

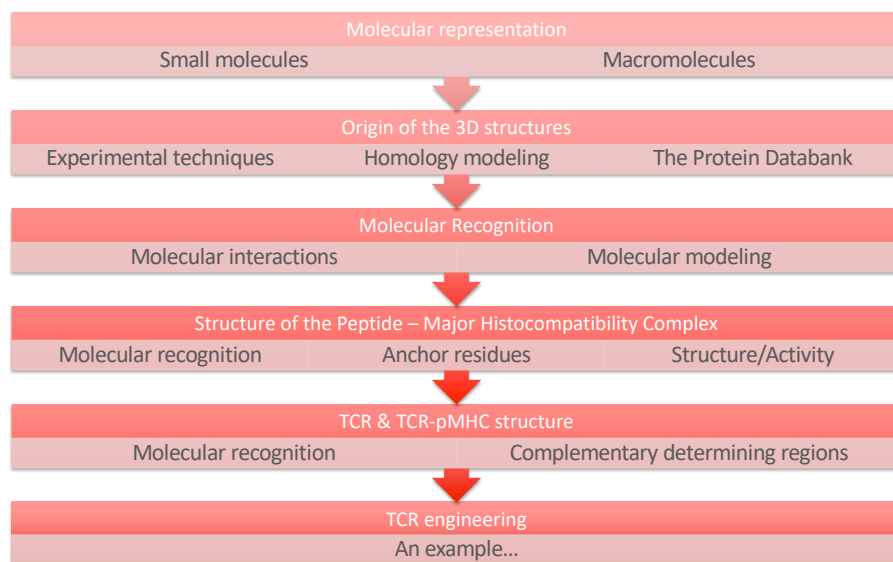
1

## Lecture and Practice Proceedings & Objectives

- Tutors are present to help you. Contact them!!!
  - Get a flavor of molecular structure and modeling
  - Acquire some basic theoretical background
  - Practice the molecular graphics techniques
  - Use them to uncover structure-activity relationship
  - Get a detailed knowledge about the 3D structures of pMHC and TCR/pMHC
- ➔ You should be able to perform simple tasks of molecular graphics and analysis

2

# Lectures & Practices Agenda



3

## Installing UCSF Chimera

In this lecture/practical you will use the software UCSF Chimera for 3D structure visualisation and analysis.

This software is:

- free for teaching or academic research
- available for the most current platforms (Windows, Mac, Linux)
- open source (you can modify it for your needs if you know how to code in python. This is out of the scope of this lecture).

You can download the latest production release here:

<https://www.cgl.ucsf.edu/chimera/download.html>

**Please, install this software on your machine.**

It will be mandatory for the practicals, but also useful for the theoretical lectures

### Download Chimera

- Daily Builds
- Snapshot Releases
- Unsupported Releases
- Old Releases
- Bug Tracking System
- Licensing Information
- Experimental Chimera Features
- Plug-ins on the Web
- Graphics Driver Bugs
- Benchmark Results
- Chimera Source Code
- Cytosin Source Code

**Tip:** We recommend ChimeraX for higher performance and many new features instead of legacy Chimera.

### Current Production Releases

- See the release notes for a list of new features and other information.
- For more recent changes, use the snapshot and daily builds; they are less tested but usually reliable.

#### 64-bit Releases:

| Platform                 | Installer, Size, and Checksum  | Date         | Notes   |
|--------------------------|--|--------------|---|
| Microsoft Windows 64-bit | chimera-1.17.3-win64.exe<br>Size: 15326400 bytes<br>MD5: 8f823a5856e5e0e0e1155295a8a8e9        | Jul 06, 2023 | Instructions<br>Documentation<br>Runs on Windows 7 or later.      |
| Mac OS X 64-bit          | chimera-1.17.3-macos64.dmg<br>Size: 16328112 bytes<br>MD5: 8f823a5856e5e0e0e1155295a8a8e9      | Jul 06, 2023 | Instructions<br>Documentation<br>Runs on Mac OS X 10.12 or later. |
| Linux 64-bit             | chimera-1.17.3-linux_x86_64.bin<br>Size: 15326400 bytes<br>MD5: 8f823a5856e5e0e0e1155295a8a8e9 | Jul 06, 2023 | Instructions<br>Documentation<br>Compiled on CentOS 5.11.         |

• 32-bit releases are no longer supported.

#### Daily Builds

- New builds are made when the code changes.
- They are untested but are usually reliable and include new bug fixes not in the production releases.

#### 64-bit Builds:

| Platform                 | Installer, Size, and Checksum   | Date         | Notes  |
|--------------------------|---|--------------|--|
| Microsoft Windows 64-bit | chimera-alpha-win64.exe<br>Size: 15326400 bytes<br>MD5: 8f823a5856e5e0e0e1155295a8a8e9                | Feb 02, 2024 | (See production version for installation instructions)<br>Runs on Windows 7 or later.<br>Release notes                 |
| Mac OS X 64-bit          | chimera-alpha-macos64.dmg<br>Size: 16328112 bytes<br>MD5: 8f823a5856e5e0e0e1155295a8a8e9              | Feb 02, 2024 | (See production version for installation instructions)<br>Runs on Mac OS X 10.12 or later.<br>Release notes            |
| Linux 64-bit             | chimera-alpha-linux_x86_64.bin<br>Size: 15326400 bytes<br>MD5: 8f823a5856e5e0e0e1155295a8a8e9         | Feb 02, 2024 | (See production version for installation instructions)<br>Compiled on CentOS 5.11.<br>Release notes                    |
| Headless Linux 64-bit    | chimera-alpha-linux_x86_64_ceresna.bin<br>Size: 14828778 bytes<br>MD5: 8f823a5856e5e0e0e1155295a8a8e9 | Feb 02, 2024 | (See production version for installation instructions)<br>For (web) servers. Compiled on CentOS 5.11.<br>Release notes |

• 32-bit builds are no longer supported.

4



## The dedicated web site

This teaching has been conceived to alternate theoretical lectures and practicals, so that you will:

- experiment yourself the visualisation and analysis of protein 3D structures
- get a 3D view of the systems mentioned in the lecture (to prevent being limited by the 2D views in the slides)

To facilitate the process, a web site has been especially conceived for this teaching. You can find it here:

<http://www.immunology-and-modelling.ch>

1. This web site will indicate you **when to switch between lecture and practicals**. For instance, you will be able to make exercises 1 and 2 just after the prologue regarding molecular representation, while exercises 3 to 5 will be made after the lecture section dedicated to MHC-I and before that of MHC-II
2. **Videos on how to execute exercises 1 to 5** have been made for your help. There are without sound, but all instructions are detailed in the booklet
3. The **booklet of the practicals and the PDF of the lecture** can be downloaded from the web site too


Videos of exercises here

Links to download lecture and exercises here

5

## The dedicated web site – Seeing the lecture in 3D

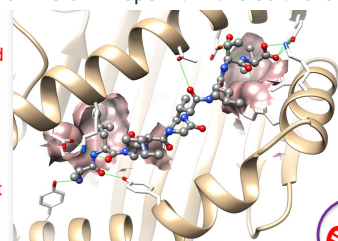
4. Once you will be trained in using UCSF Chimera, you will be able to use this tool to see in 3D, on your computer, the systems that are displayed in 2D in the lecture.

So, when you see this icon  in the lecture slide, you can check the slide number, and click on the corresponding link on the web site.

Upon clicking, the molecular system should be automatically displayed in UCSF Chimera, if your browser is managing it correctly.

Otherwise, you can download the linked file, with the .chimerax extension, on your machine. Then, double clicking on the file, or drag&dropping it on UCSF chimera will open it in this software.

Although you are encouraged to do it, to benefit from a full 3D "immersion", this is totally optional. This possibility is provided to you to go beyond 2D images in the slides. But if it does not work for you, or if you don't want to use it, it will not prevent you from understanding the lecture.



MAIT-1/Melan-A

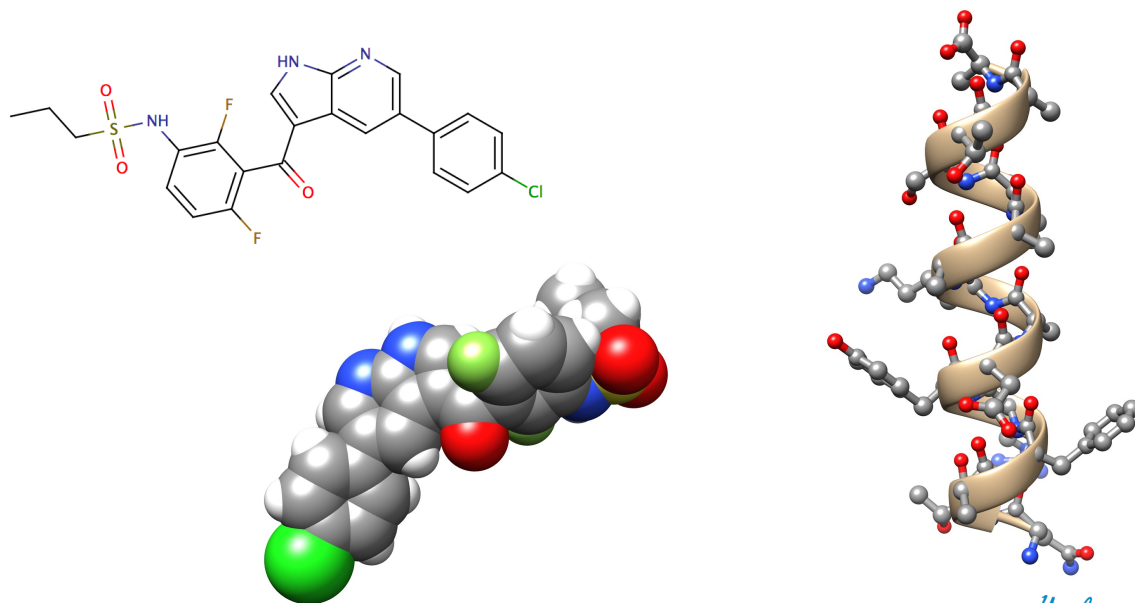
If you see the "DNA" icon, click on the link corresponding to the slide number

Videos of exercises here

Links to download lecture and exercises here

6

## Prologue: molecular representations



7

## Molecular representations – “small” molecules

Organic molecules of less than ~ 100 atoms are often referred to as “small” molecules, as opposed to biological macromolecules (i.e. proteins, DNA, etc.)

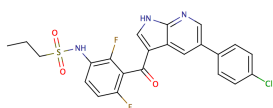
Small molecules can be represented in 1D, 2D or 3D:

Example of Vemurafenib (BRAF V600E inhibitor)

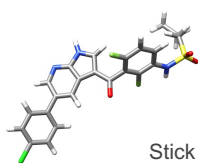
1D

**SMILES:** CCCC(=O)(=O)NC1=C(F)C(C(=O)C2=CNC3=NC=C(C=C23)C2=CC=C(Cl)C=C2)=C(F)C=C1  
**InChI:** 1S/C23H18ClF2N3O3S/c1-2-9-33(31,32)29-19-8-7-18(25)20(21(19)26)22(30)17-12-28-23-16(17)10-14(11-27-23)13-3-5-15(24)6-4-13/h3-8,10-12,29H,2,9H2,1H3,(H,27,28)

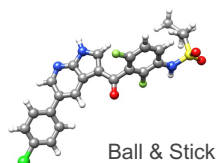
2D



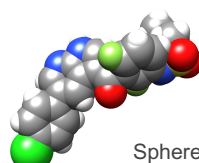
3D



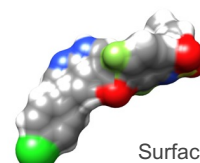
Stick



Ball & Stick



Sphere



Surface

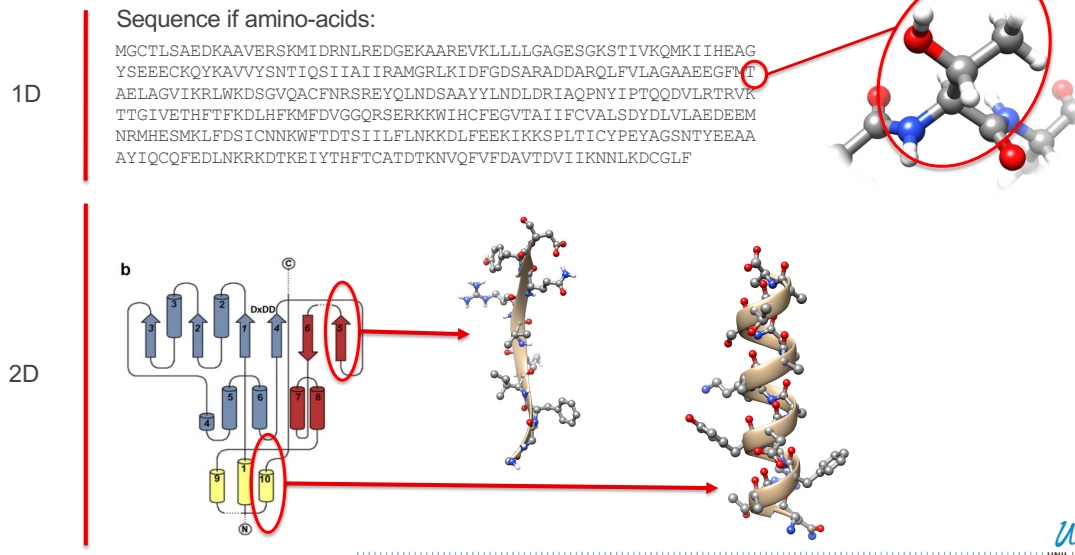
8

8

## Molecular representations – biological macromolecules

Biological macromolecules can also be represented in 1D, 2D or 3D:

Example of proteins



9

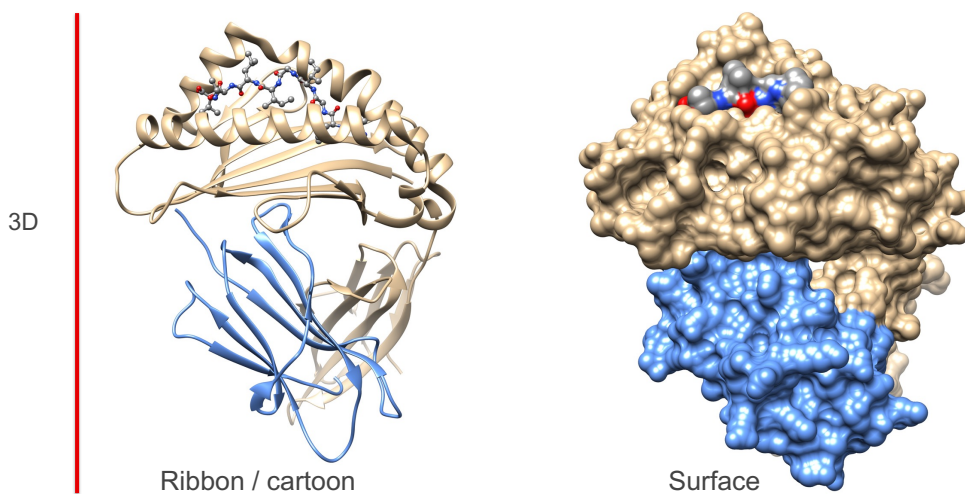
Unil  
UNIL | Université de Lausanne

9

## Molecular representations – biological macromolecules

Biological macromolecules can also be represented in 1D, 2D or 3D:

Example of proteins (Crystal structure of HLA-A2\*0201 in complex with MART-1/Melan-A)



Unil  
UNIL | Université de Lausanne

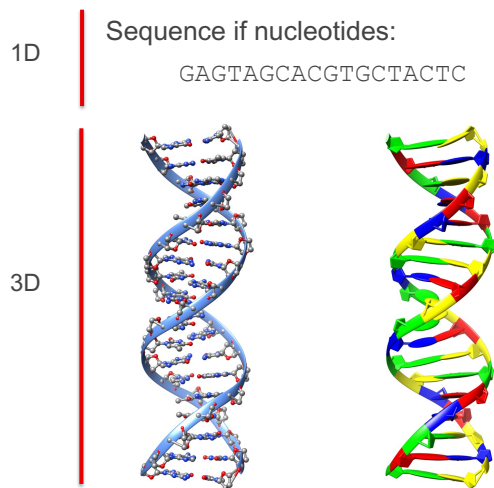
10

10

## Molecular representations – biological macromolecules

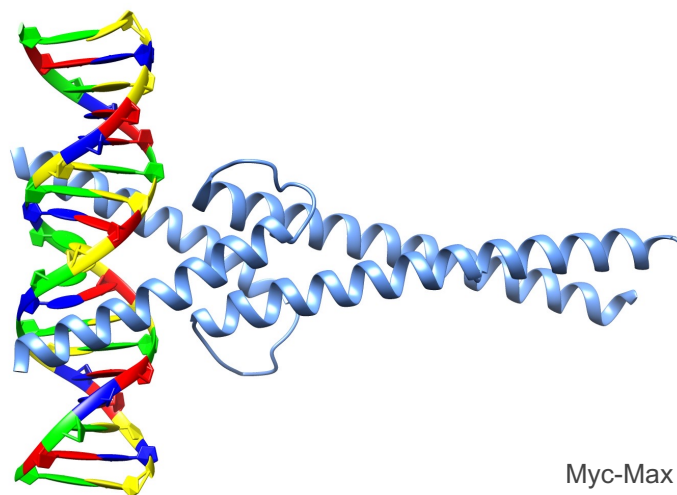
Biological macromolecules can also be represented in 1D, 2D or 3D:

Example of DNA



11

## Molecular representations – biological macromolecules



Myc-Max transcription factor

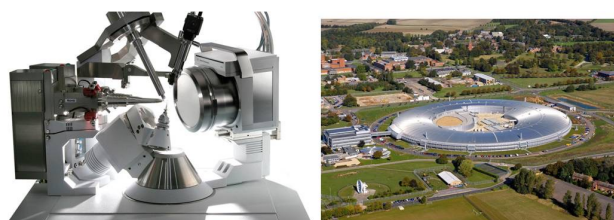
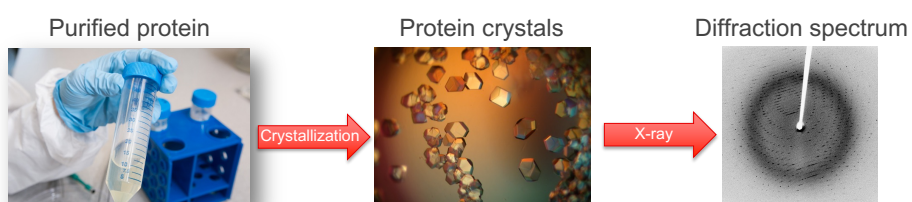
12

## Origin of the 3D structures



13

## Experimental methods – Xray crystallography

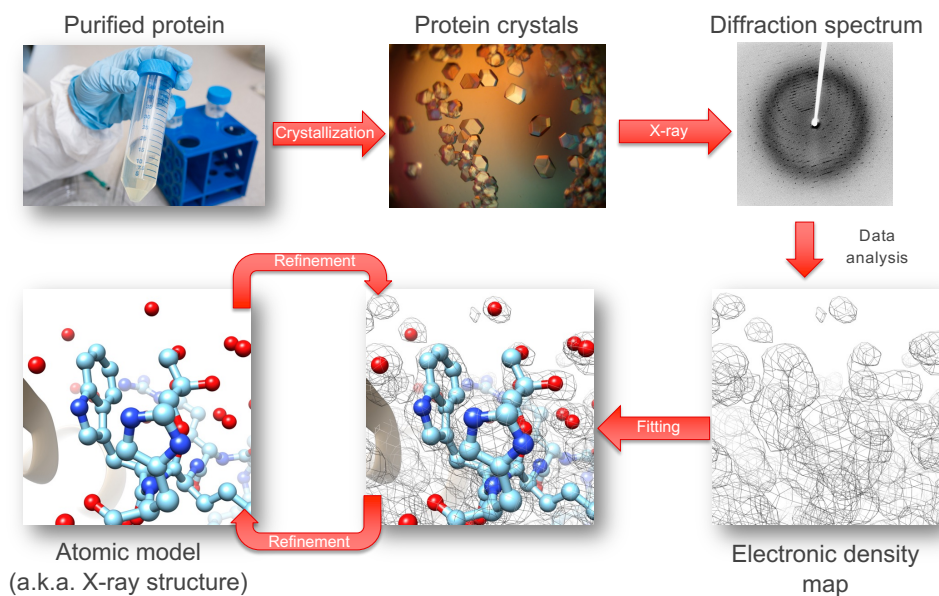


Xray diffraction

14



## Experimental methods – X-ray crystallography

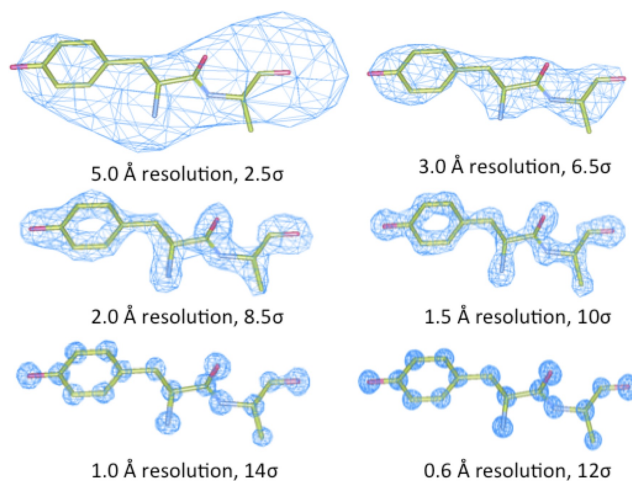


15

## Experimental methods – X-ray crystallography

### Important measures of accuracy:

- **Resolution** (in Å): measures the amount of detail that may be seen in the experimental data. The lower the better (typically around 2 Å)



Source: PLoS One. 2015 Apr 20;10(4):e0123146.

16

## Experimental methods – Xray crystallography

### 3 important measures of accuracy:

- **Resolution** (in Å): measures the amount of detail that may be seen in the experimental data. The lower the better (typically around 2 Å)
- **R-value**: measures how well the atomic model is supported by the experimental data found in the structure factor file (Perfect fit R-value = 0.0; Random fit R-value = 0.63; Typical R-value ~ 0.20) The atomic model is used to simulate a diffraction spectrum, which is compared to the experimental one.
- **R-free value**: idem than R-value, but calculated for a set of experimental data that have not been used to create the model (~10% of the data are removed before refinement, in order to be used in this test). Generally, R-free value > R-value; Typically R-free value ~ 0.26 for a good quality structure.

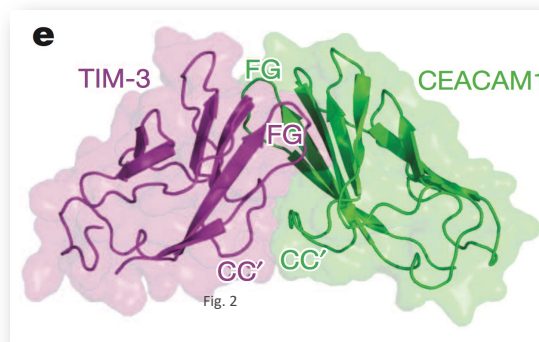
### Typical limitations:

- Hydrogen atoms are generally not visible
- Some regions are not defined (e.g. flexible loops or flexible side chains)
- X-ray structures are models. They can be totally wrong!!

## Experimental methods – Xray crystallography

X-ray structures are models. They can be totally wrong!!

Huang, Y.-H., et al. CEACAM1 regulates TIM-3-mediated tolerance and exhaustion. *Nature*, 2015, 517(7534), 386–390.



Xray structure of the complex  
CEACAM1/TIM3  
PDB ID: 4QYC  
Resolution: 3.4Å  
R-value: 0.232

Correction →

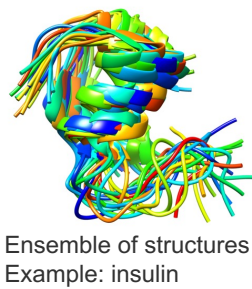
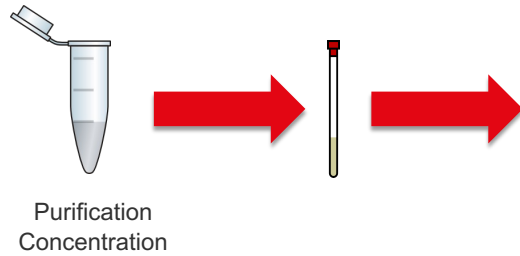
**5DZL**

Crystal structure of the protein human CEACAM1

DOI: 10.2210/pdb5dzl/pdb Entry 5DZL supersedes 4QYC

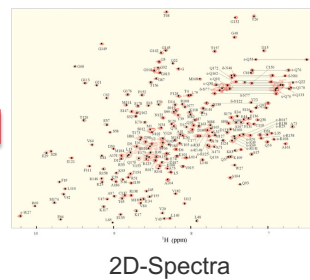
It was a homodimer of CEACAM1...!

