Analysing selectivity through multi-dimensional activity cliff analysis

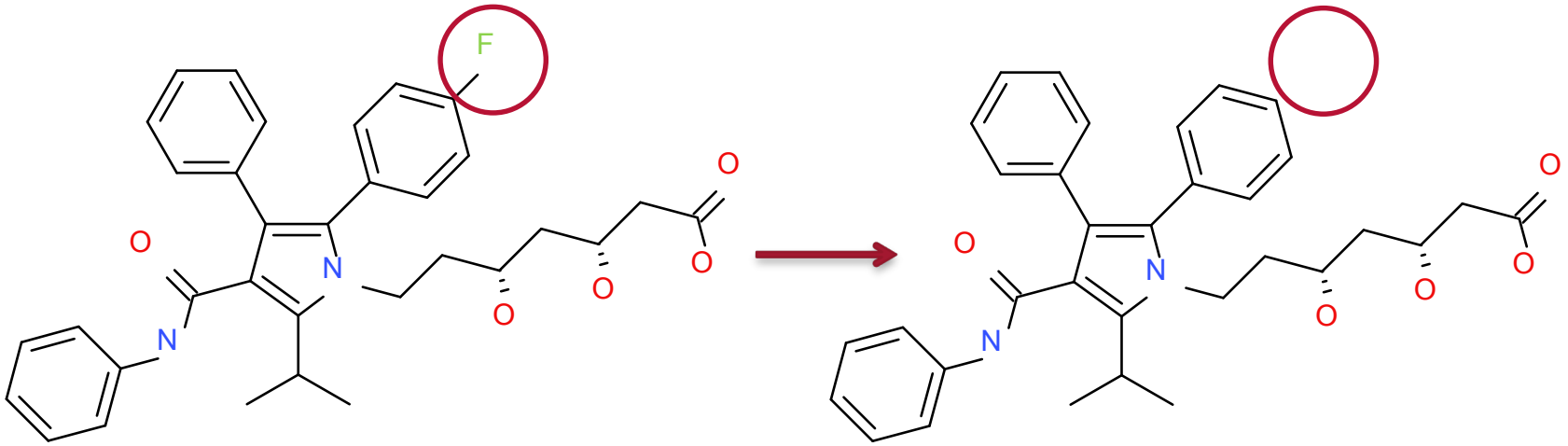
Tim Cheeseright

Cresset summary

- Growing and profitable company
  - 20% year on year growth since 2009
  - 20 People, 12 with PhDs
- Primary market pharmaceutical and biotech R&D
  - Software:
    - 14 of the top 20 pharmaceutical companies use Cresset's technology in their research programmes
  - Consultancy Services:
    - ~200 collaborative projects delivered to global clients
- Secondary markets: agrochemicals, flavours and fragrances, consumer health and fine chemicals

cresset

Drug discovery's similarity hypothesis



> Similar molecules have similar activities

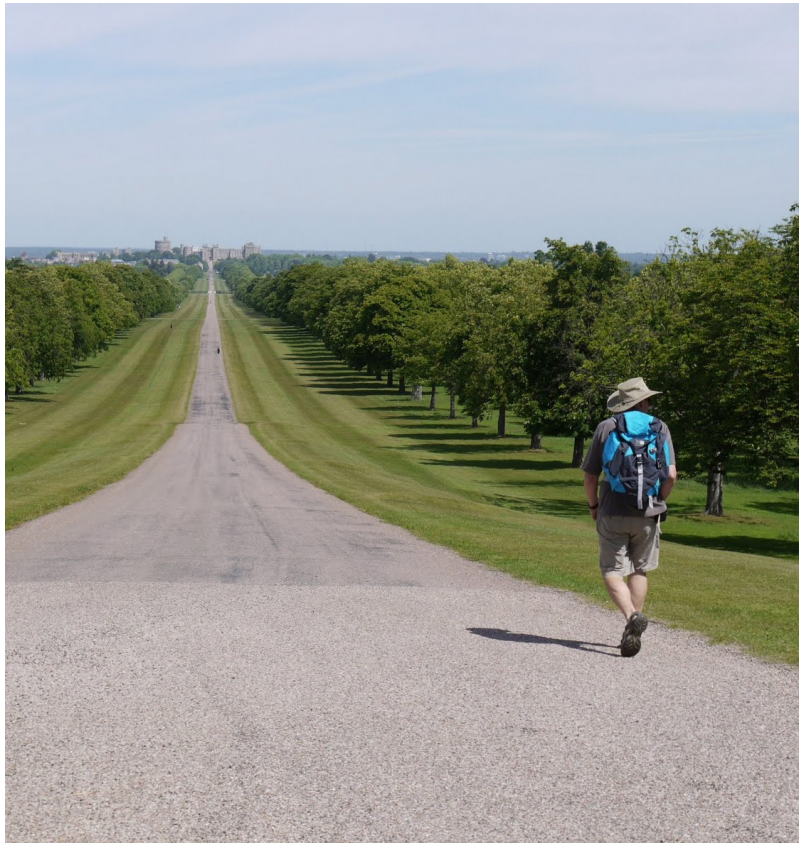
> Small changes lead to small changes

→ QSAR, virtual screening, lead optimization

(Un)Interesting SAR

What about the bits where the similarity hypothesis breaks down?

Nothing happens



Something dramatic happens

Activity Cliffs – interesting regions of SAR

- Many names:
  - Disparity (Merck 1990s)
  - SALI (Guha/Drie 2008)
  - Activity Landscapes
  - Activity Cliffs
- Definition:
  - For each pair of molecules $\kappa = \frac{Act_1 - Act_2}{Distance_{12}}$
- Usually distance = 1 – similarity
  - Similarity from 2D fingerprints, tanimoto etc
- Large K indicates an activity cliff

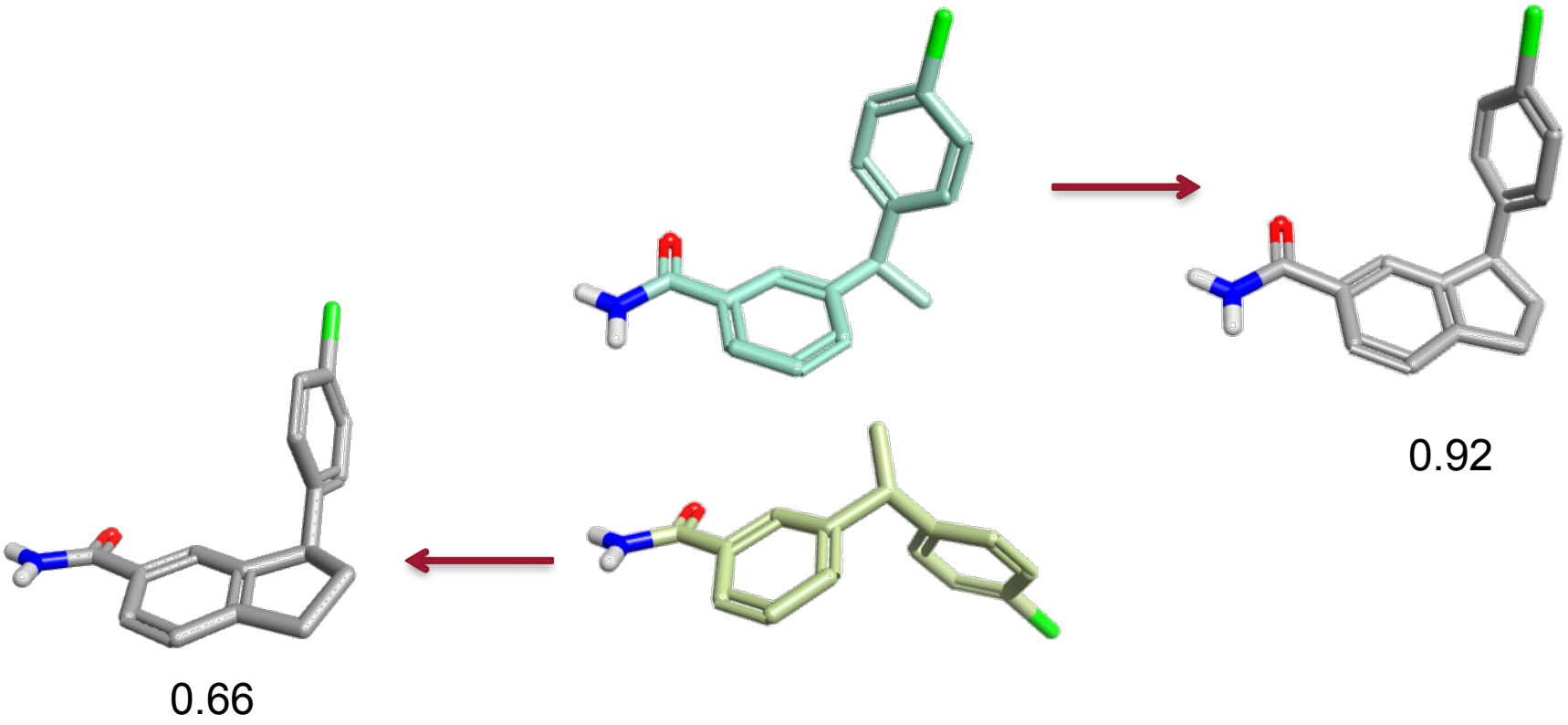
Gaining understanding of Activity Cliffs

- Activity cliffs from 2D similarity highly valuable
- But no explanation for why the cliff is present
- Without an explanation we cannot use the cliff to design new compounds with confidence
- True understanding can come from 3D metrics
  - Shape
  - Electrostatics
- What about using 3D similarity from the outset?

