



 **EuChemS**
European Chemical Society

EuChemS CompChem

European Conference on
Computational and Theoretical Chemistry

September 3-7 2023

Chair: Zoe Cournia
Co-Chair: Hans Peter Lüthi



A vertical splash of water droplets and bubbles on the left side of the slide, set against a light blue background with a subtle pattern of small white dots.

Computational Study as Guideline for Experimental Research



Branimir Bertoša
Faculty of Science, University of Zagreb, Croatia

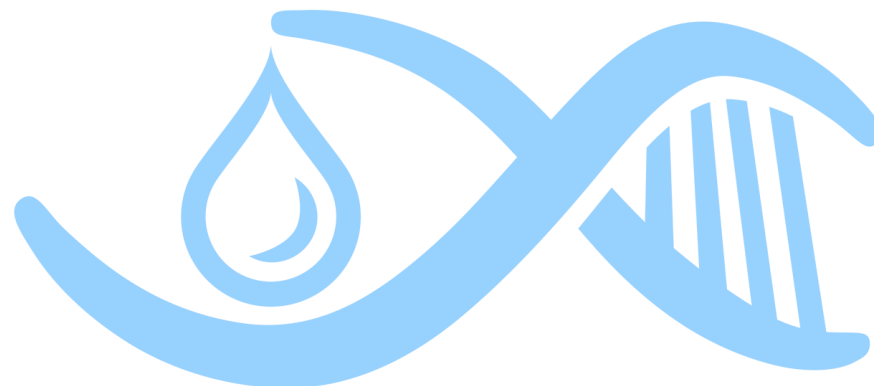


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 952110.

Consortium



Potential customer



MARILIA

Your safe water drop



Experimental
chemistry research



Startup studio,
Transfer of knowledge
from fundamental science
to commercialisation



Computational research



Assay development
Molecular biology



THE MARILIA PROJECT

A new approach for detection of pathogens in water.

THE PROBLEM

Nowadays millions of tons of inadequately or insufficiently treated sewage, industrial and agricultural waste are released directly into the groundwater, rivers, lakes, and oceans. **The water to be used for drinking or in manufacturing processes must meet specific quality requirements.**

THE CURRENT PATHOGEN DETECTION APPROACHES

The current procedures are mostly carried out manually, using **bacterial culture plating methods** or (to a much smaller degree) using **expensive molecular methods** such as Enzyme Linked Immunosorbent Assays (ELISA), reporter enzyme-dependent detection and Polymerase Chain Reaction (PCR).

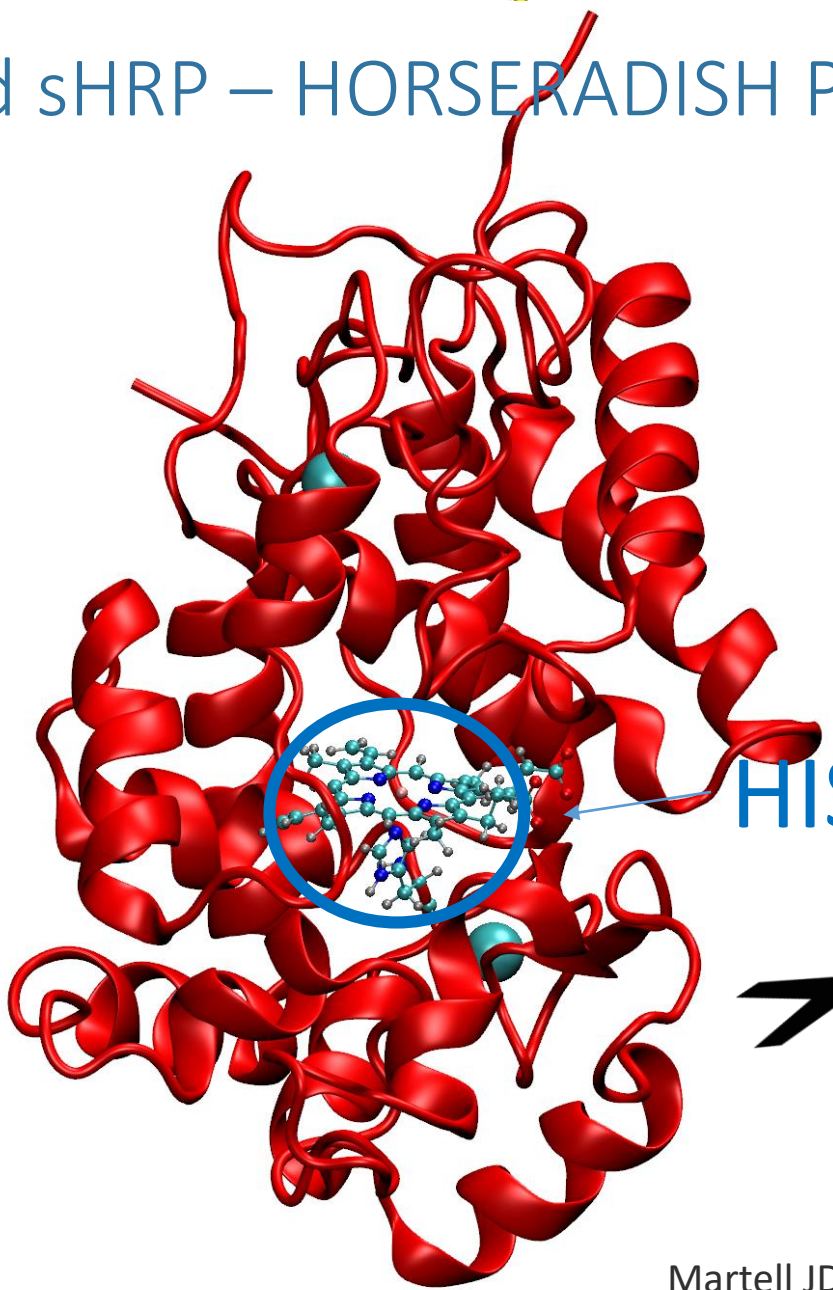
Cultivation-based approaches require enterprises to wait for **48-72 hours** for the results and skilled personnel and access to a microbiological laboratory. In contrast, molecular techniques can be implemented within automated, on-site testing devices that are able to identify bacteria within hours. However, the current molecular technologies able to detect low cell numbers require expensive reagents, consumables, and sophisticated instruments **resulting in costs of more than 50 €/test.**

THE MAIN GOAL OF MARILIA PROJECT

To develop a novel test which is cheaper (< 10 €) and faster (< 10 min) than existing ones.

HRP and sHRP – HORSE RADISH PEROXIDASE

Residue ID	Wild type	Mutation ¹
21	Threonine	Isoleucine
78	Proline	Serine
93	Arginine	Glycine
175	Asparagine	Serine
255	Asparagine	Aspartate
299	Leucine	Arginine



HIS170 – HEME

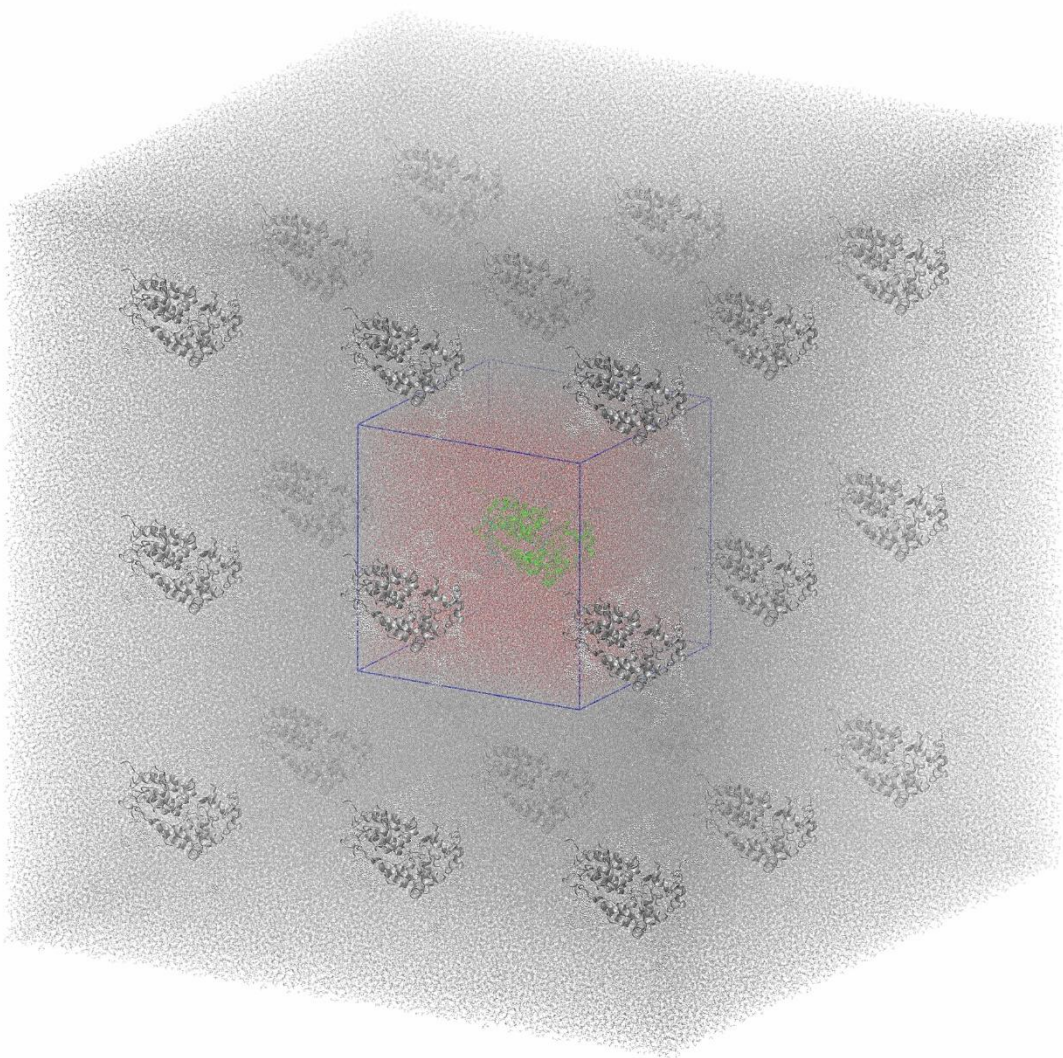
GLY213 – ASN214

Martell JD, Yamagata M, Deerinck TJ, et al. *Nat. Biotechnol.* (2016) **34(7)** 774–780.

Preparation of systems for MD simulations

Systems:

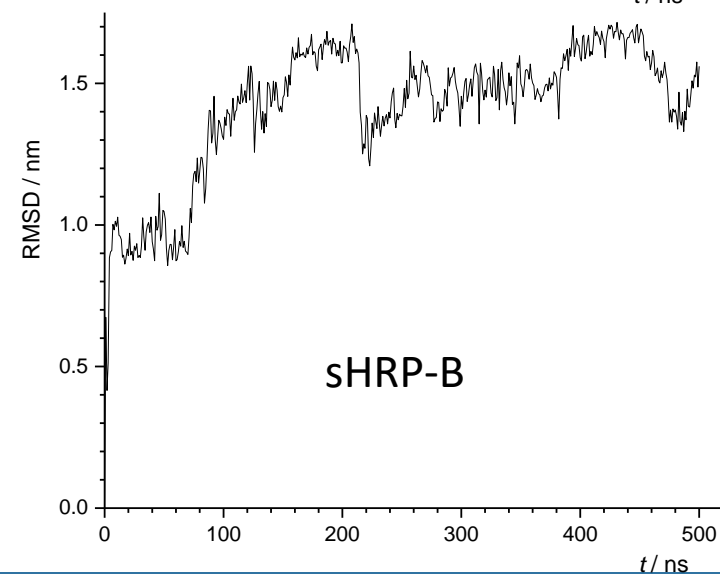
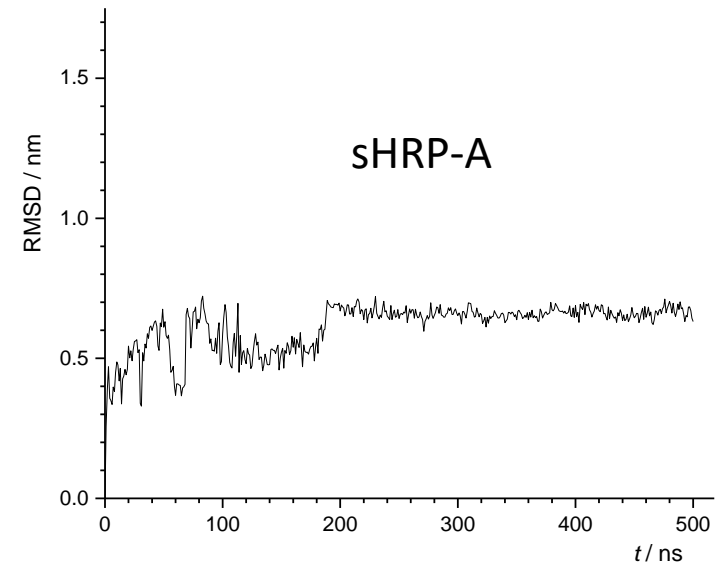
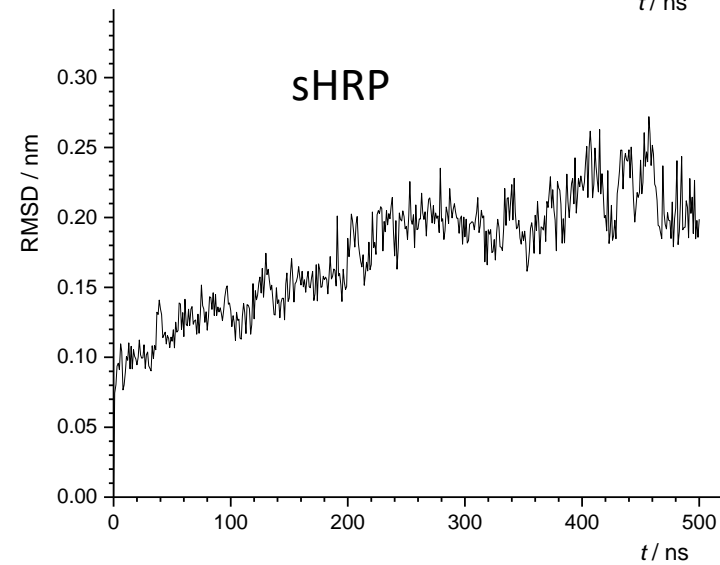
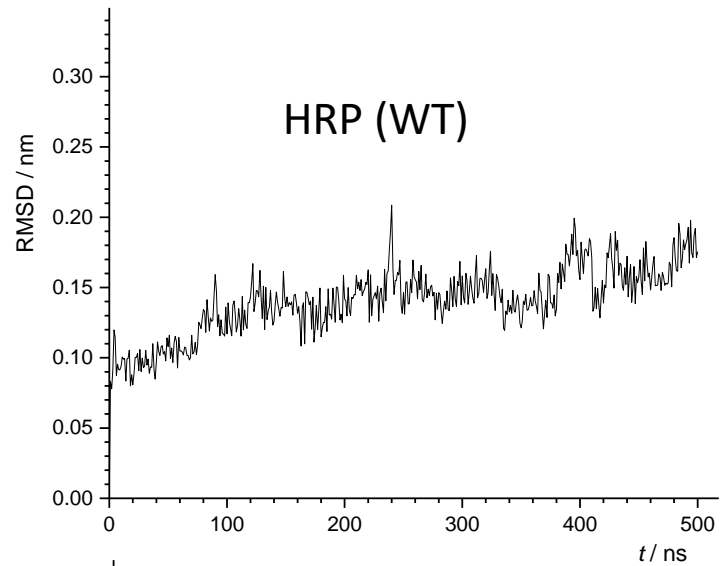
- HRP
- sHRP
- sHRP-A
- sHRP-B



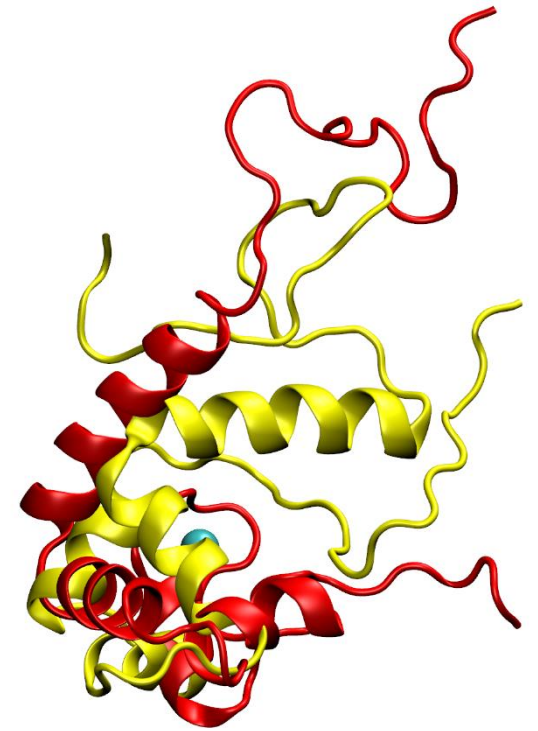
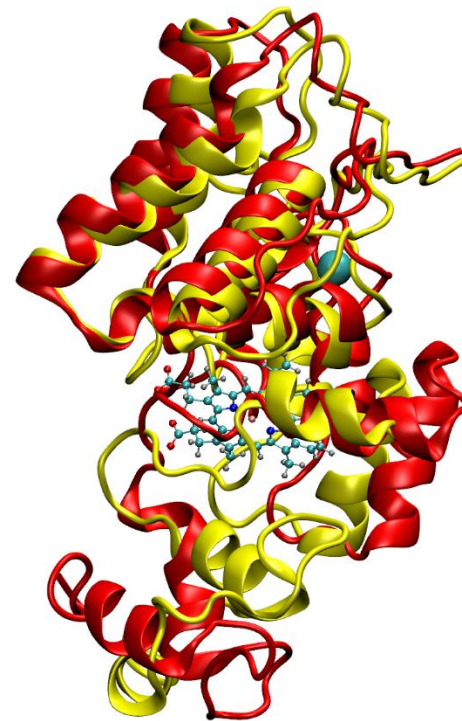
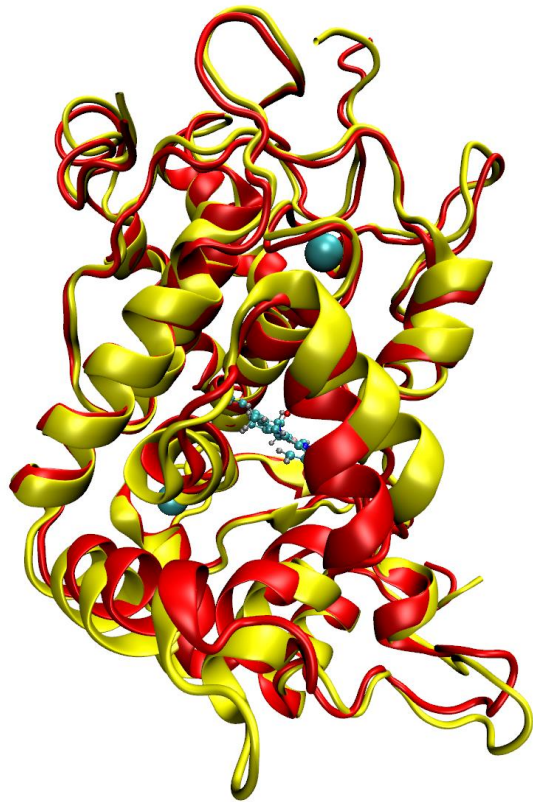
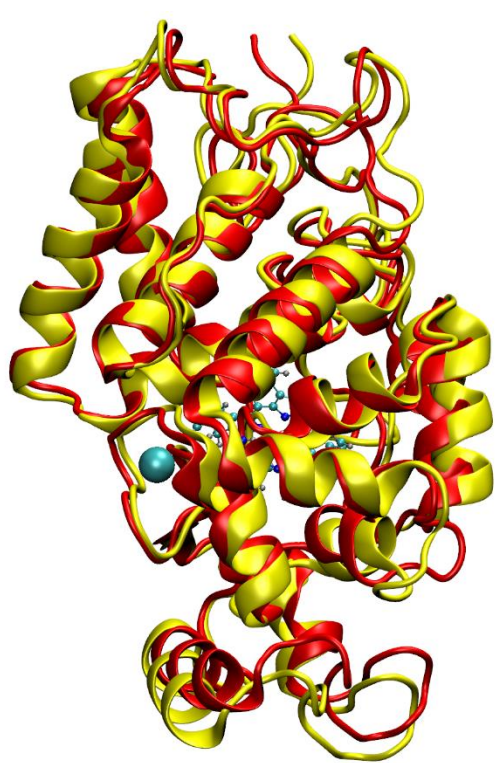
- PDB: 1h5a - the horseradish peroxidase (HRP) C1A enzyme
- System preparation: CHARMM-GUI
- Explicit solvent (TIP3P water model)
- PBC cubic box
- GROMACS with Charm36m forcefield
- Steepest descent energy minimization algorithm (5000 steps)
- Equilibration:
 - 600 ps NVT – increasing the temperature from 10 K to room temperature
 - 500 ps NPT
- Production phase: 500 ns NPT