## Experimental methods – NMR spectroscopy



Distance  
constraints

Modeling



Pros : Structure in solution

- Cons :
- Limited to small proteins
  - Low resolution
  - Highly flexible regions

Unil

UNIL | Université de Lausanne

19

19

## Experimental methods – CryoEM

### DUBOCHET'S VITRIFICATION METHOD

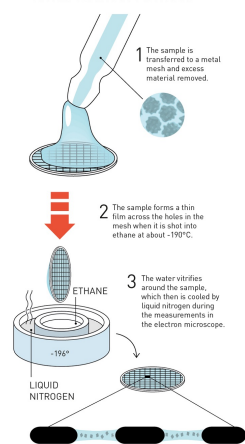
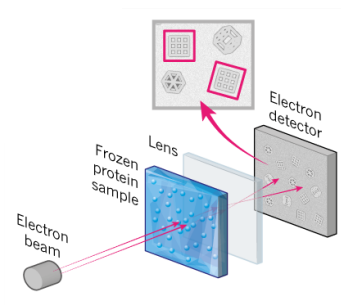
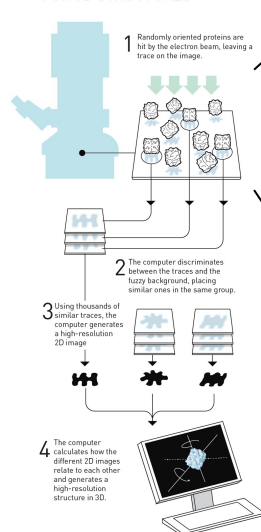


Illustration: © Johan Jarnestad/The Royal Swedish Academy of Sciences

### FRANK'S IMAGE ANALYSIS FOR 3D STRUCTURES



- Very power electronic beam
- Better resolution than light (smaller wave length)
- In vacuo in the microscope
- Frozen sample (77 K or 4 K)
- Vitrified water

Unil

UNIL | Université de Lausanne

20

20

## Experimental methods – CryoEM

### DUBOCHET'S VITRIFICATION METHOD

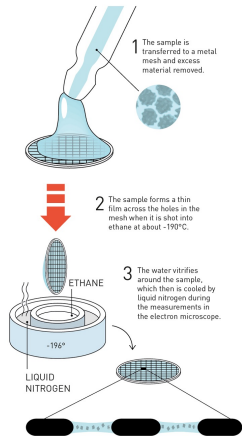
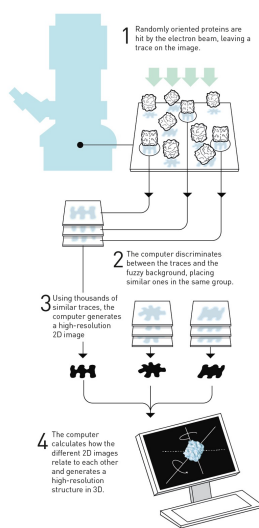


Illustration: © Johan Jarnestedt/The Royal Swedish Academy of Sciences

### FRANK'S IMAGE ANALYSIS FOR 3D STRUCTURES



Until recently:

- Only low resolution structures. Need to be used together with Xray crystallography or NMR (for example, insertion of Xray structures into the Cryo-EM density map)
- Limited to large-size systems (which can actually be seen as a pros or a cons)

Nowadays:

- Resolution close to that of Xray crystallography
- Applicable to smaller systems
- More Cryo-EM structures produced every year than NMR structures
- Capture structures in relevant states (isolated molecules, in solution, at a given salt concentration and pH)

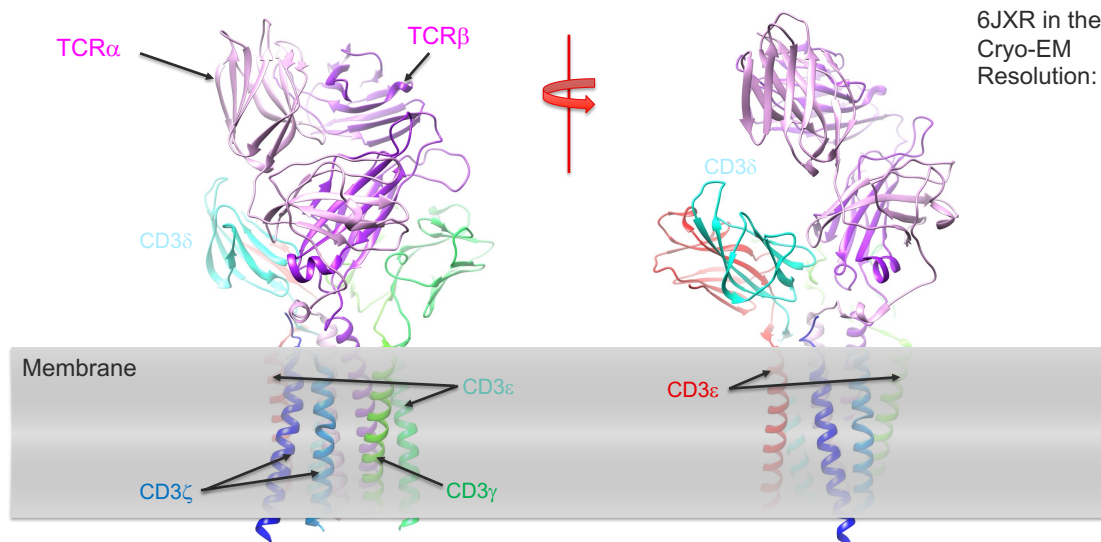
Unil

UNIL | Université de Lausanne

21

21

## Experimental methods – CryoEM



Structural basis of assembly of the human T cell receptor-CD3 complex.  
Zheng, L., Lin, J., Zhang, B., Zhu, Y., Li, N., Xie, S., Wang, Y., Gao, N., Huang, Z.  
(2019) Nature 573: 546-552

Unil

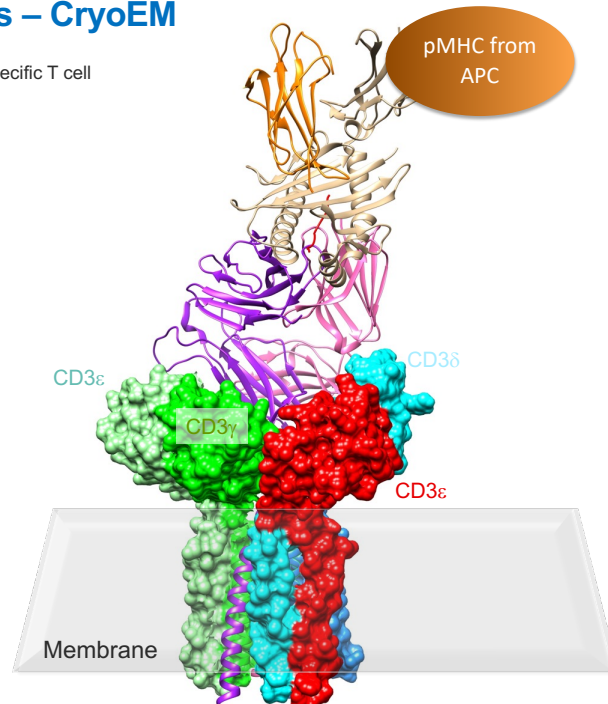
UNIL | Université de Lausanne

22

22

## Experimental methods – CryoEM

Structure of a fully assembled tumor-specific T cell receptor ligated by pMHC  
 Lukas Sušac et. al.  
 Cell, 2022, 185 (17), 3201-3213.e19



7PHR in the PDB  
 Cryo-EM  
 Resolution: 3.1 Å

## Experimental methods – CryoEM

### The Nobel Prize in Chemistry 2017



Photo: Félix Imhof © UNIL [CC BY-SA 4.0]  
**Jacques Dubochet**  
 Prize share: 1/3



Photo: B. Winkowski © Columbia University Medical Center  
**Joachim Frank**  
 Prize share: 1/3



Photo: MRC Laboratory of Molecular Biology  
**Richard Henderson**  
 Prize share: 1/3

The Nobel Prize in Chemistry 2017 was awarded to Jacques Dubochet, Joachim Frank and Richard Henderson "for developing cryo-electron microscopy for the high-resolution structure determination of biomolecules in solution".



## Experimental methods - Summary

| Technique            | Advantages  | Disadvantages  |
|----------------------|---|--|
| Xray crystallography | High resolution (1 to 3 Å )   | Requires to crystallize the protein<br>Does not allow studying transmembrane or very flexible proteins         |
| NMR                  | Does not require protein crystallization<br>~ High resolution   | Generally limited to small proteins  |
| Cryo-EM              | Does not necessitate to crystallize the protein: possible to study transmembrane proteins, and more flexible proteins than Xray.<br><br>New techniques allow studying smaller proteins, and increasing resolution | Generally limited to large proteins<br>Low resolution, 4 to 20 Å (a lot of progresses have been done recently) |

25

## Where to find experimental 3D structures? The protein databank

Experimental 3D structures are stored in the **Protein Data Bank (PDB)**

Worldwide Protein Data Bank (wwPDB)

<https://www.wwpdb.org/>

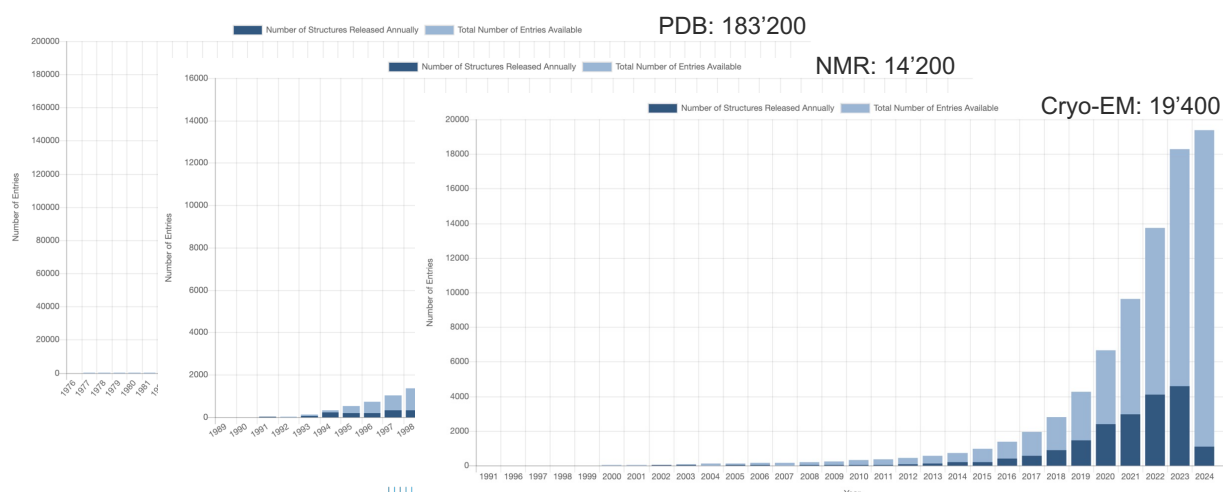
Protein Data Bank in US (PDB)

<http://www.rcsb.org/>

Protein Data Bank in Europe (PDBe)

<http://www.ebi.ac.uk/pdbe/>

217'200 structures  
in March 2024



26

## Where to find experimental 3D structures? The protein databank

<http://www.rcsb.org/>

MART-1 HLA-A2

Unil  
UNIL | Université de Lausanne

27

27

## Where to find experimental 3D structures? The protein databank

Possible to sort

25 structures  
out of 235

PDB ID

Authors

Experimental  
methods

Unil  
UNIL | Université de Lausanne

28

28

## Where to find experimental 3D structures? The protein databank

**1JF1**  
Crystal structure of HLA-A2\*0201 in complex with a decameric altered peptide ligand from the MART-1/Melan-A  
Sliz, P., Michelin, O., Cerottini, J.C., Luescher, I., Romero, P., Karplus, M., Wiley, D.C.  
(2001) J Immunol 167: 3276-3284  
Released: 2001-09-14  
Method: X-RAY DIFFRACTION 1.85 Å  
Organisms: Homo sapiens  
Macromolecule: HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, A-2 ALPHA CHAIN (protein); beta-2-microglobulin (protein); decameric peptide ligand from the MART-1/Melan-A (protein)

**1JHT**  
Crystal structure of HLA-A2\*0201 in complex with a nonameric altered peptide ligand (ALGIGLTV) from the MART-1/Melan-A.  
Sliz, P., Michelin, O., Karplus, M., Romero, P., Wiley, D.  
(2001) J Immunol 167: 3276-3284  
Released: 2001-09-14  
Method: X-RAY DIFFRACTION 2.15 Å  
Organisms: Homo sapiens  
Macromolecule: HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, A-2 ALPHA CHAIN (protein); beta-2-microglobulin (protein); nonameric peptide ligand from the MART-1/Melan-A (protein)

**5NHT**  
human 199.54-16 TCR in complex with Melan-A/MART-1 (26-35) peptide and HLA-A2  
Exertier, C., Reiser, J.-B., Lantier, V., Chouquet, A., Bonneville, M., Sauquín, X., Housset, D.  
To be published  
Released: 2018-05-16  
Method: X-RAY DIFFRACTION 3.2 Å  
Organisms: Homo sapiens  
Macromolecule: Beta-2-microglobulin (protein); HLA class I histocompatibility antigen, A-2 alpha chain (protein); Melanoma antigen recognized by T-cells 1 (protein); T-cell receptor alpha variable 12-2, T-cell receptor, sp3.4 alpha chain (protein); T-cell receptor beta variable 19, TRB protein (protein)

**5NQK**  
human 199.16 TCR in complex with Melan-A/MART-1 (26-35) peptide and HLA-

Unil  
UNIL | Université de Lausanne 29

29

## Where to find experimental 3D structures? The protein databank

**1JHT**  
Crystal structure of HLA-A2\*0201 in complex with a nonameric altered peptide ligand (ALGIGLTV) from the MART-1/Melan-A.  
PDB DOI: <https://doi.org/10.2210/pdb.1JHT/pdb>  
Classification: IMMUNE SYSTEM  
Organism(s): Homo sapiens  
Expression System: Escherichia coli  
Mutation(s): No  
Deposited: 2001-06-28 Released: 2001-09-14  
Deposition Author(s): Sliz, P., Michelin, O., Karplus, M., Romero, P., Wiley, D.  
Experimental Data Snapshot  
Method: X-RAY DIFFRACTION  
Resolution: 2.15 Å  
R-Value Free: 0.252  
R-Value Work: 0.210  
R-Value Observed: 0.210  
wwPDB Validation  
Metric: Percentile Ranks  
R-Value: 0.250  
R-Value Free: 0.252  
R-Value Work: 0.210  
R-Value Observed: 0.210  
This is version 1.3 of the entry. See complete history.  
Literature  
Download Primary Citation  
Crystal structures of two closely related but antigenically distinct HLA-A2/melanocyte-melanoma tumor-antigen peptide complexes.  
Sliz, P., Michelin, O., Cerottini, J.C., Luescher, I., Romero, P., Karplus, M., Wiley, D.C.  
(2001) J Immunol 167: 3276-3284  
PubMed: 11544315  
DOI: <https://doi.org/10.4049/jimmunol.167.6.3276>  
Primary Citation of Related Structures:  
1JF1, 1JHT  
PubMed Abstract:  
We have determined high-resolution crystal structures of the complexes of HLA-A2 molecules with two modified immunodominant peptides from the melanoma tumor-associated protein Melan-A/melanoma Ag recognized by T cells-1. The two peptides, a decamer and nonamer with overlapping sequences (ELAGIGLTV and ALGIGLTV), a...  
Organizational Affiliation:  
Department of Molecular and Cellular Biology and Howard Hughes Medical Institute, Harvard University.

Unil  
UNIL | Université de Lausanne 30

30

## Where to find experimental 3D structures? The protein databank

Online visualization

**Click here!!**

1JHT  
Crystal structure of HLA-A2\*0201 in complex with a nonameric altered peptide ligand (ALGILTLV) from the MART-1/Melan-A.

Sequence of 1JHT | Cryst... | Chain | 1: HLA CL... | A |

Structure  
Type Assembly  
Asm Id 1: Author And Softwar...  
Dynamic Bonds X Off  
Nothing Focused  
Measurements  
Structure Motif Search  
Components 1JHT  
Preset + Add  
Polymer Cartoon  
Water Ball & Stick  
Unit Cell P 21 21 21  
Density  
Quality Assessment  
Assembly Symmetry  
Export Models  
Export Animation  
Export Geometry

31

## Where to find experimental 3D structures? The protein databank

Information regarding the protein, and what is present in the experimental structure

Macromolecules

Find similar proteins by: Sequence (by identity cutoff) | 3D Structure

Entity ID: 1

| Molecule   | Chains | Sequence Length | Organism     | Details        | Image |
|--|--------|-----------------|--------------|----------------|-------|
| HLA CLASS I HISTOCOMPATIBILITY ANTIGEN A*E ALPHA CHAIN | A      | 275             | Homo sapiens | Mutation(s): 0 |       |

UniProt & NIH Common Fund Data Resources

Find proteins for P04439 (Homo sapiens)

PHAROS: P04439 GTEX: ENSG00000206503 Explore P04439 Go to UniProtKB: P04439

Entity Groups

Sequence Clusters 30% Identity 50% Identity 70% Identity 80% Identity 90% Identity 100% Identity

UniProt Group P04439

Sequence Annotations Expand

Reference Sequence 1JHT\_1

1JHT\_1  
UNIPROT P04439  
UNMODELED A  
HYDROPATHY  
DISORDER  
DISORDERED BINDING  
PFAM  
ANTIBODY DOMAIN

ID of the protein chain

Number of residues in the protein chain

Source organism

Link to this protein in Uniprot

Information on sequence, mutations, missing regions

32

## Where to find experimental 3D structures? The protein databank

Download / display

Click here!

### 1JHT

Crystal structure of HLA-A2\*0201 in complex with a peptide ligand (ALGIGILT)V

PDB DOI: 10.2222

Classification: M

Organism(s): Homo

Expression System

Mutation(s): No

Deposited: 2001

Deposition Author

Experimental Data

Method: X-RAY

Resolution: 2.15

R-Value Free: 0.2

R-Value Work: 0.2

R-Value Observed

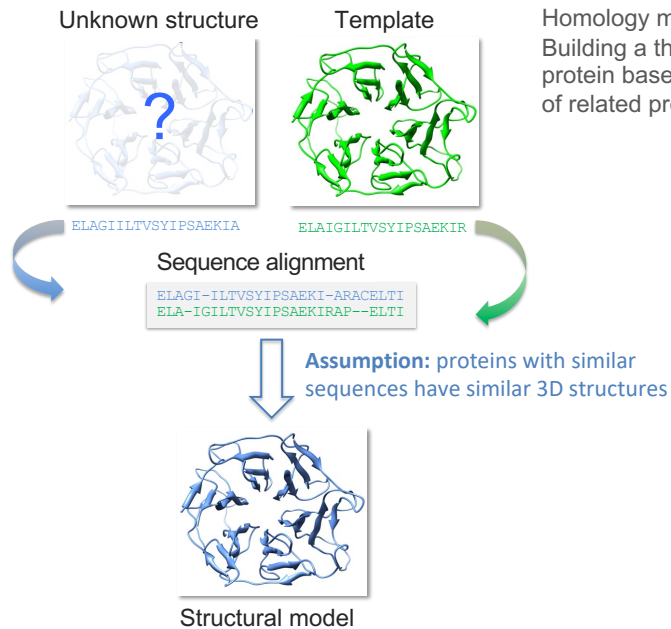
A header with information on the protein and experimental conditions

```

HEADER
    CRYSTAL STRUCTURE OF HLA-A2*0201 IN COMPLEX WITH A
    PEPTIDE LIGAND (ALGIGILT)V FROM THE HLA-
    CLASS II ANTIGEN PRESENTING CELL
    TITLE
    1JHT
    UNRESOLVED
    COMPND
    1 ALGIGILT
    2 HLA-A2*0201
    3 HLA-B*0801
    4 HLA-C*0201
    5 HLA-D*0201
    6 HLA-E*0201
    7 HLA-F*0201
    8 HLA-G*0201
    9 HLA-I*0201
    10 HLA-J*0201
    11 HLA-K*0201
    12 HLA-L*0201
    13 HLA-M*0201
    14 HLA-N*0201
    15 HLA-O*0201
    16 HLA-P*0201
    17 HLA-Q*0201
    18 HLA-R*0201
    19 HLA-S*0201
    20 HLA-T*0201
    21 HLA-U*0201
    22 HLA-V*0201
    23 HLA-W*0201
    24 HLA-X*0201
    25 HLA-Y*0201
    26 HLA-Z*0201
    27 HLA-AA*0201
    28 HLA-AB*0201
    29 HLA-AC*0201
    30 HLA-AD*0201
    31 HLA-AE*0201
    32 HLA-AF*0201
    33 HLA-AG*0201
    34 HLA-AH*0201
    35 HLA-AI*0201
    36 HLA-AJ*0201
    37 HLA-AK*0201
    38 HLA-AL*0201
    39 HLA-AM*0201
    40 HLA-AN*0201
    41 HLA-AO*0201
    42 HLA-AP*0201
    43 HLA-AQ*0201
    44 HLA-AR*0201
    45 HLA-AS*0201
    46 HLA-AT*0201
    47 HLA-AU*0201
    48 HLA-AV*0201
    49 HLA-AW*0201
    50 HLA-AX*0201
    51 HLA-AY*0201
    52 HLA-AZ*0201
    53 HLA-BA*0201
    54 HLA-BB*0201
    55 HLA-BC*0201
    56 HLA-BD*0201
    57 HLA-BE*0201
    58 HLA-BF*0201
    59 HLA-BG*0201
    60 HLA-BH*0201
    61 HLA-BI*0201
    62 HLA-BJ*0201
    63 HLA-BK*0201
    64 HLA-BL*0201
    65 HLA-BM*0201
    66 HLA-BN*0201
    67 HLA-BO*0201
    68 HLA-BP*0201
    69 HLA-BQ*0201
    70 HLA-BR*0201
    71 HLA-BS*0201
    72 HLA-BT*0201
    73 HLA-BU*0201
    74 HLA-BV*0201
    75 HLA-BW*0201
    76 HLA-BX*0201
    77 HLA-BY*0201
    78 HLA-BZ*0201
    79 HLA-CA*0201
    80 HLA-CB*0201
    81 HLA-CC*0201
    82 HLA-CD*0201
    83 HLA-CE*0201
    84 HLA-CF*0201
    85 HLA-CG*0201
    86 HLA-CH*0201
    87 HLA-CI*0201
    88 HLA-CJ*0201
    89 HLA-CK*0201
    90 HLA-CL*0201
    91 HLA-CM*0201
    92 HLA-CN*0201
    93 HLA-CO*0201
    94 HLA-CP*0201
    95 HLA-CQ*0201
    96 HLA-CR*0201
    97 HLA-CS*0201
    98 HLA-CT*0201
    99 HLA-CU*0201
    100 HLA-CV*0201
    101 HLA-CW*0201
    102 HLA-CX*0201
    103 HLA-CY*0201
    104 HLA-CZ*0201
    105 HLA-DA*0201
    106 HLA-DB*0201
    107 HLA-DC*0201
    108 HLA-DD*0201
    109 HLA-DE*0201
    110 HLA-DF*0201
    111 HLA-DG*0201
    112 HLA-DH*0201
    113 HLA-DI*0201
    114 HLA-DJ*0201
    115 HLA-DK*0201
    116 HLA-DL*0201
    117 HLA-DM*0201
    118 HLA-DN*0201
    119 HLA-DO*0201
    120 HLA-DP*0201
    121 HLA-DQ*0201
    122 HLA-DR*0201
    123 HLA-DS*0201
    124 HLA-DT*0201
    125 HLA-DU*0201
    126 HLA-DV*0201
    127 HLA-DW*0201
    128 HLA-DX*0201
    129 HLA-DY*0201
    130 HLA-DZ*0201
    131 HLA-EA*0201
    132 HLA-EB*0201
    133 HLA-EC*0201
    134 HLA-ED*0201
    135 HLA-EE*0201
    136 HLA-EF*0201
    137 HLA-EG*0201
    138 HLA-EH*0201
    139 HLA-EI*0201
    140 HLA-EJ*0201
    141 HLA-EK*0201
    142 HLA-EL*0201
    143 HLA-EM*0201
    144 HLA-EN*0201
    145 HLA-EO*0201
    146 HLA-EP*0201
    147 HLA-EQ*0201
    148 HLA-ER*0201
    149 HLA-ES*0201
    150 HLA-ET*0201
    151 HLA-EU*0201
    152 HLA-EV*0201
    153 HLA-EW*0201
    154 HLA-EX*0201
    155 HLA-EY*0201
    156 HLA-EZ*0201
    157 HLA-FA*0201
    158 HLA-FB*0201
    159 HLA-FC*0201
    160 HLA-FD*0201
    161 HLA-FE*0201
    162 HLA-FF*0201
    163 HLA-FG*0201
    164 HLA-FH*0201
    165 HLA-FI*0201
    166 HLA-FJ*0201
    167 HLA-FK*0201
    168 HLA-FL*0201
    169 HLA-FM*0201
    170 HLA-FN*0201
    171 HLA-FO*0201
    172 HLA-FP*0201
    173 HLA-FQ*0201
    174 HLA-FR*0201
    175 HLA-FS*0201
    176 HLA-FT*0201
    177 HLA-FU*0201
    178 HLA-FV*0201
    179 HLA-FW*0201
    180 HLA-FX*0201
    181 HLA-FY*0201
    182 HLA-FZ*0201
    183 HLA-GA*0201
    184 HLA-GB*0201
    185 HLA-GC*0201
    186 HLA-GD*0201
    187 HLA-GE*0201
    188 HLA-GF*0201
    189 HLA-GG*0201
    190 HLA-GH*0201
    191 HLA-GI*0201
    192 HLA-GJ*0201
    193 HLA-GK*0201
    194 HLA-GL*0201
    195 HLA-GM*0201
    196 HLA-GN*0201
    197 HLA-GO*0201
    198 HLA-GP*0201
    199 HLA-GQ*0201
    200 HLA-GR*0201
    201 HLA-GS*0201
    202 HLA-GT*0201
    203 HLA-GU*0201
    204 HLA-GV*0201
    205 HLA-GW*0201
    206 HLA-GX*0201
    207 HLA-GY*0201
    208 HLA-GZ*0201
    209 HLA-HA*0201
    210 HLA-HB*0201
    211 HLA-HC*0201
    212 HLA-HD*0201
    213 HLA-HE*0201
    214 HLA-HF*0201
    215 HLA-HG*0201
    216 HLA-HH*0201
    217 HLA-HI*0201
    218 HLA-HJ*0201
    219 HLA-HK*0201
    220 HLA-HL*0201
    221 HLA-HM*0201
    222 HLA-HN*0201
    223 HLA-HO*0201
    224 HLA-HP*0201
    225 HLA-HQ*0201
    226 HLA-HR*0201
    227 HLA-HS*0201
    228 HLA-HT*0201
    229 HLA-HU*0201
    230 HLA-HV*0201
    231 HLA-HW*0201
    232 HLA-HX*0201
    233 HLA-HY*0201
    234 HLA-HZ*0201
    235 HLA-IA*0201
    236 HLA-IB*0201
    237 HLA-IC*0201
    238 HLA-ID*0201
    239 HLA-IE*0201
    240 HLA-IF*0201
    241 HLA-IG*0201
    242 HLA-IH*0201
    243 HLA-II*0201
    244 HLA-IL*0201
    245 HLA-IM*0201
    246 HLA-IN*0201
    247 HLA-IO*0201
    248 HLA-IP*0201
    249 HLA-IQ*0201
    250 HLA-IR*0201
    251 HLA-IS*0201
    252 HLA-IT*0201
    253 HLA-IU*0201
    254 HLA-IV*0201
    255 HLA-IW*0201
    256 HLA-IX*0201
    257 HLA-IY*0201
    258 HLA-IZ*0201
    259 HLA-JA*0201
    260 HLA-JB*0201
    261 HLA-JC*0201
    262 HLA-JD*0201
    263 HLA-JE*0201
    264 HLA-JF*0201
    265 HLA-JG*0201
    266 HLA-JH*0201
    267 HLA-JI*0201
    268 HLA-JJ*0201
    269 HLA-JK*0201
    270 HLA-JL*0201
    271 HLA-JM*0201
    272 HLA-JN*0201
    273 HLA-JO*0201
    274 HLA-JP*0201
    275 HLA-JQ*0201
    276 HLA-JR*0201
    277 HLA-JS*0201
    278 HLA-JT*0201
    279 HLA-JU*0201
    280 HLA-JV*0201
    281 HLA-JW*0201
    282 HLA-JX*0201
    283 HLA-JY*0201
    284 HLA-JZ*0201
    285 HLA-KA*0201
    286 HLA-KB*0201
    287 HLA-KC*0201
    288 HLA-KD*0201
    289 HLA-KE*0201
    290 HLA-KF*0201
    291 HLA-KG*0201
    292 HLA-KH*0201
    293 HLA-KI*0201
    294 HLA-KJ*0201
    295 HLA-KK*0201
    296 HLA-KL*0201
    297 HLA-KM*0201
    298 HLA-KN*0201
    299 HLA-KO*0201
    300 HLA-KP*0201
    301 HLA-KQ*0201
    302 HLA-KR*0201
    303 HLA-KS*0201
    304 HLA-KT*0201
    305 HLA-KU*0201
    306 HLA-KV*0201
    307 HLA-KW*0201
    308 HLA-KX*0201
    309 HLA-KY*0201
    310 HLA-KZ*0201
    311 HLA-LA*0201
    312 HLA-LB*0201
    313 HLA-LC*0201
    314 HLA-LD*0201
    315 HLA-LE*0201
    316 HLA-LF*0201
    317 HLA-LG*0201
    318 HLA-LH*0201
    319 HLA-LI*0201
    320 HLA-LJ*0201
    321 HLA-LK*0201
    322 HLA-LL*0201
    323 HLA-LM*0201
    324 HLA-LN*0201
    325 HLA-LO*0201
    326 HLA-LP*0201
    327 HLA-LQ*0201
    328 HLA-LR*0201
    329 HLA-LS*0201
    330 HLA-LT*0201
    331 HLA-LU*0201
    332 HLA-LV*0201
    333 HLA-LW*0201
    334 HLA-LX*0201
    335 HLA-LY*0201
    336 HLA-LZ*0201
    337 HLA-MA*0201
    338 HLA-MB*0201
    339 HLA-MC*0201
    340 HLA-MD*0201
    341 HLA-ME*0201
    342 HLA-MF*0201
    343 HLA-MG*0201
    344 HLA-MH*0201
    345 HLA-MI*0201
    346 HLA-MJ*0201
    347 HLA-MK*0201
    348 HLA-ML*0201
    349 HLA-MN*0201
    350 HLA-MO*0201
    351 HLA-MP*0201
    352 HLA-MQ*0201
    353 HLA-MR*0201
    354 HLA-MS*0201
    355 HLA-MT*0201
    356 HLA-MU*0201
    357 HLA-MV*0201
    358 HLA-MW*0201
    359 HLA-MX*0201
    360 HLA-MY*0201
    361 HLA-MZ*0201
    362 HLA-NA*0201
    363 HLA-NB*0201
    364 HLA-NC*0201
    365 HLA-ND*0201
    366 HLA-NE*0201
    367 HLA-NF*0201
    368 HLA-NG*0201
    369 HLA-NH*0201
    370 HLA-NI*0201
    371 HLA-NJ*0201
    372 HLA-NK*0201
    373 HLA-NL*0201
    374 HLA-NM*0201
    375 HLA-NO*0201
    376 HLA-NP*0201
    377 HLA-NQ*0201
    378 HLA-NR*0201
    379 HLA-NS*0201
    380 HLA-NT*0201
    381 HLA-NU*0201
    382 HLA-NV*0201
    383 HLA-NW*0201
    384 HLA-NX*0201
    385 HLA-NY*0201
    386 HLA-NZ*0201
    387 HLA-OA*0201
    388 HLA-OB*0201
    389 HLA-OC*0201
    390 HLA-OD*0201
    391 HLA-OE*0201
    392 HLA-OF*0201
    393 HLA-OG*0201
    394 HLA-OH*0201
    395 HLA-OI*0201
    396 HLA-OJ*0201
    397 HLA-OK*0201
    398 HLA-OL*0201
    399 HLA-OM*0201
    400 HLA-ON*0201
    401 HLA-OO*0201
    402 HLA-OP*0201
    403 HLA-OQ*0201
    404 HLA-OR*0201
    405 HLA-OS*0201
    406 HLA-OT*0201
    407 HLA-OU*0201
    408 HLA-OV*0201
    409 HLA-OW*0201
    410 HLA-OX*0201
    411 HLA-OY*0201
    412 HLA-OZ*0201
    413 HLA-PA*0201
    414 HLA-PB*0201
    415 HLA-PC*0201
    416 HLA-PD*0201
    417 HLA-PE*0201
    418 HLA-PF*0201
    419 HLA-PG*0201
    420 HLA-PH*0201
    421 HLA-PI*0201
    422 HLA-PJ*0201
    423 HLA-PK*0201
    424 HLA-PL*0201
    425 HLA-PM*0201
    426 HLA-PN*0201
    427 HLA-PO*0201
    428 HLA-PP*0201
    429 HLA-PQ*0201
    430 HLA-PR*0201
    431 HLA-PS*0201
    432 HLA-PT*0201
    433 HLA-PU*0201
    434 HLA-PV*0201
    435 HLA-PW*0201
    436 HLA-PX*0201
    437 HLA-PY*0201
    438 HLA-PZ*0201
    439 HLA-QA*0201
    440 HLA-QB*0201
    441 HLA-QC*0201
    442 HLA-QD*0201
    443 HLA-QE*0201
    444 HLA-QF*0201
    445 HLA-QG*0201
    446 HLA-QH*0201
    447 HLA-QI*0201
    448 HLA-QJ*0201
    449 HLA-QK*0201
    450 HLA-QL*0201
    451 HLA-QM*0201
    452 HLA-QN*0201
    453 HLA-QO*0201
    454 HLA-QP*0201
    455 HLA-QQ*0201
    456 HLA-QR*0201
    457 HLA-QS*0201
    458 HLA-QT*0201
    459 HLA-QU*0201
    460 HLA-QV*0201
    461 HLA-QW*0201
    462 HLA-QX*0201
    463 HLA-QY*0201
    464 HLA-QZ*0201
    465 HLA-RA*0201
    466 HLA-RB*0201
    467 HLA-RC*0201
    468 HLA-RD*0201
    469 HLA-RE*0201
    470 HLA-RF*0201
    471 HLA-RG*0201
    472 HLA-RH*0201
    473 HLA-RI*0201
    474 HLA-RJ*0201
    475 HLA-RK*0201
    476 HLA-RL*0201
    477 HLA-RM*0201
    478 HLA-RN*0201
    479 HLA-RO*0201
    480 HLA-RP*0201
    481 HLA-RQ*0201
    482 HLA-RR*0201
    483 HLA-RS*0201
    484 HLA-RT*0201
    485 HLA-RU*0201
    486 HLA-RV*0201
    487 HLA-RW*0201
    488 HLA-RX*0201
    489 HLA-RY*0201
    490 HLA-RZ*0201
    491 HLA-SA*0201
    492 HLA-SB*0201
    493 HLA-SC*0201
    494 HLA-SD*0201
    495 HLA-SE*0201
    496 HLA-SF*0201
    497 HLA-SG*0201
    498 HLA-SH*0201
    499 HLA-SI*0201
    500 HLA-SJ*0201
    501 HLA-SK*0201
    502 HLA-SL*0201
    503 HLA-SM*0201
    504 HLA-SN*0201
    505 HLA-SO*0201
    506 HLA-SP*0201
    507 HLA-SQ*0201
    508 HLA-SR*0201
    509 HLA-SS*0201
    510 HLA-ST*0201
    511 HLA-SU*0201
    512 HLA-SV*0201
    513 HLA-SW*0201
    514 HLA-SX*0201
    515 HLA-SY*0201
    516 HLA-SZ*0201
    517 HLA-TA*0201
    518 HLA-TB*0201
    519 HLA-TC*0201
    520 HLA-TD*0201
    521 HLA-TE*0201
    522 HLA-TF*0201
    523 HLA-TG*0201
    524 HLA-TH*0201
    525 HLA-TI*0201
    526 HLA-TJ*0201
    527 HLA-TK*0201
    528 HLA-TL*0201
    529 HLA-TM*0201
    530 HLA-TN*0201
    531 HLA-TO*0201
    532 HLA-TP*0201
    533 HLA-TQ*0201
    534 HLA-TR*0201
    535 HLA-TS*0201
    536 HLA-TT*0201
    537 HLA-TU*0201
    538 HLA-TV*0201
    539 HLA-TW*0201
    540 HLA-TX*0201
    541 HLA-TY*0201
    542 HLA-TZ*0201
    543 HLA-UA*0201
    544 HLA-UB*0201
    545 HLA-UC*0201
    546 HLA-UD*0201
    547 HLA-UE*0201
    548 HLA-UF*0201
    549 HLA-UG*0201
    550 HLA-UH*0201
    551 HLA-UI*0201
    552 HLA-UJ*0201
    553 HLA-UK*0201
    554 HLA-UL*0201
    555 HLA-UM*0201
    556 HLA-UN*0201
    557 HLA-UO*0201
    558 HLA-UP*0201
    559 HLA-UQ*0201
    560 HLA-UR*0201
    561 HLA-US*0201
    562 HLA-UT*0201
    563 HLA-UV*0201
    564 HLA-UW*0201
    565 HLA-UX*0201
    566 HLA-UY*0201
    567 HLA-UZ*0201
    568 HLA-VA*0201
    569 HLA-VB*0201
    570 HLA-VC*0201
    571 HLA-VD*0201
    572 HLA-VE*0201
    573 HLA-VF*0201
    574 HLA-VG*0201
    575 HLA-VH*0201
    576 HLA-VI*0201
    577 HLA-VJ*0201
    578 HLA-VK*0201
    579 HLA-VL*0201
    580 HLA-VM*0201
    581 HLA-VN*0201
    582 HLA-VO*0201
    583 HLA-VP*0201
    584 HLA-VQ*0201
    585 HLA-VR*0201
    586 HLA-VS*0201
    587 HLA-VT*0201
    588 HLA-VU*0201
    589 HLA-VV*0201
    590 HLA-VW*0201
    591 HLA-VX*0201
    592 HLA-VY*0201
    593 HLA-VZ*0201
    594 HLA-WA*0201
    595 HLA-WB*0201
    596 HLA-WC*0201
    597 HLA-WD*0201
    598 HLA-WE*0201
    599 HLA-WF*0201
    600 HLA-WG*0201
    601 HLA-WH*0201
    602 HLA-WI*0201
    603 HLA-WJ*0201
    604 HLA-WK*0201
    605 HLA-WL*0201
    606 HLA-WM*0201
    607 HLA-WN*0201
    608 HLA-WO*0201
    609 HLA-WP*0201
    610 HLA-WQ*0201
    611 HLA-WR*0201
    612 HLA-WS*0201
    613 HLA-WT*0201
    614 HLA-WU*0201
    615 HLA-WV*0201
    616 HLA-WX*0201
    617 HLA-WY*0201
    618 HLA-WZ*0201
    619 HLA-XA*0201
    620 HLA-XB*0201
    621 HLA-XC*0201
    622 HLA-XD*0201
    623 HLA-XE*0201
    624 HLA-XF*0201
    625 HLA-XG*0201
    626 HLA-XH*0201
    627 HLA-XI*0201
    628 HLA-XJ*0201
    629 HLA-XK*0201
    630 HLA-XL*0201
    631 HLA-XM*0201
    632 HLA-XN*0201
    633 HLA-XO*0201
    634 HLA-XP*0201
    635 HLA-XQ*0201
    636 HLA-XR*0201
    637 HLA-XS*0201
    638 HLA-XT*0201
    639 HLA-XU*0201
    640 HLA-XV*0201
    641 HLA-XW*0201
    642 HLA-XX*0201
    643 HLA-XY*0201
    644 HLA-XZ*0201
    645 HLA-YA*0201
    646 HLA-YB*0201
    647 HLA-YC*0201
    648 HLA-YD*0201
    649 HLA-YE*0201
    650 HLA-YF*0201
    651 HLA-YG*0201
    652 HLA-YH*0201
    653 HLA-YI*0201
    654 HLA-YJ*0201
    655 HLA-YK*0201
    656 HLA-YL*0201
    657 HLA-YM*0201
    658 HLA-YN*0201
    659 HLA-YO*0201
    660 HLA-YP*0201
    661 HLA-YQ*0201
    662 HLA-YR*0201
    663 HLA-YS*0201
    664 HLA-YT*0201
    665 HLA-YU*0201
    666 HLA-YV*0201
    667 HLA-YW*0201
    668 HLA-YX*0201
    669 HLA-YY*0201
    670 HLA-YZ*0201
    671 HLA-ZA*0201
    672 HLA-ZB*0201
    673 HLA-ZC*0201
    674 HLA-ZD*0201
    675 HLA-ZE*0201
    676 HLA-ZF*0201
    677 HLA-ZG*0201
    678 HLA-ZH*0201
    679 HLA-ZI*0201
    680 HLA-ZJ*0201
    681 HLA-ZK*0201
    682 HLA-ZL*0201
    683 HLA-ZM*0201
    684 HLA-ZN*0201
    685 HLA-ZO*0201
    686 HLA-ZP*0201
    687 HLA-ZQ*0201
    688 HLA-ZR*0201
    689 HLA-ZS*0201
    690 HLA-ZT*0201
    691 HLA-ZU*0201
    692 HLA-ZV*0201
    693 HLA-ZW*0201
    694 HLA-ZX*0201
    695 HLA-ZY*0201
    696 HLA-ZZ*0201
    697 HLA-AA*0201
    698 HLA-AB*0201
    699 HLA-AC*0201
    700 HLA-AD*0201
    701 HLA-AE*0201
    702 HLA-AF*0201
    703 HLA-AG*0201
    704 HLA-AH*0201
    705 HLA-AI*0201
    706 HLA-AJ*0201
    707 HLA-AK*0201
    708 HLA-AL*0201
    709 HLA-AM*0201
    710 HLA-AN*0201
    711 HLA-AO*0201
    712 HLA-AP*0201
    713 HLA-AQ*0201
    714 HLA-AR*0201
    715 HLA-AS*0201
    716 HLA-AT*0201
    717 HLA-AU*0201
    718 HLA-AV*0201
    719 HLA-AW*0201
    720 HLA-AX*0201
    721 HLA-AY*0201
    722 HLA-AZ*0201
    723 HLA-BA*0201
    724 HLA-BB*0201
    725 HLA-BC*0201
    726 HLA-BD*0201
    727 HLA-BE*0201
    728 HLA-BF*0201
    729 HLA-BG*0201
    730 HLA-BH*0201
    731 HLA-BI*0201
    732 HLA-BJ*0201
    733 HLA-BK*0201
    734 HLA-BL*0201
    735 HLA-BM*0201
    736 HLA-BN*0201
    737 HLA-BO*0201
    738 HLA-BP*0201
    739 HLA-BQ*0201
    740 HLA-BR*0201
    741 HLA-BS*0201
    742 HLA-BT*0201
    743 HLA-BU*0201
    744 HLA-BV*0201
    745 HLA-BW*0201
    746 HLA-BX*0201
    747 HLA-BY*0201
    748 HLA-BZ*0201
    749 HLA-CA*0201
    750 HLA-CB*0201
    751 HLA-CC*0201
    752 HLA-CD*0201
    753 HLA-CE*0201
    754 HLA-CF*0201
    755 HLA-CG*0201
    756 HLA-CH*0201
    757 HLA-CI*0201
    758 HLA-CJ*0201
    759 HLA-CK*0201
    760 HLA-CL*0201
    761 HLA-CM*0201
    762 HLA-CN*0201
    763 HLA-CO*0201
    764 HLA-CP*0201
    765 HLA-CQ*0201
    766 HLA-CR*0201
    767 HLA-CS*0201
    768 HLA-CT*0201
    769 HLA-CU*0201
    770 HLA-CV*0201
    771 HLA-CW*0
```



