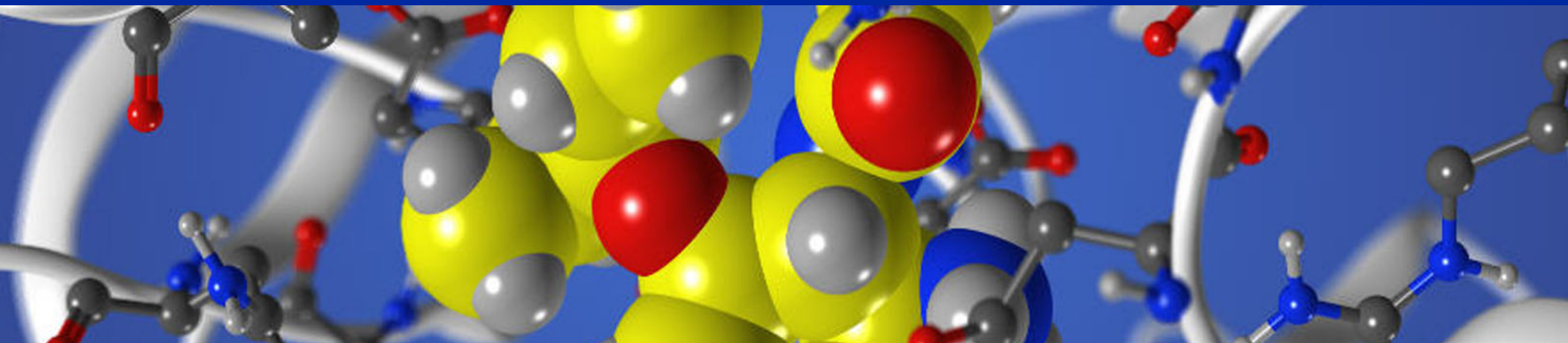
Analyzing Search Hits with ChemAxon's Markush Enumeration Tool

Dr. Guy de Weck - pRED Informatics



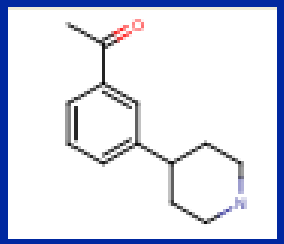
General workflow for a Novelty Search

ChemAxon's Markush Enumeration Tool

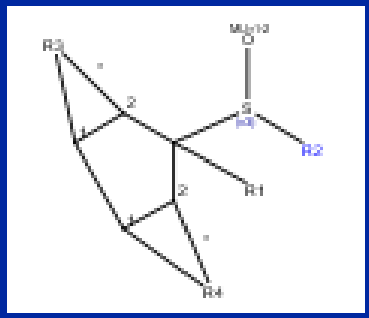
Suggested Improvements

General Workflow for a Novelty Search

- Idea for potentially new chemical structures
- Verifying patentability in various databases:
  - MMS, MARPAT, REGISTRY, Reaxys, …
- Analyzing search hits
  - Tedious work: determining structural proximity of a searched substructure compared to the claimed structures

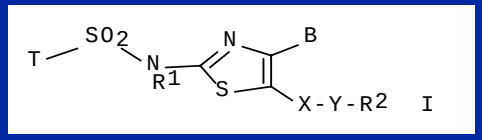
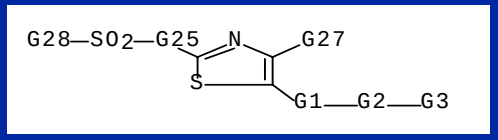
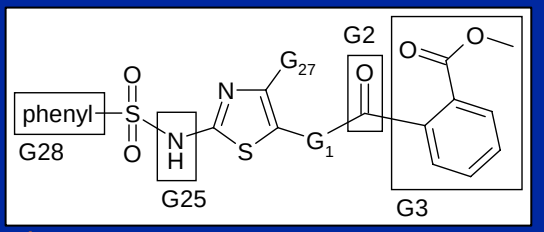
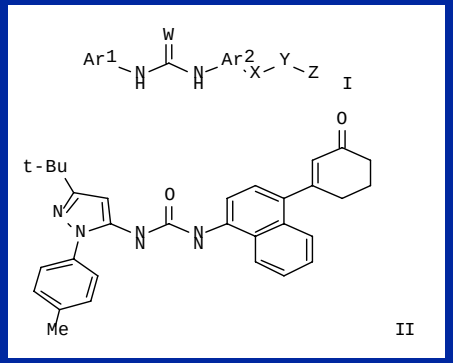
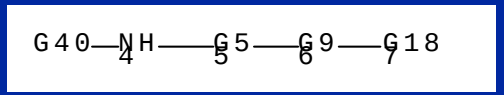
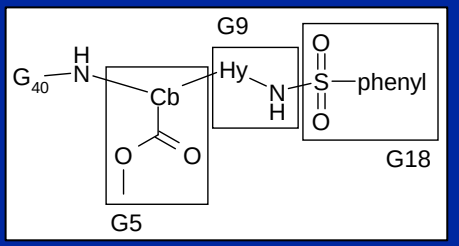
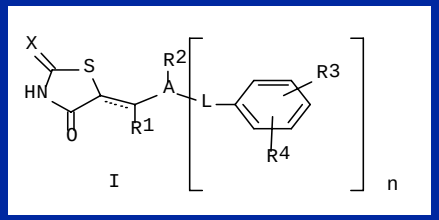
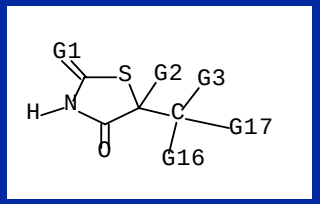