Stability of simulated systems



Stability of simulated systems

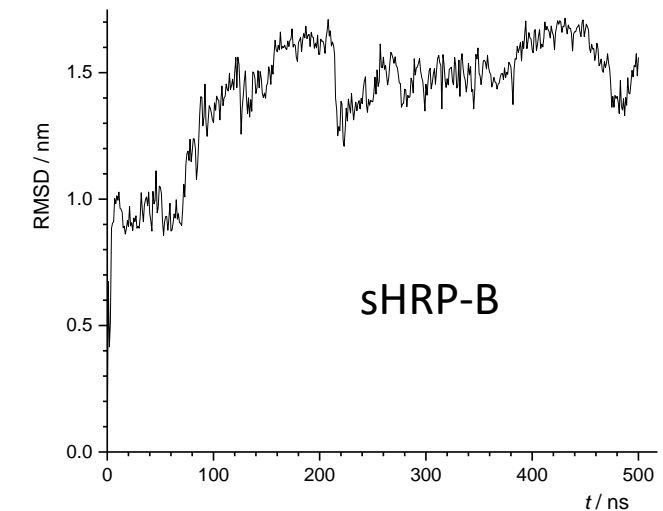
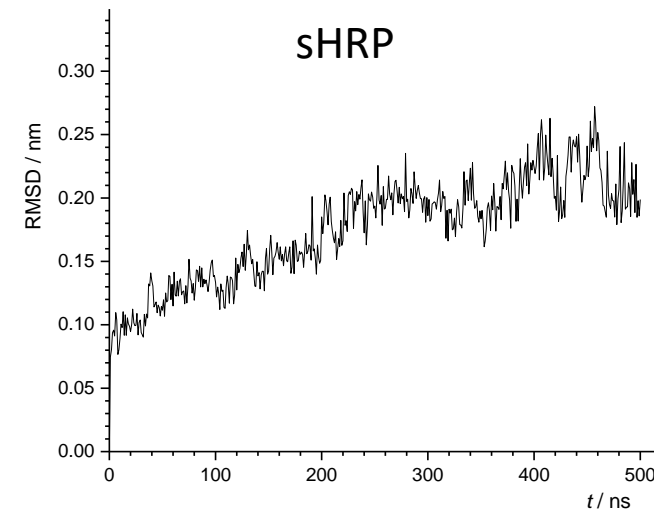
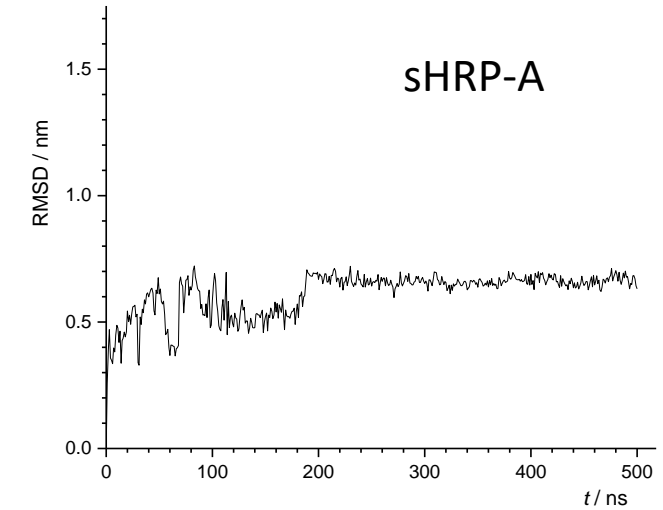
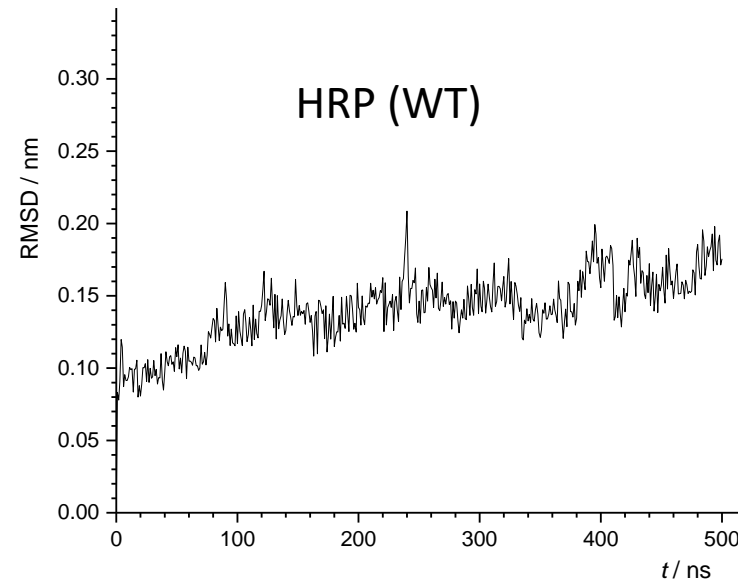


RED – starting structure

YELLOW – structure at the end of simulation

Stability of simulated systems

- systems are **not completely stabilised** during 500 ns of MD simulations
- **experiments** pointed to the **questionable stability** of various forms of the enzyme
- effect of **glycosylation** ?

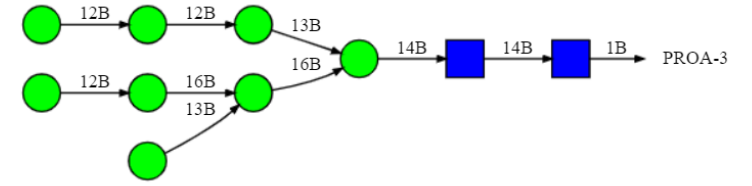


Glycosylation

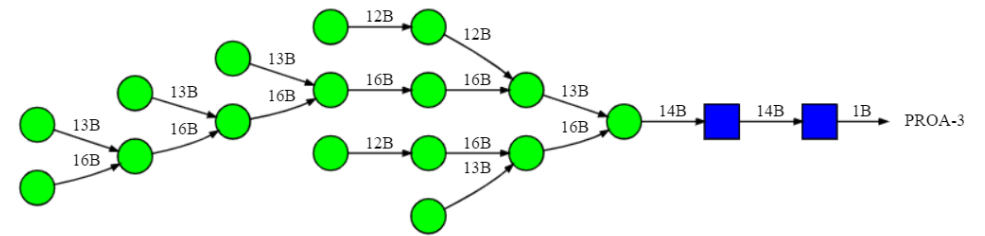
Type of glycosylation?

ASN-X-THR/SER

(X is any amino acid residue other than proline or aspartic acid)

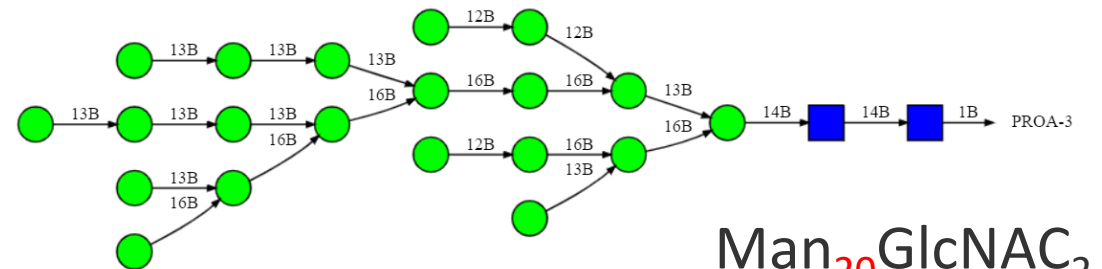


Man₈GlcNAC₂



Man₁₆GlcNAC₂

■ Glc
● Man



Man₂₀GlcNAC₂

F. W. Krainer, C. Gmeiner, L. Neusch, M. Windwarder, R. Pletzenauer, C. Herwig, F. Altmann, A. Glieder, O. Spadiut, **2013**, DOI 10.1038/srep03279
E. L. Wu, X. Cheng, S. Jo, H. Rui, K. C. Song, E. M. Dávila-Contreras, Y. Qi, J. Lee, V. Monje-Galvan, R. M. Venable, J. B. Klauda, W. Im, *J. Comput. Chem.* **2014**, 35, 1997–2004.

Glycosylation

ASN13

ASN57

ASN158

ASN186

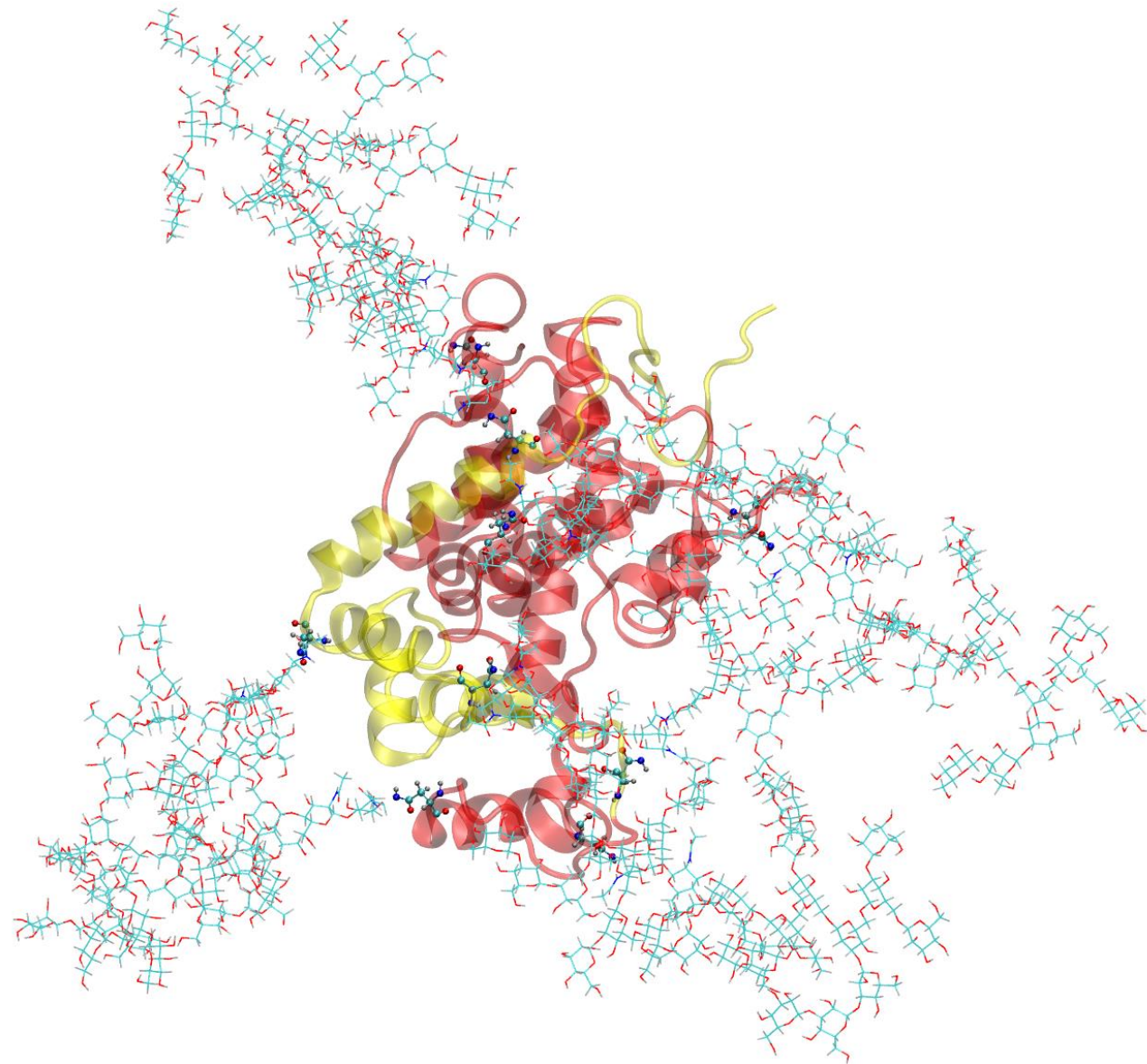
ASN198

ASN214

ASN255

ASN268

ASN286



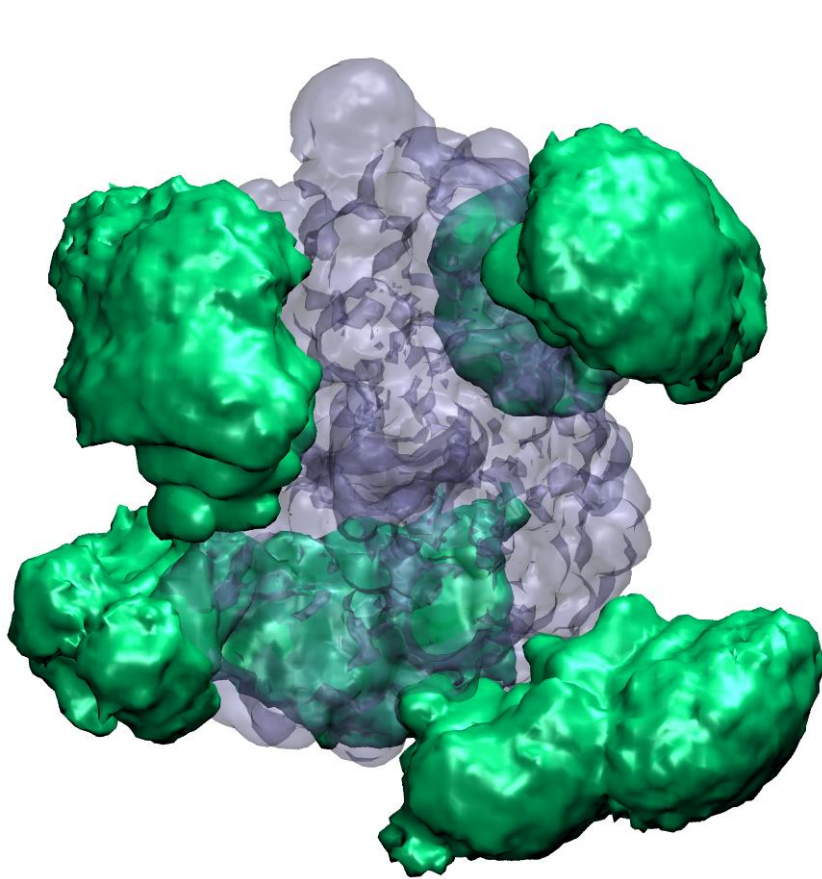
Glycosylation

8 systems were prepared and subjected to 500 ns of MD simulations

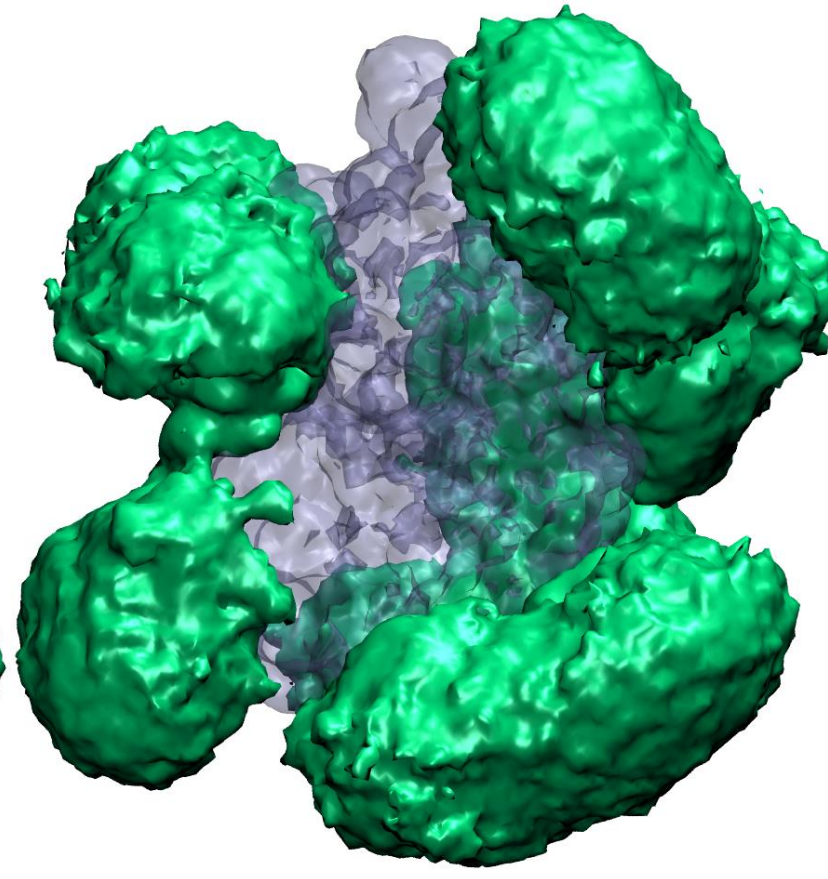
System	Glycan branching type			
HRP	NO glycan	Man ₈ GlcNAc ₂	Man ₁₆ GlcNAc ₂	Man ₂₀ GlcNAc ₂
sHRP	NO glycan	Man ₈ GlcNAc ₂	Man ₁₆ GlcNAc ₂	Man ₂₀ GlcNAc ₂

Glycosylation

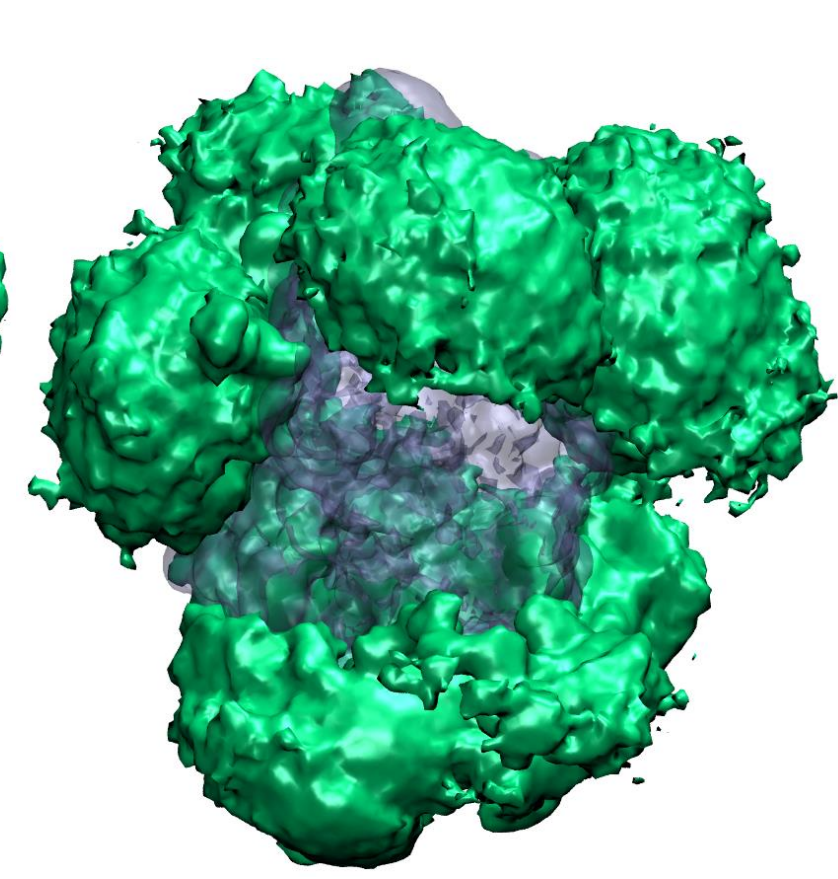
Influence of the degree of glycosylation



