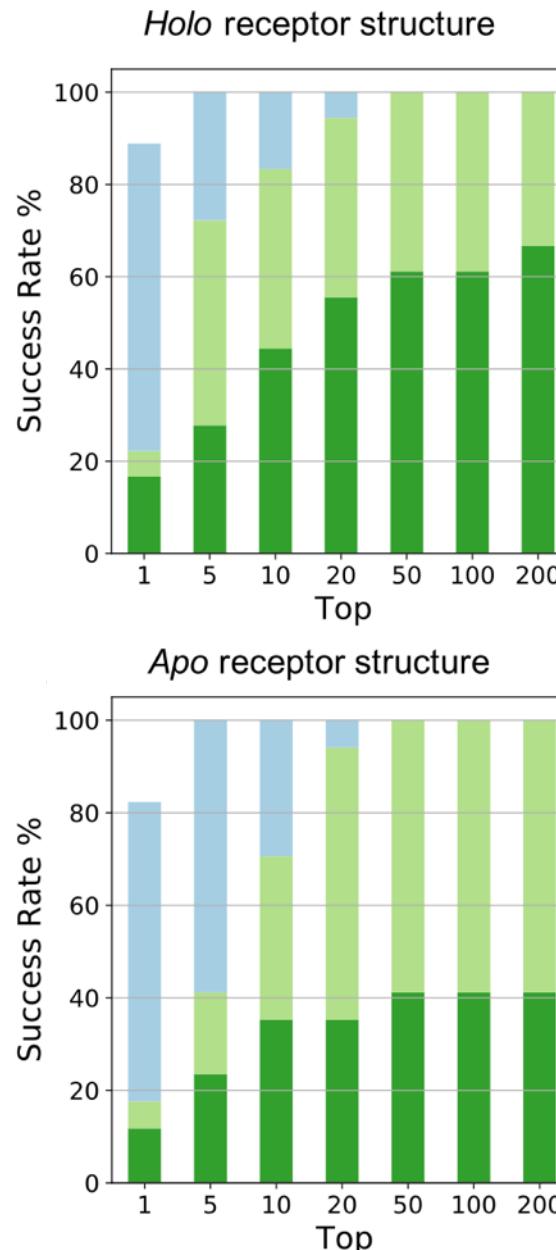
Ensemble docking protocol

- Optimized HADDOCK 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

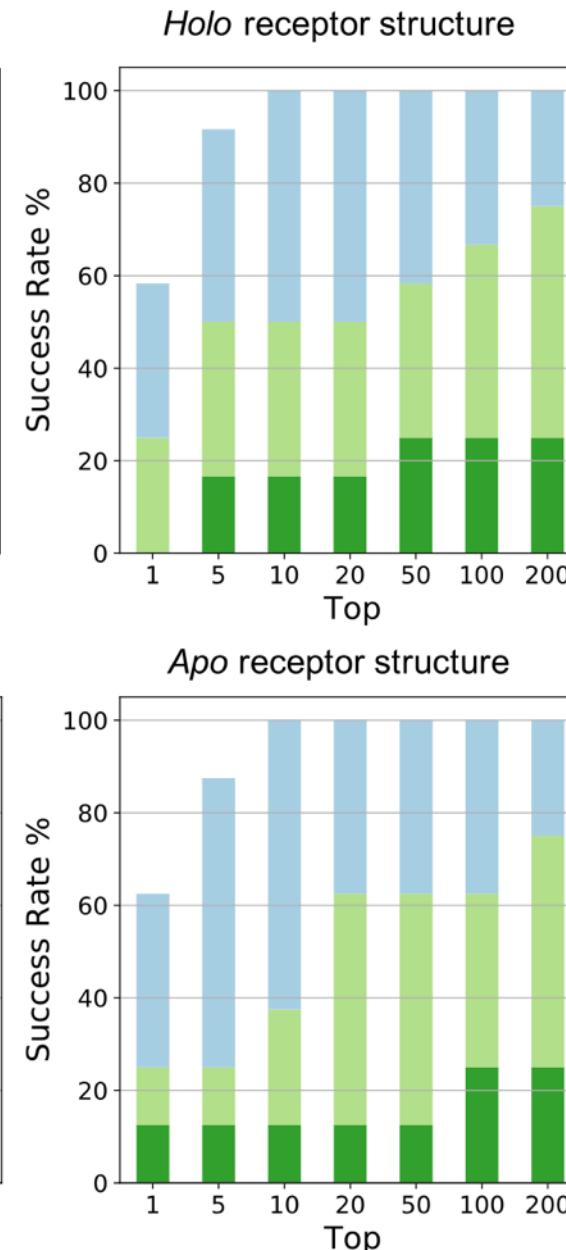


Overall performance

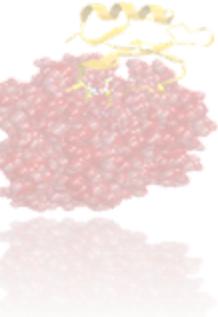
Backbone dataset (#18)