Searched substructure

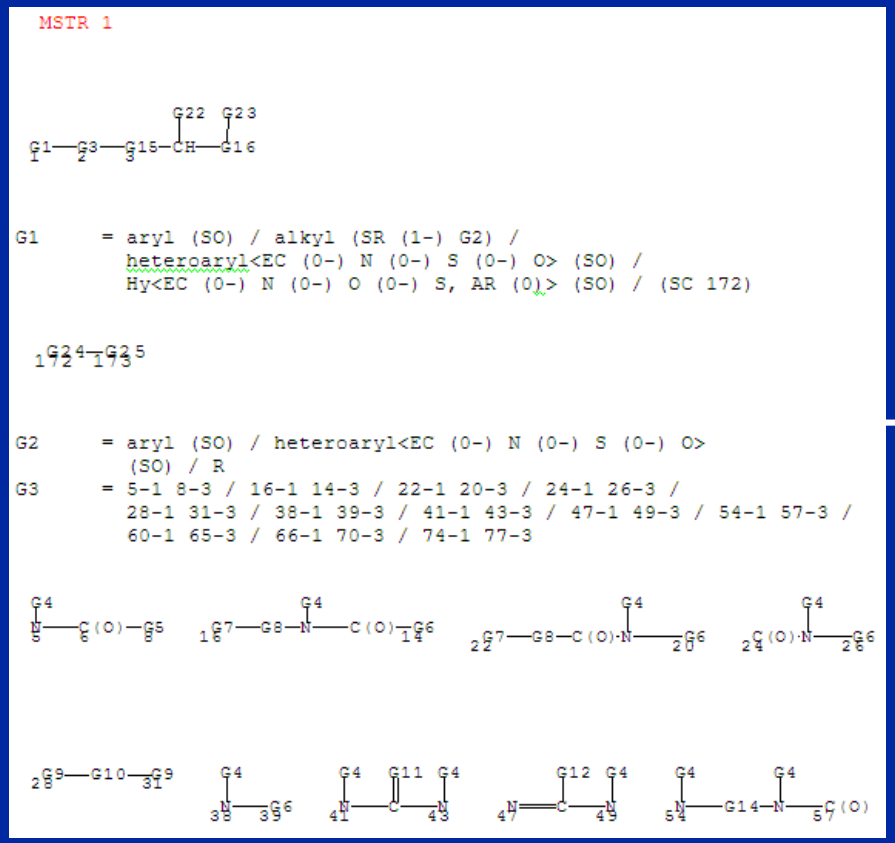


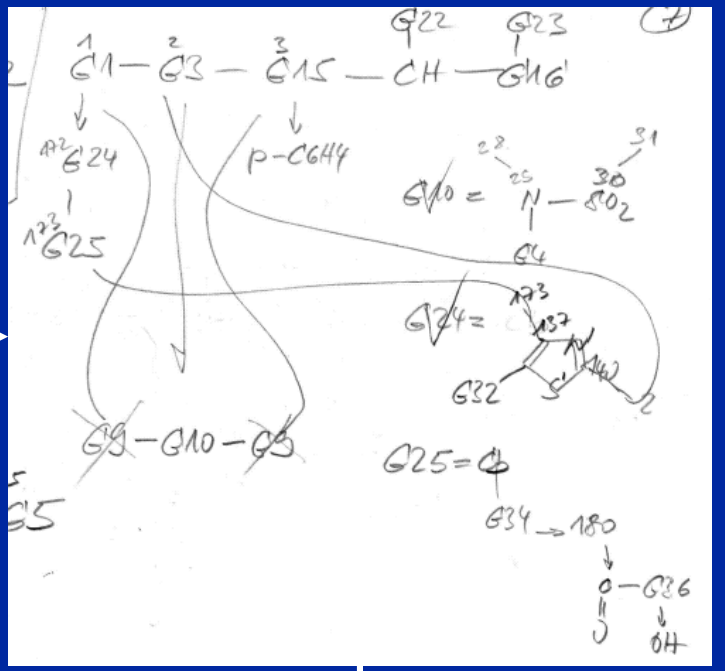
Markush structure

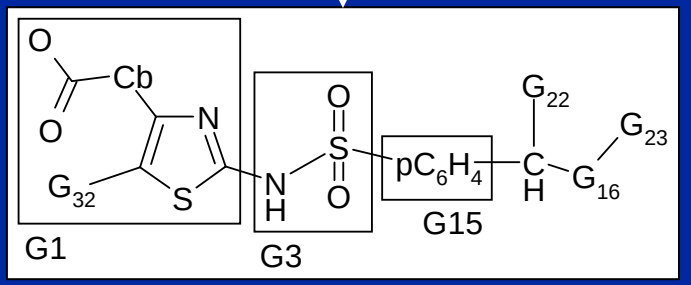
General Workflow for a Novelty Search

Graphic Image	Markush structure	Structural proximity
		 close
		 medium
		G17 contains the searched substructure distant