## And when there is no experimental structure? Homology modeling



Homology modeling:  
Building a theoretical model of the 3D structure of a protein based on experimentally known 3D structure of related proteins.

Programs et web servers:

- Modeller
- I-Tasser
- Robetta
- HHPred
- ...

Databases of structural models:

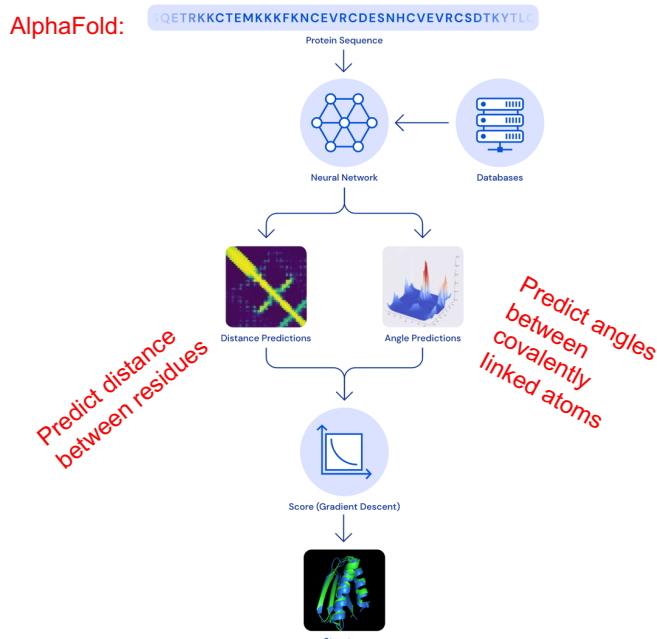
- Swiss-model
- Modbase
- ...

Unil

35

35

## And when there is no experimental structure? Deep Learning



Predict distance between residues:

Sequence of the target protein



Sequence alignment of all apperented proteins

```

C T S Y P I K L M D F E R T S W Q A P R I M T G H K
C S S Y P I K L M D W E R T S W Q A P R I C T G Y K
C Q S Y P L K L M D F E R T S W Q V P R I P T G H K
C N S Y P L K L M D C E R T S W Q V P R I D T G C K
C S S Y P I K L M D F E R T S W Q A P R I F T G H K
C D S Y P V K L M D F E R T S W Q L P R I G T G H K
C C S Y P I K L M D K E R T S W Q A P R I M T G E K
C S S Y P A K L M D F E R T S W Q L P R I K T G H K
C T S Y P I K L M D D E R T S W Q A P R I L T G R K

```

Correlated mutations

Correlated mutations

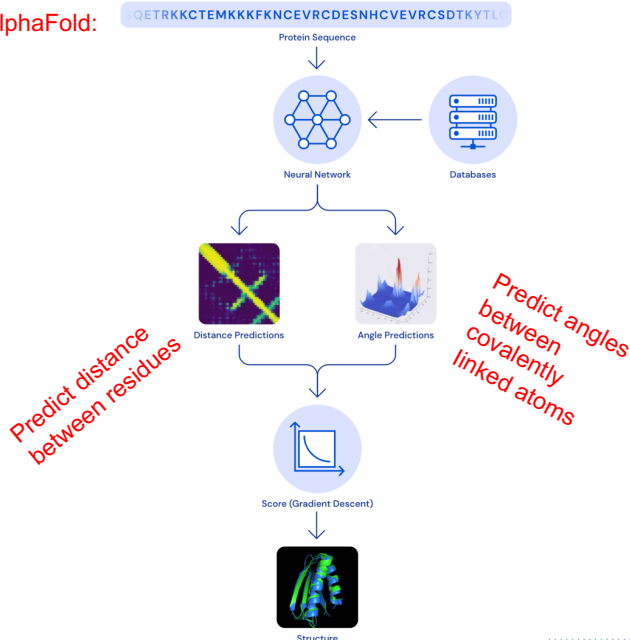
Unil

36

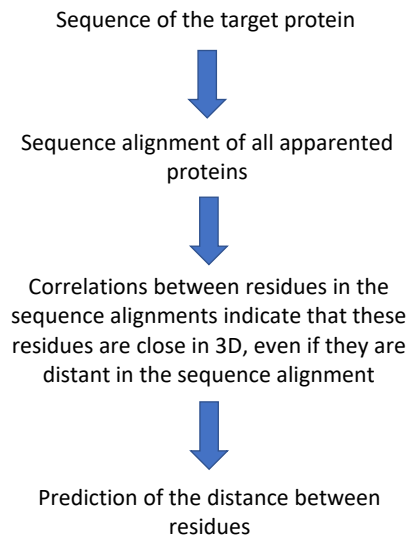
36

## And when there is no experimental structure? Deep Learning

AlphaFold:



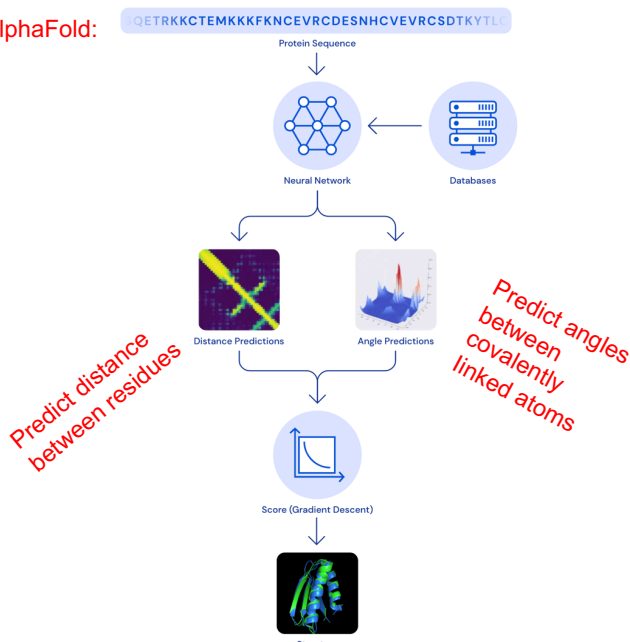
Predict distance between residues:



37

## And when there is no experimental structure? Deep Learning

AlphaFold:



Free database of AlphaFold models:

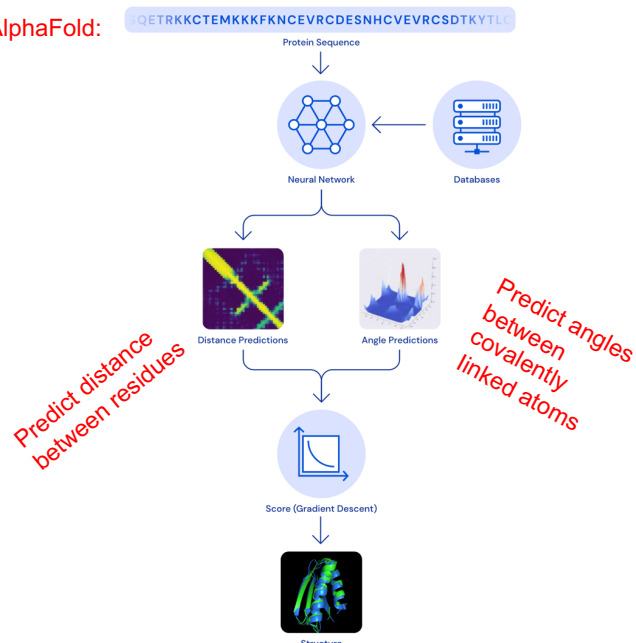
<https://alphafold.ebi.ac.uk>

The screenshot shows the homepage of the AlphaFold Protein Structure Database. The header includes navigation links like 'Home', 'About', 'FAQs', and 'Downloads'. The main content area features the title 'AlphaFold Protein Structure Database' and a search bar. Below the search bar, it states 'AlphaFold DB provides open access to over 200 million protein structure predictions to accelerate scientific research.' A 'Background' section at the bottom describes AlphaFold as an AI system developed by DeepMind that predicts a protein's 3D structure from its amino acid sequence.

38

## And when there is no experimental structure? Deep Learning

AlphaFold:



### WARNINGS!!!!

- AlphaFold models can contain errors and should be interpreted with care
- When existing, experimental structures should be preferred to AlphaFold models
- Although it can provide a good overall structure of a TCR or TCRpMHC, **AlphaFold is not good at predicting the right conformation of the Complementary Determining Loops** (see later)

Possible to use dedicated tools, like

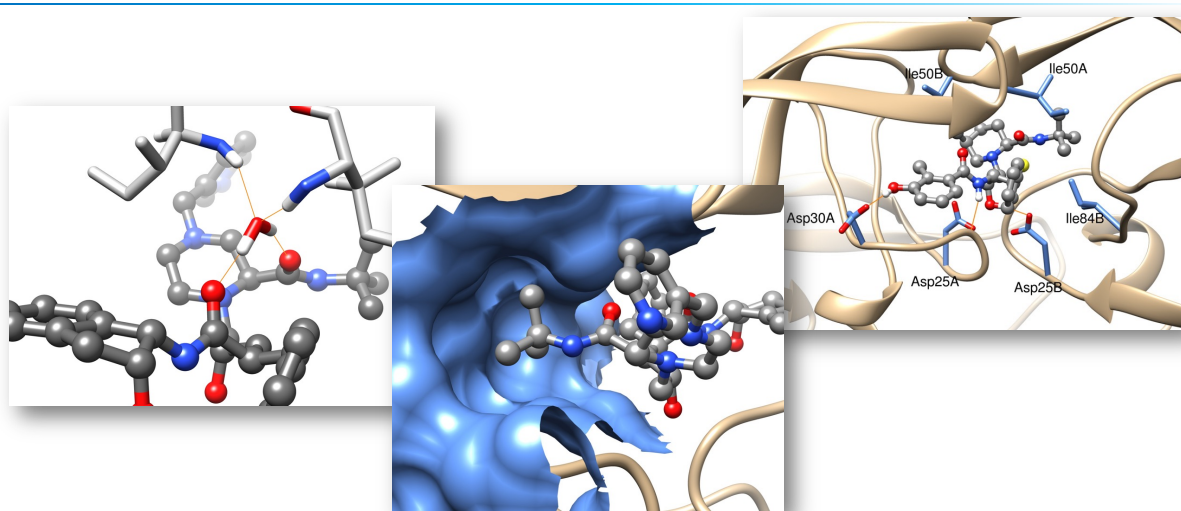
- TCRmodel2, <https://tcmodel.ibbr.umd.edu>
- TCRdock, <https://github.com/phbradley/TCRdock>

Unil

39

39

## Molecular Recognition



Unil

40

40

## Molecular recognition

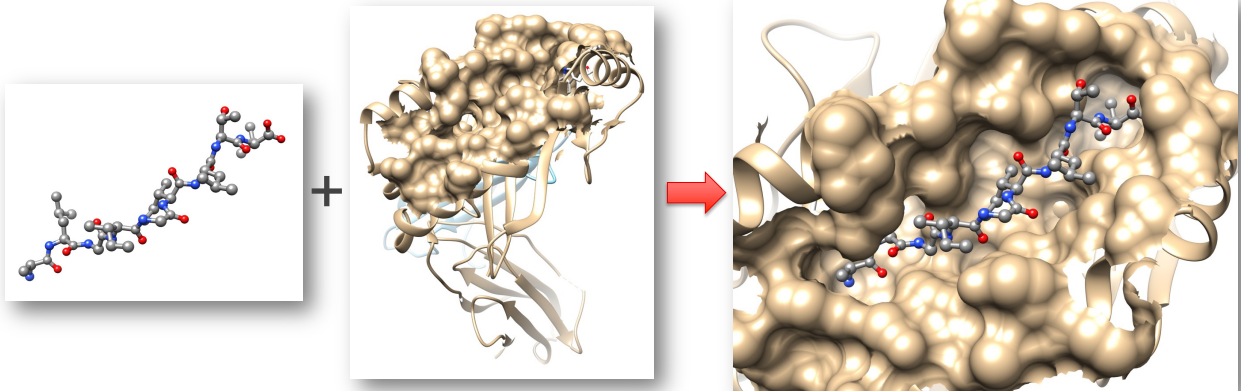
Molecular interactions



Molecular recognition



Biological response



41

41

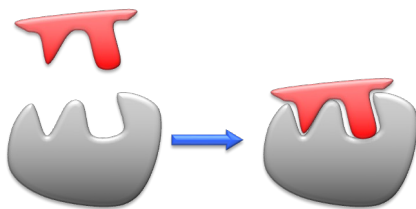
## Molecular recognition – Historical models

### “Lock and key” model.

Emil Fischer in the 1890s.

The protein has a particular shape into which the ligand fits exactly.

Ligand

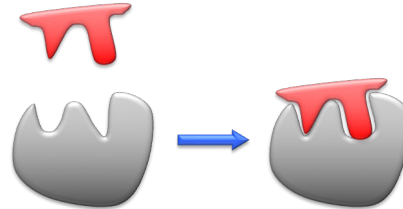


Receptor

### Induced fit model

Daniel Koshland 1958.

The binding site of the macromolecule is flexible and its shape can be modified as the ligand interacts with it.



### Molecular recognition:

Collection of **interactions** between molecules that govern their **binding**.Qualitative **nature** of the interactions?Quantitative **intensity** of the molecular recognition?

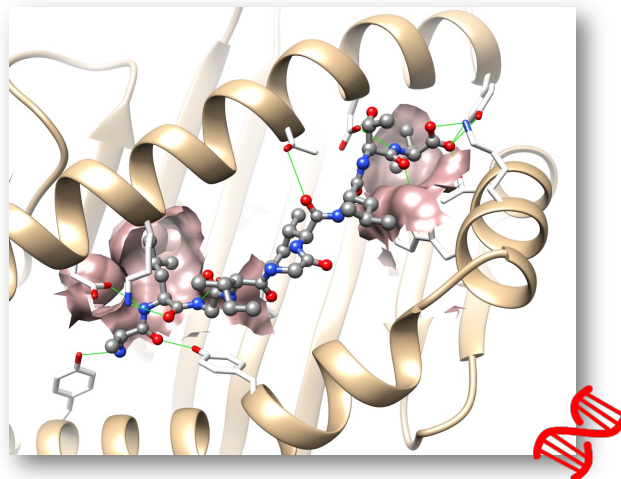
42

42

## Molecular recognition - type of interactions

### Non covalent interactions between atoms :

- non-polar interactions (shape recognition)
- electrostatic interactions (salt bridge and hydrogen bond)
- $\pi$  interactions
- metal/ion interactions



Crystal structure of HLA-A2\*0201 in complex with MART-1/Melan-A

43

## Molecular recognition - type of interactions

### Non Polar:

Ala, Val, Leu, Ile,  
Pro, Met, ~Cys

### Polar:

Ser, Thr, Asn, Gln,  
Tyr, His, Trp, ~Cys

### Aromatic:

Phe, Tyr, Trp, His

### Negatively charged:

Asp, Glu

### Positively charged:

Arg, Lys, ~His

## A GUIDE TO THE TWENTY COMMON AMINO ACIDS

AMINO ACIDS ARE THE BUILDING BLOCKS OF PROTEINS IN LIVING ORGANISMS. THERE ARE OVER 500 AMINO ACIDS FOUND IN NATURE - HOWEVER, THE HUMAN GENETIC CODE ONLY DIRECTLY ENCODES 20. 'ESSENTIAL' AMINO ACIDS MUST BE OBTAINED FROM THE DIET, WHILST NON-ESSENTIAL AMINO ACIDS CAN BE SYNTHESISED IN THE BODY.

**Chart key:** ● ALIPHATIC ● AROMATIC ● ACIDIC ● BASIC ● HYDROXYLIC ● SULFUR-CONTAINING ● AMIDIC ○ NON-ESSENTIAL ○ ESSENTIAL

|   |  |   |   |   |  |
|---|--|---|---|---|--|
| <br><b>ALANINE</b> (A)<br><small>GCT, GCC, GCA, GCG</small> | <br><b>GLYCINE</b> (G)<br><small>GGT, GGC, GGA, GGG</small>          | <br><b>ISOLEUCINE</b> (I)<br><small>ATT, ATC, ATA</small>     | <br><b>LEUCINE</b> (L)<br><small>CTT, CTC, CTA, CTG, TTA, TTG</small> | <br><b>PROLINE</b> (P)<br><small>CCT, CCG, CCA, CCG</small> | <br><b>VALINE</b> (V)<br><small>GTT, GTC, GTA, GTG</small>             |
| <br><b>PHENYLALANINE</b> (F)<br><small>TTT, TTC</small>     | <br><b>TRYPTOPHAN</b> (W)<br><small>TGG</small>                      | <br><b>TYROSINE</b> (Y)<br><small>TAC, TAT</small>            | <br><b>ASPARTIC ACID</b> (D)<br><small>GAT, GAC</small>               | <br><b>GLUTAMIC ACID</b> (E)<br><small>GAA, GAG</small>     | <br><b>ARGININE</b> (R)<br><small>CGT, CGC, CGA, CGG, ACA, AGG</small> |
| <br><b>LYSINE</b> (K)<br><small>AAT, AAG</small>            | <br><b>SERINE</b> (S)<br><small>TCT, TCC, TCA, TCG, AGT, AGC</small> | <br><b>THREONINE</b> (T)<br><small>ACT, ACC, ACA, ACG</small> | <br><b>CYSTEINE</b> (C)<br><small>TGT, TGC</small>                    | <br><b>METHIONINE</b> (M)<br><small>ATG, AAT</small>        | <br><b>ASPARAGINE</b> (N)<br><small>AAT, AAC</small>                   |
| <br><b>GLUTAMINE</b> (Q)<br><small>CAA, CAG</small>         |  |   |   |   |  |

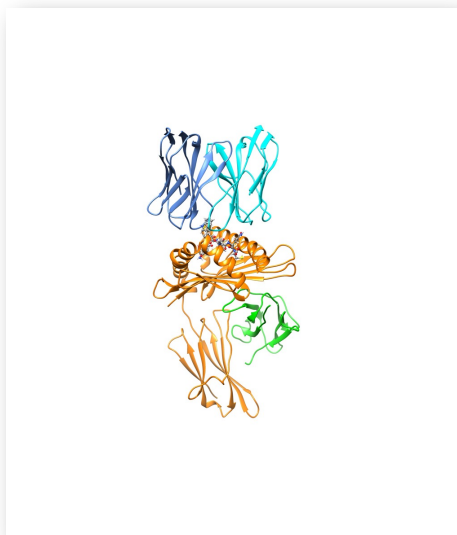
**Note:** This chart only shows those amino acids for which the human genetic code directly codes for. Selenocysteine is often referred to as the 21st amino acid, but is encoded in a special manner. In some cases, distinguishing between asparagine/aspartic acid and glutamine/glutamic acid is difficult. In these cases, the codes asx (B) and glx (Z) are respectively used.