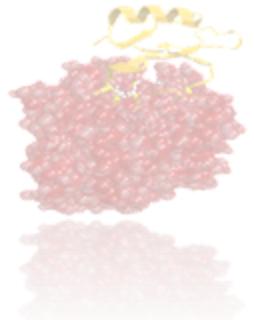
Disulphide dataset (#12)



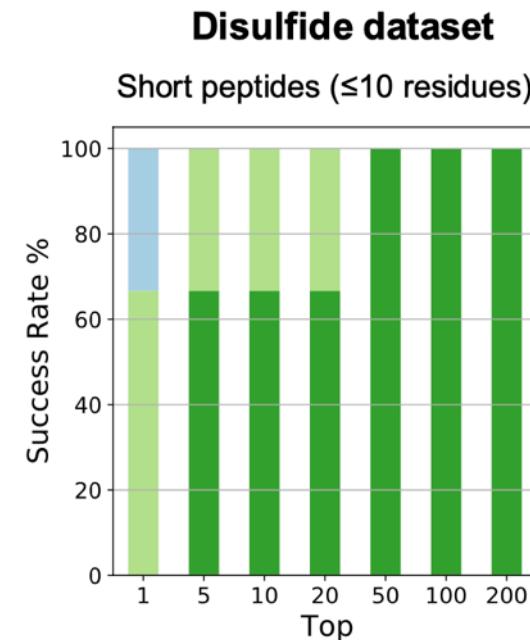
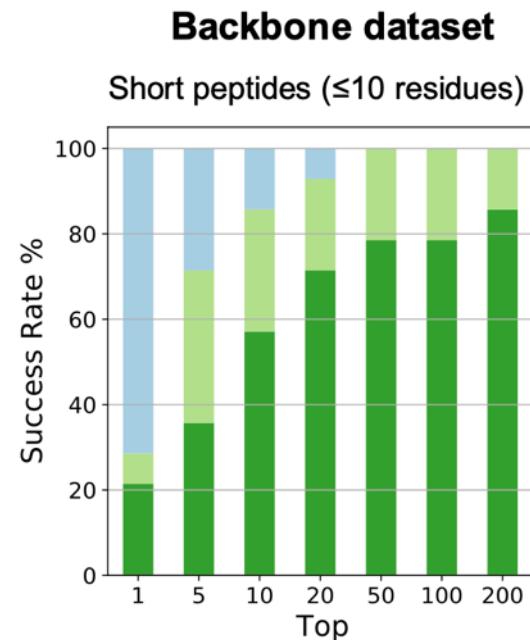
fnat ≥ 0.8
fnat ≥ 0.5
fnat ≥ 0.2



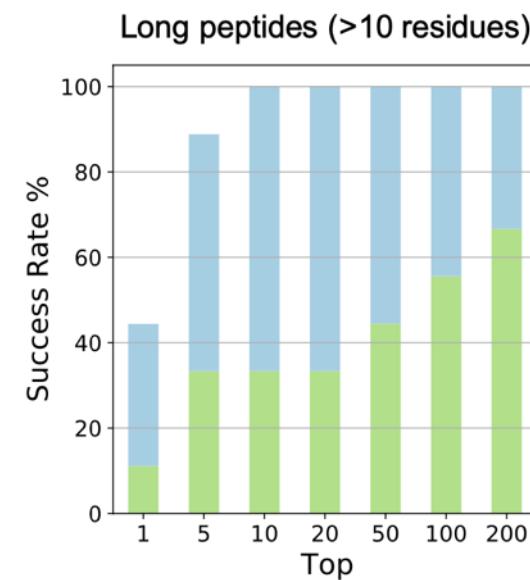
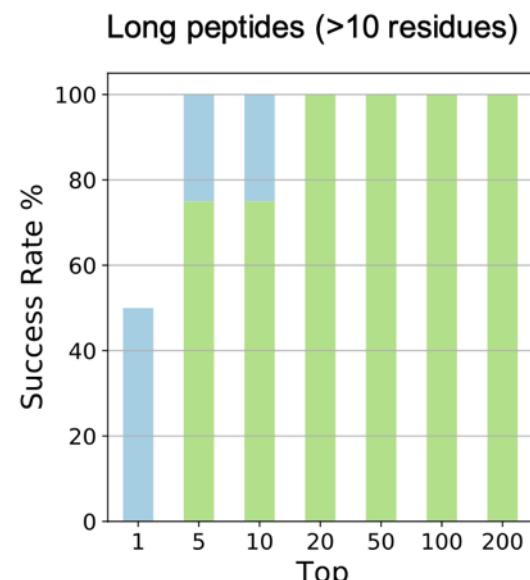
Performance vs peptide length



