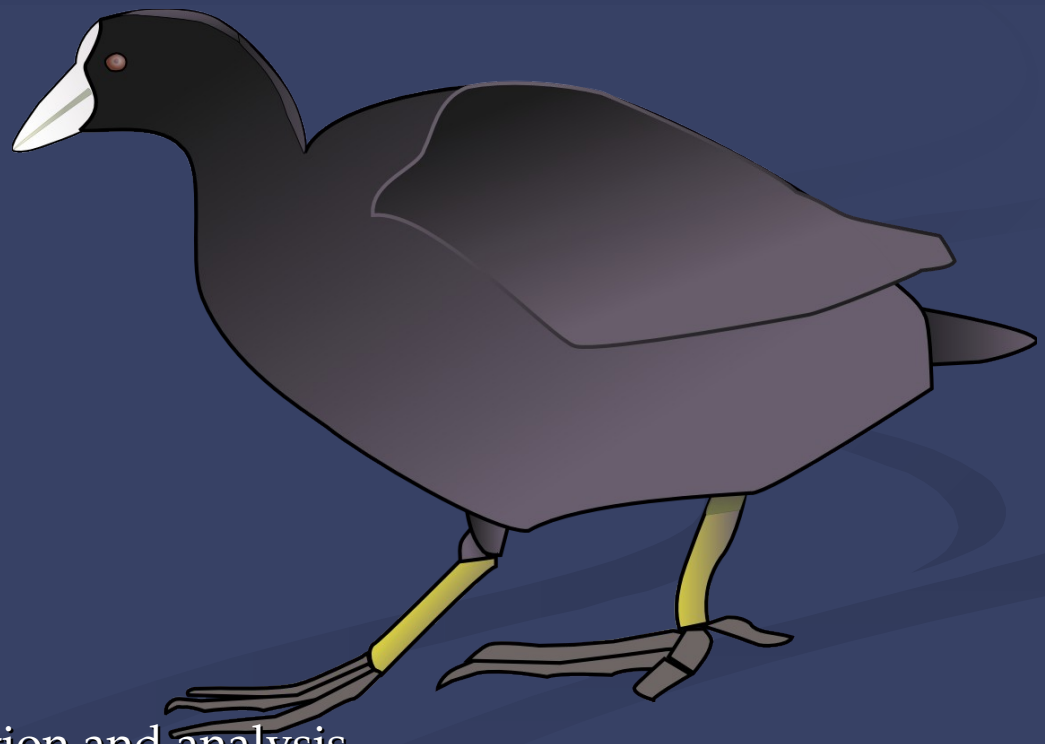

Modern Methods of
Ligand Refinement
and Analysis
Using Coot
Judit Debreczeni

Ligands in Coot

- Importing and building ligands from scratch
  - prodrg, libcheck
  - Monomer library search, Sbase search
  - 2D Ligand builder
  - Atom name and torsion match
- Ligand fitting
- Validation
  - Mogul
- Representation
  - Bond odrders
  - Surfaces
- Analysis
  - Molprobity
  - Lidia:
    Ligand display for interaction and analysis

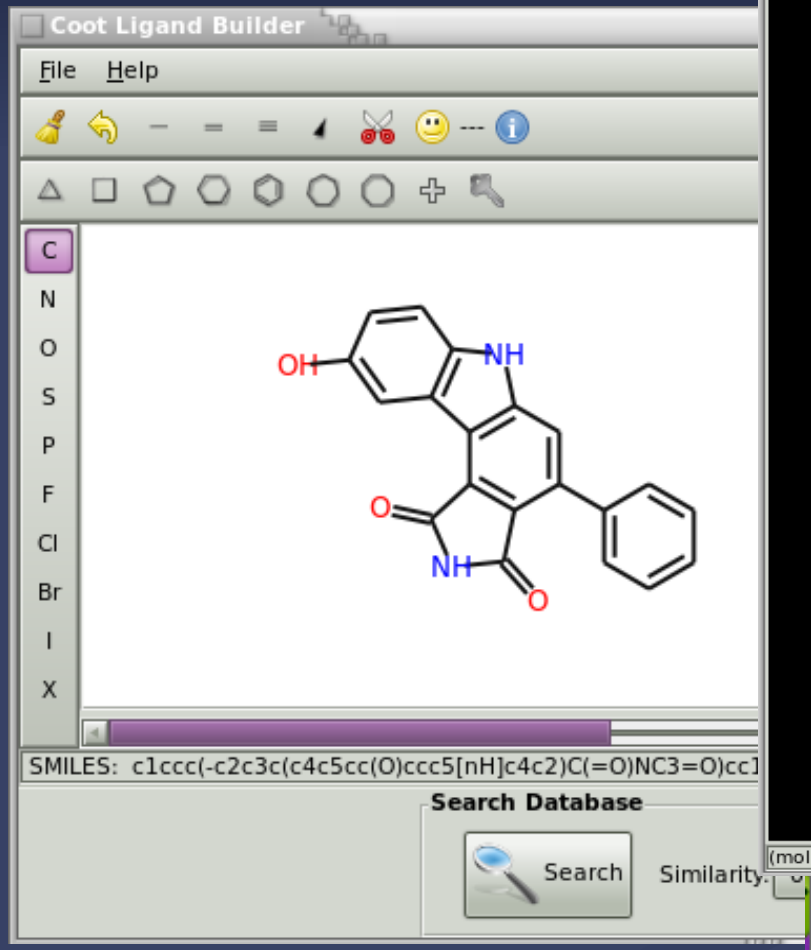


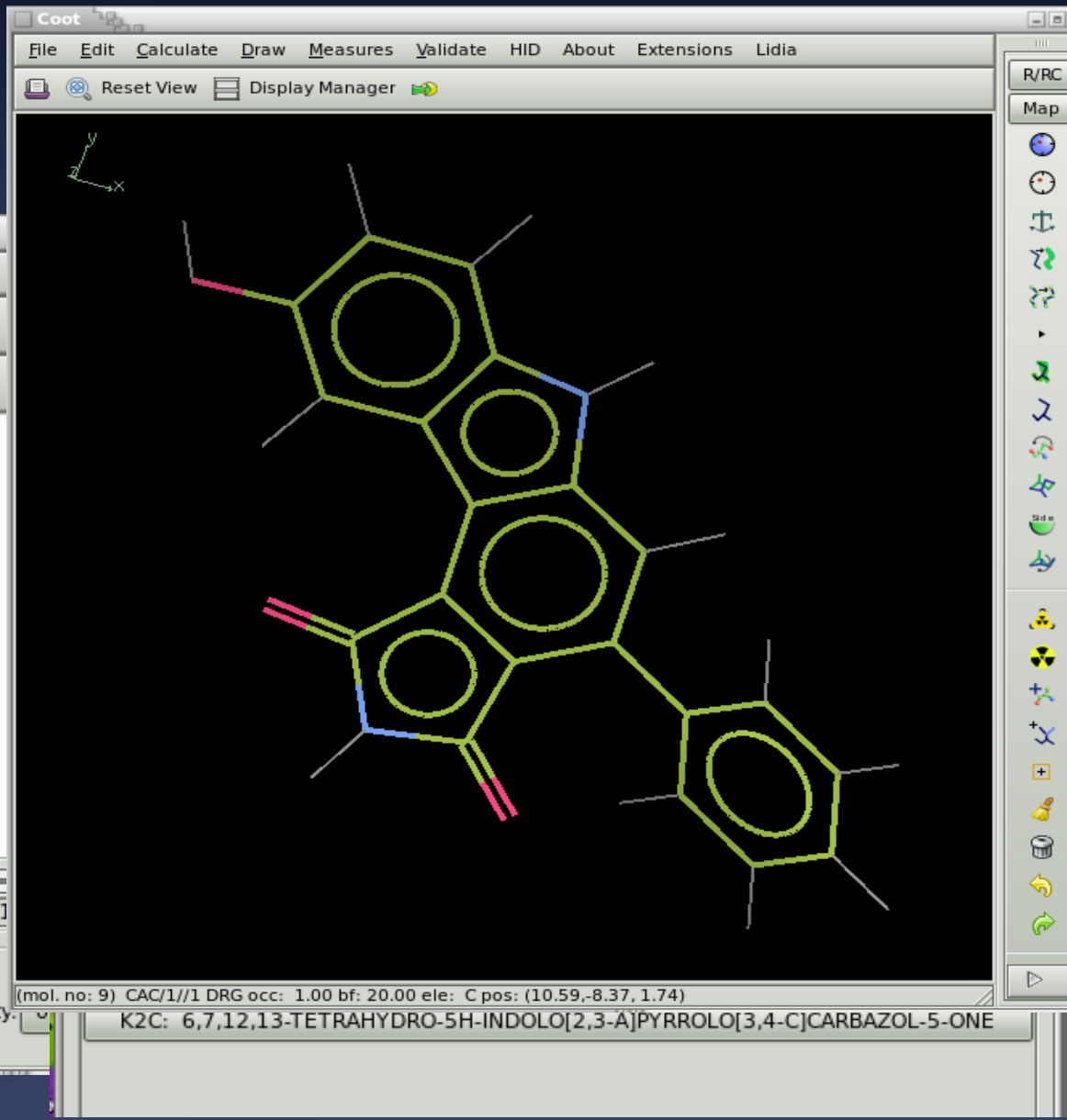
Getting started...