44

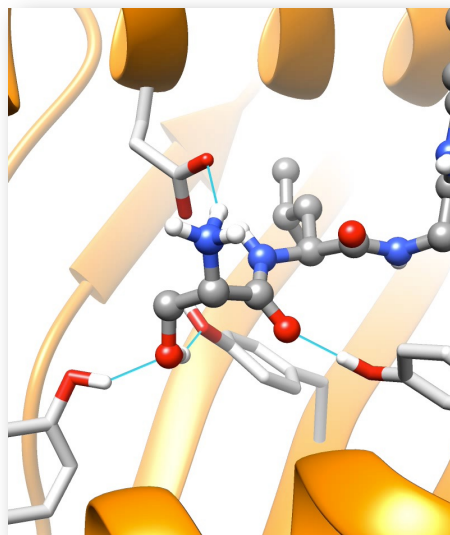


## Molecular recognition – introduction to molecular mechanics

Structure determination



Biological events



Unil

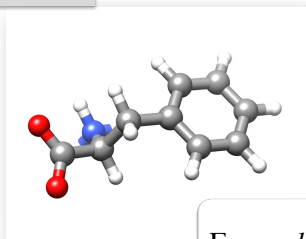
UNIL | Université de Lausanne

45

45

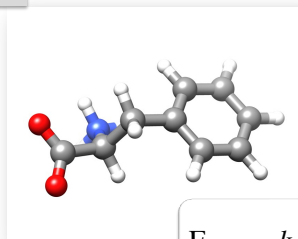
## Molecular recognition – introduction to molecular mechanics

Bond length



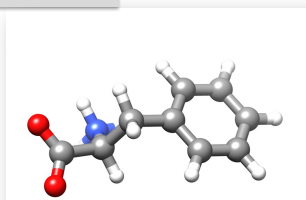
$$E_{bond} = k_b (b - b_0)^2$$

Bond angle



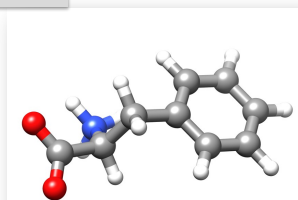
$$E_{angle} = k_\theta (\theta - \theta_0)^2$$

Dihedral angle



$$E_{dihedral} = k_\varphi (1 + \cos(n\varphi - \delta))$$

Improper angle



$$E_{improper} = k_\omega (\omega - \omega_0)^2$$

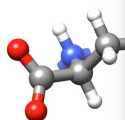
Molecular dynamics is decomposed into elementary motions

46

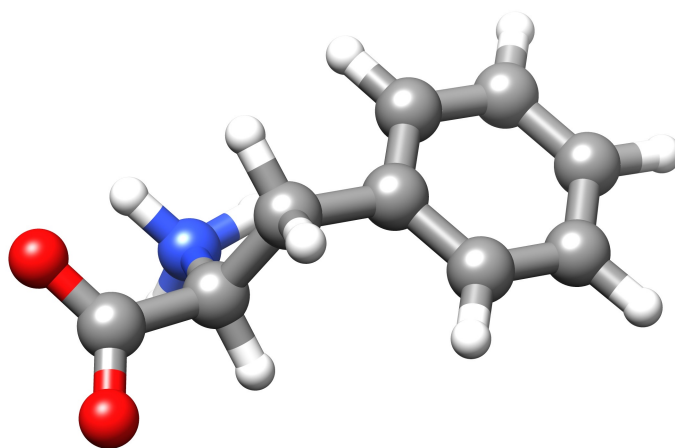
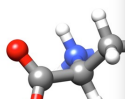
46

## Molecular recognition – introduction to molecular mechanics

Bond length



Dihedral angle



Molecular dynamics is decomposed into elementary motions

$$E = k_{\theta}(\theta - \theta_0)^2$$

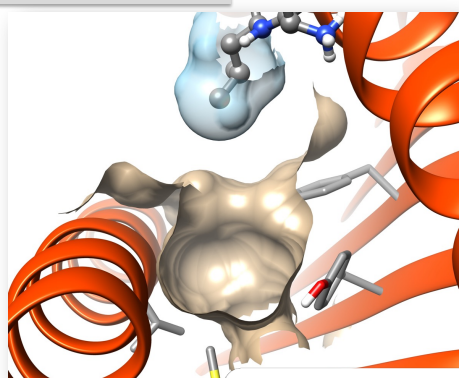
$$E_{\text{bonded}} = \sum_{\text{bonds}} k_b (b - b_0)^2 + \sum_{\text{angles}} k_{\theta} (\theta - \theta_0)^2 + \sum_{\text{dihedrals}} k_{\varphi} (1 + \cos(n\varphi - \delta)) + \sum_{\text{impropers}} k_{\omega} (\omega - \omega_0)^2$$

47

## Molecular recognition – Molecular interactions

Molecular recognition is driven by non-polar and electrostatic interactions

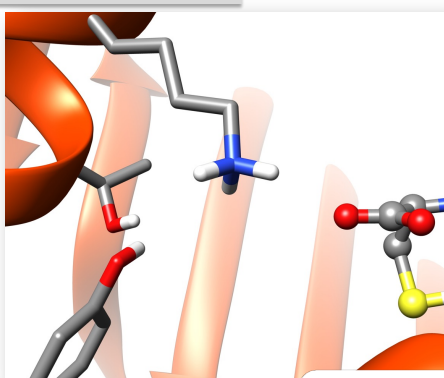
Non-polar interactions



$$E_{\text{vdW}} = \epsilon \left[ \left( \frac{r_m}{r_{ij}} \right)^{12} - 2 \left( \frac{r_m}{r_{ij}} \right)^6 \right]$$

→ Shape complementarity

Electrostatic interactions



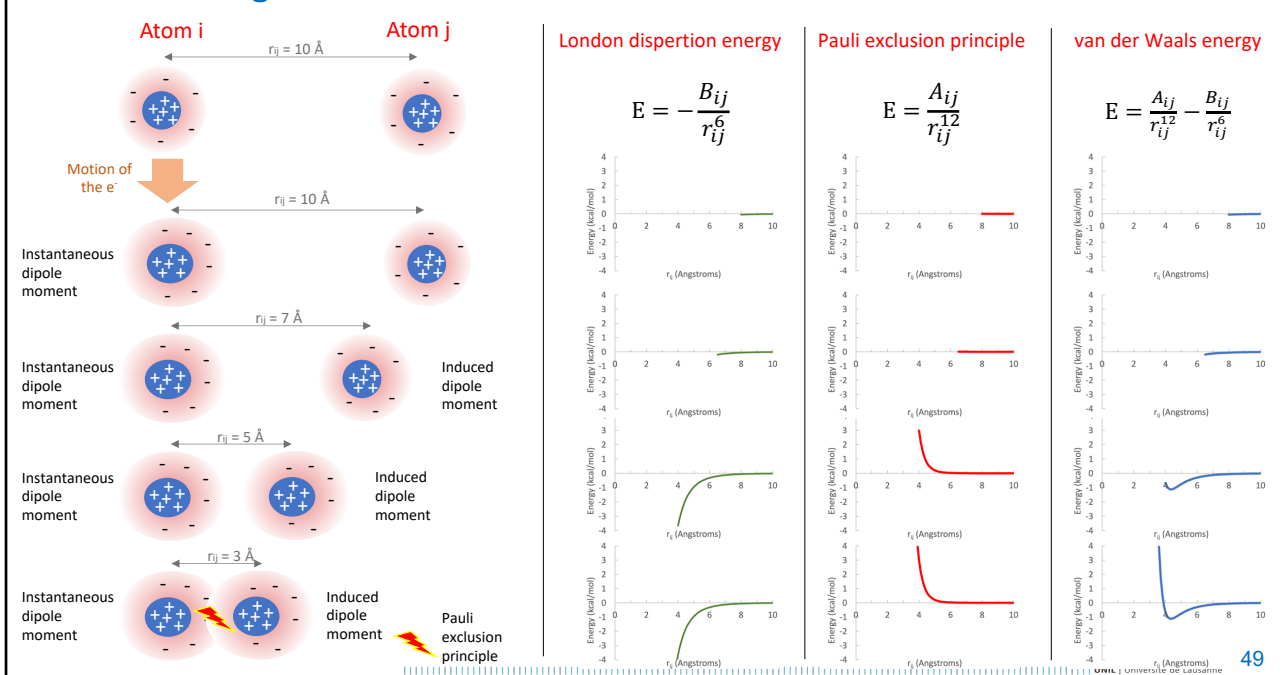
$$E_{\text{elec}} = \frac{q_i q_j}{4\pi\epsilon_0\epsilon r_{ij}}$$

→ Specificity

Unil 48  
UNIL | Université de Lausanne

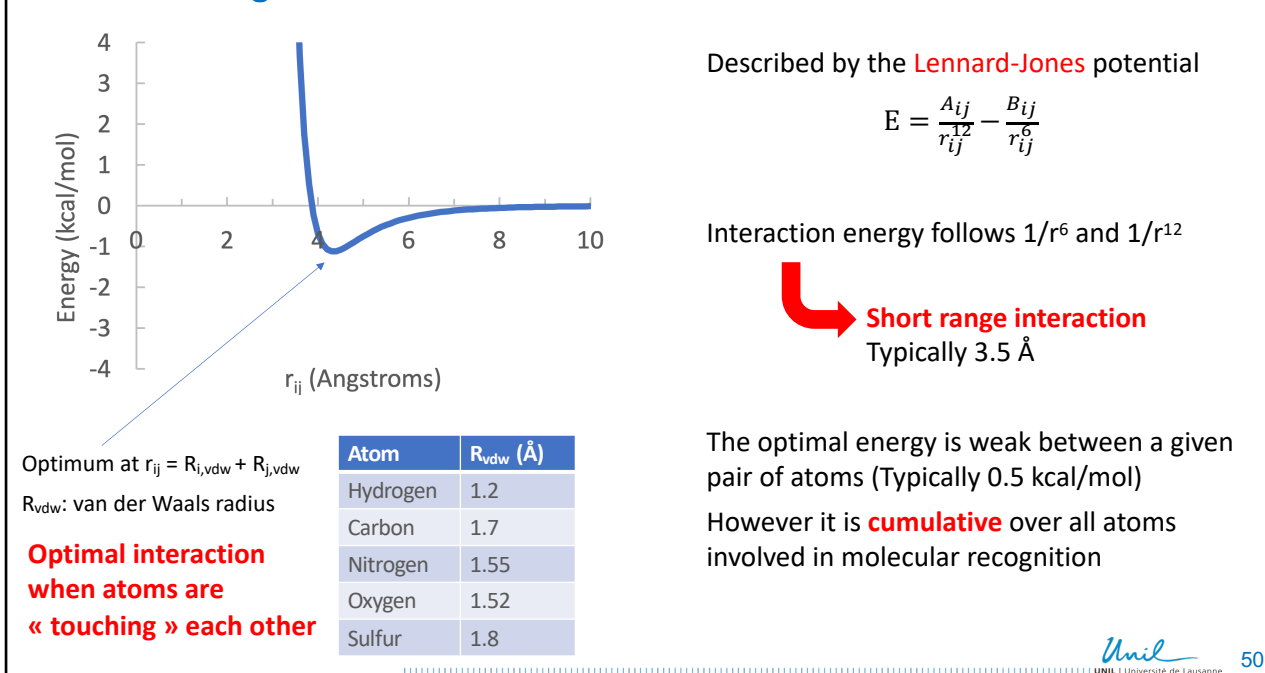
48

## Molecular recognition – Van der Waals interactions



49


## Molecular recognition – Van der Waals interactions



50

## Molecular recognition – Van der Waals interactions

Do not require charges or partial charges on atoms

 **van der Waals interactions are considered as non-polar interactions**  
... even though they are electrostatic by nature

Interactions particularly **important for non-polar residues**:

- Alanine, Valine, Leucine, Isoleucine, Proline
- Cysteine, Methionine
- Phenylalanine, Tyrosine, Tryptophan

## Molecular recognition – Van der Waals interactions

Each atom tries to be positioned at optimal distance from its neighbors

2 atoms



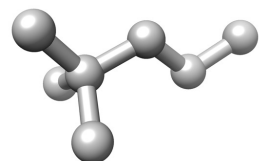
3 atoms



4 atoms

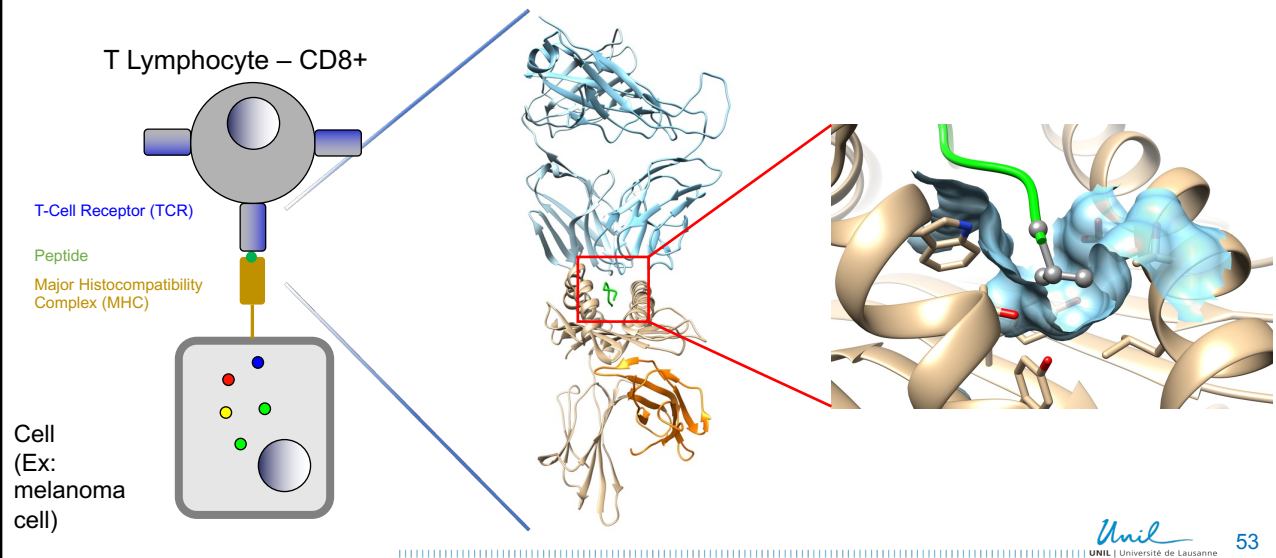


However, in molecules, atoms are also linked via covalent bonds, which force a geometry...



## Molecular recognition – Van der Waals interactions

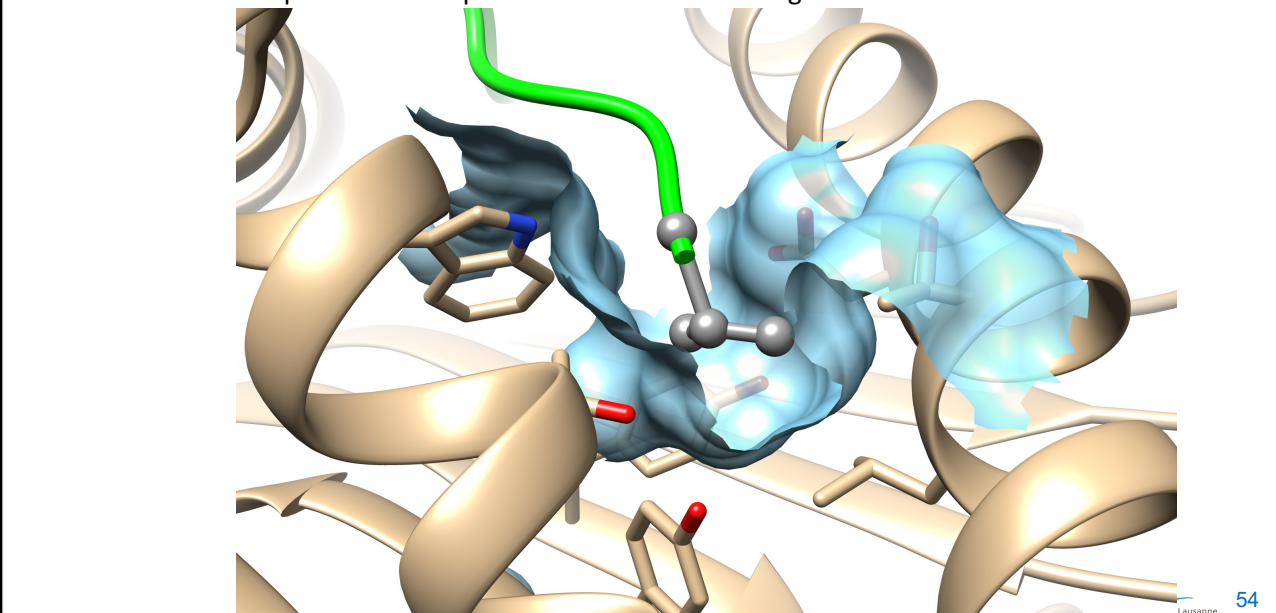
Each atom tries to be positioned at optimal distance from its neighbors



53

## Molecular recognition – Van der Waals interactions

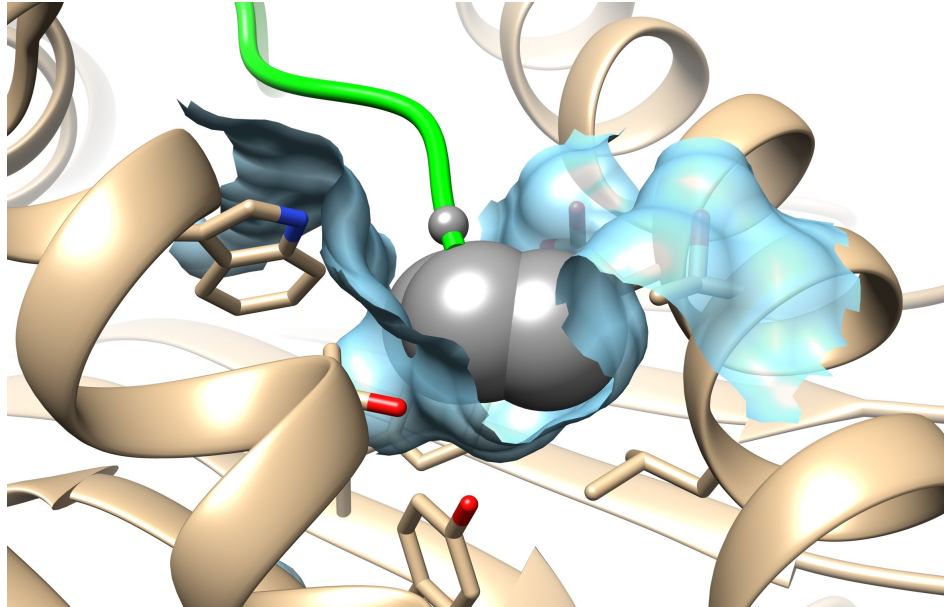
Each atom tries to be positioned at optimal distance from its neighbors



54

## Molecular recognition – Van der Waals interactions

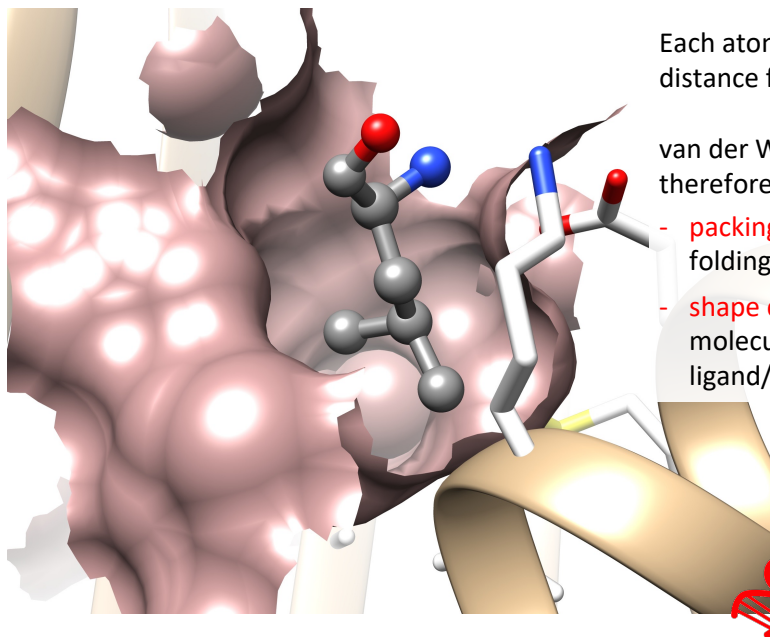
Each atom tries to be positioned at optimal distance from its neighbors



55

55

## Molecular recognition – Van der Waals interactions



Each atom tries to be positioned at optimal distance from its neighbors

van der Waals interactions contribute therefore to:

- **packing of atoms** (and macromolecule folding)
- **shape complementarity** between binding molecules (example: protein/protein or ligand/protéine complexes)

Unil  
UNIL | Université de Lausanne 56

56



## Molecular recognition – Electrostatic interactions

The interaction between two point charges in a uniform medium is described by the **Coulomb law**



Coulomb energy

$$E_{\text{Coul}} = \frac{1}{4\pi\epsilon_0\epsilon} \frac{q_i q_j}{r_{ij}}$$

$\epsilon_0$  : dielectric constant of vacuo

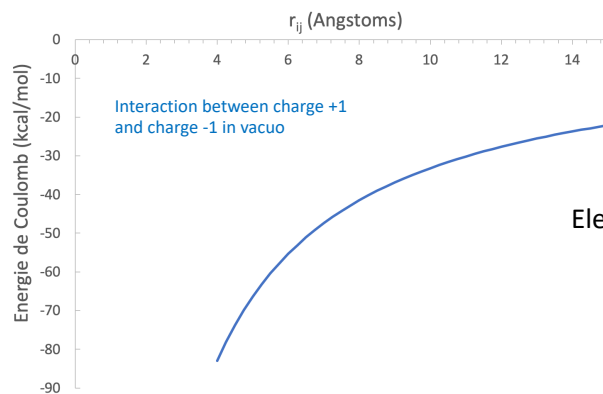
$$\frac{1}{4\pi\epsilon_0} = 332 \text{ (kcal/mol) } \text{\AA}^2 / q_e^2$$

$\epsilon$ : dielectric constant of medium

ex:  $\epsilon_{\text{(vacuo)}} = 1$  ;  $\epsilon_{\text{(water)}} = 80$

Interaction between charges +1 et -1 at 5 Å :

- -66 kcal/mol in vacuo
- -0.8 kcal/mol in water



Electrostatic interaction energy follows a  $1/r$  expression

**Long range interaction**

57

## Molecular recognition – Electrostatic interactions

Electrostatic interactions can involve:

- Integer charge – integer charge

Called **ionic interactions**.

At short distance ( $\sim 4/5$  Å), ionic interactions are called **salt bridges**.



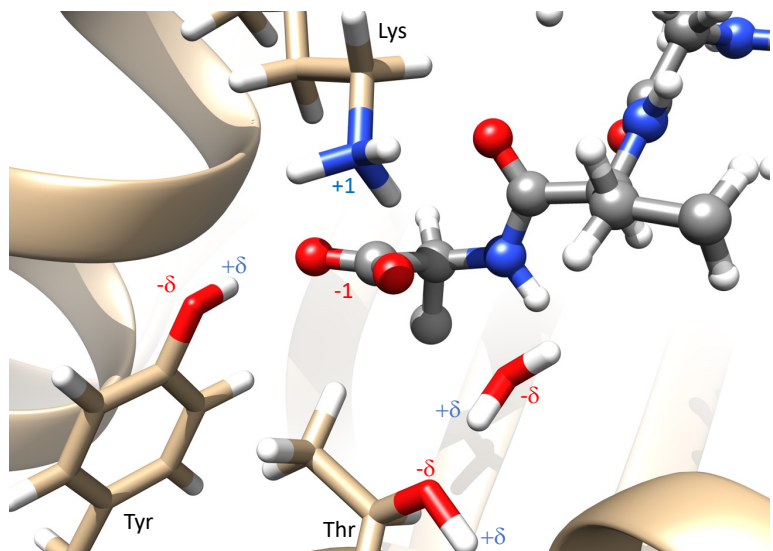
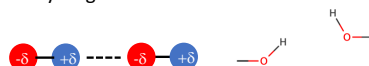
- Integer charge – permanent dipole

Ex: charged assisted hydrogen bond



- Permanent dipole – permanent dipole

Ex: hydrogen bond



58

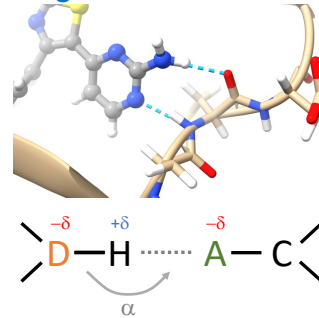
## Molecular recognition – Electrostatic interactions – Hydrogen bonds

Typically between two dipoles:

- D-H where D is the hydrogen bond **donor**
- A-C where A is the hydrogen bond **acceptor** and C a carbon atom

**Extremely frequent** in proteins and nucleic acids

Important factor of the architecture of bio-macromolecules



Typical distances in hydrogen bonds:

- Between H and A :  $\sim 1.95 \text{ \AA}$
- Between A and D : O – O :  $2.50 - 2.70 \text{ \AA}$   
O – N :  $2.75 - 2.85 \text{ \AA}$   
N – N :  $2.70 - 3.00 \text{ \AA}$

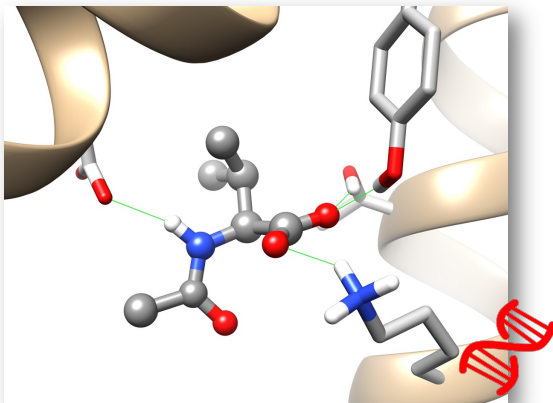
Angle  $\alpha$  depends on atom types and atom hybridization

Unil

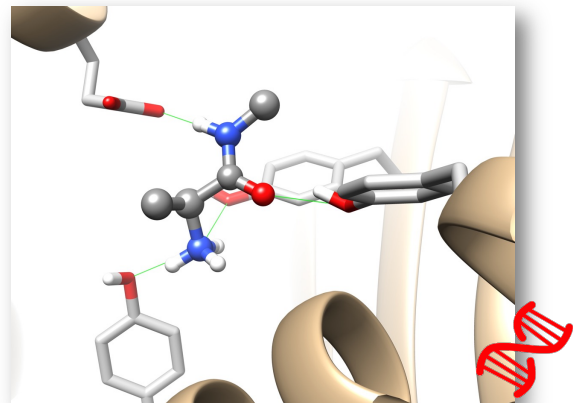
59

59

## Molecular recognition – Electrostatic interactions – Hydrogen bonds



Ex: H-bonds between residue Val9 of MART-1/Melan-A and pocket F of HLA-A2\*0201



Ex: H-bonds between residue Ala1 of MART-1/Melan-A and pocket A of HLA-A2\*0201

Electrostatic interactions are **local and directional** (H-bonds even more than salt bridges)

➔ **Directionality / locality of interactions**  
**Specificity of molecular recognition**