<10 a.a
(#10)



>10 a.a
(#4)



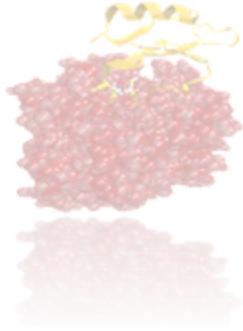
<10 a.a
(#3)

>10 a.a
(#9)

- 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



Universiteit Utrecht

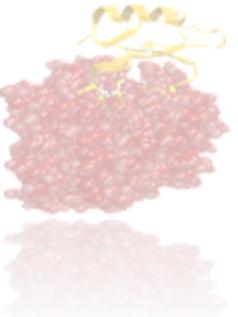
[Faculty of Science
Chemistry]



Overview

- Integrative modelling with HADDOCK
 - Introducing HADDOCK3
 - AI-based antibody-antigen modelling
 - Protein-cyclic peptides
 - Protein-glycans
 - Conclusions & perspectives

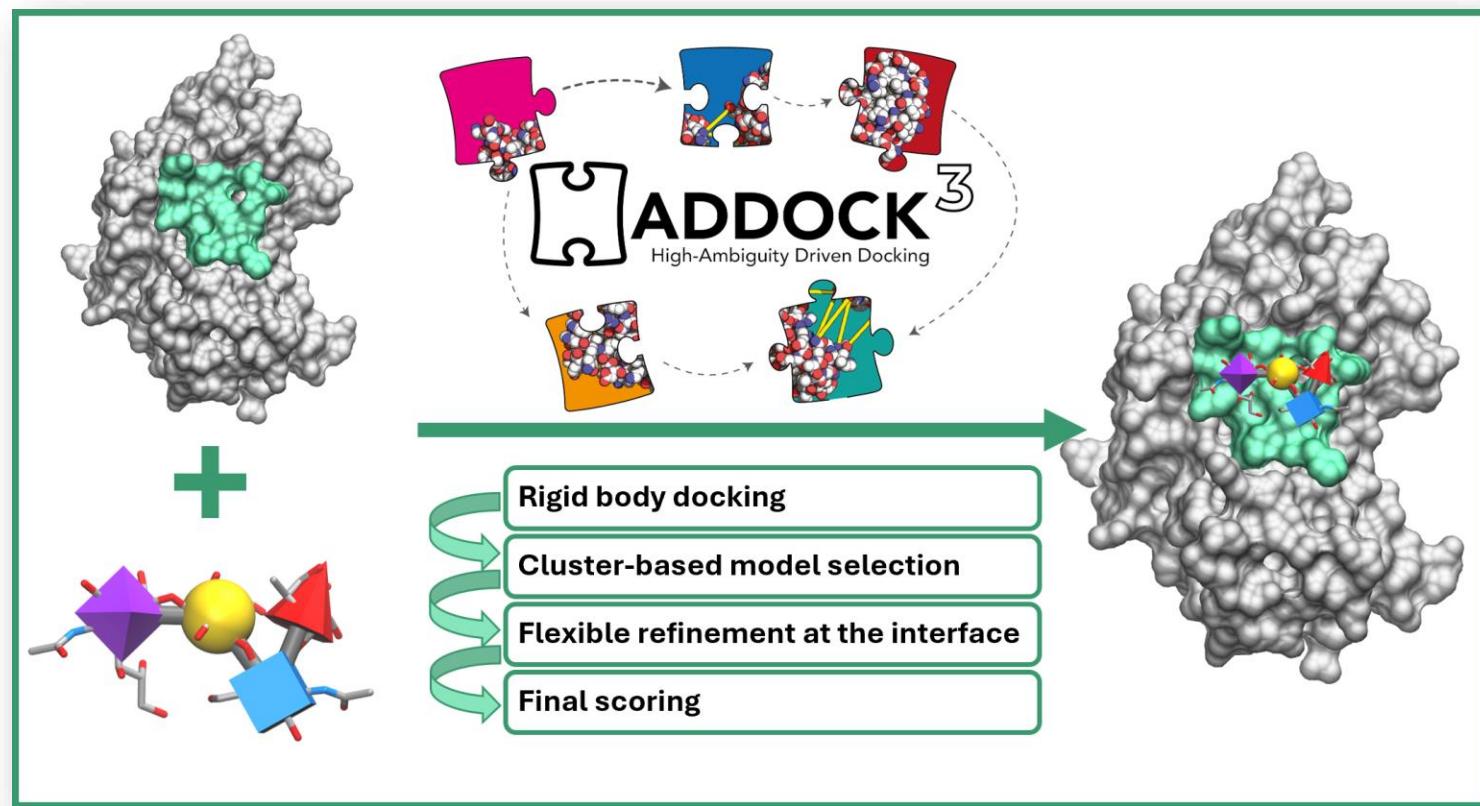
Modelling protein-glycan complexes with HADDOCK