General Workflow for a Novelty Search



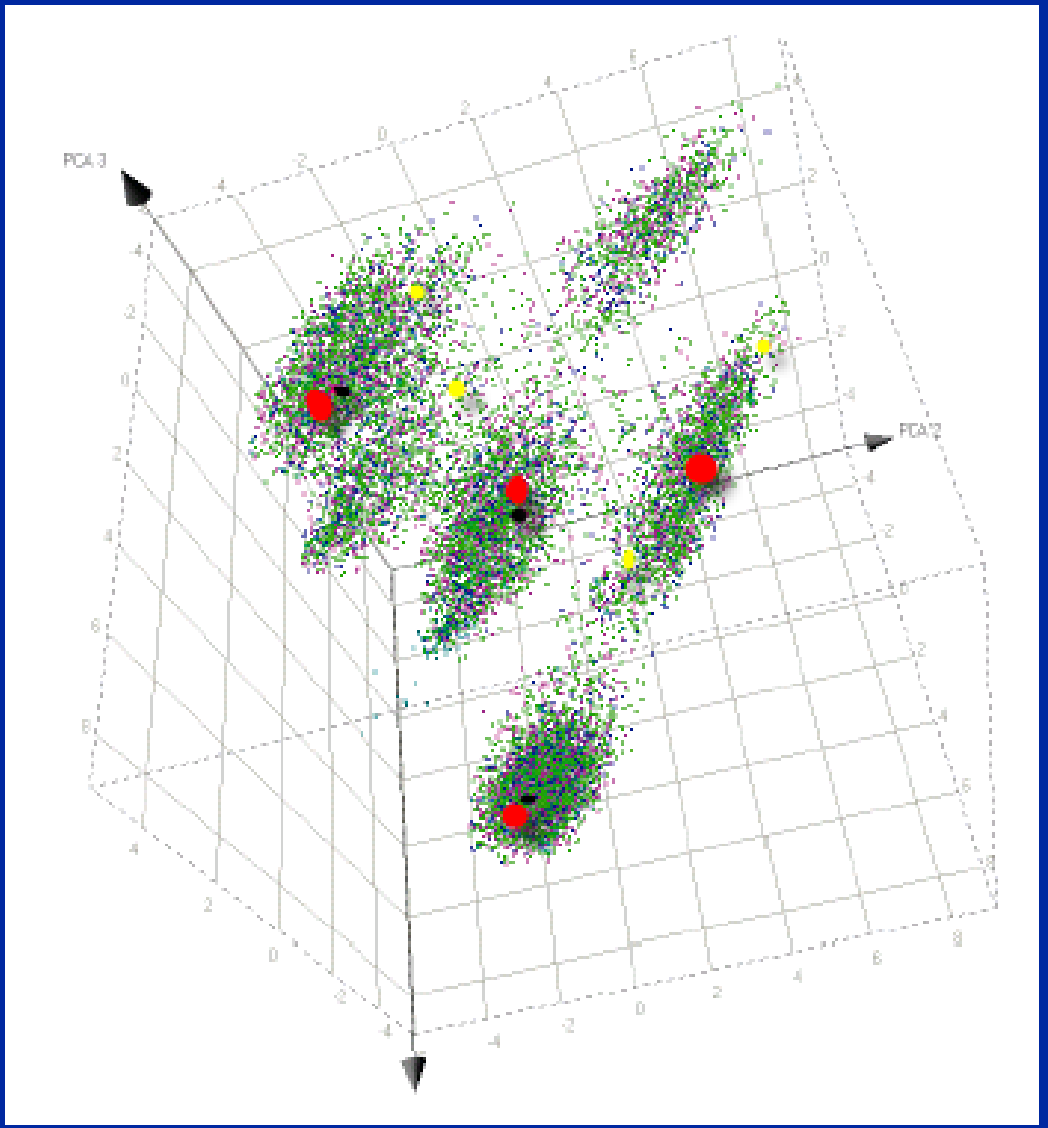




- Defining structural proximity is an extremely time consuming process!

General Workflow for a Novelty Search

- A graphical representation of structural proximity:
  - Markush space density of a patent
  - Red: exemplified and prophetic structures
  - Yellow: distant structural proximity
  - Black: close structural proximity
- Markush space from decrIPt, PCA of normalized fingerprints



