



CEITEC



Central European Institute of Technology
BRNO | CZECH REPUBLIC

Biomacromolecular structure analysis - detection of patterns

Radka Svobodová
CEITEC, Masaryk University



EUROPEAN UNION
EUROPEAN REGIONAL DEVELOPMENT FUND
INVESTING IN YOUR FUTURE

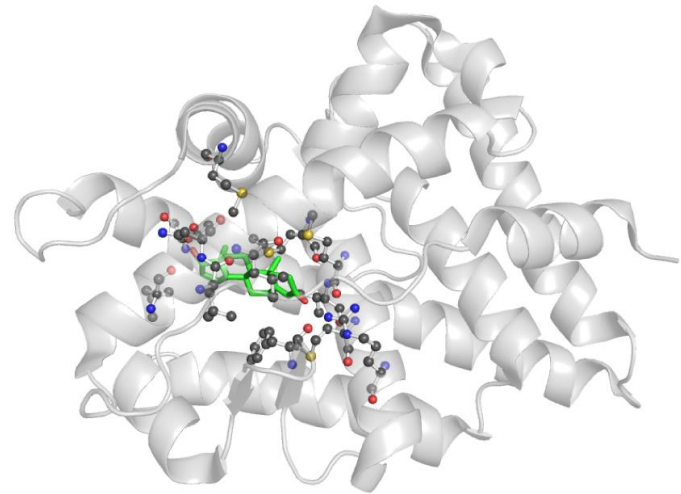
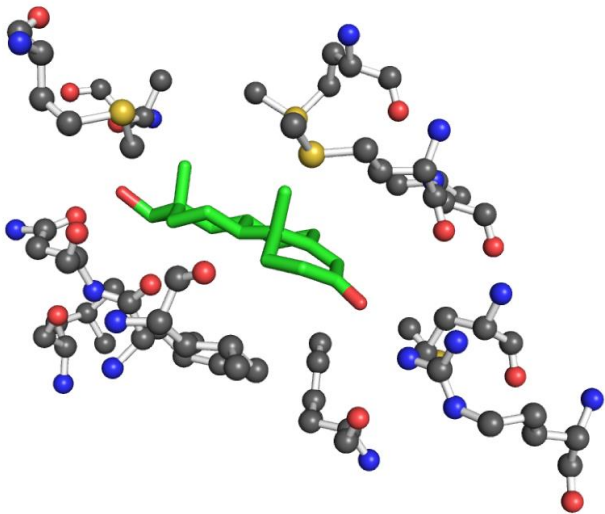


2007-13
OP Research and
Development for Innovation

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.

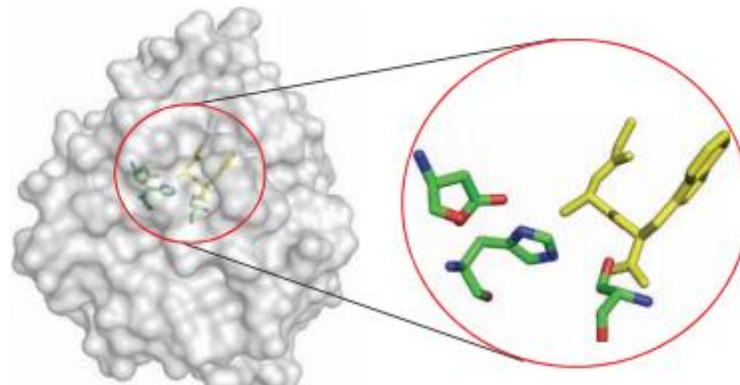
Detection: Why to detect biomacromolecular parts?

