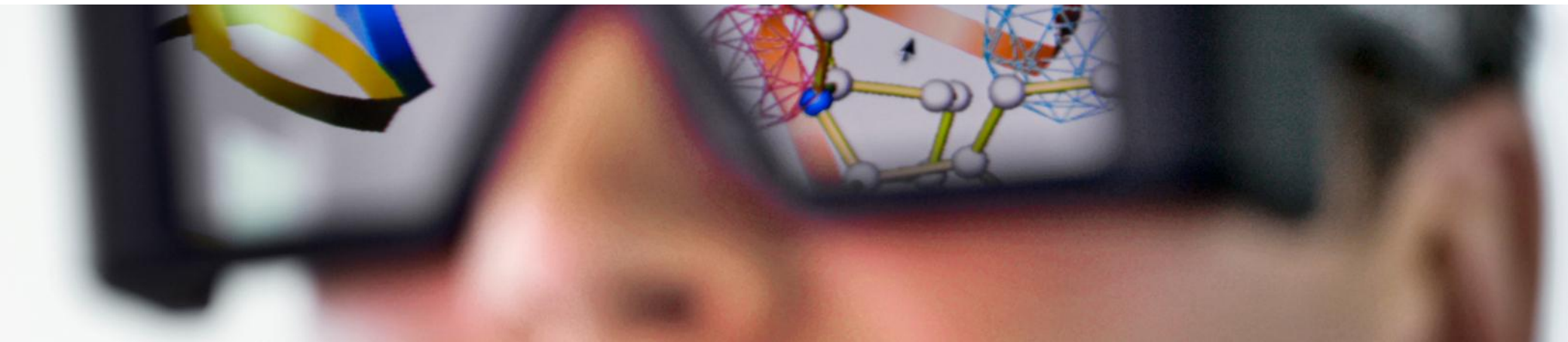


---

# **SAR Analyzer – Intuitive Access To SAR Information**

*Lisa Peltason, Daniel Stoffler*

*Cheminformatics & Statistics, Roche Basel*



# SAR Analyzer

## Goals

- Facilitate navigation of complex data landscapes
- Provide intuitive access to SAR information
- Support decision-making process
- **Not:** predictions

## Approach

- Offer portfolio of interactive tools
- Focus on information-rich visualization & intuitive interaction
- Combine & adapt existing methods

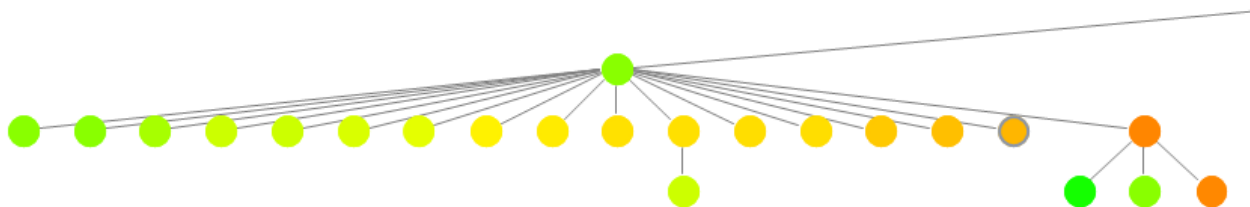
## Implementation

- Java Eclipse RCP application
- Modular plug-in architecture
- Standalone prototype applications
- JChem/Marvin/Screen Java API for data structures, representation, and algorithms

# SAR Tree

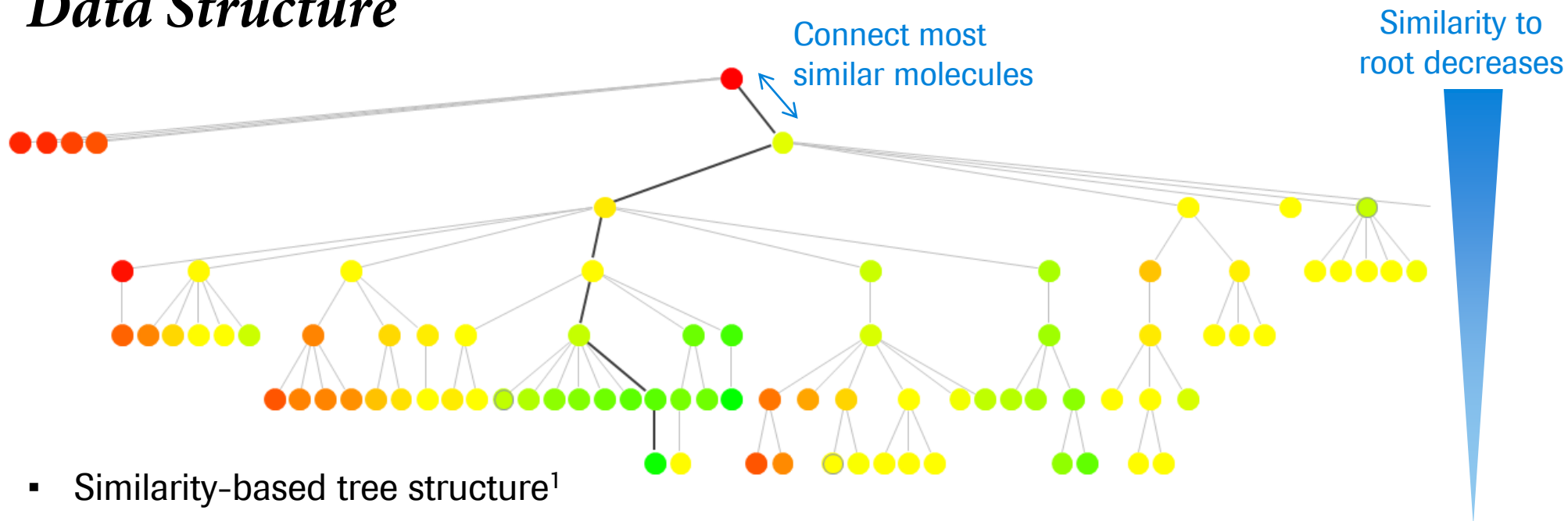
## Questions

- How do chemical modifications alter biological activity?
- Are there systematic SAR trends?
- How is activity related to molecular properties?
- Can we detect activity cliffs?



