



Central European Institute of Technology  
BRNO | CZECH REPUBLIC

# Biomacromolecular structure analysis

Radka Svobodová  
CEITEC, Masaryk University



EUROPEAN UNION  
EUROPEAN REGIONAL DEVELOPMENT FUND  
INVESTING IN YOUR FUTURE

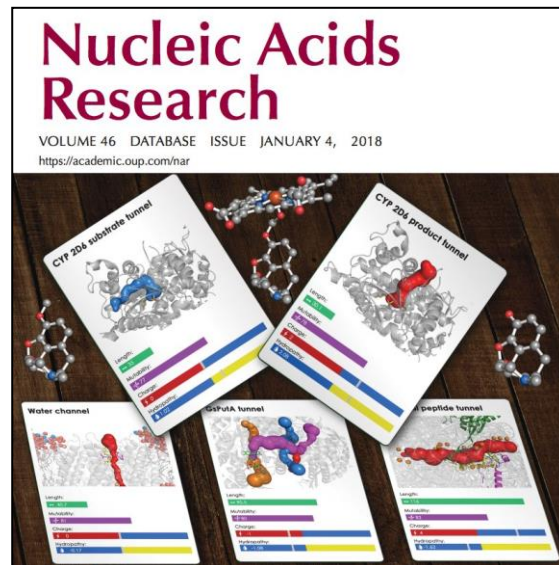


**OP Research and  
Development for Innovation**



# MOLE and ChannelsDB – summary of past

- **MOLE** – software tool for detection and characterization of channels
- **ChannelDB** – interactive database of channels with expected and annotated biological relevancy.
- Both include channel visualisation via LiteMol



Pravda L., Sehnal D., Svobodová R., Navrátilová V., Toušek D., Berka K., ... & Koča J. (2017). *ChannelsDB: database of biomacromolecular tunnels and pores*. **Nucleic acids research**, 46(D1), D399-D405.

# MOLE and ChannelsDB – news

MOLEonline update finished & published

**MOLEonline: Web application for detection of channels and pores and calculation of their geometrical and physico-chemical properties**