Using 3D similarity

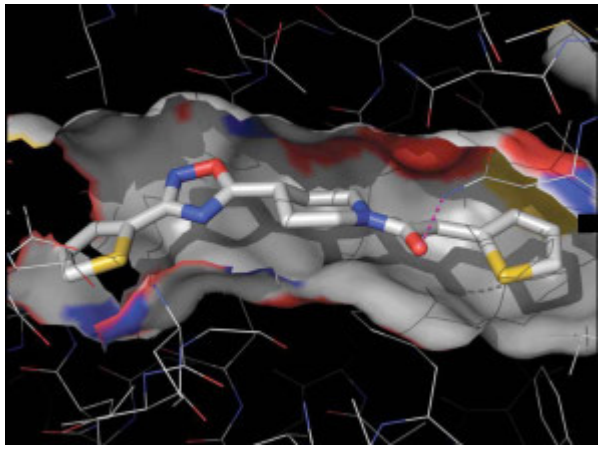
> 2D metrics are easy: 1:1 map to topology

> 3D is defined for **conformers**, not for **molecules**



Context is everything

- Don’t need/want **generic** 3D similarity
  - Have activity context – bound to the protein



- Align all molecules to known bioactive reference conformer
- Provides a conformation context to each molecule

3D disparity

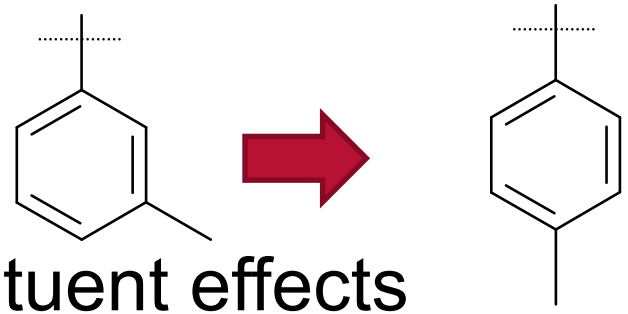
1. Generate conformers
2. Align to reference(s)
3. Calculate 3D similarity matrix on aligned conformations

    What 3D properties do we want to capture?

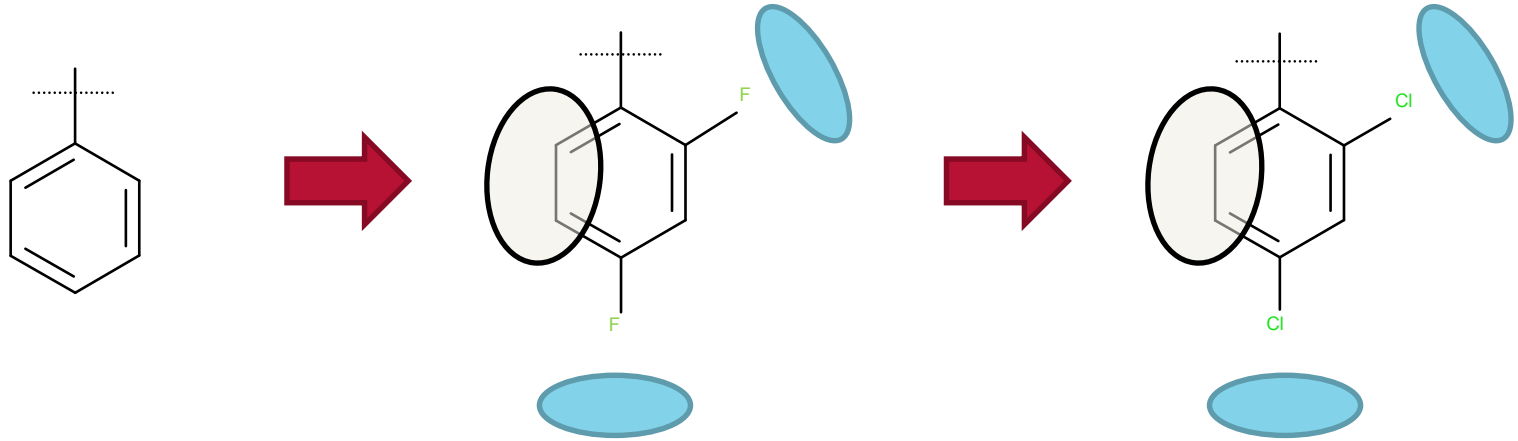
cresset

Properties of a 3D similarity

> Shape / Sterics



> Electrostatics – substituent effects

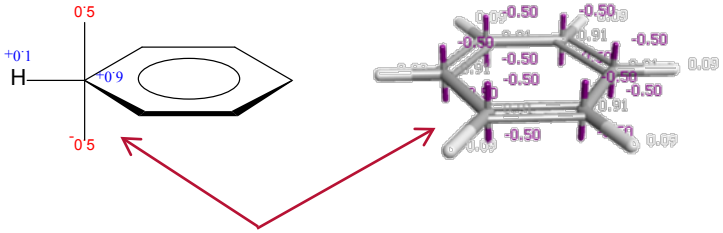


Changes to potential interactions from new atoms

Changes induced in retained portions

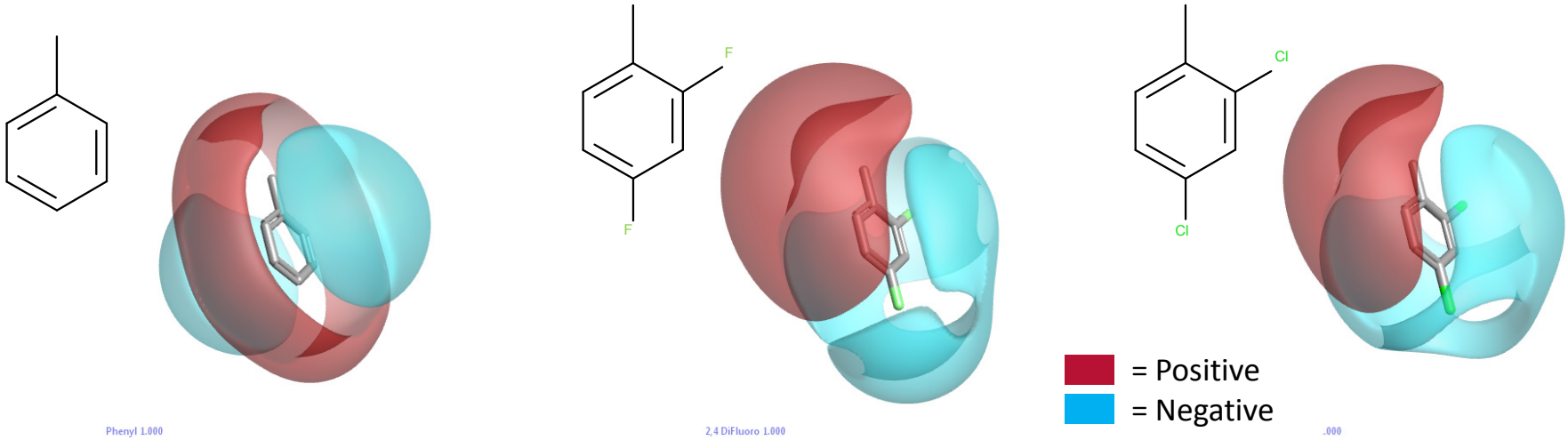
Detailed electrostatics from XED

> eXtended Electron Distribution gives detailed electrostatic interaction patterns

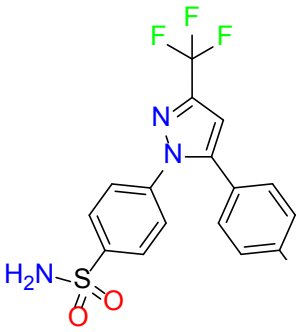


Separation of π- and σ- charges enables modelling of substituent effects

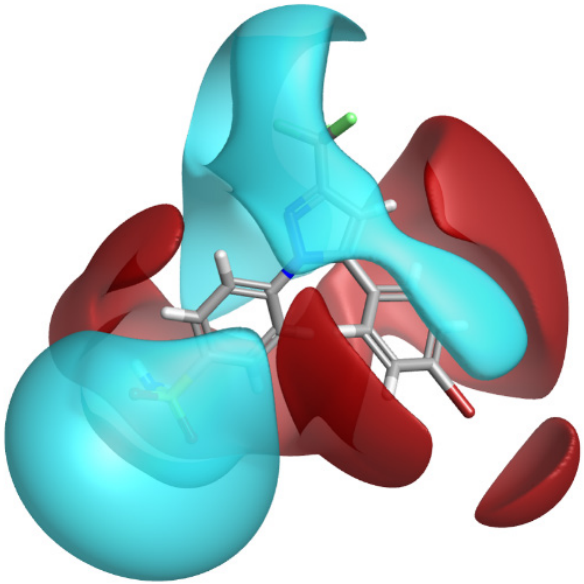
XED adds p-orbitals to get detailed representation of atoms



Field points



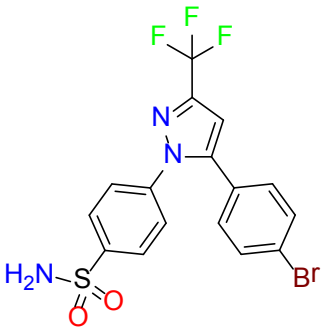
MIP contains too much information to use computationally in a reasonable time



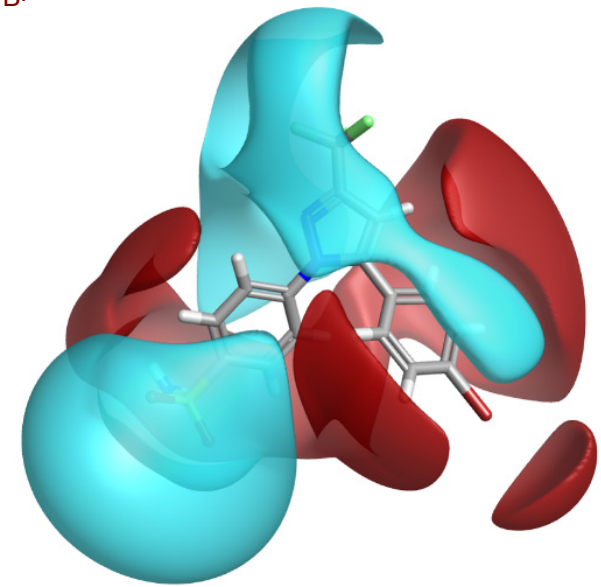
3D Molecular Electrostatic Interaction Potential (MIP)