Anna Ranaudo



Marco Giulini

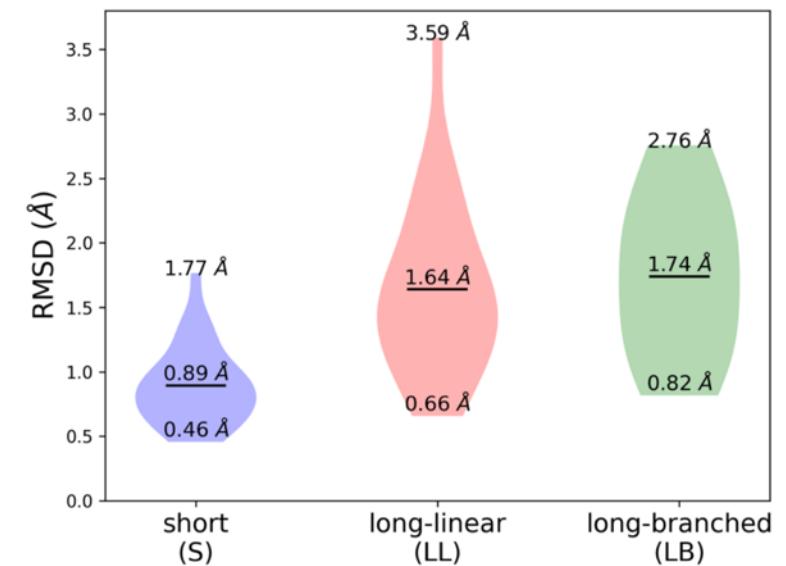
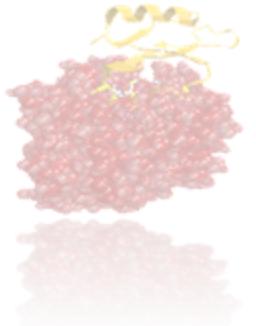


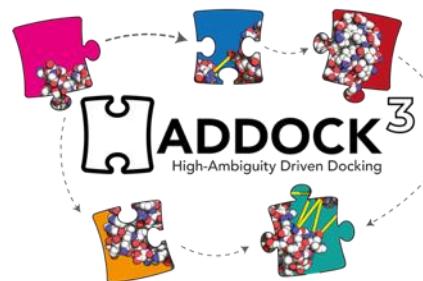
Universiteit Utrecht

[Faculty of Science
Chemistry]

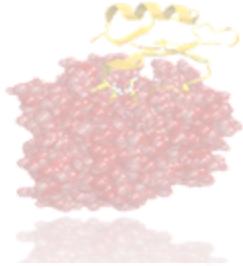
Dataset

- **89 non-redundant bound protein-protein glycan complexes**
- **55 unbound case** (i.e. apo-protein structure available)
 - **47 linear and 8 branched glycans**
 - **25 glycans composed of three or fewer monosaccharide units**
 - **30 with more than three units**
- **<https://github.com/haddocking/protein-glycans>**
- **Glycans conformations generated with the GLYCAM server (<https://glycam.org>)**