$\text{Man}_8\text{GlcNAc}_2$



$\text{Man}_{16}\text{GlcNAc}_2$



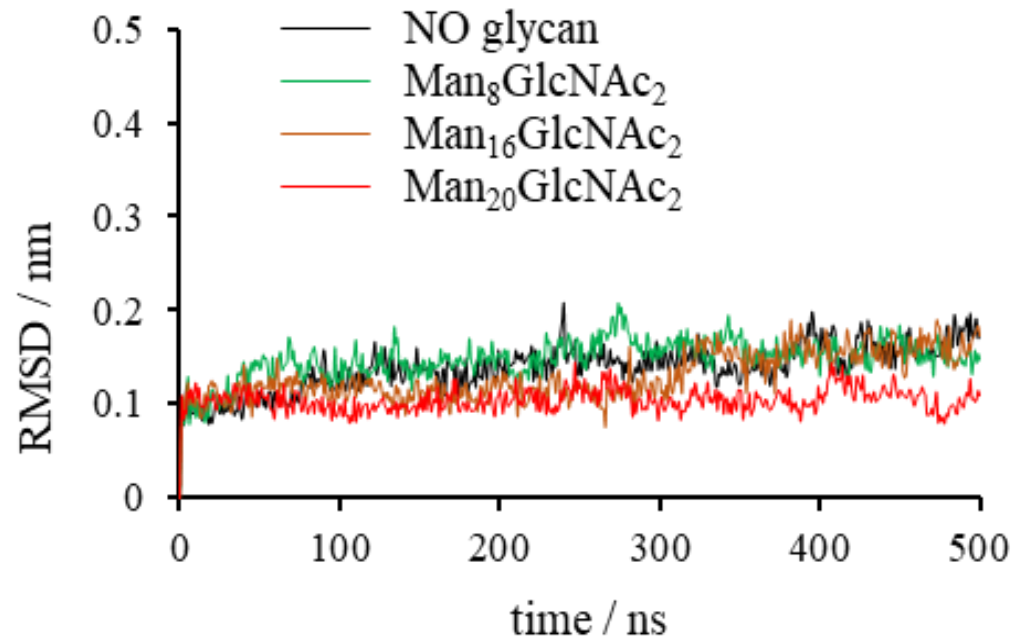
$\text{Man}_{20}\text{GlcNAc}_2$

Effect of *N*-glycosylation on protein structural properties

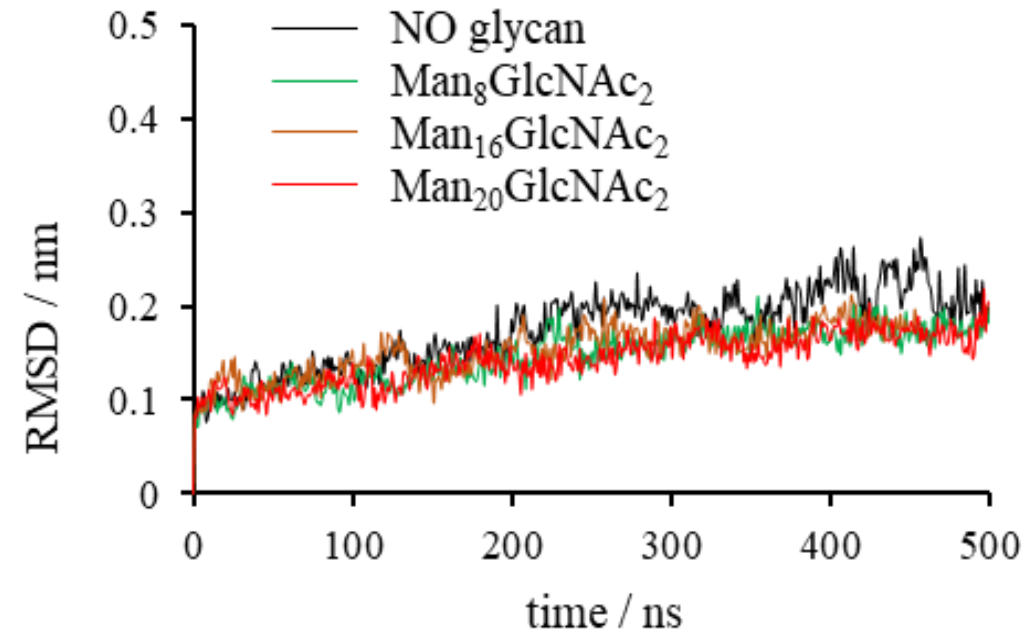
Glycosylation

RMSD – glycosylation **increases stability** of the protein

a) HRP



b) sHRP

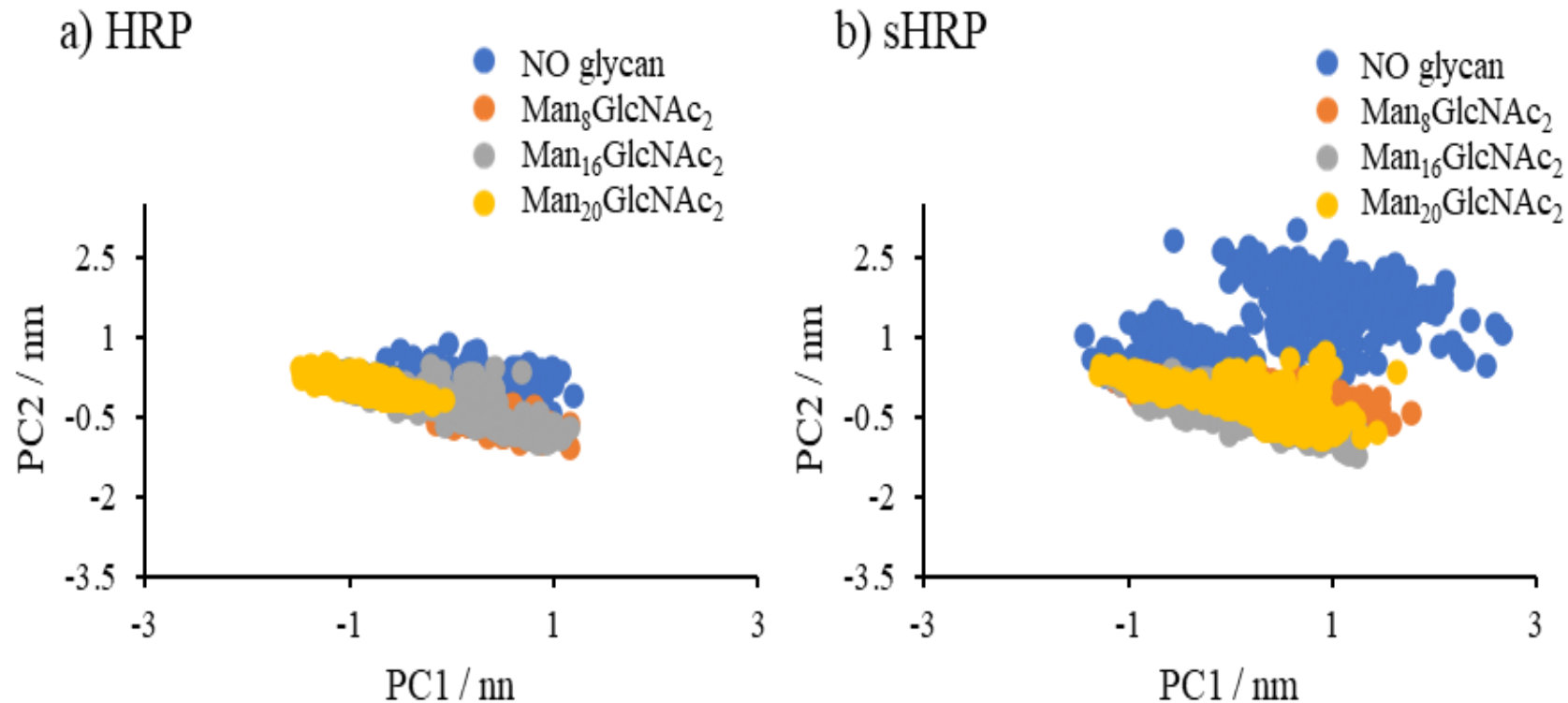


S. Škulj, A. Barišić, N. Mutter, O. Spadiut, I. Barišić, B. Bertoša, *Computational and Structural Biotechnology Journal* **20** (2022) 3096–3105.



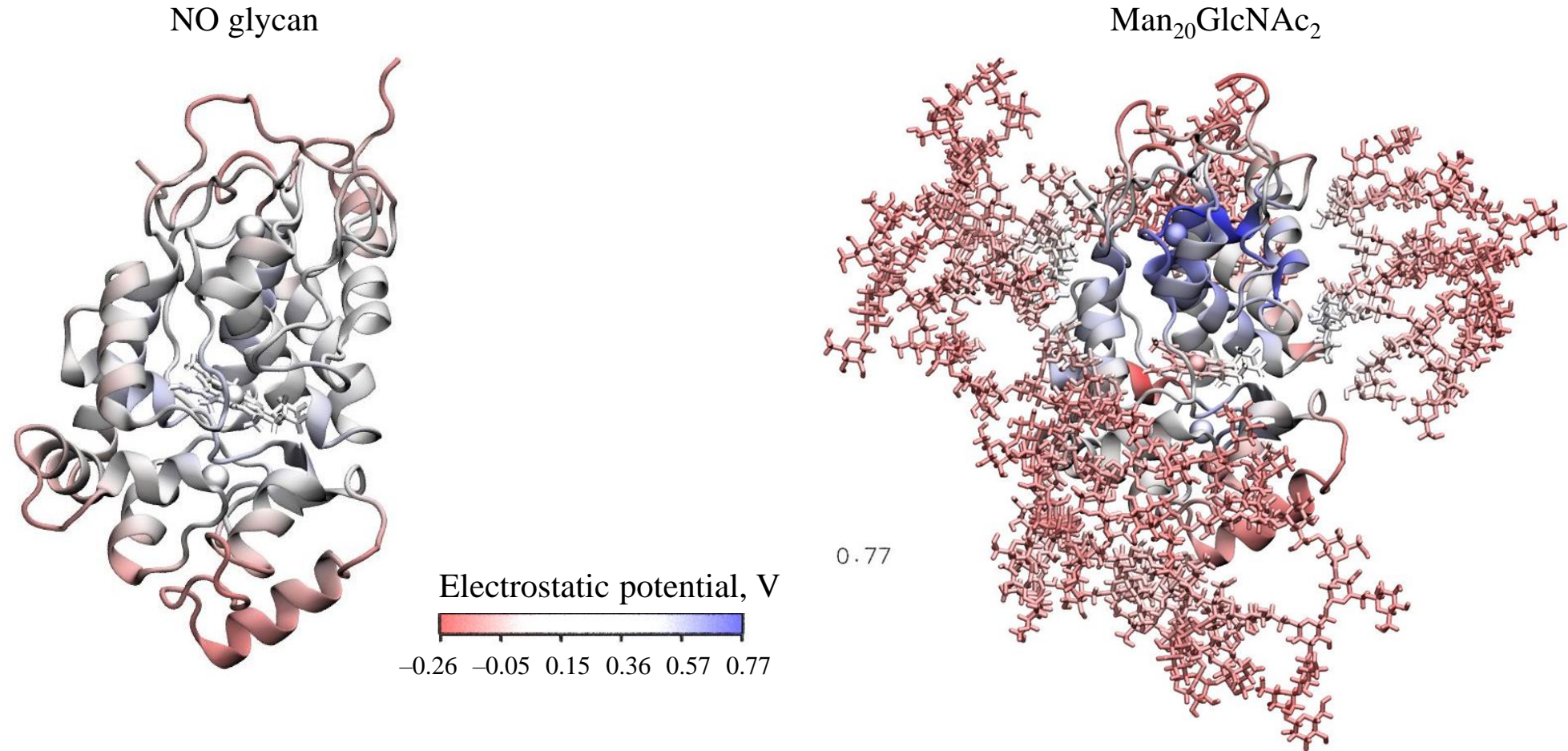
Glycosylation

PCA – Principal Component Analysis – shows structural variations are decreased due to glycosylation



Effect of *N*-glycosylation on protein electrostatic potential

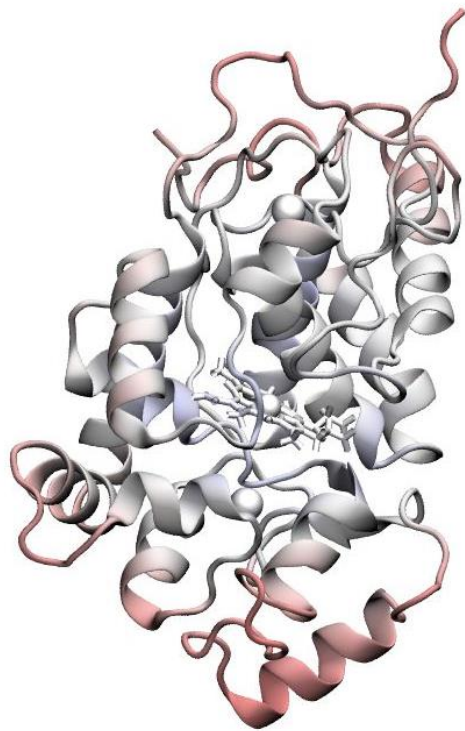
Effect of Glycosylation



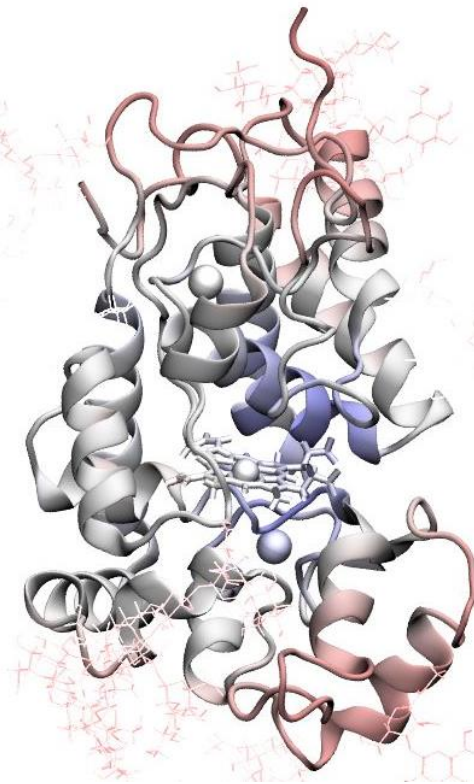
S. Škulj, A. Barišić, N. Mutter, O. Spadiut, I. Barišić, B. Bertoša, *Computational and Structural Biotechnology Journal* **20** (2022) 3096–3105.

Glycosylation - Electrostatic potential

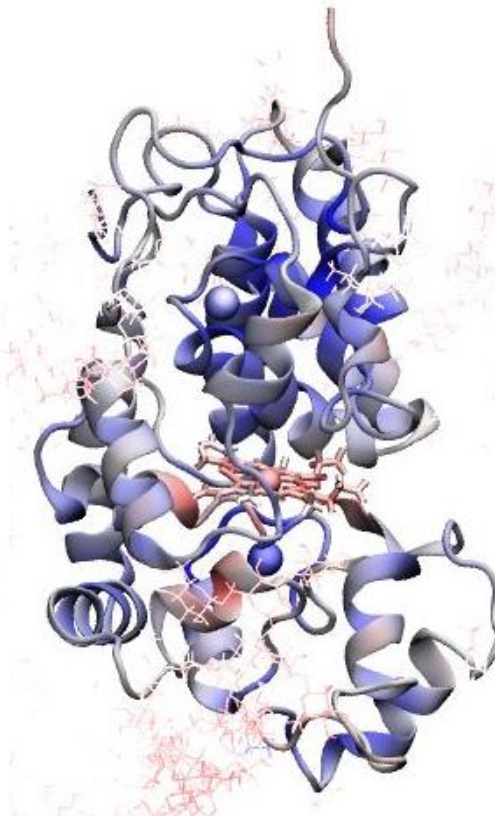
HRP NO



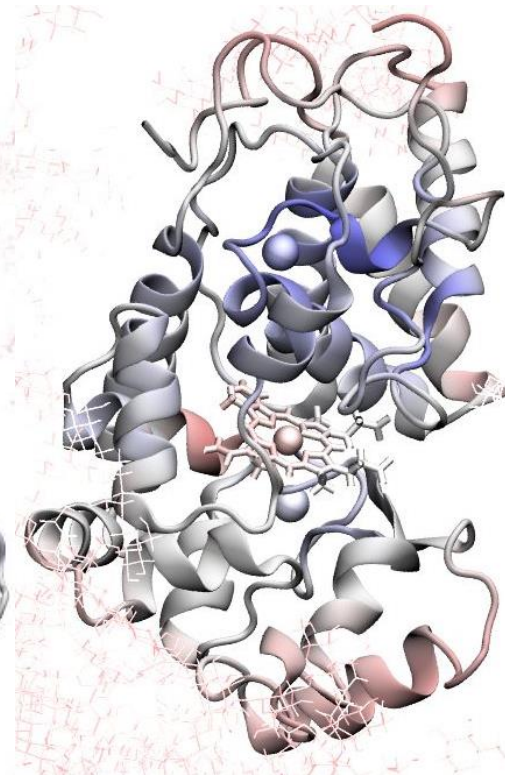
HRP Man₈GlcNAc₂



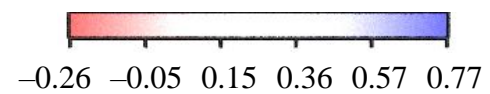
HRP Man₁₆GlcNAc₂



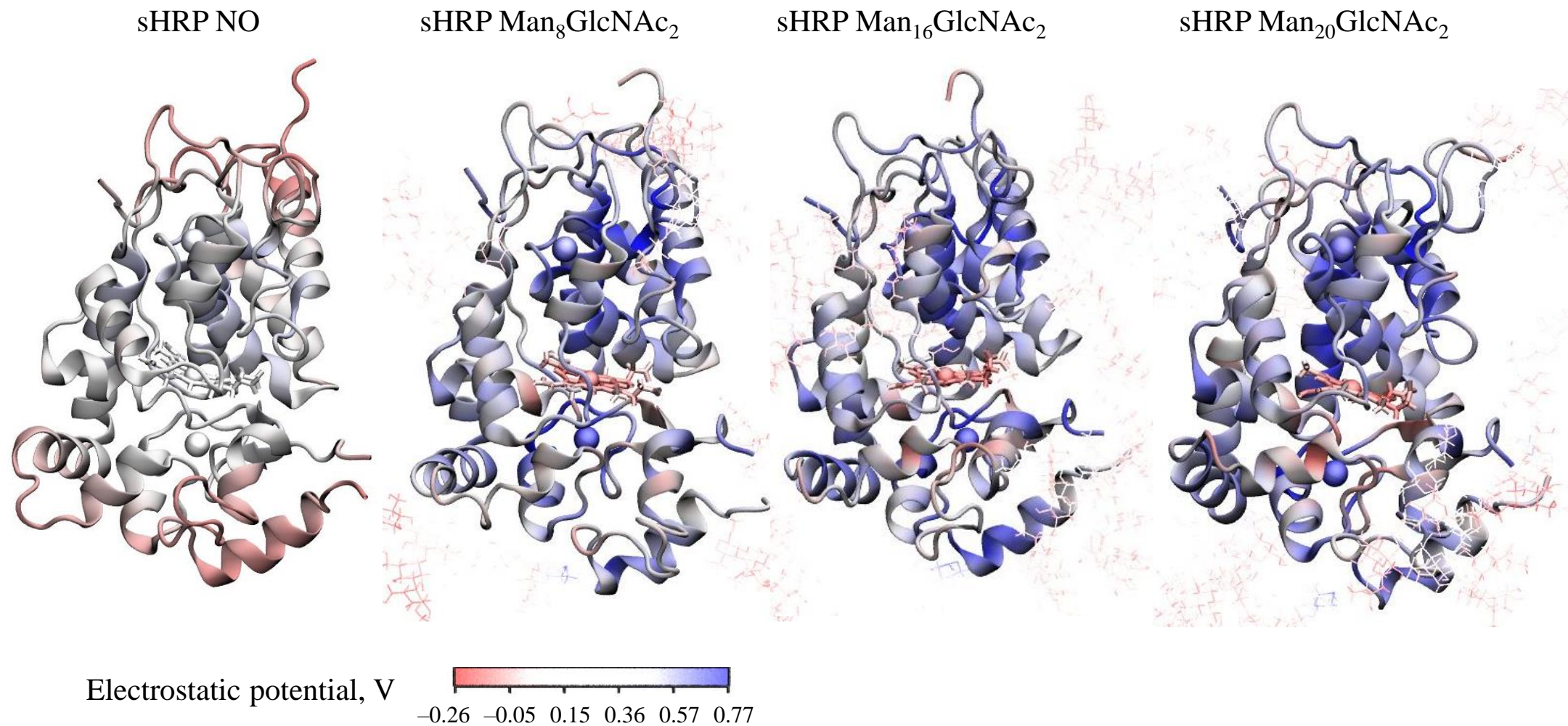
HRP Man₂₀GlcNAc₂



Electrostatic potential, V

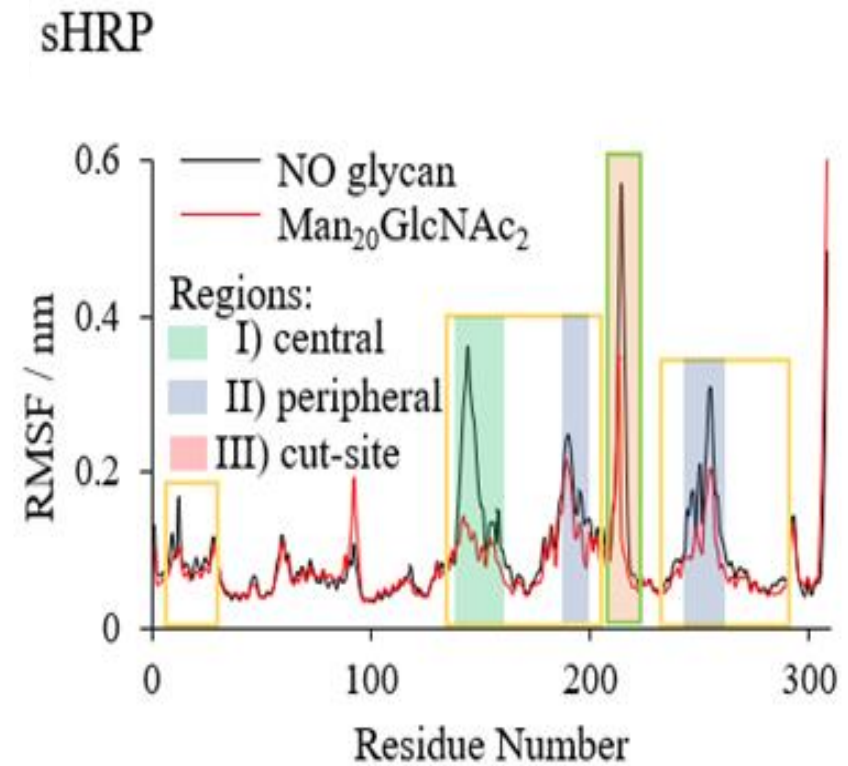
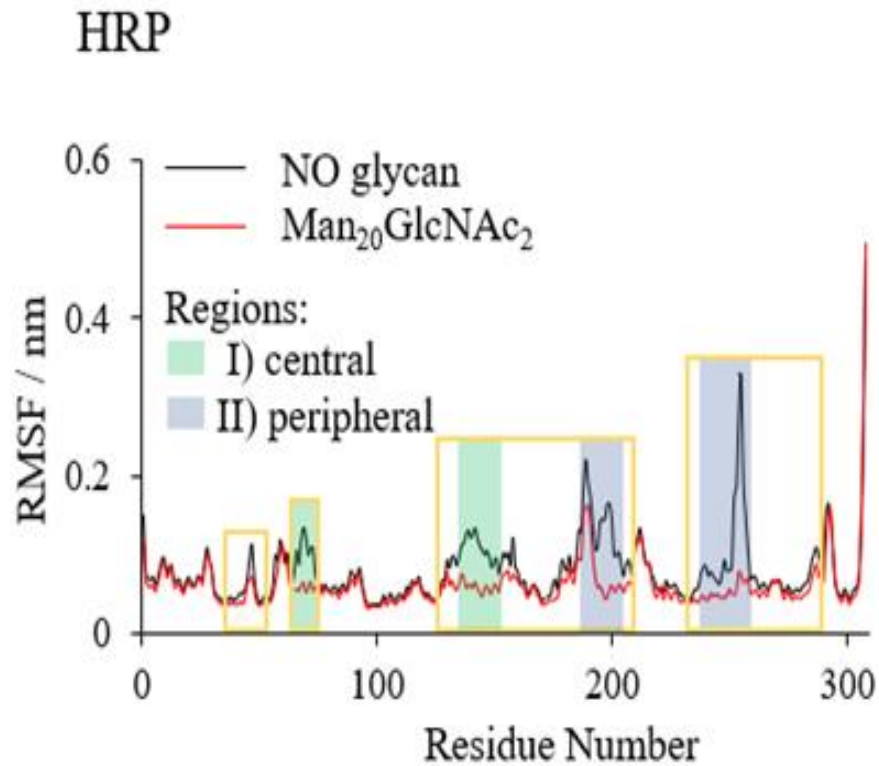