- Coordinates (pdb) and dictionary (cif)
  (Import CIF dictionary)
- libcheck interface (smi)
  (SMILES...)
- monomer library search
  (Get monomer..., Search monomer library...)
- Restraints editor gui
- prodrg interface
  (Prodrgify this residue...)
- 2D editor
- sbase search

2D Ligand Builder

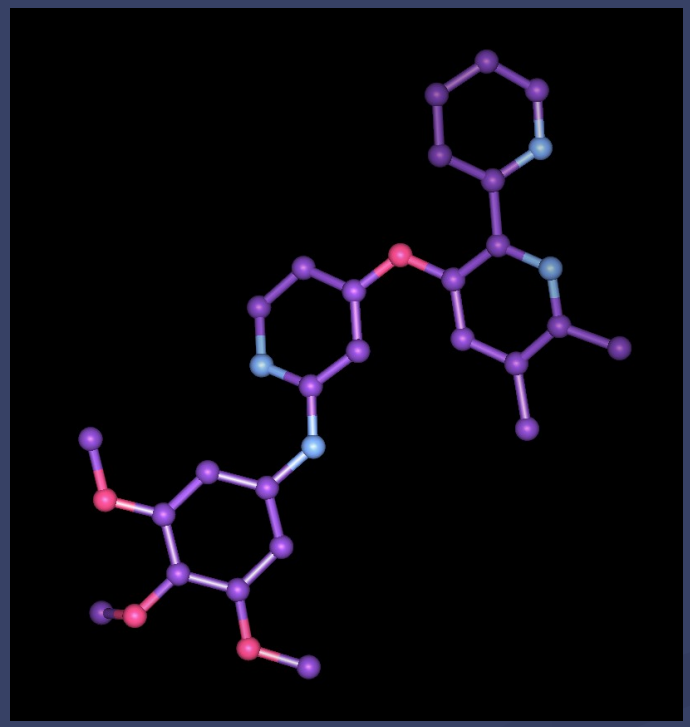
- Free sketch
- sbase search

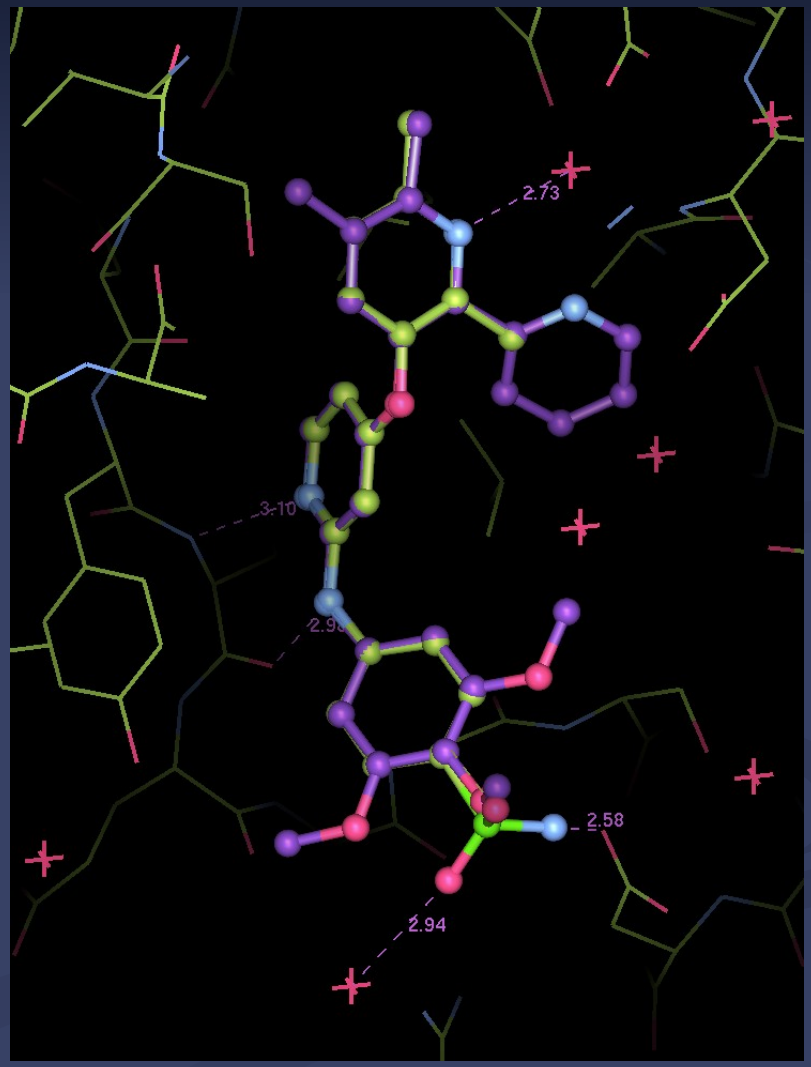




Yesterday's ligands...

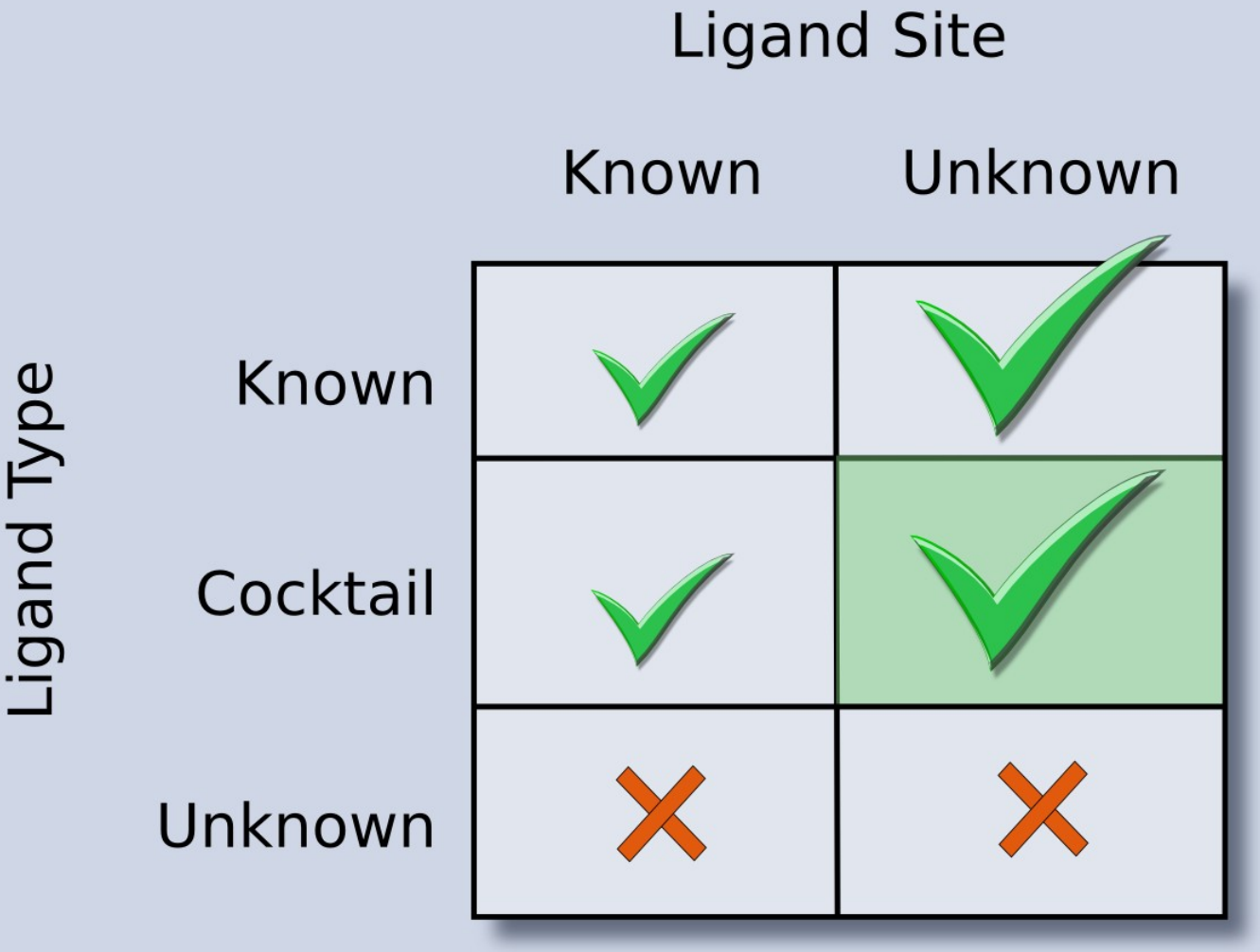
- Atom name matching
- Torsion matching
- Ligand overlay