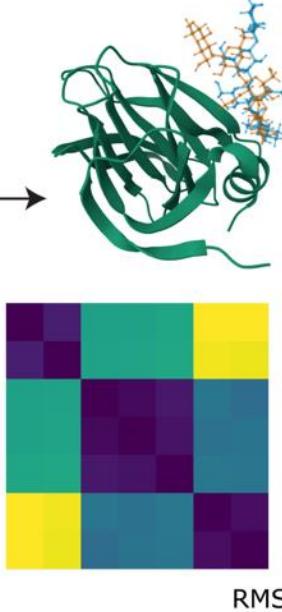
protocol



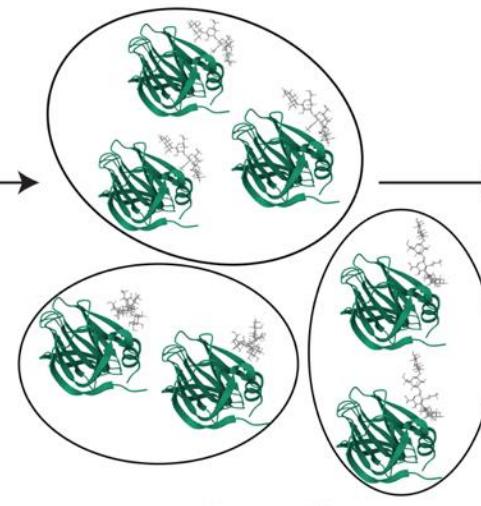
Rigid body docking



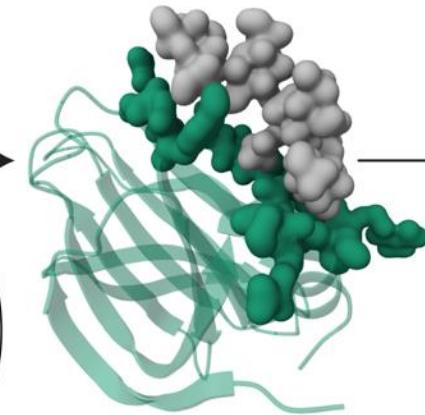
RMSD matrix calculation



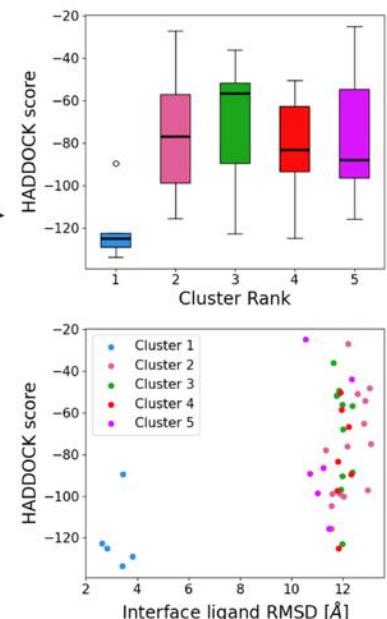
Cluster-based model selection



Flexible refinement at the interface



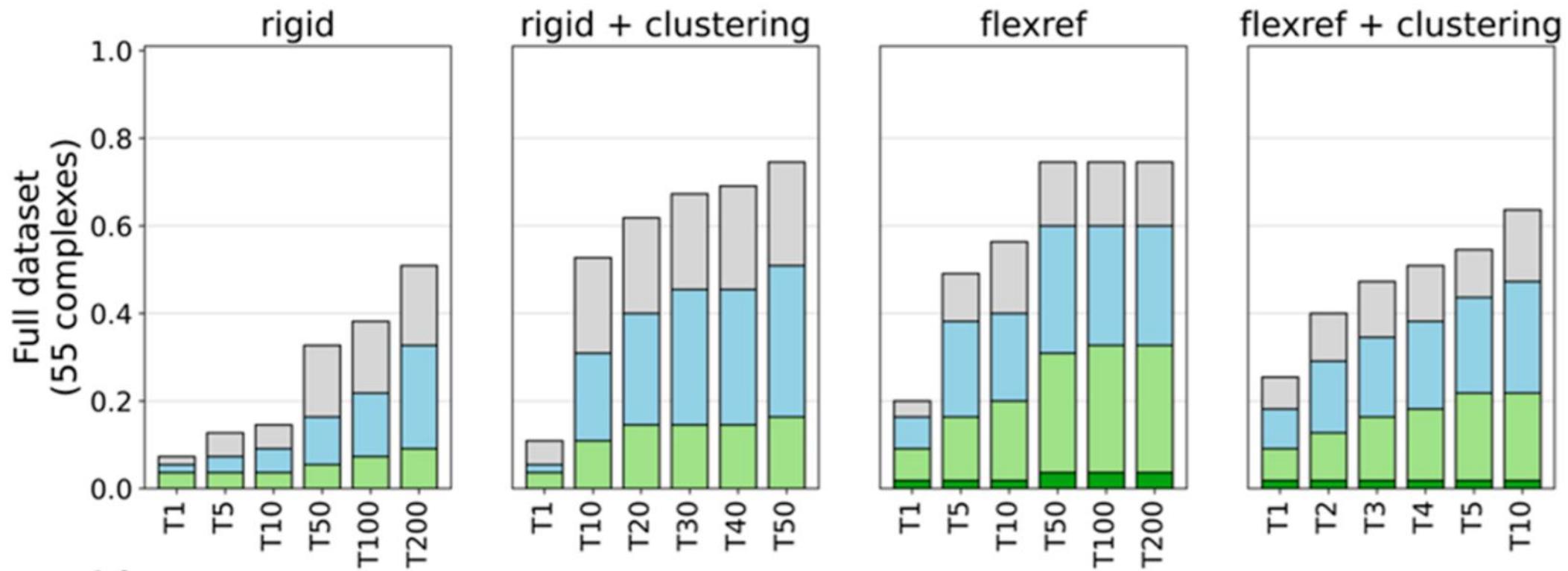