Glycosylation - Electrostatic potential



Effect of *N*-glycosylation on protein dynamical properties

Glycosylation - fluctuations

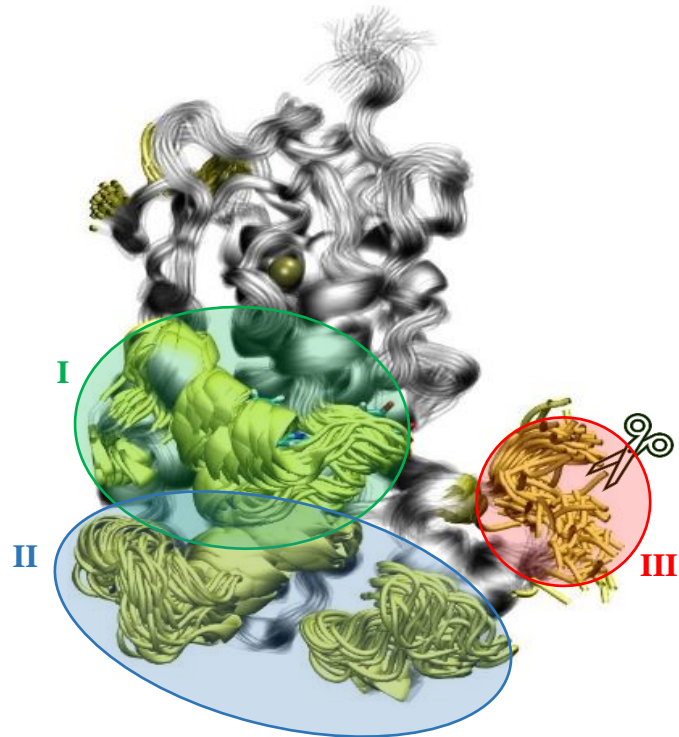


System		NO glycan	Man ₈ GlcNAc ₂	Man ₁₆ GlcNAc ₂	Man ₂₀ GlcNAc ₂
HRP	AVERAGE RMSF × 10 ⁻² / nm	7.7 ± 3.7	6.6 ± 3.1	6.9 ± 2.9	5.7 ± 2.2
sHRP	AVERAGE RMSF × 10 ⁻² / nm	9.5 ± 7.0	7.6 ± 5.0	7.9 ± 4.1	7.4 ± 3.9

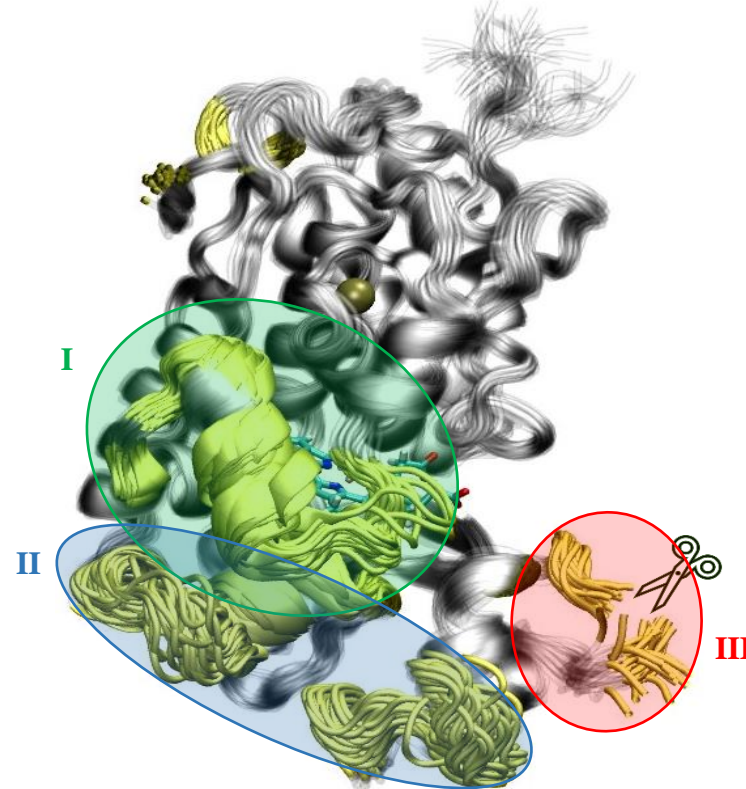


Glycosylation - fluctuations

sHRP – NO glycan



sHRP – Man₂₀GlcNAc₂



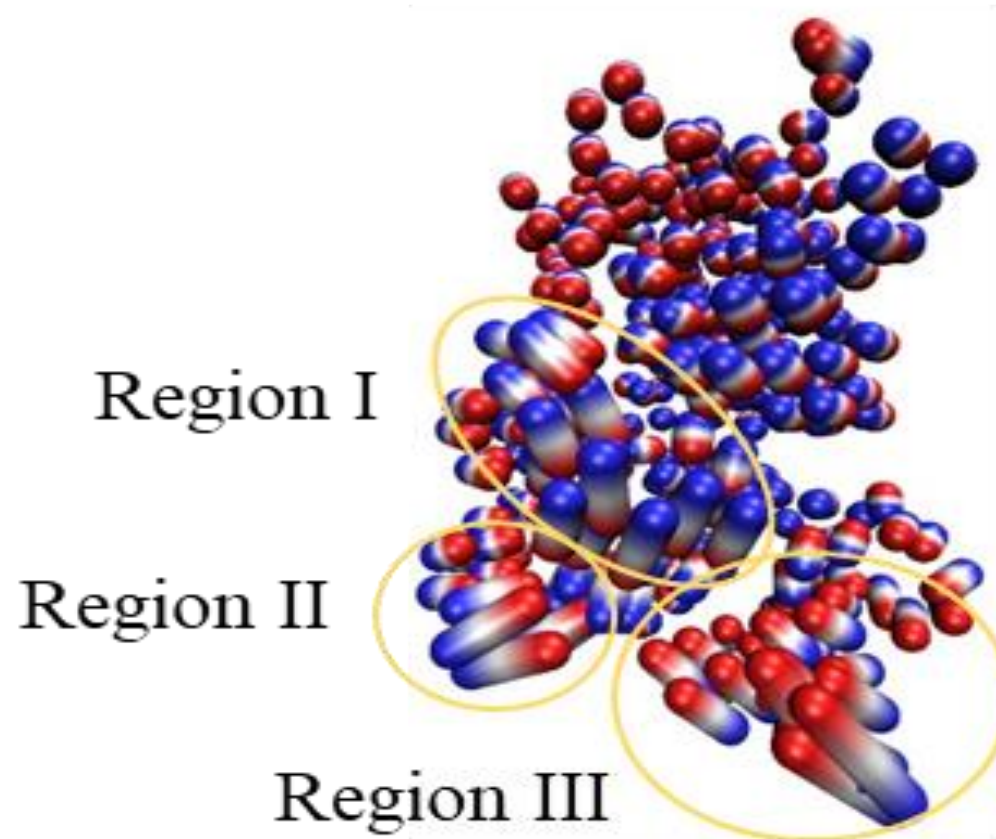
Residues:

- I) central region: 140-151, 155-158, 160
- II) peripheral region: 189-199, 244-247, 249-261
- III) cut-site region: 213-217

➤ Glycosylation - fluctuations

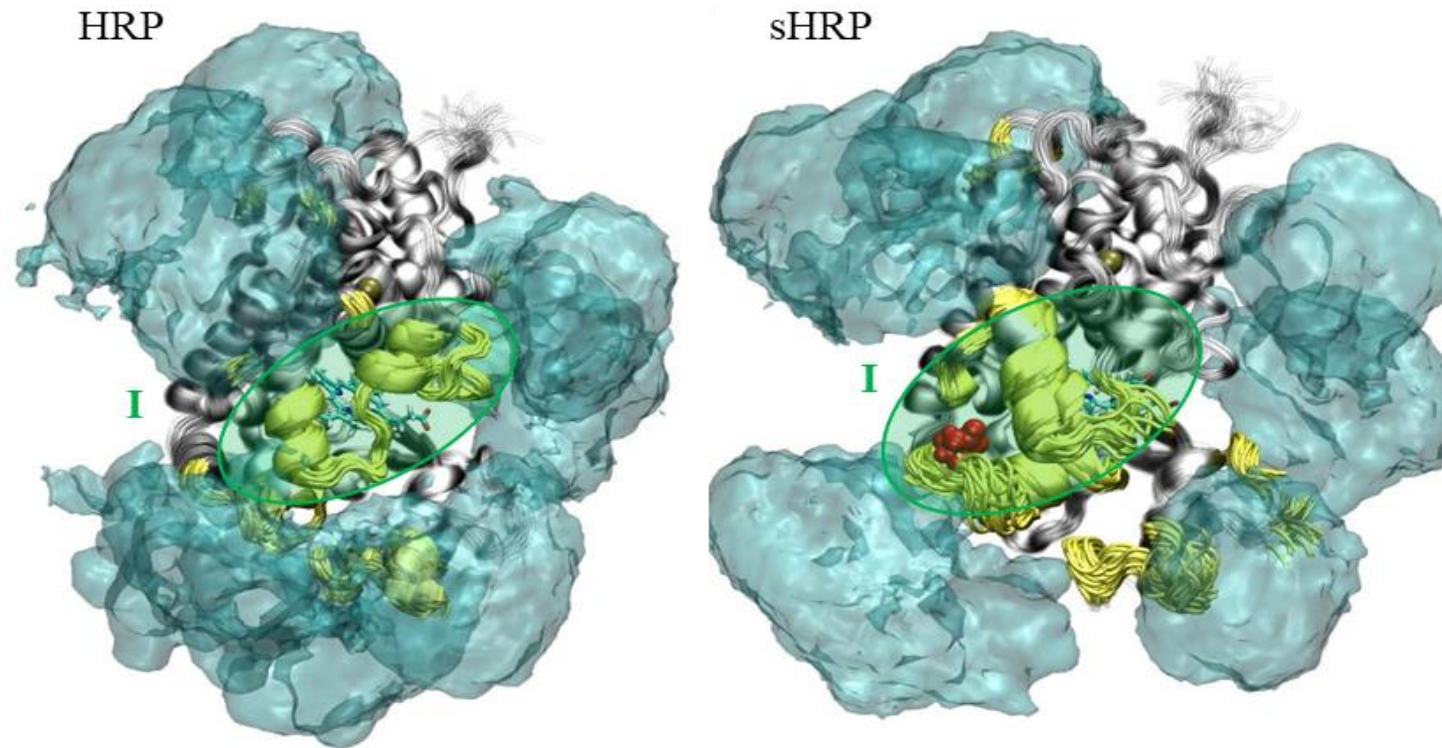
PCA – Principal Component Analysis – the same regions were identified

Movement along first eigenvector PC1



➤ Glycosylation – fluctuations of glycans

PROPAGATED EFFECT¹ – glycosylation is on the surface of the protein, but it affects fluctuations of central region of the protein



1. Lee, H., Qi, Y. & Im, W. Effects of *N*-glycosylation on protein conformation and dynamics: Protein Data Bank analysis and molecular dynamics simulation study. *Sci Rep* 5 (2015) 8926

➤ Glycosylation

Conclusions:¹

- glycosylation **increases stability** of protein's **structure**
- glycosylation **induces polarisation** of the protein **electrostatic potential**
- glycosylation **decreases** protein **fluctuations**
- glycosylation effects are **propagated** from the surface to the distance protein regions

Further experimental research was conducted using glycosylated form of HRP!

1. S. Škulj, A. Barišić, N. Mutter, O. Spadiut, I. Barišić, B. Bertoša, *Computational and Structural Biotechnology Journal* **20** (2022) 3096–3105.

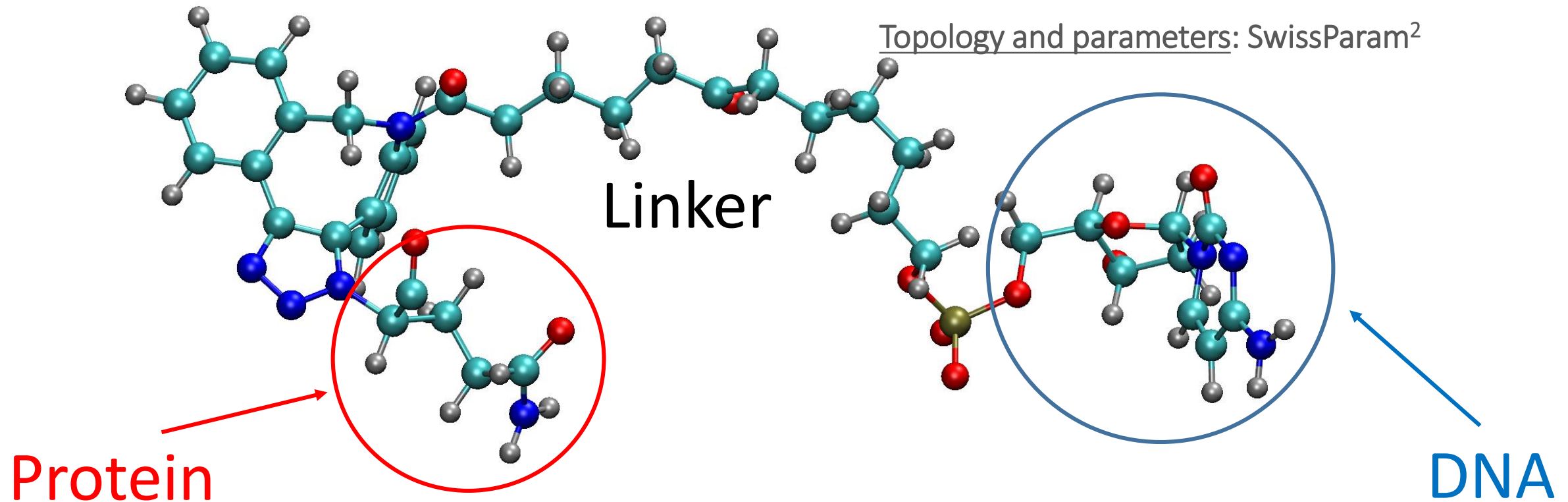
Linker parametrisation

Parametrisation of linker:

Geometry minimization: Gaussian09¹

QM (DFT) in implicit water using SMD/B3LYP/6-31g(d)
level of theory

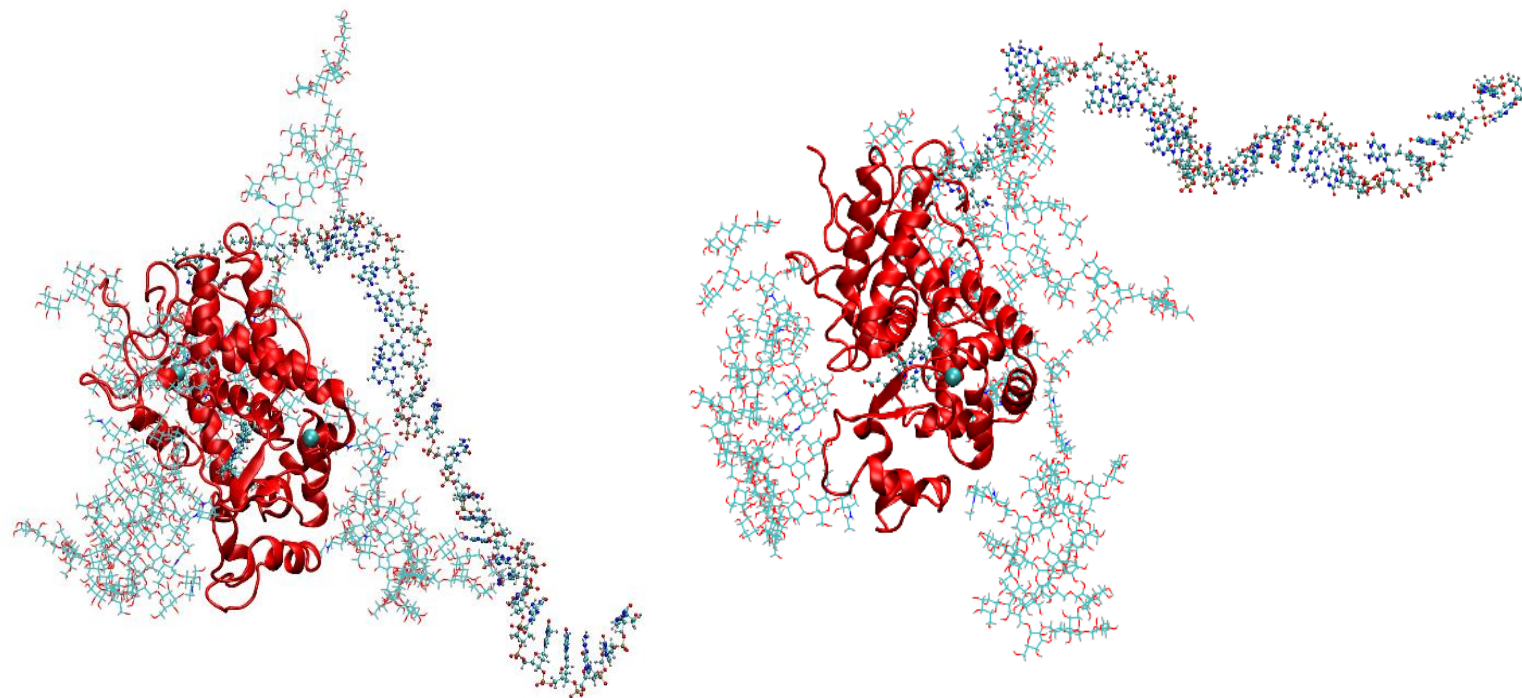
Topology and parameters: SwissParam²



1. Frisch, M. J. et al. Gaussian 09; Gaussian, Inc.: Wallingford, CT, (2009).

2. Zoete, V., Cuendet, M. A., Grosdidier, A. & Michielin, O. SwissParam: a fast force field generation tool for small organic molecules. *J. Comput. Chem.* **32**, 2359–2368 (2011).

ssDNA-HRP – Starting structure



HRP information:

- HRP (WT)
- $\text{Man}_{16}\text{GlcNAc}_2$



MARILIA

Your safe water drop

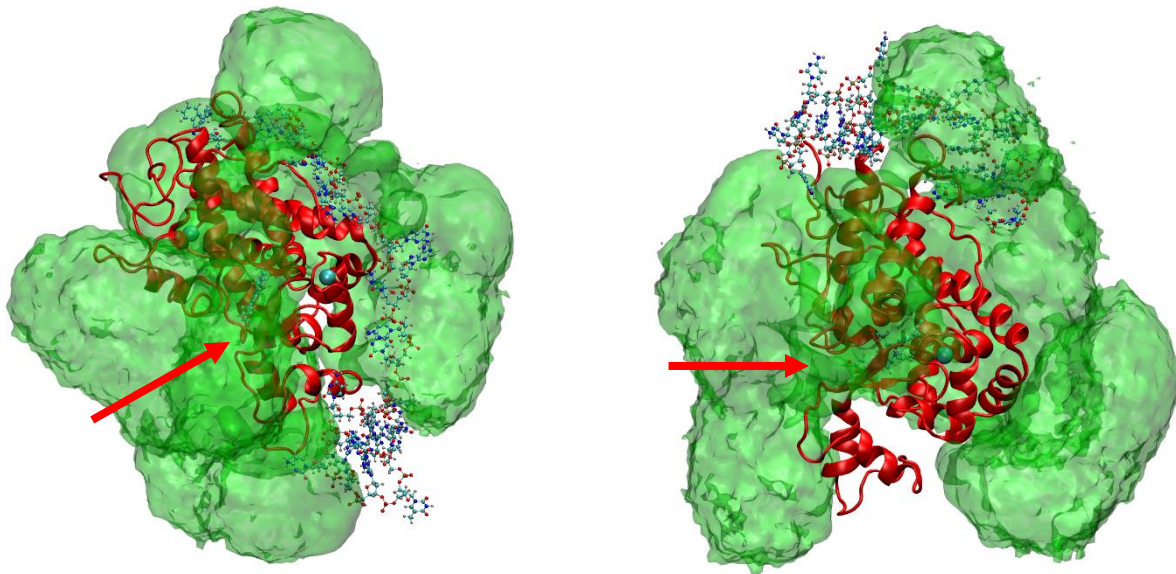
Glycosylated HRP protein with single strand DNA

