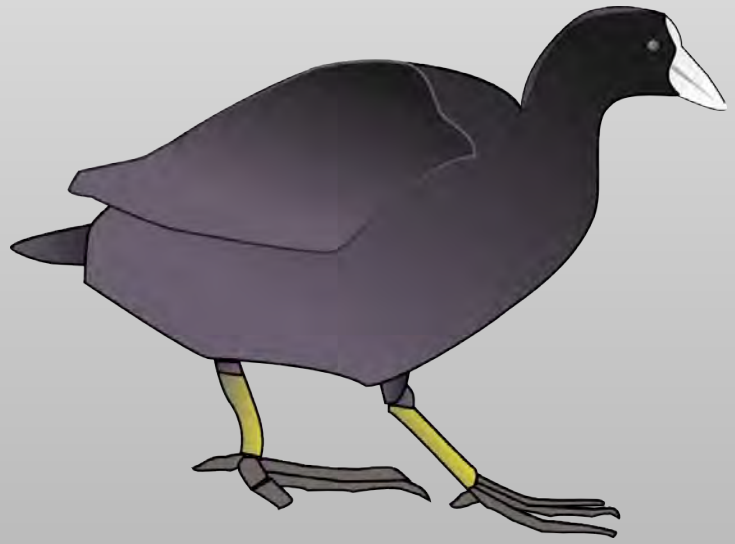
Ligands, Validation, … and more model-building in *Coot*



Bernhard Lohkamp

Karolinska Institutet
Bernhard.Lohkamp@ki.se

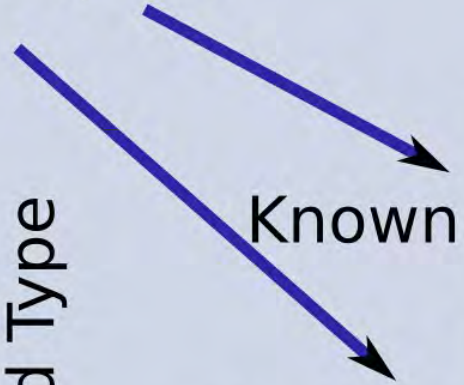
Tsukuba PF November 2014

Ligands in Coot

- Ligand fitting
- Importing and building ligand from scratch
  - PRODRG, LIBCHECK
- Representation
- Analysis
- Validation

Ligand building

Optimize with "ligand expert"options



Ligand Site

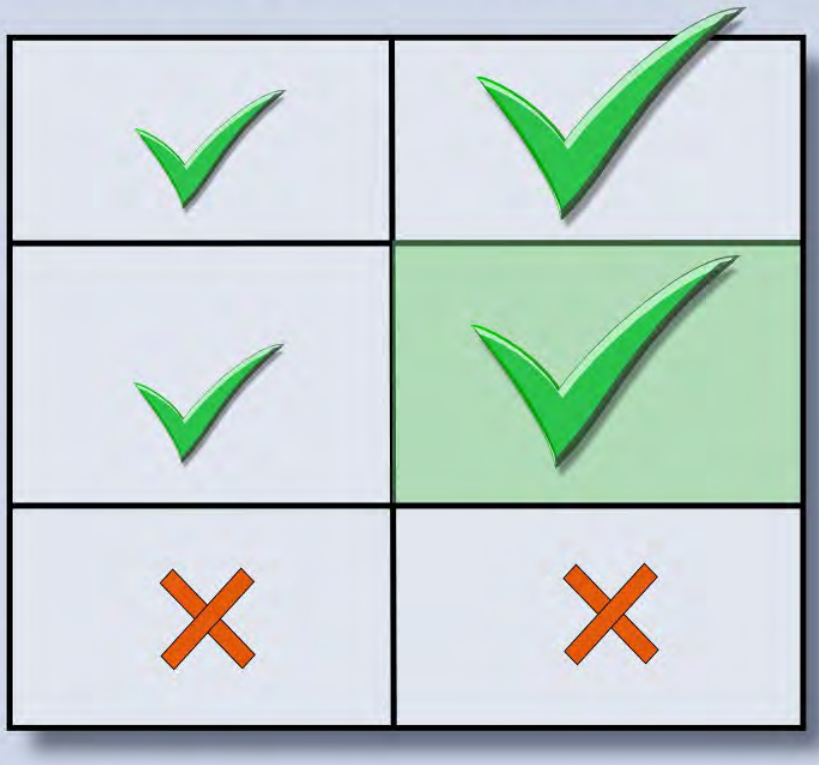
Known Unknown

Ligand Type

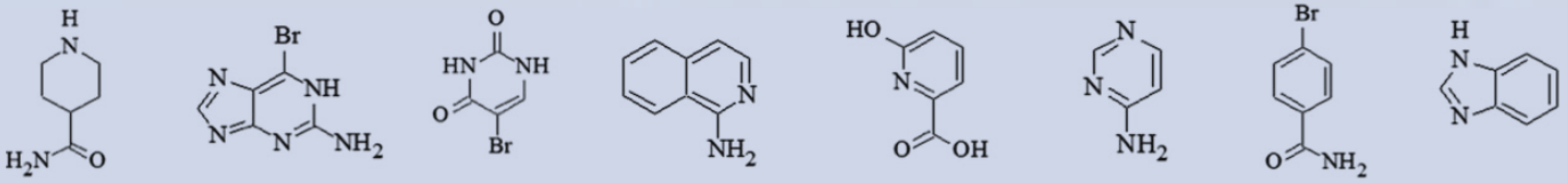
Known

Cocktail

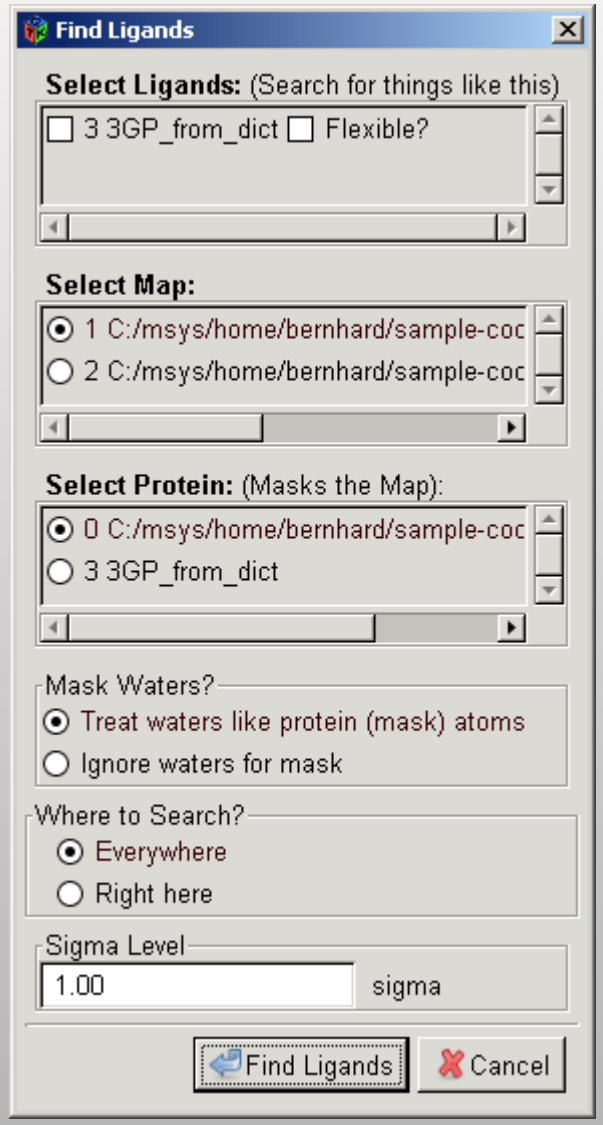
Unknown



Cocktail Examples



Ligand search in Coot



Ligand Fitting

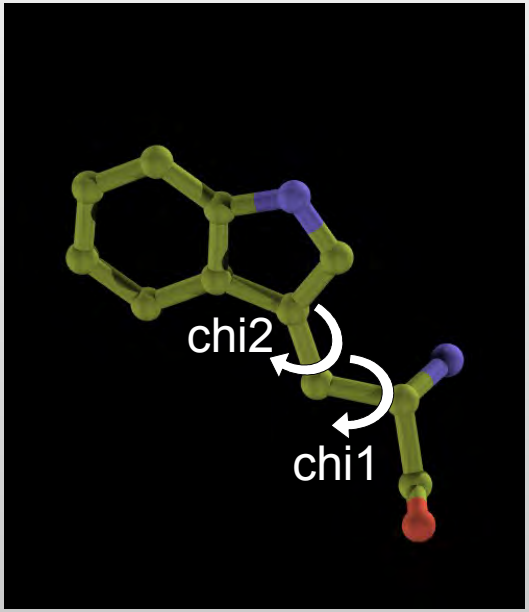
- *c.f.* Oldfield (2001) *Acta Cryst. D* X-LIGAND
- Somewhat different torsion search algorithm
- Build in crystal-space

REFMAC Monomer Library
chem_comp_bond

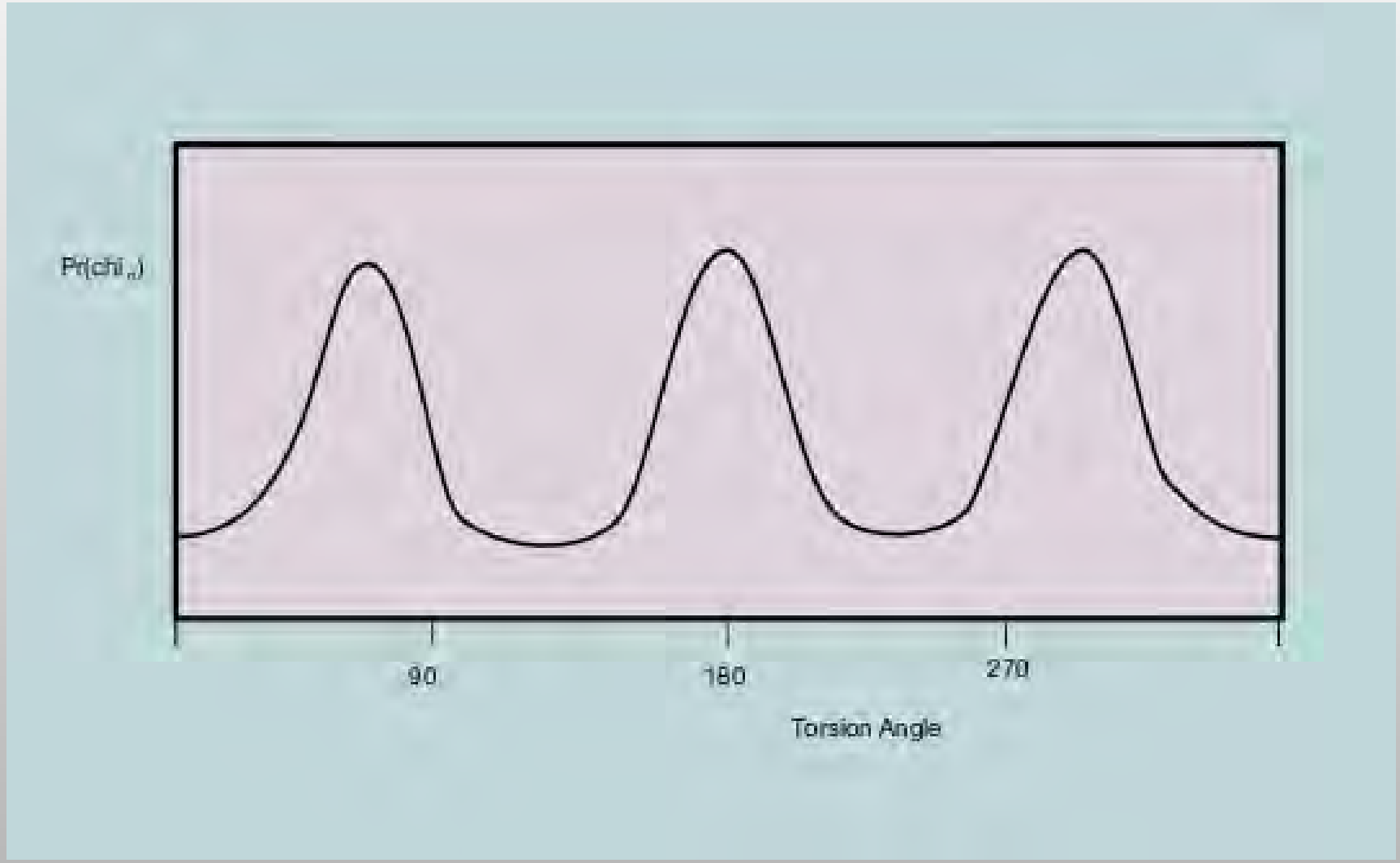
```
loop_
_chem_comp_bond.comp_id
_chem_comp_bond.atom_id_1
_chem_comp_bond.atom_id_2
_chem_comp_bond.type
_chem_comp_bond.value_dist
_chem_comp_bond.value_dist_esd
 ALA      N    H       single     0.860    0.020
 ALA      N    CA      single     1.458    0.019
 ALA      CA   HA      single     0.980    0.020
 ALA      CA   CB      single     1.521    0.033
 ALA      CB   HB1     single     0.960    0.020
 ALA      CB   HB2     single     0.960    0.020
 ALA      CB   HB3     single     0.960    0.020
 ALA      CA   C       single     1.525    0.021
 ALA      C    O       double     1.231    0.020
```

REFMAC Monomer Library
chem_comp_tor

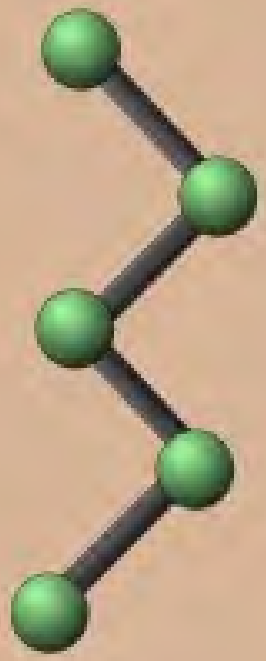
```
loop_
_chem_comp_tor.comp_id
_chem_comp_tor.id
_chem_comp_tor.atom_id_1
_chem_comp_tor.atom_id_2
_chem_comp_tor.atom_id_3
_chem_comp_tor.atom_id_4
_chem_comp_tor.value_angle
_chem_comp_tor.value_angle_esd
_chem_comp_tor.period
 TRP   chi1  N   CA  CB  CG  180.000   15.000   3
 TRP   chi2  CA  CB  CG  CD1  90.000   20.000   2
```



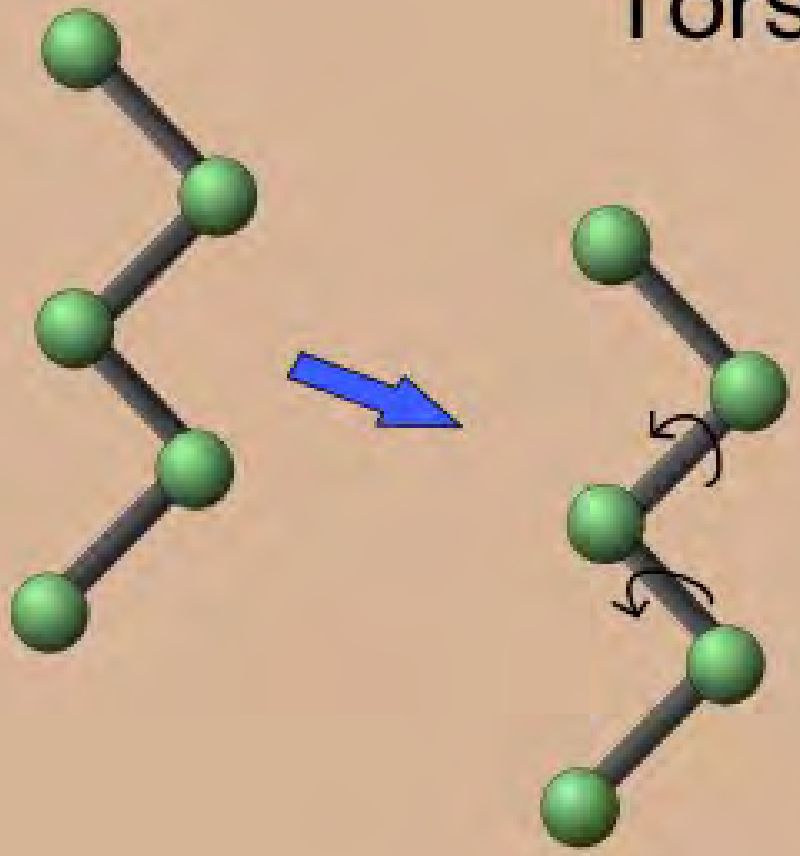
Ligand Torsionable Angle Probability from CIF file



