



Oslo  
University

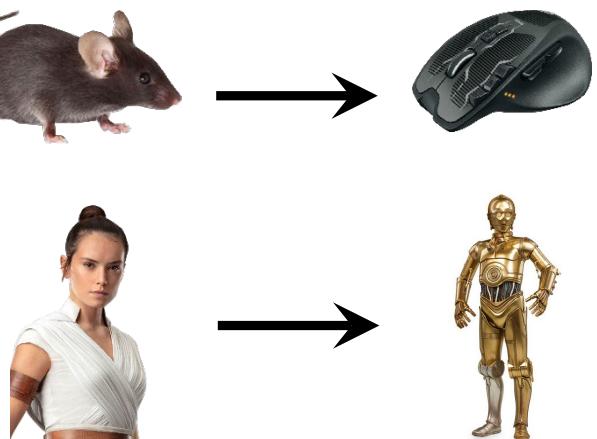
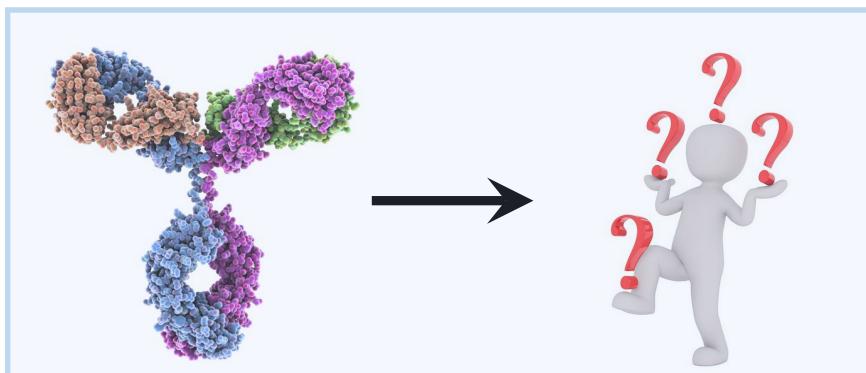
# Absolut!

Unconstrained generation of **synthetic 3D antibody-antigen structures** to guide **development of machine learning methodology** for real-world antibody specificity prediction

[github.com/csi-greifflab/Absolut/](https://github.com/csi-greifflab/Absolut/)

Philippe A. ROBERT

May 2022, AIRR-C Meeting



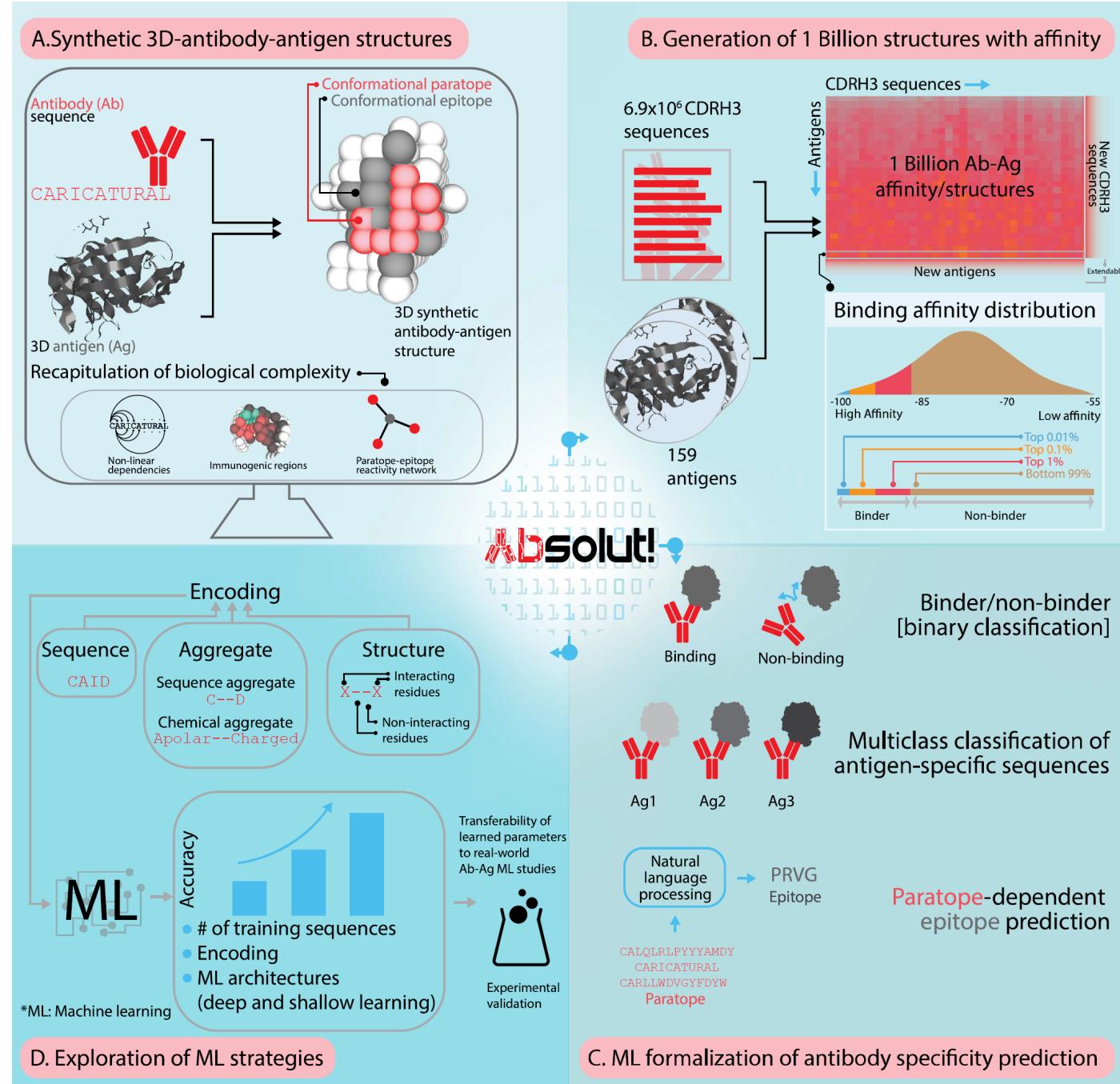
*Victor Greiff* group / Computational and Systems Immunology Lab.