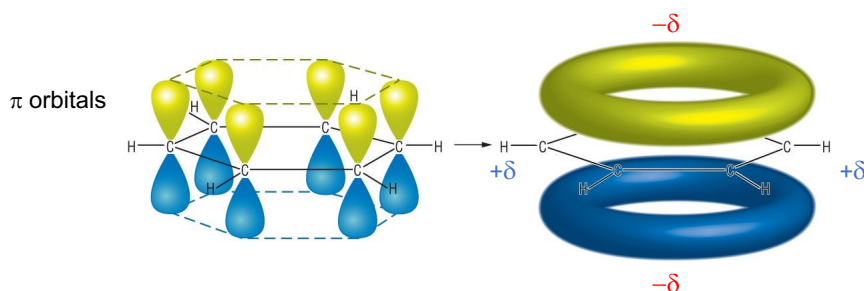
Unil

60

60

## Molecular recognition – $\pi$ interactions

Electronic structure of benzene:



Aromatic cycles (Phenyl, Tyrosine, Tryptophan & Histidine) can interact with:

- Other aromatic cycles (stacking)
- Metals
- Polar groups
- Hydrogen bond donors

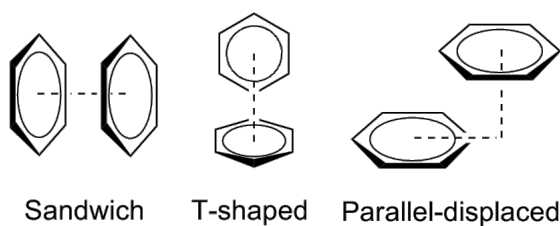
Unil

UNIL | Université de Lausanne

61

61

## Molecular recognition – $\pi$ interactions



(source: Wikipedia)

T-shaped and parallel-displaced  $\pi$ - $\pi$  interactions are the most frequent

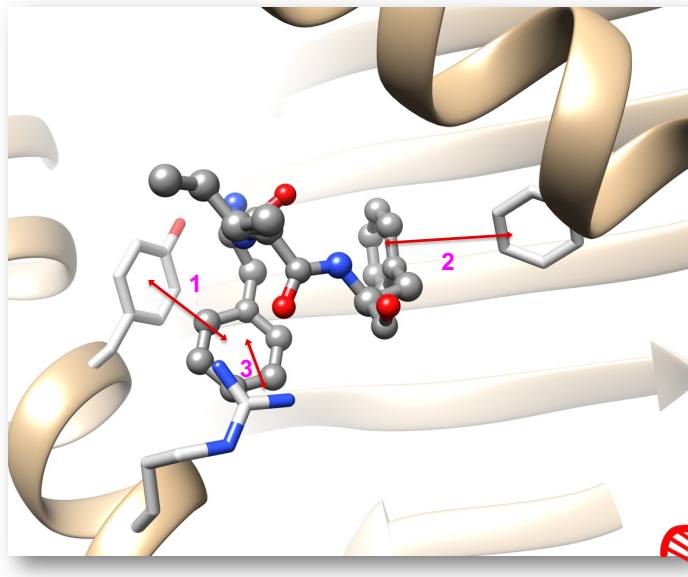
Unil

UNIL | Université de Lausanne

62

62

## Molecular recognition – $\pi$ interactions



1.  $\pi$  - stacking
2. T - staking (the two aromatic cycles are orthogonal)
3. Cation -  $\pi$  interaction

Ex:  $\pi$  interactions between murine coronavirus epitope RCEIFANI and H-2Kb (4PV8 in PDB)

Unil

UNIL | Université de Lausanne

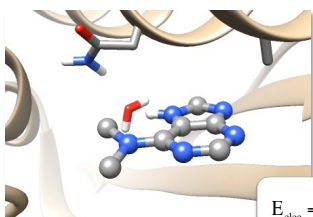
63

63

## Molecular recognition – Other factors

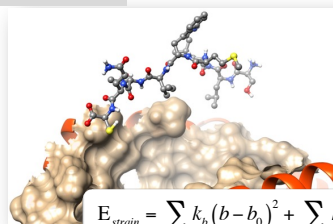
Many other factors impact the molecular recognition and binding affinity

### Water bridges



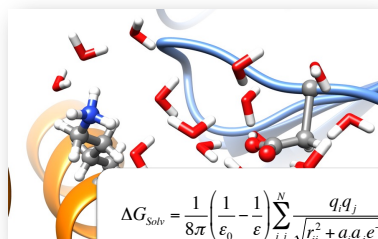
$$E_{\text{elec}} = \frac{q_i q_j}{4\pi\epsilon_0\epsilon r_{ij}}$$

### Conformational changes



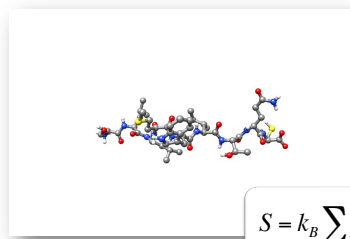
$$E_{\text{strain}} = \sum_{\text{bonds}} k_b (b - b_0)^2 + \sum_{\text{angles}} k_\theta (\theta - \theta_0)^2 + \dots$$

### Desolvation and elec. shielding



$$\Delta G_{\text{solv}} = \frac{1}{8\pi} \left( \frac{1}{\epsilon_0} - \frac{1}{\epsilon} \right) \sum_{i,j} \frac{q_i q_j}{\sqrt{r_{ij}^2 + a_i a_j e^{-D}}} , D = \left( \frac{r_{ij}}{2\sqrt{a_i a_j}} \right)^2$$

### Entropy changes



$$S = k_B \sum p_i \ln(p_i)$$

UNIL | Université de Lausanne

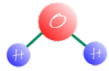
64

64

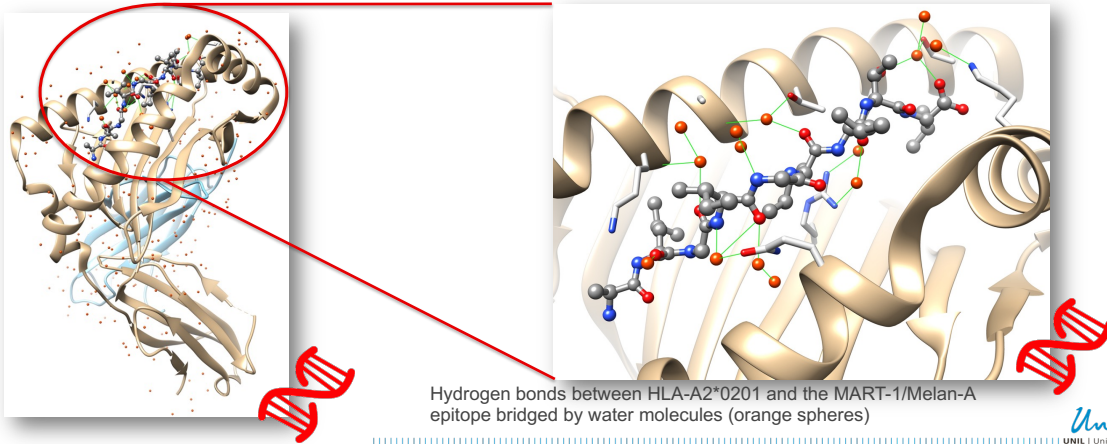
## Molecular recognition – Other factors – Water

Molecular recognition between small molecule and protein takes place in an **aqueous environment**.

### Discrete water molecules



- Bridge interactions through H-bonds or OH...  $\pi$   $\rightarrow$  favorable to binding.
- Displacement from the protein cavity  $\rightarrow$  favorable to binding.

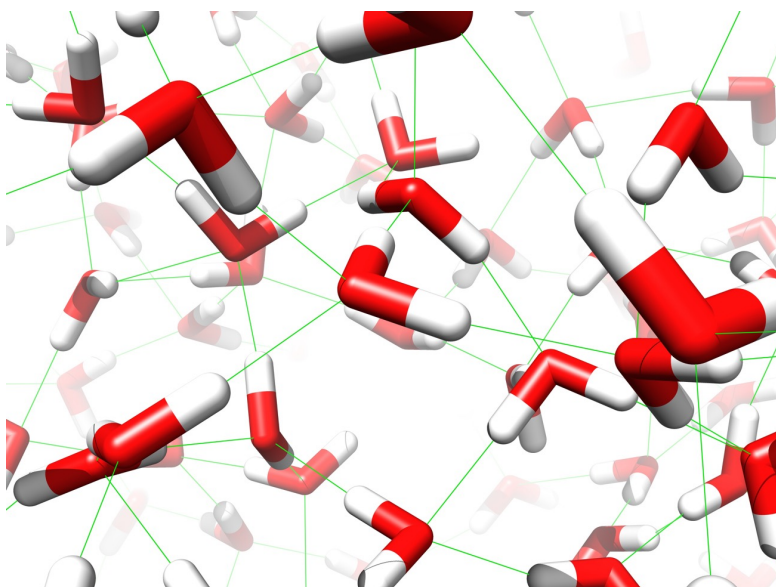


Unil  
UNIL | Université de Lausanne

65

65

## Molecular recognition – Other factors – Water – Hydrophobic effect



Water structure is stabilized by hydrogen bonds and dipole interactions

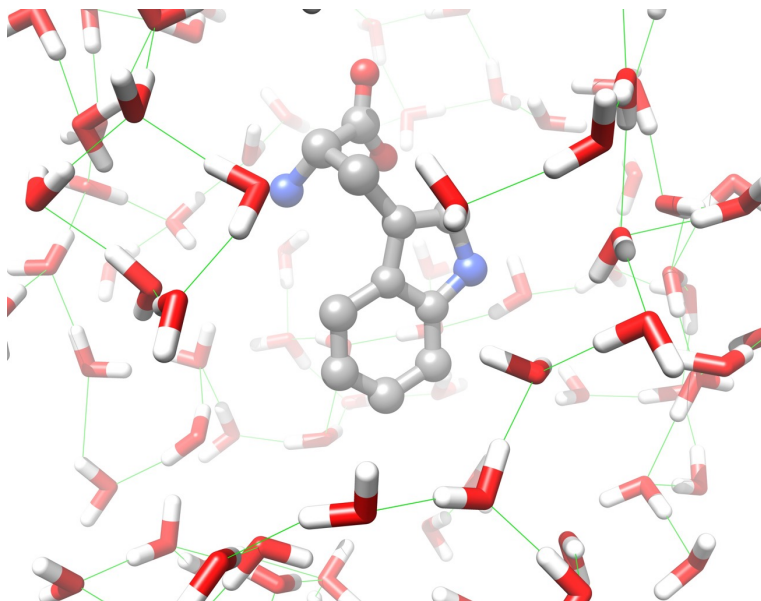
Unil  
UNIL | Université de Lausanne

66

66



## Molecular recognition – Other factors – Water – Hydrophobic effect

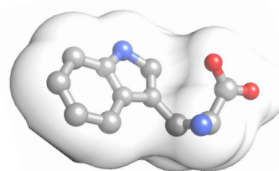


The presence of a solute decreases water-water interactions

Non-polar solvation energy is proportional to the solvent accessible surface area (SASA) for large molecules:

$$E = \sigma \times \text{SASA}$$

$$\sigma = 0.025 \text{ kcal/\AA}^2$$



Unil

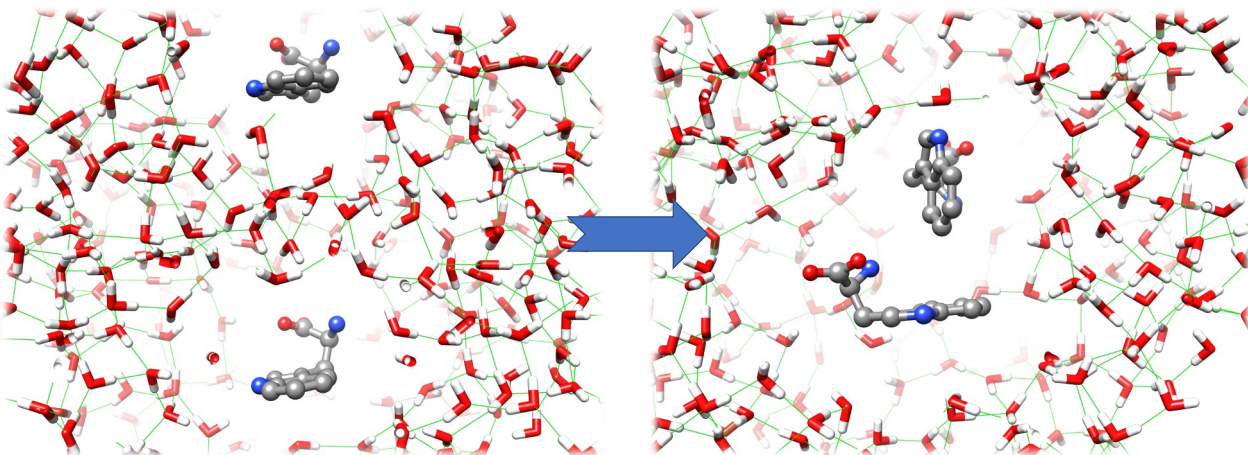
UNIL | Université de Lausanne

67

67

## Molecular recognition – Other factors – Water – Hydrophobic effect

Solutes aggregate to limit their deleterious on water structure



$$\text{Energy of non-polar desolvation: } \Delta G_{np} = \sigma \times \Delta \text{SASA}$$

The solvent-accessible surface area of aggregated solutes is lower than the sum of those of the separated solutes ( $\Delta \text{SASA} < 0$ ).  $\Delta G_{np}$  is therefore favorable to aggregation (binding of solutes)

Unil

UNIL | Université de Lausanne

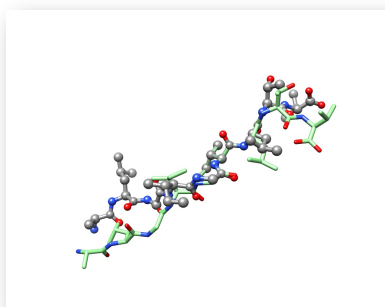
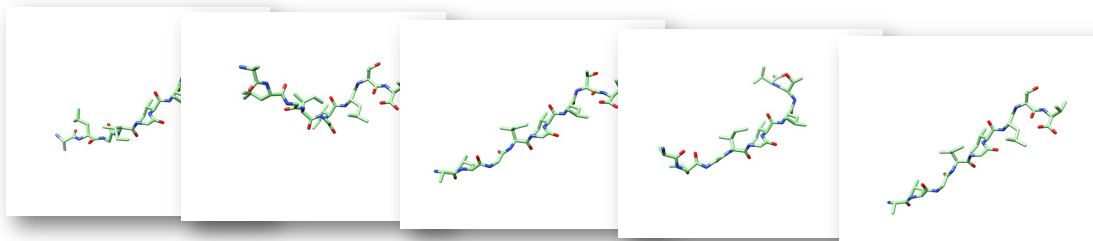
68

68



## Molecular recognition – Other factors – Conformational changes

Molecules have many conformations (conformers)



Peptide **bioactive** conformation (geometry as bound to the protein)

does **NOT** correspond to

**Lowest energy** conformation (most stable geometry in solution)

BUT is a low energy conformation (within 3 to 5 kcal/mol)

Bioactive conformation (in protein)

Lowest energy conformation (in solution)

Unil

UNIL | Université de Lausanne

69

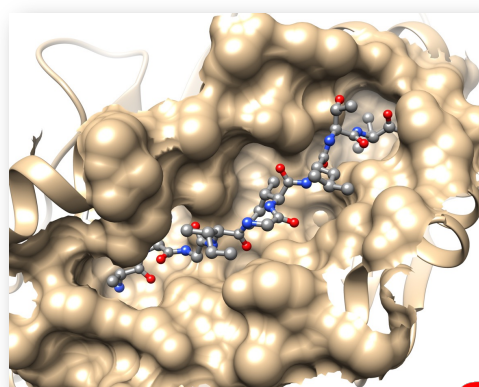
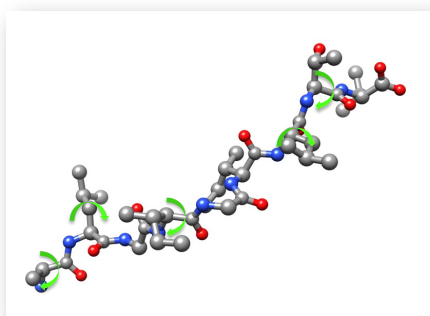
69

## Molecular recognition – Other factors – Entropy changes

Entropy is a measure of disorder. Nature likes disorder!

Loss of entropic energy when entropy (disorder) decreases.

Gain of entropic energy when entropy (disorder) increases.



Two main events upon ligand binding to protein:

- **Conformational** degrees of **freedom** (rotatable bonds) are **blocked**: **unfavorable**!
- **Water** molecules are **kicked-out** from the protein binding site to **bulk**: **favorable**!

Unil

UNIL | Université de Lausanne

70

70

## Molecular recognition – Summary

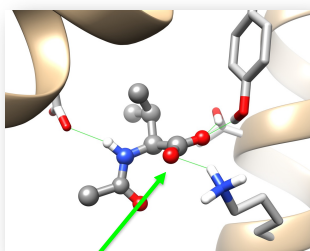
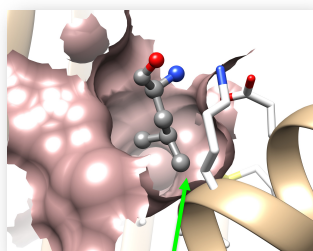
| Category                | Interaction           | Distance    | Residues involved   | Remarks   |
|-------------------------|-----------------------|-------------|---|---|
| Electrostatic           | Ionic (charge-charge) | Long range  | Arg, Lys, Asp, Glu<br>His (if charged)                      | Called salt bridge at short distance  |
|                         | Hydrogen bond         | Short range | Arg, Lys, Asp, Glu<br>His, Tyr<br>Ser, Thr, Asn, Gln<br>Cys | Directionality / locality of interactions<br>Specificity of molecular recognition |
|                         | $\pi$ interaction     | Short range | Phe, Tyr, Trp, His  |   |
| Electrostatic/Non-polar | Van der Waals         | Short range | Ala, Val, Ile, Leu, Pro,<br>Cys, Met<br>Phe, Tyr, Trp, His  | Packing of atoms<br>Shape complementarity   |
| Non-polar               | Hydrophobic effect    | -           | All   | Solute aggregation  |



71

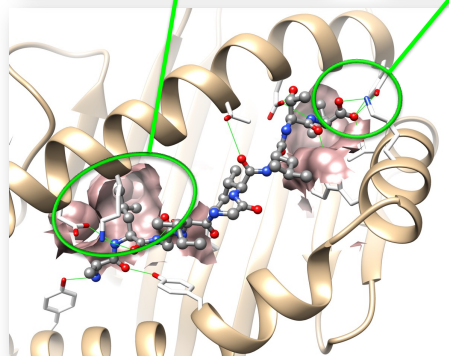
71

## Molecular recognition – Potency and specificity



Various and numerous ligand-protein interactions:

- local and directional interactions
- shape complementarity



**Specificity**

(Limits number/nature of possible epitopes)

**Affinity/potency**

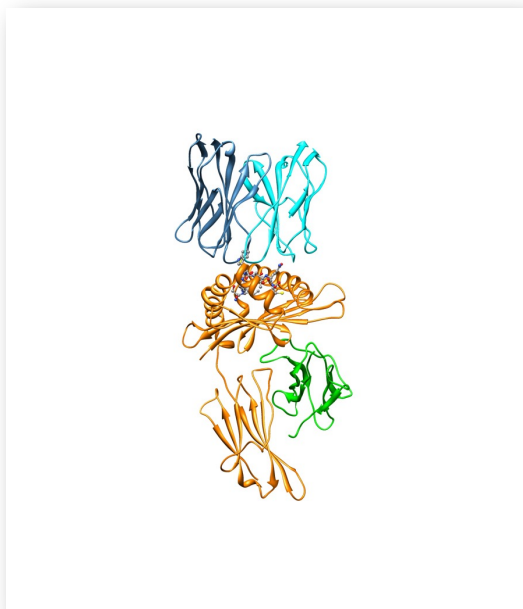
(Increased epitope recognition)



72

72

## Molecular recognition - Molecular Mechanics - Molecular Dynamics



- Adding explicit droplet of water:

**System** solvated with explicit water molecules (TIP3P model):

- ~ 29,500 water molecules
- ~ 100,000 atoms in total

- **Molecular Dynamics (MD)**

Atom motions are calculated to follow Newton's equation of motion, at **300 K** and **1 atm**.

Typical simulation times: from **0.5 ns** to ~ **100 ns** ( $10^{-9}$  s).

→ Simulation closer to physiological reality, but more computationally intensive

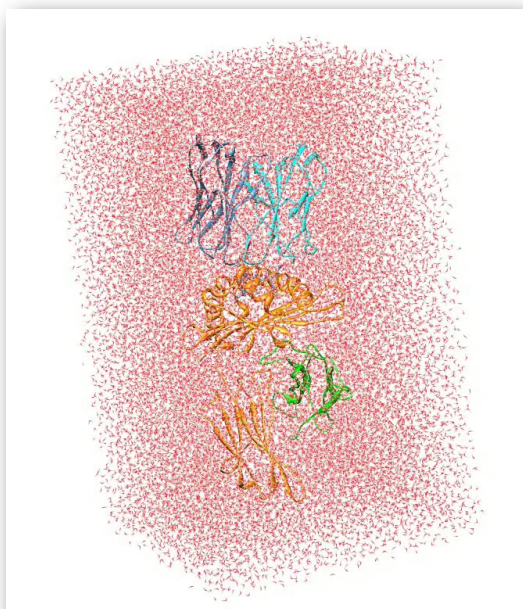
Unil

UNIL | Université de Lausanne

73

73

## Molecular recognition - Molecular Mechanics - Molecular Dynamics



- Adding explicit droplet of water:

**System** solvated with explicit water molecules (TIP3P model):

- ~ 29,500 water molecules
- ~ 100,000 atoms in total

- **Molecular Dynamics (MD)**

Atom motions are calculated to follow Newton's equation of motion, at **300 K** and **1 atm**.

Typical simulation times: from **0.5 ns** to ~ **100 ns** ( $10^{-9}$  s).

→ Simulation closer to physiological reality, but more computationally intensive

Unil

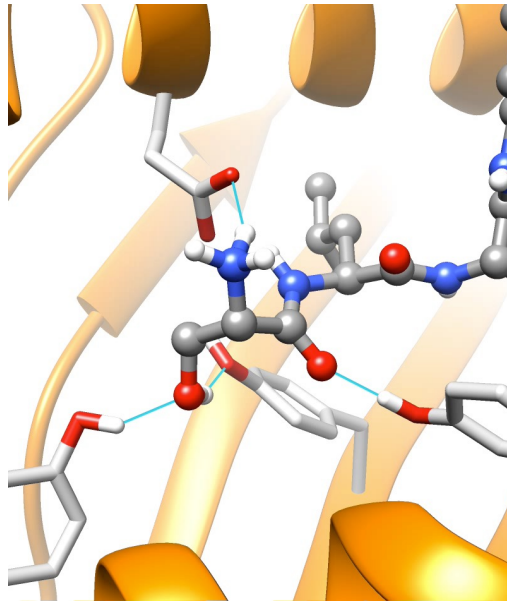
UNIL | Université de Lausanne

74

74

## Molecular recognition - Molecular Mechanics - Molecular Dynamics

Typical motions in a peptide/MHC complex at room temperature:

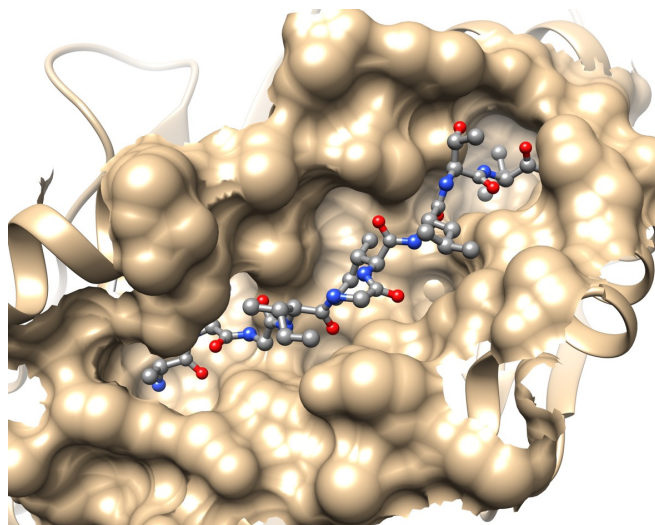


Unil  
UNIL | Université de Lausanne

75

75

## 3D structure of peptides bound to the Major Histocompatibility Complex Class I (pMHC)



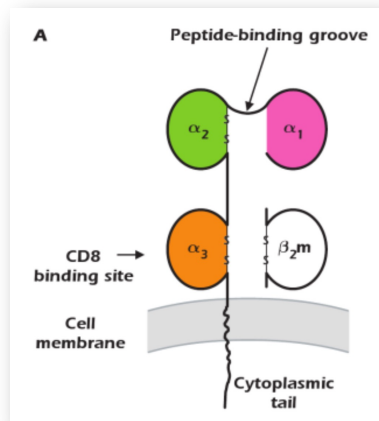
Unil  
UNIL | Université de Lausanne

76

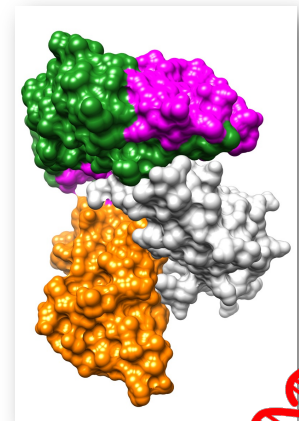
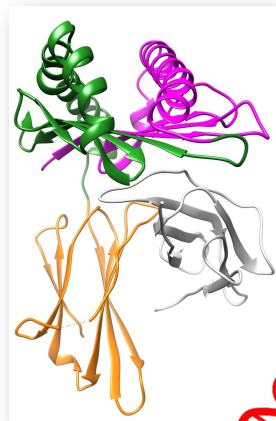
76

## Structure of pMHC Class I

- Transmembrane glycoprotein (43 kDa)
- 3 domains  $\alpha_1$ ,  $\alpha_2$  and  $\alpha_3$
- Deep groove to bind a peptide (8-10 residues)
- Expressed at cell surface
- Non-covalent association with  $\beta_2$  microglobulin ( $\beta_2m$ )



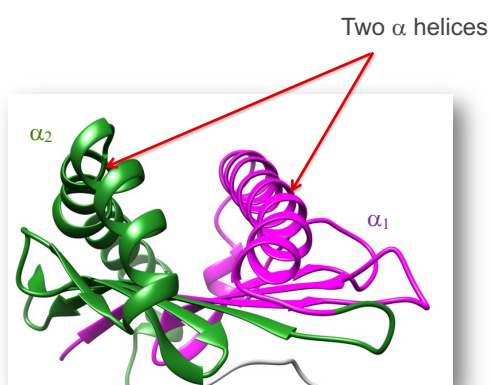
From "Immunology: A Short Course"  
by Richard Coico, Geoffrey Sunshine