= Positive
= Negative

Field points

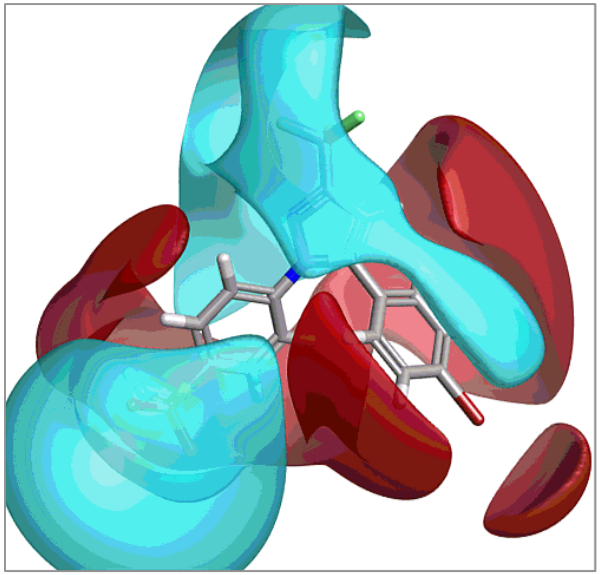


MIP contains too much information to use computationally in a reasonable time



3D Molecular Electrostatic Interaction Potential (MIP)

Field Points provide computationally tractable framework for electrostatic similarity

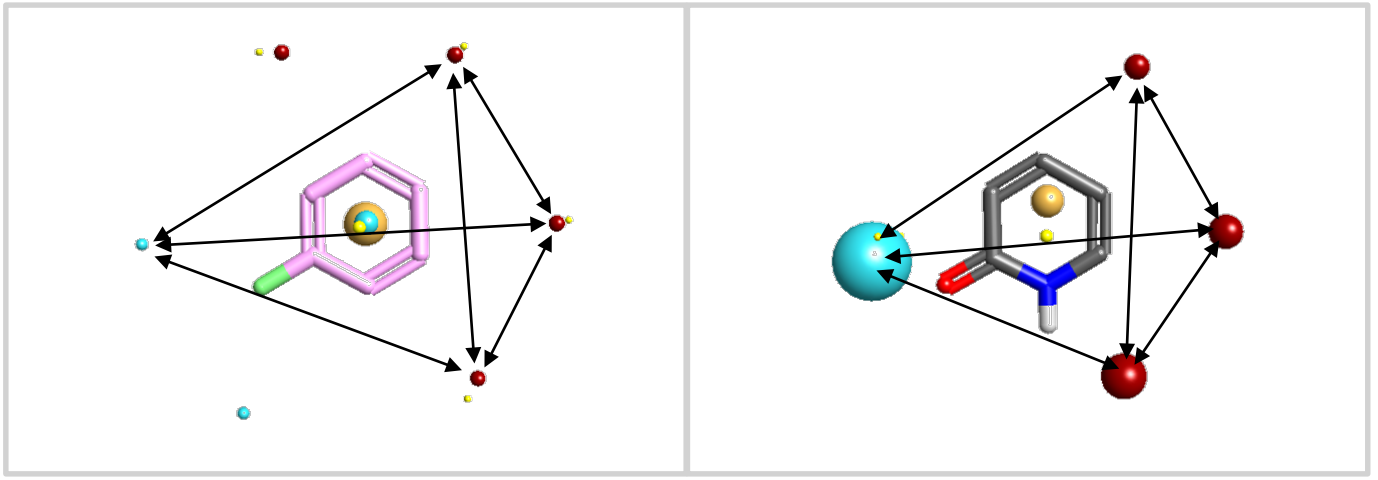


Field Points

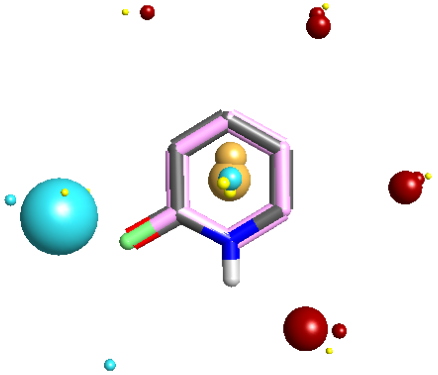
= Positive
= Negative
= Shape
= Hydrophobic

Alignment, scoring and comparisons

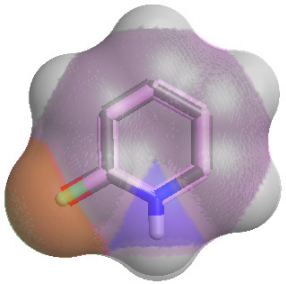
Clique based alignment



Fields
0.66



Cheeseright et al,
J. Chem Inf. Mod., 2006, 665



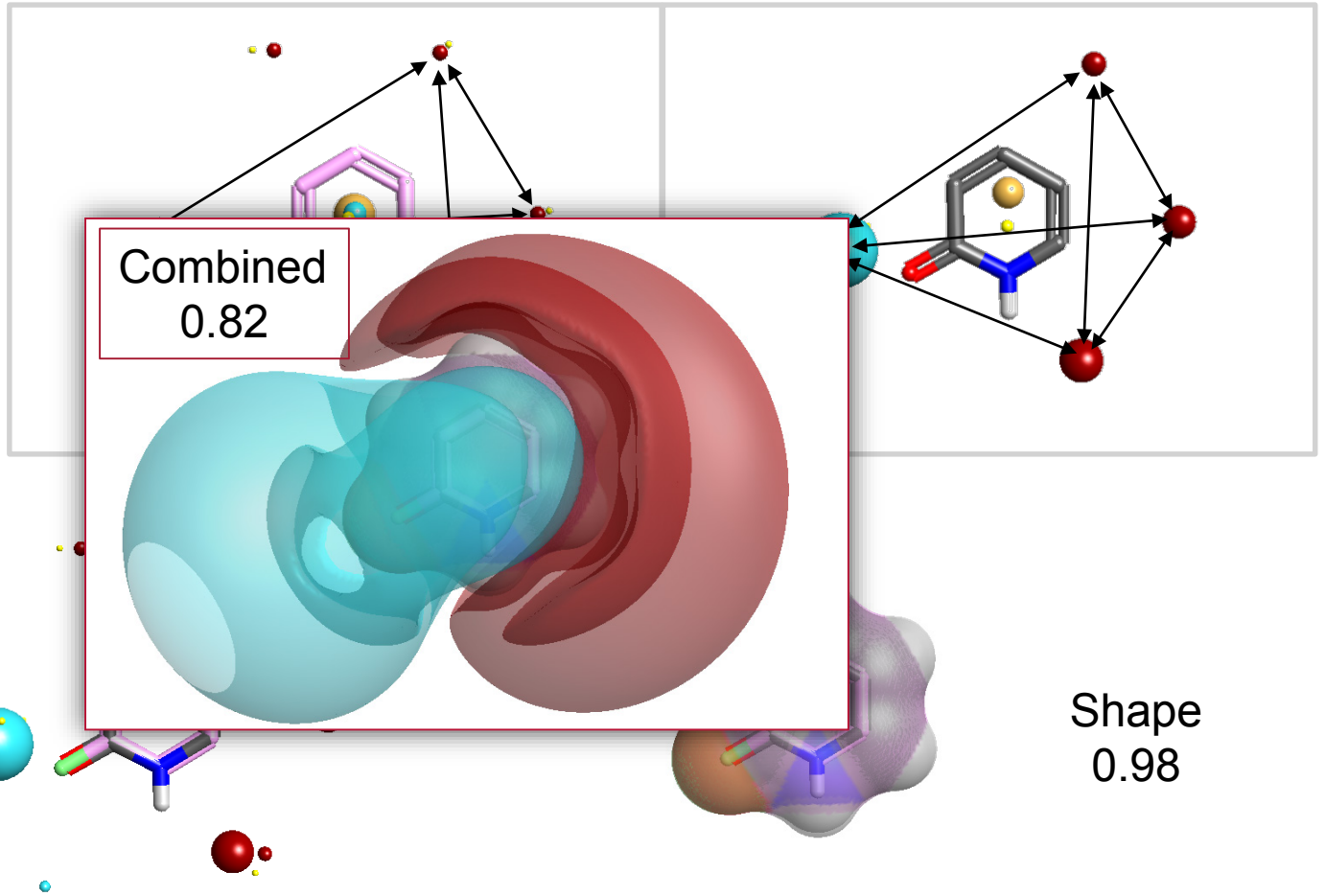
Shape
0.98

Grant, Gallardo, Pickup,
J. Comp. Chem., 1996, 1653

cresset

Alignment, scoring and comparisons

Clique based alignment



Fields
0.66

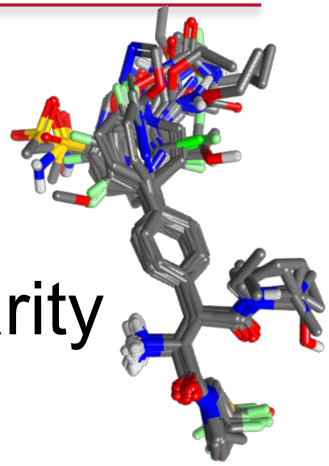
Shape
0.98

Cheeseright et al,
J. Chem Inf. Mod., 2006, 665

Grant, Gallardo, Pickup,
J. Comp. Chem., 1996, 1653

cresset

3D disparity workflow



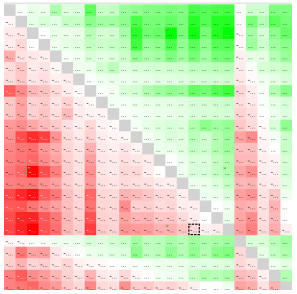
1. Generate conformers
2. Align to reference(s)
3. Calculate 3D shape & electrostatic similarity matrix
   - > Allow small movements
4. Calculate disparity matrix from similarity numbers

$$\frac{Act_1 - Act_2}{Distance_{12}}$$

   - > Similarity cutoff of 0.95 (Distance cutoff of 0.05)
5. Visualize
   - > Difficult – 100 molecules gives 4950 pairs!