Dpt. of Immunology, University of Oslo

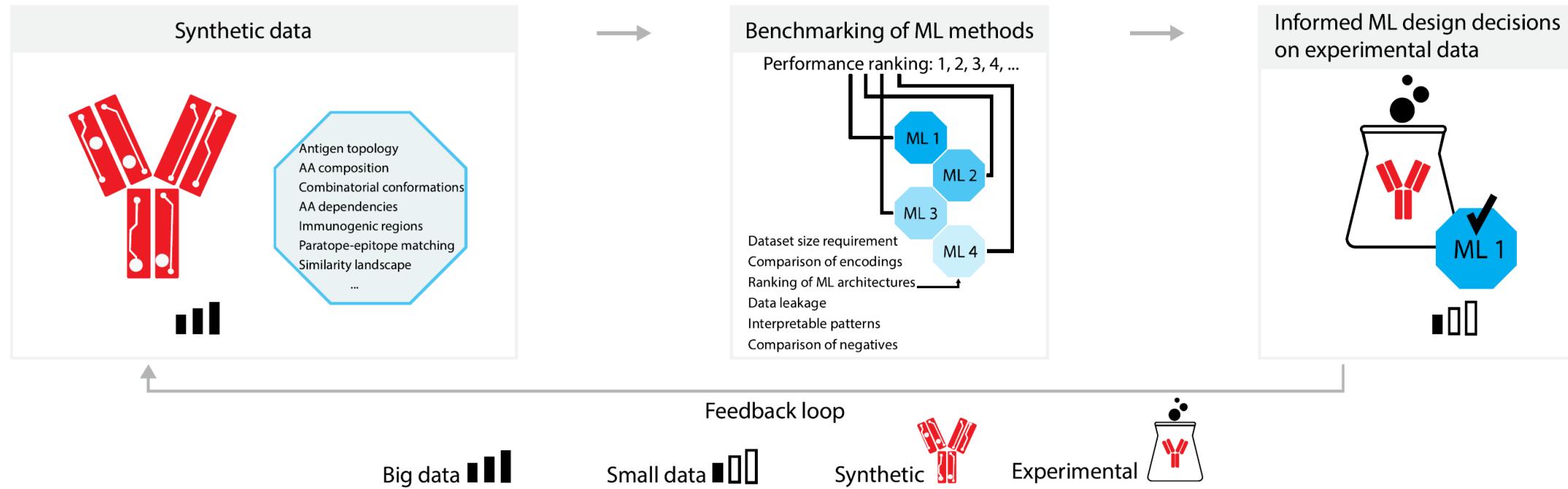
UiO : University of Oslo



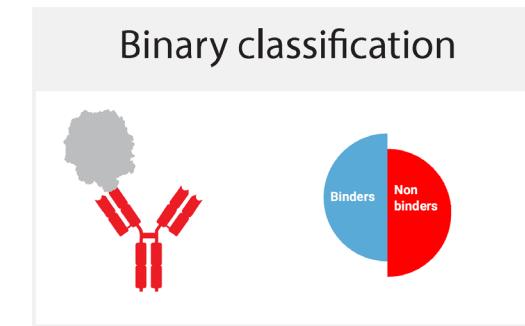
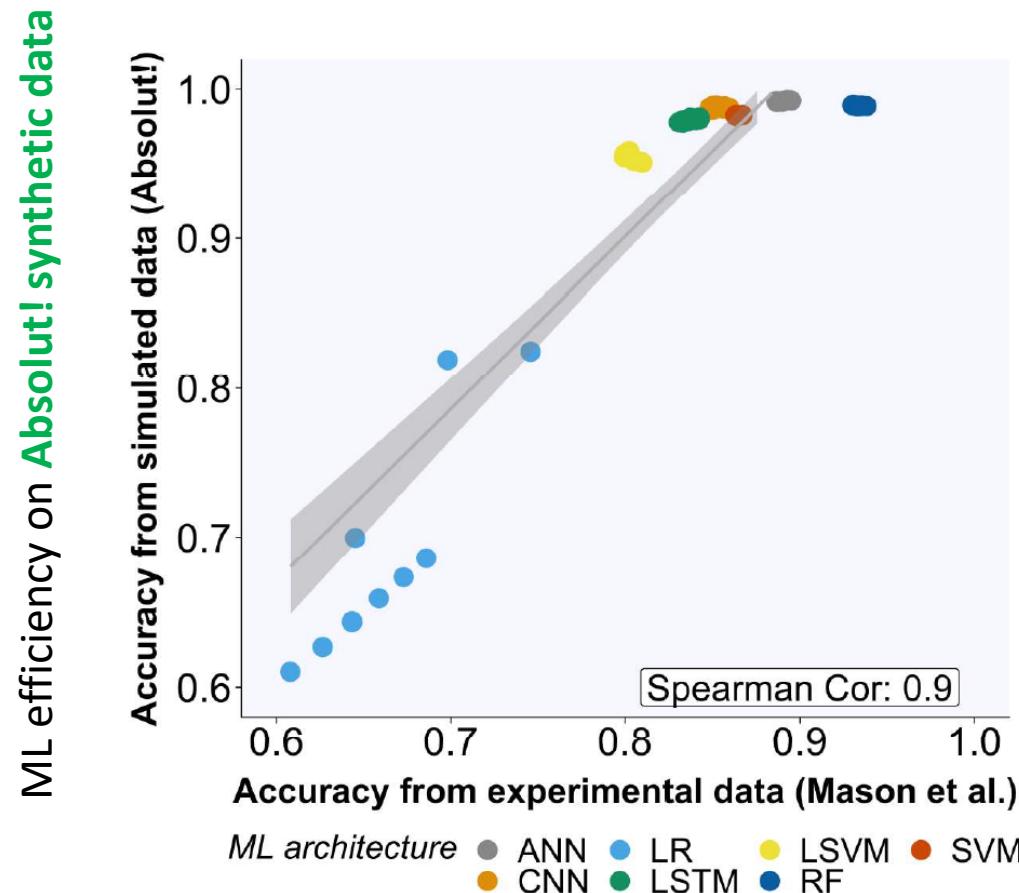
GREIFFLAB  
Immunity | AI | Therapeutics  
Laboratory for Computational  
and Systems Immunology



## Aim: Rank ML methods in the synthetic world for application in the real-world



Aim: Rank ML methods in the synthetic world for application in the real-world



ML efficiency on **real-world data**  
(on high affinity antibody sequences binary classification)