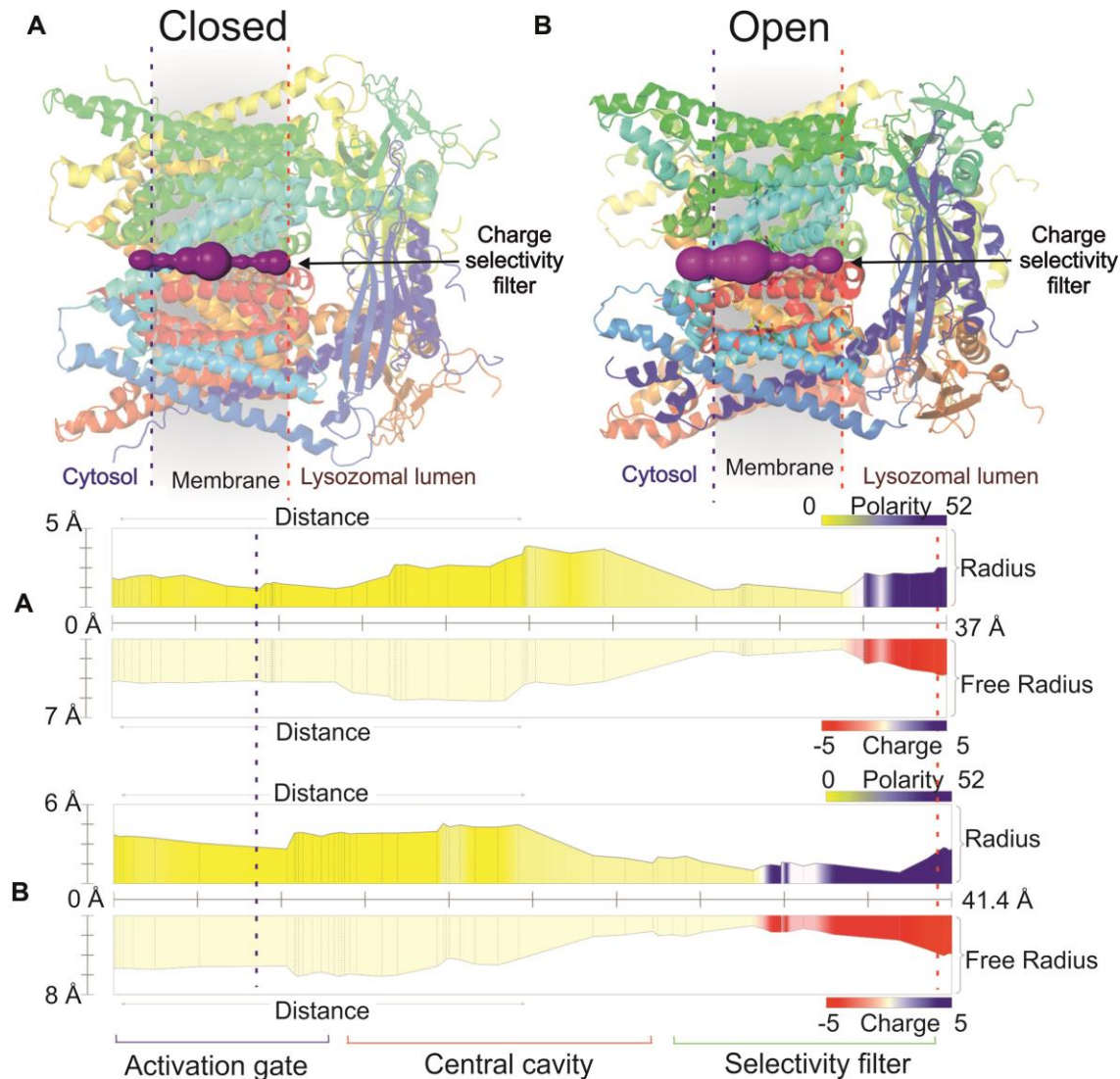
**New features:**

- Automatic detection of transmembrane pores
- Visualization of properties on the channel's profile
- Interconnection with other bioinformatics tools (PDBe, CSA, ChannelsDB, OPM, UniProt) and data transfer from them
- Integration of LiteMol suite

Pravda L., Sehnal D., Toušek D., Navrátilová V., Bazgier V., Berka K., ... & Koča J., Otyepka M. (2018). *MOLEonline: a web-based tool for analyzing channels, tunnels and pores (2018 update)*. **Nucleic acids research**.

# MOLE and ChannelsDB – news

## Use case: Transient receptor potential mucolipin 1



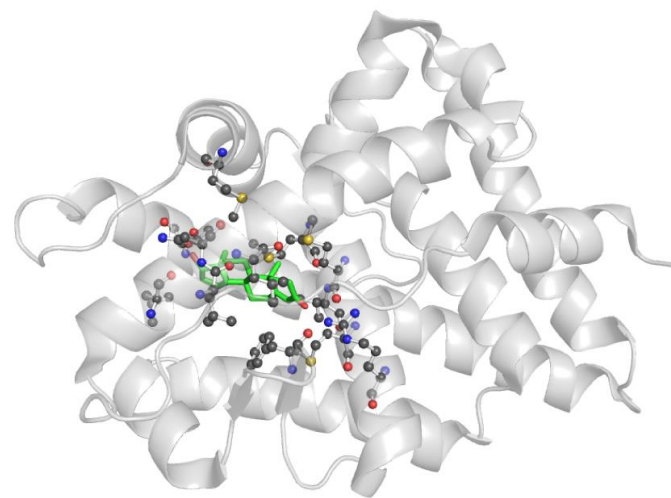
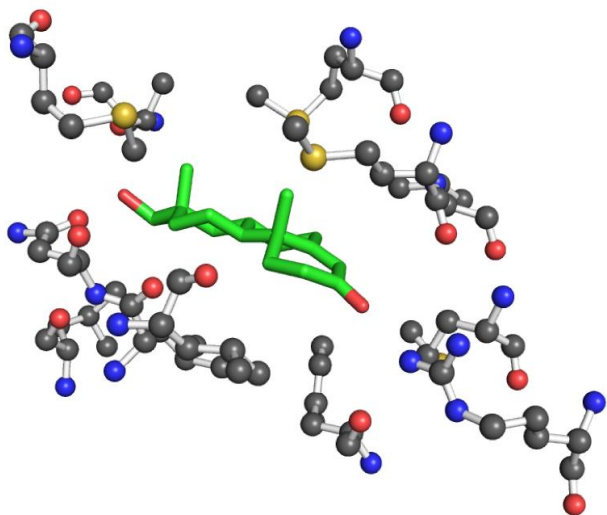
(A) The channel (closed conformation) of TRPML1 (PDB ID: 5WJ5). The blue and red dashed lines delimit the membrane region. The profiles depict polarity and charge along the pore.