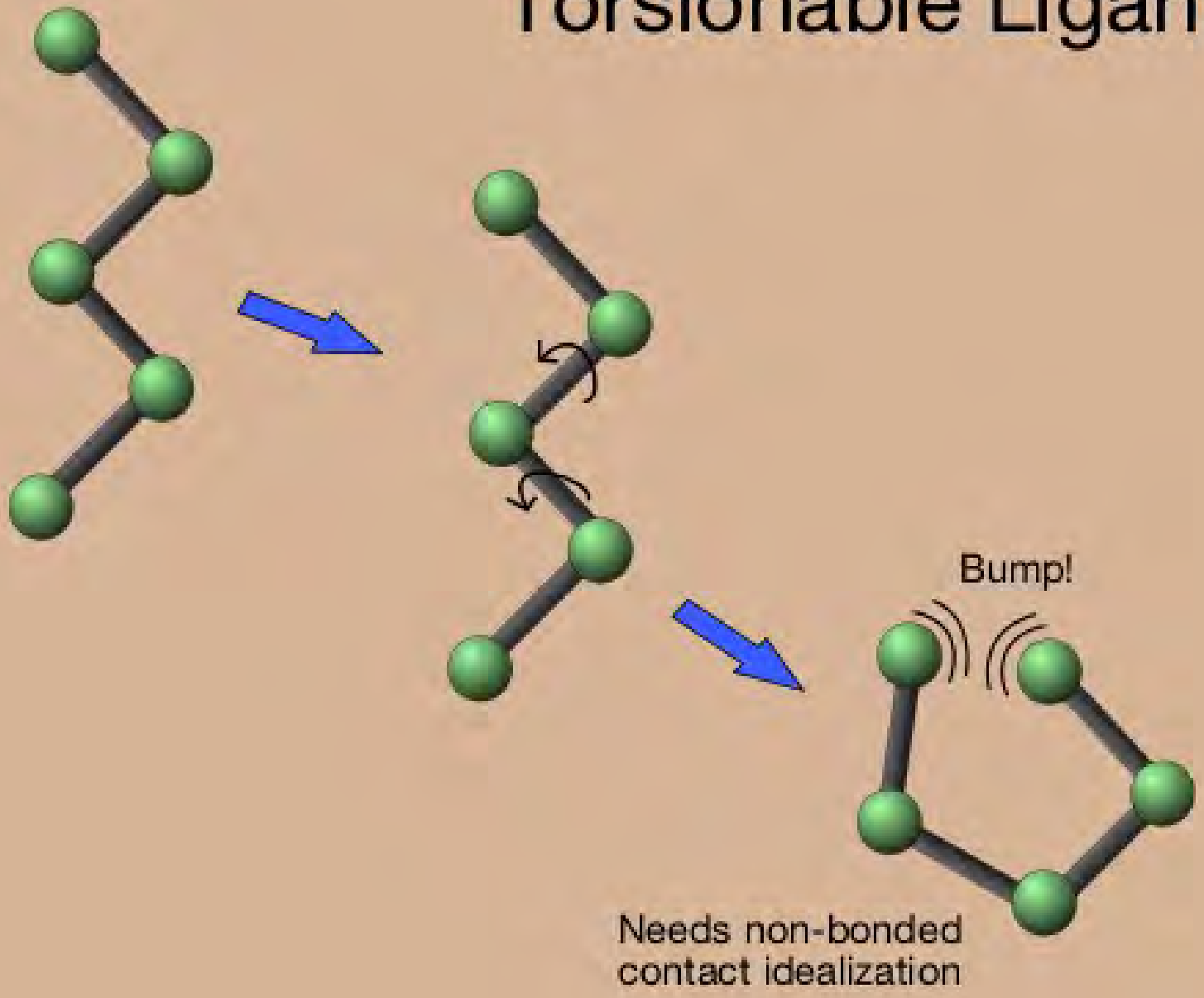
Torsionable Ligands



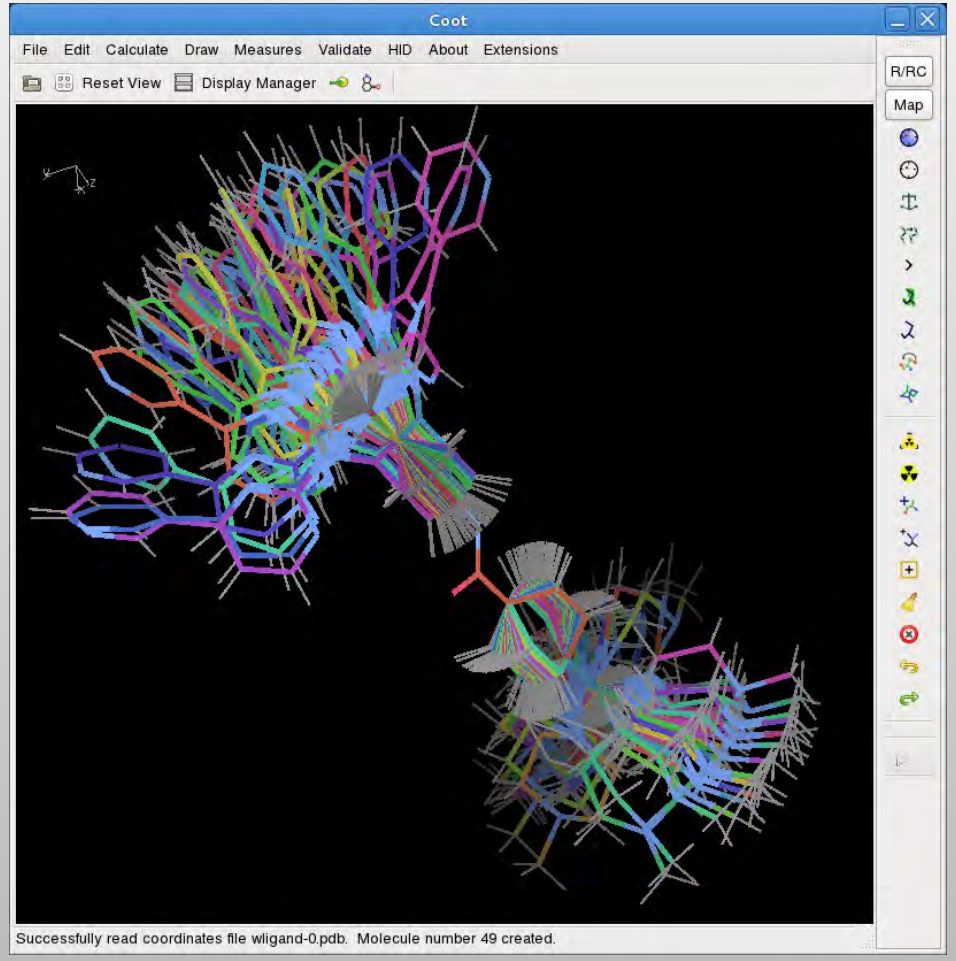
Torsionable Ligands



Torsionable Ligands

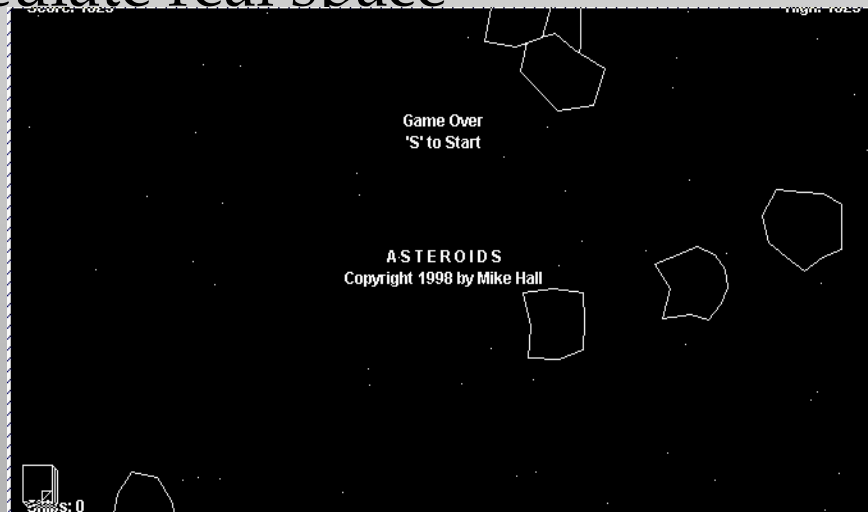


Conformer Generation



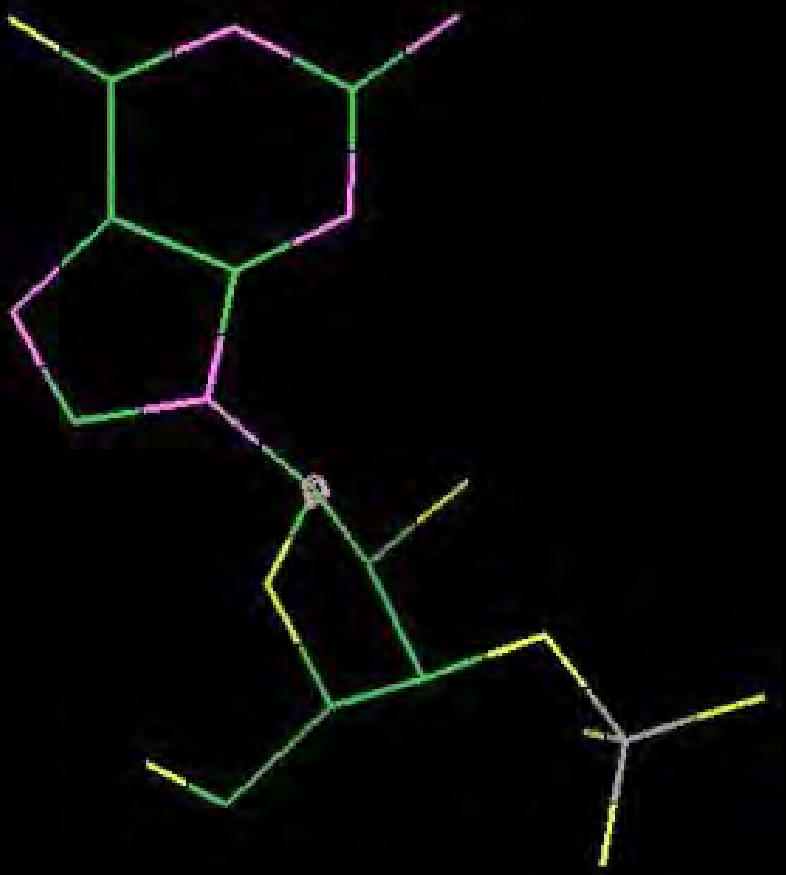
Crystal Space

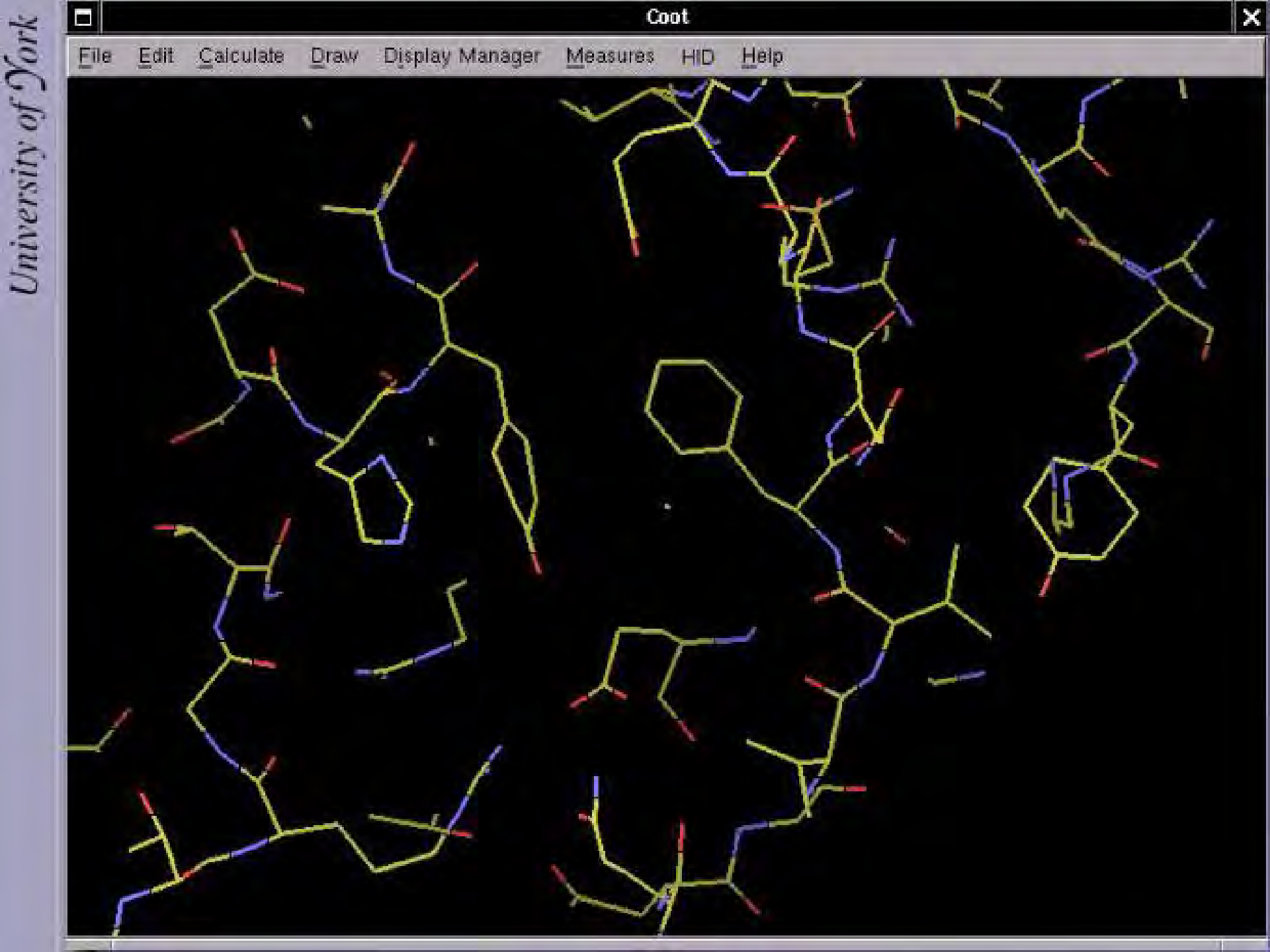
- Build in “crystal space”
  - Like real-space, but wrapped by crystal symmetry
    - Like “Asteroids”
  - Assures only one real-space representation of map features
    - Build everything only once
    - No symmetry clashing
  - However, more difficult to calculate real space geometries
    - …such as bonds, torsions