500 ns simulation

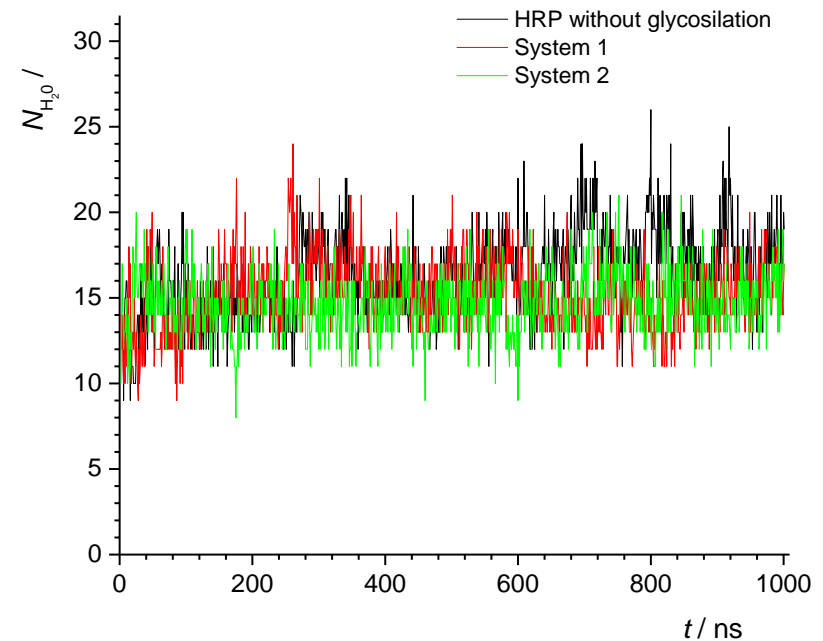


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 952110.

ssDNA-HRP – End of simulation

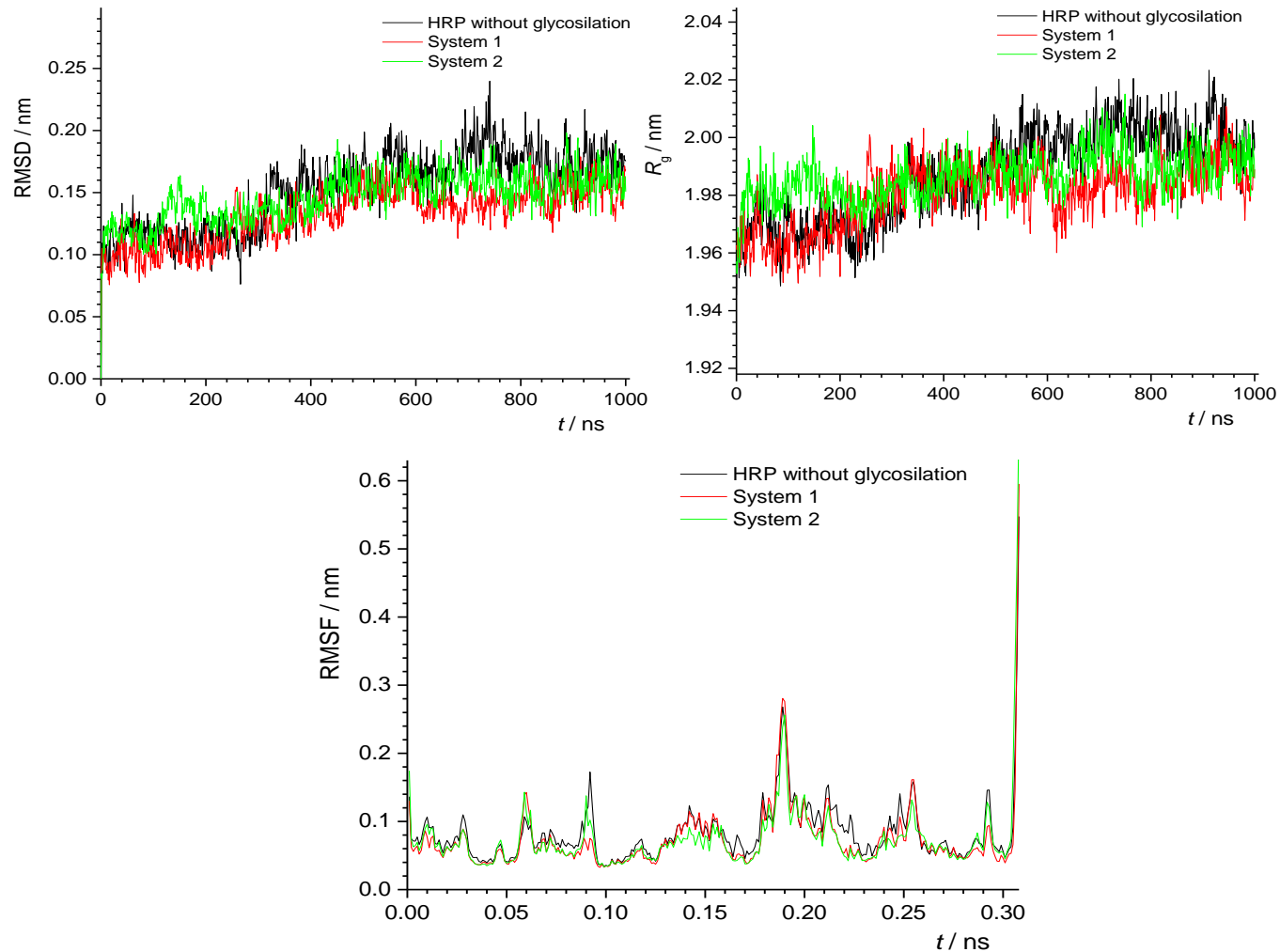


Structures after 1 μ s of MD simulations

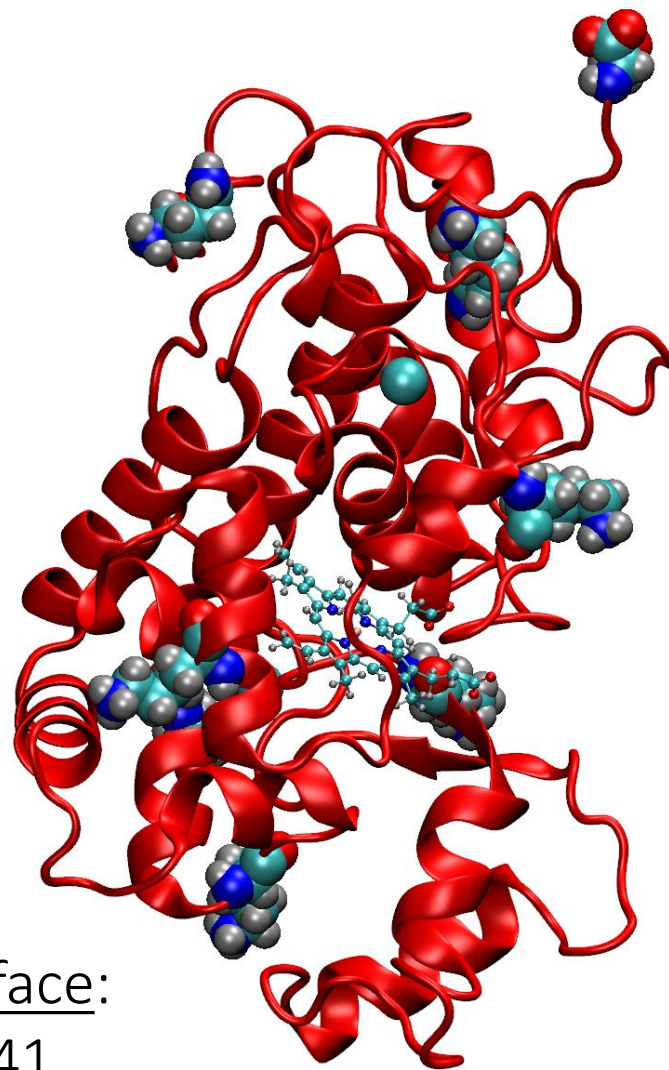


Number of water molecules ≤ 0.5 nm from the active site entrance.

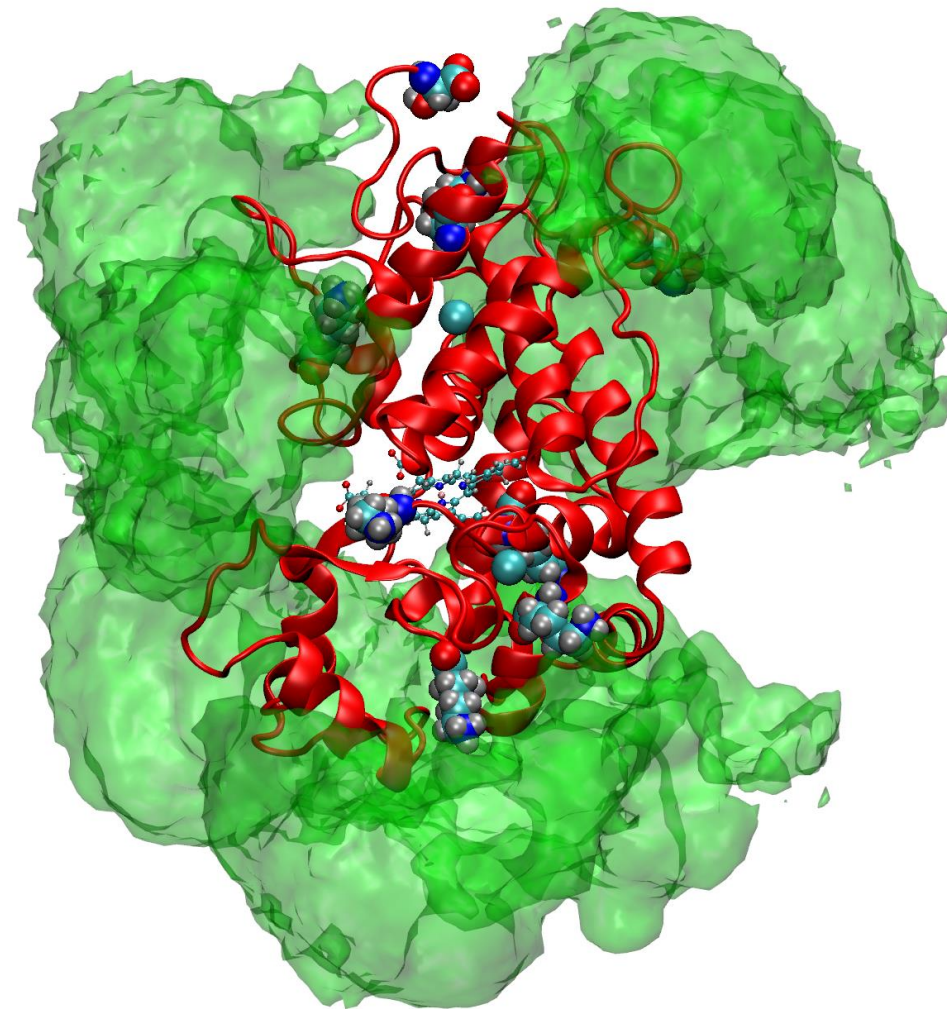
ssDNA-HRP – End of simulation



➤ Lysine availability for oligonucleotide attachment

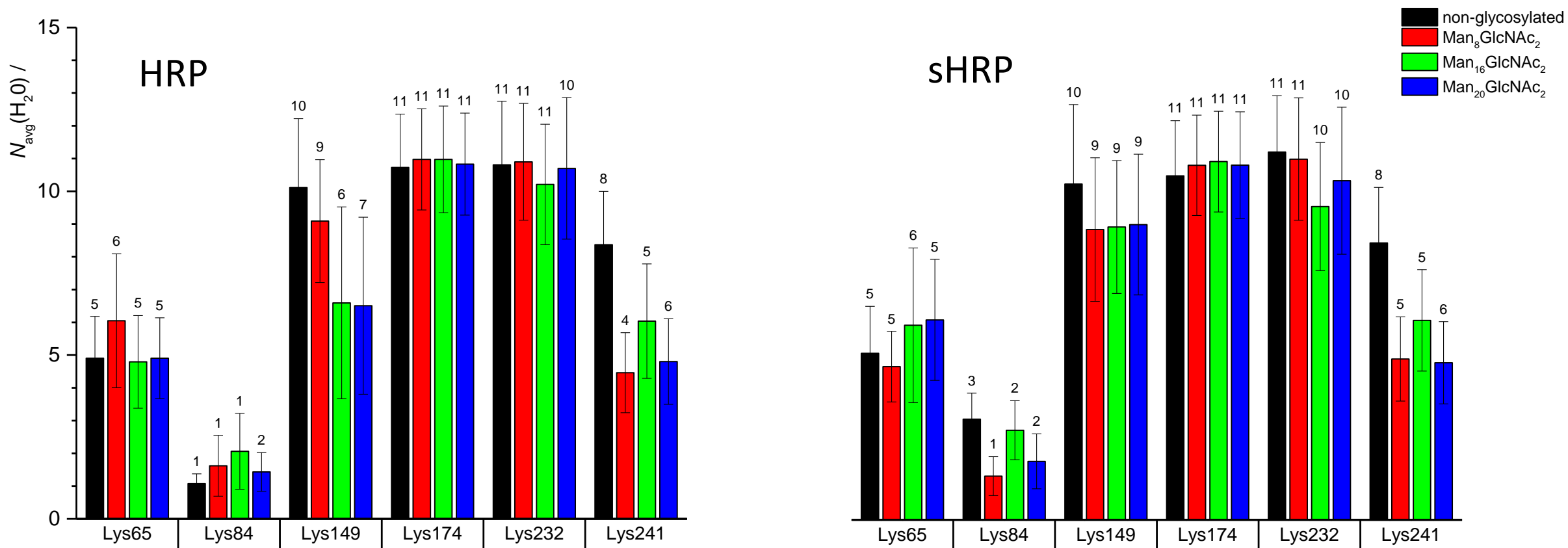


6 Lys on HRP (sHRP) surface:
65, 84, 149, 174, 232, 241



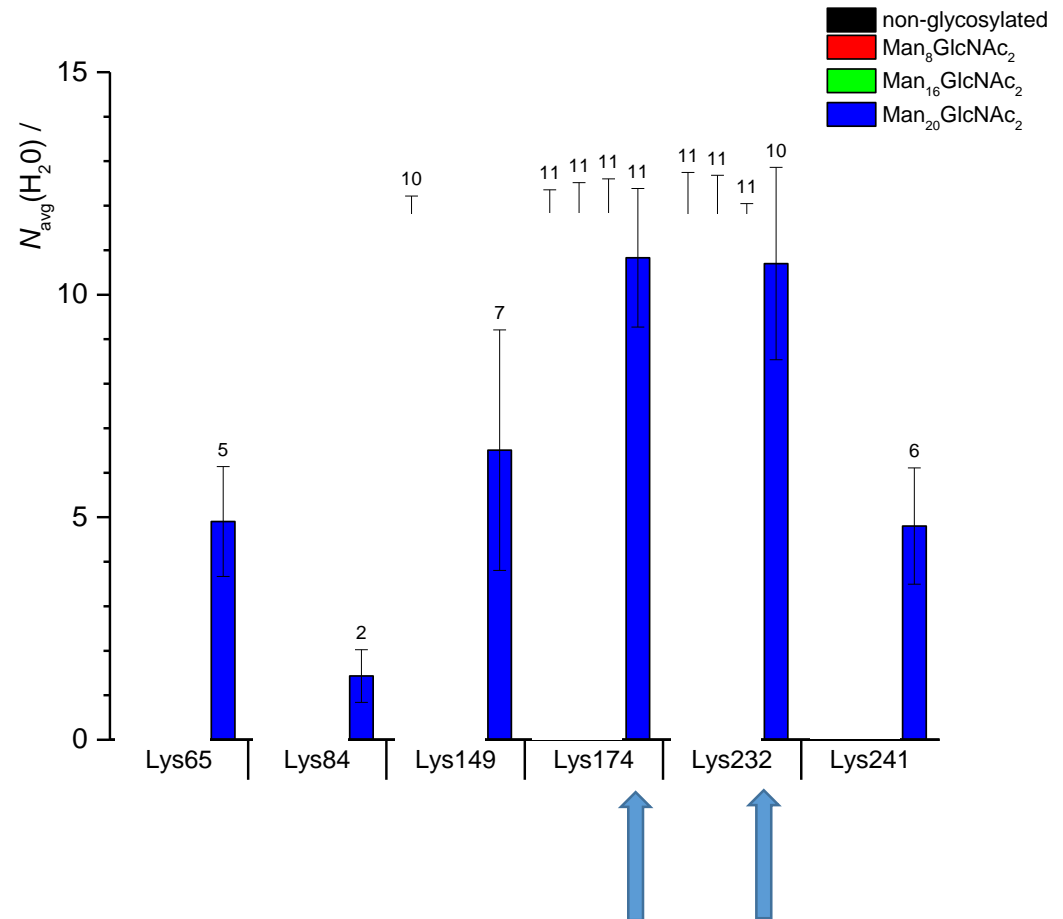
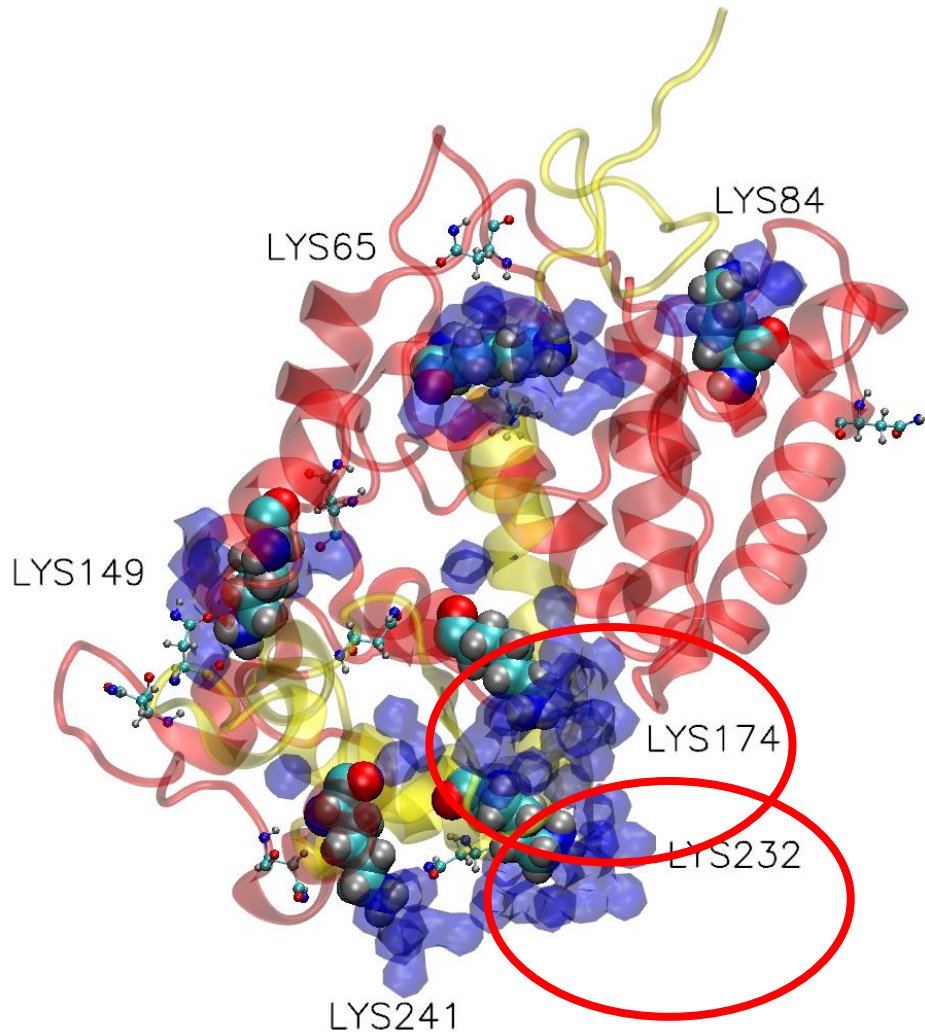
Lysine availability for oligonucleotide attachment

Lysine exposure to the solvent was quantitatively measured by counting the number of water molecules in Lys vicinity during MD simulations.