# SAR Tree

## Data Structure

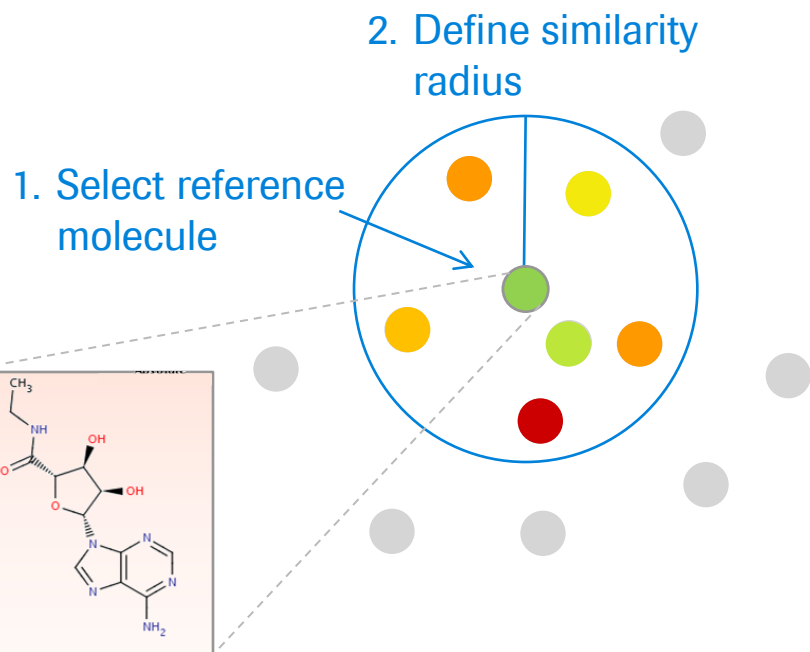


- Similarity-based tree structure<sup>1</sup>
  - Each node represents one molecule
  - Continuous color spectrum according to activity / molecular properties
  - Edges connect closely related molecules (“Nearest neighbors”) (1)
  - Similarity to the root compound gradually decreases in top-down direction (2)
- Pathways in the tree visualize SAR information:
  - Gradual structural changes departing from root molecule along with changes in activity / properties

[1] “Similarity-Potency Trees”: Wawer M; Bajorath J; *J. Chem. Inf. Model.* **2010**, 50, 1395-1409.

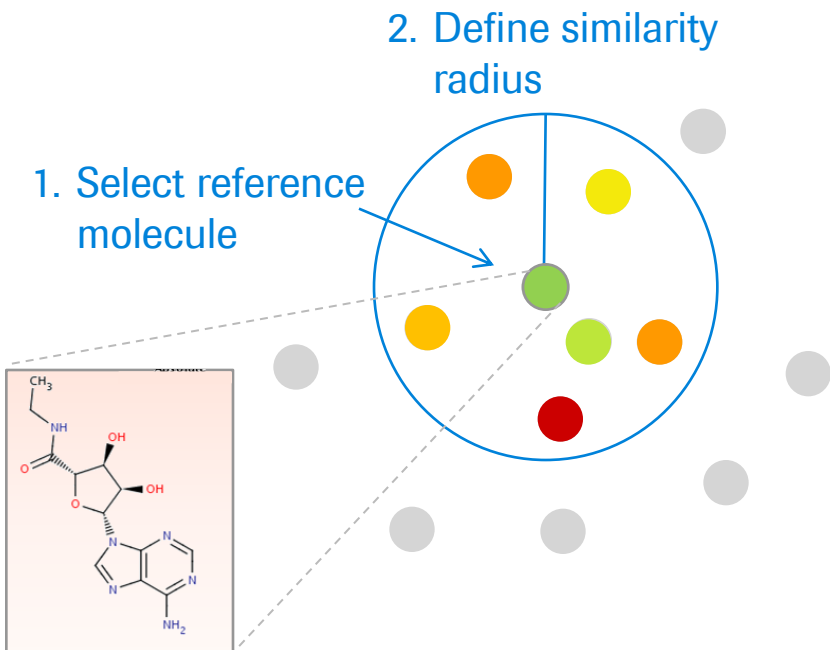
# SAR Tree

## Tree Construction



# SAR Tree

## Tree Construction



## Similarity Calculation

- ChemAxon fingerprints
  - Hashed chemical fingerprint
  - Pharmacophore fingerprint
  - Custom feature list fingerprint from input\*
  - ECFP / FCFP
  - BCUT descriptors
- Pairwise similarity calculation using Tanimoto, Tversky, or normalized Euclidean dissimilarity:

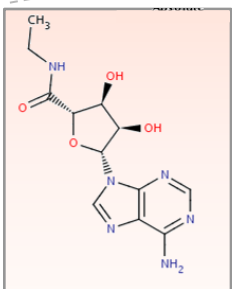
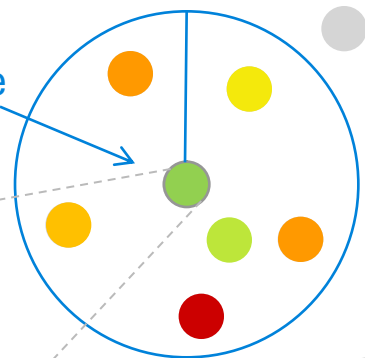
```
sim = 1 -
thisDesc.getDissimilarity(otherDesc)
```