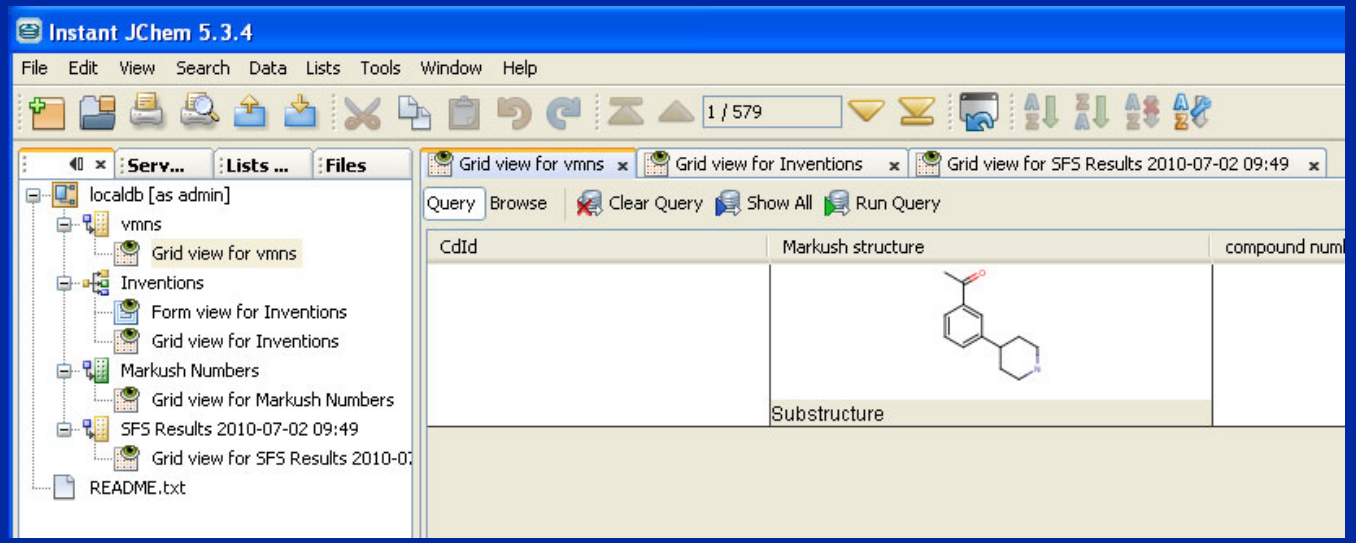
General Workflow for a Novelty Search

ChemAxon's Markush Enumeration Tool

Suggested Improvements

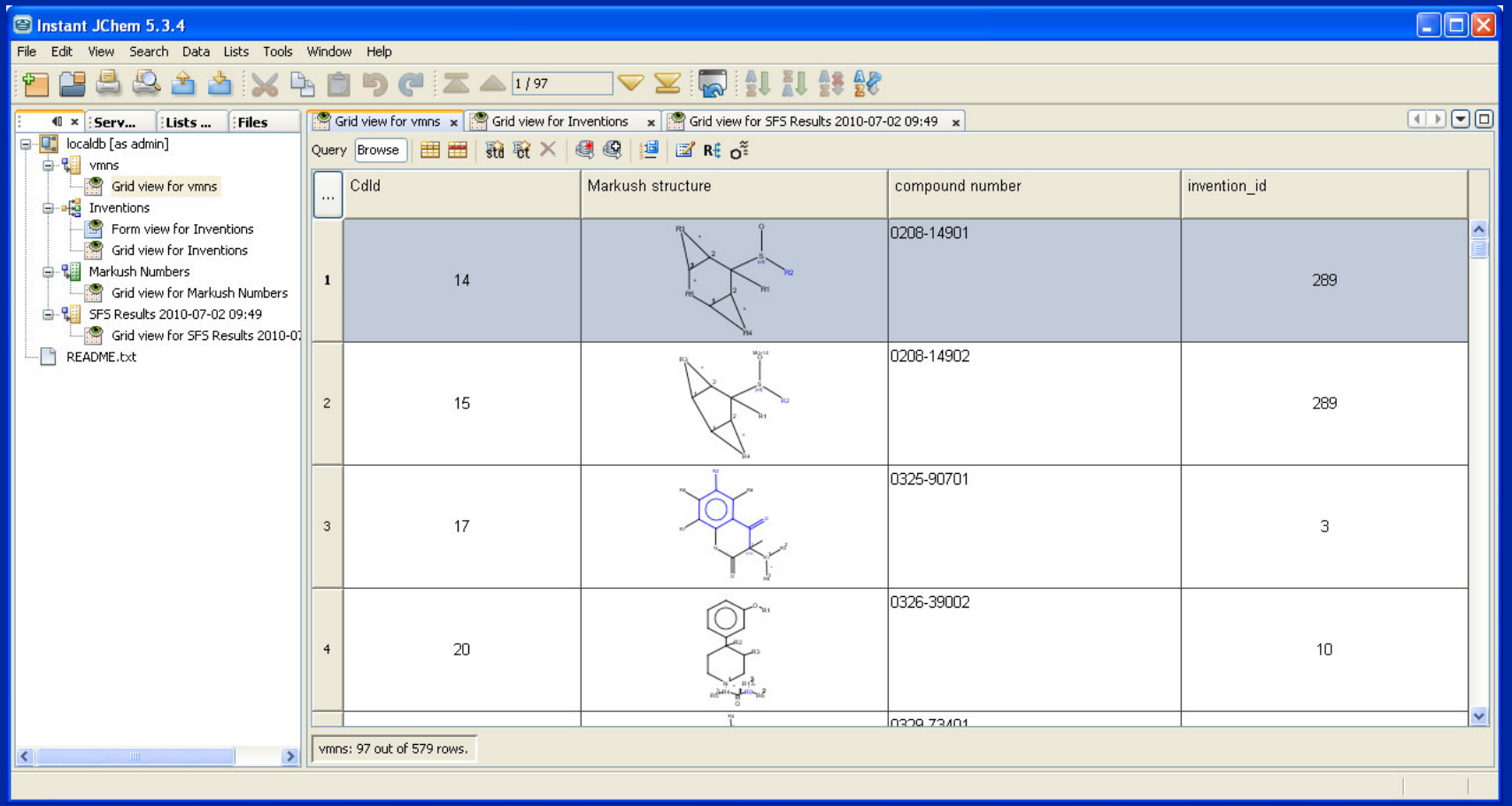
ChemAxon's Markush Enumeration Tool

- How can ChemAxon's Markush enumeration tool help defining structural proximity?
- The 300 patents and their corresponding 579 Markush structures were loaded in Instant JChem and searched by substructure.