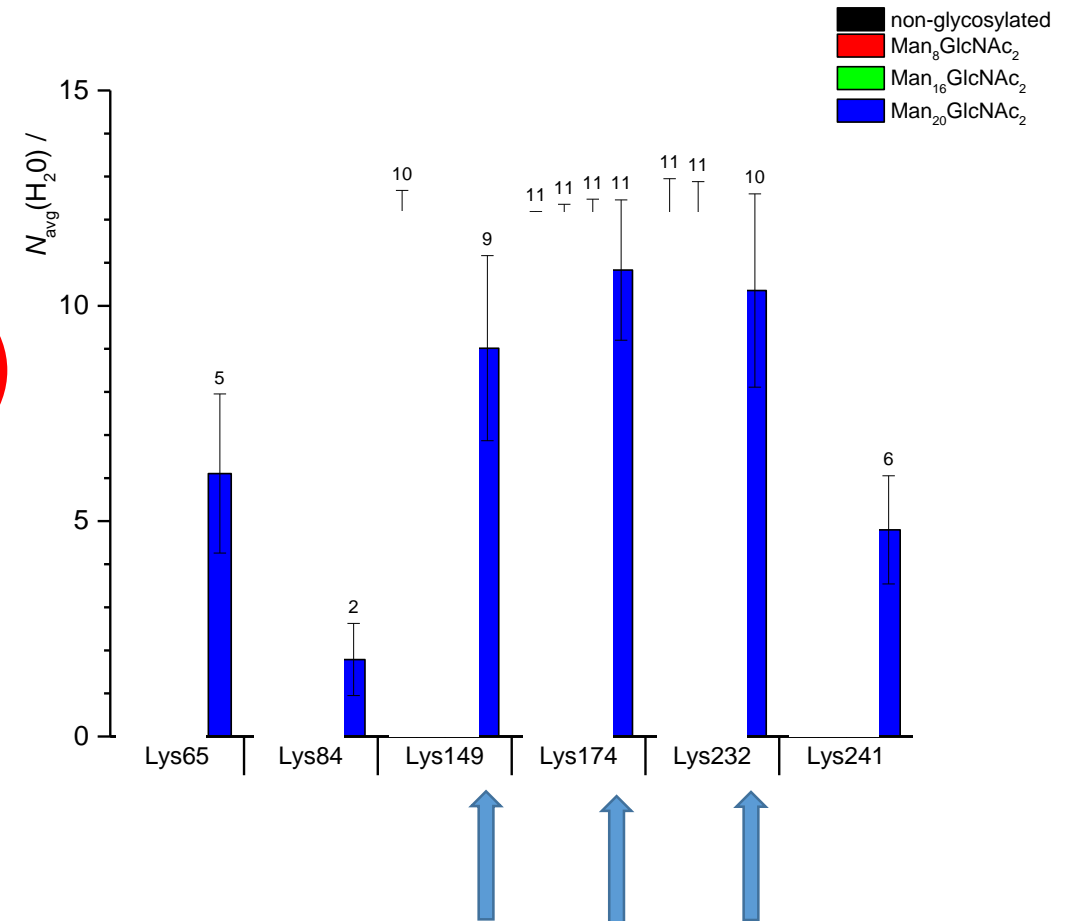
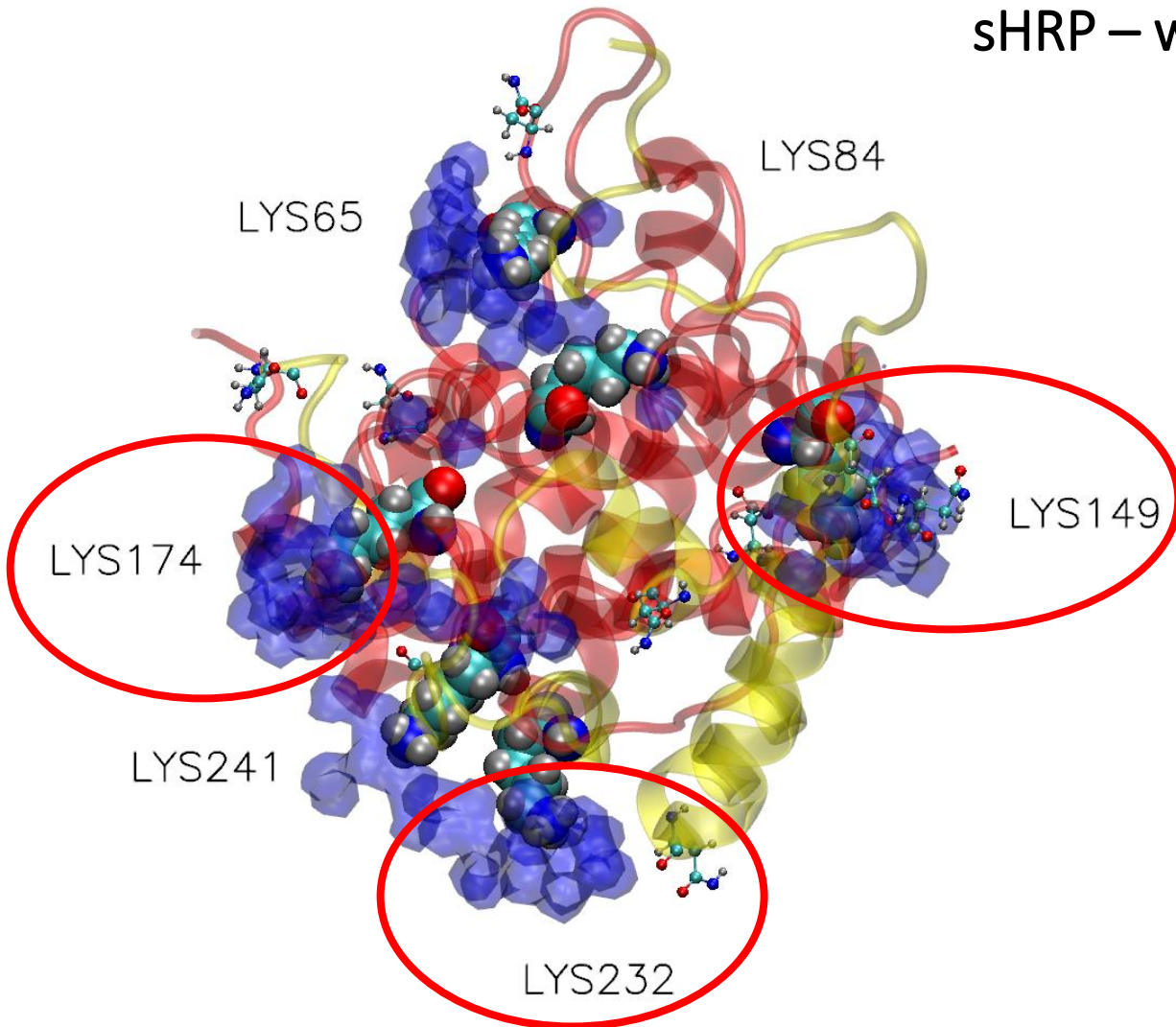
3.3.1 Lysine availability for oligonucleotide attachment

HRP (WT) – with full glycosylation



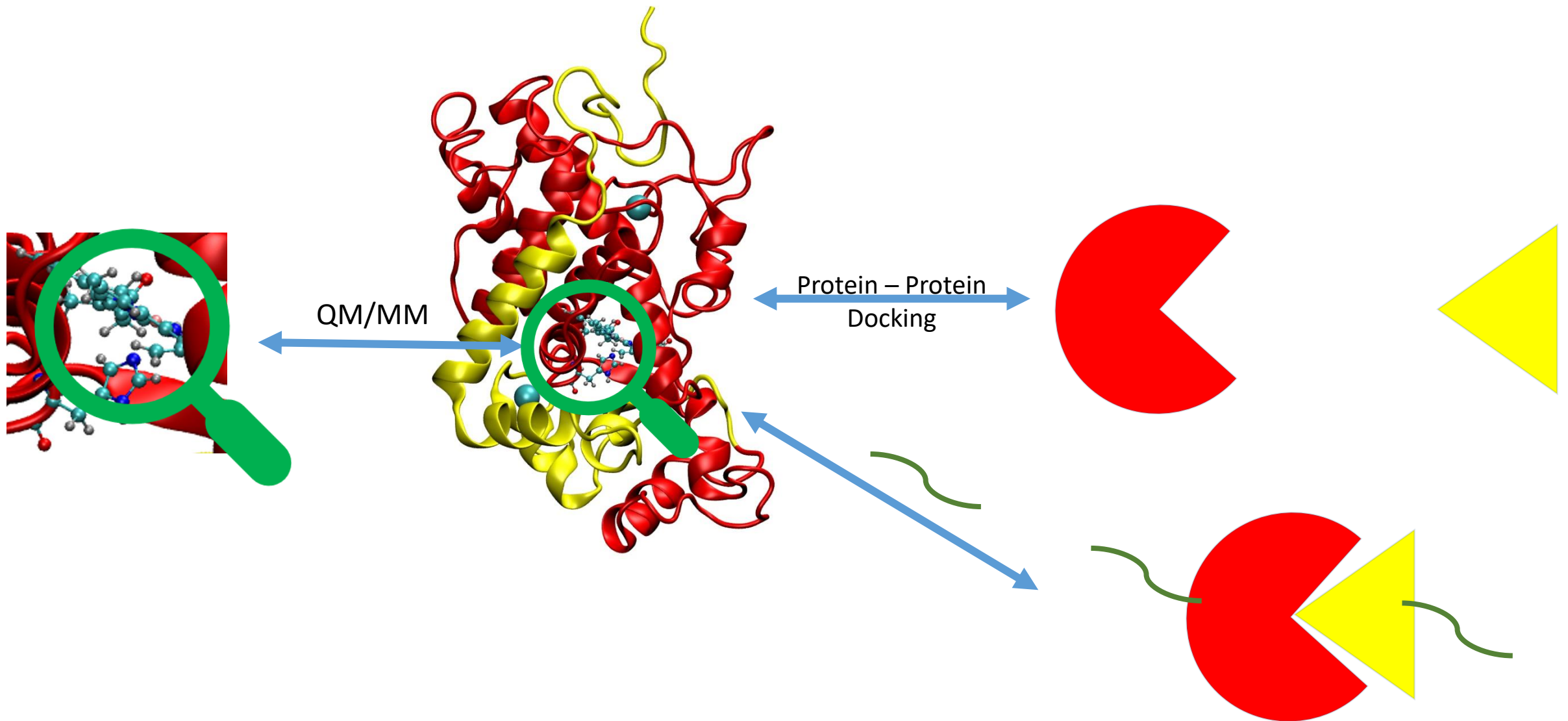
3.3.1 Lysine availability for oligonucleotide attachment

sHRP – with full glycosylation



Future steps

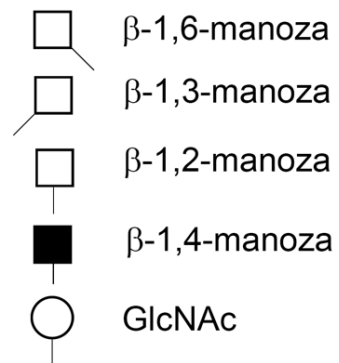
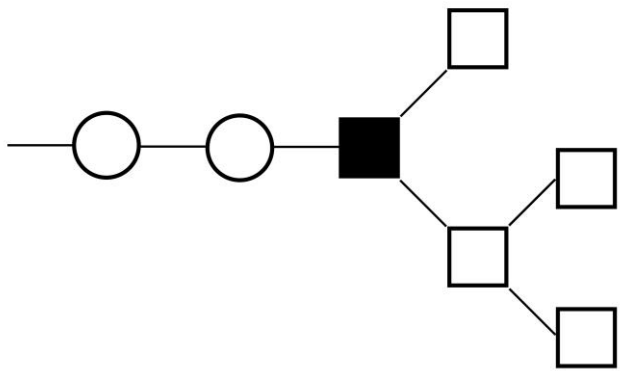
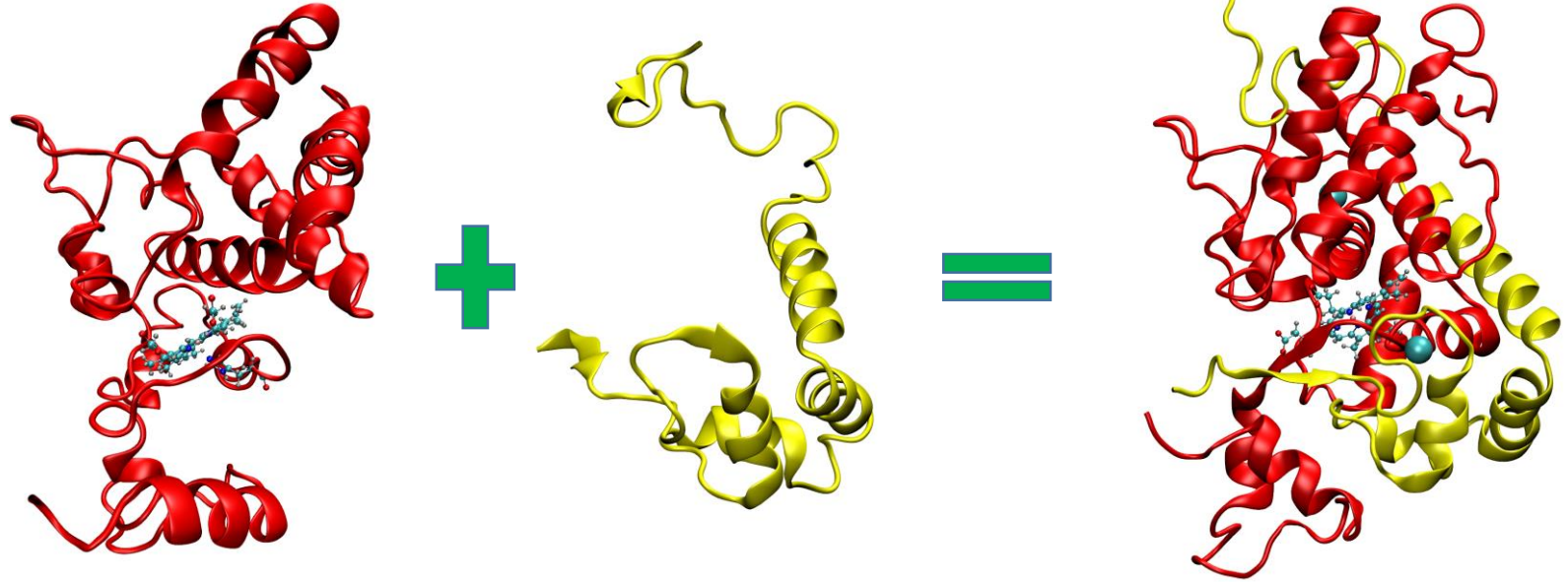
Future steps



Preparation for protein-protein docking

HRP information:

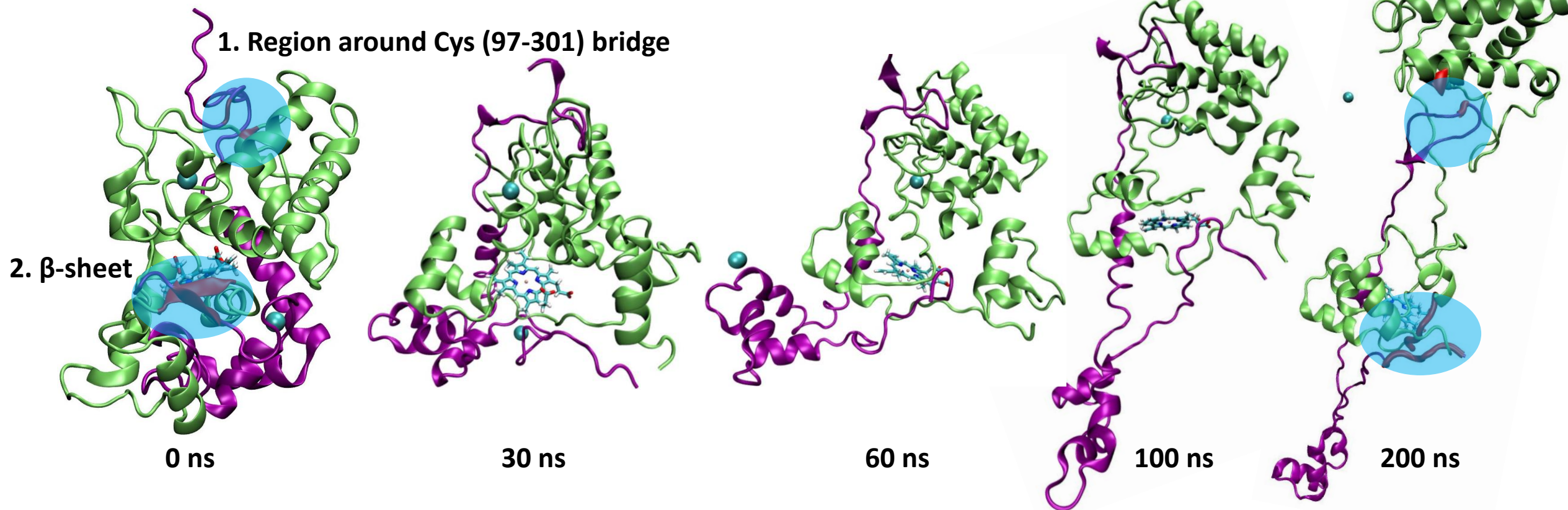
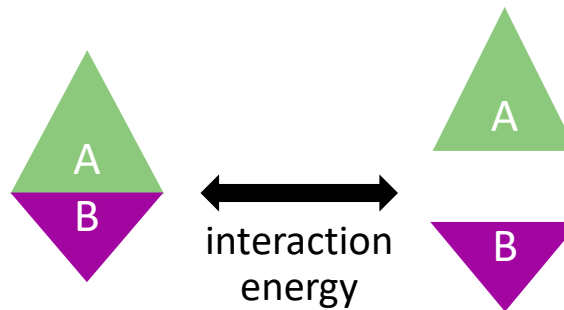
- Forms:
 - HRP
 - sHRP
 - **sHRPa**
 - **sHRPb**
- without glycolisation
- $\text{Man}_5\text{GlcNAc}_2$



Metadynamics

Enhanced MD simulation – Metadynamics

- interaction energy
- pulling center of mass (COM) of both subunits

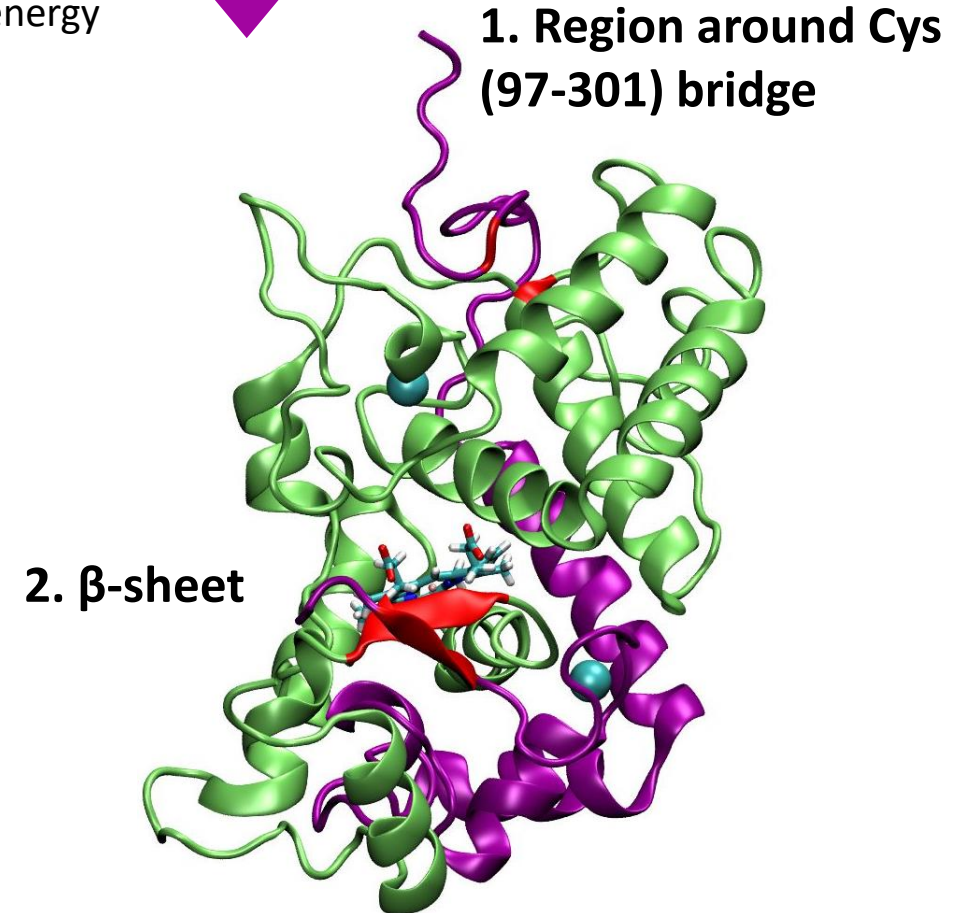
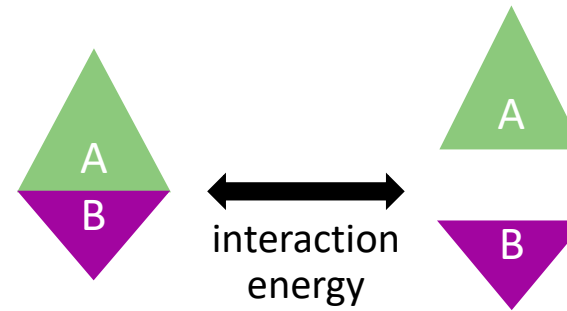


Metadynamics

Enhanced MD simulation – Metadynamics

- interaction energy
- pulling:
 1. Center of mass (COM) of both subunits
 2. Region around Cys (97-301) bridge
 3. β -sheet

simulations in progress...