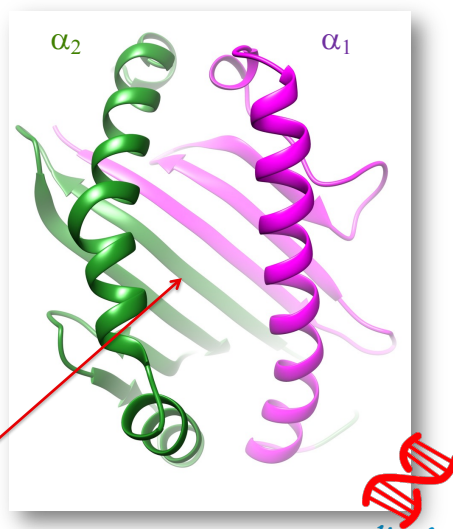
Unil 77  
UNIL | Université de Lausanne

77

## Structure of pMHC Class I - The binding pocket



$\beta$  sheet formed of 8  $\beta$  strands

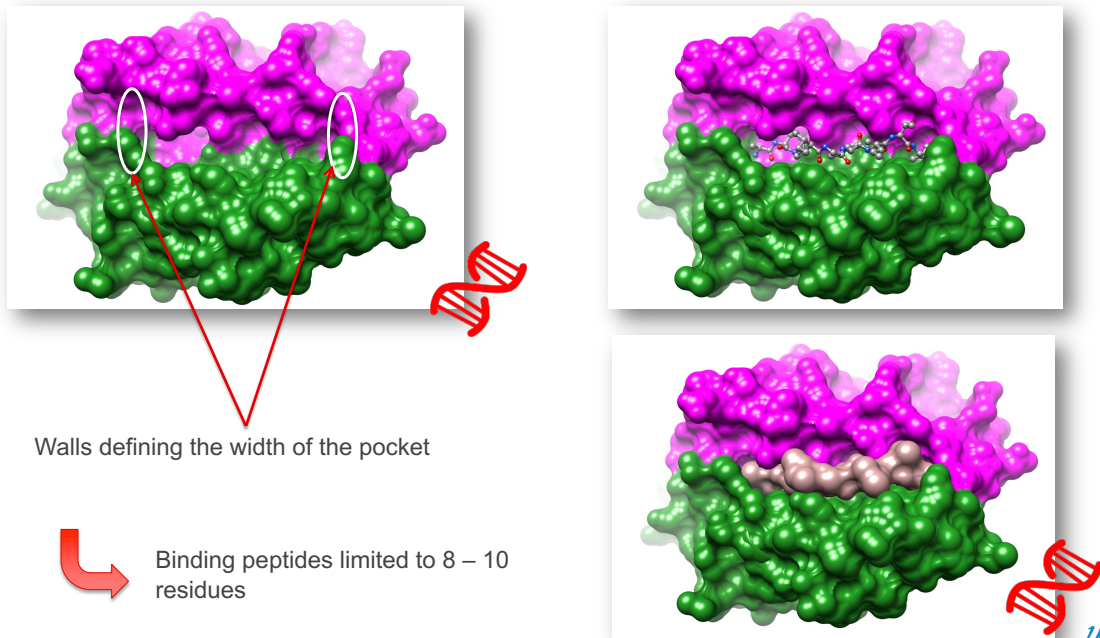


Unil 78  
UNIL | Université de Lausanne

78



### Structure of pMHC Class I - The binding pocket

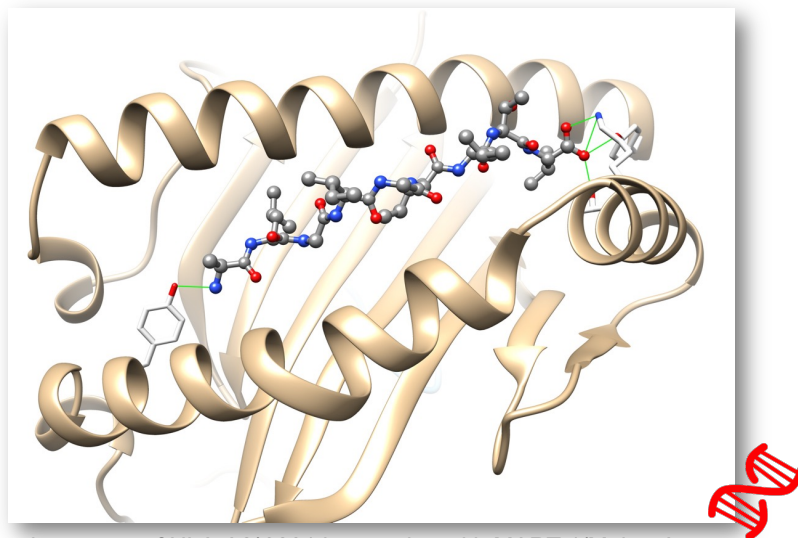


79

79

### Structure of pMHC Class I - Peptide binding to MHC groove

N and C-termini of the peptide are stabilized by H-bonds networks



Crystal structure of HLA-A2\*0201 in complex with MART-1/Melan-A

Unil  
UNIL | Université de Lausanne

80

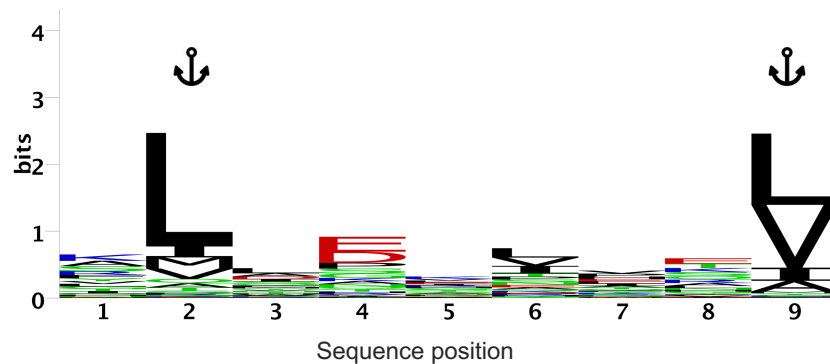
80



## Structure of pMHC Class I - Peptide binding to MHC groove

- The specificity of the peptide binding is provided by the **anchor residues**, i.e. invariant or closely related residues in the peptide sequence.
- A peptide binding to MHC Class I has typically 2 main anchor residues and 2/3 secondary anchor residues.
- Other positions are variable.

### Sequence logo for HLA-A2\*02:01



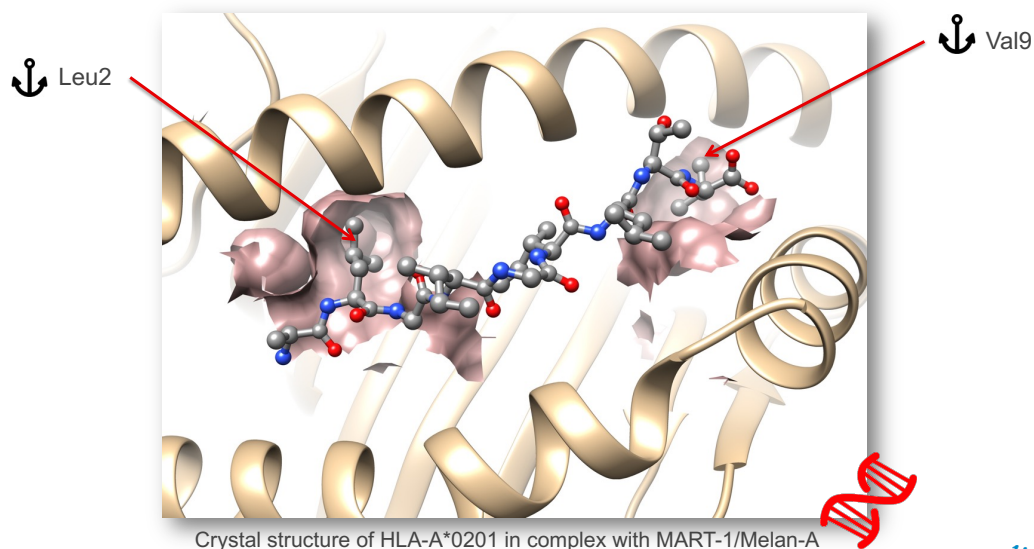
Unil  
UNIL | Université de Lausanne

81

81

## Structure of pMHC Class I - Peptide binding to MHC groove

Anchor residues make strong interactions with the MHC binding pockets



Crystal structure of HLA-A\*0201 in complex with MART-1/Melan-A

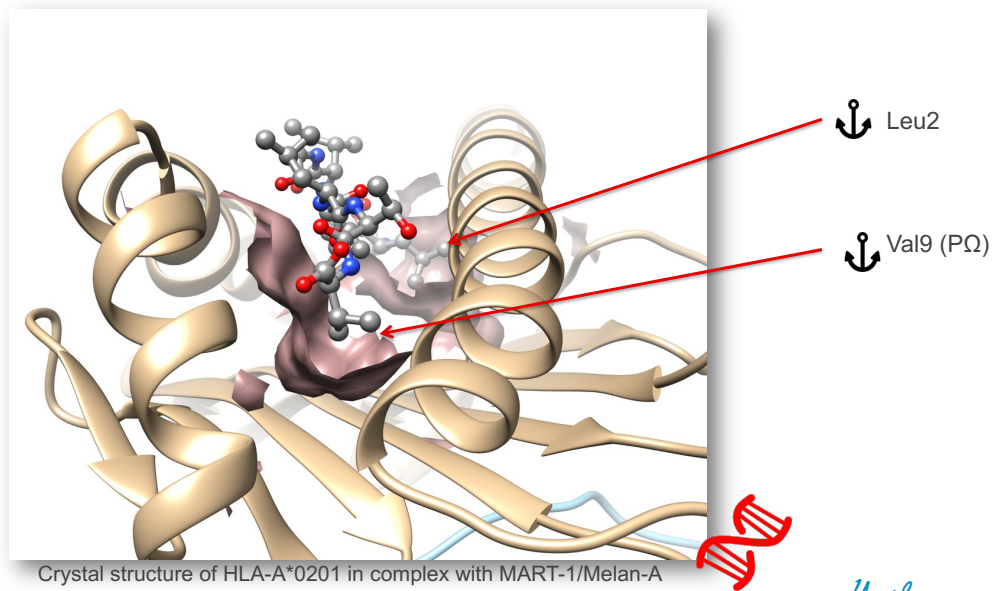
Unil  
UNIL | Université de Lausanne

82

82

## Structure of pMHC Class I - Peptide binding to MHC groove

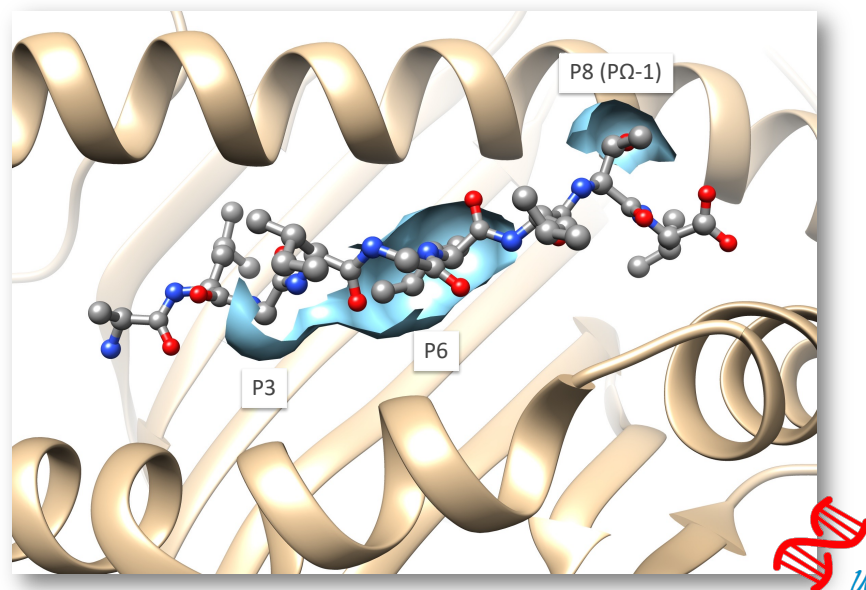
Anchor residues make strong interactions with the MHC binding pockets



83

## Structure of pMHC Class I - Peptide binding to MHC groove

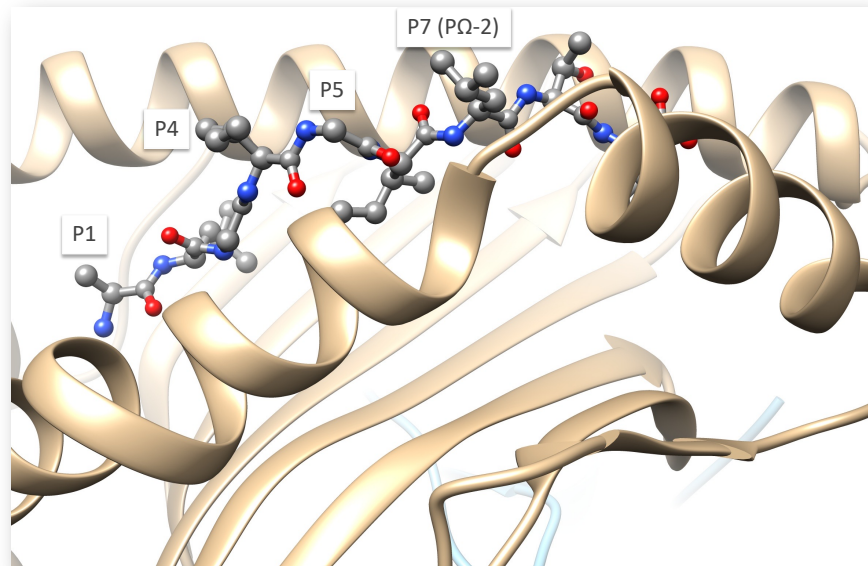
Secondary anchor residues also make interactions with the MHC binding pockets, but are more variable



84

## Structure of pMHC Class I - Peptide binding to MHC groove

Other positions are facing the solvent (or the TCR) and are variable



Unil

UNIL | Université de Lausanne

85

85

## Structure of pMHC Class I - Peptide binding to MHC groove

### Exercise 4 of the booklet

- Can you explain the difference in binding affinities of these two peptides for HLA-A\*0201?

- AAGIGILTV : 60  $\mu$ M. PDB ID.: 2GUO
- ALGIGILTV : 1.5  $\mu$ M. PDB ID.: 1JHT

Unil

UNIL | Université de Lausanne

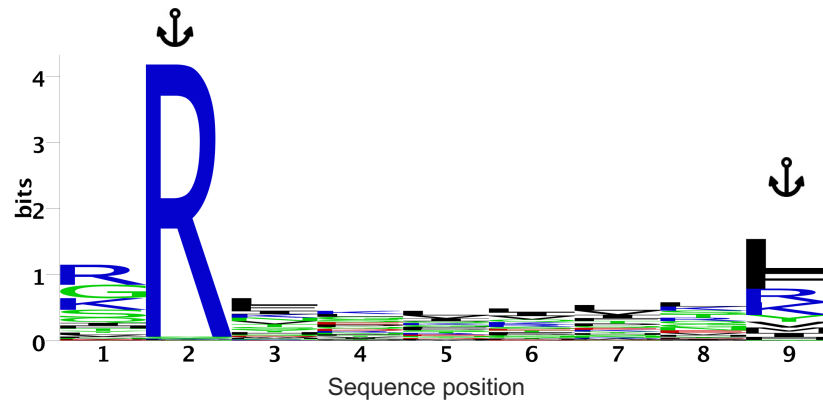
86

86

## Structure of pMHC Class I - Peptide binding to MHC groove

### Exercise 5 of the booklet

Sequence logo for **HLA-B\*2705**:



Bassani-Sternberg, [...] Gfeller. (2017). *PLoS Computational Biology*, 13(8), e1005725.

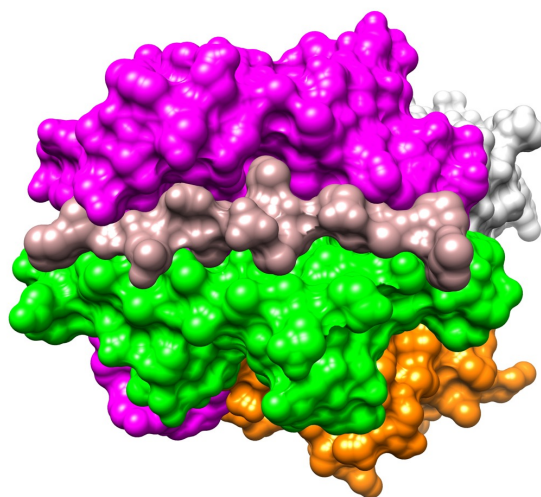
- What are the preferred amino-acids for the two anchor residues?
- Is it different from that of HLA-A\*0201?
- Using the PDB and UCSF Chimera, can you explain this?

Unil  
UNIL | Université de Lausanne

87

87

## 3D structure of peptides bound to the Major Histocompatibility Complex Class II



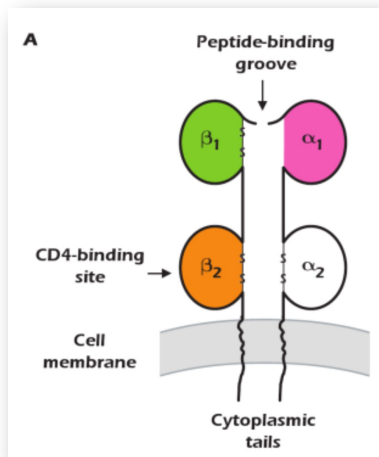
Unil  
UNIL | Université de Lausanne

88

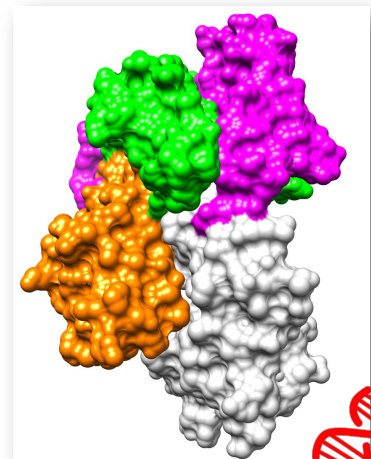
88

## Structure of pMHC Class II

- Transmembrane glycoprotein
- 2 chains  $\alpha$  (35 kDa) and  $\beta$  (28 kDa), forming 4 domains,  $\alpha_1$ ,  $\alpha_2$ ,  $\beta_1$  and  $\beta_2$ .
- Deep groove to bind a peptide (12-18 residues) between  $\alpha_1$  and  $\beta_1$
- Expressed at cell surface



From "Immunology: A Short Course"  
by Richard Colco, Geoffrey Sunshine

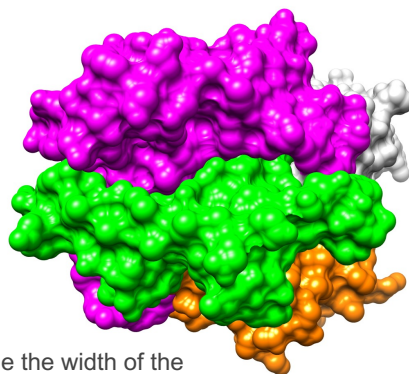


Unil  
UNIL | Université de Lausanne

89

89

## Structure of pMHC Class II – The binding groove

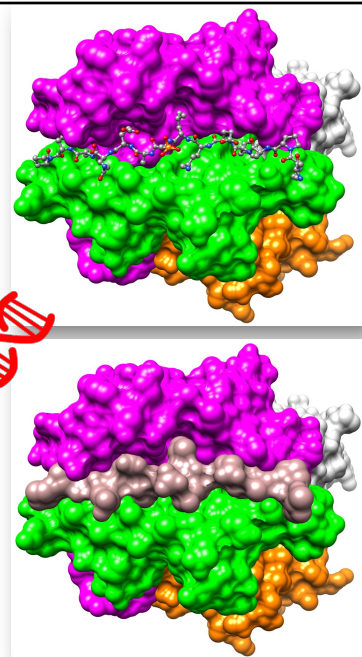


No walls to define the width of the pocket



Can adapt very large peptides, 12 – 18 residues

Structure of MHC class II molecule HLA-DR1 in complex with phosphopeptide MART-1 (15 residues)



Unil  
UNIL | Université de Lausanne

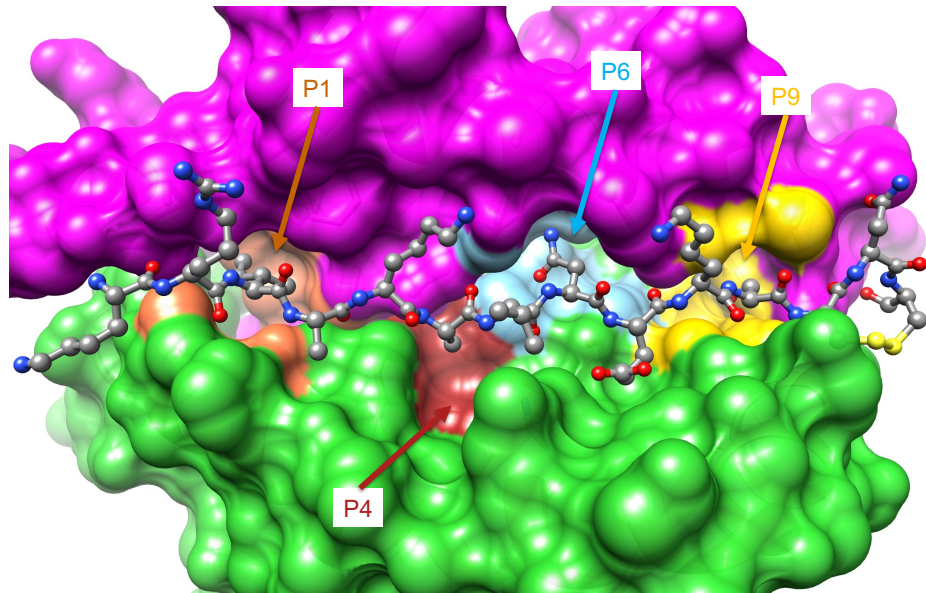
90

90



## Structure of pMHC Class II – The binding pockets

3 to 4 major anchoring pockets to accommodate the primary peptide anchor residues



91

## Structure of pMHC Class II – The binding pocket

A large number of peptide residues interacting with MHC-II

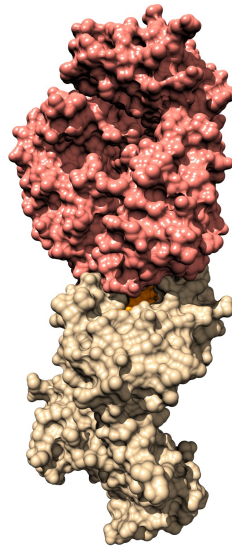
### Exercise 6 of the booklet

- Load structures 3L6F, 1FYT and 1AQD in UCSF Chimera
- Only keep chains A, B and C
- What is the common HLA protein?
- Show chains A and B as ribbons (hiding residues) and chains C as ball and stick (hiding the ribbon)
- The peptides are composed of how many residues ?
- What can we say about the position of the backbone of the peptides?
- Use the "Tools/Structure Comparison/Match->Align" tool to align the peptide sequences (chains C, only). Use the "Headers" menu to remove "RMSD" from the header, and add "Conservation", "Consensus" and "Charge variation".
- What are the most conserved positions?
- Are they facing HLA or the solvent?
- What are the residues defining the pocket in which they are bound?
- What type of interactions are taking place?

92



## 3D structure of the complex between the T-cell receptor (TCR) and pMHC



Unil

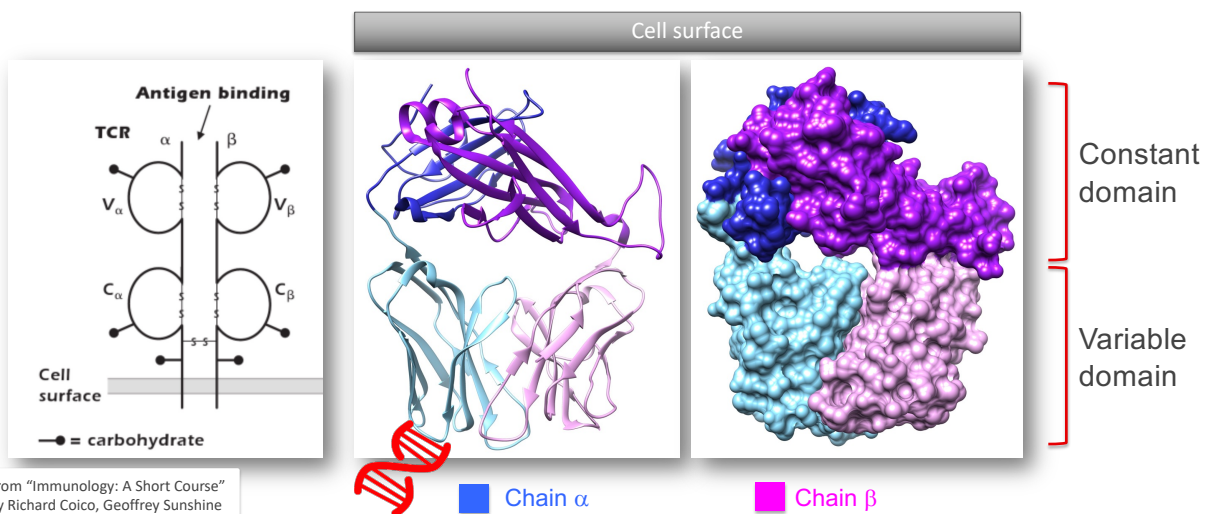
UNIL | Université de Lausanne

93

93

## TCR 3D structure

- Membrane-anchored heterodimeric protein
- 2 chains  $\alpha$  and  $\beta$
- Each chain is composed of 2 extracellular domains: a variable domain V, and a constant domain C



Unil

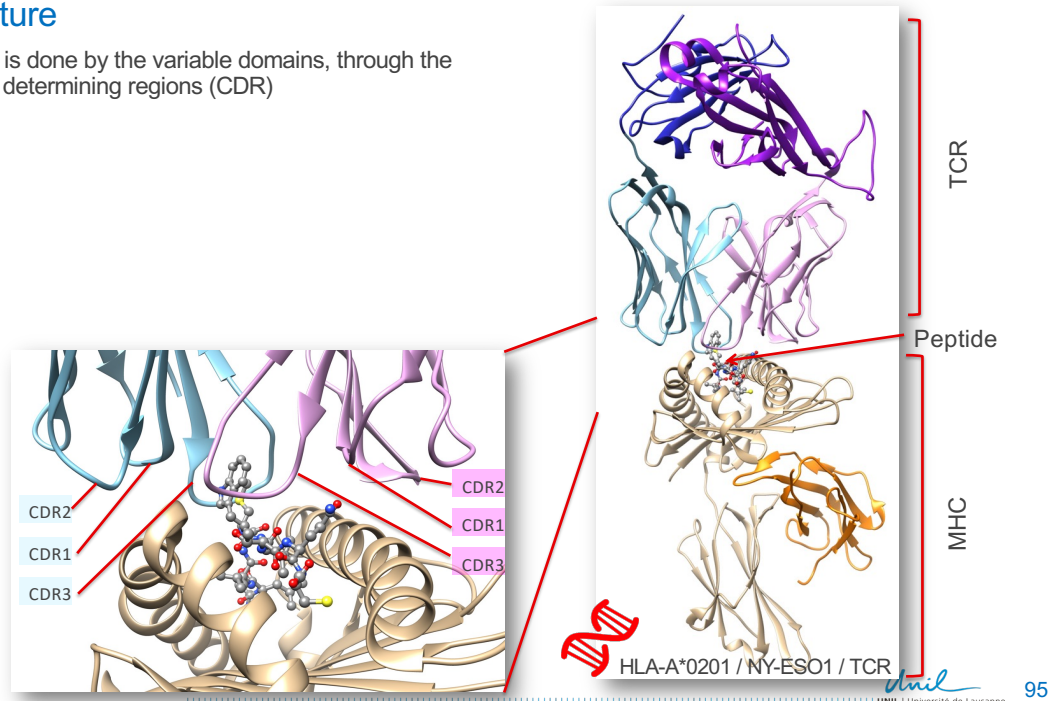
UNIL | Université de Lausanne

94

94

## TCR 3D structure

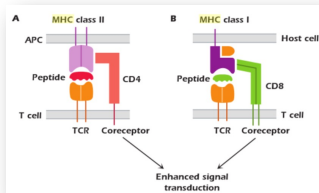
- Antigen binding is done by the variable domains, through the complementary determining regions (CDR)



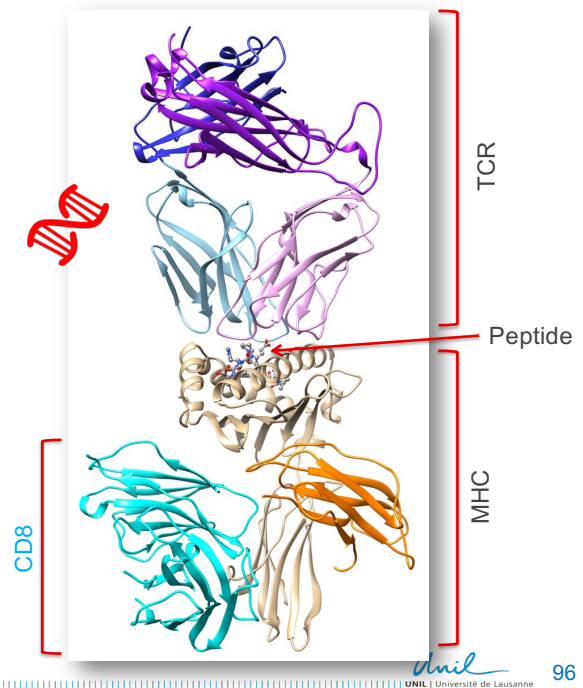
95

## TCR 3D structure

- Antigen binding is done by the variable domains, through the complementary determining regions (CDR)



From "Immunology: A Short Course" by Richard Coico, Geoffrey Sunshine



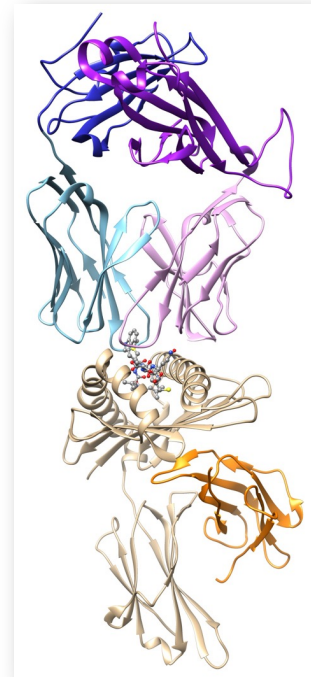
96

## TCR 3D structure

### Exercise 7 of the booklet

- Load the 2BNR PDB structure into UCSF chimera
- Produce a molecular representation similar to the one you see on the right
 

|  |  |
|--|--|
| Residues in the CDR for TCR $\alpha$ : | <ul style="list-style-type: none"> <li>• CDR1: 28 - 32</li> <li>• CDR2: 51 - 55</li> <li>• CDR3: 94 - 101</li> </ul> |
| Residues in the CDR for TCR $\beta$ :  | <ul style="list-style-type: none"> <li>• CDR1: 25 - 29</li> <li>• CDR2: 49 - 53</li> <li>• CDR3: 94 - 100</li> </ul> |
- Display the atoms of the CDRs for TCR $\alpha$  and TCR $\beta$
- What are the CDRs that make most of the contacts with the peptide epitope? And with the MHC?
- What are the TCR residues that make contact with peptide Trp5? What types of interactions are taking place?
- What MHC residues are close to TCR $\beta$  Ala51?
- If you wanted to change residue 51 of TCR $\beta$ , to increase the affinity of TCR for pMHC, what mutation would you introduce? Why?