\* Using F. Pitschi's work as template

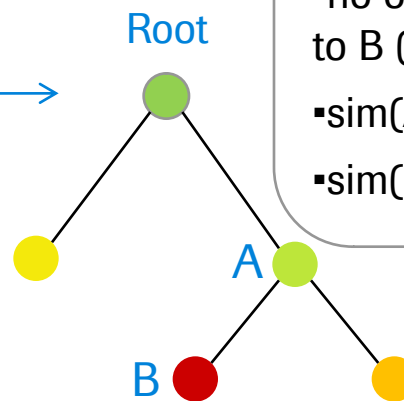
# SAR Tree

## Tree Construction

1. Select reference molecule
2. Define similarity radius



Reference molecule



### 3. Construction rules

Connect molecule B to A only if

- no other molecule is more similar to B (nearest neighbors) (1)
- $\text{sim}(A,B) > \text{threshold}$
- $\text{sim}(B,\text{root}) < \text{sim}(A,\text{root})$  (2)

# SAR Tree Prototype

The screenshot displays the SAR Viewer interface with several key components:

- MoleculeTable:** A table with columns for 'Index' and 'Structure', showing chemical structures and their corresponding indices.
- Sar Tree:** A hierarchical tree visualization where nodes are colored (yellow, red, green) to represent different SAR features. A callout box highlights 'Tree and path display' with the text: 'Interactive visualizations using Eclipse Zest toolkit'.
- Path:** A detailed view of a path through the tree, showing the chemical structure of 'piC50\_A2a' and its similarity scores (0.86, 0.9, 0.98) relative to other compounds.
- Analysis Panel:** A control panel on the right for tree construction, including:
  - 'Select property for tree construction:' dropdown set to 'piC50\_A2a'.
  - 'Transform to logarithmic scale' checkbox.
  - 'Tree size parameters' section.
  - 'Tree similarity parameters' section with sliders for 'Min. % similarity for compounds in tree' (set to 60.0) and 'Min. % similarity for connected compounds' (set to 70.0).
  - 'Update tree' button.

## Molecule data spreadsheet

- SWT-based JFace table
- Static molecule images

Mol.toBinFormat("png, wireframe")

## Tree and path display

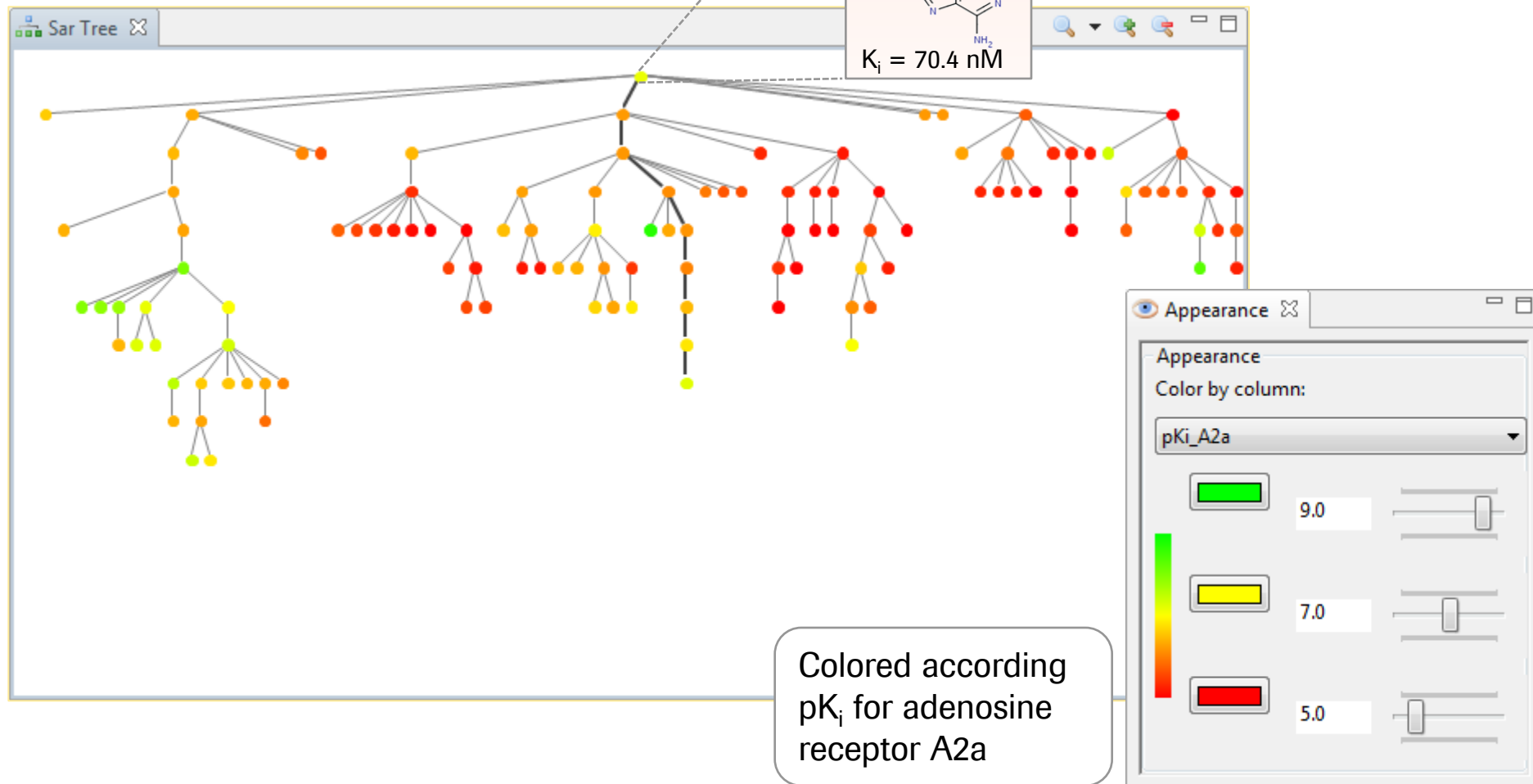
- Interactive visualizations using Eclipse Zest toolkit