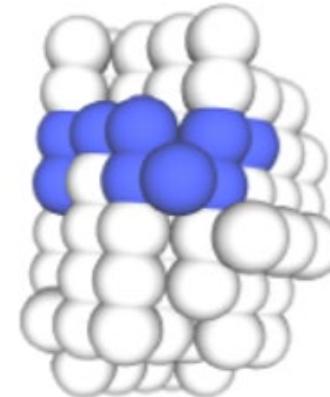
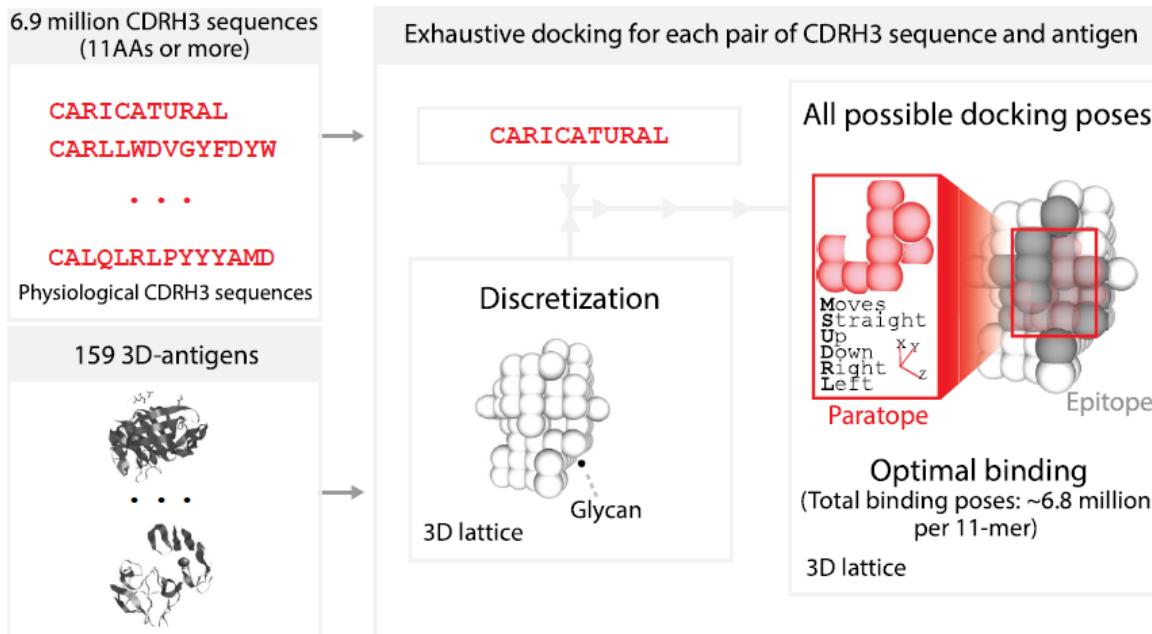
# Absolut! generates synthetic antibody-antigen structures in the lattice world

## Inputs:

CDRH3 sequence  
Antigen structure

## Outputs:

Binding Energy  
Binding structure



UDRUSUDRDU 137187

**Large-scale datasets suitable for benchmarking:**

- Sequence-based ML models
- Structure-based ML models (Amino-Acid level)

## Step 0: Installation

### AbsolutNoLib

Only works with pre-existing 159 antigens

```
sudo apt-get install build-essential #for g++ and make  
git clone https://github.com/csi-greifflab/Absolut  
cd Absolut/src/  
make  
. /AbsolutNoLib
```

### Absolut

Can include more antigens from their PDB ('discretization')