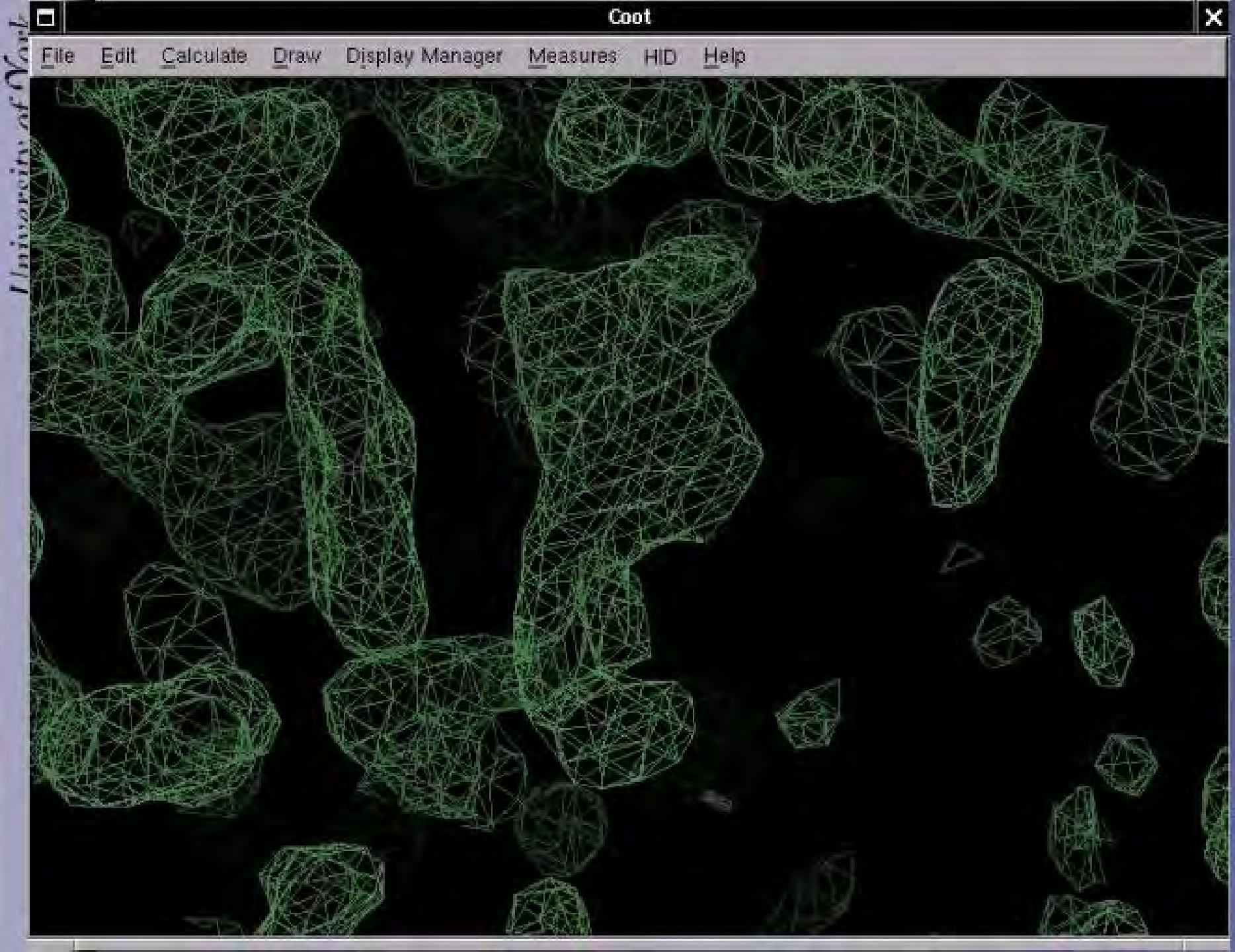


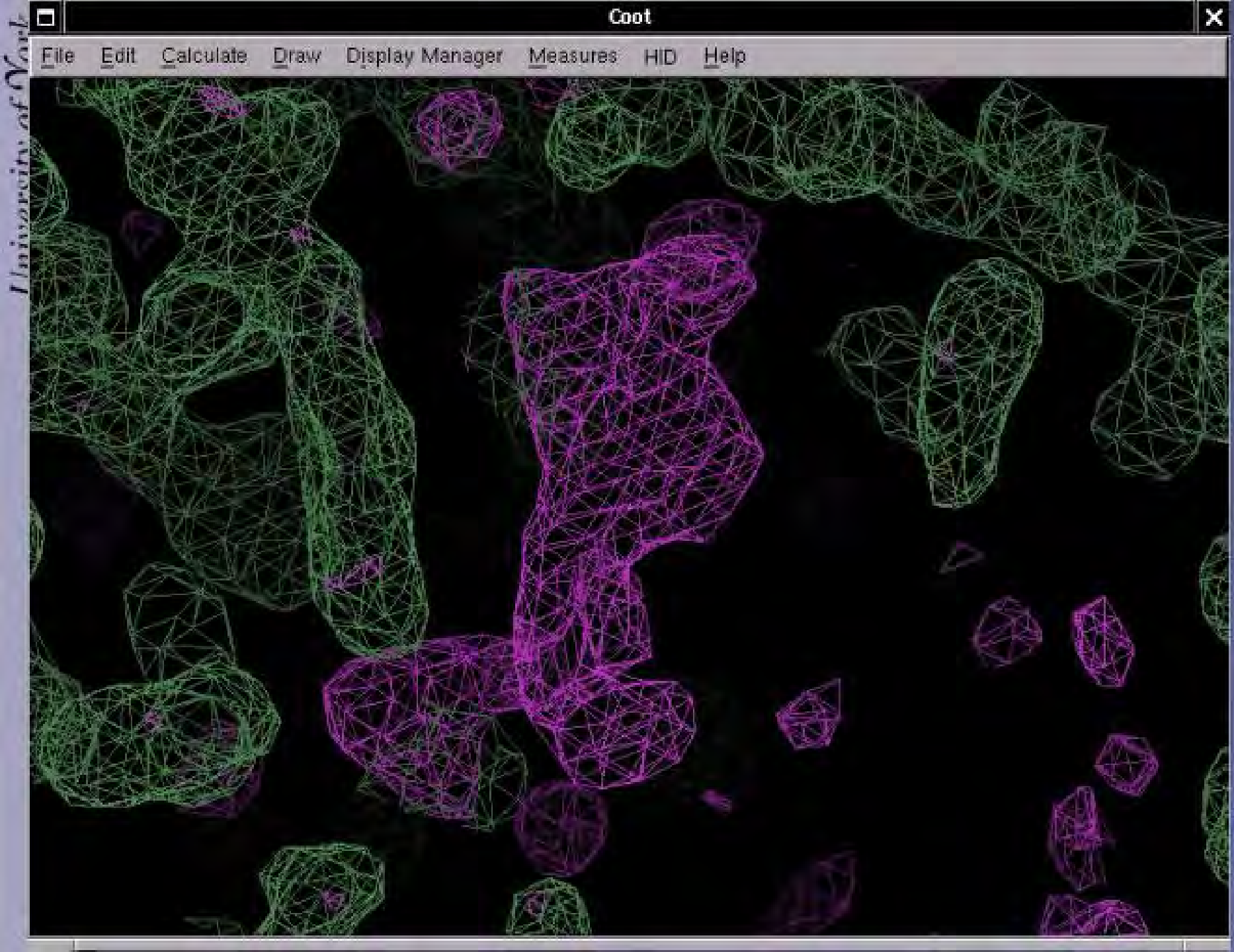
Coot

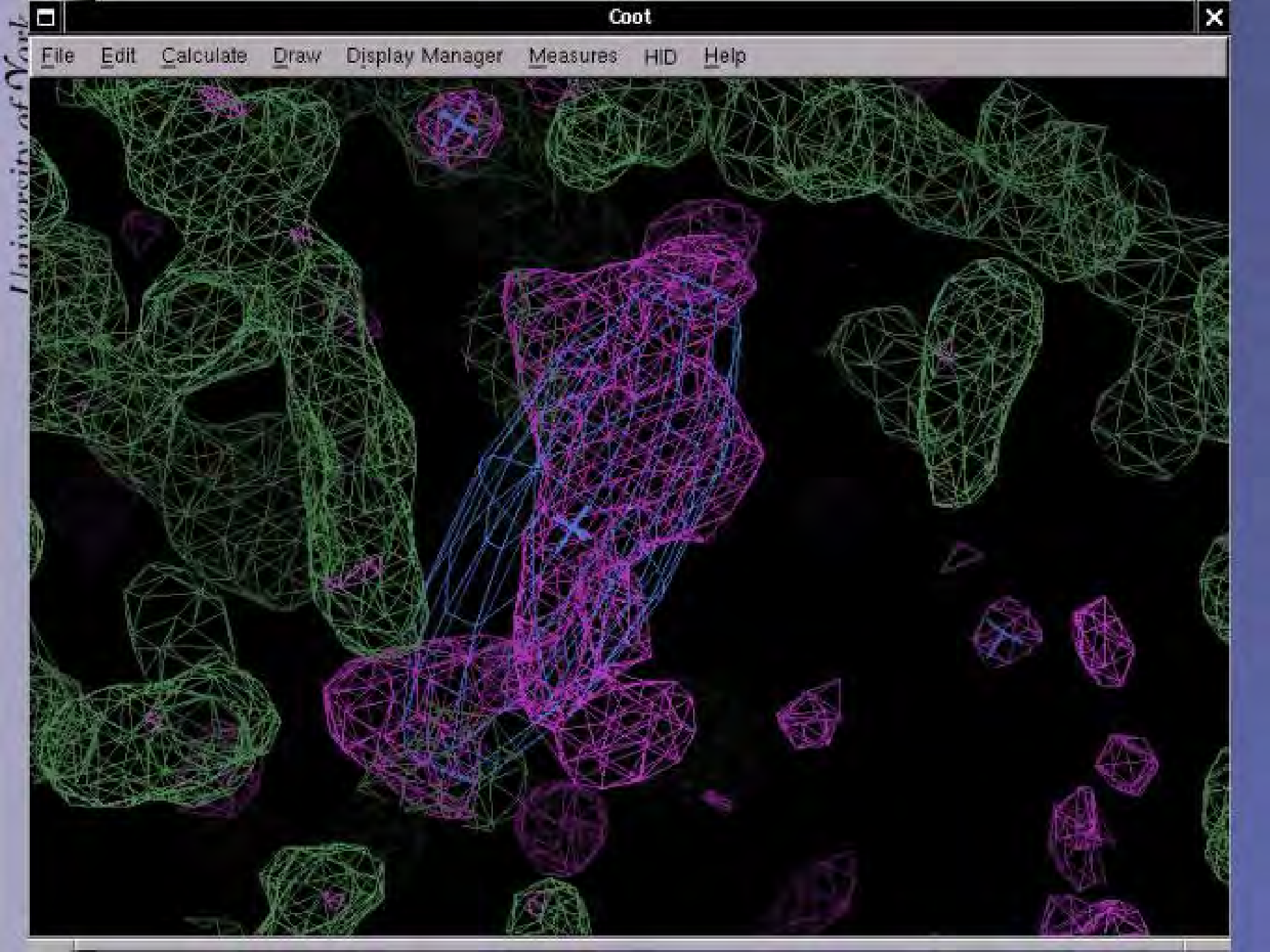
File Edit Calculate Draw Display Manager Measures HID Help

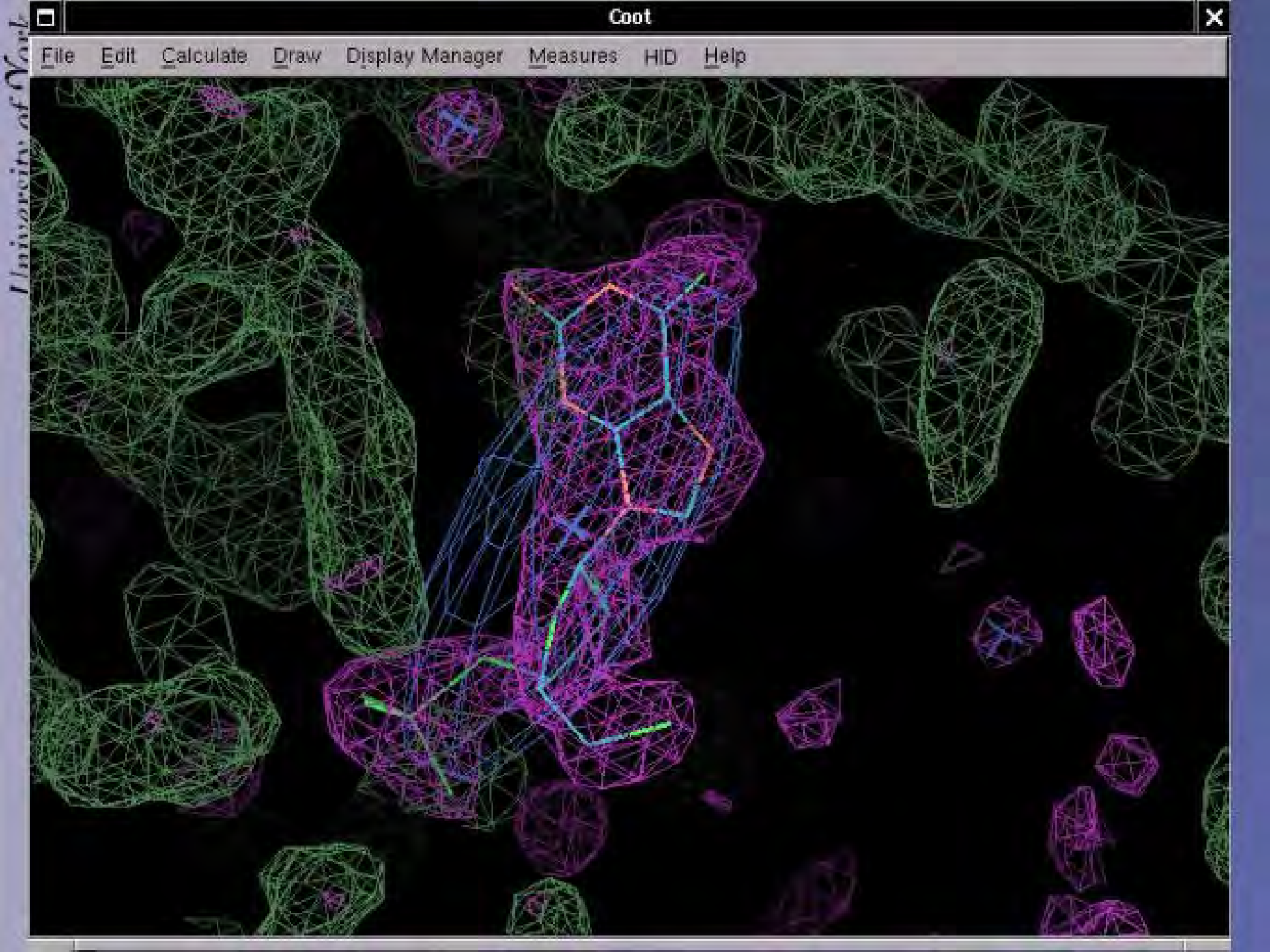


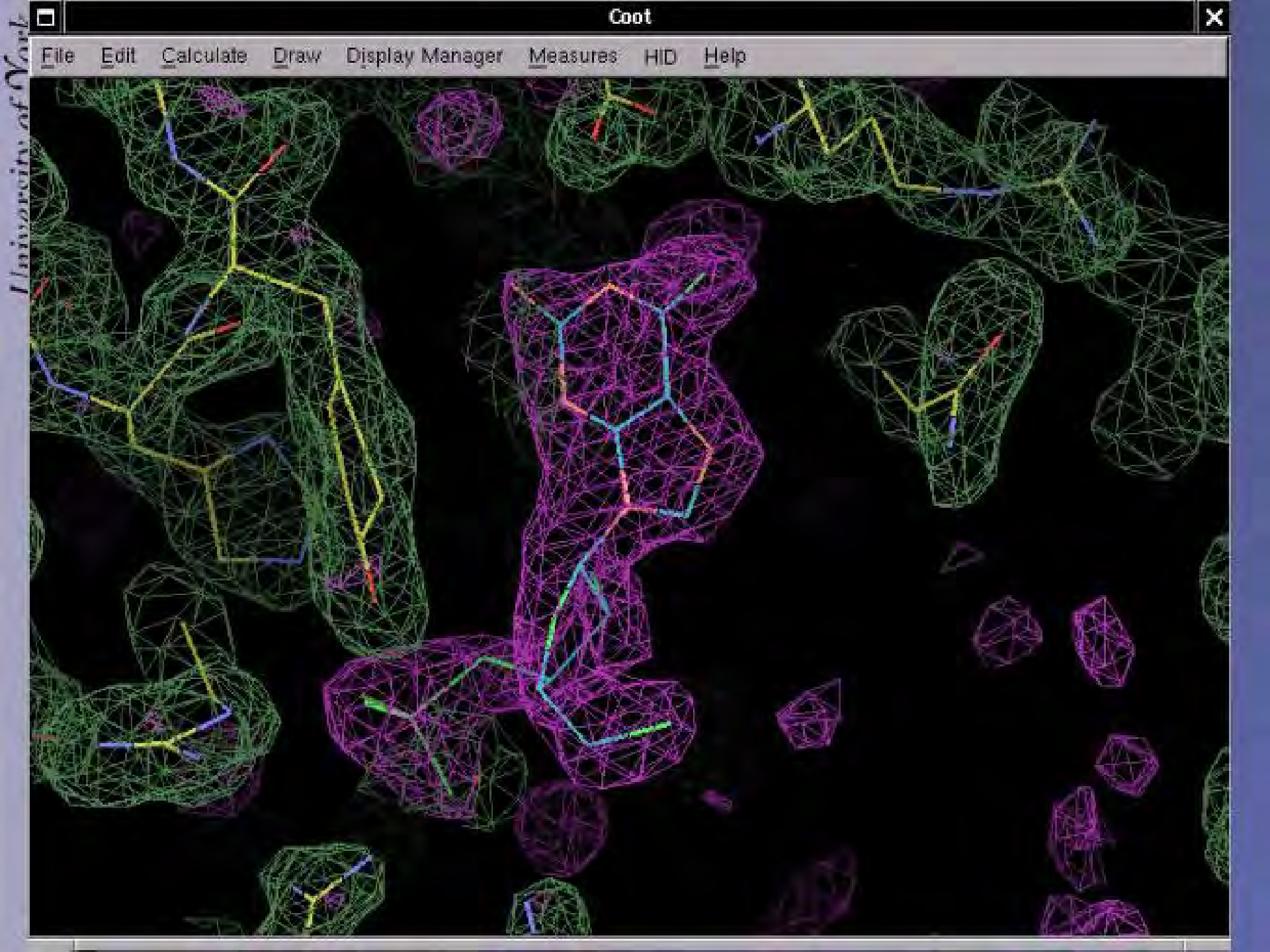

University of York
Coot
File
Edit
Calculate
Draw
Display Manager
Measures
HID
Help


Coot
File
Edit
Calculate
Draw
Display Manager
Measures
HID
Help
University of York


Coot
File
Edit
Calculate
Draw
Display Manager
Measures
HID
Help
University of York


Coot
File
Edit
Calculate
Draw
Display Manager
Measures
HID
Help
University of York


Coot
File
Edit
Calculate
Draw
Display Manager
Measures
HID
Help
University of York

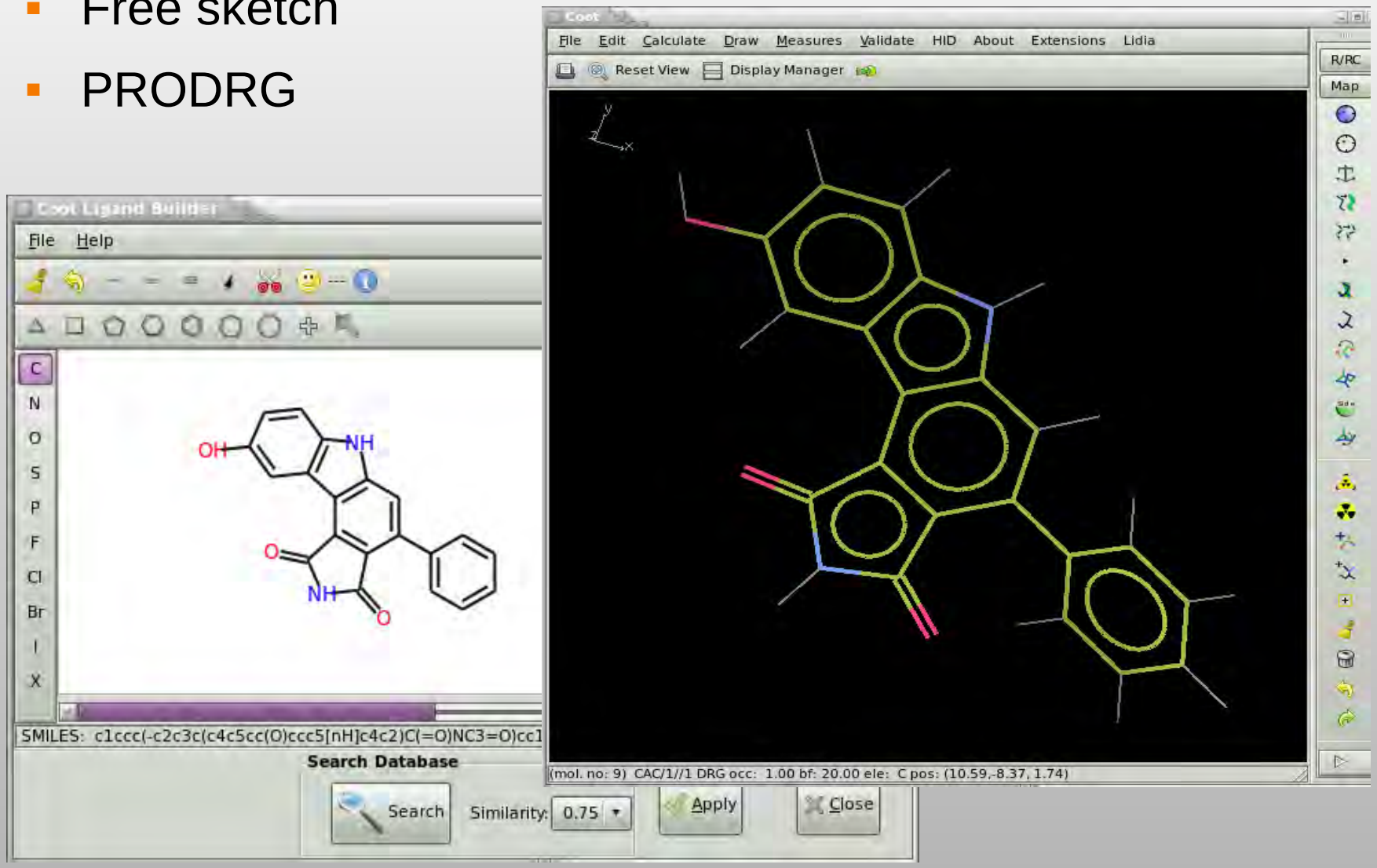

Coot
File
Edit
Calculate
Draw
Display Manager
Measures
HID
Help
University of York

Conformation Idealization

- Each conformer is passed through the “Regularization” function of Coot
  - Non-bonded terms included
- Better to have hydrogen atoms on the model
- Slows things down a good deal…
  - May not be the best method to explore conformational variability for many rotatable bonds