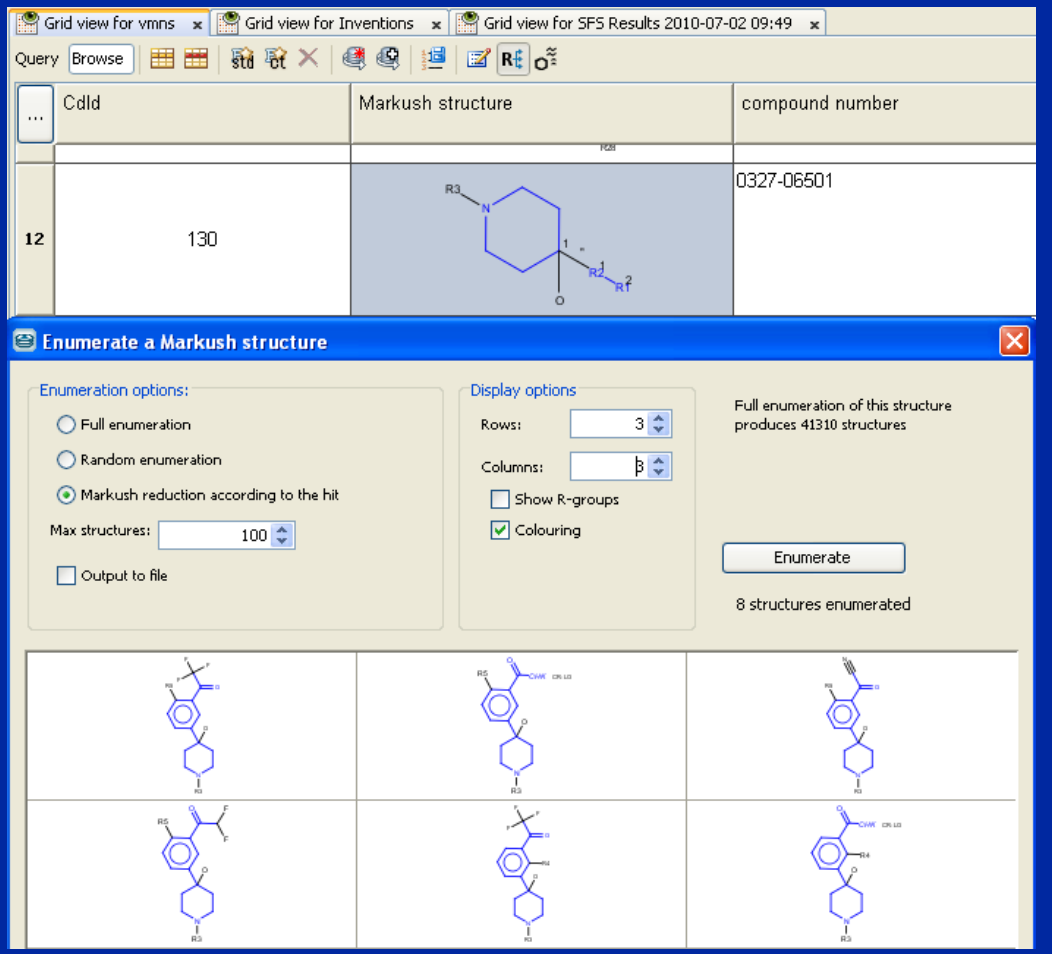
ChemAxon's Markush Enumeration Tool

- This substructure search retrieved 97 Markush structures.