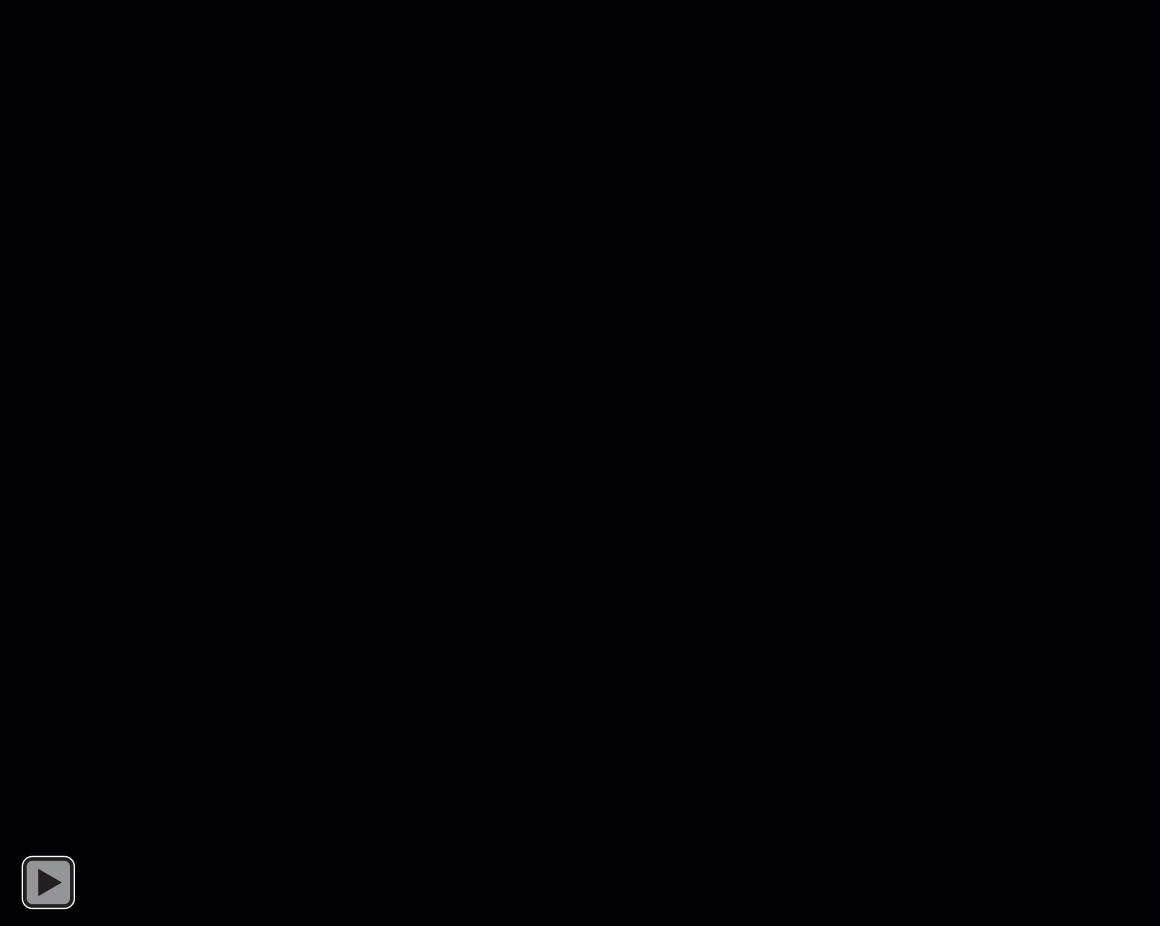
```
sudo apt-get install build-essential #for g++ and make  
sudo apt-get install libgsl-dev  
sudo apt-get install curl  
sudo apt-get install libgsl-dev  
sudo apt-get install freeglut3-dev  
sudo apt-get install qtbase5-dev  
sudo apt-get install qtcreator  
sudo apt-get install libqt5svg5  
sudo apt-get install libqt5printsupport5  
sudo apt-get install rasmol  
# sudo apt-get install python-is-python3
```

```
git clone https://github.com/csi-greifflab/Absolut  
cd Absolut/src/bin/  
qmake .. /Absolut/Absolut.pro  
make  
. /Absolut
```