## Control panel

- Display and calculation settings

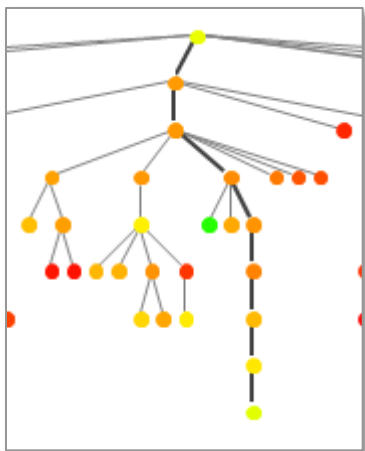
# SAR Tree Analysis

Adenosine receptor data set from ChEMBL<sup>2</sup>

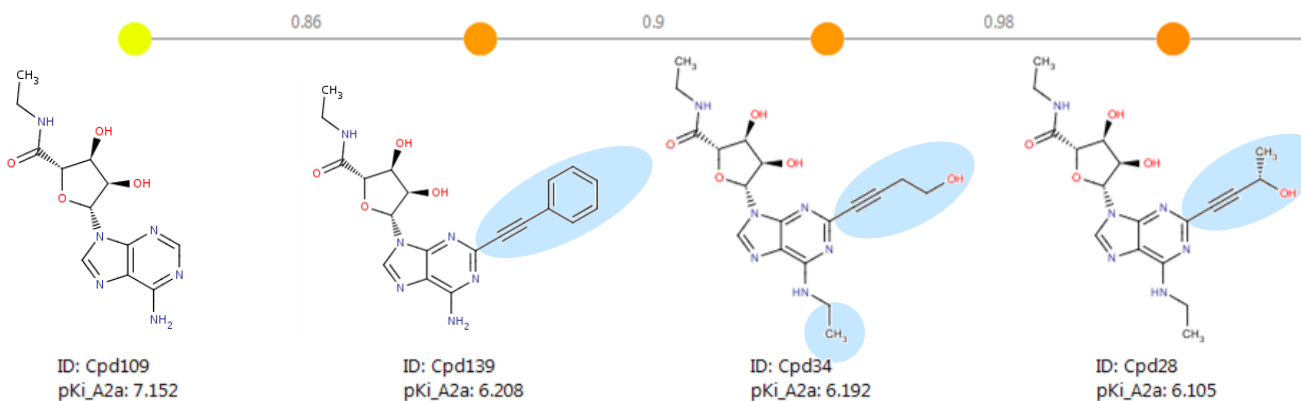