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)`

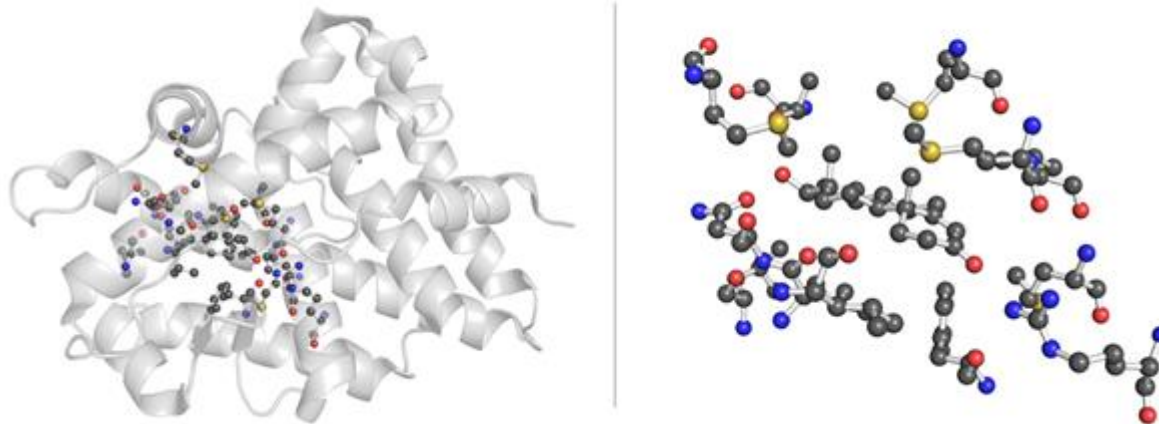
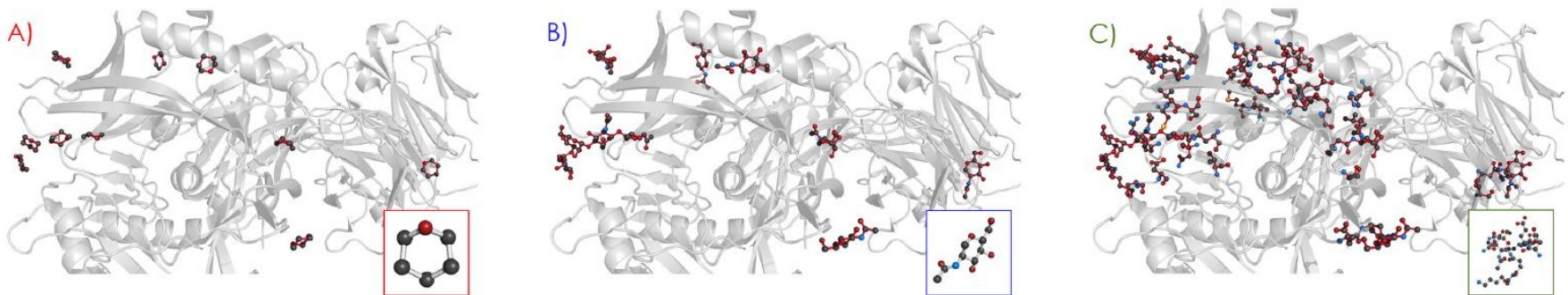


Figure: Detection of testosterone (TES) and its 4 Å large surrounding via PatternQuery: A query and a picture of the detected fragment.

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>):

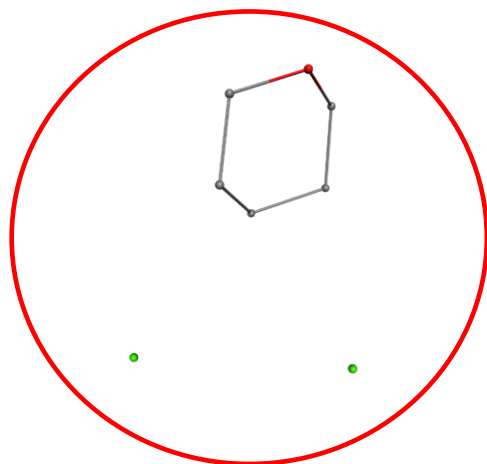


```
Rings(5 * ['C'] + ['O']). ConnectedResidues(0). AmbientResidues(4)
```

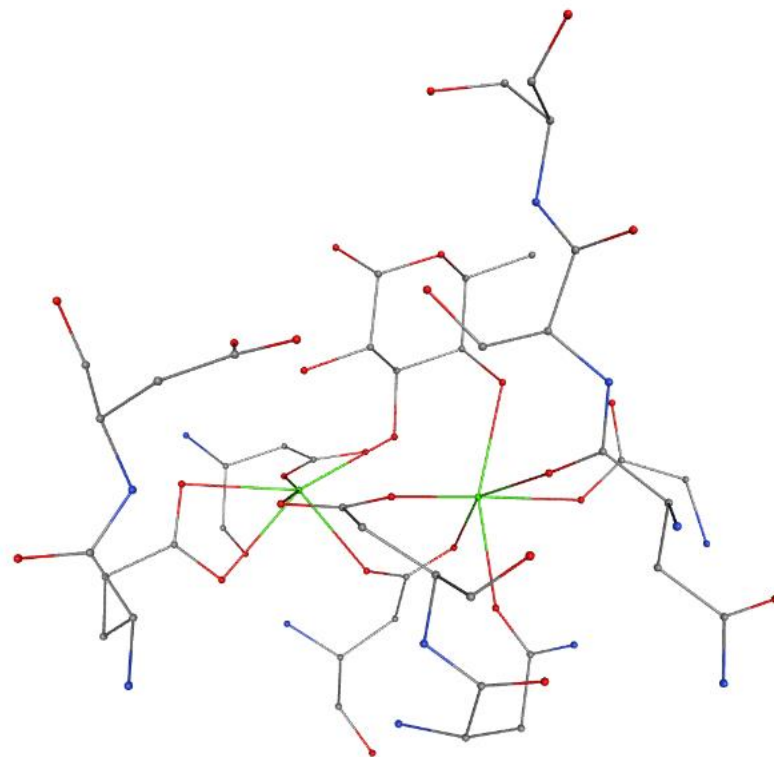
A) First, the query identifies a pyranose moiety (a ring composed of 5 carbons and an oxygen atom). B) Then, all residues which include this pattern in their structure are identified. C) Finally, all the residues that are at most 4Å from any of the pyranose containing residues are detected as well.