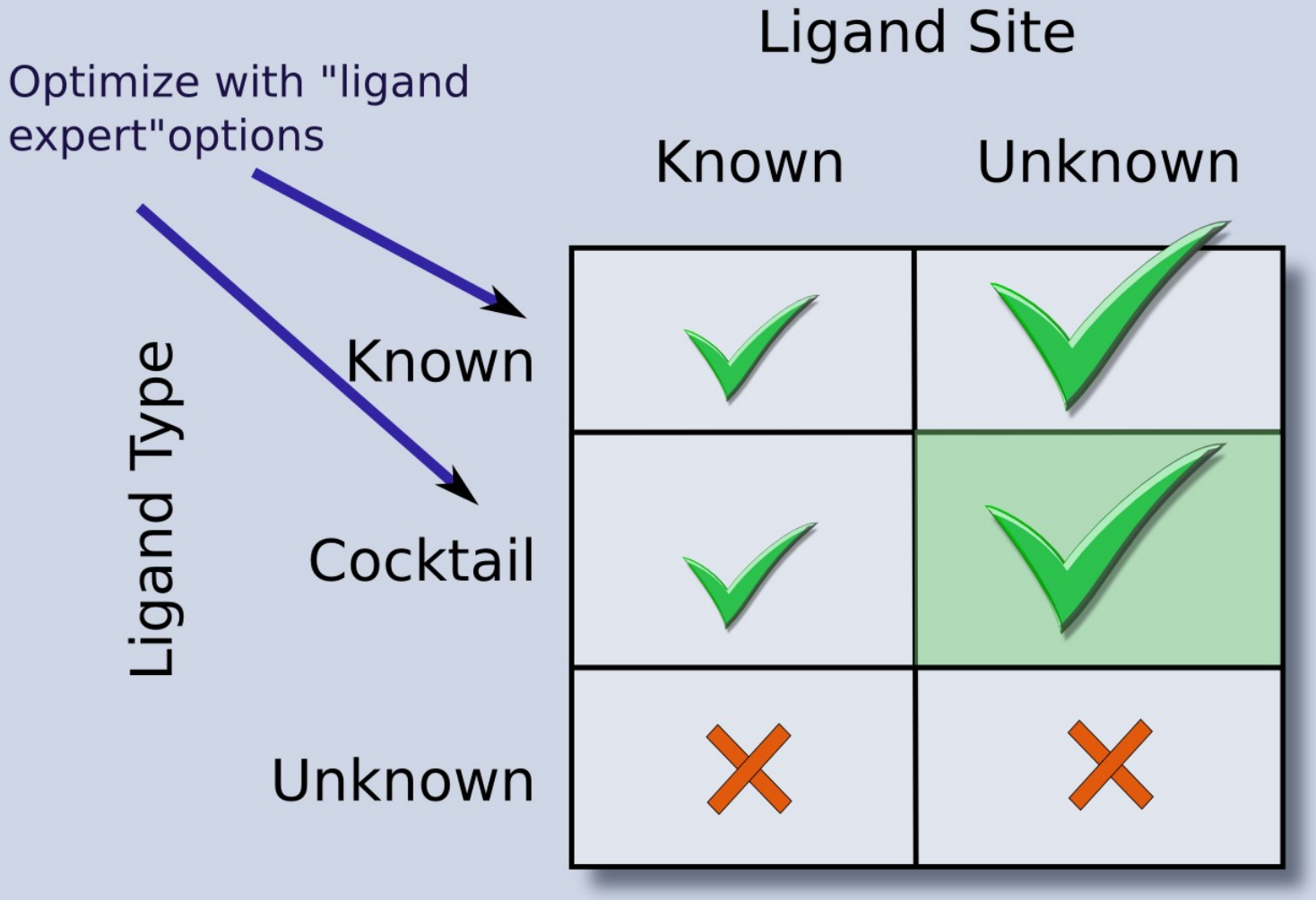




Ligand Fitting

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


Ligand Site
Known
Unknown
Ligand Type
Known
Cocktail
Unknown


Optimize with "ligand expert"options
Ligand Site
Known
Unknown
Ligand Type
Known
Cocktail
Unknown

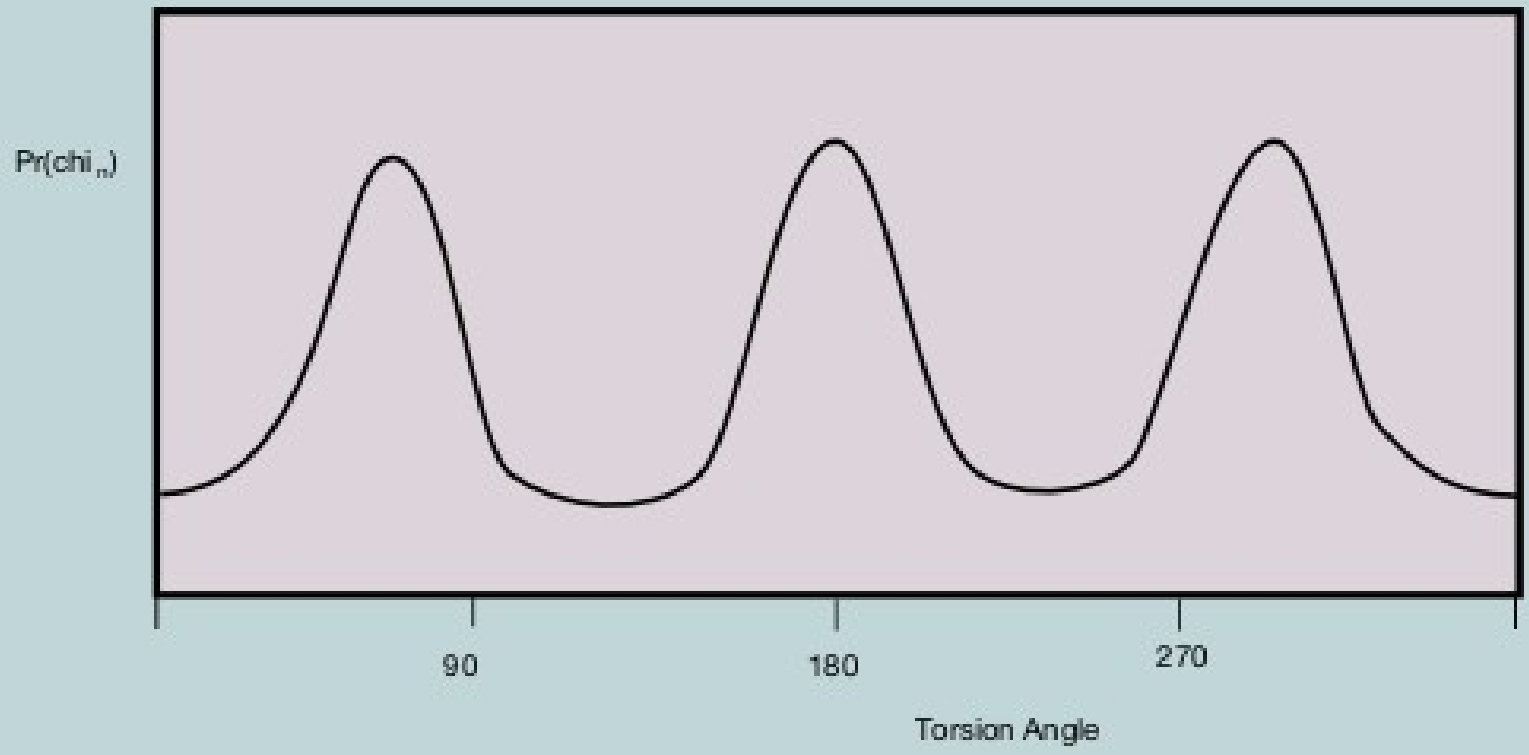
REFMAC Monomer Library chem_comp_bond

```
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
```