(B) The channel in open conformation (PDB ID: 5WJ9). The pore is divided into three main parts: Activation gate; Central cavity and Selectivity filter. The figure was generated using Pymol software.

# PatternQuery: Description and detection of fragments

- Describes a fragment based on its topology and geometry
- Extracts the fragment from all structures in Protein Data Bank

`Residues("TES").AmbientResidues(4)`



Sehnal D., Pravda L., Svobodová R., Ionescu C.-M., Koča J. (2015) *PatternQuery: web application for fast detection of biomacromolecular structural patterns in the entire Protein Data Bank*. **Nucleic Acids Res.**, **43**, W383–W388.

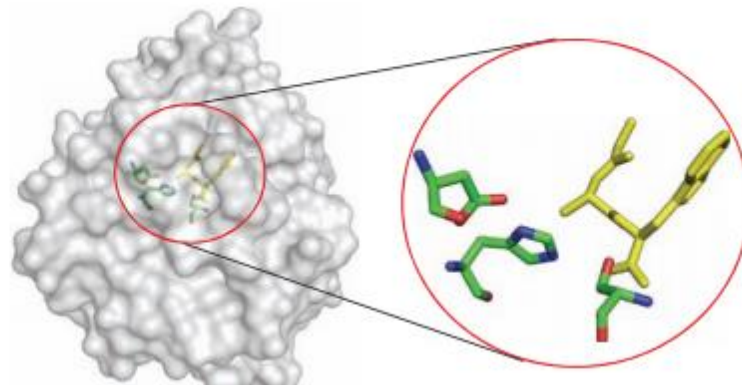
# Detection: Why to detect biomacromolecular parts?

## Fragments:

- Patterns for drug design
- Comparison of biomacromolecules
- Understanding / discovery of biomacromolecule function

## Channels:

- Key objects for biomacromolecule function
- Influence the binding site selectivity (only some substrate can path through)





# Detection: How to detect biomacromolecular fragments?

- **Methodology:**
  - Describe a fragment via a defined expression (query)
  - Find all suitable fragments
- **Tools:** PatternQuery, RASMOT-3D PRO, Promotif, Prosite, IMAAAGine, PDBeMotif, SPRITE& ASSAM, 3Dfit, SPASM, Protein segment finder