Visualization

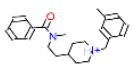
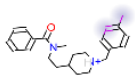
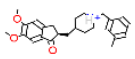
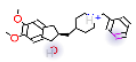
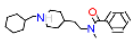
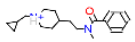
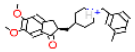
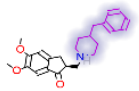
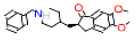
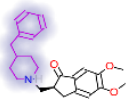
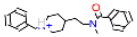
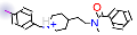
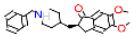
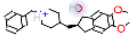
- Existing ways to visualize
  - Table & Matrix views



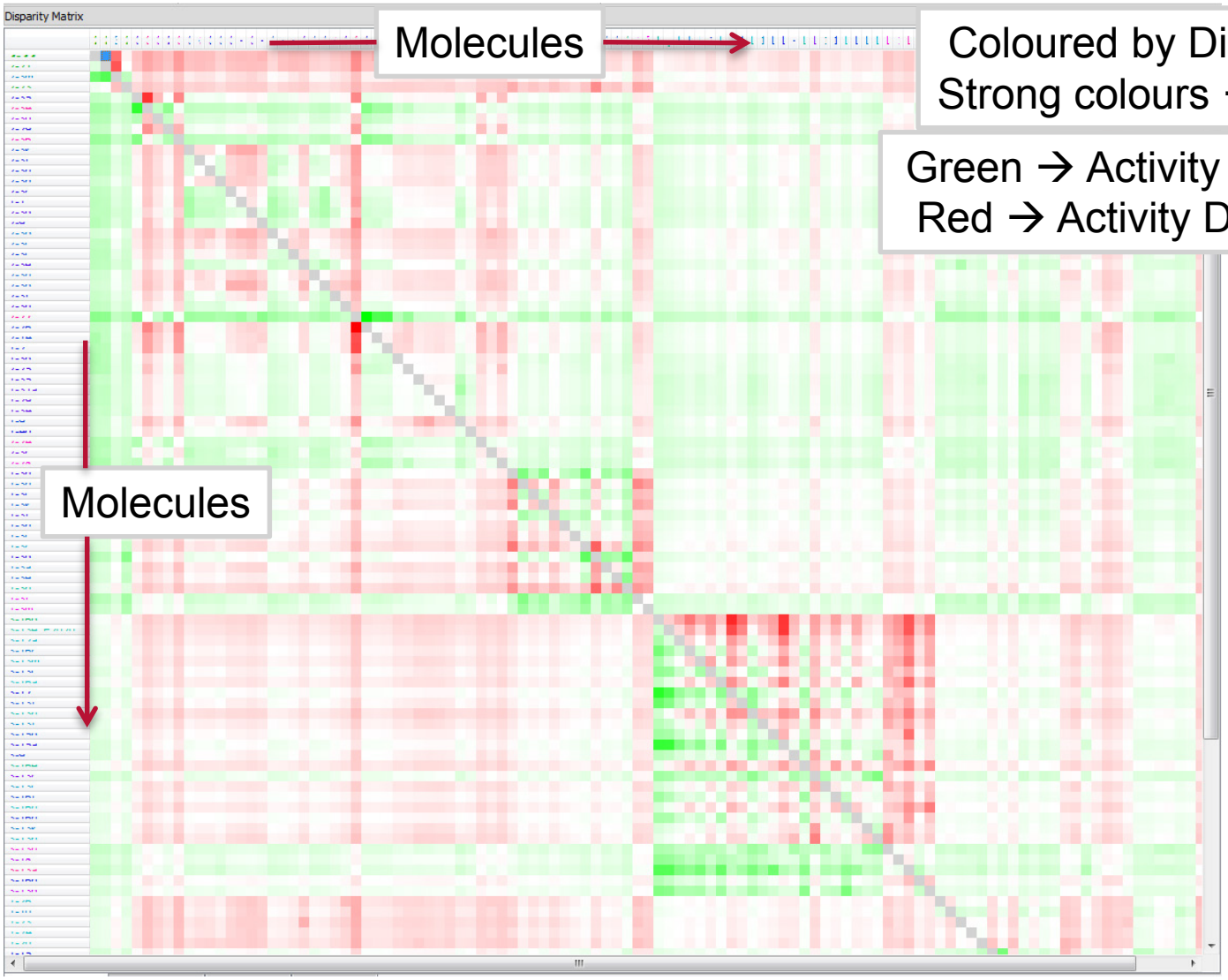
cresset

Top pairs table

Disparity Table

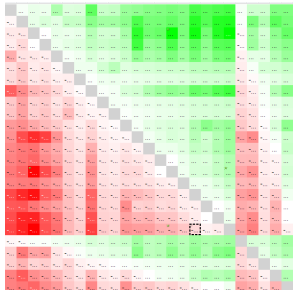
Good	Good Activity	Bad	Bad Activity	Disparity	Similarity	Fav	Δ Activity	Δ LE	Δ LLE	Δ TPSA	Δ SlogP	2D Sim
2-26 	6.84	2-27 	4.39	-49	0.951	☆	-2.45	-0.094	-0.094	0	0	0.776
3-16b 	8.7	3-17 	6.52	-41.2	0.947	☆	-2.18	-0.067	-0.07	3.2	-0.1	0.633
2-35 	6.39	2-34 	4.42	-39.4	0.959	☆	-1.97	-0.055	-0.026	0	-1.2	0.77
3-16b 	8.7	3-15a 	6.32	-38.7	0.939	☆	-2.38	-0.074	-0.057	0	-0.7	0.701
3-13e_E2020 	8.24	3-15a 	6.32	-38.4	0.952	☆	-1.92	-0.069	-0.056	0	-0.3	0.835
1-2 	6.77	2-27 	4.39	-34.5	0.931	☆	-2.38	-0.102	-0.108	0	0.3	0.791
3-13e_E2020 	8.24	3-17 	6.52	-34.4	0.963	☆	-1.72	-0.061	-0.07	3.2	0.2	0.762

Disparity matrix

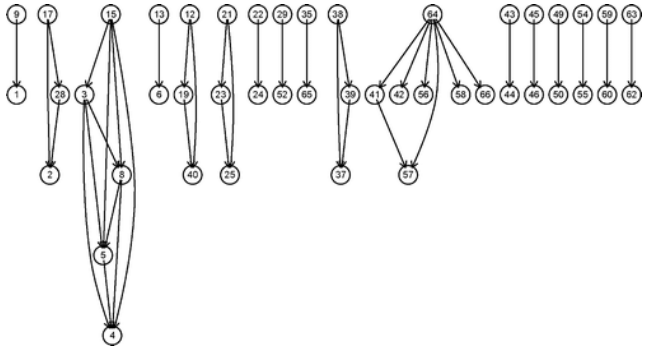


Visualization

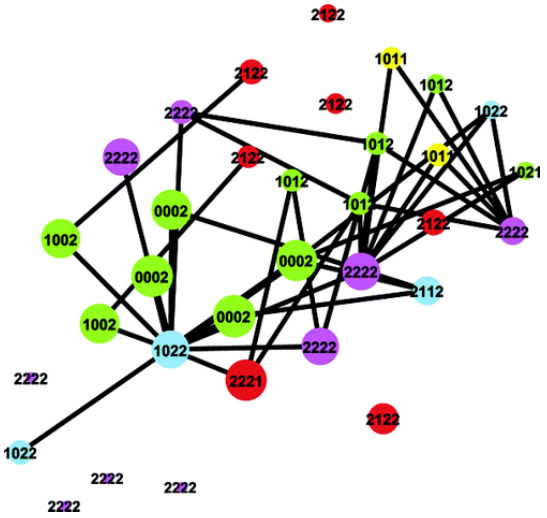
- Existing ways to visualize
  - Table & Matrix views