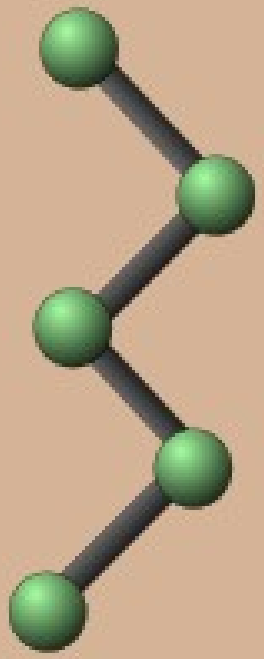
REFMAC Monomer Library chem_comp_tor

```
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
 ADP   var_1 O2A PA  O3A    PB    59.999   20.000   1
 ADP   var_2 PA  O3A PB     O1B   60.003   20.000   1
 ADP   var_3 O2A PA  O5*    C5*   -59.997  20.000   1
 ADP   var_4 PA  O5* C5*    C4*   180.000  20.000   1
 ADP   var_5 O5* C5* C4*    C3*   176.890  20.000   3
 ADP   var_6 C5* C4* O4*    C1*   150.000  20.000   1
 ADP   var_7 C5* C4* C3*    C2*  -150.000  20.000   3
```

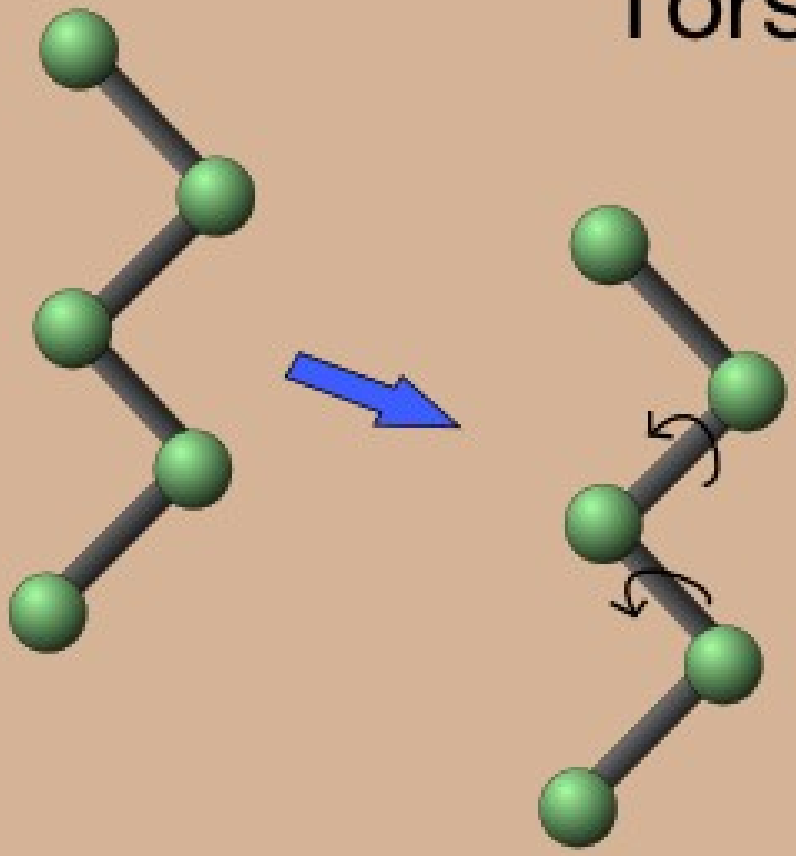
Ligand Torsionable Angle Probability from CIF file