Colored according  
pK<sub>i</sub> for adenosine  
receptor A2a

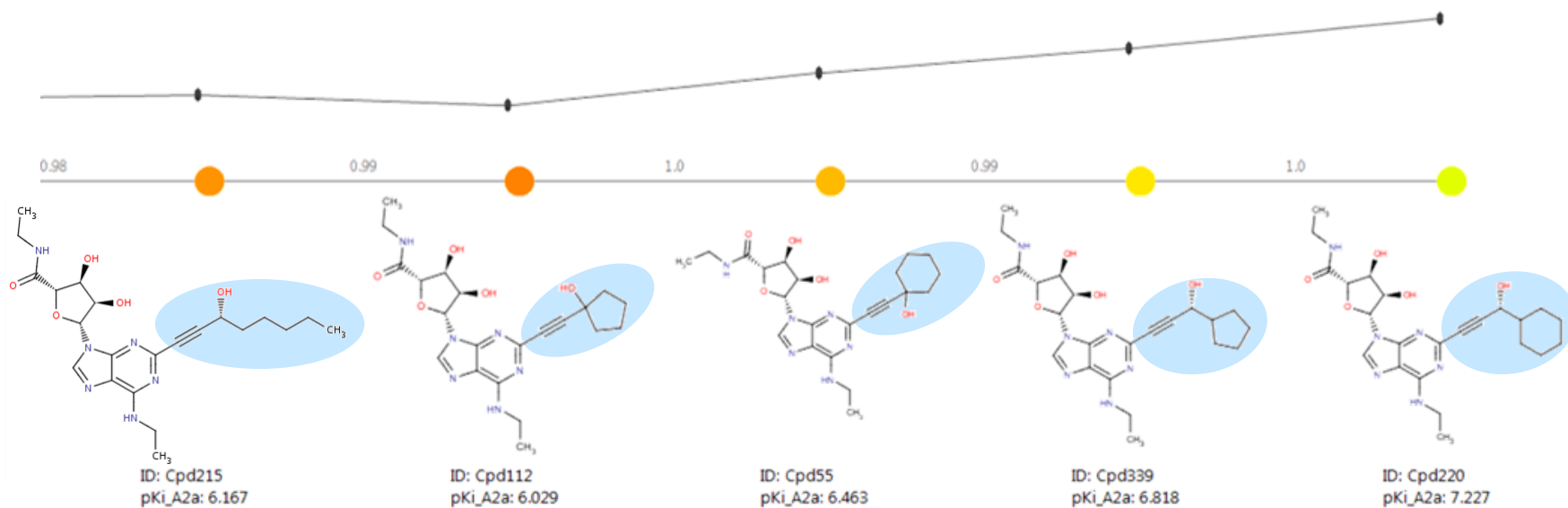
# SAR Path Analysis



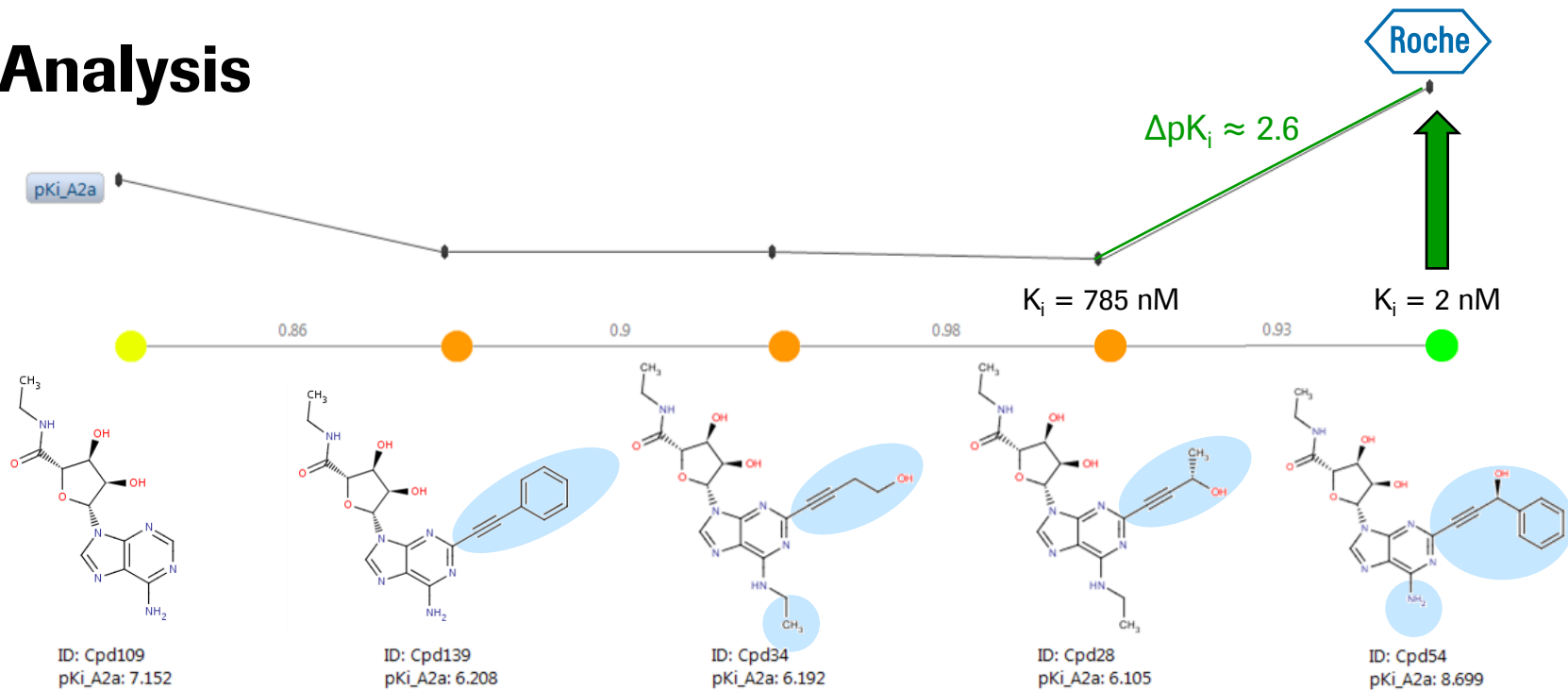
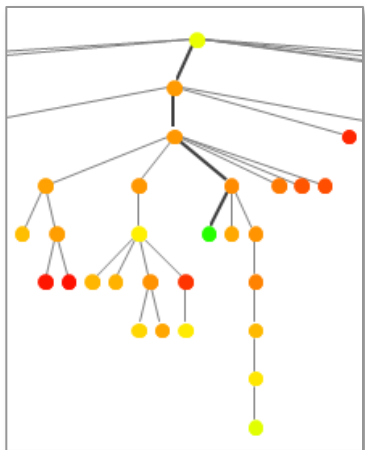
pKi\_A2a