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)
```



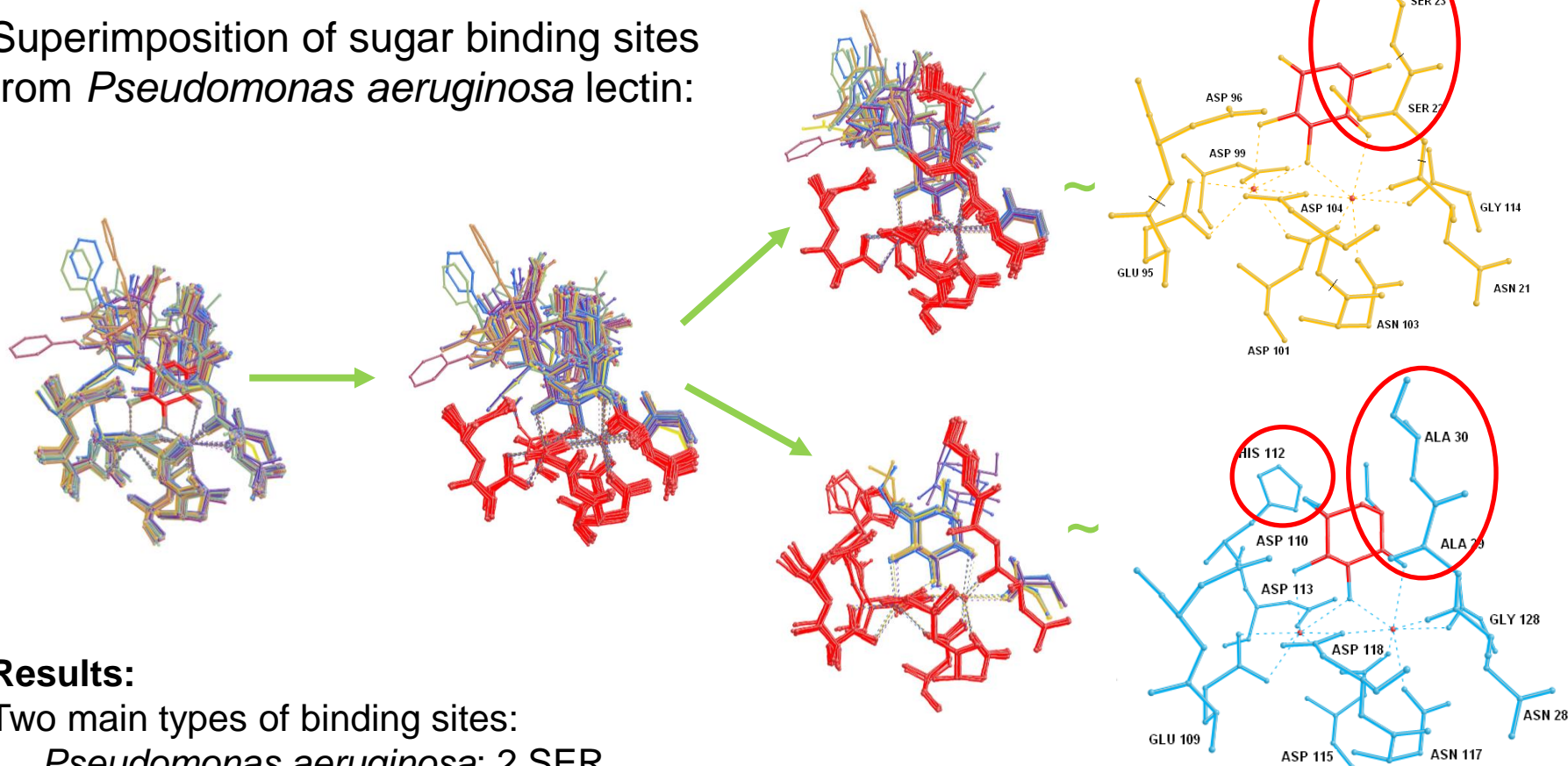
Results (38 entries from PDB):

- *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

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*