## Step 0: Installation

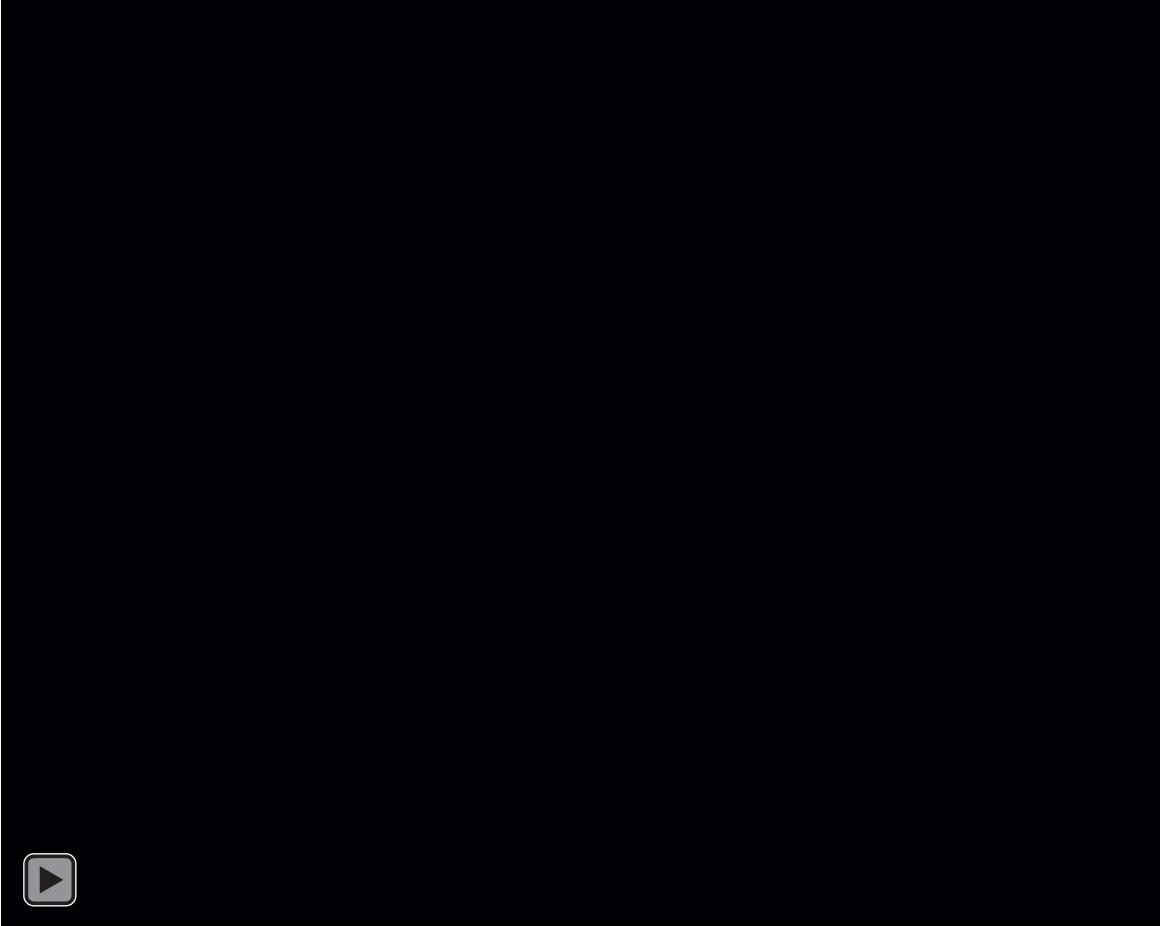
### AbsolutNoLib

Only works with pre-existing 159 antigens

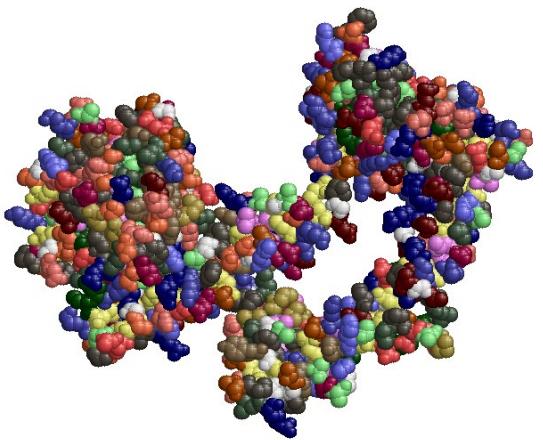


### Absolut

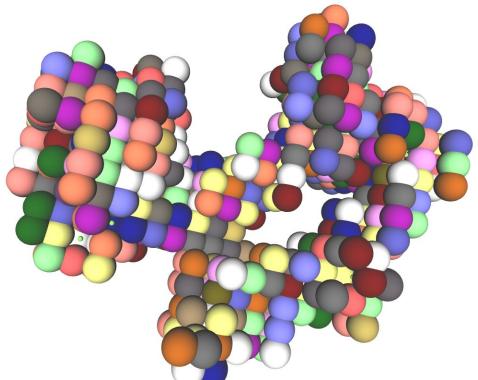
Can include more antigens from their PDB ('discretization')



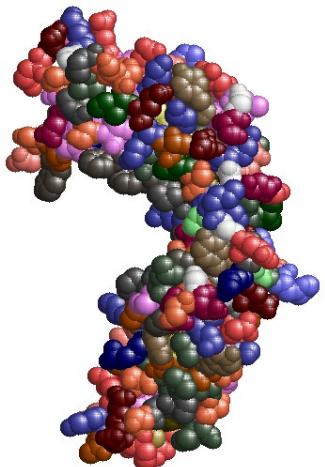
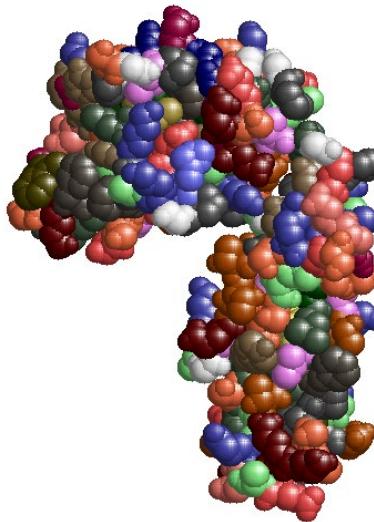
## Step 1: Discretization of real antigen structures