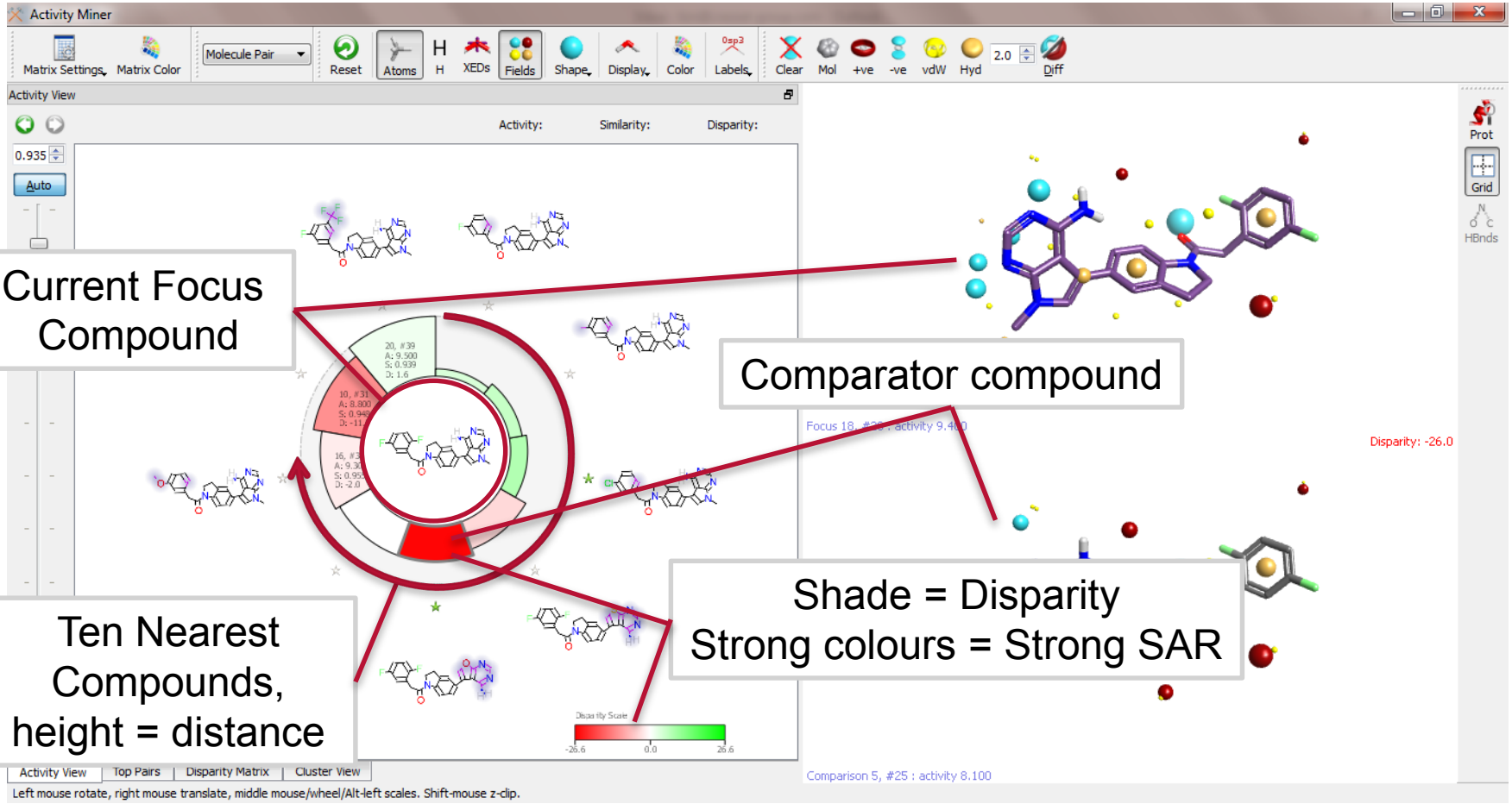
  - Graph view (Guha/van Drie 2008)



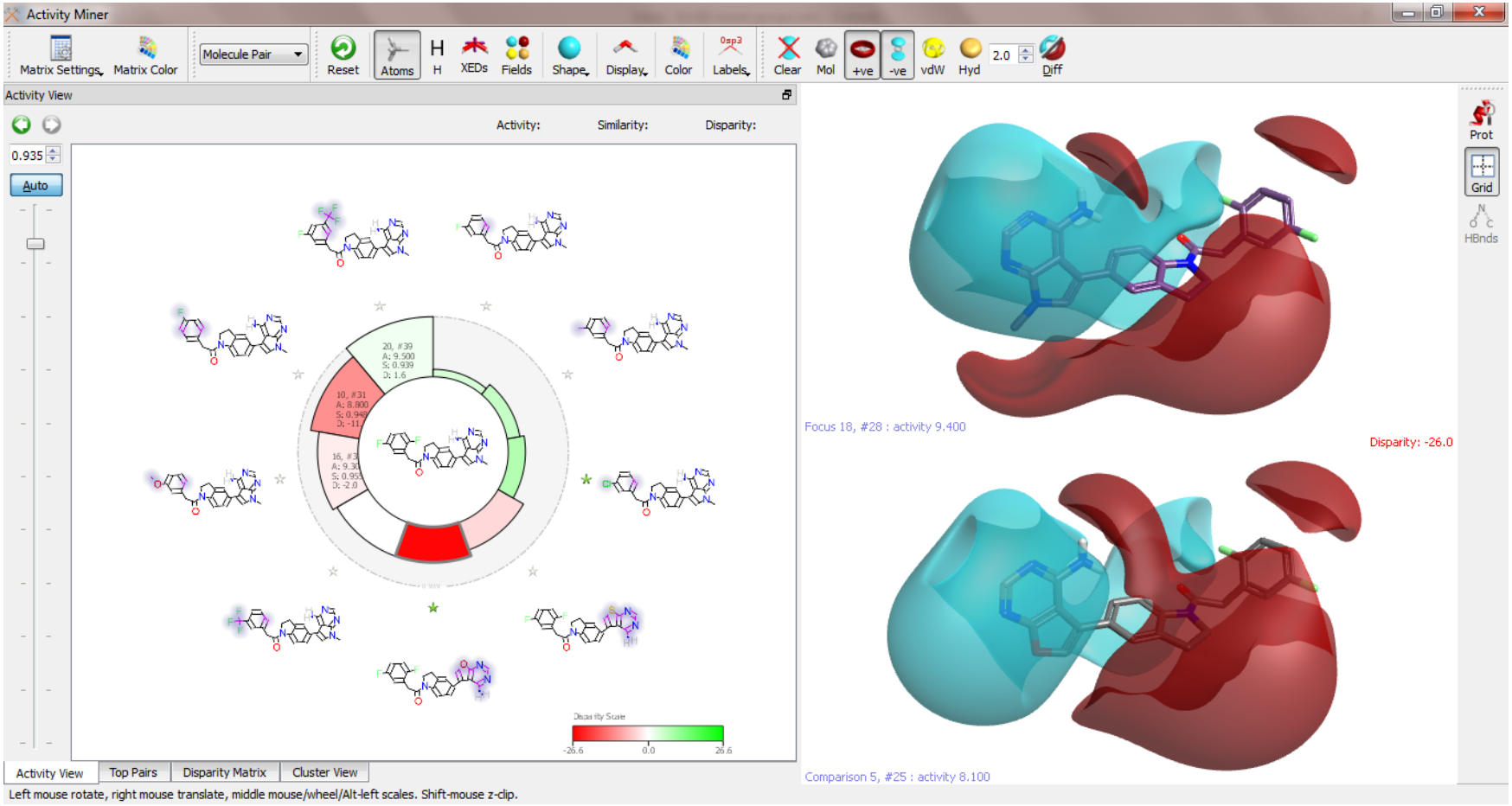
  - Activity landscapes (Bajorath)