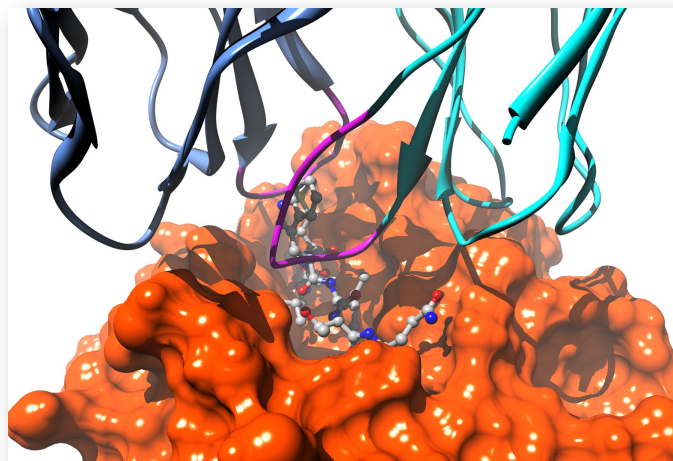


Unil  
UNIL | Université de Lausanne

97

97

## TCR engineering – an example

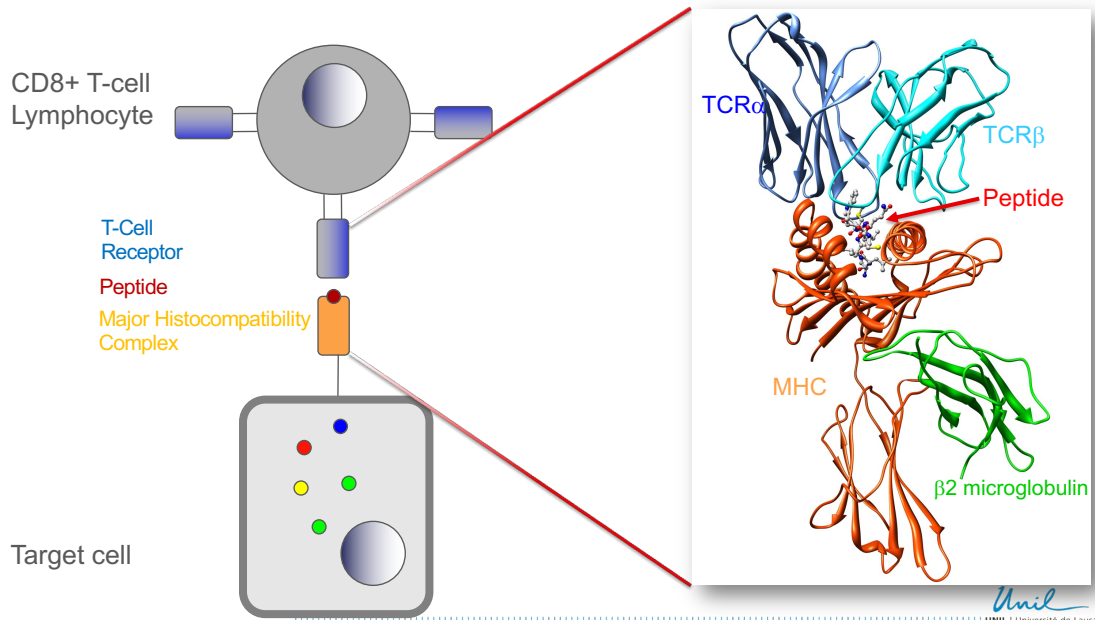


Unil  
UNIL | Université de Lausanne

98

98

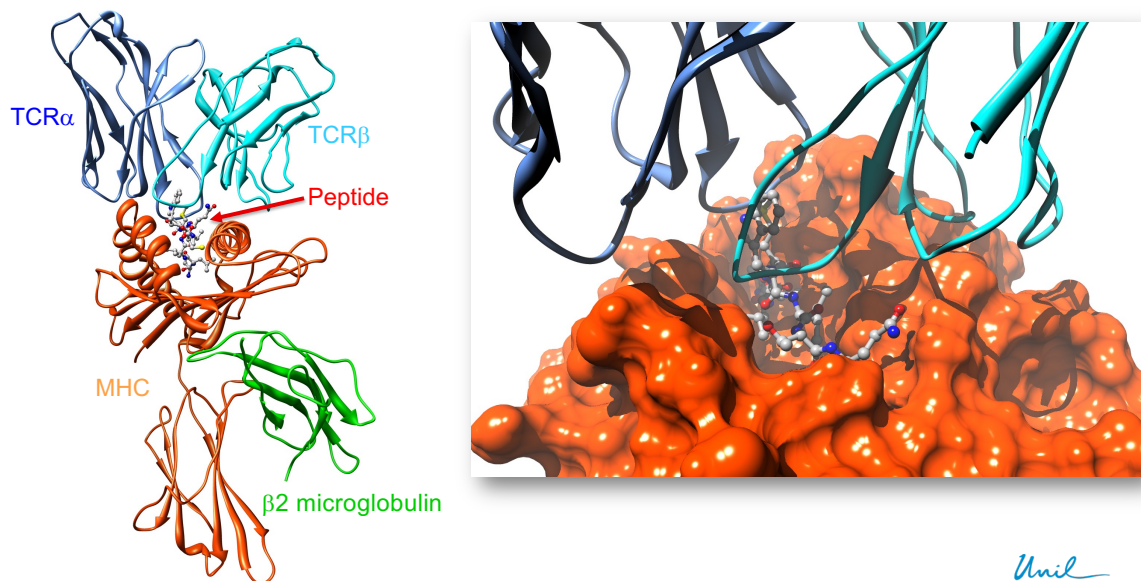
## TCR engineering – an example



99

## TCR engineering – 3D structure

TCR recognizes pMHC through complementary determining regions (CDR) loops

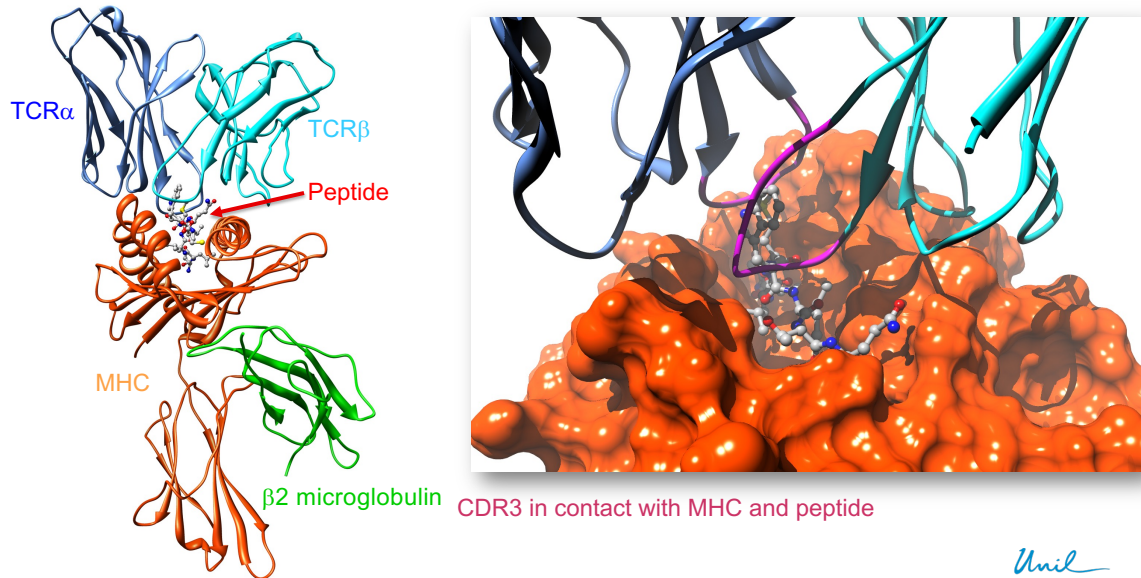


100



## TCR engineering – 3D structure

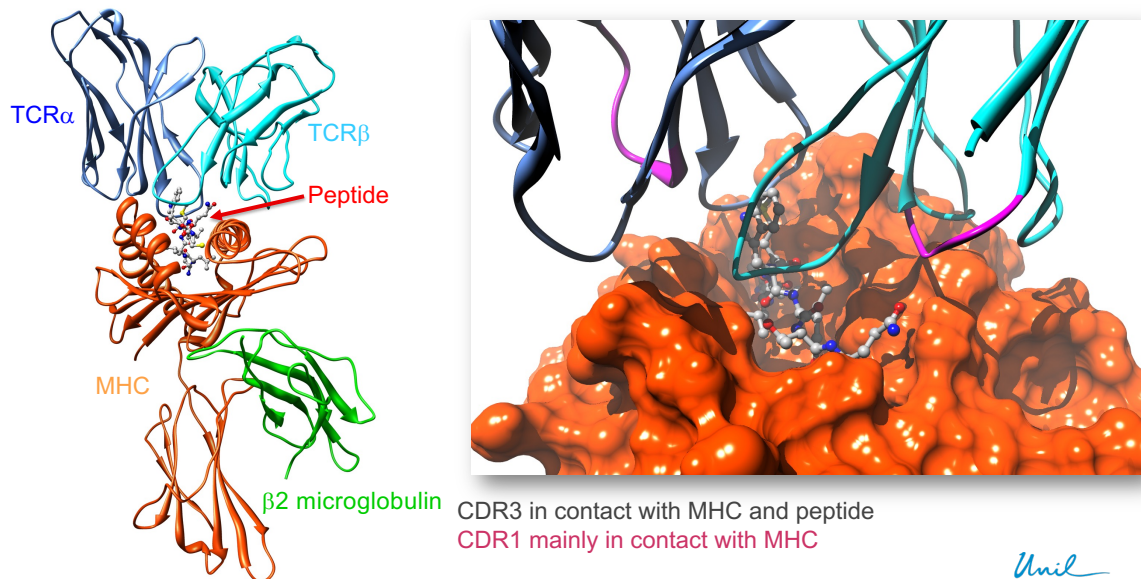
TCR recognizes pMHC through complementary determining regions (CDR) loops



101

## TCR engineering – 3D structure

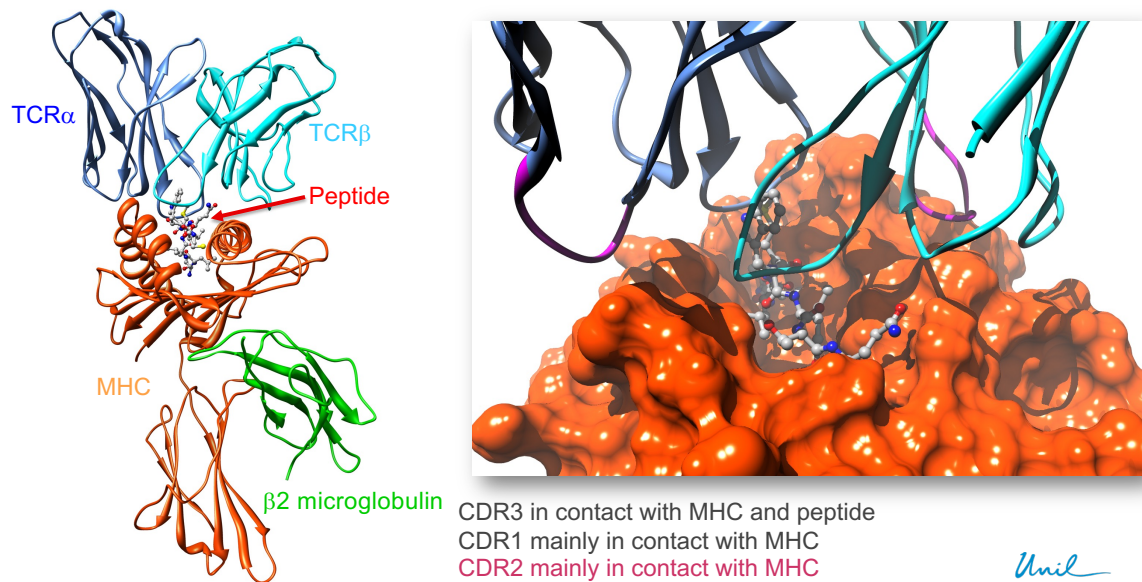
TCR recognizes pMHC through complementary determining regions (CDR) loops



102

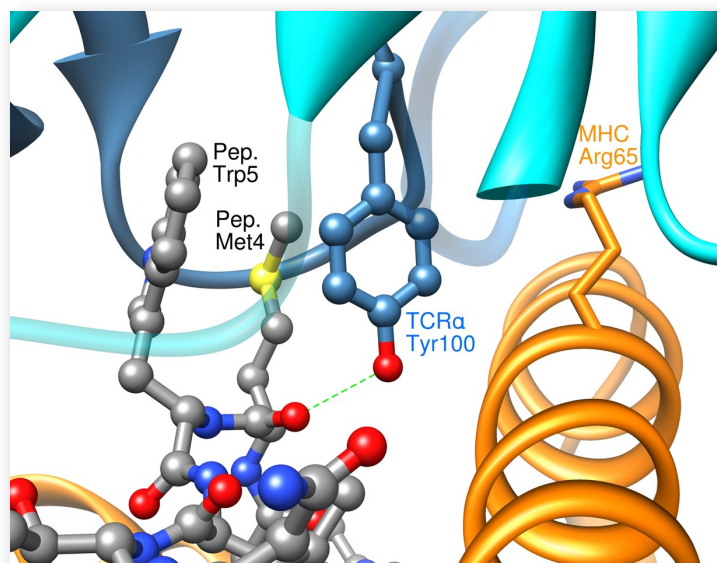
## TCR engineering – 3D structure

TCR recognizes pMHC through complementary determining regions (CDR) loops



103

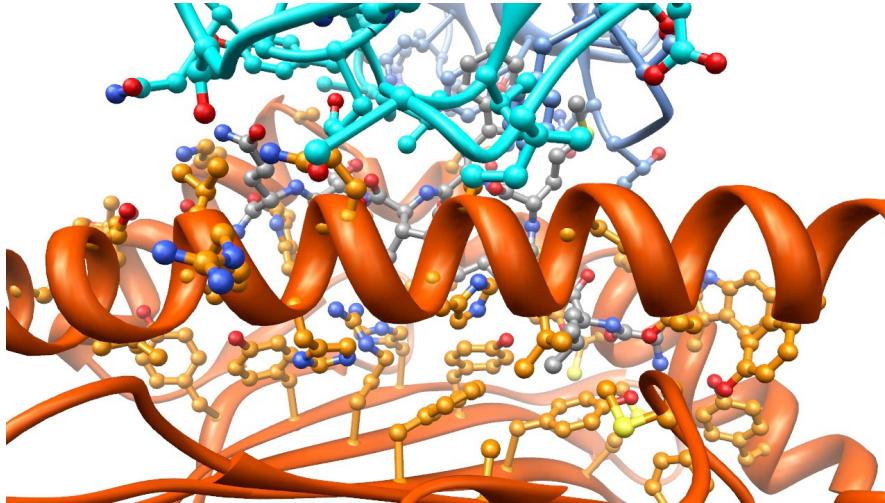
## TCR engineering – Binding free energy



104



## TCR engineering – Binding free energy



Visually: “important” interactions everywhere



Need for a physics-based method to estimate quantitatively the importance of each residue/interaction

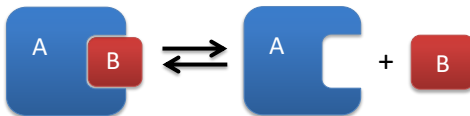
Unil

105

105

## TCR engineering – Binding free energy, $\Delta G_{\text{bind}}$

Link between experiment and modeling

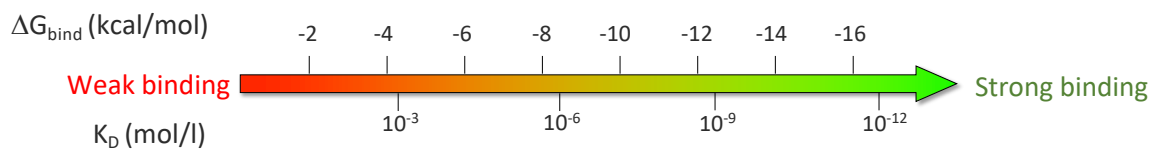


$$K_D = \frac{[A][B]}{[AB]}$$

$K_D$  : dissociation constant

Accessible by  
computer-  
aided methods

$$\Delta G_{\text{bind}} = RT \ln(K_D) = \Delta H - T\Delta S$$



Unil

106

106

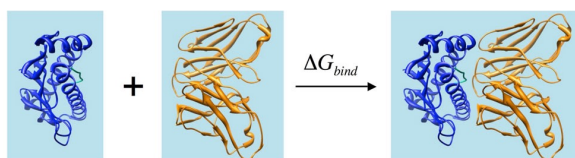
## TCR engineering – Binding free energy, $\Delta G_{\text{bind}}$

In silico methods to estimate  $\Delta G_{\text{bind}}$

**MM-GBSA:**

$$\langle \Delta G_{\text{bind}} \rangle = \langle \Delta H_{\text{bind}}^{\text{gas}} \rangle + \langle \Delta G_{\text{desolv}} \rangle - T \langle \Delta S \rangle$$

Provides contribution of **each atom** to the association strength



Zoete, V., Meuwly, M., & Karplus, M. *Proteins*, **2005**, 61, 79–93.

Zoete, V.\*, Meuwly, M.\* *J. Comput. Chem.*, **2006**, 27, 1843–1857.

MM-GBSA: Molecular Mechanics Generalized Born Surface Area

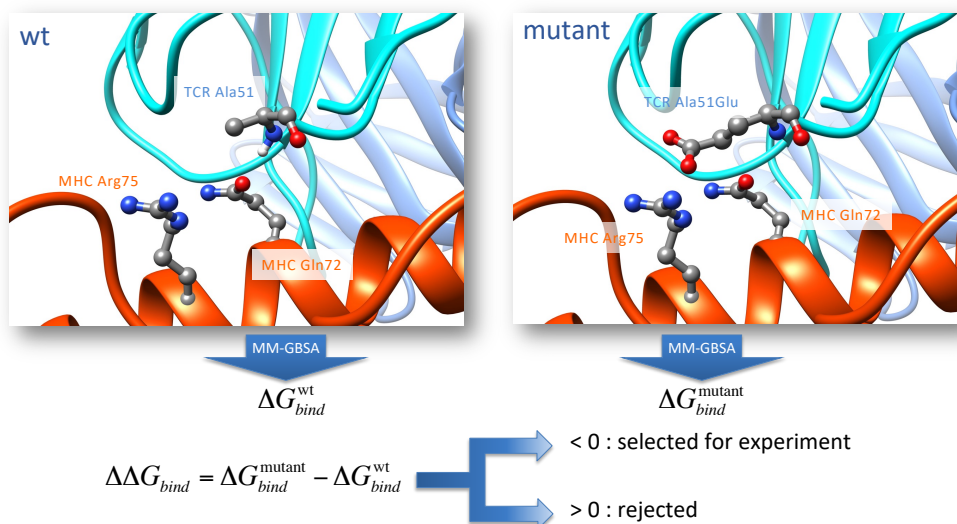
Unil  
UNIL | Université de Lausanne

107

107

## TCR engineering – Binding free energy, $\Delta G_{\text{bind}}$

Using  $\Delta G_{\text{bind}}$  to select mutations for experimental assay



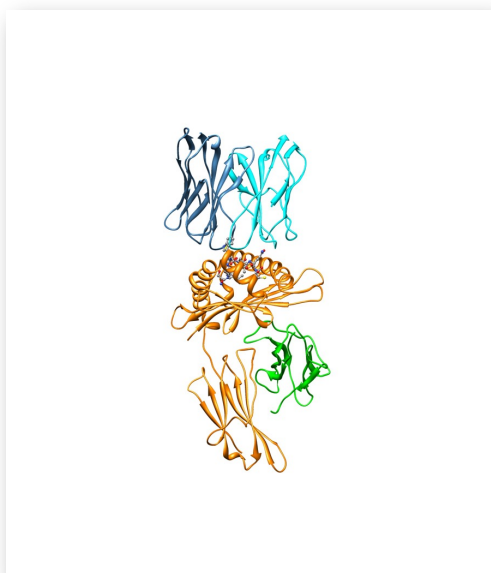
Zoete, V., Irving, M. B., & Michielin, O. MM-GBSA binding free energy decomposition and T cell receptor engineering. *J. Molec. Rec.*, **2010**, 23, 142–152.

Unil  
UNIL | Université de Lausanne

108

108

## TCR engineering – The computational pipeline



### Molecular Dynamics (MD) simulation:

System solvated with explicit water molecules (TIP3P model):

- ~ 29,500 water molecules
- ~ 100,000 atoms in total

Atom motions are calculated to follow Newton's equation of motion, at 300 K and 1 atm.

Typical simulation times: from 0.5 ns to ~ 100 ns

Energy terms averaged over 200 to 500 frames extracted from the MD simulation

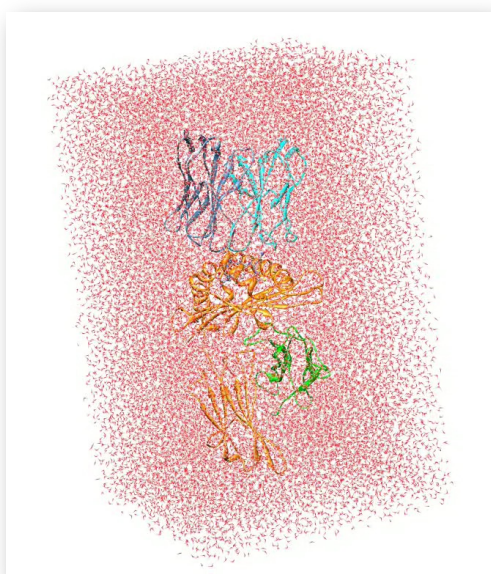
Unil

UNIL | Université de Lausanne

109

109

## TCR engineering – The computational pipeline



### Molecular Dynamics (MD) simulation:

System solvated with explicit water molecules (TIP3P model):

- ~ 29,500 water molecules
- ~ 100,000 atoms in total

Atom motions are calculated to follow Newton's equation of motion, at 300 K and 1 atm.

Typical simulation times: from 0.5 ns to ~ 100 ns

Energy terms averaged over 200 to 500 frames extracted from the MD simulation

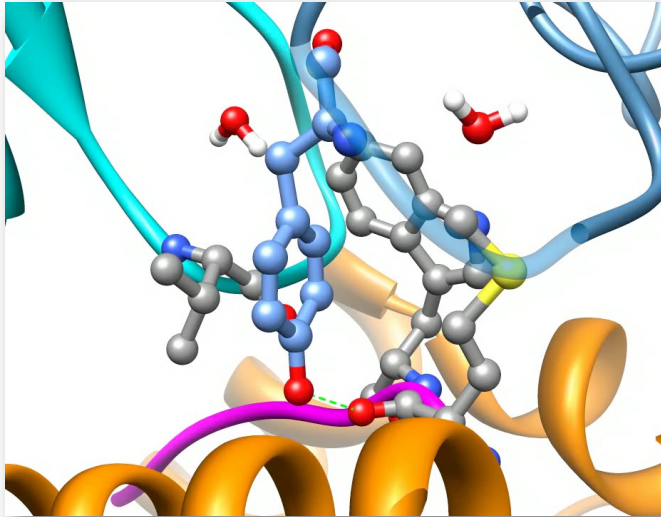
Unil

UNIL | Université de Lausanne

110

110

## TCR engineering – The computational pipeline



### Molecular Dynamics (MD) simulation:

System solvated with explicit water molecules (TIP3P model):

- ~ 29,500 water molecules
- ~ 100,000 atoms in total

Atom motions are calculated to follow Newton's equation of motion, at 300 K and 1 atm.