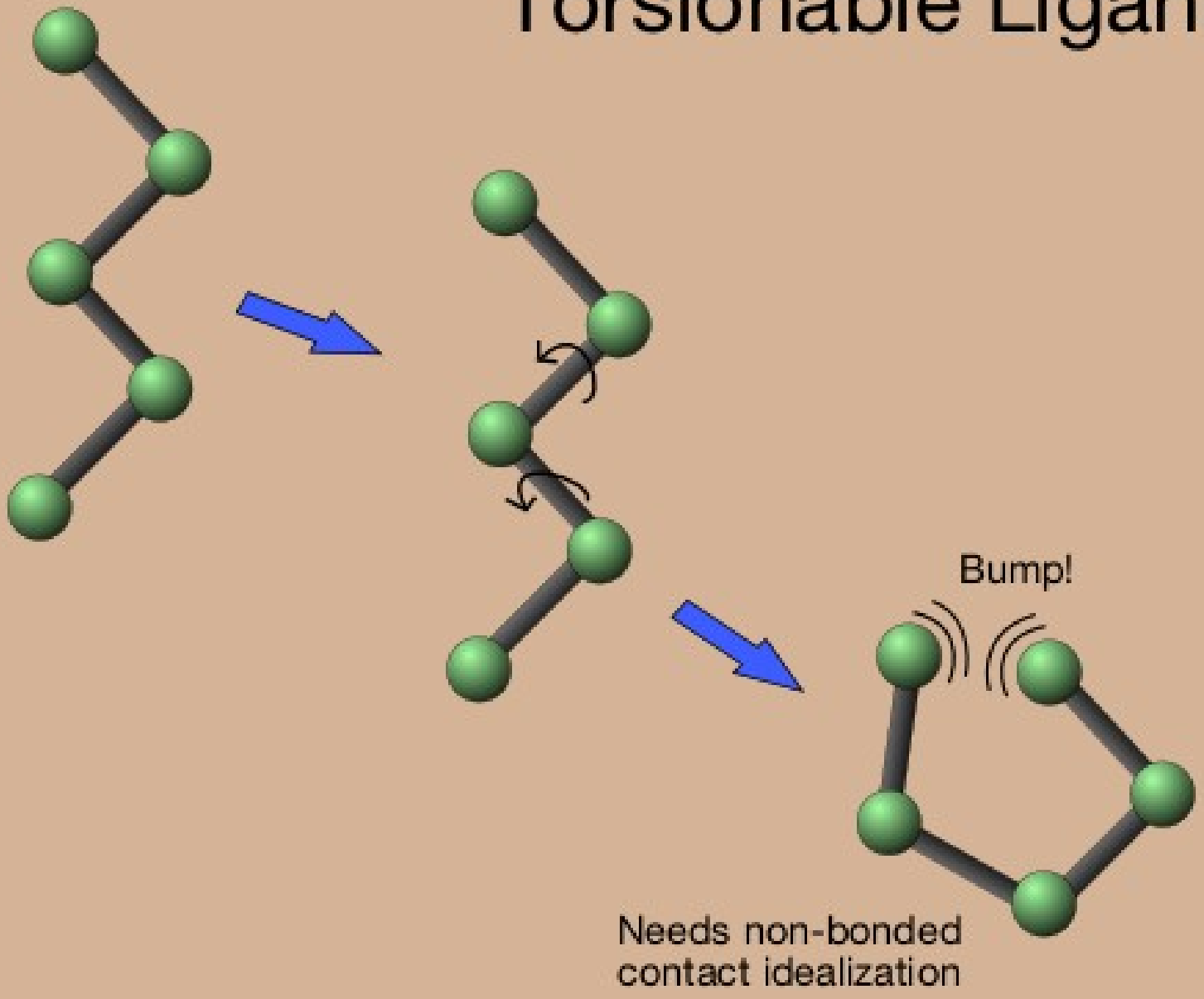
Torsionable Ligands



Torsionable Ligands



Torsionable Ligands



Crystal Space

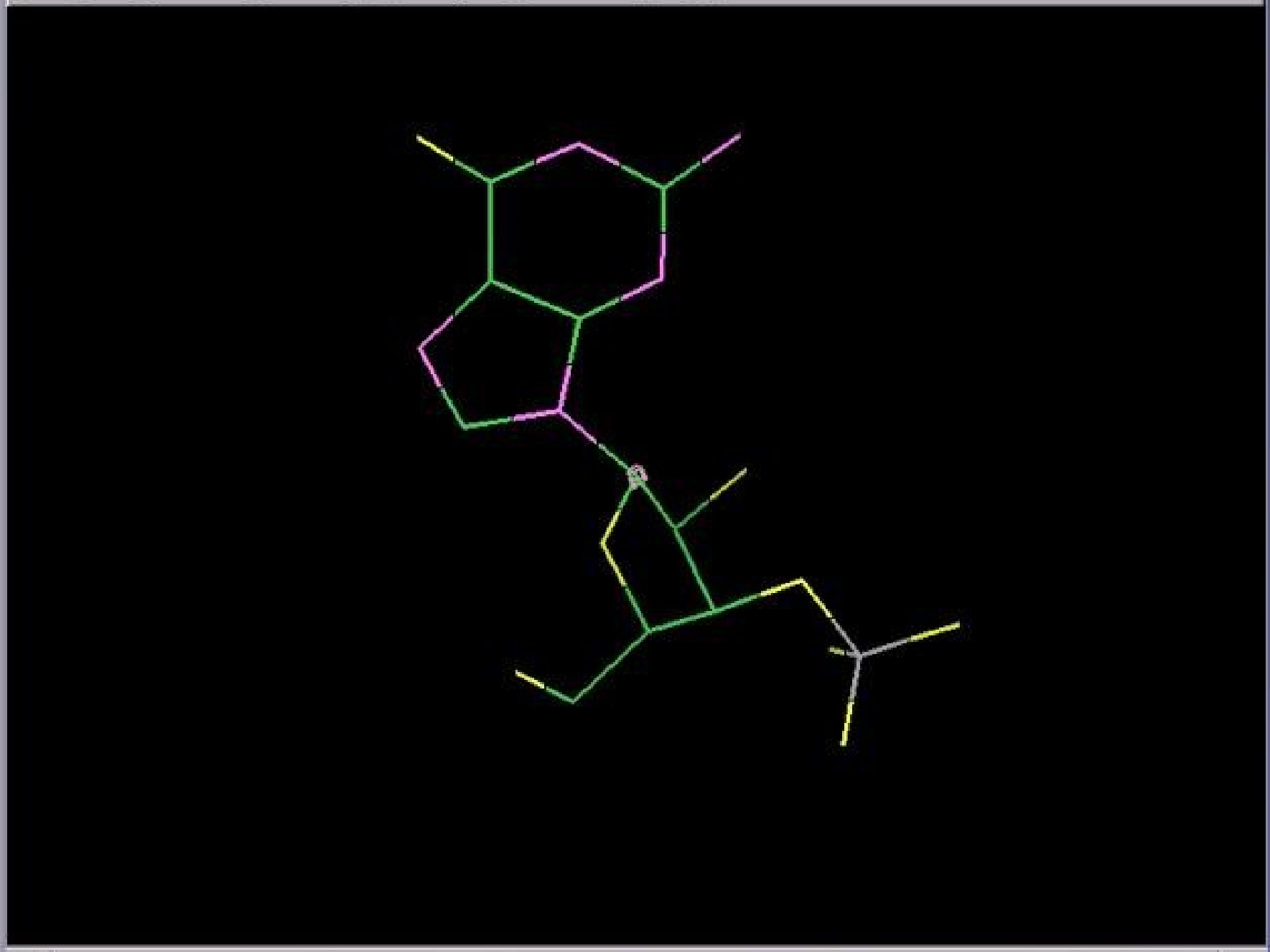
- Build in “crystal space”
  - Like real-space, but wrapped by crystal symmetry
    - Like “Asteroids”
  - Assures only one real-space representation of map features
    - Build everything only once,
    - No symmetry clashing
  - However, more difficult to calculate real space geometries
    - …such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