`Residues("TES").AmbientResidues(4)`

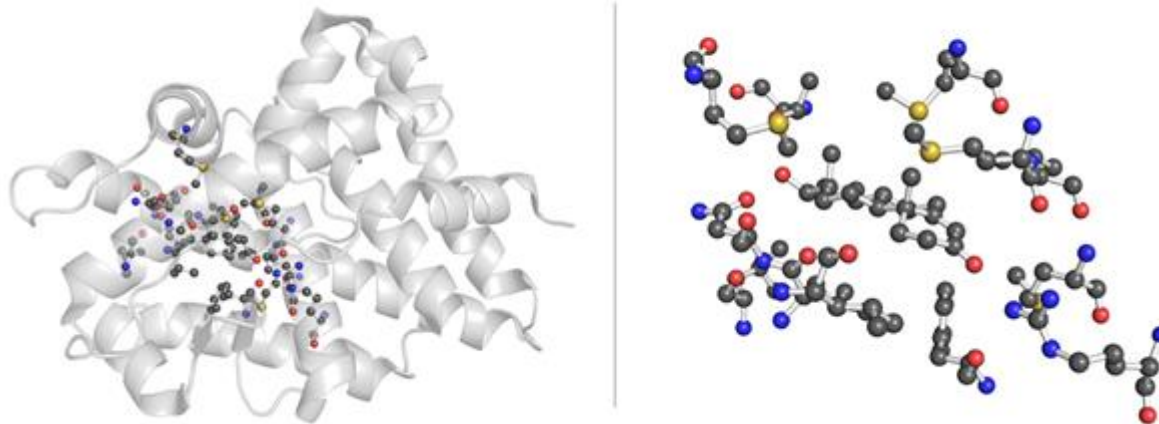
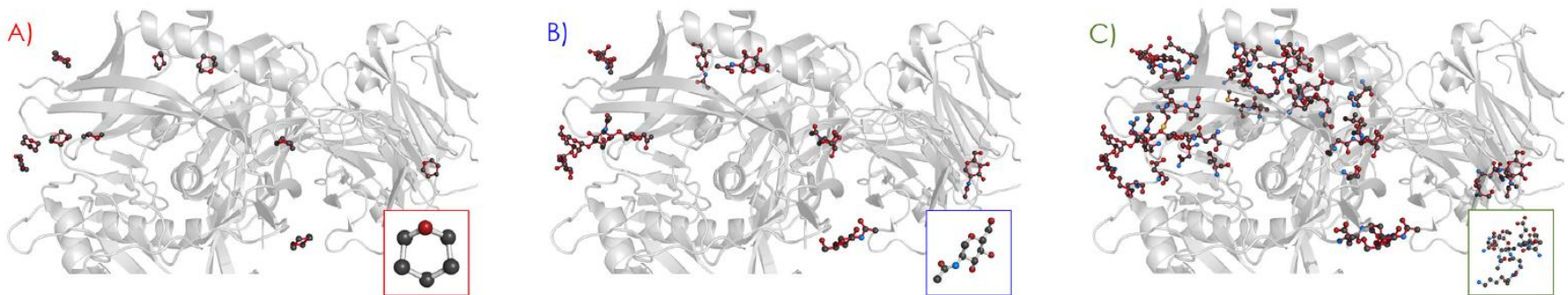


Figure: Detection of testosterone (TES) and its 4 Å large surrounding via PatternQuery: A query and a picture of the detected fragment.

# Detection: Example – binding pocket detection

## Detection of fragments within 3U7Y

- **Glycoprotein gp160** from Human immunodeficiency virus 1 in complex with Homo sapiens immunoglobulins (PDB ID 3u7y).
- **Goal:** Detect a binding pocket of any residue containing a pyranose
- **Results** - via PatternQuery (<http://ncbr.muni.cz/PatternQuery>):



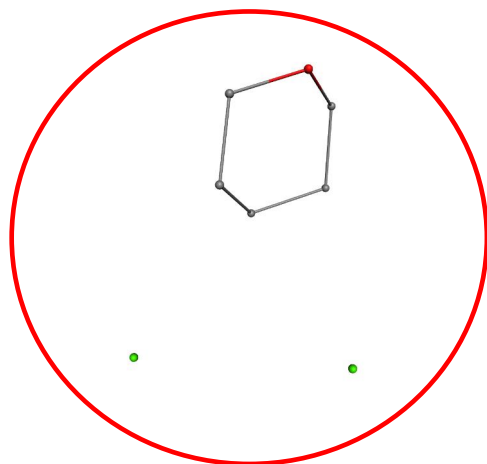
```
Rings(5 * ['C'] + ['O']). ConnectedResidues(0). AmbientResidues(4)
```

A) First, the query identifies a pyranose moiety (a ring composed of 5 carbons and an oxygen atom). B) Then, all residues which include this pattern in their structure are identified. C) Finally, all the residues that are at most 4Å from any of the pyranose containing residues are detected as well.