→

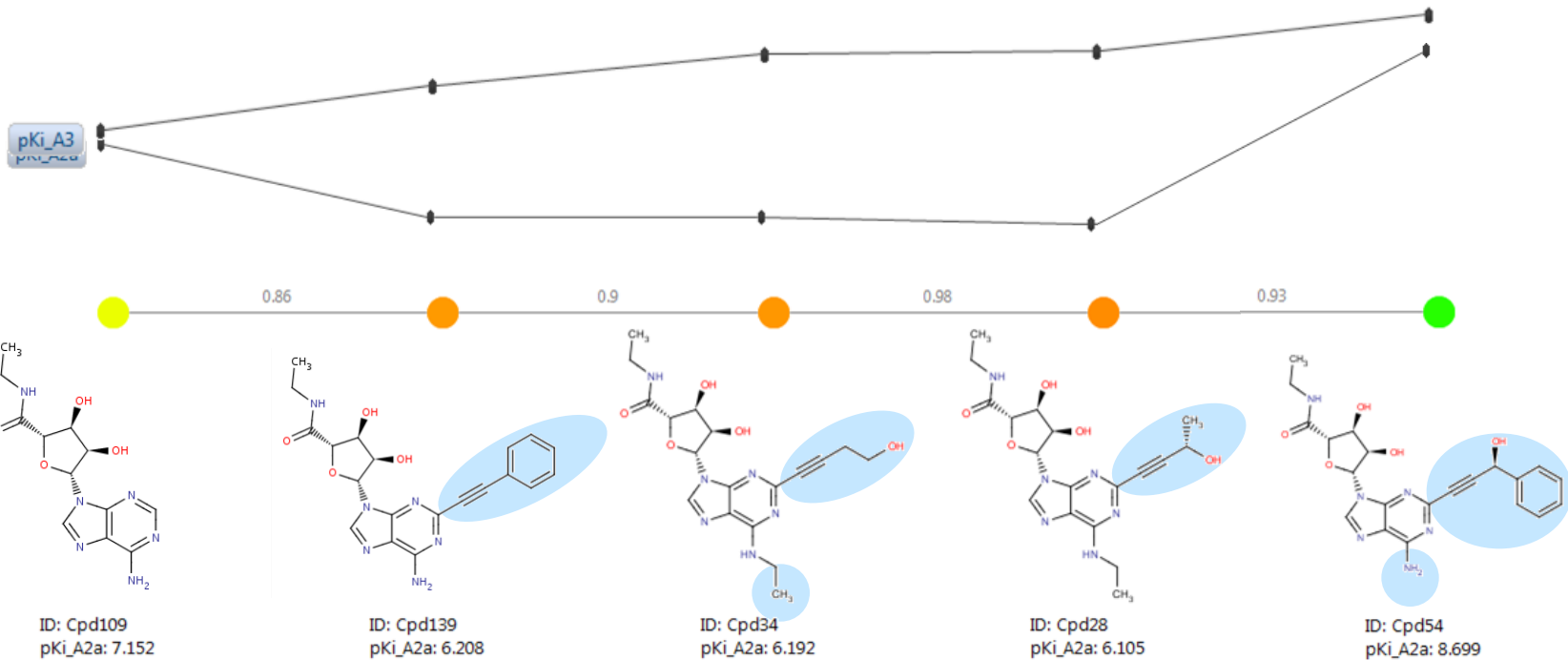
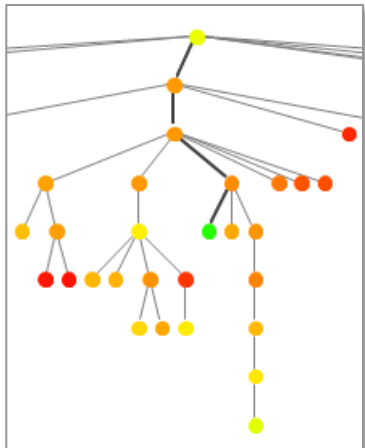