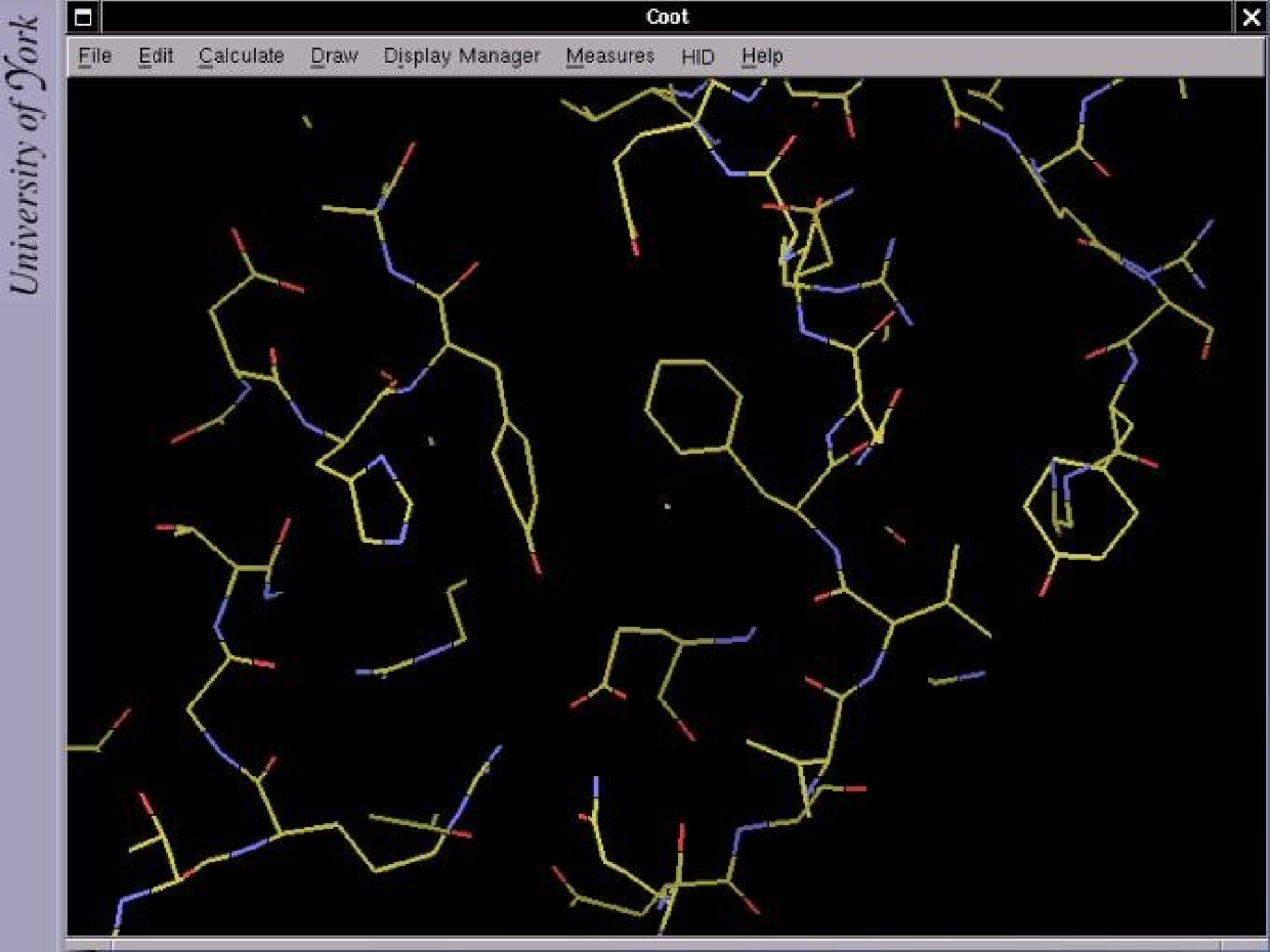
Crystal Space

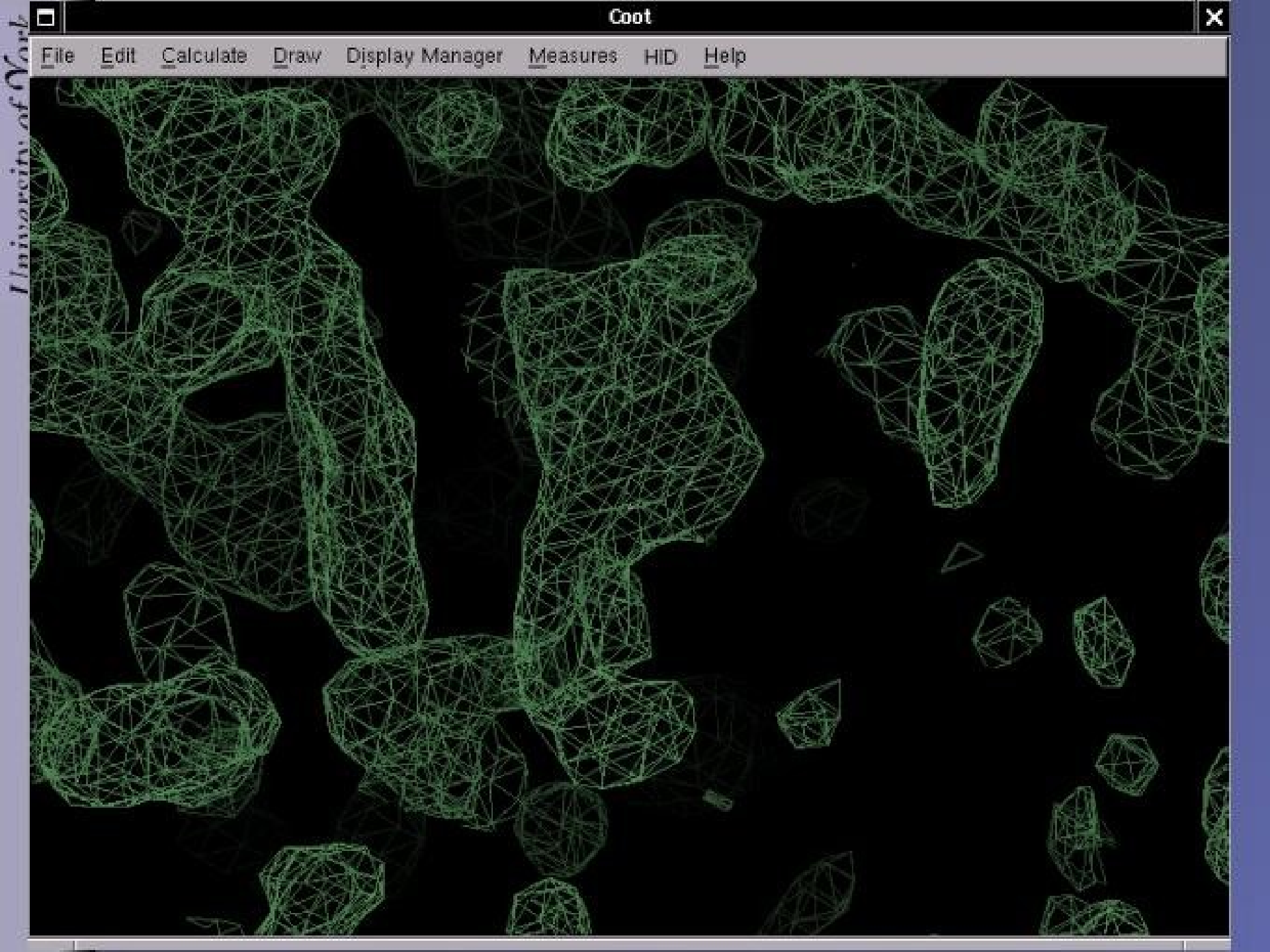
- Building in crystal space is good:
  - We don't need to define where the protein is and create an extended map that surrounds it
  - We don't have to worry about the relative position of the ligand and the protein
    - Unknown "BORDER" parameter
  - We find (and fit) each site exactly once
    - No symmetry problems