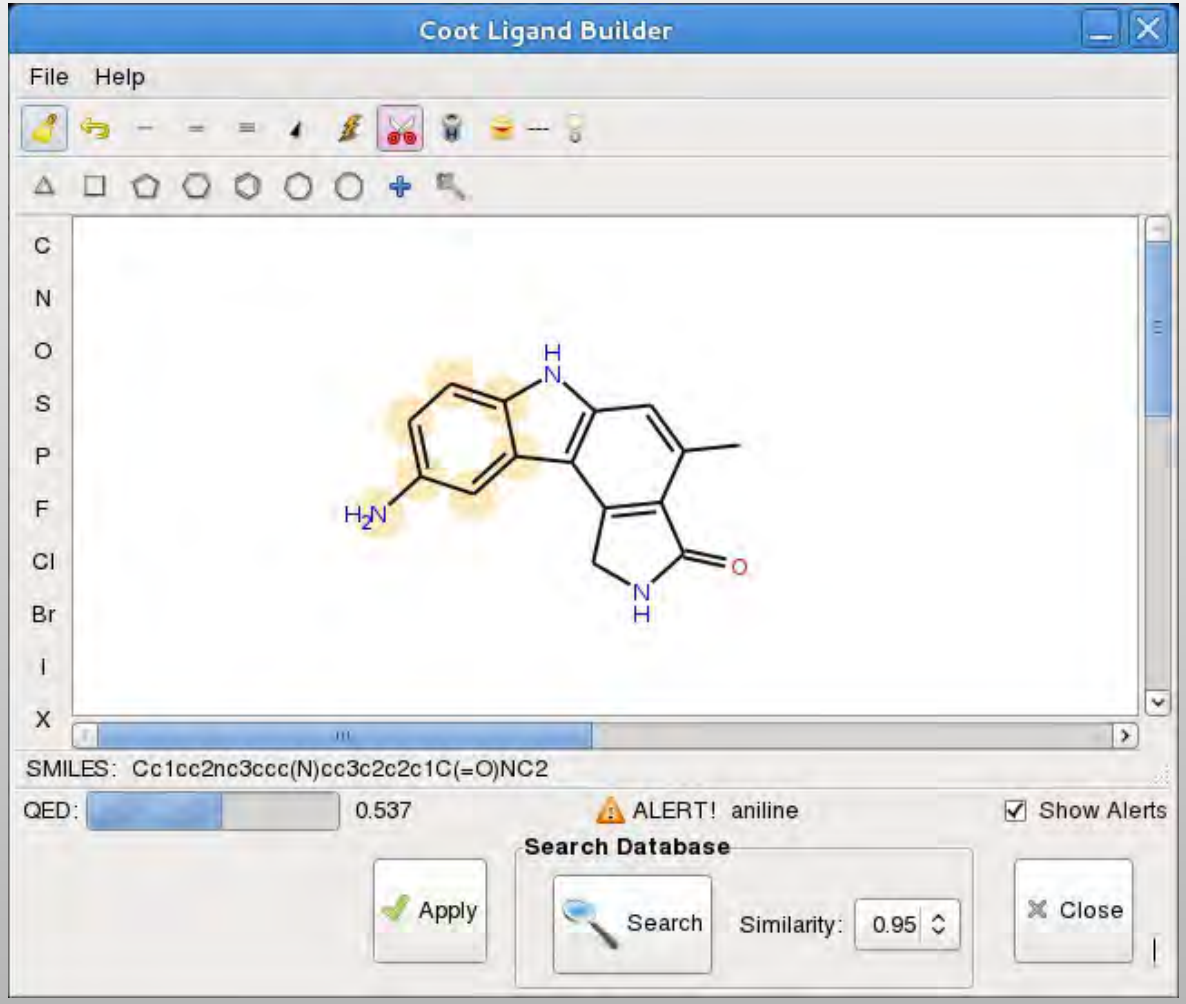
2D Ligand Builder

- Free sketch
- PRODRG



2D Sketcher

- Structural Alerts



200 built-in alerts with the ability to any additions

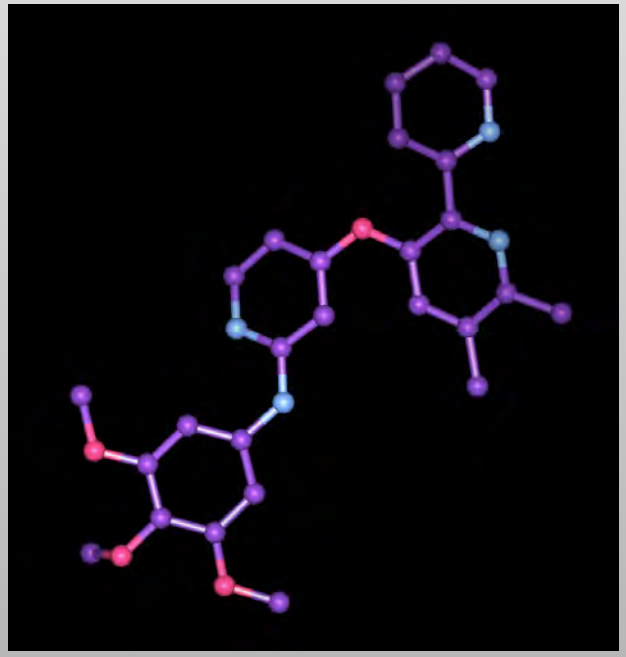
Ligand Overlay

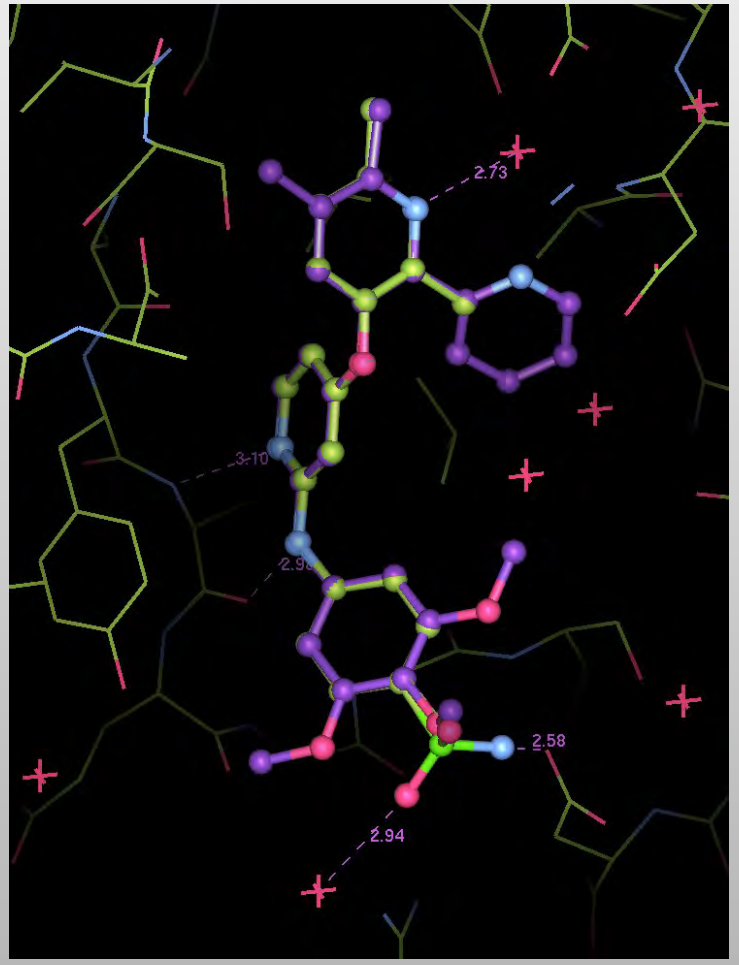
- Algorithm and Code by Eugene Krissinel
- Tries to overlay different ligands/monomers by graph matching
- Useful for “database” ligands where atom names are not selected by hand
- Has been used as the basis of the function which “mutates” residues to alternative monomer types
  - e.g. phosphorylation

”Yesterday's” Ligand

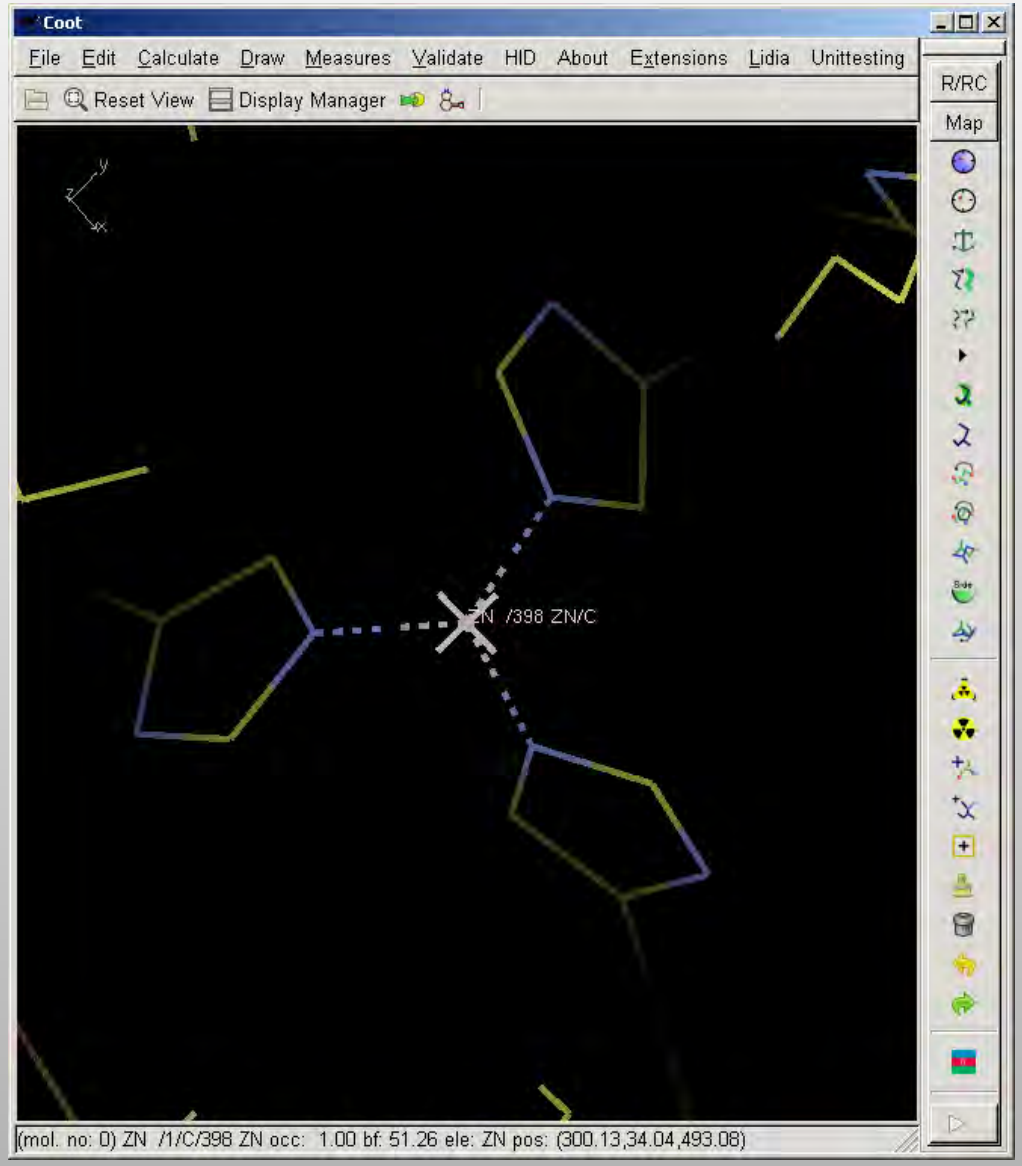
Common subgraph isomorphism, Krissinel & Henrick (2004)

- Atom name matching
- Torsion matching
- Ligand overlay

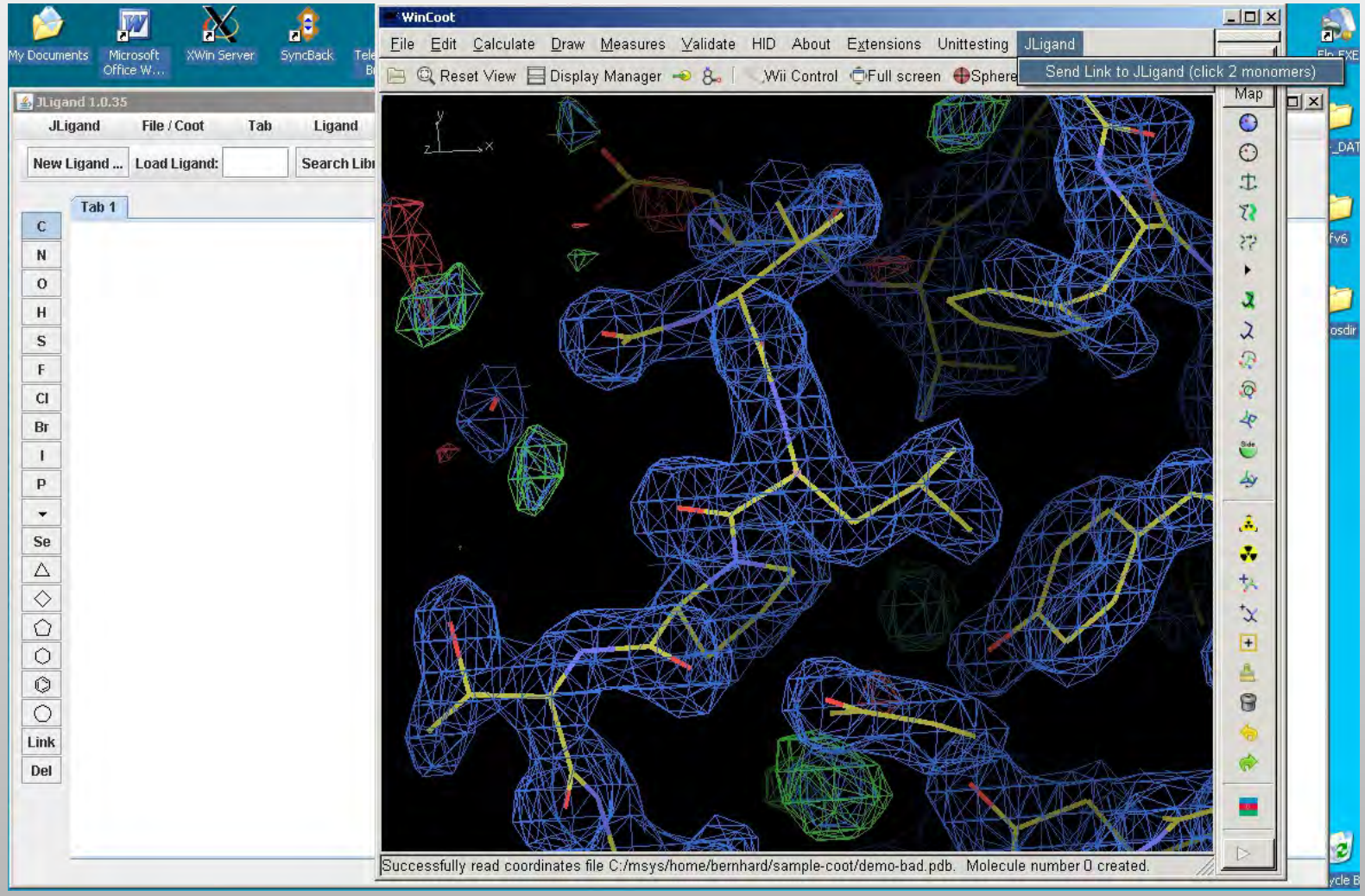




Displaying LINKs



Interface with JLigand to create LINKs



First: Extensions → Modelling → Launch JLigand

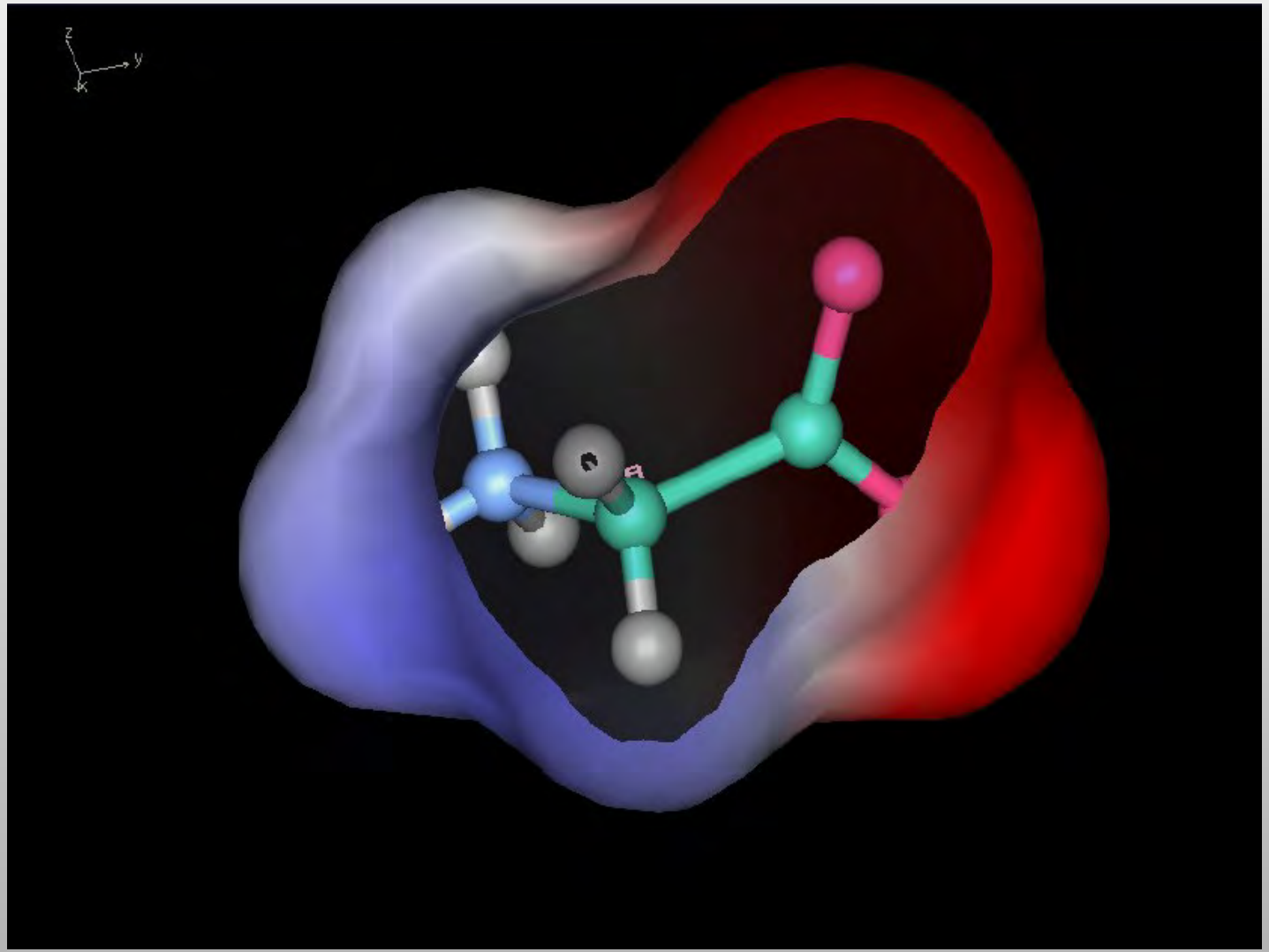
Ligand representation

Ligand Represenation

- Bond orders (from dictionary)