Activity View

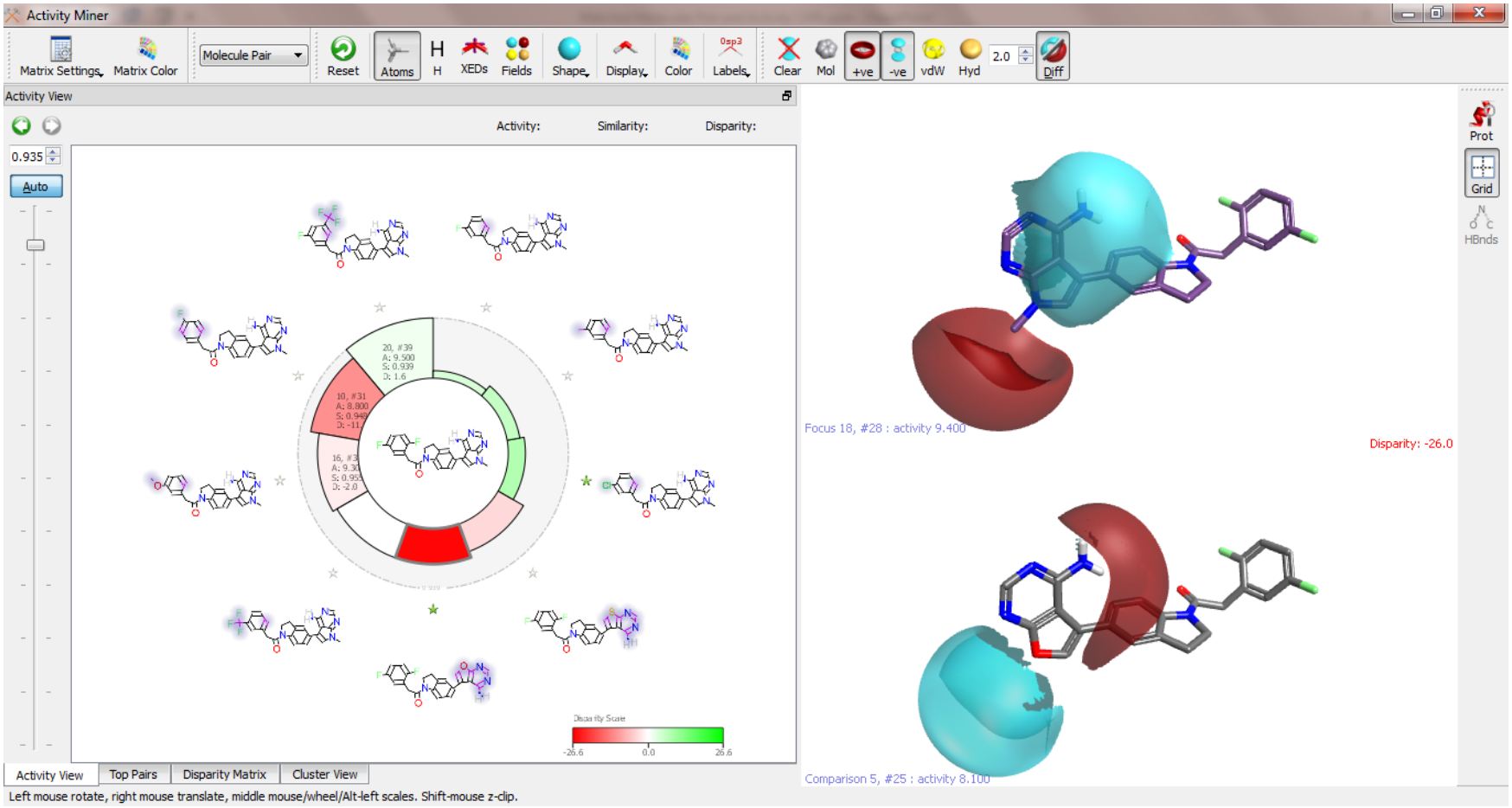


Electrostatic comparison



Electrostatic comparison

Difference plot – Regions where each molecule has stronger electrostatics



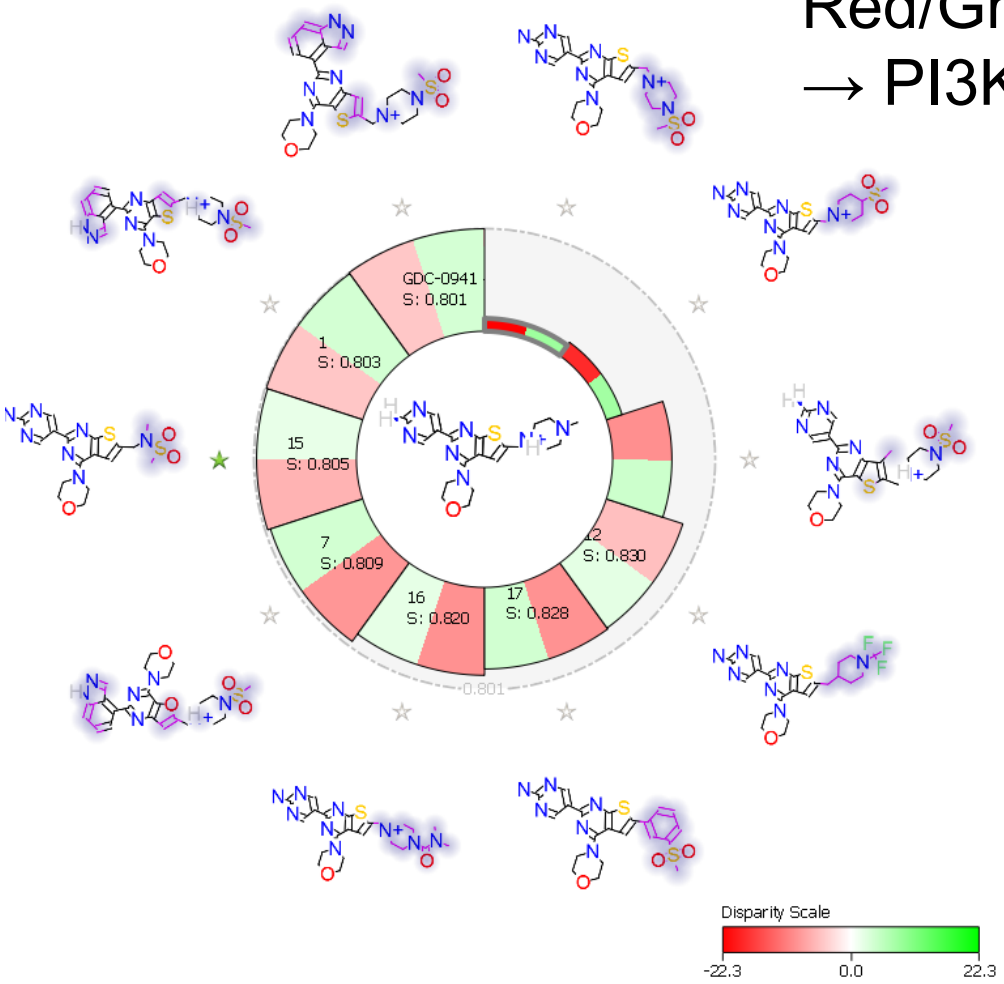
Selectivity Cliffs

- Selectivity often as important as potency
- Look at what structural changes caused large changes in selectivity
- Use Selectivity Endpoint as Activity?

$$\kappa \approx \frac{\Delta\, Selectivity}{\Delta\, Structure} = \frac{\left(\frac{Activity_\beta}{Activity_a}\right)_A - \left(\frac{Activity_\beta}{Activity_a}\right)_B}{(1 - Similarity)}$$

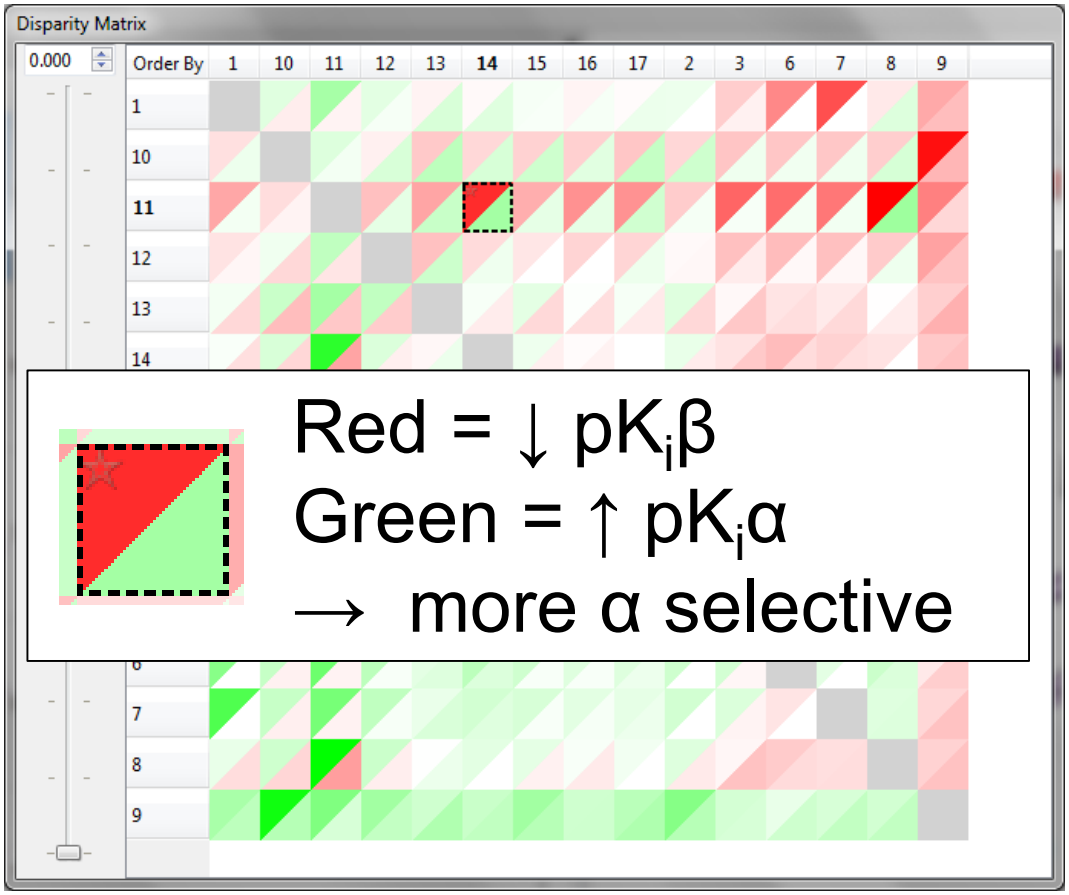
- What about 3 activities?
- How would we visualize that?