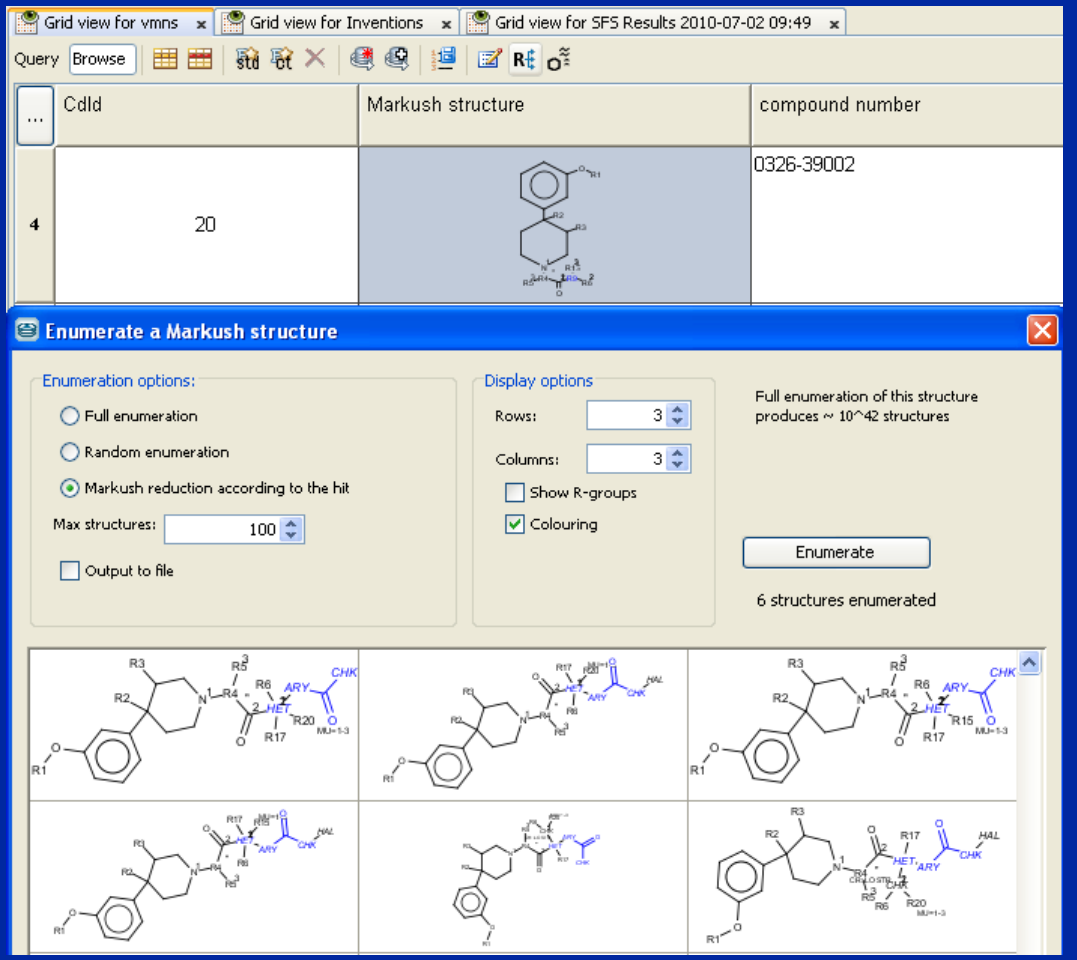
ChemAxon's Markush Enumeration Tool

- An example of close structural proximity using Markush reduction according to the hit:
- Full enumeration: 41310 structures.
- Markush reduction according to the hit: 8 structures.
- The highlighted part suggests close structural proximity of the searched substructures with the exemplified/prophetic structures.



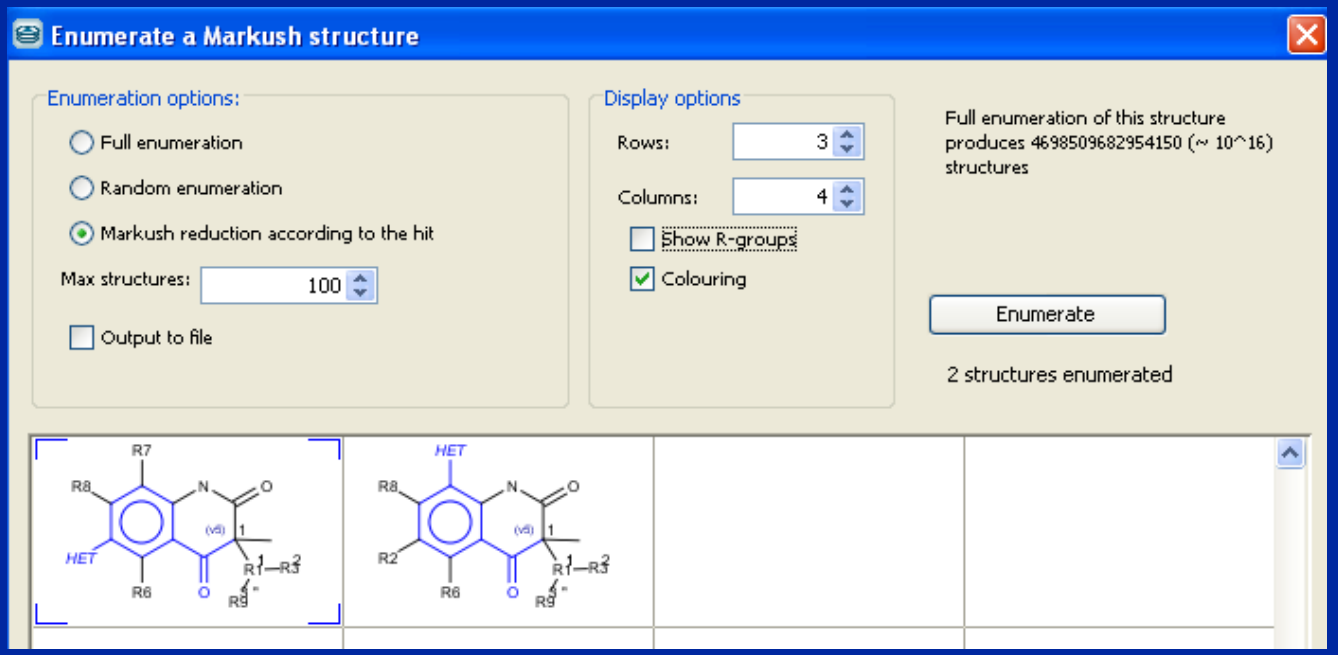
ChemAxon's Markush Enumeration Tool

- An example of distant structural proximity using Markush reduction according to the hit:
- Full enumeration: 10^42 structures.
- Markush reduction according to the hit: 6 structures.
- The highlighted part suggests distant structural proximity of the searched substructures with the exemplified/prophetic structures.



ChemAxon's Markush Enumeration Tool

- An example of medium structural proximity using Markush reduction according to the hit:



- Full enumeration: 10^16 structures.
- Markush reduction according to the hit: 2 structures.
- In this case subsequent partial enumeration steps are needed to define the structural proximity.

ChemAxon's Markush Enumeration Tool

- Summary:
  - Patent records having a distant structural proximity relative to the searched substructure are easily identified with ChemAxon's Markush enumeration tool. This is saving time in the novelty search workflow.
  - Patent records with a medium to close structural proximity relative to the searched substructure need to be analyzed in more detail:
    - Repeated partial enumeration by one or more R groups within the Markush enumeration tool is needed.