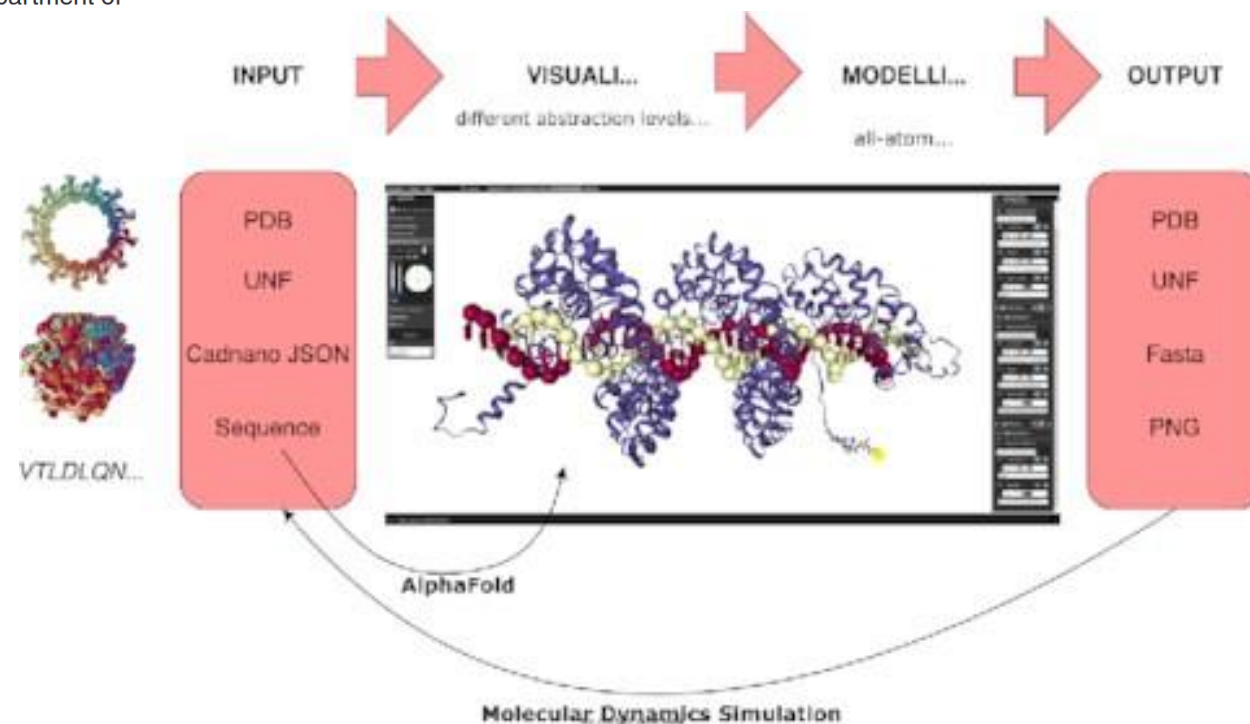


CATANA: an online modelling environment for proteins and nucleic acid nanostructures

David Kuřák^{1,2,3}, Lucas Melo^{1,2}, Fabian Schroeder^{1,2}, Zoe Jelic-Matošević⁴, Natalie Mutter¹, Branimir Bertoša⁴ and Ivan Barišić^{1,2,*}

¹Molecular Diagnostics, AIT Austrian Institute of Technology, 1210 Vienna, Austria, ²Eko Refugium, 47240 Slunj, Croatia, ³Visitlab, Faculty of Informatics, Masaryk University, Brno 602 00, Czech Republic and ⁴Department of Chemistry, Faculty of Science, University of Zagreb, Horvatovac 102a, HR-10000 Zagreb, Croatia

Received March 23, 2022; Revised April 19, 2022; Editorial Decision April 23, 2022; Accepted May 10, 2022



Advantages

- **User friendly** with all features needed for modelling and visualization
- **Online** tool
- Implementation of **Alphafold**
- Modelling (building and visualization) of a **various range of molecular systems** from small (several oligonucleotides) to large (protein-DNA complexes, DNA origami)
- Relaxation of unrealistic structures
- Preparation of systems for **all-atom MD simulations** with different force fields
- Preparation of systems for **coarse-grained MD simulations**

Disadvantages- Opportunities for improvement:

- Protonation
- Trajectory analyses
- non-standard structures for biological systems
- ...

Testing CATANA:

Several systems were prepared for simulations from crystal structures (or using standard procedure) and using CATANA

Approaches

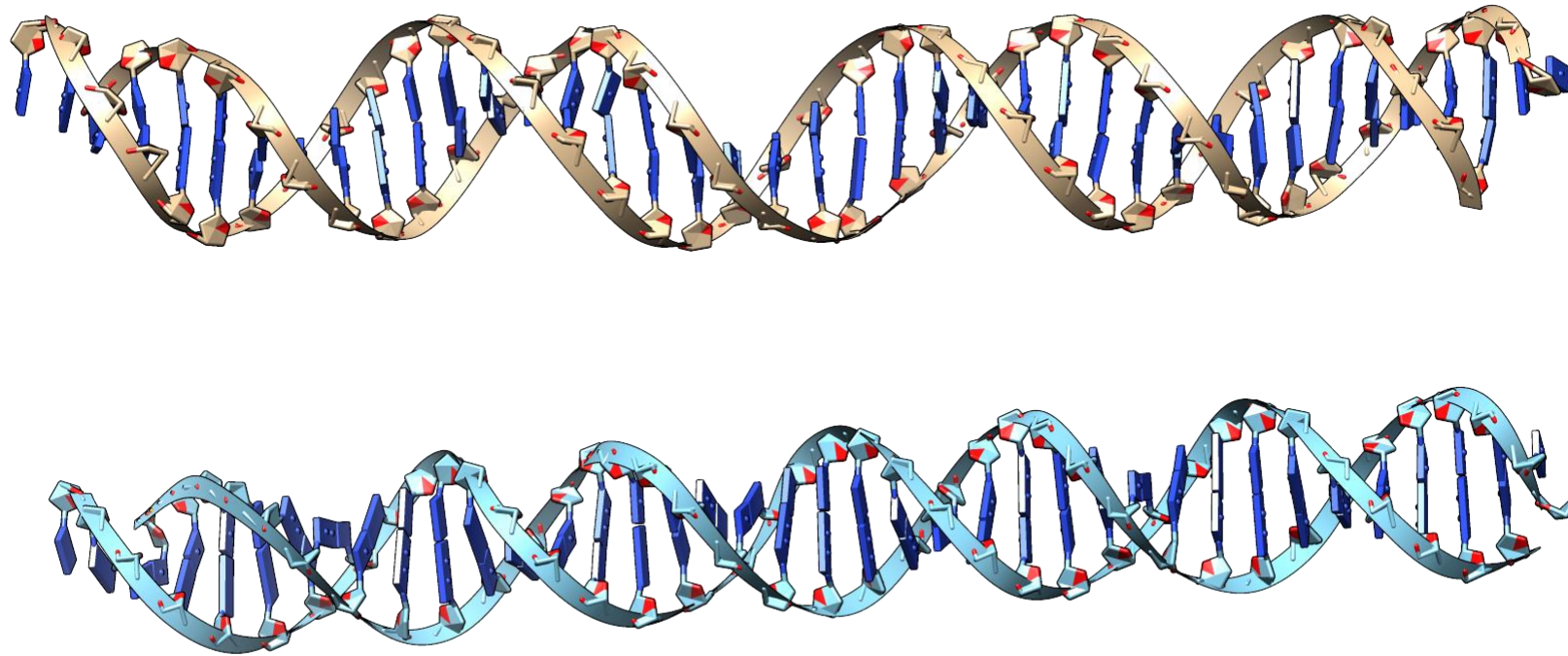
- All-atom MD simulations with different force fields
- Coarse-grained MD simulations

Goals of simulations:

- to test the reliability of structures prepared in CATANA as the starting structures for MD simulations

➤ DNA sequence – crystal structure vs CATANA

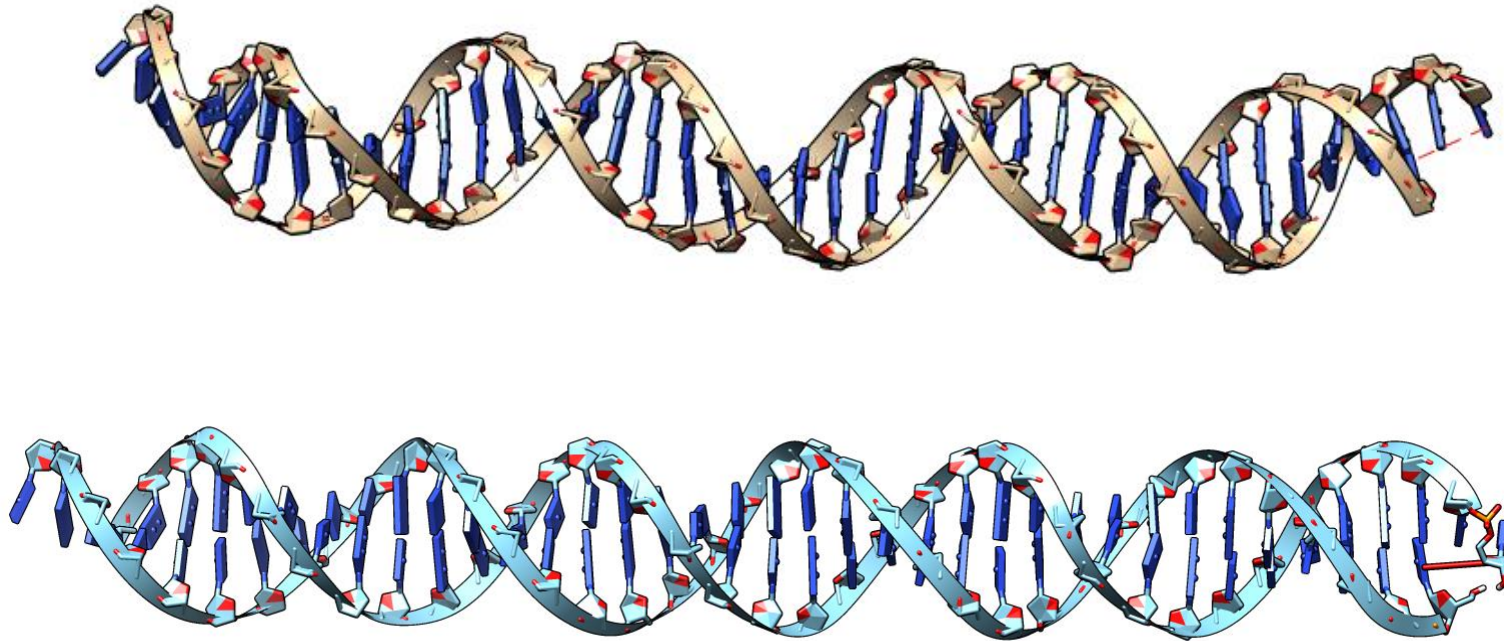
- Extracted from the **crystal structure** of TAL protein (PDB ID: 3UGM)



- Built in **CATANA**

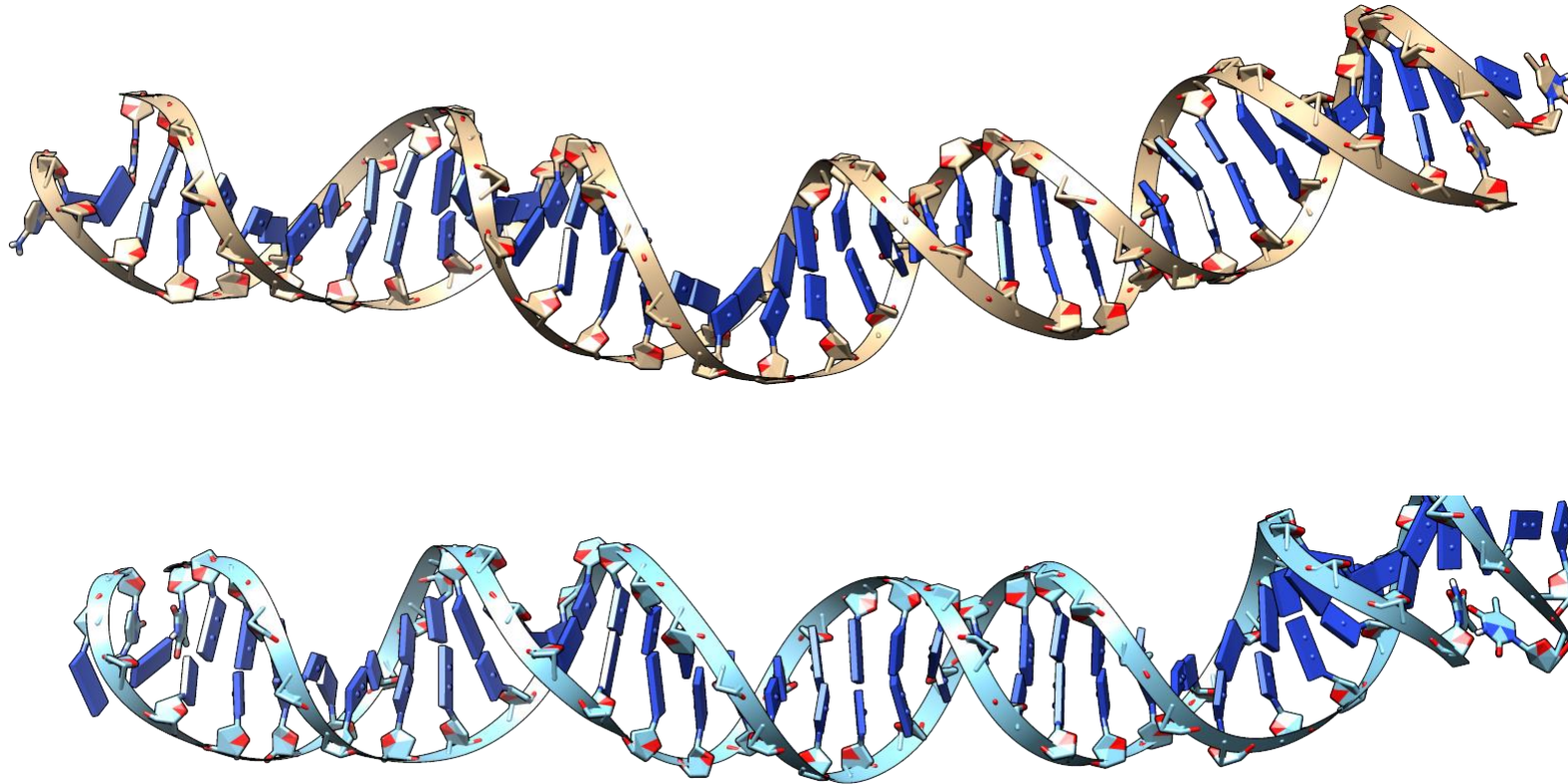
➤ DNA sequence – crystal structure vs CATANA

- After geometry optimisation (energy minimisation)