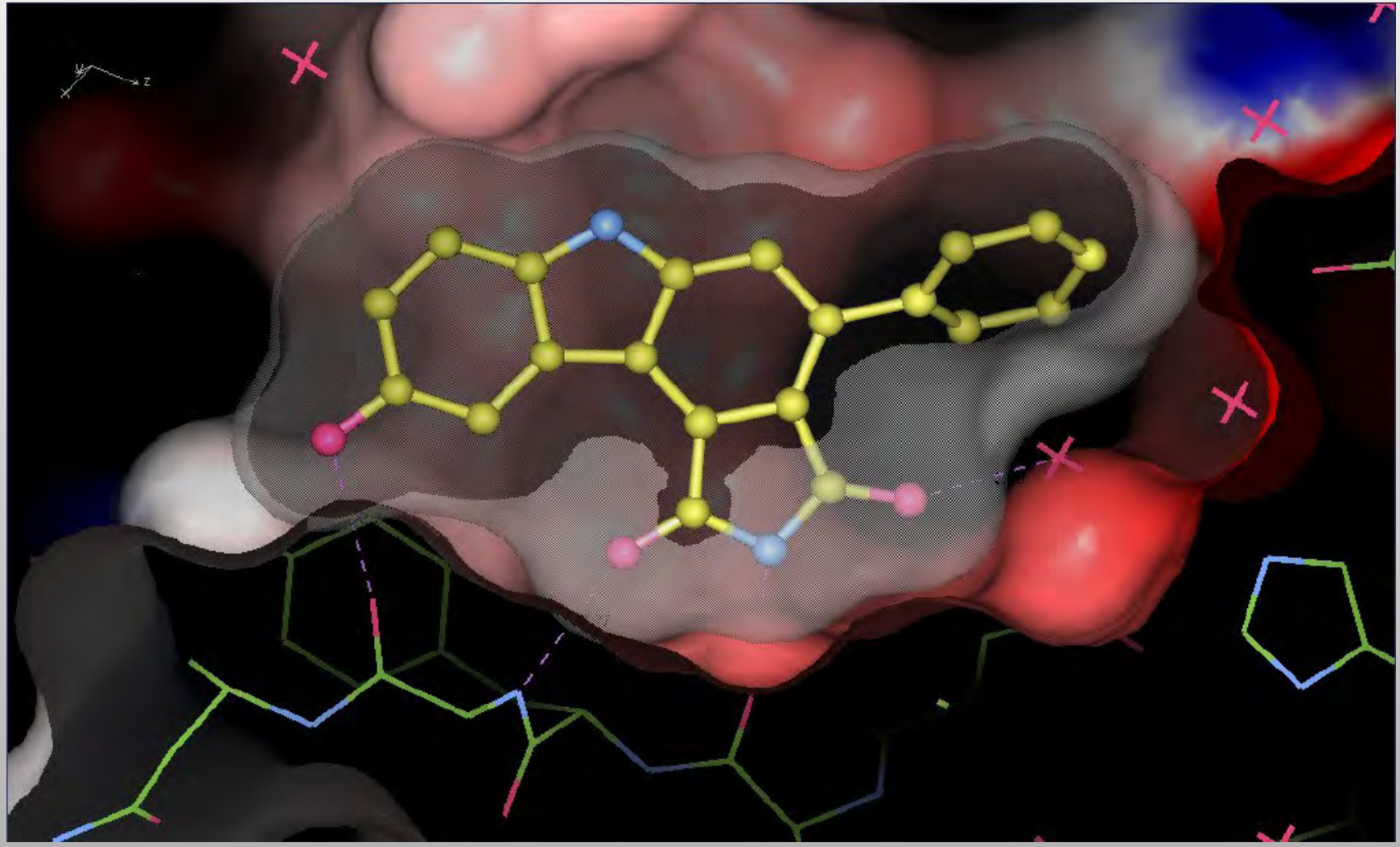
Surfaces

- Partial charges

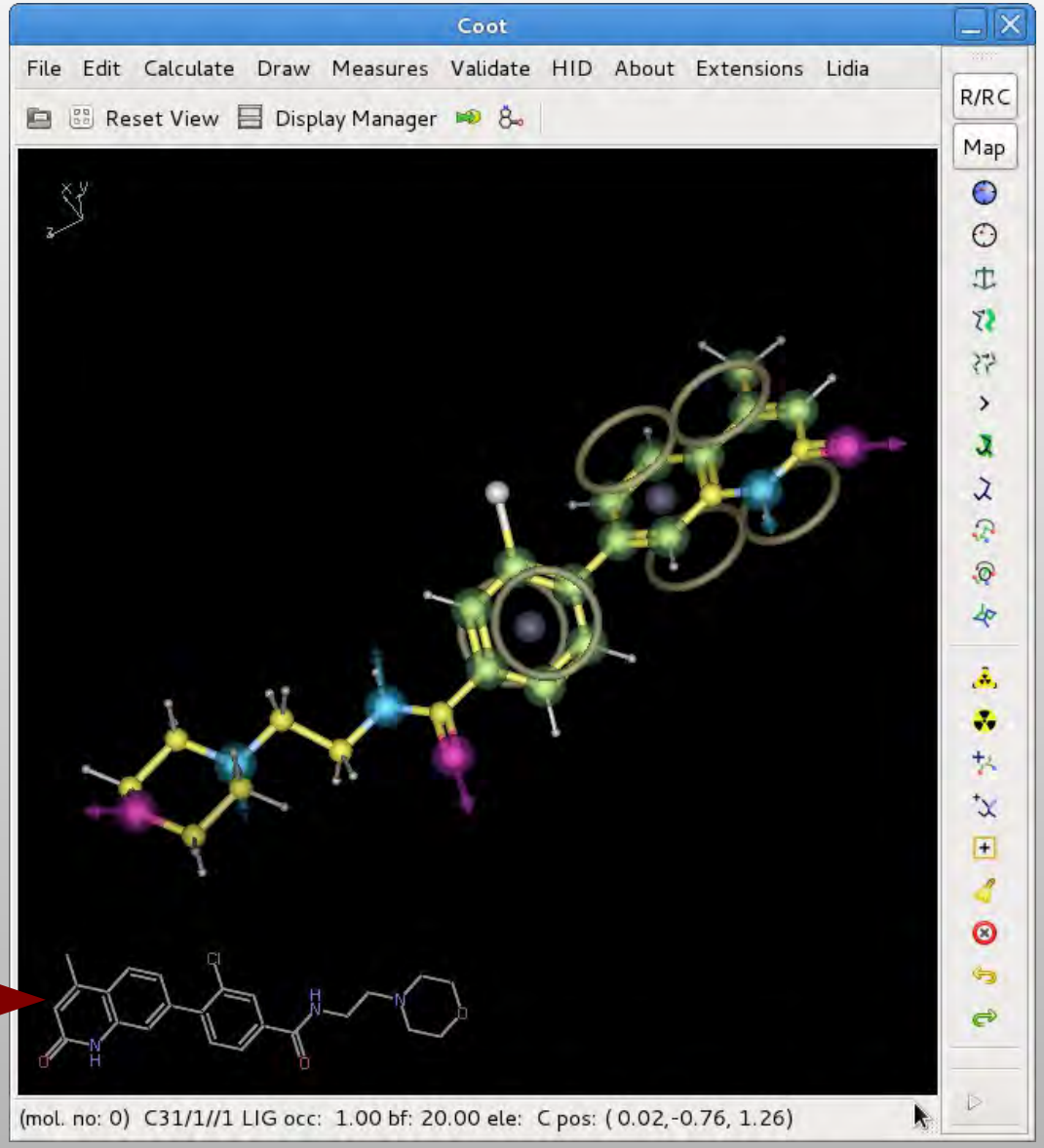


Surfaces

- Transparent surfaces – surface complementarity



Chemical Features

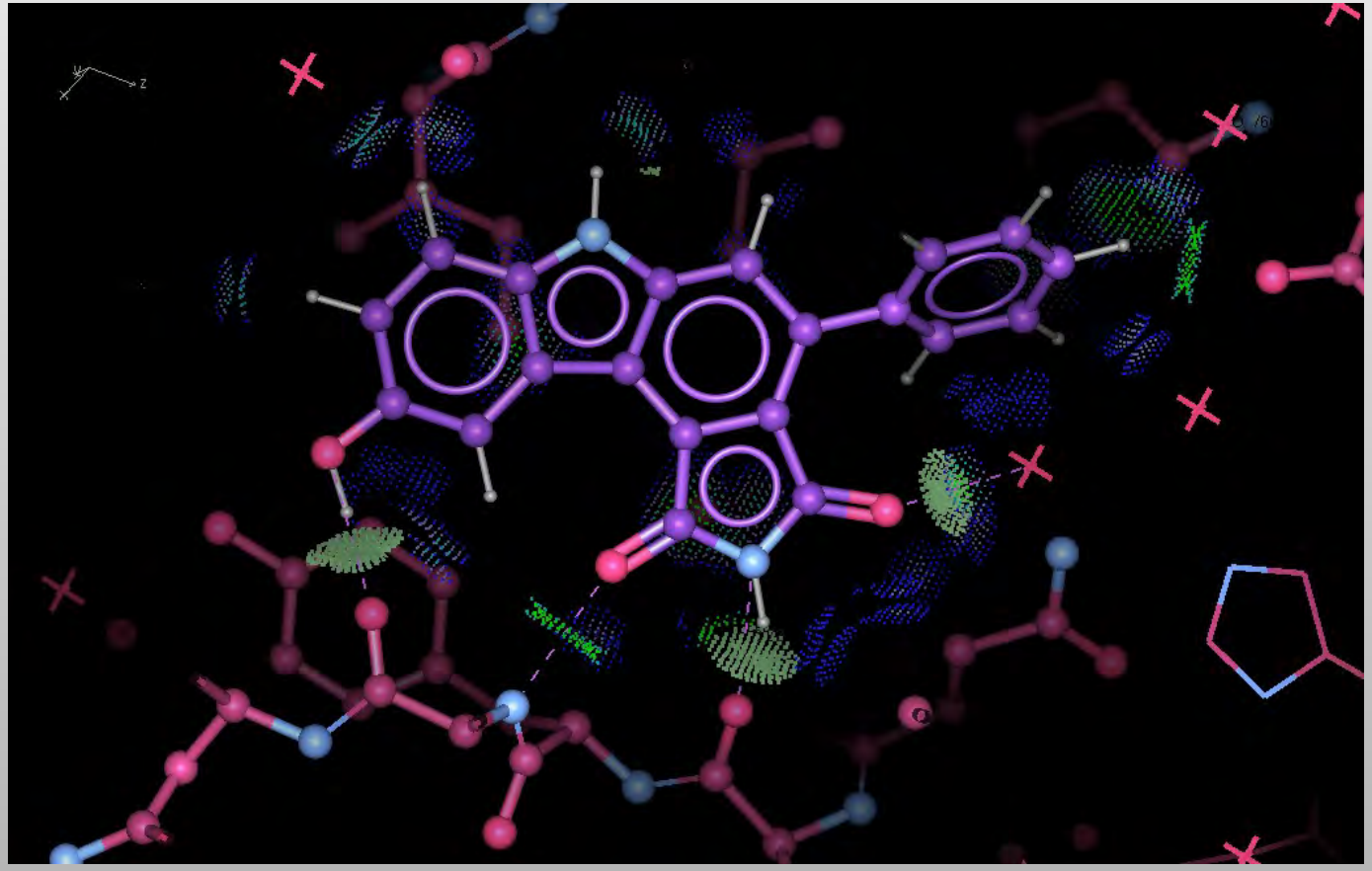


and chemical
diagram thumbnails

Ligand representation

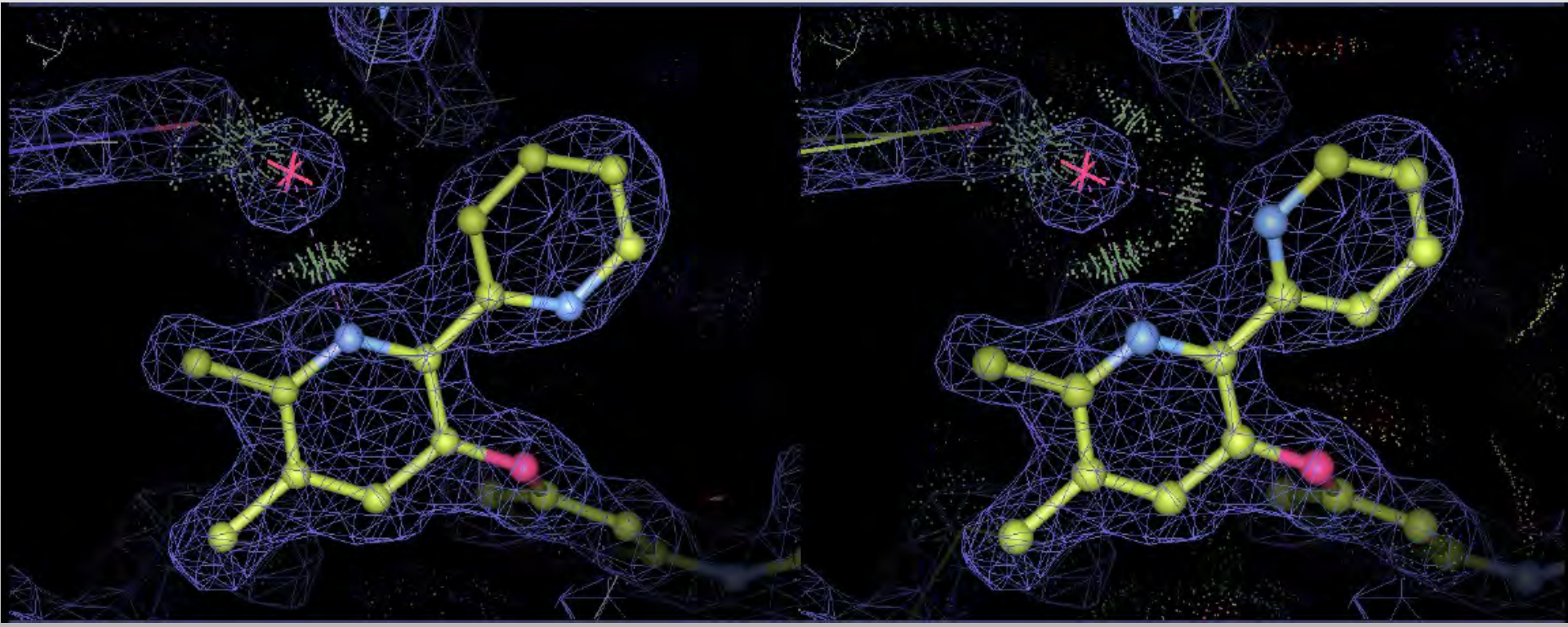
Binding mode analysis

- Binding site highlighting
- Isolated Molprobity probe dots



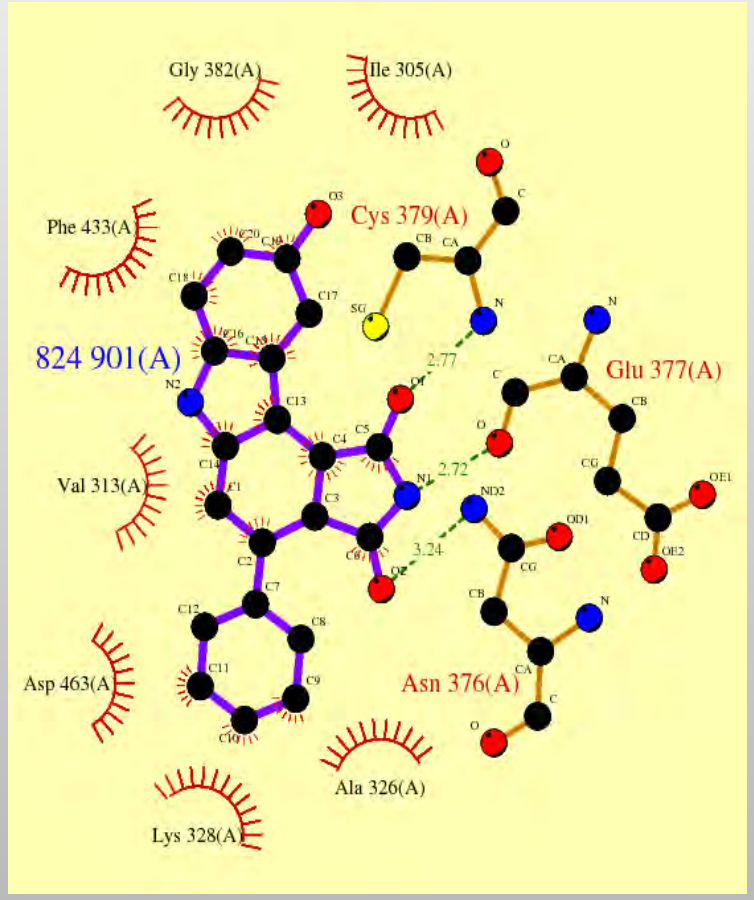
Binding mode analysis

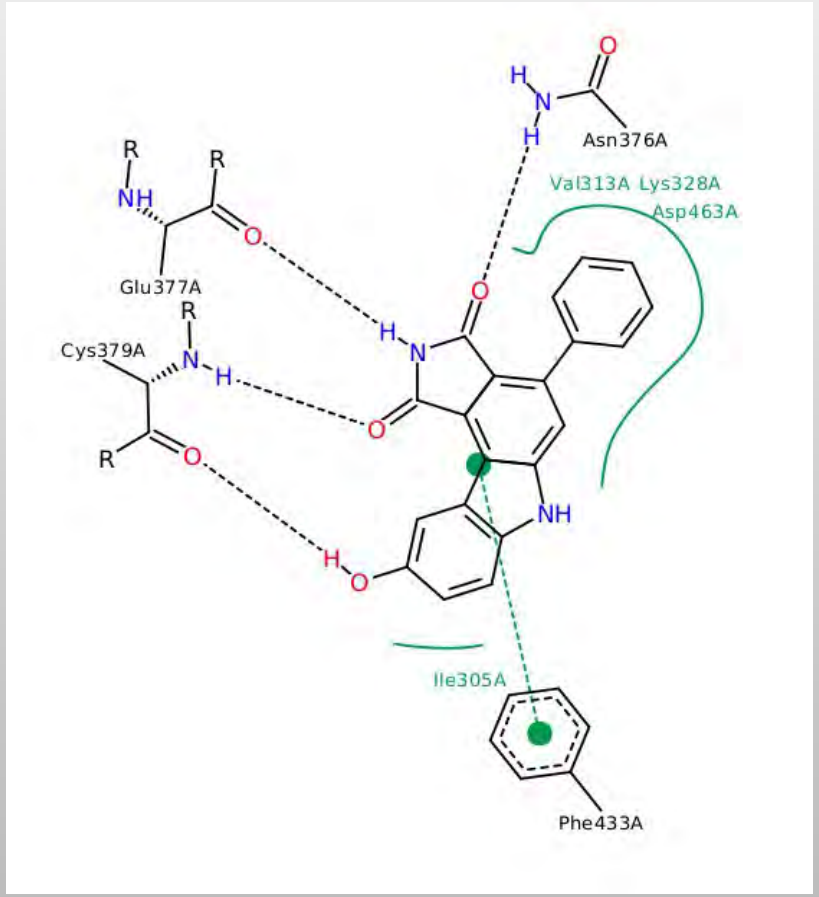
- Binding site highlighting
- Isolated molprobity dots



Ligand Environment Layout

- 2d Ligand pocket layout (ligplot, poseview)





Can we do better? - Interactivity?