Activity View – 2 activities



Red/Green = ↓ $pK_i\beta$, ↑ $pK_i\alpha$
→ PI3Kα selective

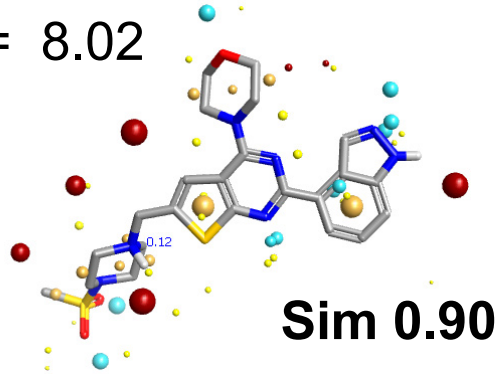
Selectivity matrices – 2 activities



GDC-0941

$pK_i\alpha$ = 9.06

$pK_i\beta$= 8.02

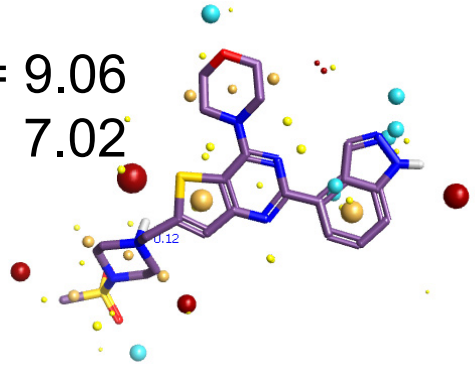


Sim 0.90

6

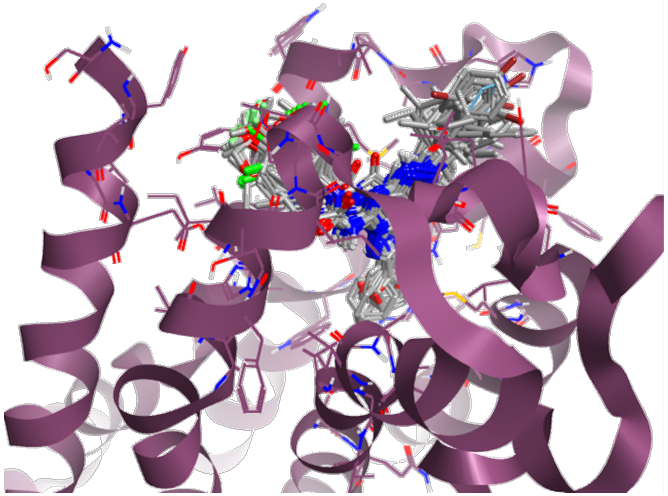
$pK_i\alpha$ = 9.06

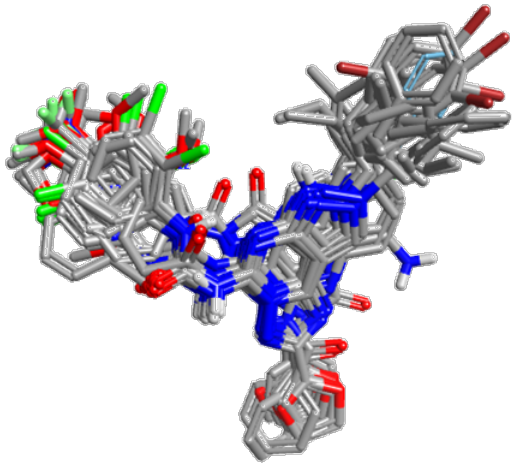
$pK_i\beta$= 7.02



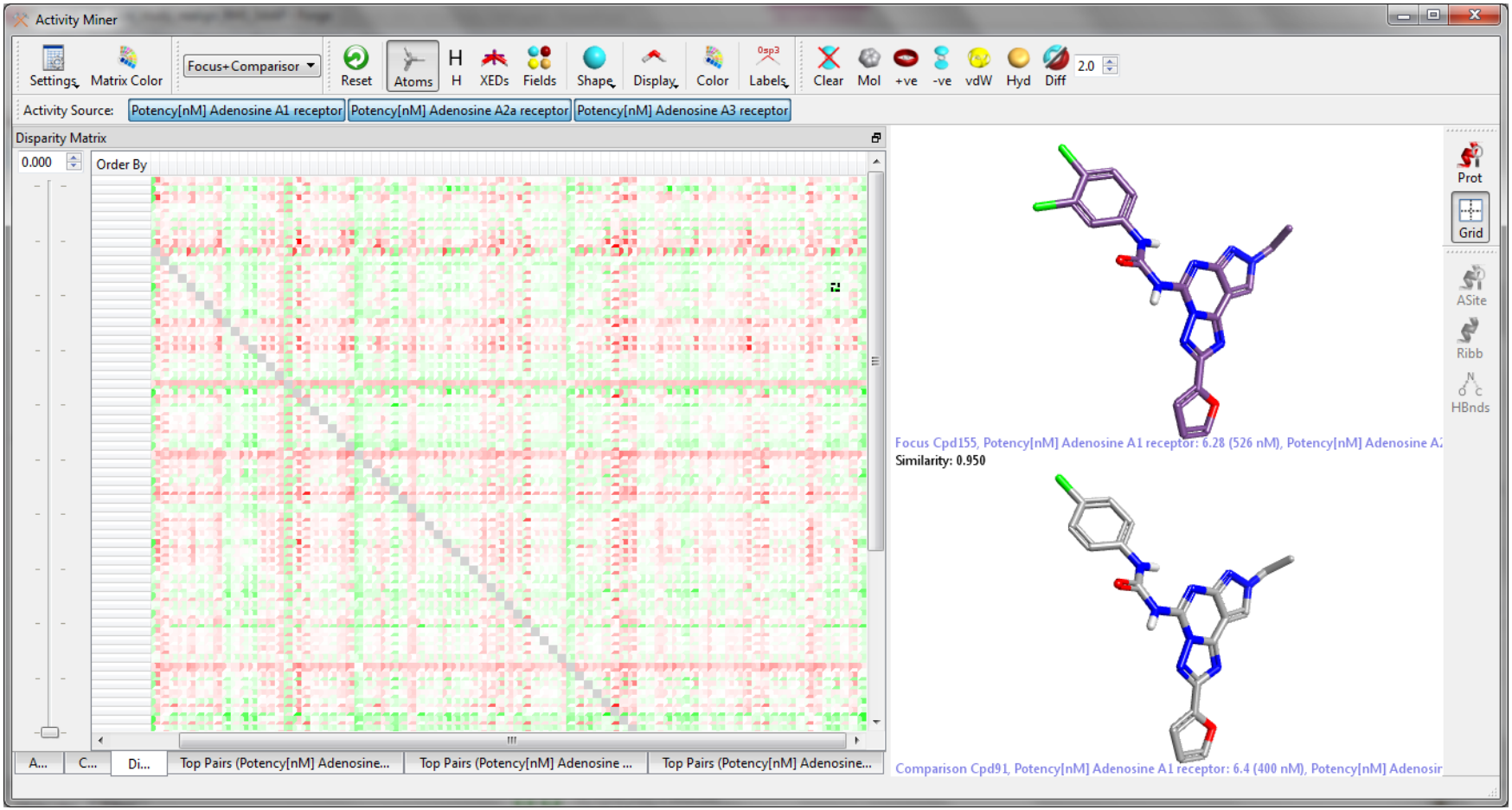
Application to Adenosine Receptor Antagonists

- Data set from Bajorath J. Chem. Inf. Model 51 258-266 2011
- 3 Activities – A1, A2a, A3 receptors
- Ligands aligned to x-ray structures 3PWH, 3EML
- 89 cmpd sub-set with high 3D similarity (>0.7)