# PatternQuery - Detecion of key parts

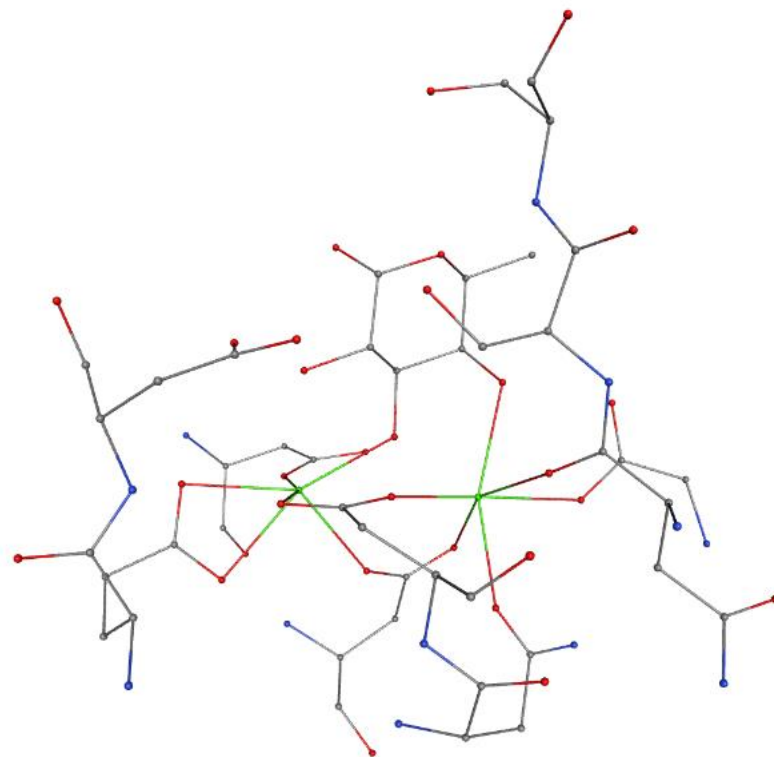
Sugar binding site from *Pseudomonas aeruginosa* lectin



```
Near (4,  
  Rings (5 * ["C"] + ["O"]),  
  Near (4, Atoms ("Ca"), Atoms ("Ca"))  
)  
.AmbientResidues (4)
```