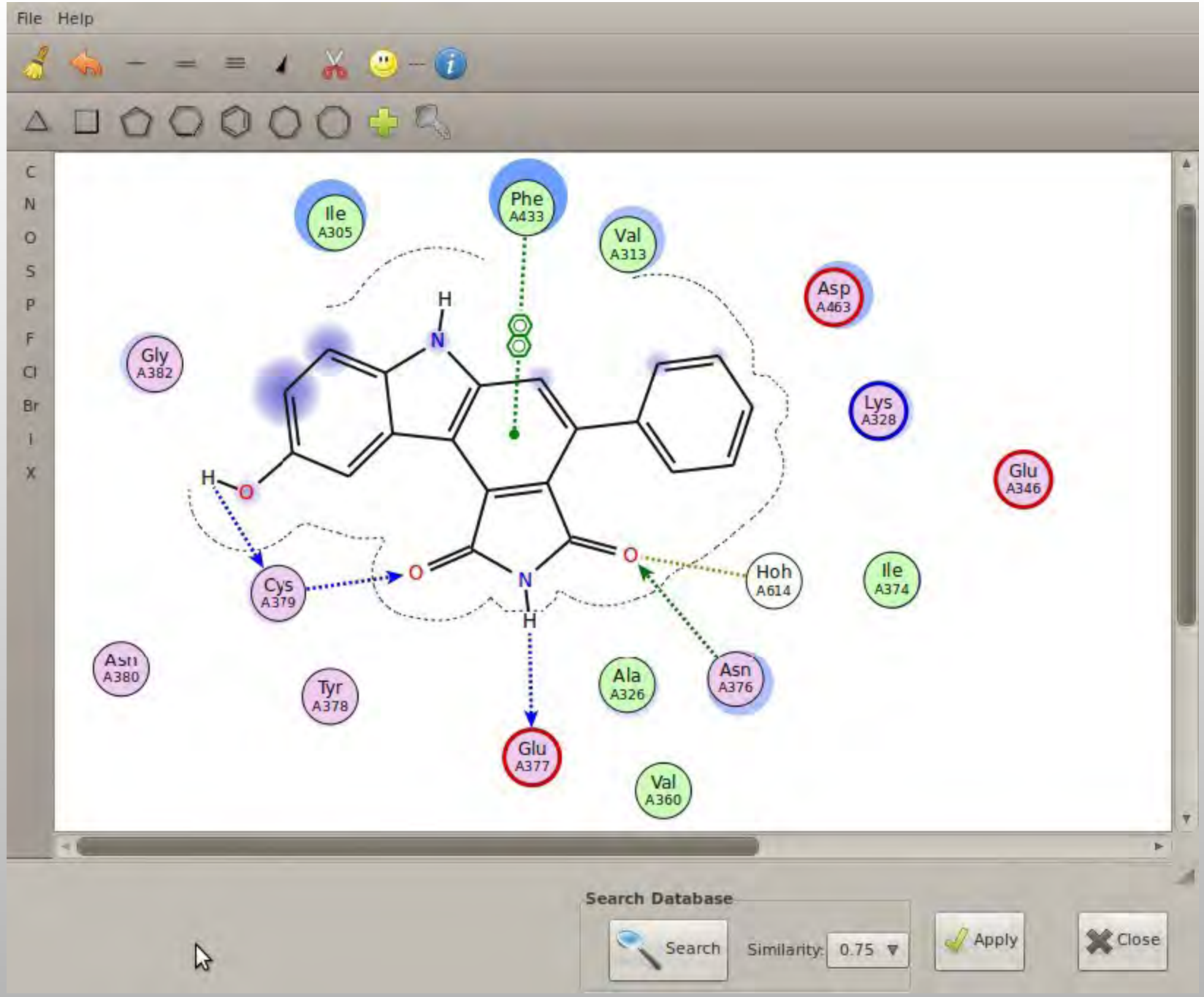
Ligand Environment Layout

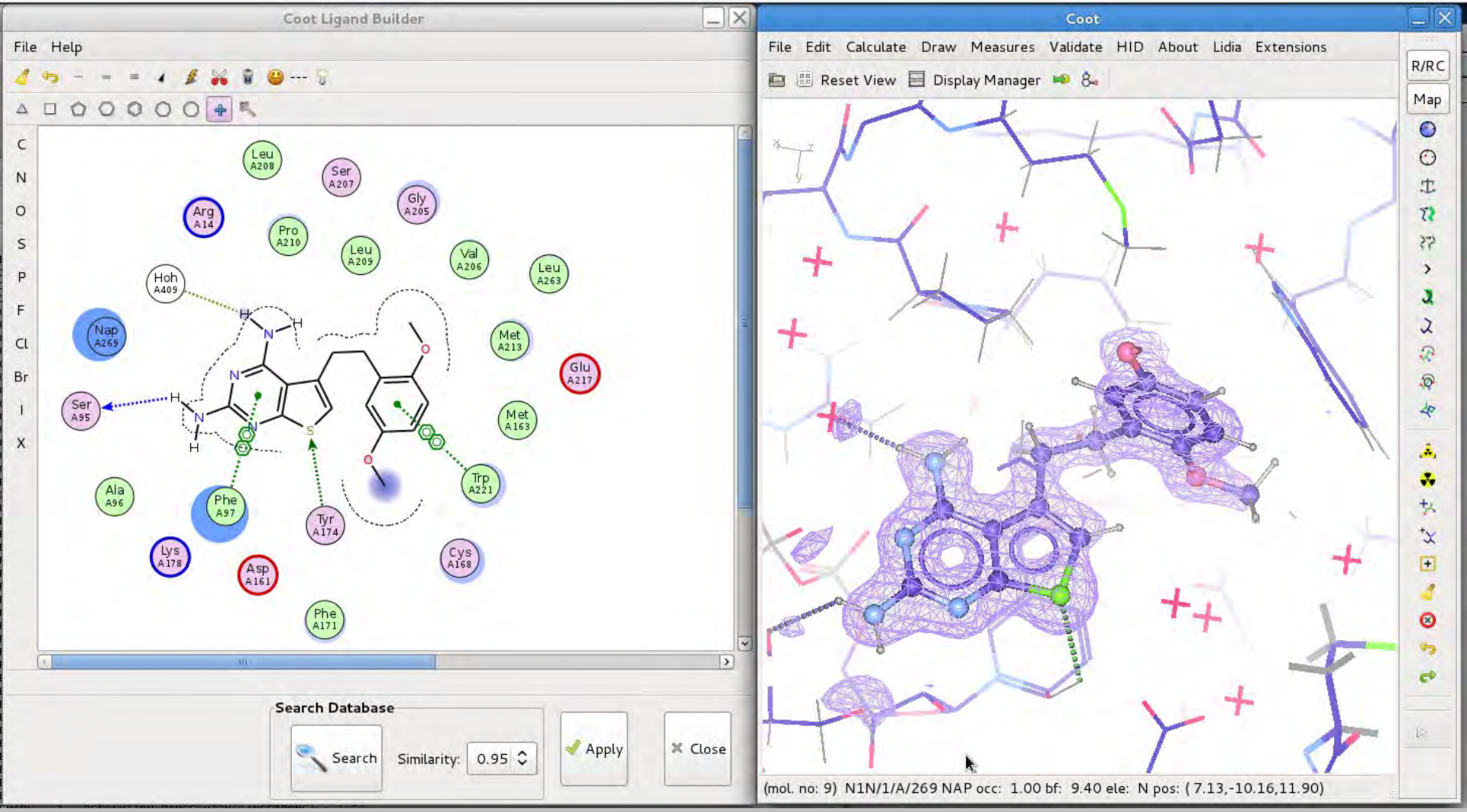
- Binding pocket residues
- Interactions
- Solvent accessibility
  - halos
- Solvent exclusion by ligand
- Substitution contour

Ligand Enviroment Layout

- Considerations
  - 2D placement and distances should reflect 3D metrics (as much as possible)
    - H-bonded residues should be close the atoms to which they are bonded
  - Residues should not overlap the ligand
  - Residues should not overlap each other
  - *c.f.* Clark & Labute (2007)

    (work in progress)


File Help
C
N
O
S
P
F
Cl
Br
I
X
Ile A305
Phe A433
Val A313
Asp A463
Gly A382
Lys A328
Glu A346
Cys A379
Hoh A614
Ile A374
Asn A380
Tyr A378
Ala A326
Asn A376
Glu A377
Val A360
Search Database
Search
Similarity: 0.75
Apply
Close


Coot Ligand Builder
File Help
Leu A208
Ser A207
Gly A205
Arg A14
Pro A210
Leu A209
Val A206
Leu A263
Hoh A409
Nap A269
Met A213
Glu A217
Ser A95
Met A163
Trp A221
Ala A96
Phe A97
Tyr A174
Lys A178
Asp A161
Cys A168
Phe A171
Search Database
Search
Similarity: 0.95
Apply
Close
Coot
File Edit Calculate Draw Measures Validate HID About Lidia Extensions
Reset View
Display Manager
R/RC
Map
(mol. no: 9) N1N/1/A/269 NAP occ: 1.00 bf: 9.40 ele: N pos: (7.13,-10.16,11.90)

Ligand futures

- Ligand validation with CSD (Mogul or similar)
- Dynamic hydrogen bonding and non-bonding contacts
- pyrogen?!
- Ace-drg

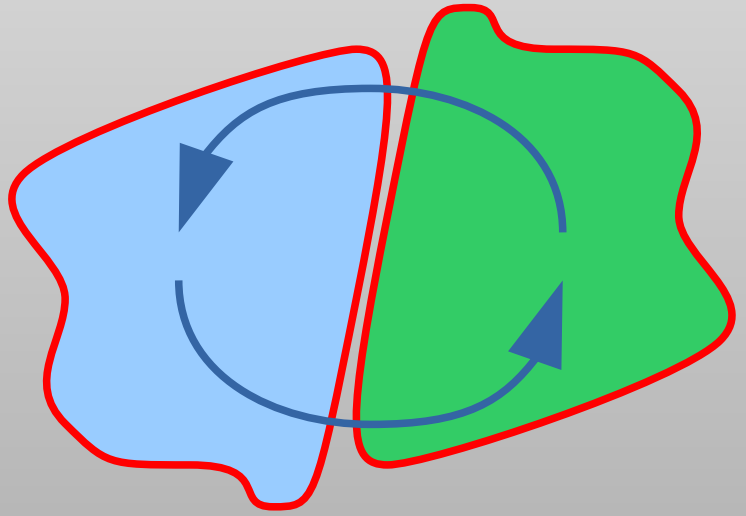
…and more?

- NCS
- Validation

Handling NCS…

What is Non-Crystallographic Symmetry (NCS)?

- 2 or more copies of a molecule in the unit cell not related by crystallographic symmetry
- NCS related molecules provide different representations of the same molecule

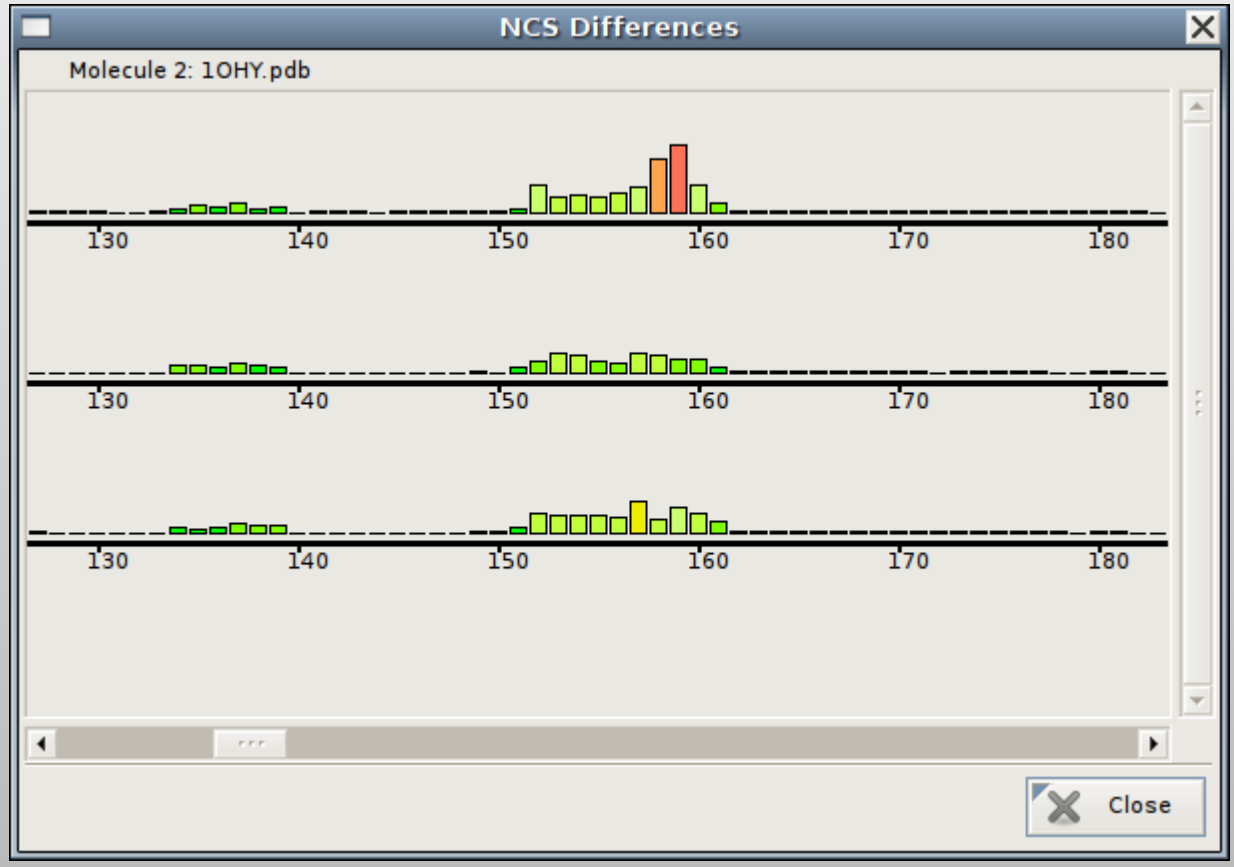


Handling NCS

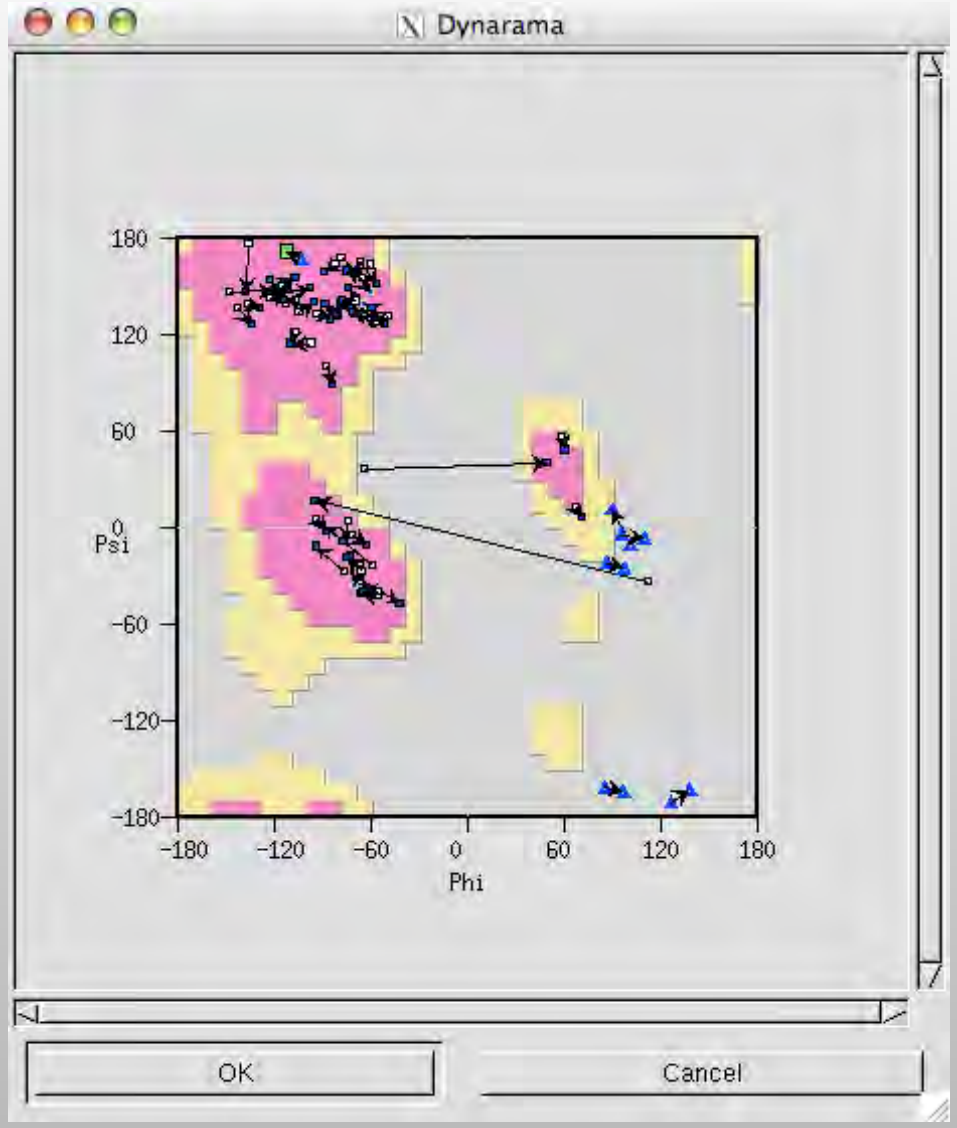
Typical Scenario:

- I have done an LSQ overlap of my NCS-related molecules and from the graph, have seen significant deviations in the positions of some side-chains.
- Why are they different?