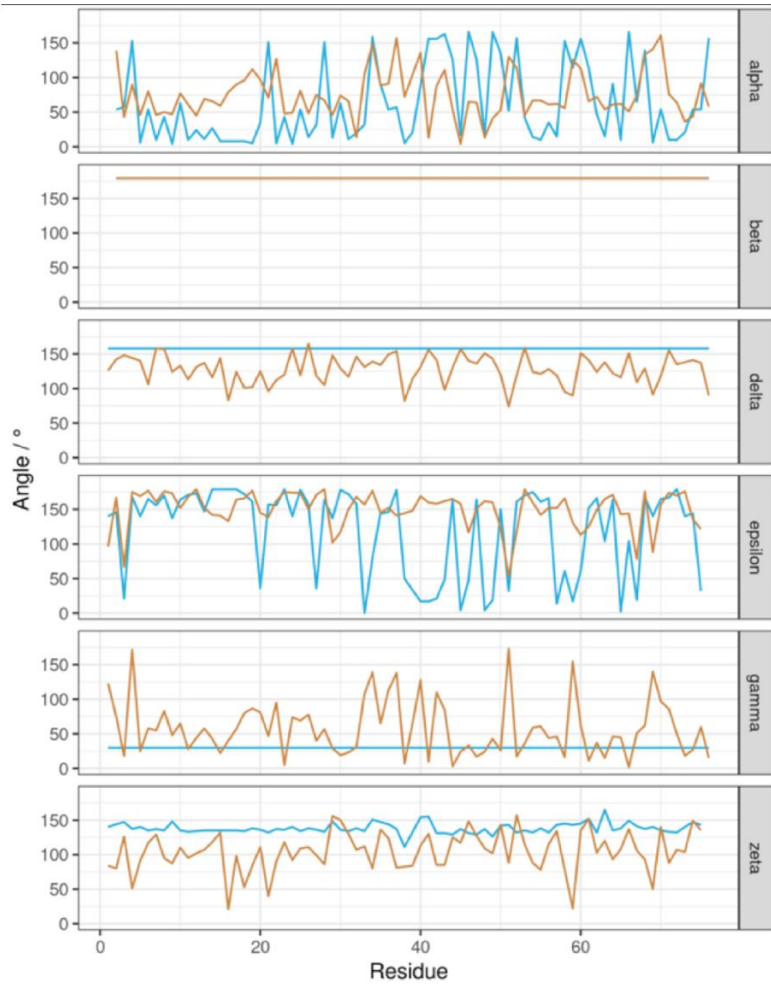


➤ DNA sequence – crystal structure vs CATANA

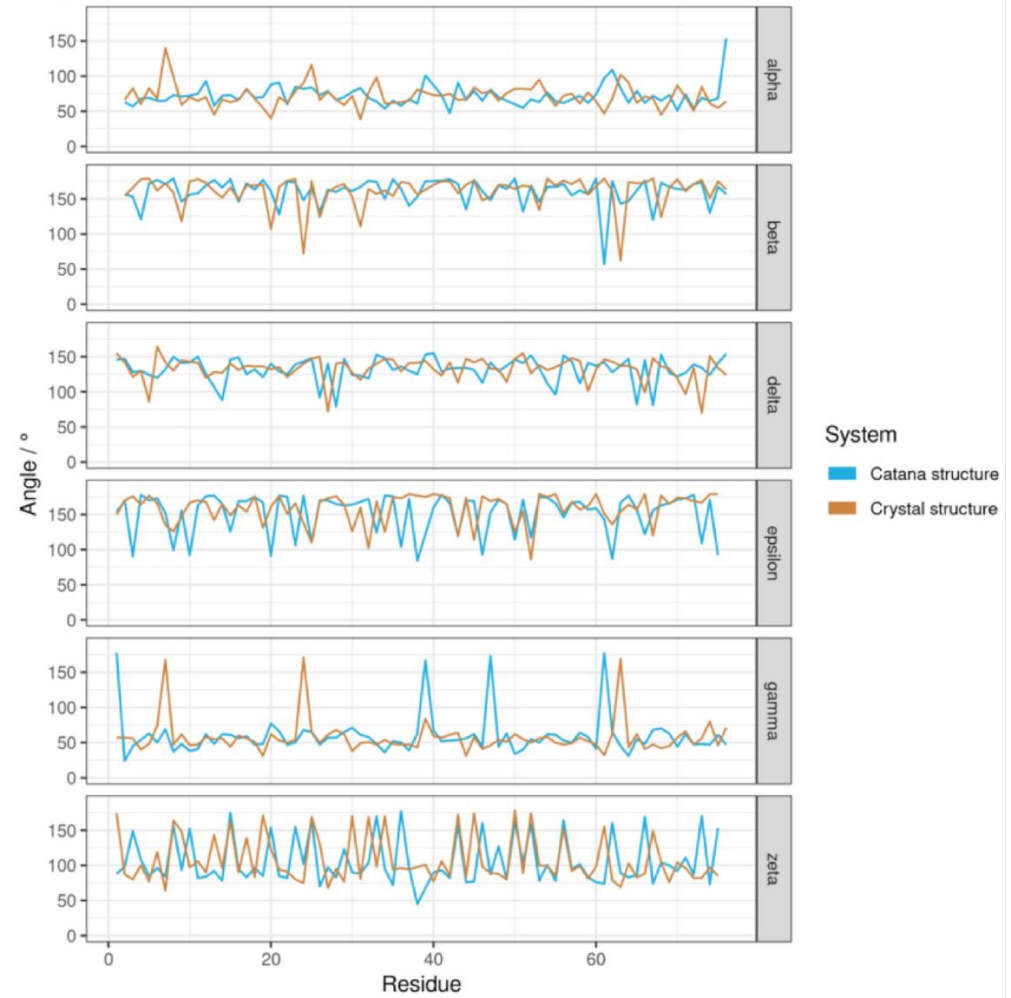
- After 10 ns of MD simulation



DNA sequence – crystal structure vs CATANA



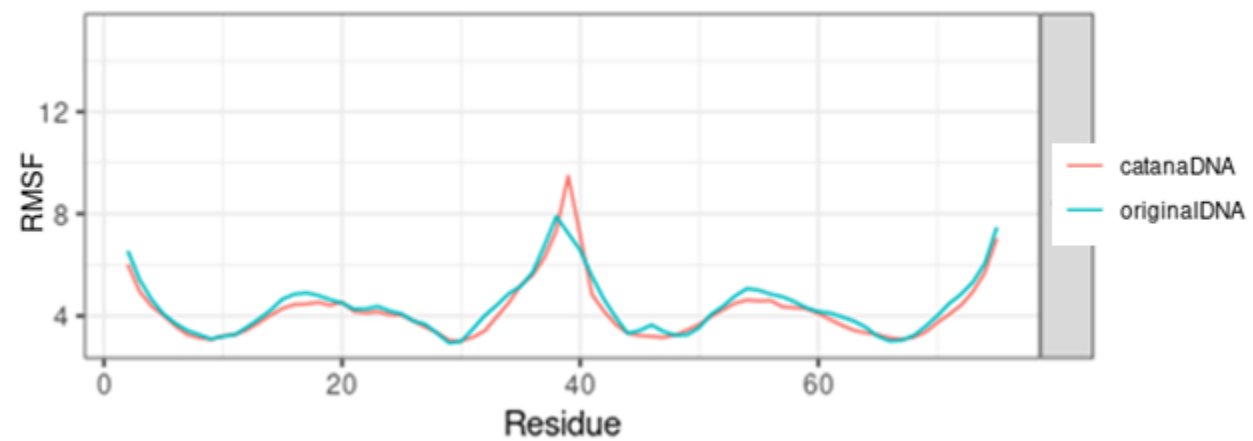
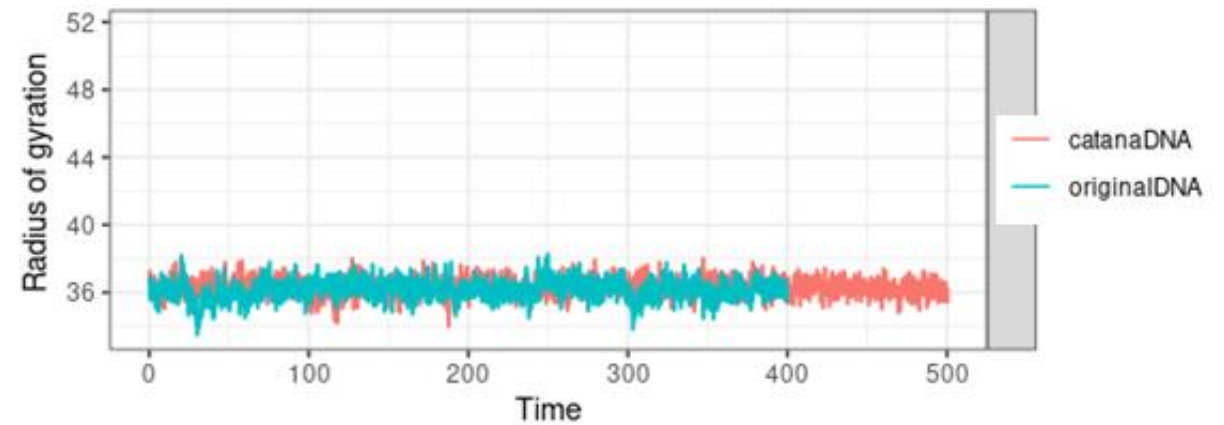
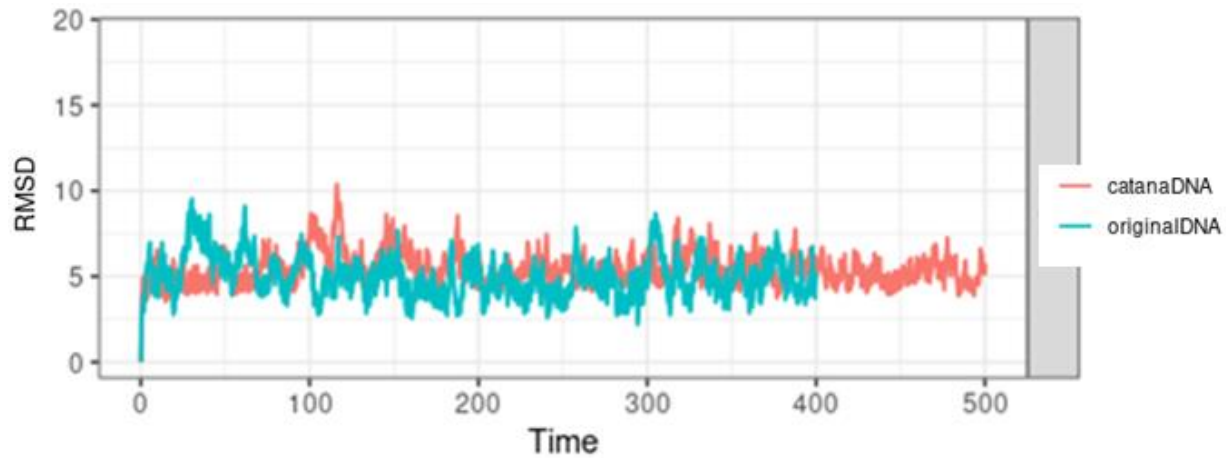
Initial values



After 10 ns of MD simulation

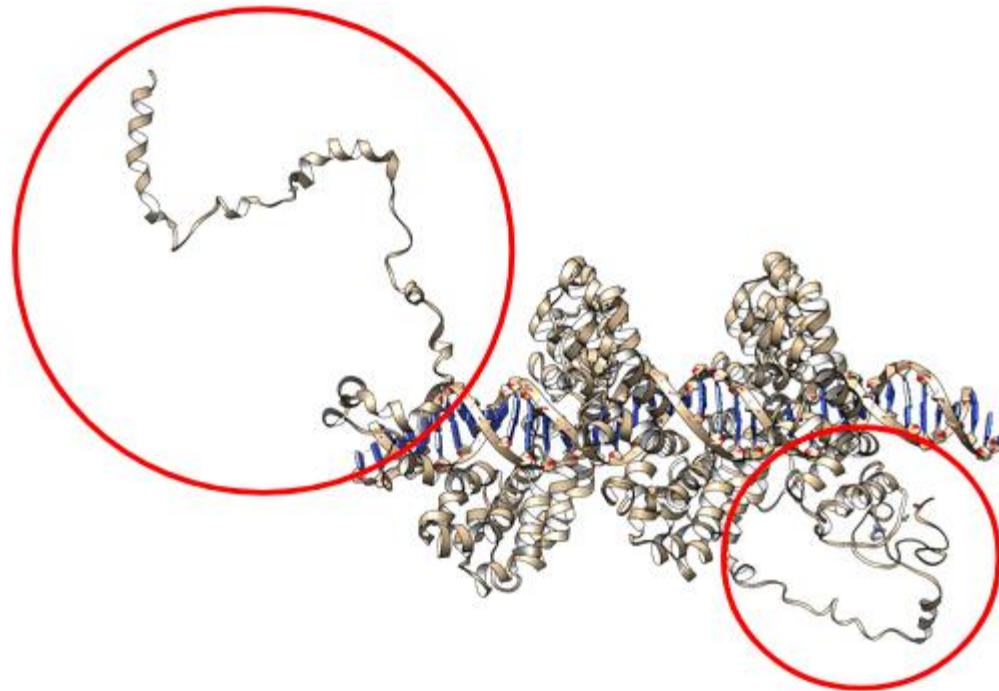
DNA sequence – crystal structure vs CATANA

- MD simulation



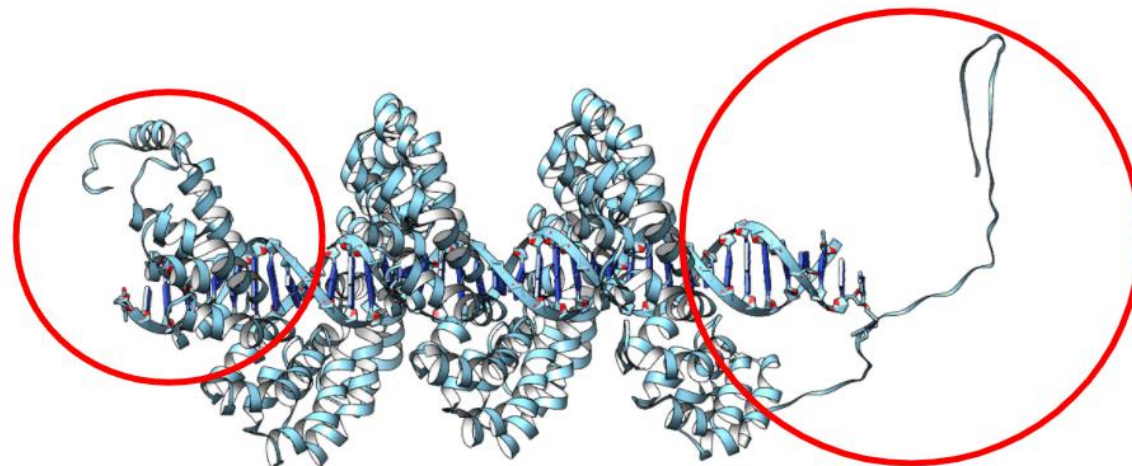
➤ TAL protein in complex with DNA (crystal structure – 3UGM)

- Missing residues modelled in **Charmm GUI**
- All-atom MD simulations



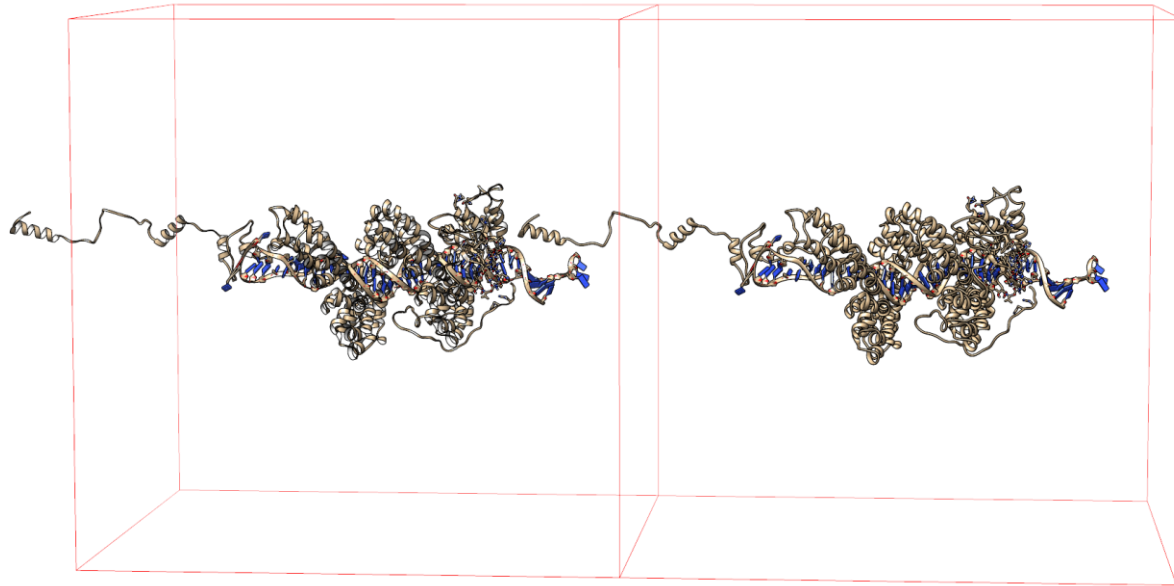
➤ TAL protein in complex with DNA built in Catana

- Missing residues modelled with **Alphafold** (from **Catana**)
- All-atom MD simulations
- In addition, the Catana structure with a **random DNA** sequence was prepared and simulated



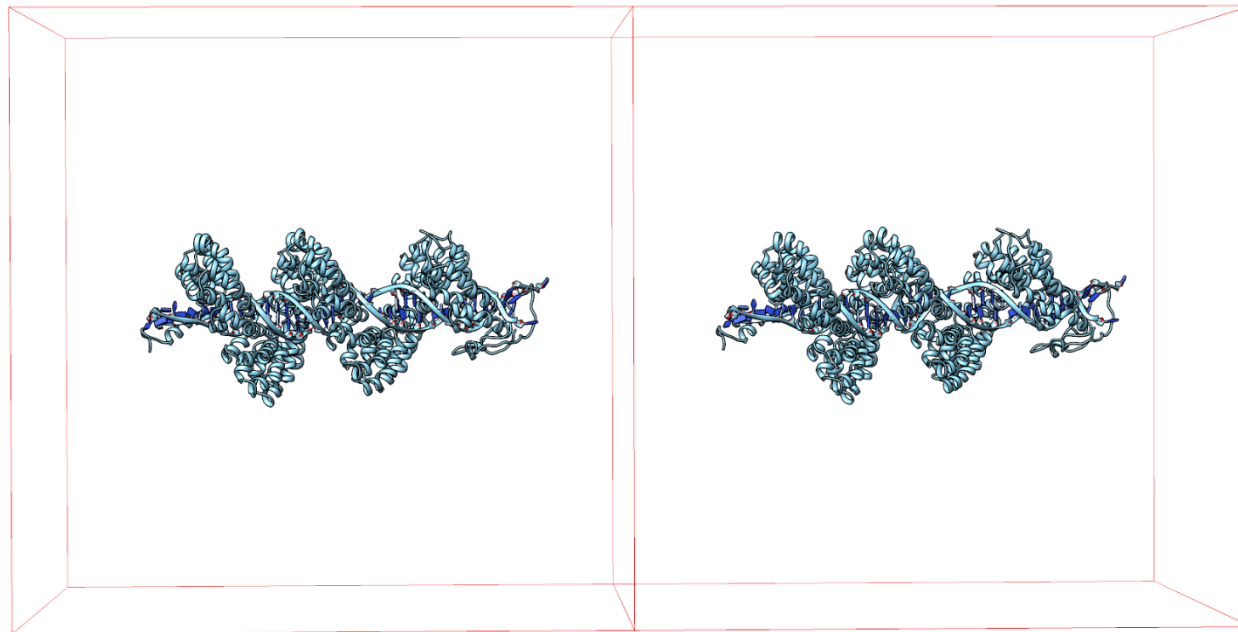
➤ TAL protein in complex with DNA (crystal structure –3UGM)

- Missing residues modelled in **Charmm GUI**
- All-atom MD simulations

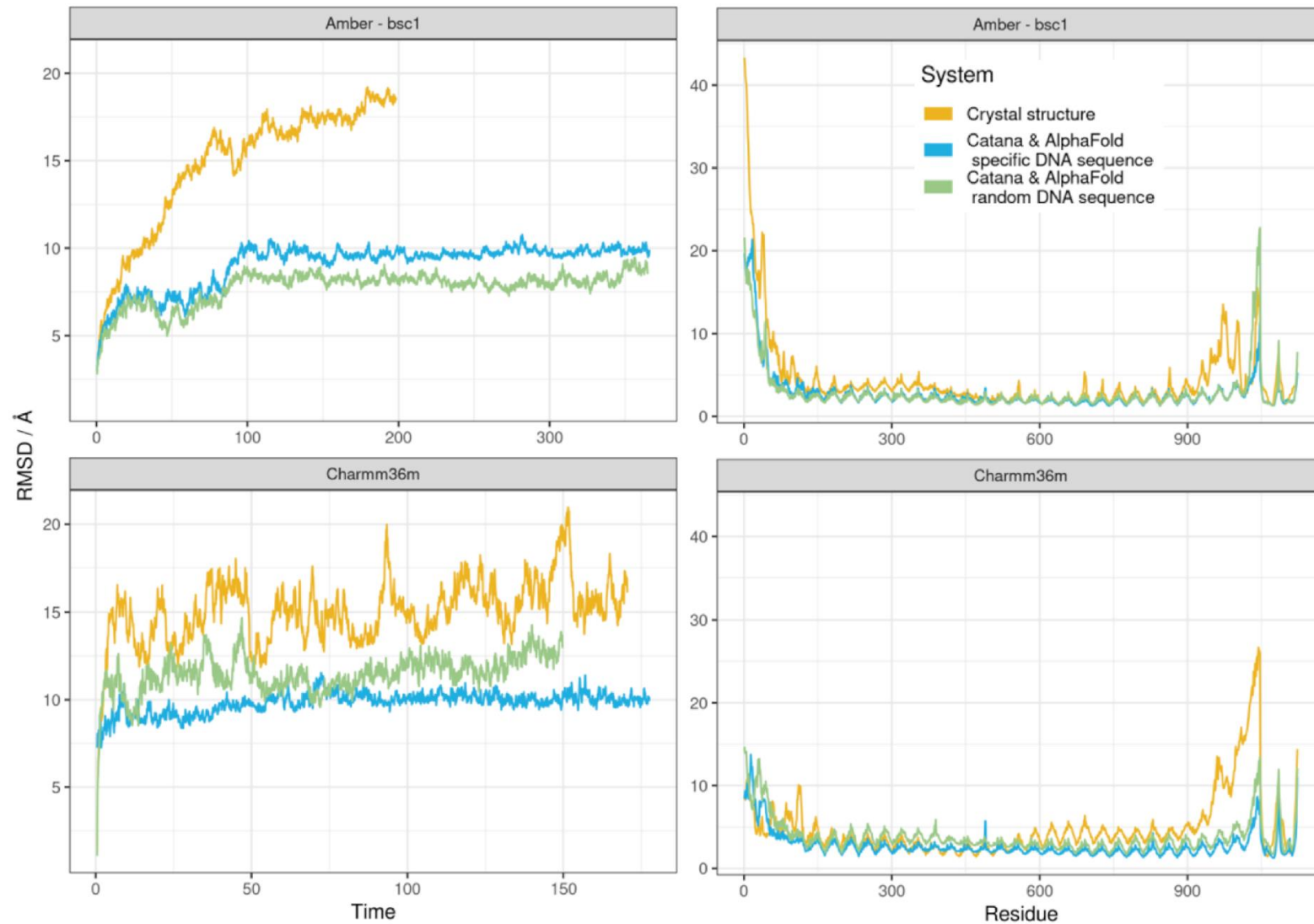


➤ TAL protein in complex with DNA built in Catana

- Missing residues modelled with **Alphafold** (from **Catana**)
- All-atom MD simulations

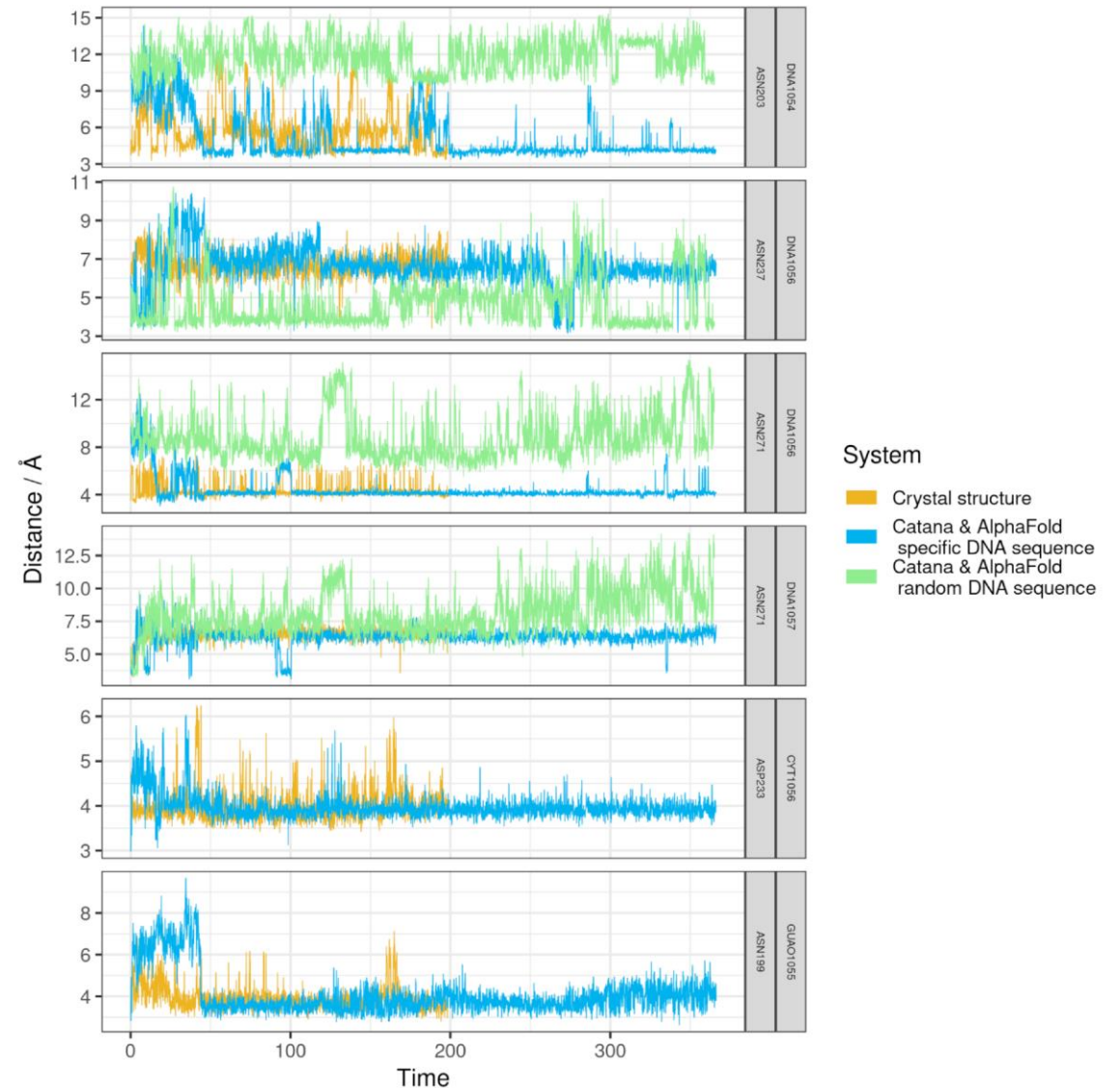


TAL protein in complex with DNA built in Catana



TAL protein in complex with DNA built in Catana

- Protein–DNA specific interactions (**not present with random DNA** sequence, but are present with **Catana** and **X-ray** structures)





All collaborators on the MARILIA project





MARILIA

Your safe water drop

THANK YOU FOR YOUR TIME



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 952110.