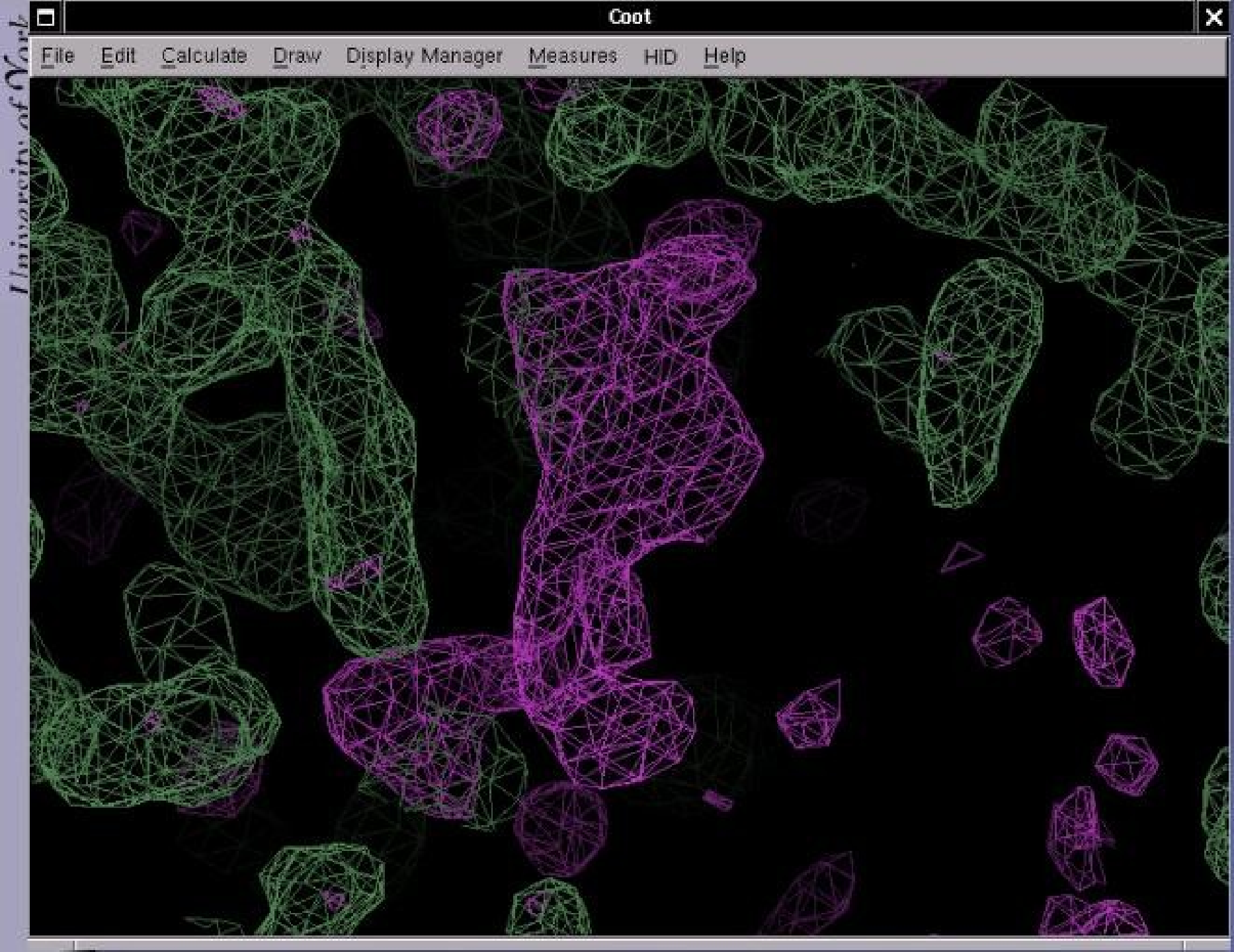
Coot

File Edit Calculate Draw Display Manager Measures HID Help



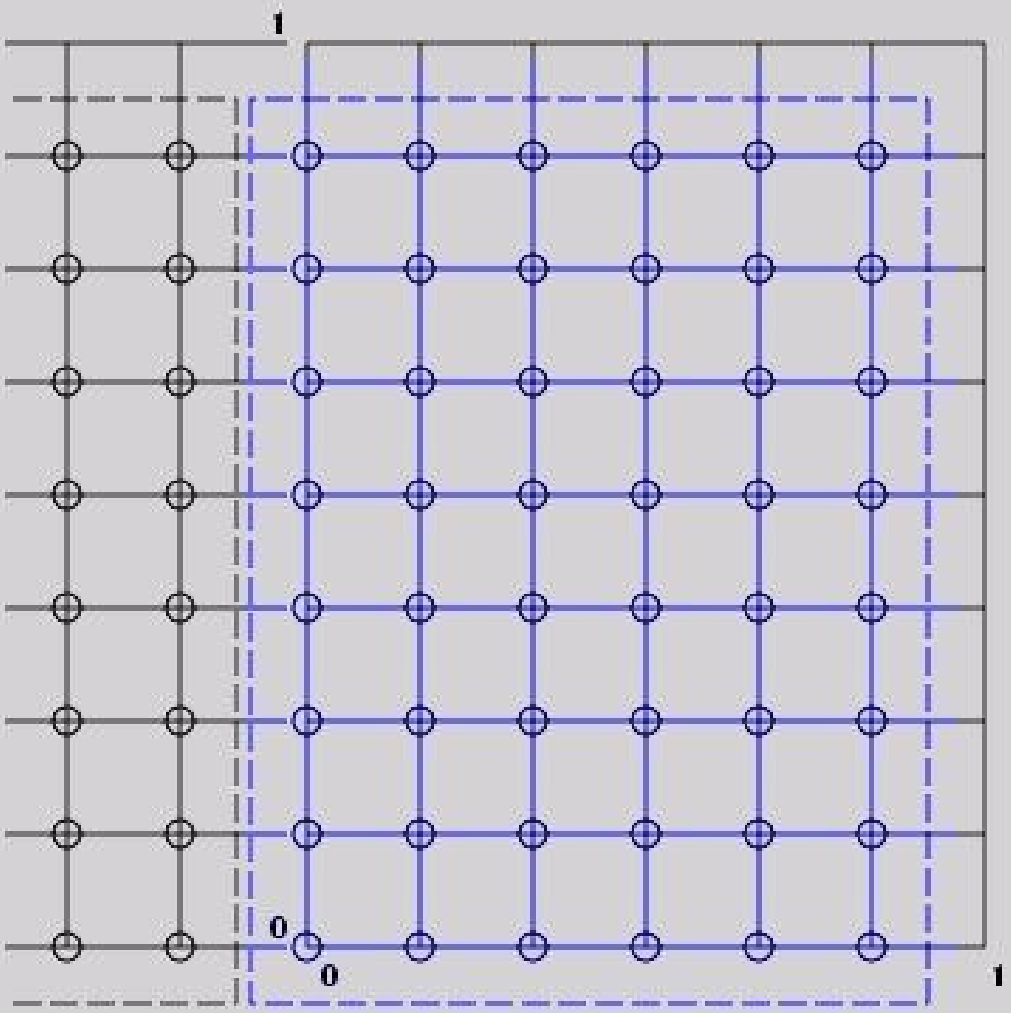

University of York
Coot
File
Edit
Calculate
Draw
Display Manager
Measures
HID
Help

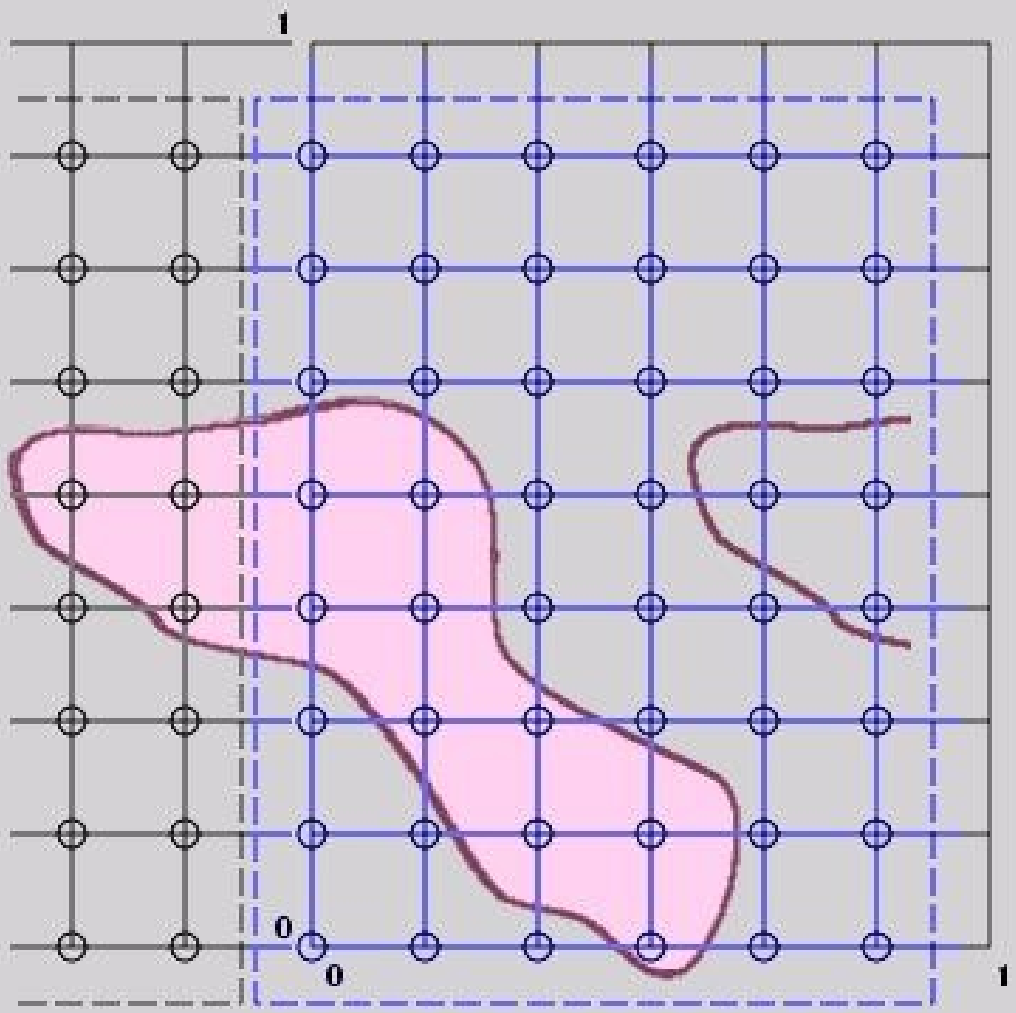

Coot
File
Edit
Calculate
Draw
Display Manager
Measures
HID
Help
University of York

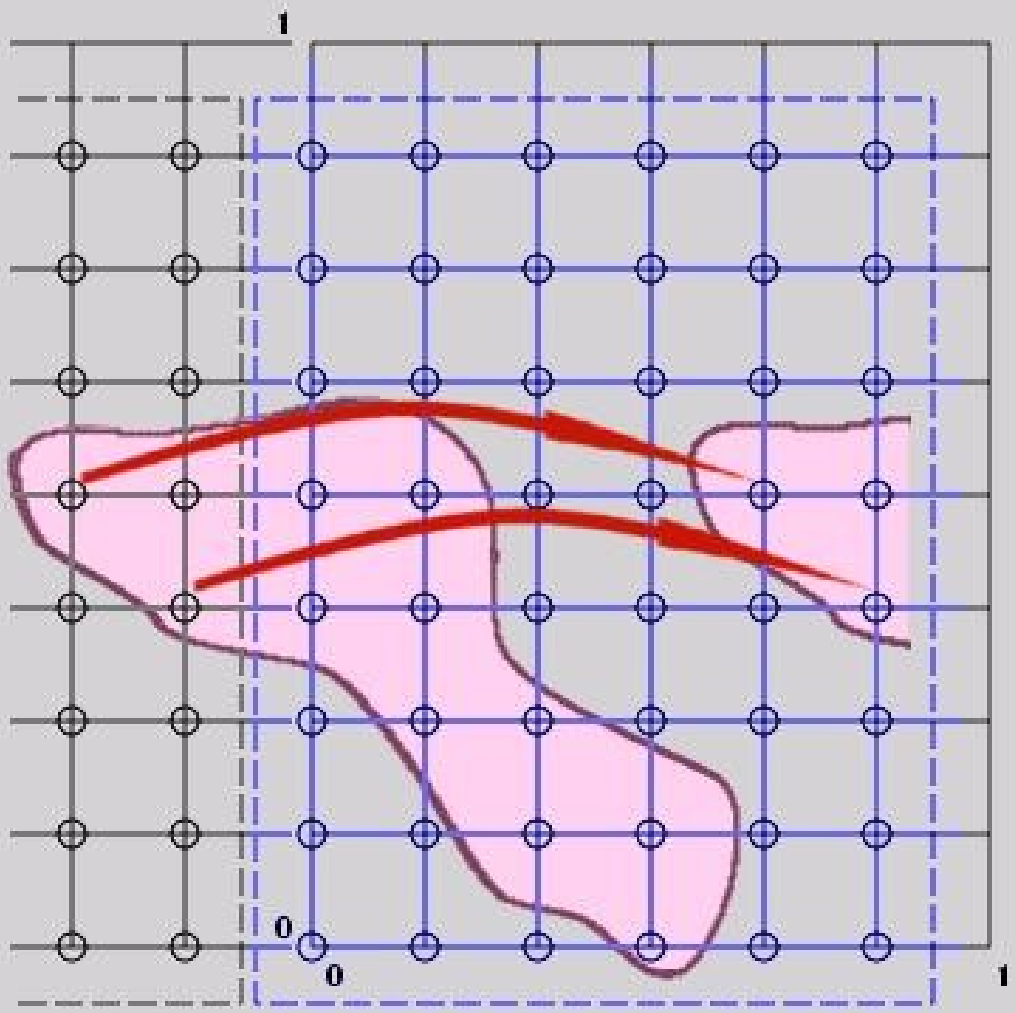

Coot
File
Edit
Calculate
Draw
Display Manager
Measures
HID
Help
University of York