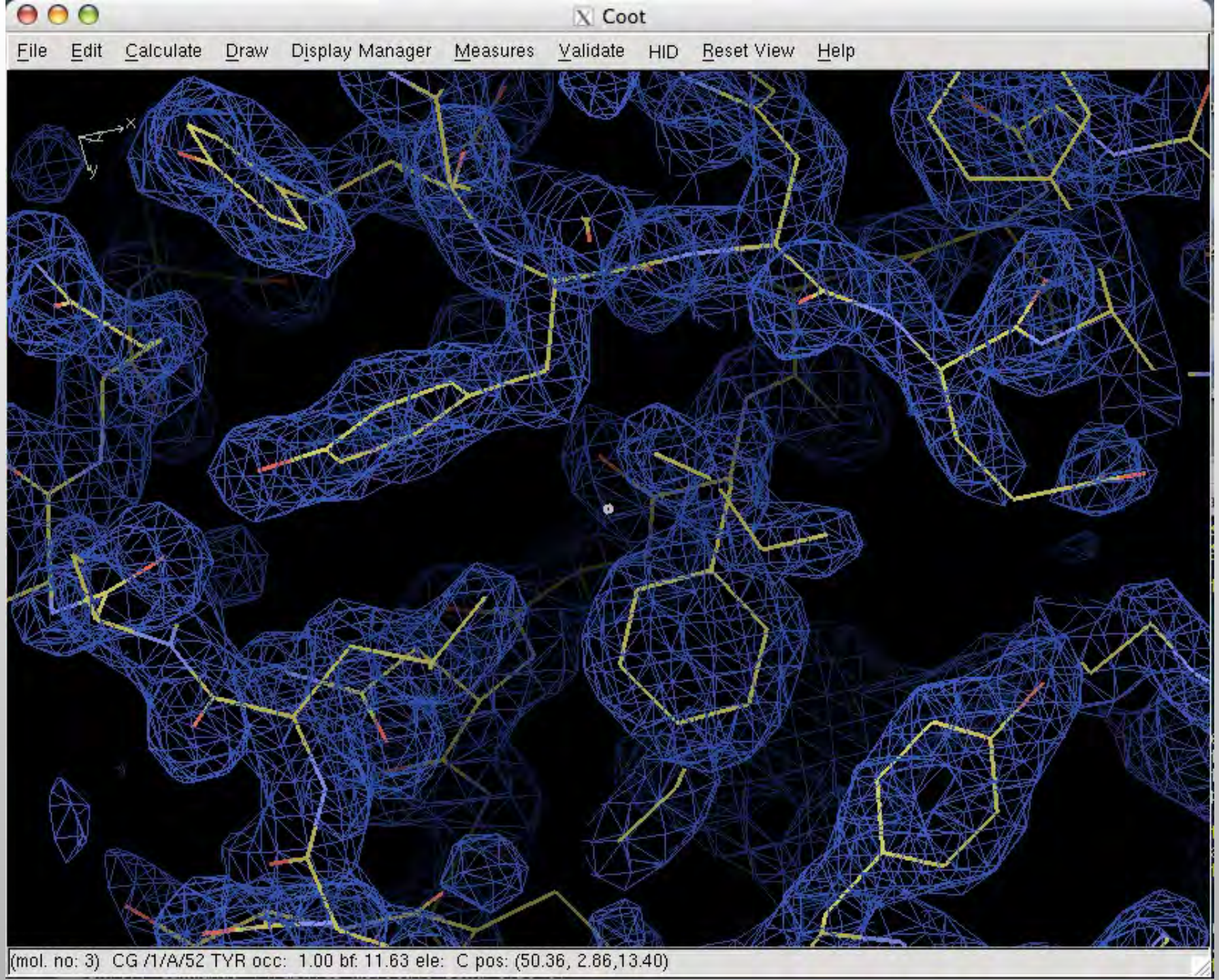
...or NCS Differences graph



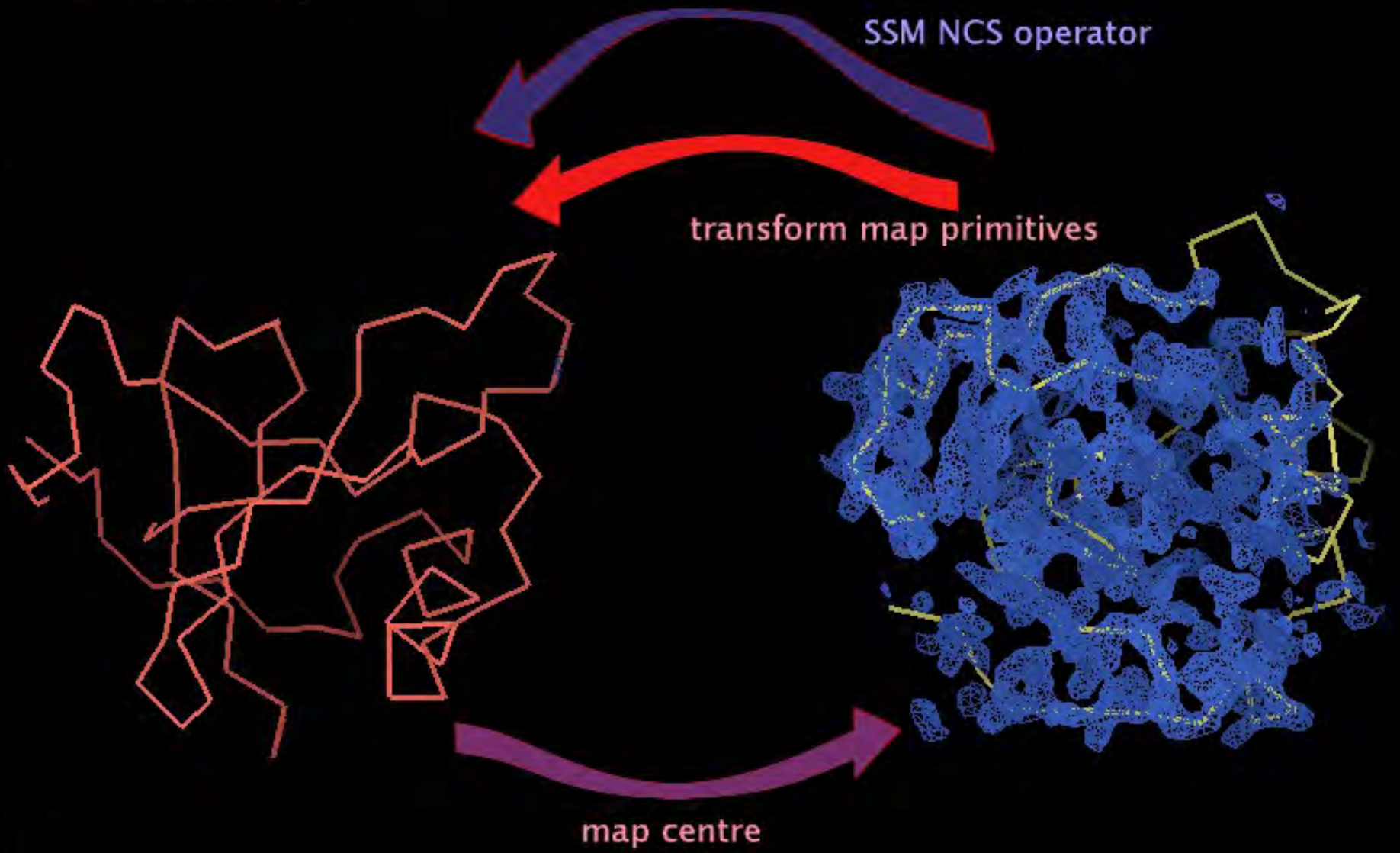
…or Kleywegt Plots[*]

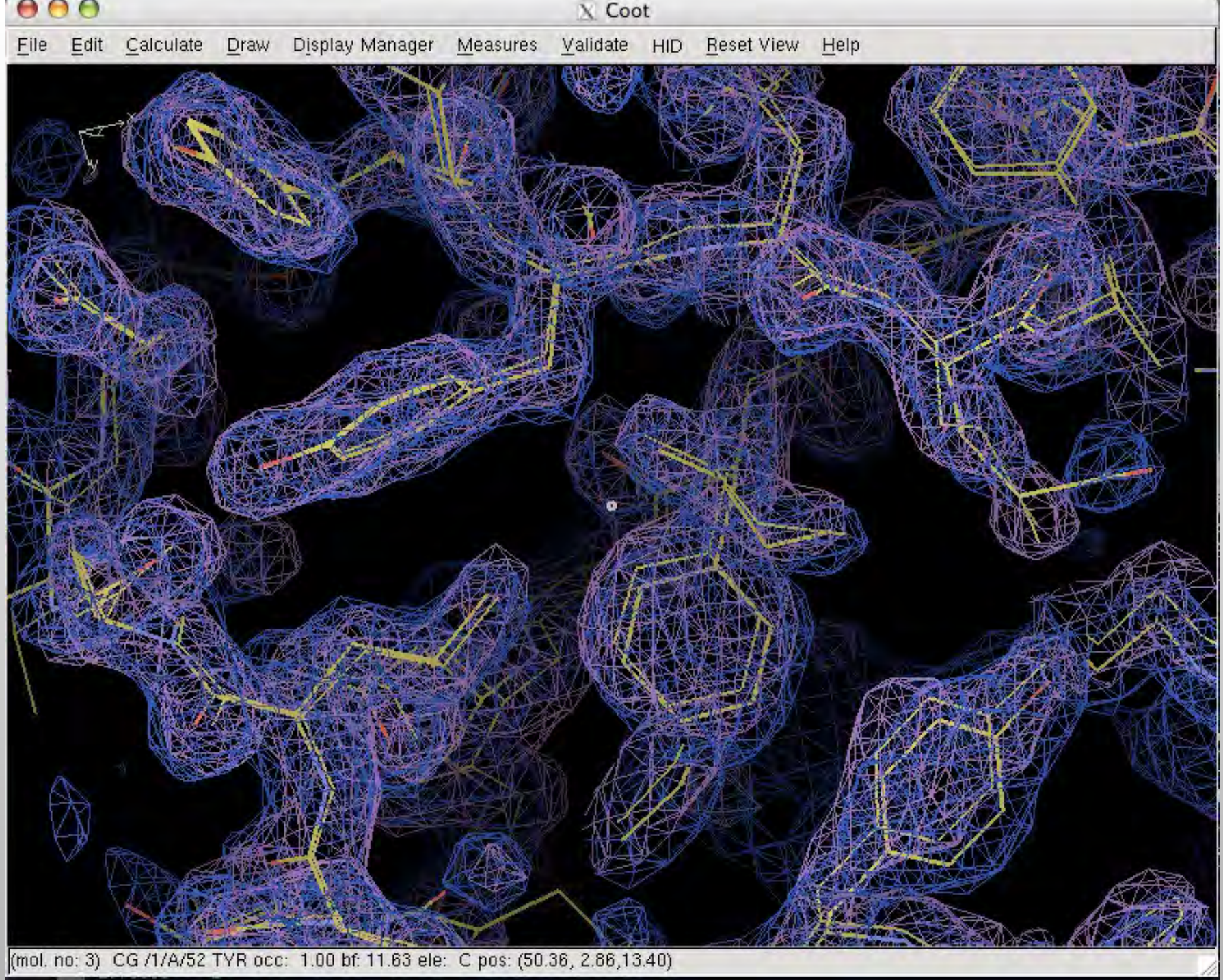


[*] Named by George Sheldrick


Coot
File Edit Calculate Draw Display Manager Measures Validate HID Reset View Help
(mol. no: 3) CG /1/A/52 TYR occ: 1.00 bf: 11.63 ele: C pos: (50.36, 2.86,13.40)

NCS Overlays



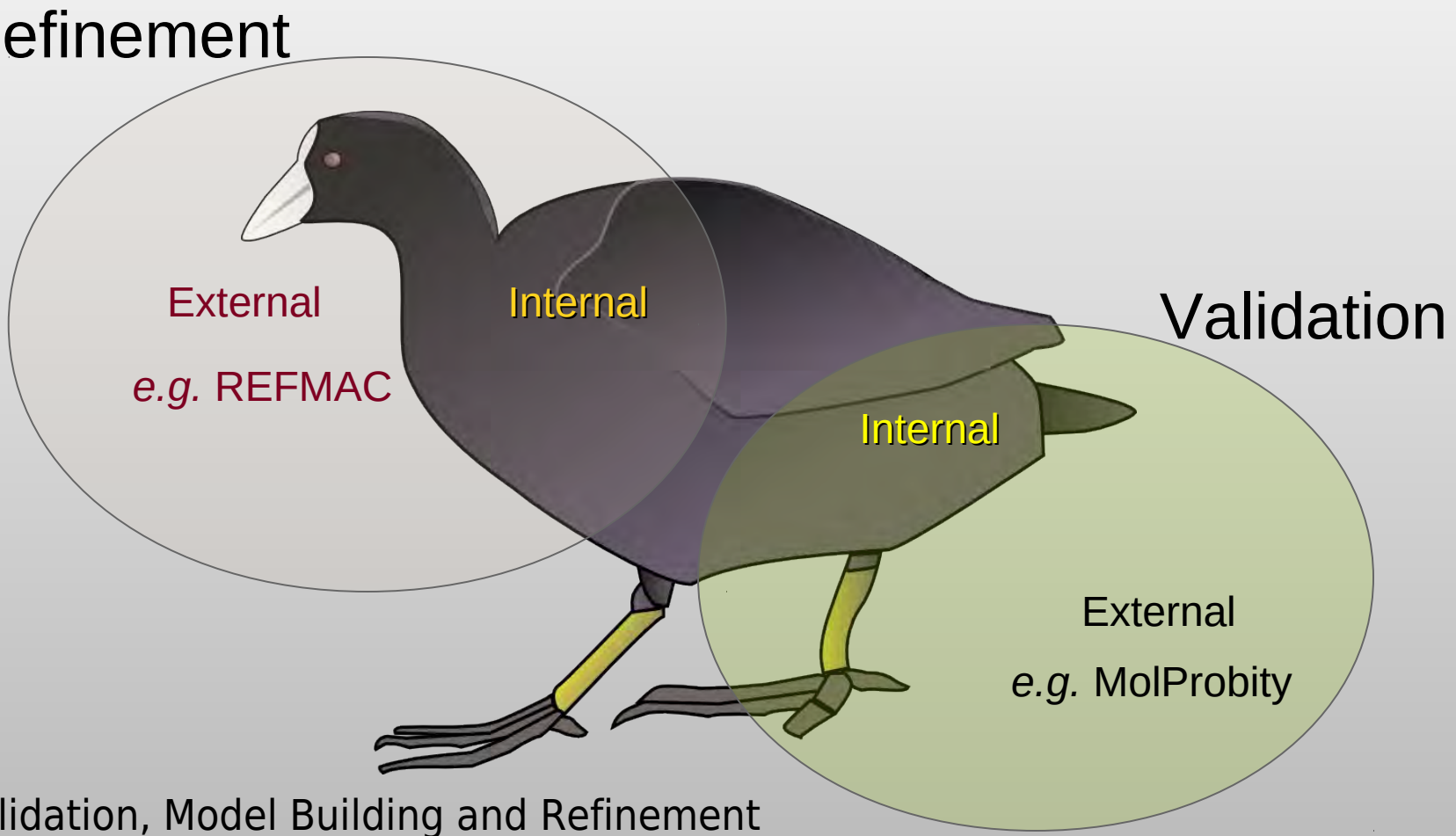

Coot
File Edit Calculate Draw Display Manager Measures Validate HID Reset View Help
(mol. no: 3) CG /1/A/52 TYR occ: 1.00 bf: 11.63 ele: C pos: (50.36, 2.86,13.40)

NCS Model-modification Tools

- Automatic detection of NCS
  - And their operators
- Copy Master NCS molecule to others
  - Applies NCS transformation
- Copy NCS Master residue-range
- Change NCS Master chain
- NCS Skipping ('o' key)

Validation...

Feature Integration



Validation, Model Building and Refinement should be used together

What is Validation?

- Comparison of various aspects of the model with pre-conceived notions of “good quality”
  - Includes **unrestrained** and **restrained** criteria
  - Many aspects of validation overlap with refinement and model-building
  - Global and local indicators

Why Validate?