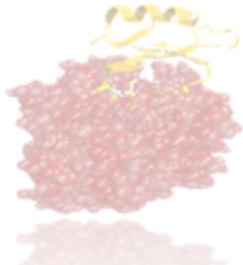
Cluster-based scoring



Weight of van der Waals energy term increased to 1 for rigidbody



Unbound results



Near acceptable ($3 \text{ \AA} < \text{IL-RMSD} \leq 4 \text{ \AA}$)

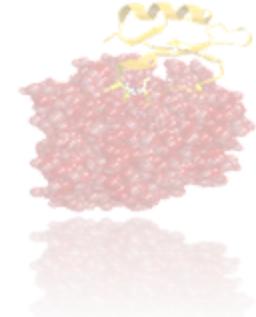
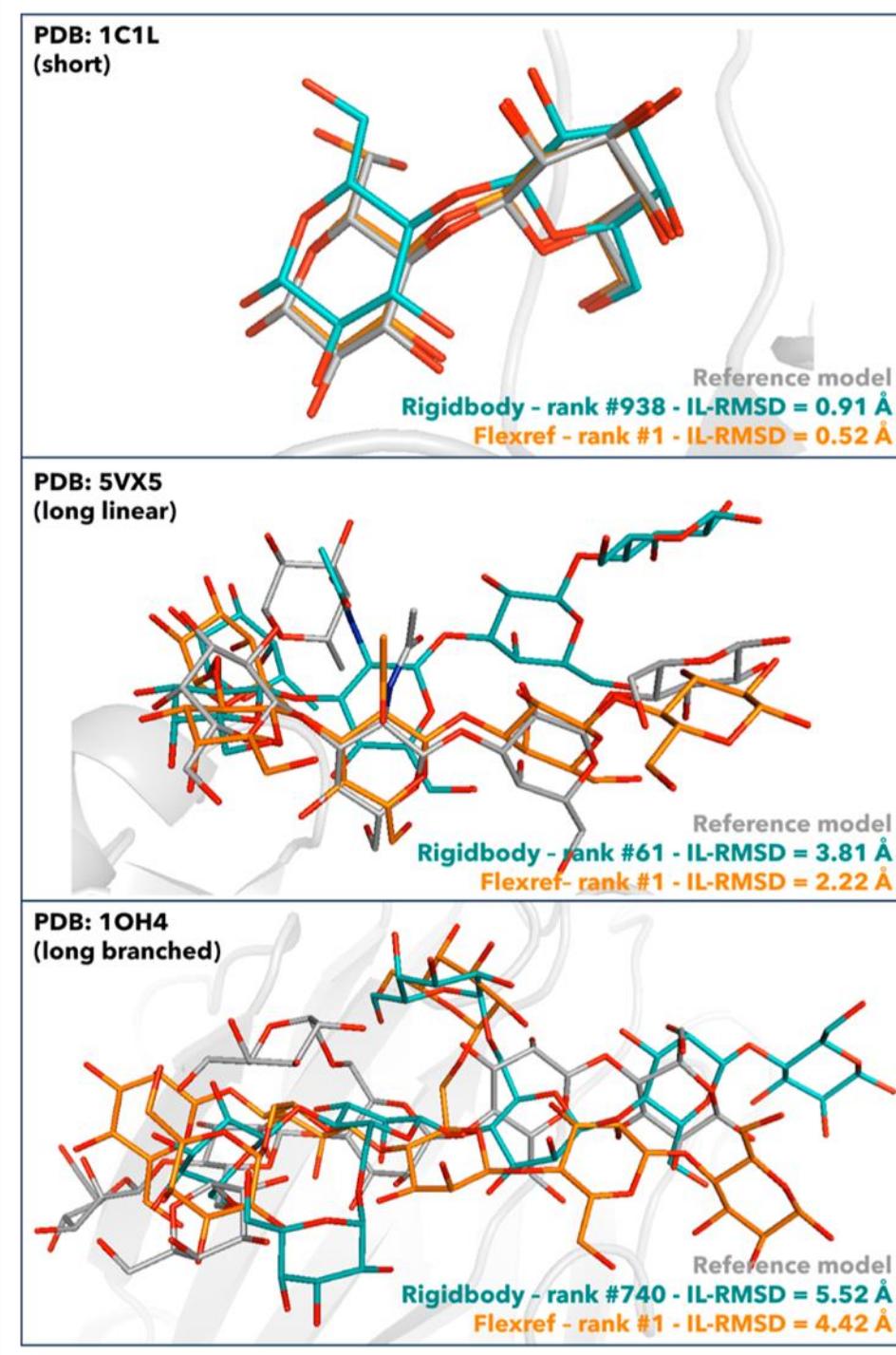
Acceptable ($2 \text{ \AA} < \text{IL-RMSD} \leq 3 \text{ \AA}$)