Roche

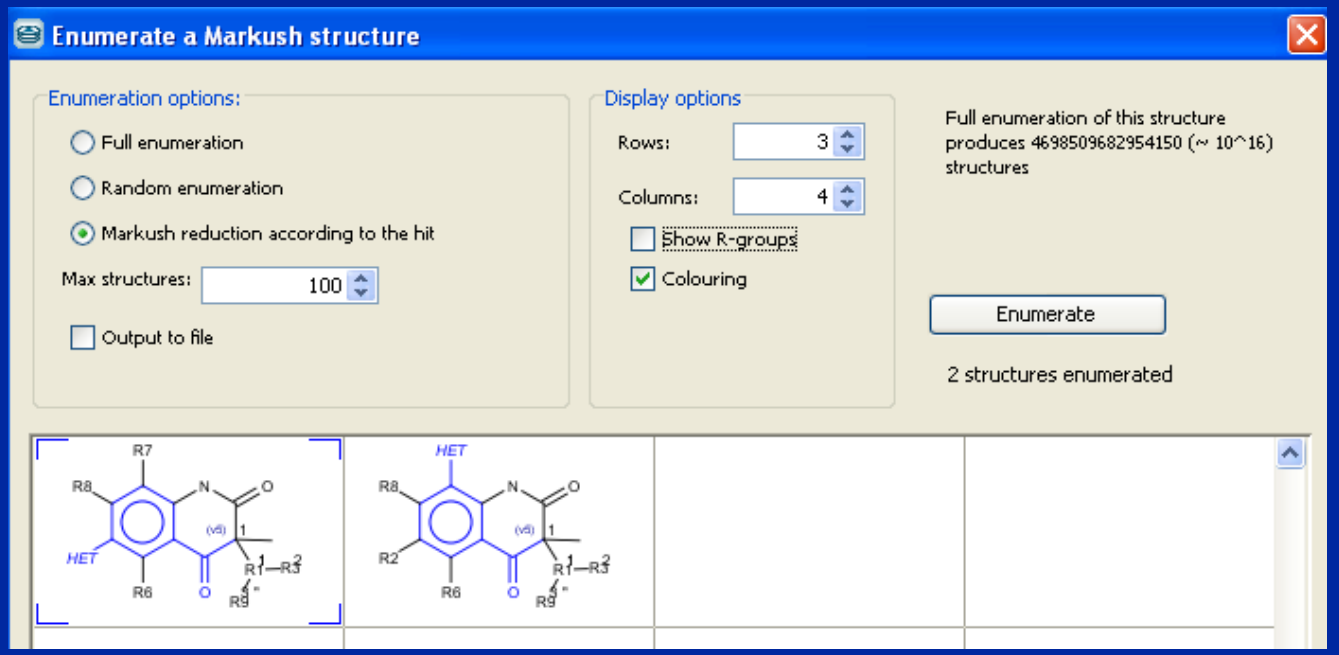
Novelty Search Workflow

ChemAxon's Markush Enumeration Tool

Suggested Improvements

Suggested Improvements

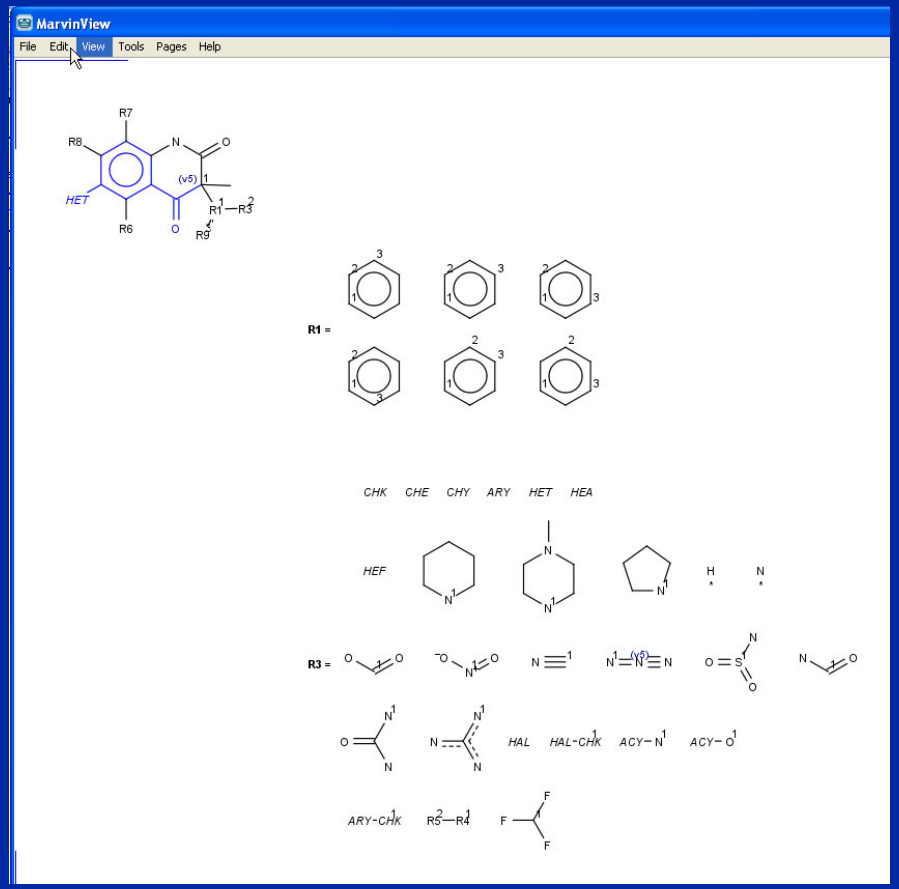
- Repeated partial enumeration by one or more R groups within the Markush enumeration tool.