- Model-building is error-prone
  - (although automated methods seem to do better)
- Map interpretation is subjective
    - Someone else did the model-building
- The model was built several years ago
  - and the notion of “good quality” has changed
- Deposition requires validation

A “good” model

- Makes statistical sense
  - The reciprocal space representation agrees tolerably well with the observations (R-factor)
    - No meaningful difference map peaks (no under- or over-fitting)
- Makes chemical sense
  - Model geometry is consistent with the restraints
  - Ramachandran Plot has less than 1% outliers
  - A good **clashscore**
- Make physical sense
    - Crystal packing and contacts
- Makes biological sense
  - Residues in chemically sensible environment
  - Is consistent (on the whole) with external biochemistry observations (active site residues)

Quick Bayes

- Bayes Eq:
- $Pr(model|data) \propto Pr(data|model) * Pr(model)$
- $Pr(data|model)$ is also called the model likelihood, $L_M(model|data)$

Validation Tools - Pr(model)

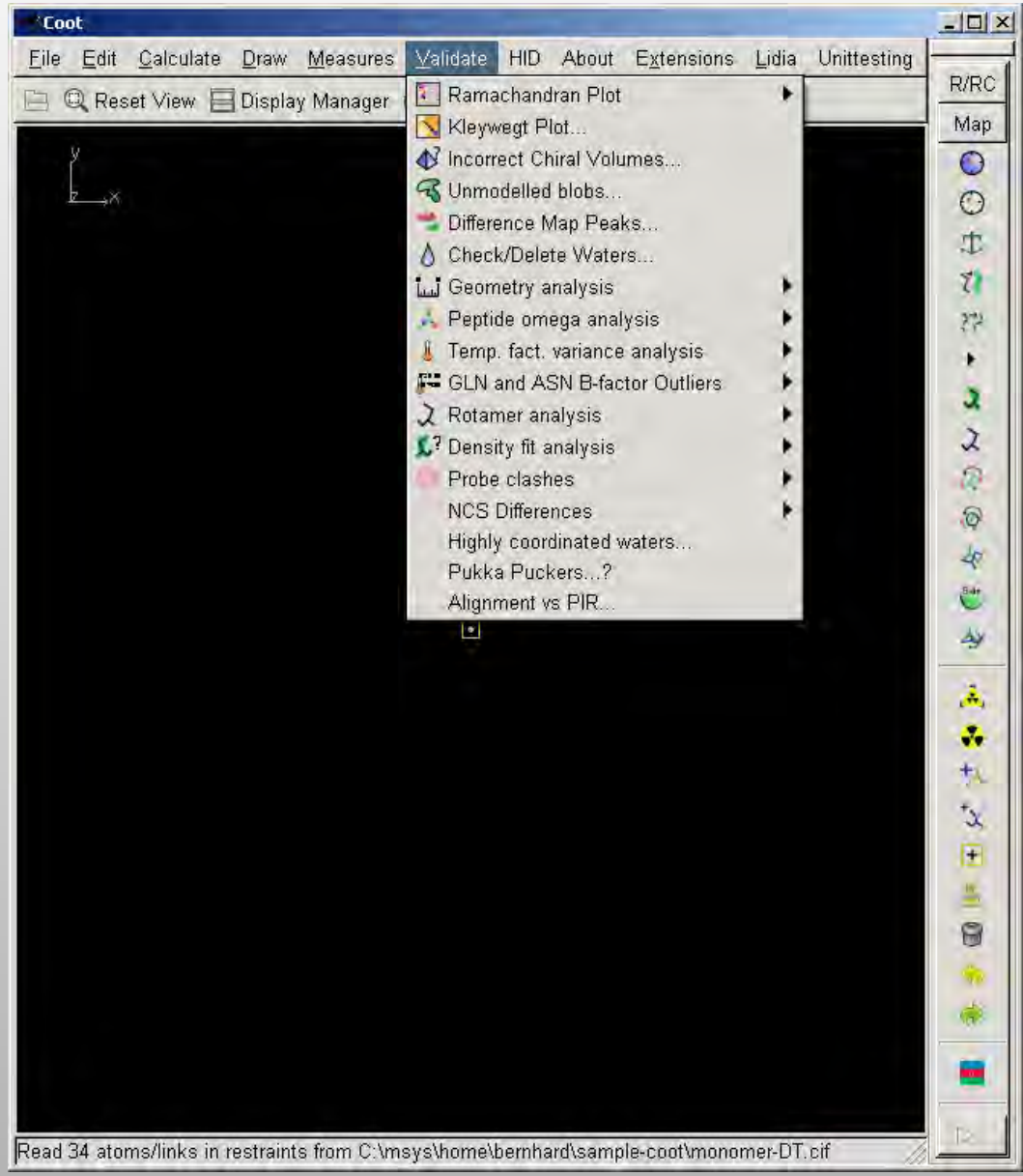
- Ramachandran Plot
  - Kleywegt Plot (NCS differences)
- Geometry Analysis
- Peptide ω Analysis
- Temperature Factor Analysis
- Rotamer Analysis
- Sequence Analysis
- Clashes

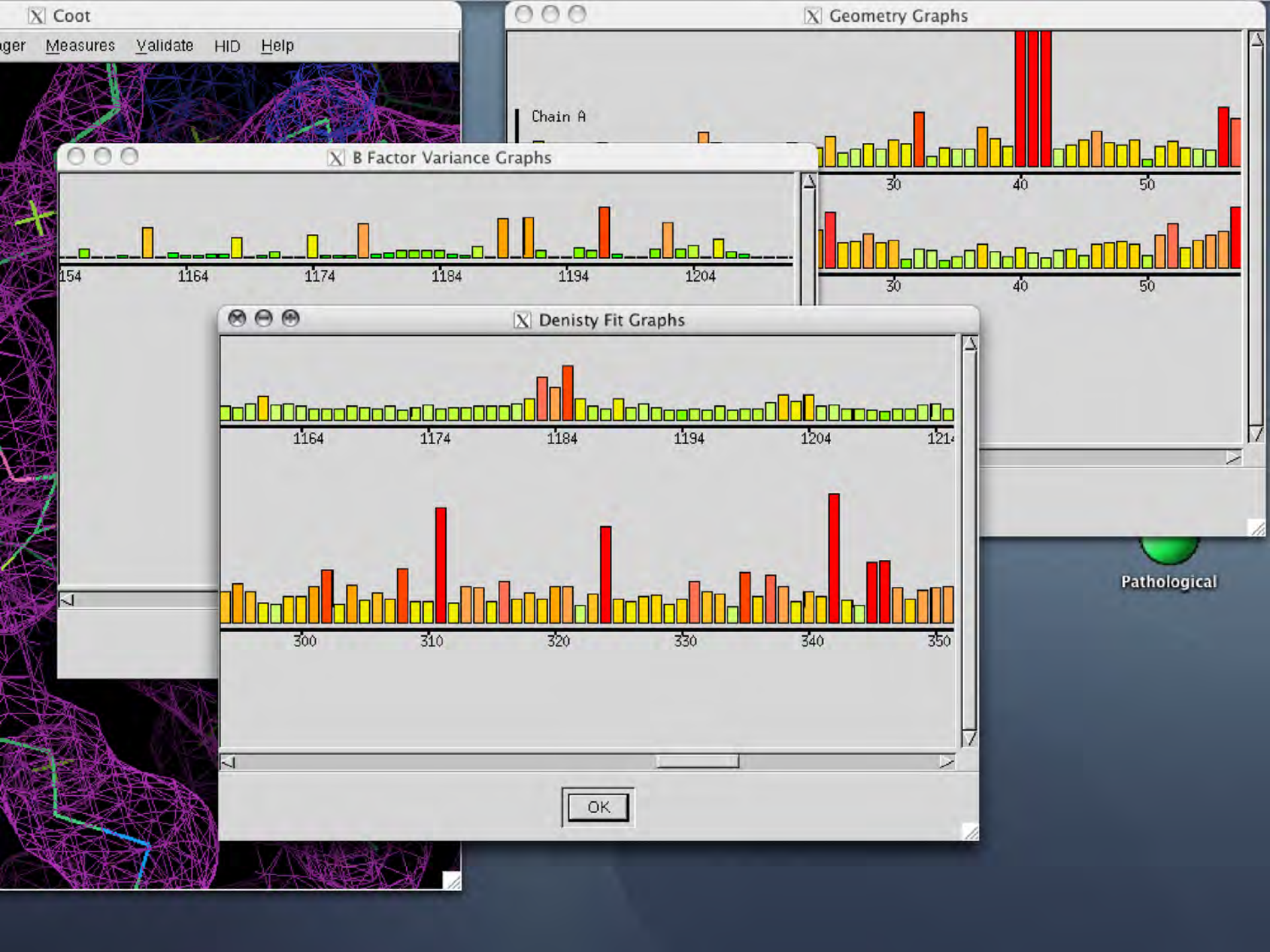
Validation Tools - Pr(data|model)

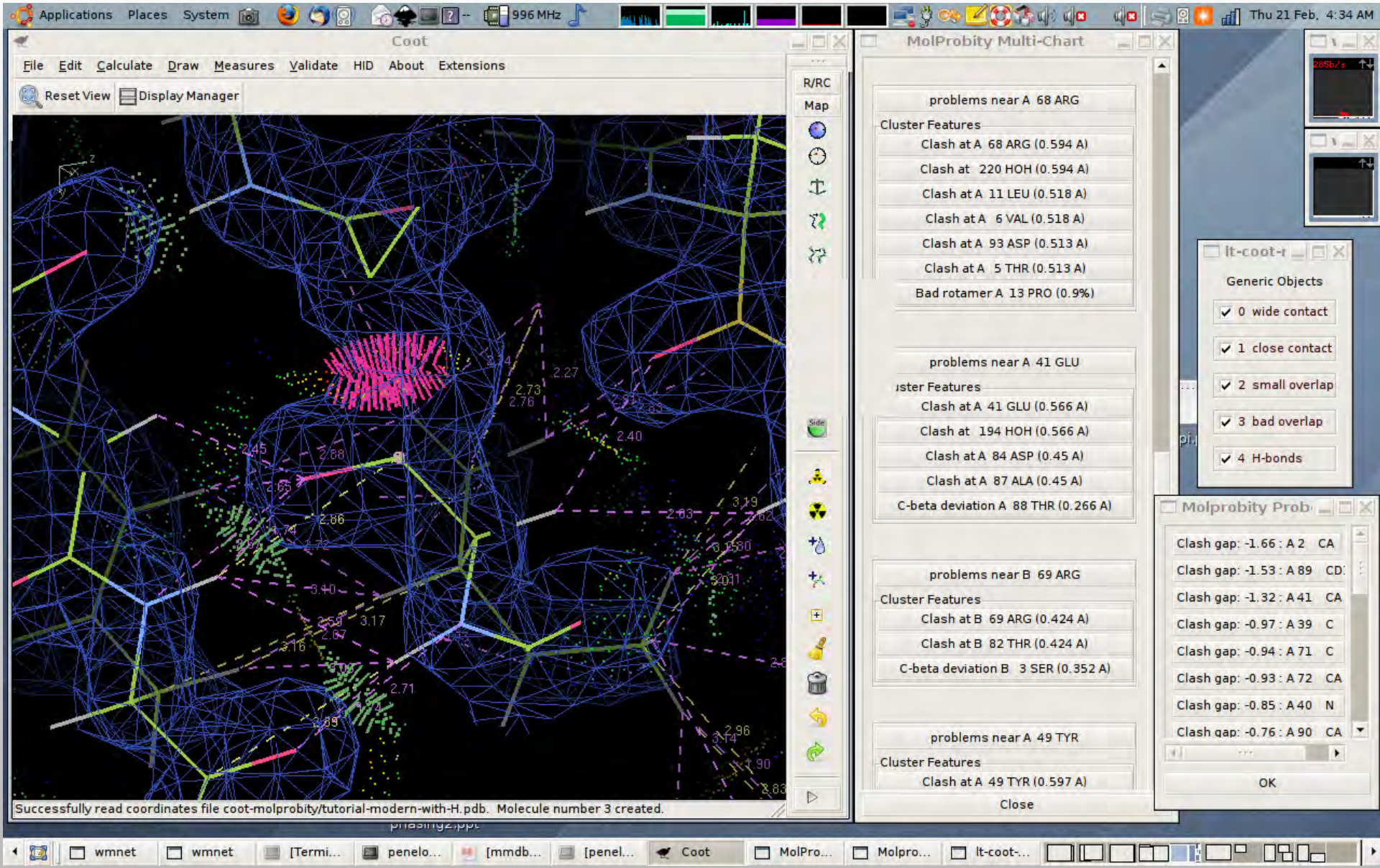
- Density Fit Analysis
- Difference Map Peaks
  - Variance analysis at water positions
- Unmodelled blobs

Validation tools

- Geometry Distortion
  - Bonds, Angles, Planes
- B-factor variance
- Density Fit
- Peptide Omega angle
- Rotamers
- Water check
  - By distance/map density/B-factor
  - By difference map variance
- Un-modelled density blobs
- Ramachandran Plot




Coot
ager Measures Validate HID Help
Geometry Graphs
Chain A
30
40
50
30
40
50
B Factor Variance Graphs
154
1164
1174
1184
1194
1204
Denisty Fit Graphs
1164
1174
1184
1194
1204
300
310
320
330
340
350
OK
Pathological


Applications Places System 996 MHz Thu 21 Feb, 4:34 AM
Coot
File Edit Calculate Draw Measures Validate HID About Extensions
Reset View Display Manager
R/RC
Map
Side
Successfully read coordinates file coot-molprobity/tutorial-modern-with-H.pdb. Molecule number 3 created.
MolProbity Multi-Chart
problems near A 68 ARG
Cluster Features
Clash at A 68 ARG (0.594 A)
Clash at 220 HOH (0.594 A)
Clash at A 11 LEU (0.518 A)
Clash at A 6 VAL (0.518 A)
Clash at A 93 ASP (0.513 A)
Clash at A 5 THR (0.513 A)
Bad rotamer A 13 PRO (0.9%)
problems near A 41 GLU
ıster Features
Clash at A 41 GLU (0.566 A)
Clash at 194 HOH (0.566 A)
Clash at A 84 ASP (0.45 A)
Clash at A 87 ALA (0.45 A)
C-beta deviation A 88 THR (0.266 A)
problems near B 69 ARG
Cluster Features
Clash at B 69 ARG (0.424 A)
Clash at B 82 THR (0.424 A)
C-beta deviation B 3 SER (0.352 A)
problems near A 49 TYR
Cluster Features
Clash at A 49 TYR (0.597 A)
Close
lt-coot-r
Generic Objects
0 wide contact
1 close contact
2 small overlap
3 bad overlap
4 H-bonds
Molprobity Prob
Clash gap: -1.66 : A 2 CA
Clash gap: -1.53 : A 89 CD
Clash gap: -1.32 : A 41 CA
Clash gap: -0.97 : A 39 C
Clash gap: -0.94 : A 71 C
Clash gap: -0.93 : A 72 CA
Clash gap: -0.85 : A 40 N
Clash gap: -0.76 : A 90 CA
OK
wmnet wmnet [Termi... penelo... [mmdb... [penel... Coot MolPro... Molpro... lt-coot-...

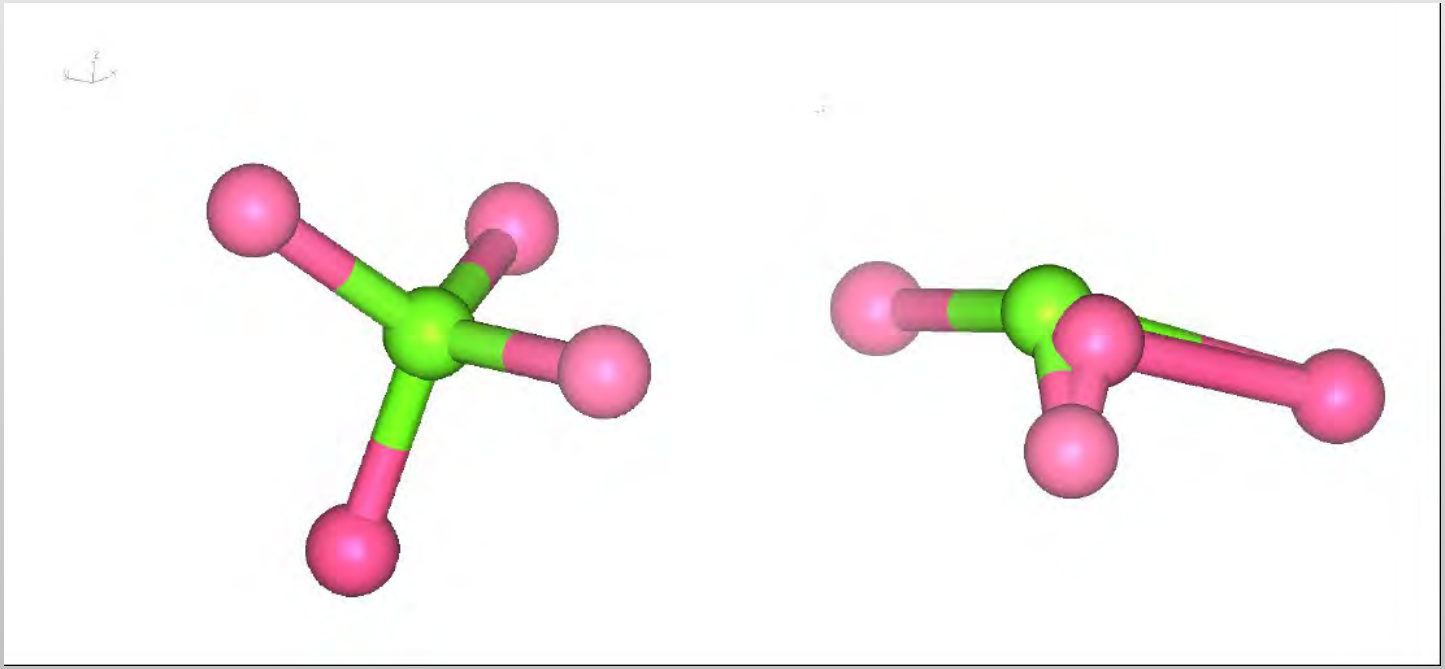
Other Programs

- Molprobity Suite
  - molprobity.biochem.duke.edu
- WHATCHECK
- VERIFY-3D

Ligand validation

Ligand validation
Why validate?

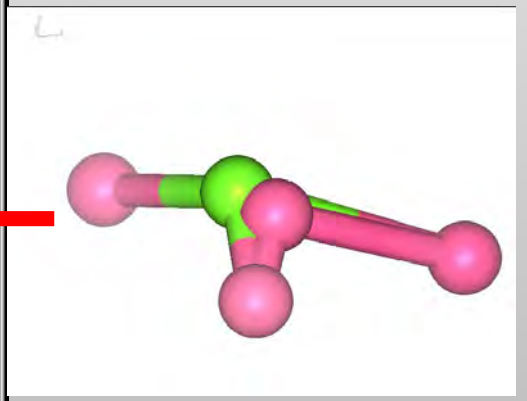
- 70.000 structures in PDB
- 11.000 'ligands' in 50.000 structures



- R/Rfree 0.15/0.19

Ligand validation

- How to validate ligand geometry?
  - Compare observed structure to restraints
  - Coot: Validate → Geometry analysis



Parametrisation issues... (what if they are wrong?)

- Perfect refinement with incorrect parameters → distorted structure
- Need independent parameters to test against

How to validate ligand geometry?

- QM
  - CPU hungry
  - In vaccuo
  - Low energy ≠ (?) bound ligand
- PDB (e.g. ValLigURL)
  - Good cofactor structures
  - Less useful for novel ligands
  - (occ.) questionable quality
- Databases [Cambridge Structural Database (CSD) or COD]
  - Using small molecule geometries
  - e.g. Mogul

Further information

- Coot WIKI
  - http://strucbio.biologie.uni-konstanz.de/ccp4wiki/index.php/Coot
- Coot BB (mailing list)
  - http://www.jiscmail.ac.uk/lists/coot.html
- Coot documentation
  - http://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/web/docs/

Acknowledgements

- Paul Emsley
- Kevin Cowtan
- Eleanor Dodson
- Keith Wilson

http://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/

or

Google: Coot

or for WinCoot

http://www.ysbl.ac.uk/~lohkamp/coot

- Libraries, dictionaries
  - Alexei Vagin, Eugene Krissinel, Stuart McNicholas
  - Dunbrack, Richardsons
- Coot Builders and Testers
  - William Scott, Ezra Peisach
  - York YSBL, Dundee, Glasgow (early adopters)
  - Coot Mailing List subscribers