Medium ($1 \text{ \AA} < \text{IL-RMSD} \leq 2 \text{ \AA}$)

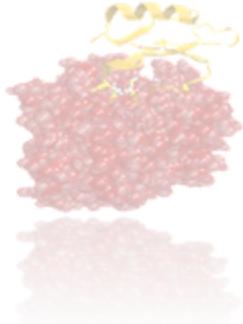
High ($\text{IL-RMSD} \leq 1 \text{ \AA}$)

Impact of the flexible refinement

- Improves both conformation (lower RMSDs) and scoring



Conclusions



- **New HADDOCK3 protocol with clustering after rigidbody increases the overall performance** (from 33% to 51% selected acceptable models)
- **Flexible refinement improves both ranking and model quality**
- **There is still room for improvement** (Top 5 performance after flexible refinement is 44%)
- **Long and branched glycans remain challenging to model**

JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

Open Access

This article is licensed under CC-BY 4.0

Article

pubs.acs.org/jcim <https://pubs.acs.org/jcim?ref=pdf>

Modeling Protein–Glycan Interactions with HADDOCK

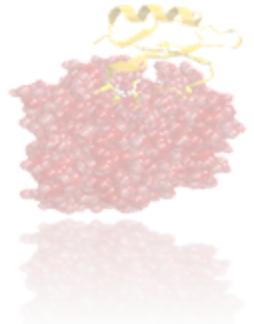
Anna Ranaudo, Marco Giulini, Angela Pelissou Ayuso, and Alexandre M. J. J. Bonvin*

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

Overview

- **Introduction**
 - **Information-driven docking with HADDOCK**
 - **Introducing HADDOCK3**
 - **AI-based antibody-antigen modelling**
 - **Protein cyclic-peptides**
 - **Protein-glycans**
 - **Conclusions & perspectives**

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...)
- **Moving toward flexible and 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