## Results (38 entries from PDB):

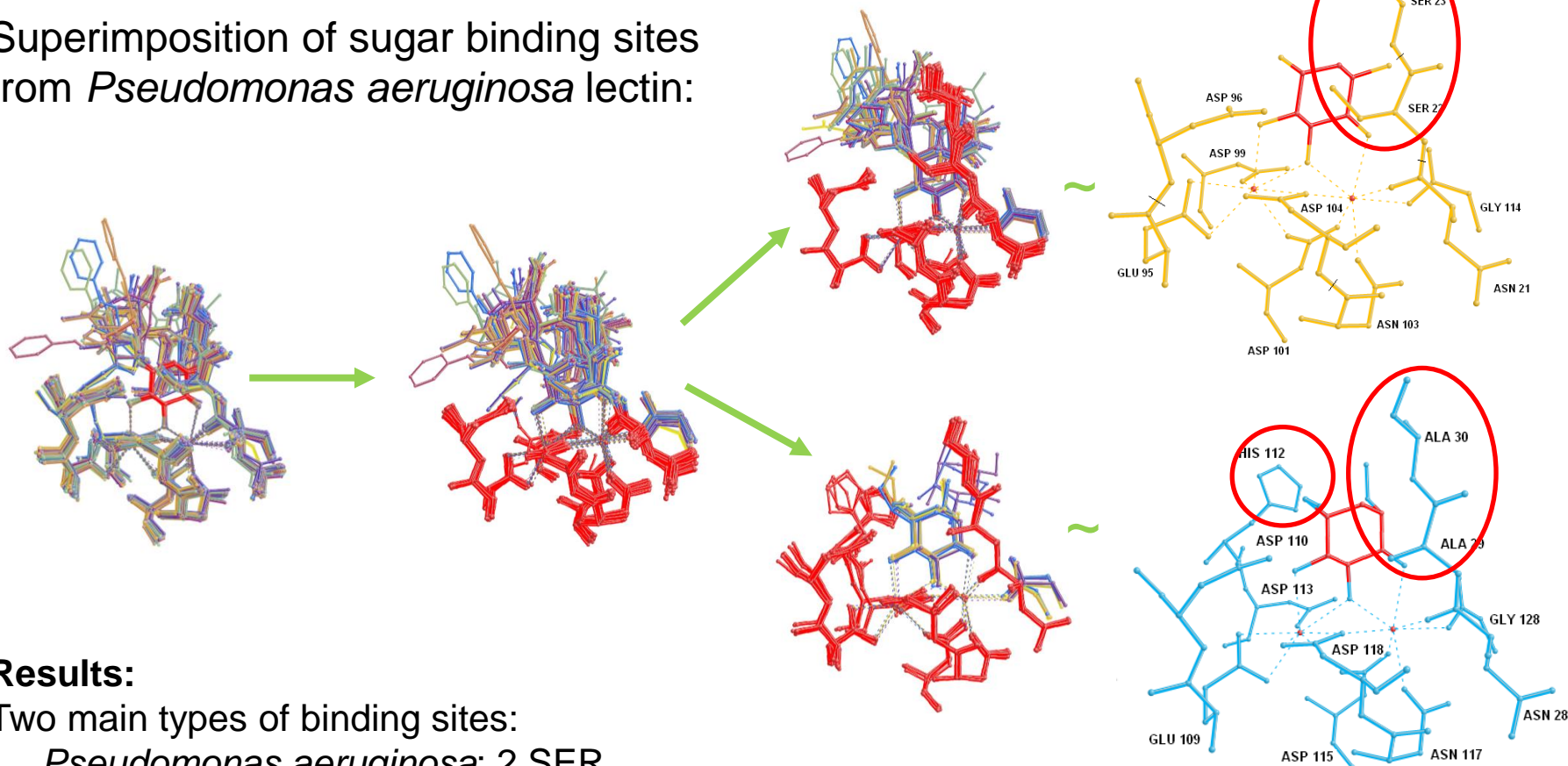
- *Pseudomonas aeruginosa* (27 entries) or its synthetic construct (2 entries)
- *Burkholderia cenocepacia* (4 entries)
- *Ralstonia solanacearum* (2 entries)
- *Chromobacterium violaceum* (2 entries)
- *Bacillus subtilis* (1 entry)



Software **PatternQuery**:  
Sehnal, Pravda, Svobodová, ..., Koča, **Nucl.**  
**Acids Res.** (2015)

# Comparison

Superimposition of sugar binding sites from *Pseudomonas aeruginosa* lectin:



## Results:

Two main types of binding sites:

- *Pseudomonas aeruginosa*: 2 SER
- *Burkholderia cenocepacia* and *Ralstonia solanacearum*: 2 ALA + HIS

Plus their combination:

- *Chromobacterium violaceum*, *Pseudomonas aeruginosa* mutants: 1 ALA, 1 SER

# Detection: Example – sugar-binding site detection

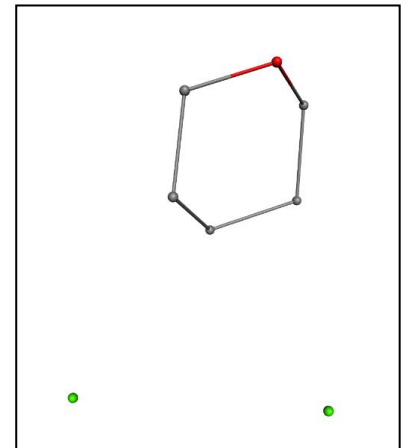
## Detection of *Pseudomonas aeruginosa* Lec-B sugar binding site

### Motivation:

- *Pseudomonas aeruginosa* is an opportunistic pathogen associated with a number of chronic infections.
- It forms a biofilm enabling it to survive both the response of the host immune system, and antibiotic treatment.
- One of the cornerstones of biofilm formation is the presence of sugar-binding protein LecB (PA-IIL).
- Its inhibition is considered to be a promising approach for anti-pseudomonadal treatment.