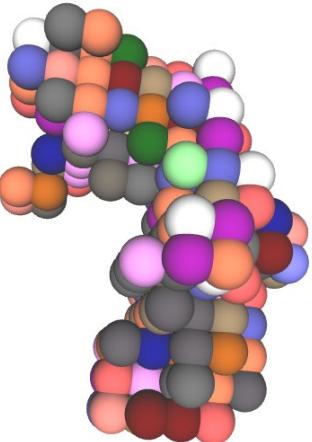
EGFR (PDB 1YY9)



'factor IV' (PDB 1NLO)



RA autoantigen (PDB 1ADQ)

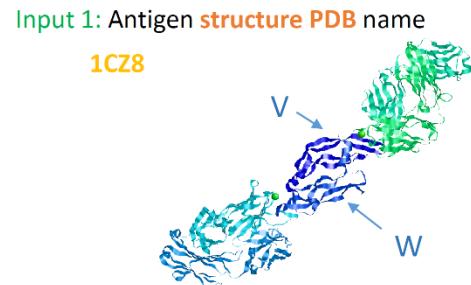


MSP plasm. Falcip. (PDB 1OB1)

## Step 1: Discretization of real antigen structures

with GUI (need full **Absolut**)

`./Absolut discretize`



Input 2: Chains in the PDB to discretize  
**VW** One letter per chains, no separation

(optional) Input 3: Resolution of the 3D lattice  
Default value: 5.25 Å

(optional) Input 4: Method of discretization

Tells which points from the PDB are transformed:

**CA** for Carbon Alpha

**CoM** for Centroid center of the side-chain only,

**FuC** for fused center of the whole AA [default value]

without GUI (still need the full **Absolut**)

`./Absolut discretize 1CZ8 VW 5.25 FuC`

or (step by step):

`./Absolut discretize_delete_insertions 1ADQ A`  
output: `1ADQdeIns.pdb` (or `1ADQ.pdb` if no insertions)

`./Absolut discretize_latfit_to_discrete_pdb 1ADQ A 1ADQdeIns.pdb 5.25 FuC`

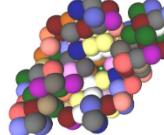
output: `1ADQ_AdiscretizedFuC5.25.pdb`

`./Absolut discretize_discrete_pdb_to_lattice 1ADQ A 1ADQ_AdiscretizedFuC5.25.pdb`

output: `1ADQ_Ainlattice.txt`

Output 1: Best representation of the selected chains in the 3D lattice:

**1CZ8\_VWInLattice.txt**



Number of subchains: 2

Structure of each subchain:

133152 SUDDLLUURRDRDSDLUDRDRDDLUULUDRRRLUSRDUDDUULDRURRLUDDLRLRUDLRLRSSDLRSSLRSRU  
LDRULRLSRU  
133160 USRSRUUDDLLLUUDLSURUDLSUURRLURRLUDRUULSDRSLSUDLRURUUDRLSURDRRLUSDRDLUUDURSLR  
LRDRLRLURR

AA sequence (concatenated, all subchains):

VVKFMVDVYQRSYCHPIETLVDFQEYYPDEIEYIFKPSCVPLMRCGGCCNDEGLECPVTEESNITMQIMRIKPHQGQHIGEMSFLQHNKCECRPKVVKFMDVYQ  
RSYCHPIETLVDFQEYYPDEIEYIFKPSCVPLMRCGGCCNDEGLECPVTEESNITMQIMRIKPHQGQHIGEMSFLQHNKCECRPK

> List additional inaccessible positions

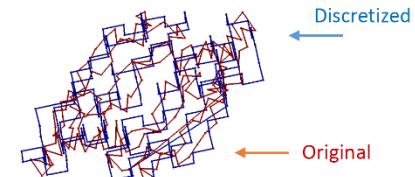
124833	33 30 30	124898	34 31 30
128929	33 30 31	128999	39 31 31
132900	36 28 32	132903	39 28 32
133028	36 30 32	133031	39 30 32
137122	34 30 33		

C++ code to add this structure in the library:

string agStruct = "SULDDLLU...";

**1CZ8discretized5.25FuC.pdb**

PDB containing the original chains in free space  
and their lattice version [direct output of latfit]



Side Outputs:

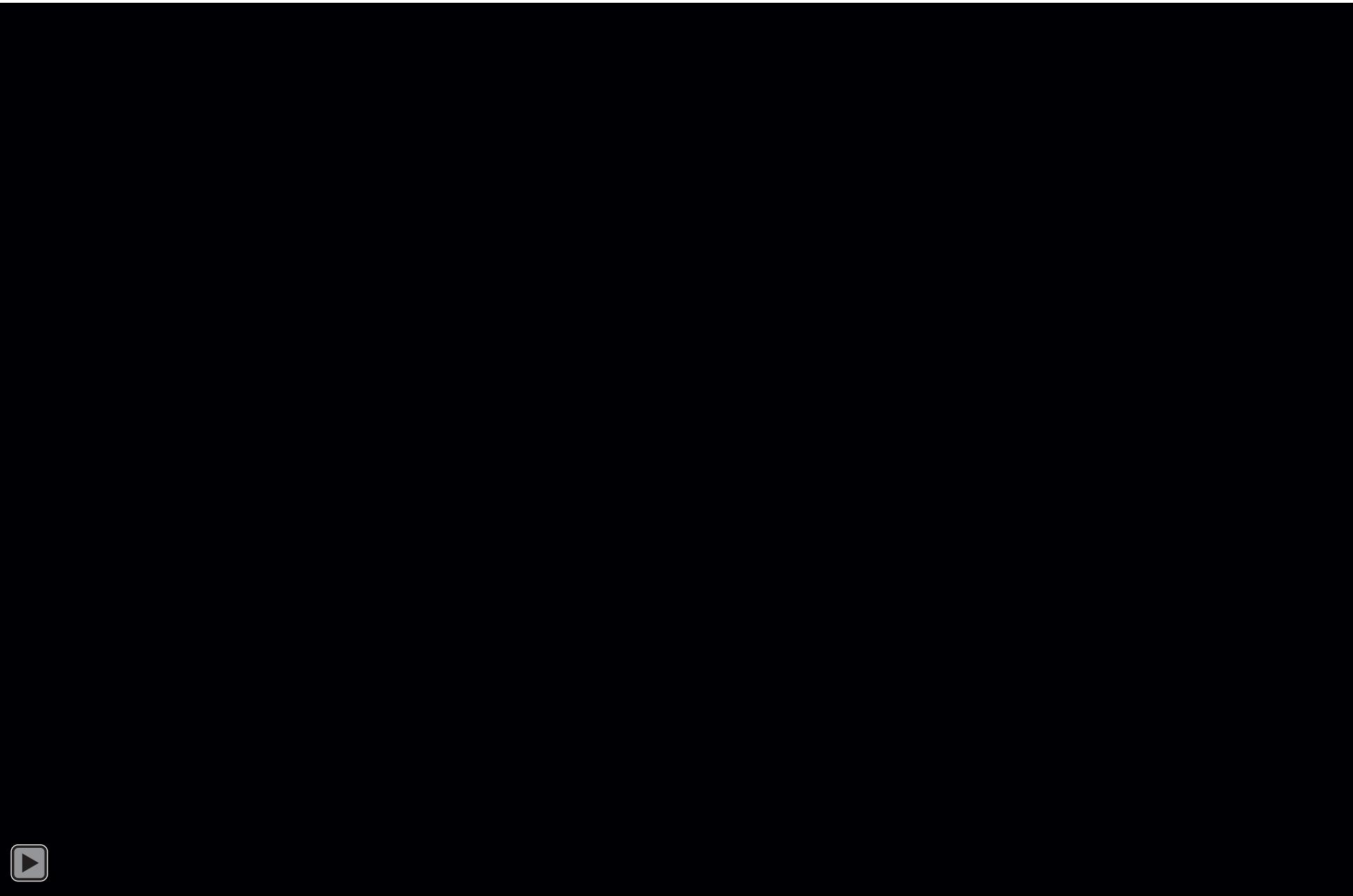
**1CZ8deIns.txt**

Original PDB without insertions

**1CZ8\_VWprepared.pdb**

Only chains of interest from original PDB [used as input for latfit]

## Step 1: Discretization of real antigen structures



## Step 1: Adding the new antigens in the library

- Find and open Absolut/src/antigenLib.cpp
- Copy the C++ code obtained from the discretization using the GUI or the inLattice.txt file
- Has the antigen 1 loop or more?

Yes? => Paste the code in the single-loop part (line ~183)

No? => Paste the code in the multiple-loops part (line ~1278)

- Recompile Absolut:

```
cd src/  
rm AbsolutNoLib  
make
```

```
cd src/bin  
rm Absolut  
make
```