# SAR Path Analysis



Activity cliff

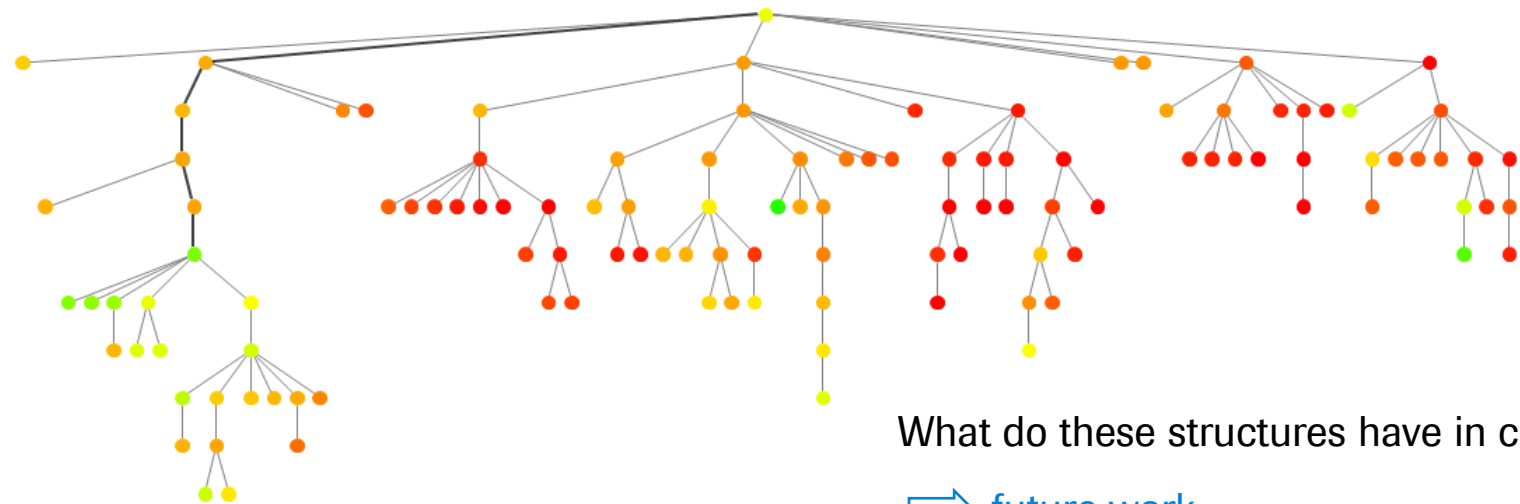
# Multi-Property Display



# SAR Tree Comparison

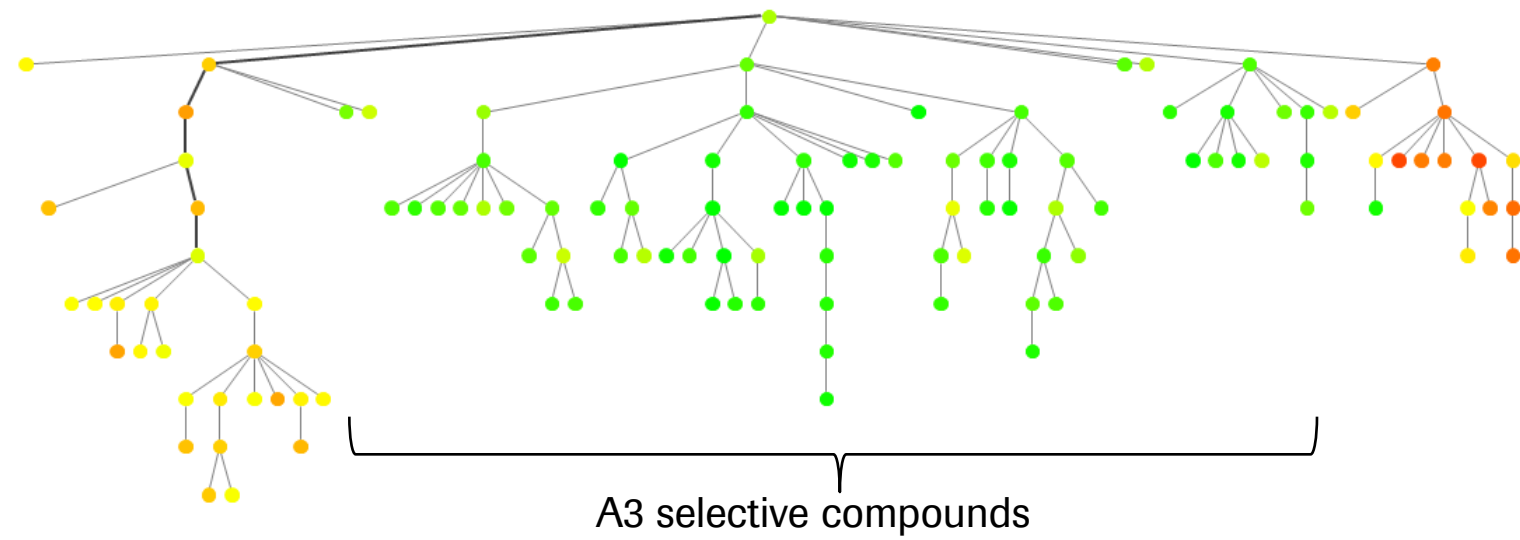
## Selectivity Data

Adenosine receptor A2a



What do these structures have in common?  
⇒ future work

Adenosine receptor A3



A3 selective compounds

# Conclusions

## *Key Principles*

- Keep it simple!
- Flexible and intuitive interface
- Real-time interaction with the data
- *Support*, not try to *dictate* the expert's workflow

# Acknowledgements

- Francesca Milletti, Jörg Degen (Cheminformatics & Statistics)
- Mark Rogers-Evans, Katrin Groebke-Zbinden, Alexander Mayweg, Wolfgang Guba, Thomas Luebbers (Discovery Chemistry)
- Ralph Haffner, Martin Strahm, David Herzig, Venus So (pRED Informatics)
- Martin Stahl (Business Owner)
- Margret Assfalg (Project Sponsor)