netherlands
eScience center

bioExcel
Centre of Excellence for Computational Biomolecular Research



bonvinlab.org/people

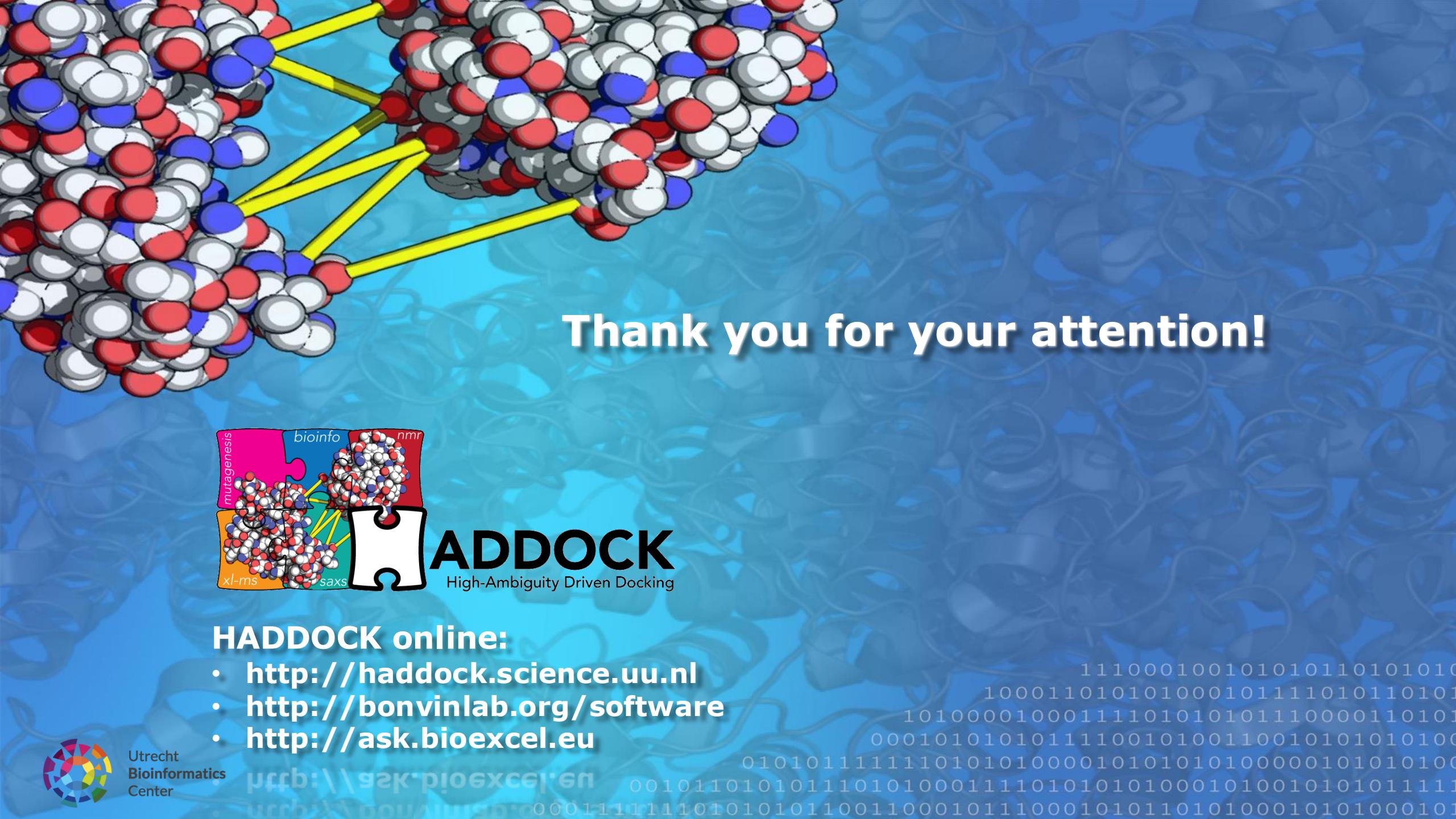
instruct
ERIC

Acknowledgments: the HADDOCK developers over the years

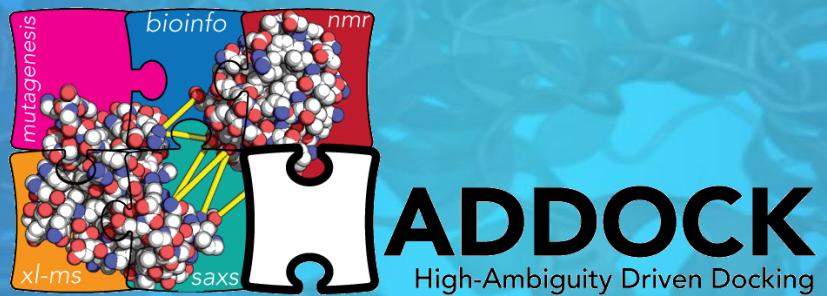


HADDOCK 20 years - Huizen NL, November 8-9 2023

110101010
010110101
000110101
101010100
101010100
101011111



Thank you for your attention!



HADDOCK online:

- <http://haddock.science.uu.nl>
- <http://bonvinlab.org/software>
- <http://ask.bioexcel.eu>