### Description of the binding site:

- A pair of Ca atoms (distance  $< 4 \text{ \AA}$ )
- They are near to ( $< 4 \text{ \AA}$ ) the pyran ring (= ring containing 5 C and 1 O).
- Include its surrounding (residues closer than  $4 \text{ \AA}$ )



# Detection: Example – sugar-binding site detection

## Detection of *Pseudomonas aeruginosa* Lec-B sugar binding site

**Pattern:**

```
Near(4,  
    Rings(5 * ["C"] + ["O"]),  
    Near(4, Atoms("Ca"), Atoms("Ca"))  
)  
.AmbientResidues(4)
```

### Results:

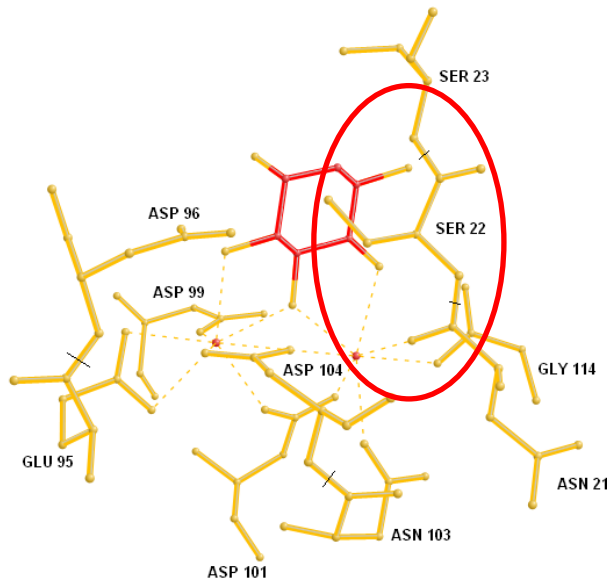
The binding site is contained in 40 PDB entries from 4 organisms (bacteria):

- *Pseudomonas aeruginosa*
- *Burkholderia cenocepacia*
- *Ralstonia solanacearum*
- *Chromobacterium violaceum*

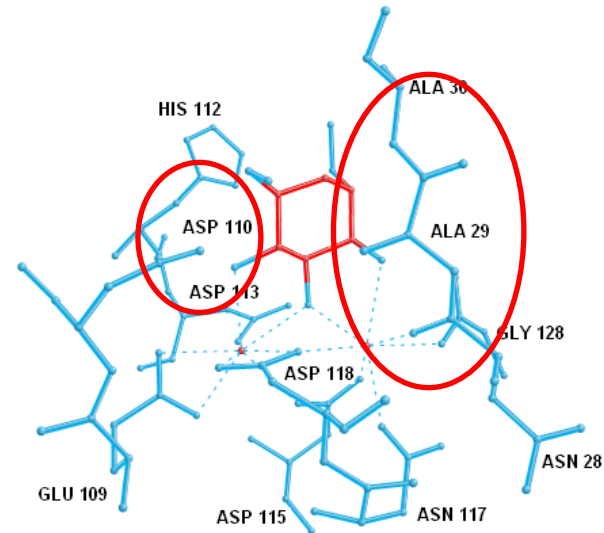
# Detection: Example – sugar-binding site detection

## Detection of *Pseudomonas aeruginosa* Lec-B sugar binding site

**Results:** The binding site is very conserved and it has two structure variants:



*Pseudomonas aeruginosa*  
*Chromobacterium violaceum*



*Burkholderia cenocepacia*  
*Ralstonia solanacearum*

**Conclusion:** The anti-pseudomonadal treatment have to target the above described binding sites (with higher affinity than the sugar).