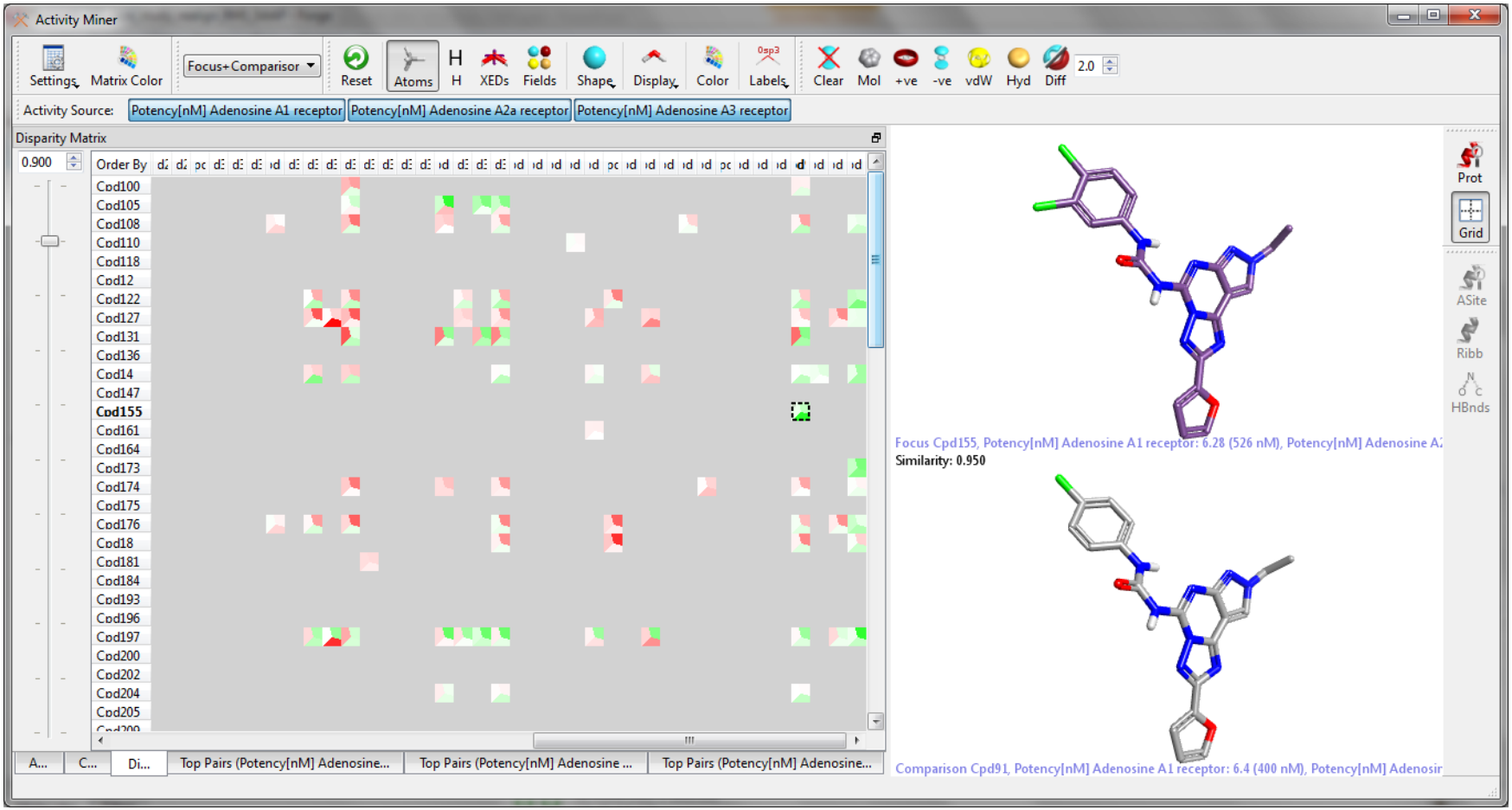




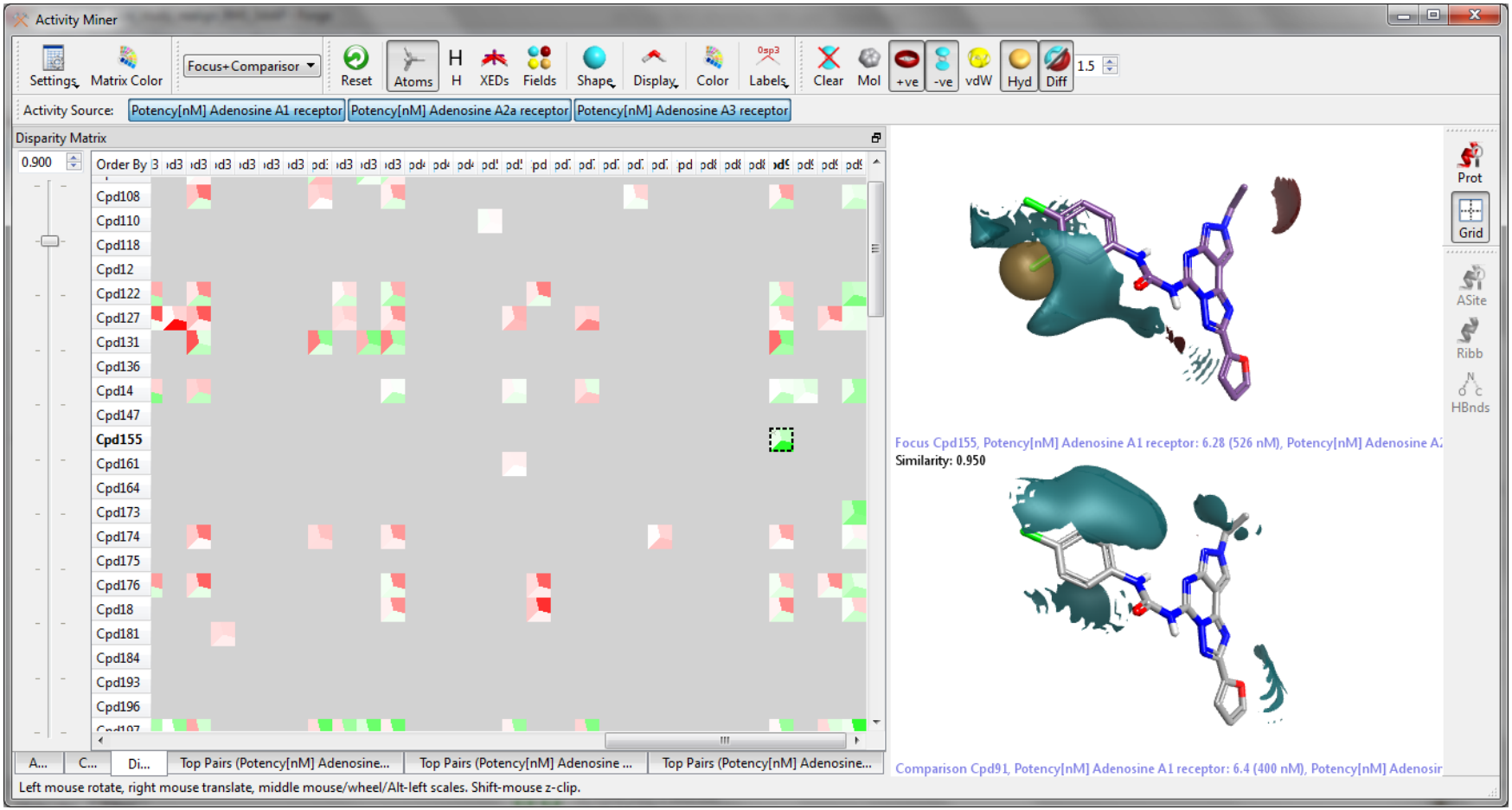
Disparity Matrix – 11,748 data points



Disparity Matrix – focus on highly similar pairs



Why?

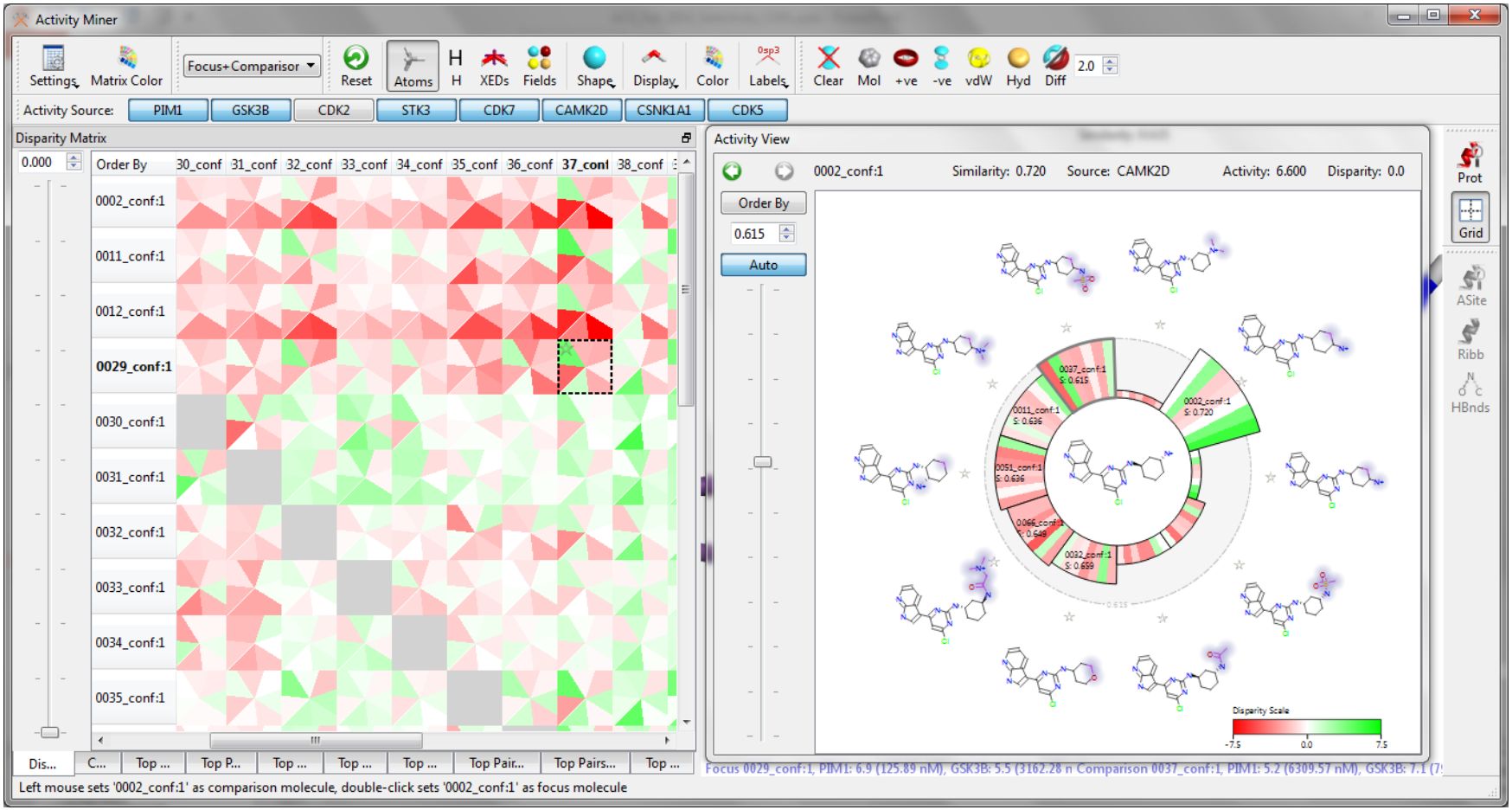

Activity Miner
Settings
Matrix Color
Focus+Comparisor
Reset
Atoms
H
XEDs
Fields
Shape
Display
Color
Labels
Clear
Mol
+ve
-ve
vdW
Hyd
Diff
1.5
Activity Source:
Potency[nM] Adenosine A1 receptor
Potency[nM] Adenosine A2a receptor
Potency[nM] Adenosine A3 receptor
Disparity Matrix
0.900
Order By
Cpd108
Cpd110
Cpd118
Cpd12
Cpd122
Cpd127
Cpd131
Cpd136
Cpd14
Cpd147
Cpd155
Cpd161
Cpd164
Cpd173
Cpd174
Cpd175
Cpd176
Cpd18
Cpd181
Cpd184
Cpd193
Cpd196
Prot
Grid
ASite
Ribb
HBnds
Focus Cpd155, Potency[nM] Adenosine A1 receptor: 6.28 (526 nM), Potency[nM] Adenosine A
Similarity: 0.950
Comparison Cpd91, Potency[nM] Adenosine A1 receptor: 6.4 (400 nM), Potency[nM] Adenosir
A...
C...
Di...
Top Pairs (Potency[nM] Adenosine...
Top Pairs (Potency[nM] Adenosine ...
Top Pairs (Potency[nM] Adenosine...
Left mouse rotate, right mouse translate, middle mouse/wheel/Alt-left scales. Shift-mouse z-clip.

Limitations

- 2 Activities work well
- 3 is OK

- 7 is too many!

Limitations



Conclusions

- Activity Cliff/Disparity analysis provides quick insights into SAR
  - Focus on understanding the reason for a cliff
  - Drive design decisions
- Multiple ways to navigate the data
  - Compound focus
  - Most significant changes
  - Global overview
  - Cluster analysis
- 2D and 3D both useful
  - 2D provides insights into conformational changes
  - 3D provides insights into electrostatic effects
- Visualizing multiple activities simultaneously allows selectivity analysis
  - Large amounts of data difficult to visualize

Acknowledgements

> Mark Mackey
> Nigel Palmer
> Rae Lawrence
> Susana Tomasio
> Giovanna Tedesco

Thank you!

Questions Welcomed



Follow Cresset


tim@cresset-group.com