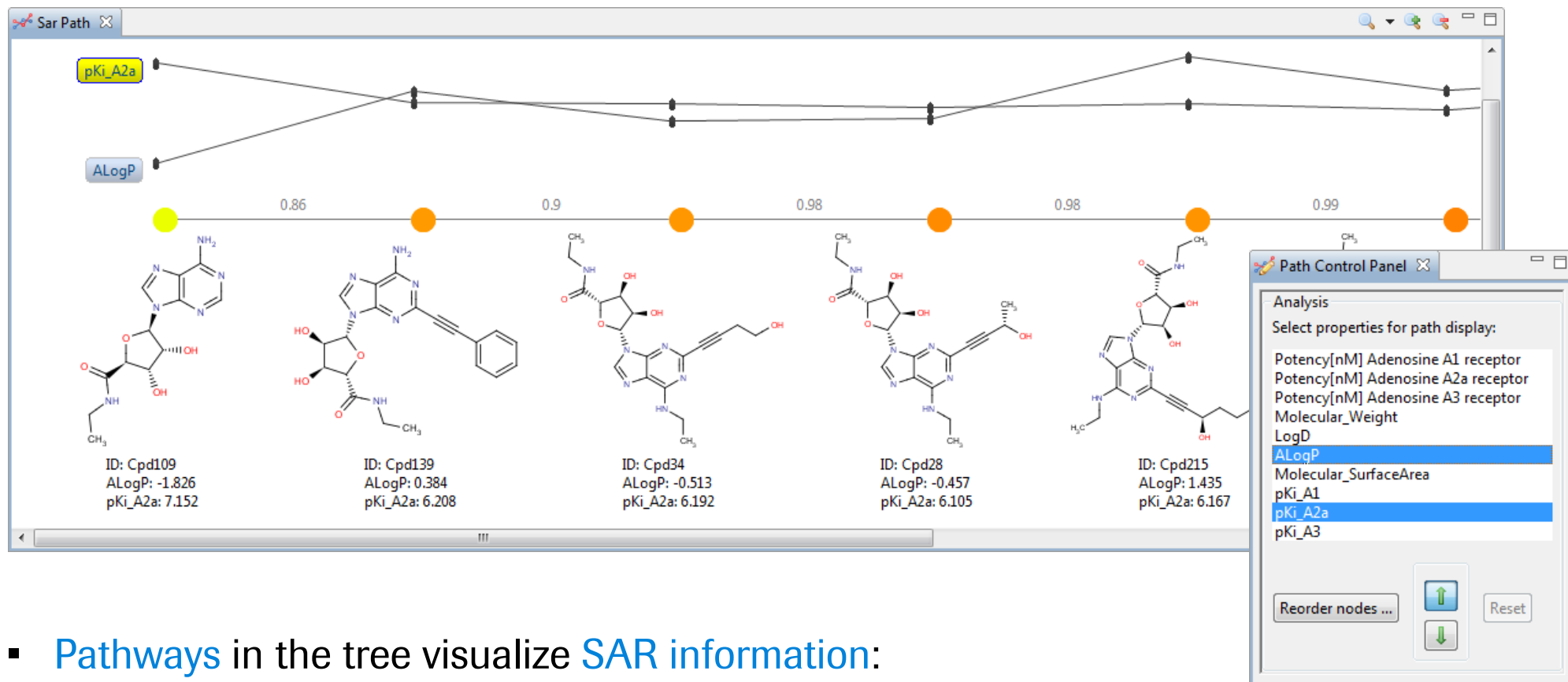
*We Innovate Healthcare*

# Backup



# SAR Tree

## Navigation Along Paths



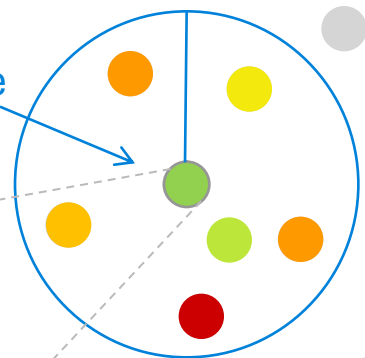
- Pathways in the tree visualize SAR information:
  - Gradual structural changes departing from root molecule along with changes in activity / properties

# SAR Tree

## Tree Construction

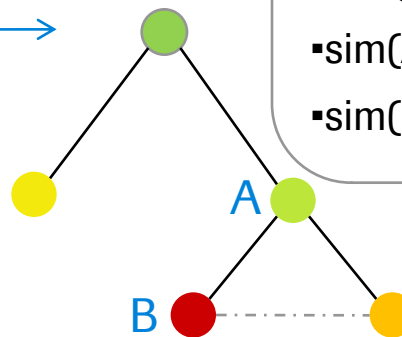
1. Select reference molecule

2. Define similarity radius



Reference molecule

Root



3. Construction rules

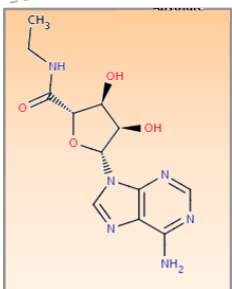
Connect molecule B to A only if

- no other molecule is more similar to B (nearest neighbors) (1)
- $\text{sim}(A,B) > \text{threshold}$
- $\text{sim}(B,\text{root}) < \text{sim}(A,\text{root})$  (2)

4. Tie break rules

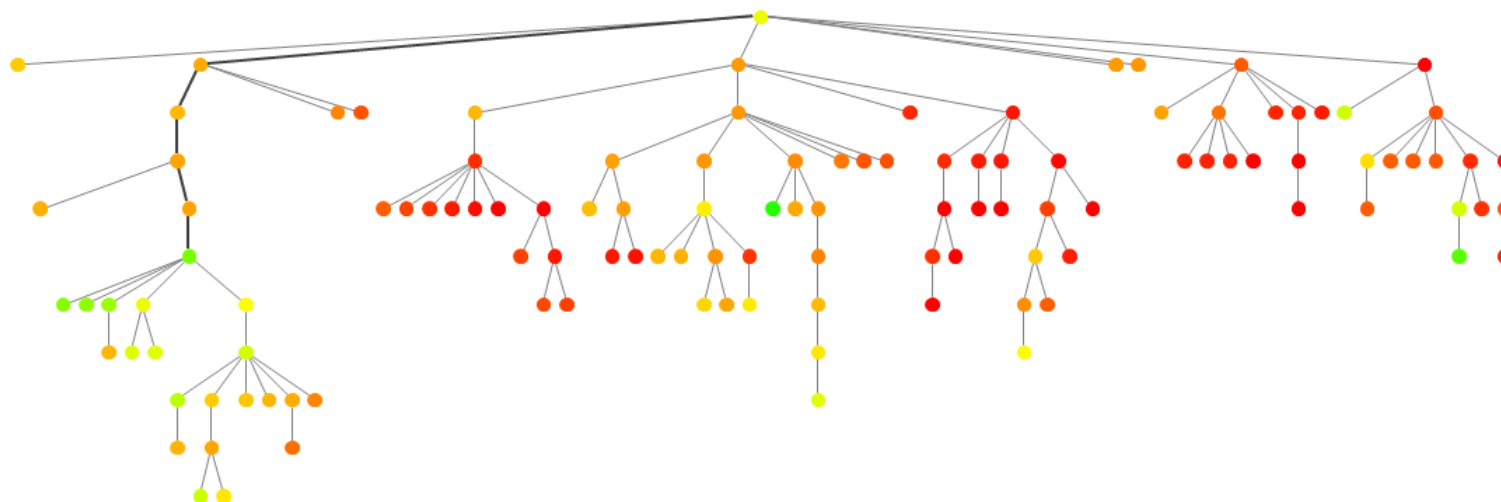
If there are several possible parents for a molecule

- keep only parent(s) with most similar property value(s)
- keep parent that is closest to the root



# Future Work

Adenosine receptor A2a



- Mine data for common patterns
  - Annotate subtrees with common structural features
  - Find substructures that are enriched in good / bad compounds
- Enable data searching & selection
- Modification of tree structure
  - Manual editing (already implemented)
  - Refine tree construction
- Support decision-making process
  - Store & share session information
  - User annotations