  - Possibility of selecting one or more R-groups and enumerating these R-groups.

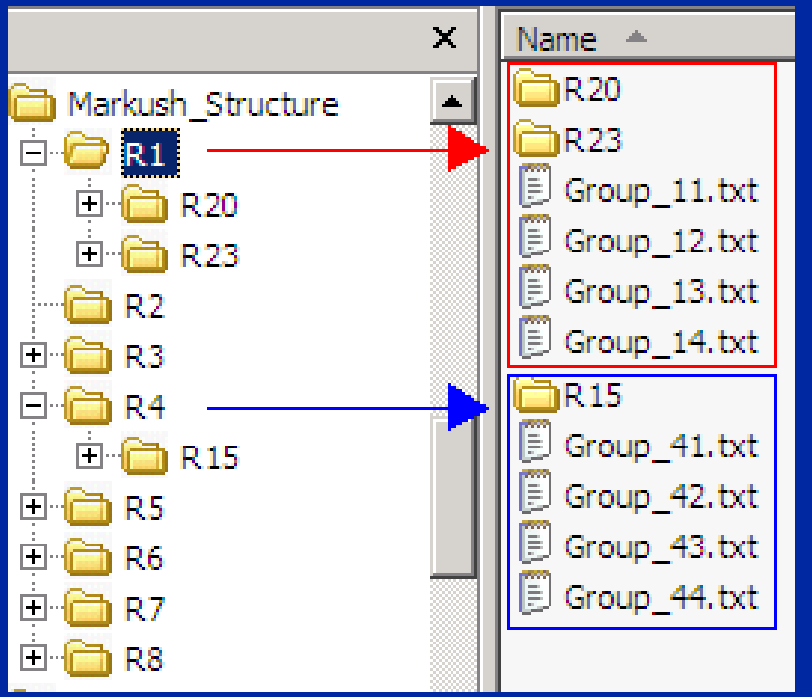
Suggested Improvements

- Display of enumerated structures when “Show R-groups” is selected:
  - The (partially) enumerated structure must be visible at all times
  - Only selected R-groups should be enumerated.



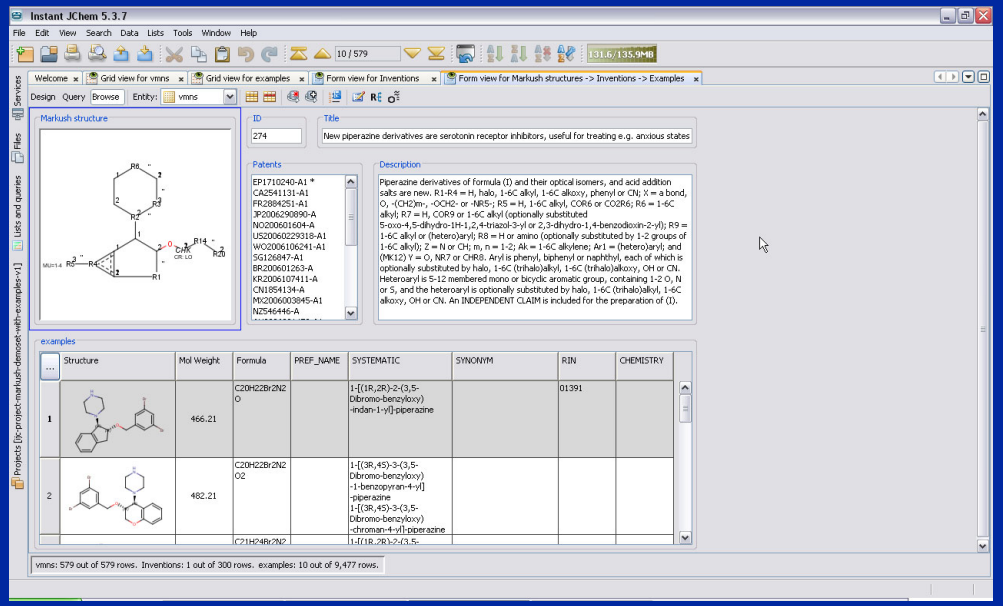
Suggested Improvements

- Explorer like display of enumerated structures: (similar to the Markush display in Torus from digital Chemistry)
- Multiple selection of R-groups
- Additional suggestions:
  - Displaying search and filtering steps: action thread on top and results below.



Suggested Improvements

- Display of exemplified / prophetic structures.
  - Substructure search on the set of exemplified/prophetic structures of the corresponding patent.
  - Highlighting of the searched substructure



(Screen shot provided by ChemAxon)

Suggested Improvements

- Running novelty searches in MMS out of ChemAxon's Instant JChem.
- Federated novelty search over all relevant sources?

Acknowledgements

Thomson Reuters:	Steve Hajkowski
ChemAxon:	Szabolcs Csepregi, Alex Allardyce, Douglas Drake
Roche:	Torsten Schindler

We Innovate Healthcare