Typical simulation times: from 0.5 ns to ~ 100 ns

Energy terms averaged over 200 to 500 frames extracted from the MD simulation

Unil

111

111

## TCR engineering – The computational pipeline

3D structure of the wild-type TCR-pMHC complex

PBC MD simulation  
MM-GBSA

$\Delta G_{bind}^{res}$   $\Delta G_{bind}^{res,bb}$   $\Delta G_{bind}^{res,sc}$   
& structural data

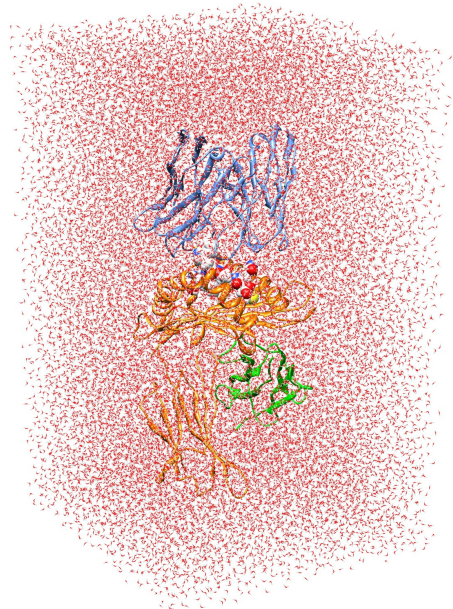
Rotamer library

3D structural models of possible TCR mutations

SBC MD simulation  
MM-GBSA

$\Delta\Delta G_{bind}$  for TCR mutations

Mutations selected for  
expression, purification and  
experimental testing



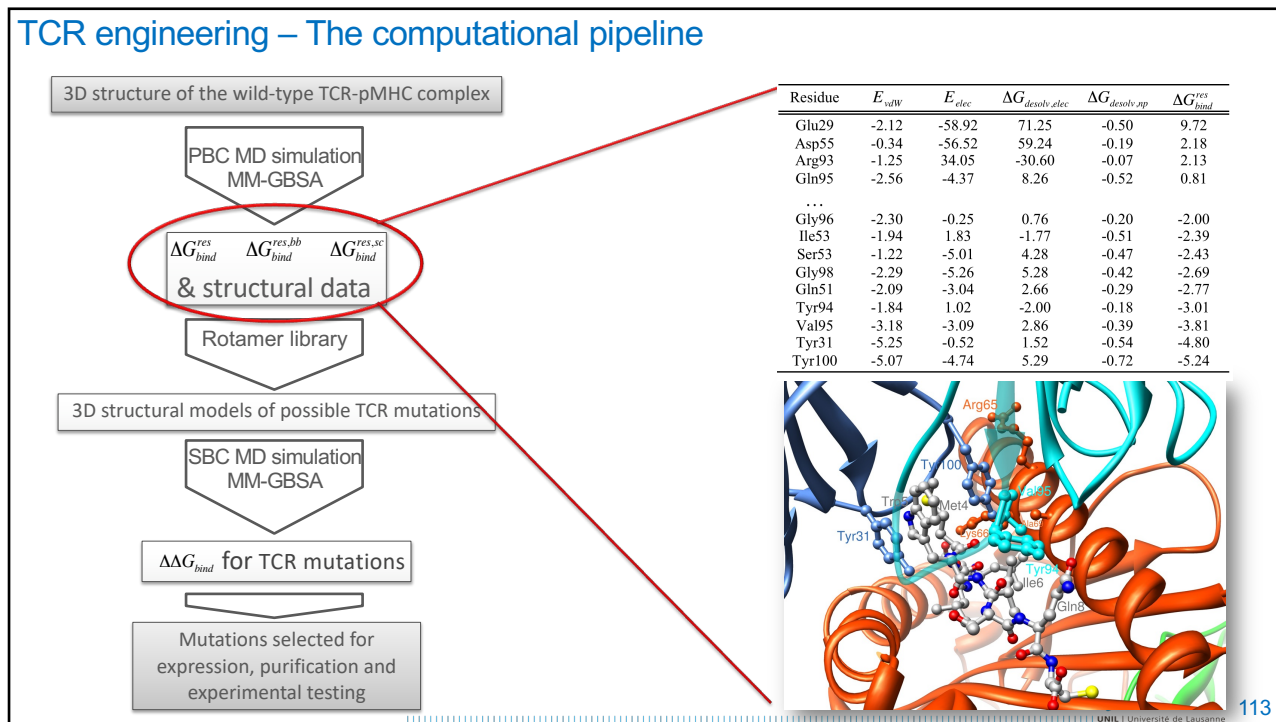
Unil

112

112

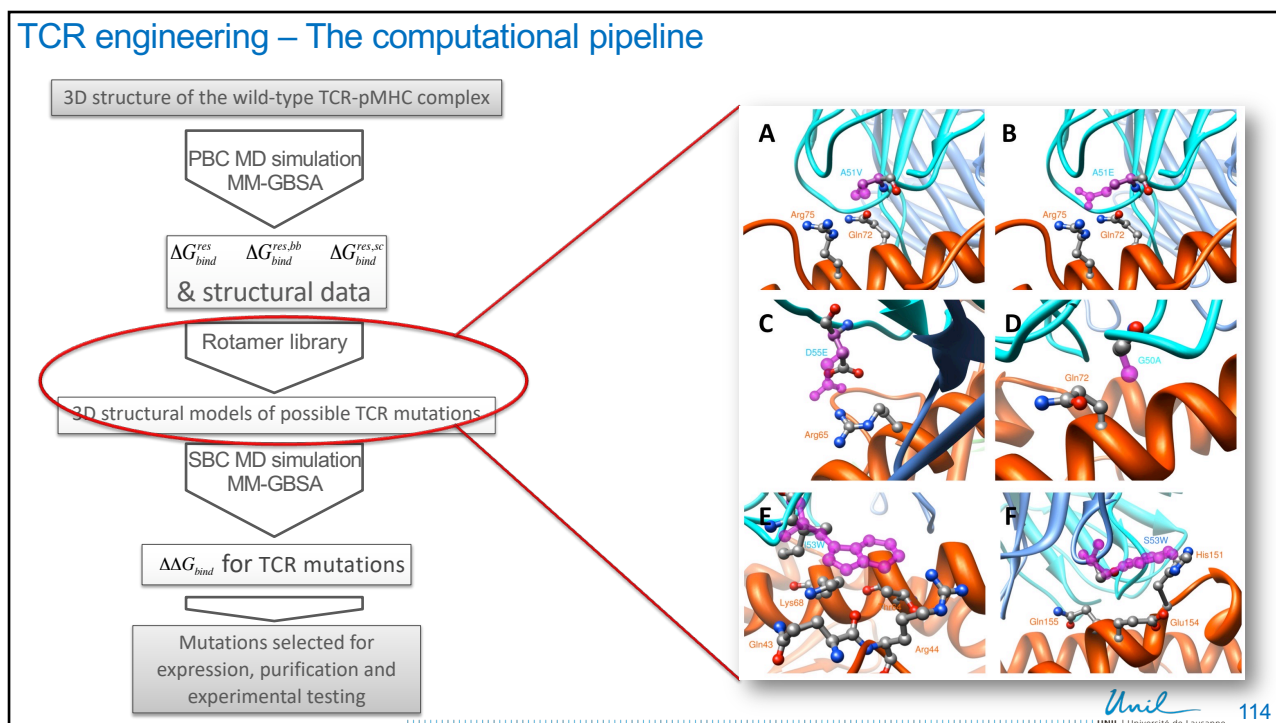


## TCR engineering – The computational pipeline



113

## TCR engineering – The computational pipeline

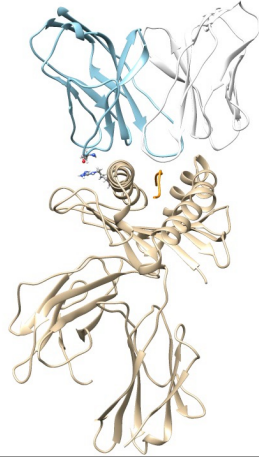


114

## TCR engineering – The computational pipeline

### Increasing affinity

A51E



Gain in binding free energy:  
-7.3 kcal/mol

Unil

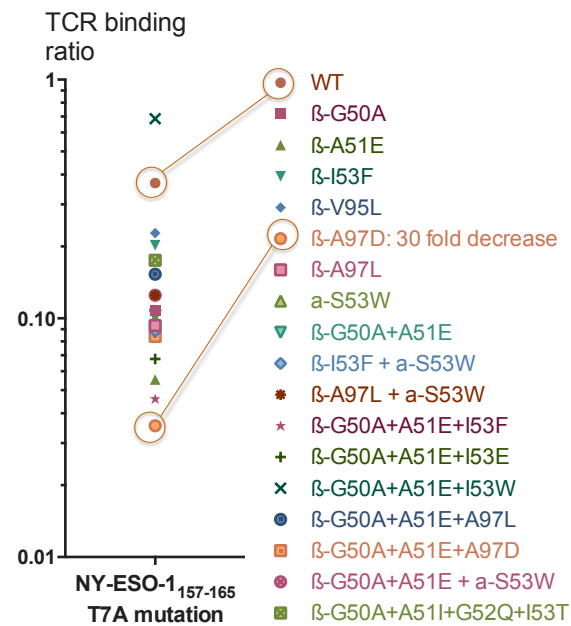
115

115

## TCR engineering – The computational pipeline

### Increasing selectivity

A97D



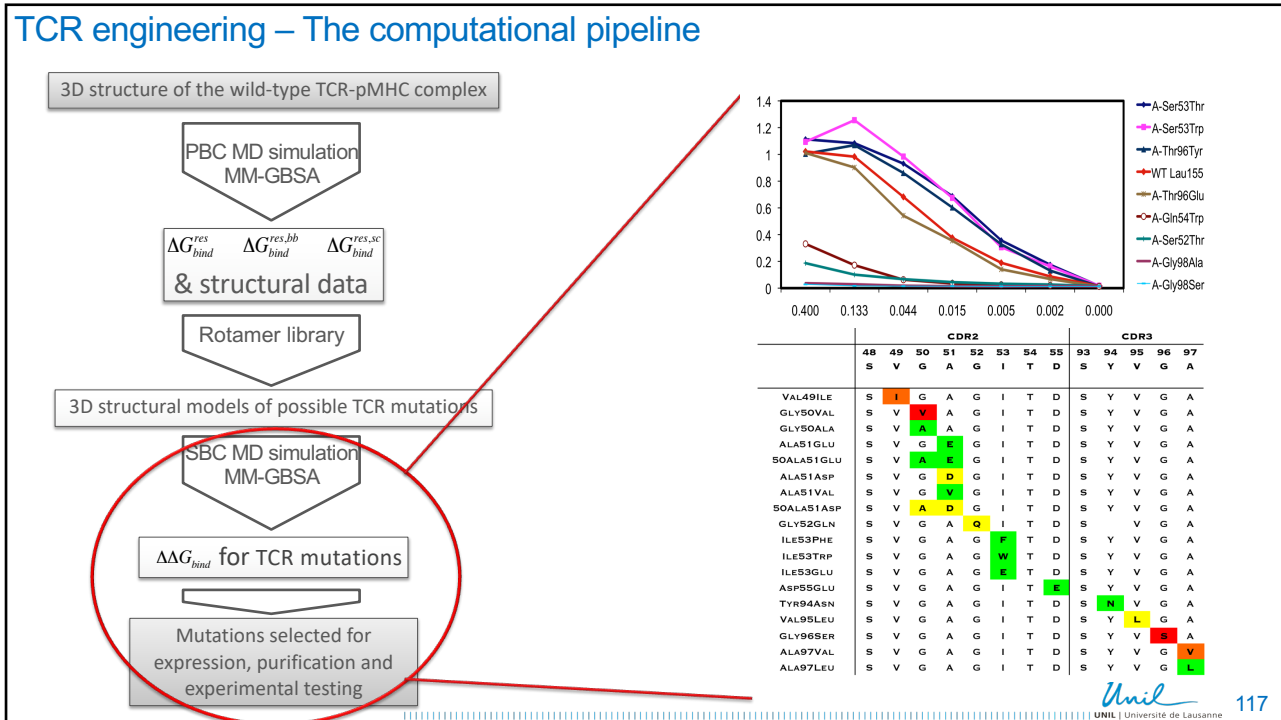
Unil

116

116



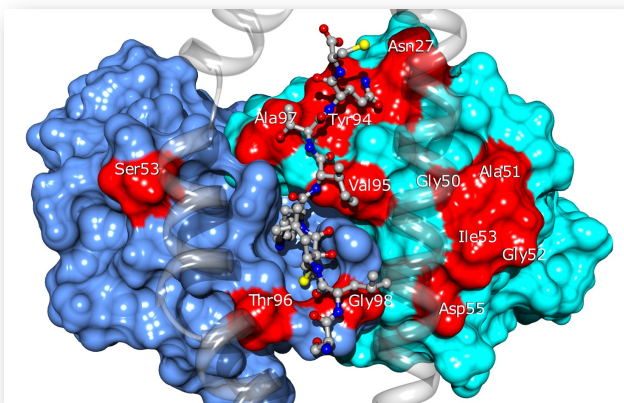
## TCR engineering – The computational pipeline



117

## TCR engineering – Application to BC1 TCR targeting NY-ESO-1

- 24 single/double mutants tested (M. Irving)
- 13 (54 %) were more active than the wt TCR
- up to 56 fold increase for single mutations
- 150 fold increase for TCR V $\beta$  G50A/A51E/A97L + V $\alpha$  S53W



Unil  
UNIL | Université de Lausanne

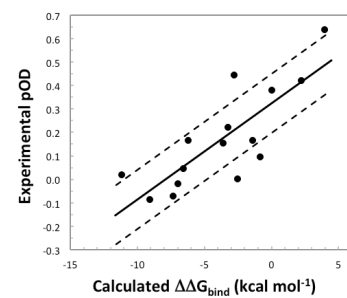
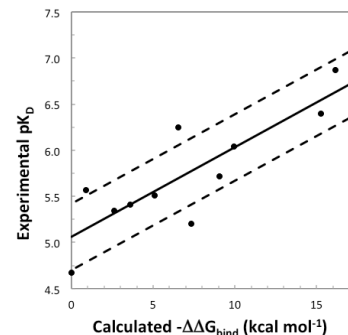
118

## TCR engineering – Application to BC1 TCR targeting NY-ESO-1

- 24 single/double mutants tested (M. Irving)
- 13 (54 %) were more active than the wt TCR
- up to 56 fold increase for single mutations
- 150 fold increase for TCR V $\beta$  G50A/A51E/A97L + V $\alpha$  S53W
- good correlation between calculated binding free energies and experimental results

$$pOD = 0.32 (\pm 0.04) + 0.041 (\pm 0.008) \Delta\Delta G_{bind} \quad R=0.81$$

$$pK_D = 5.05 (\pm 0.18) - 0.098 (\pm 0.021) \Delta\Delta G_{bind} \quad R=0.82$$



Anil

UNIL | Université de Lausanne

119

119

## TCR engineering – Application to BC1 TCR targeting NY-ESO-1

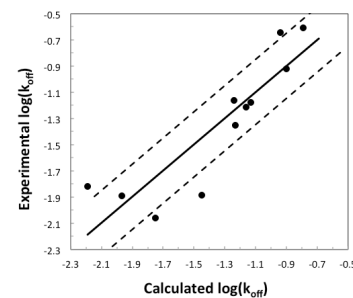
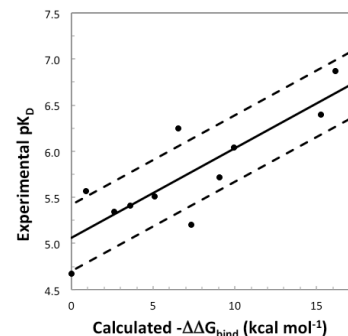
- 24 single/double mutants tested (M. Irving)
- 13 (54 %) were more active than the wt TCR
- up to 56 fold increase for single mutations
- 150 fold increase for TCR V $\beta$  G50A/A51E/A97L + V $\alpha$  S53W
- good correlation between calculated binding free energies and experimental results

$$pOD = 0.32 (\pm 0.04) + 0.041 (\pm 0.008) \Delta\Delta G_{bind} \quad R=0.81$$

$$pK_D = 5.05 (\pm 0.18) - 0.098 (\pm 0.021) \Delta\Delta G_{bind} \quad R=0.82$$

- good correlation between calculated energies and experimental  $k_{off}$  ( $R=0.88$ )

$$\log(k_{off}) = -0.94 (\pm 0.13) + 0.13 (\pm 0.02) \Delta\Delta G_{non-polar} + 0.044 (\pm 0.016) \Delta\Delta G_{polar}$$



Anil

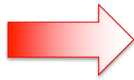
UNIL | Université de Lausanne

120

120

## TCR engineering – Application to BC1 TCR targeting NY-ESO-1

- unfitted approach: can be applied to other systems



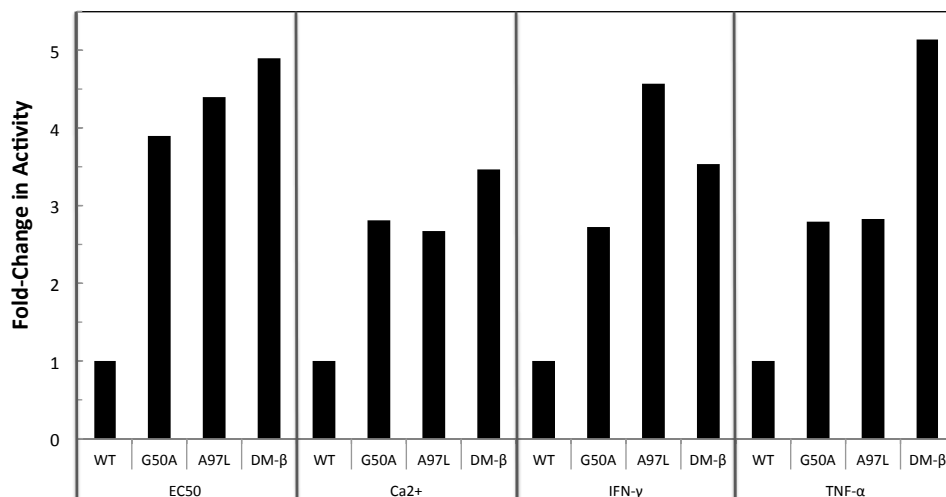
Applied to Melan-A antigen with 73% success rate

- step-by-step modeling approach: incremental improvements in TCR affinity, while minimizing the loss of specificity
- mutations toward both non polar and polar mutants
- up to 56 fold affinity increase for single mutants. Some single mutations are compatible. Combinations of them led to 150 fold increase in affinity compared to WT
- correlation between experimental affinity and calculated binding free energy
- correlation between  $k_{off}$  and calculated energy terms
- **no cross reactivity**

121

## TCR engineering – Application to BC1 TCR targeting NY-ESO-1

- both T-cell proliferation after antigenic challenge and tumor cell killing were significantly improved



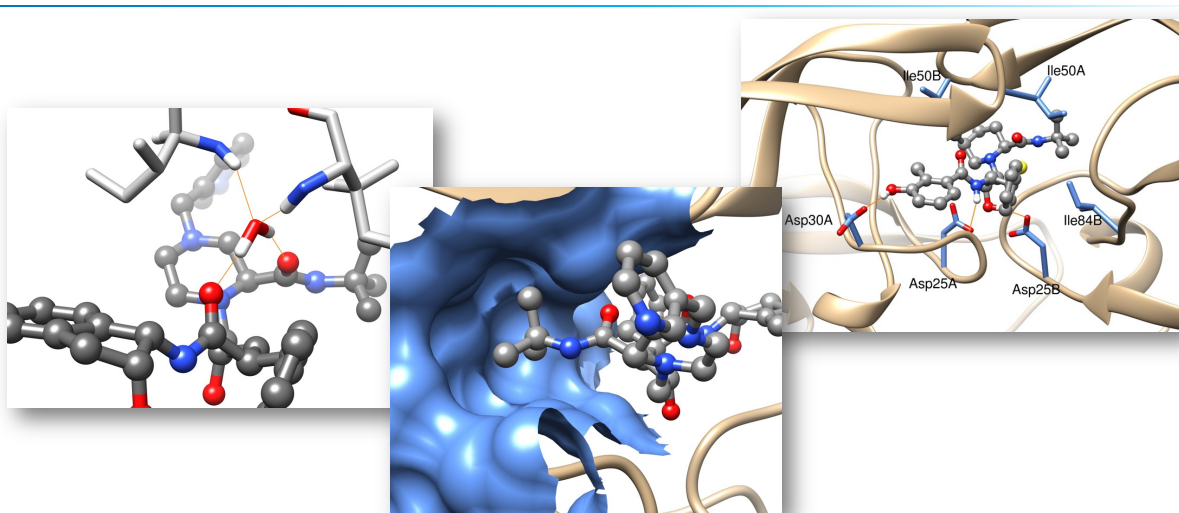
Mouse model / Clinical trial at CHUV

122

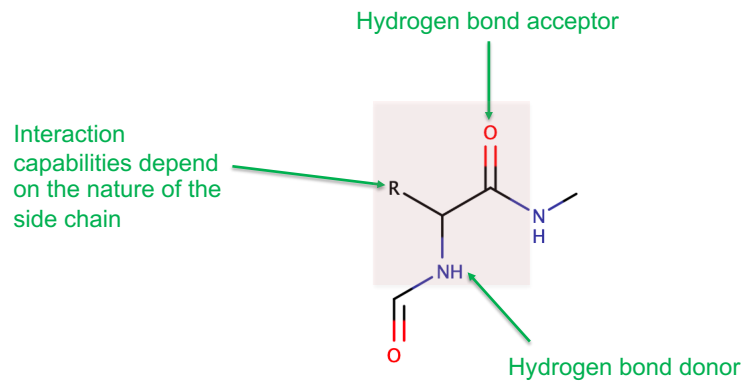
## The End!

If you have questions: [vincent.zoete@unil.ch](mailto:vincent.zoete@unil.ch)

## Molecular Recognition



## Molecular recognition – Possible interactions per amino acids

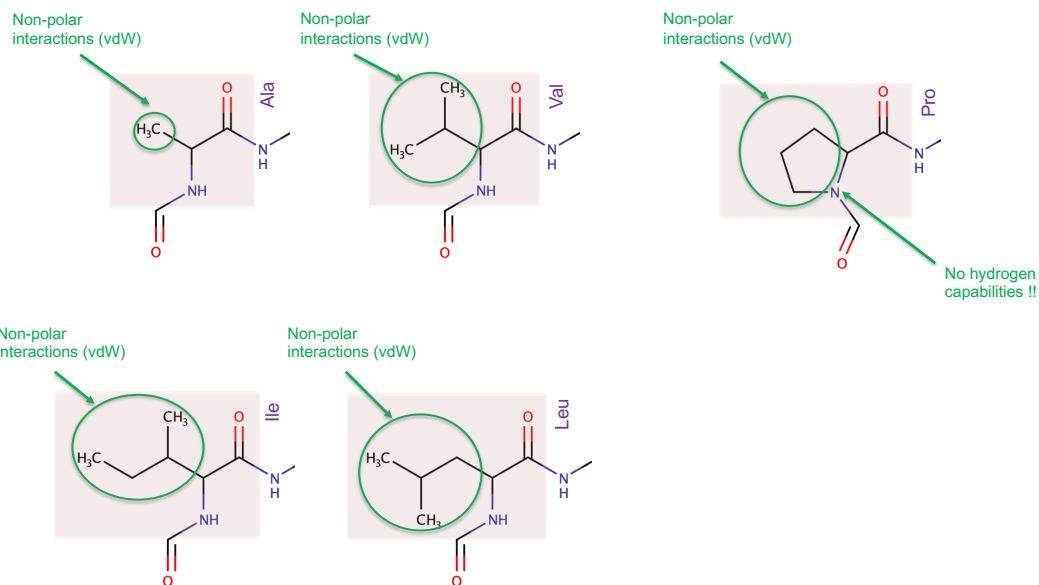


Unil  
UNIL | Université de Lausanne

12

125

## Molecular recognition – Possible interactions per amino acids

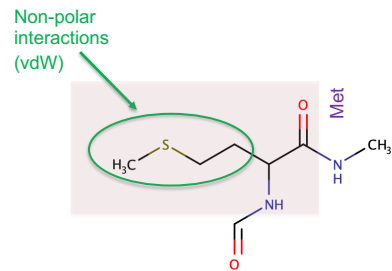
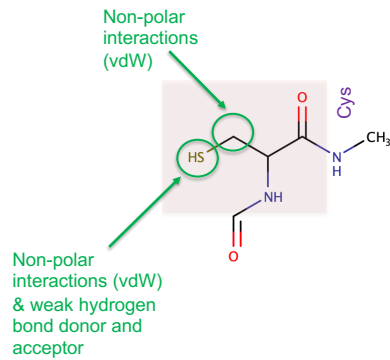


Unil  
UNIL | Université de Lausanne

12

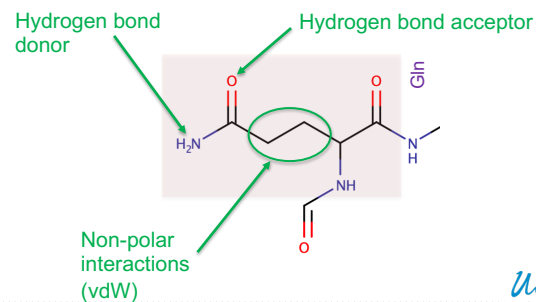
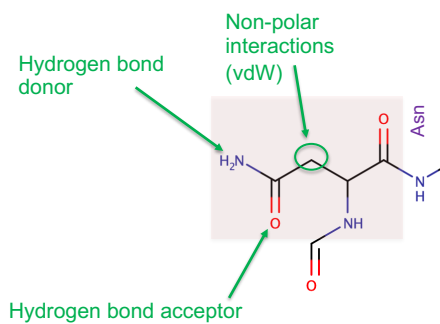
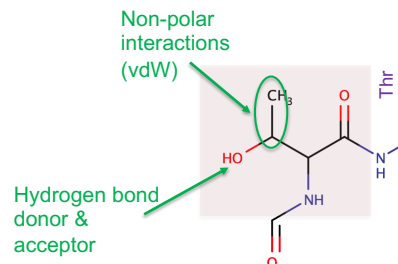
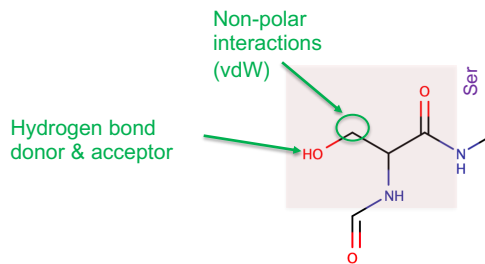
126

## Molecular recognition – Possible interactions per amino acids



127

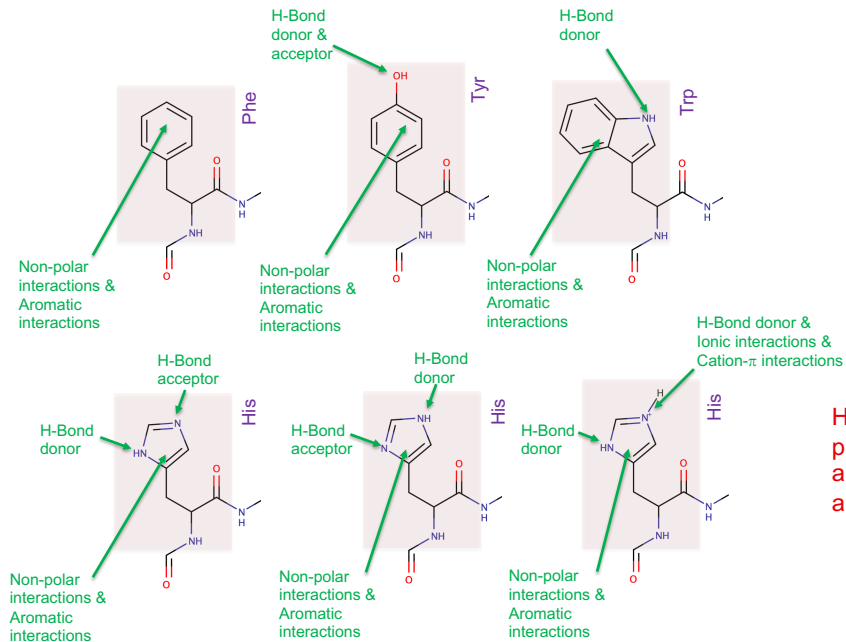
## Molecular recognition – Possible interactions per amino acids



128



## Molecular recognition – Possible interactions per amino acids

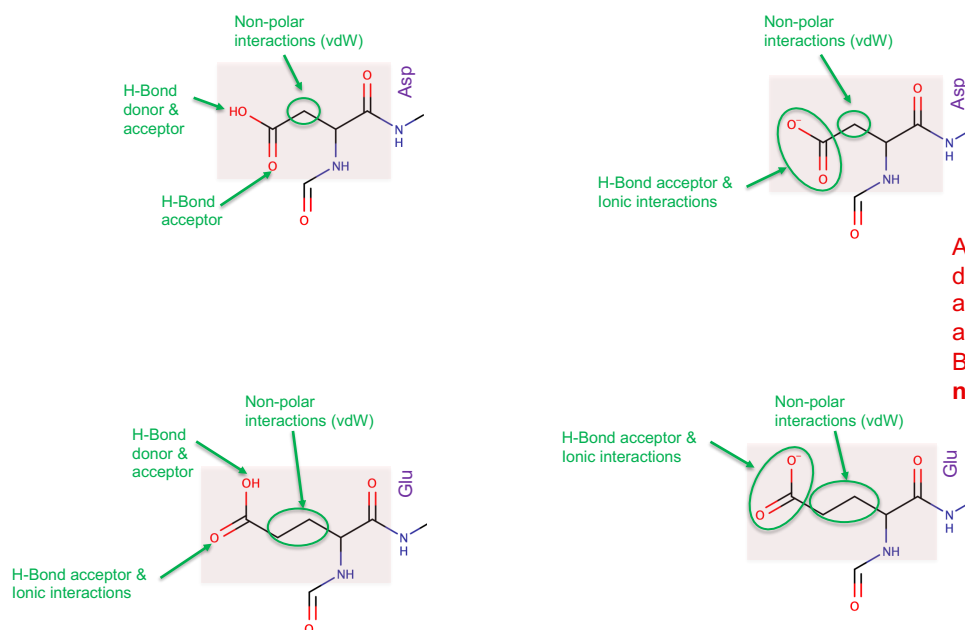


Unil

12

129

## Molecular recognition – Possible interactions per amino acids

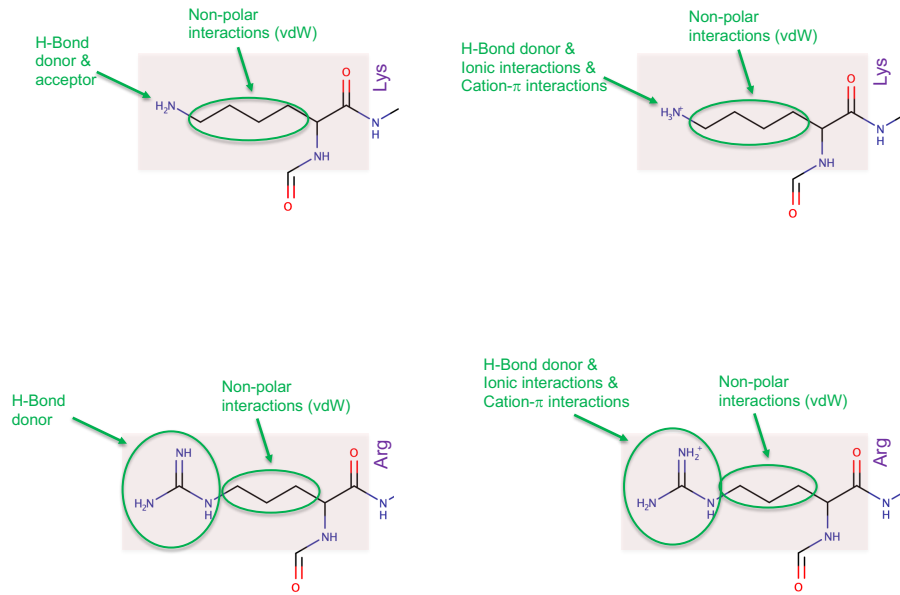


Unil

13

130

## Molecular recognition – Possible interactions per amino acids



Arg and Lys exist in 2 different protonation states as a function of the pH and the environment  
But they are **generally positive**