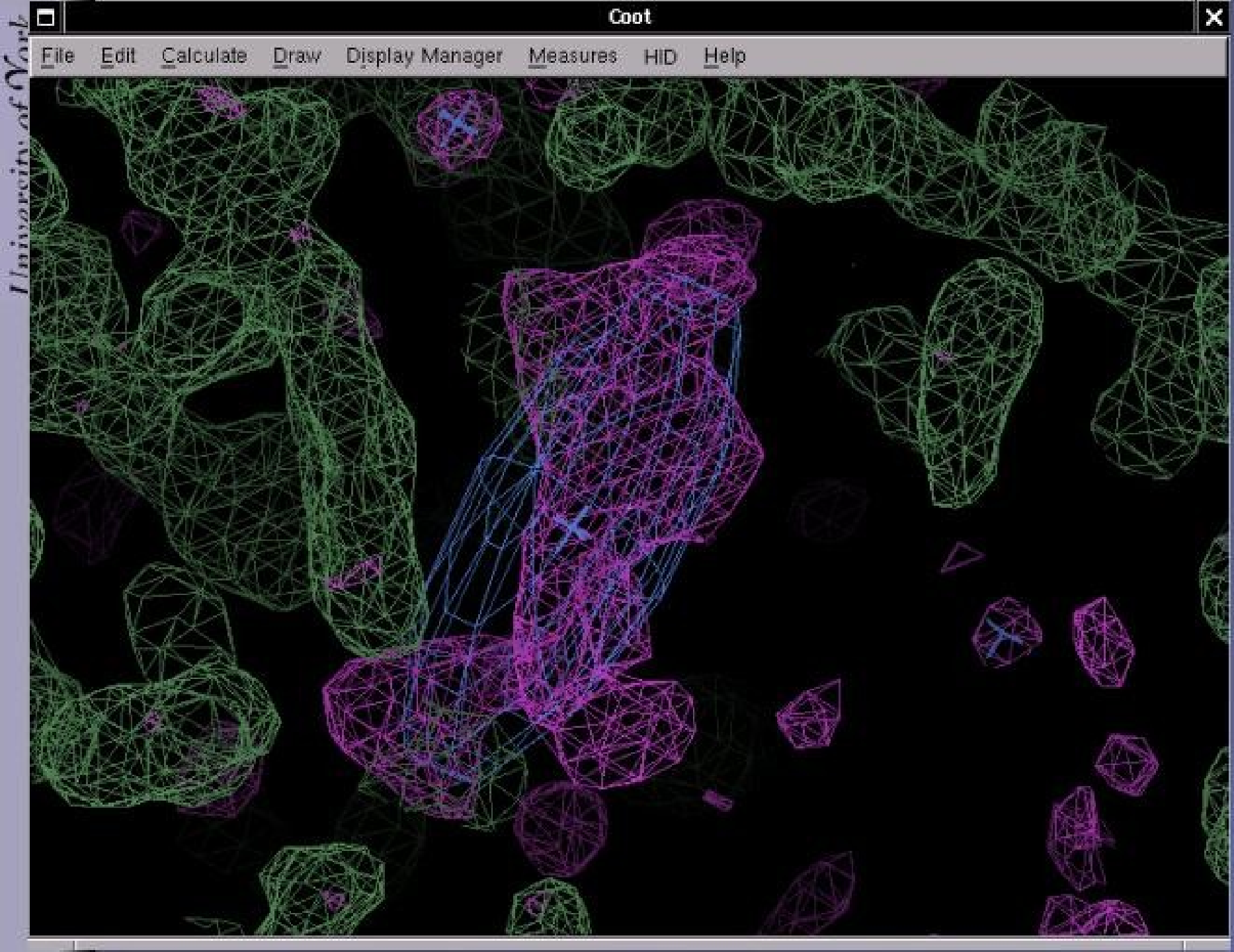
Clipper Map Mapping

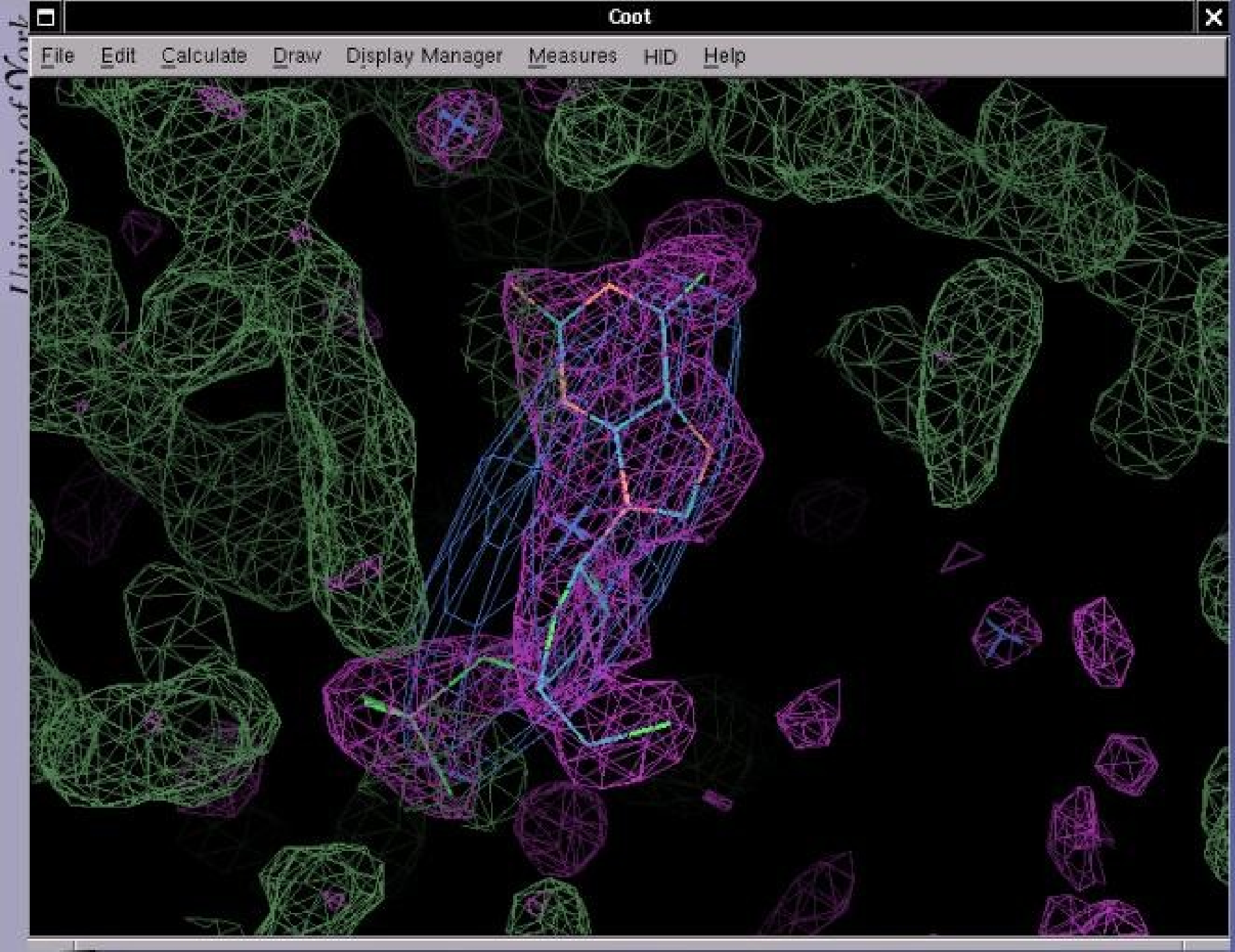
- Clipper maps
  - Appear to be “infinite”
  - Density value can be queried anywhere in space

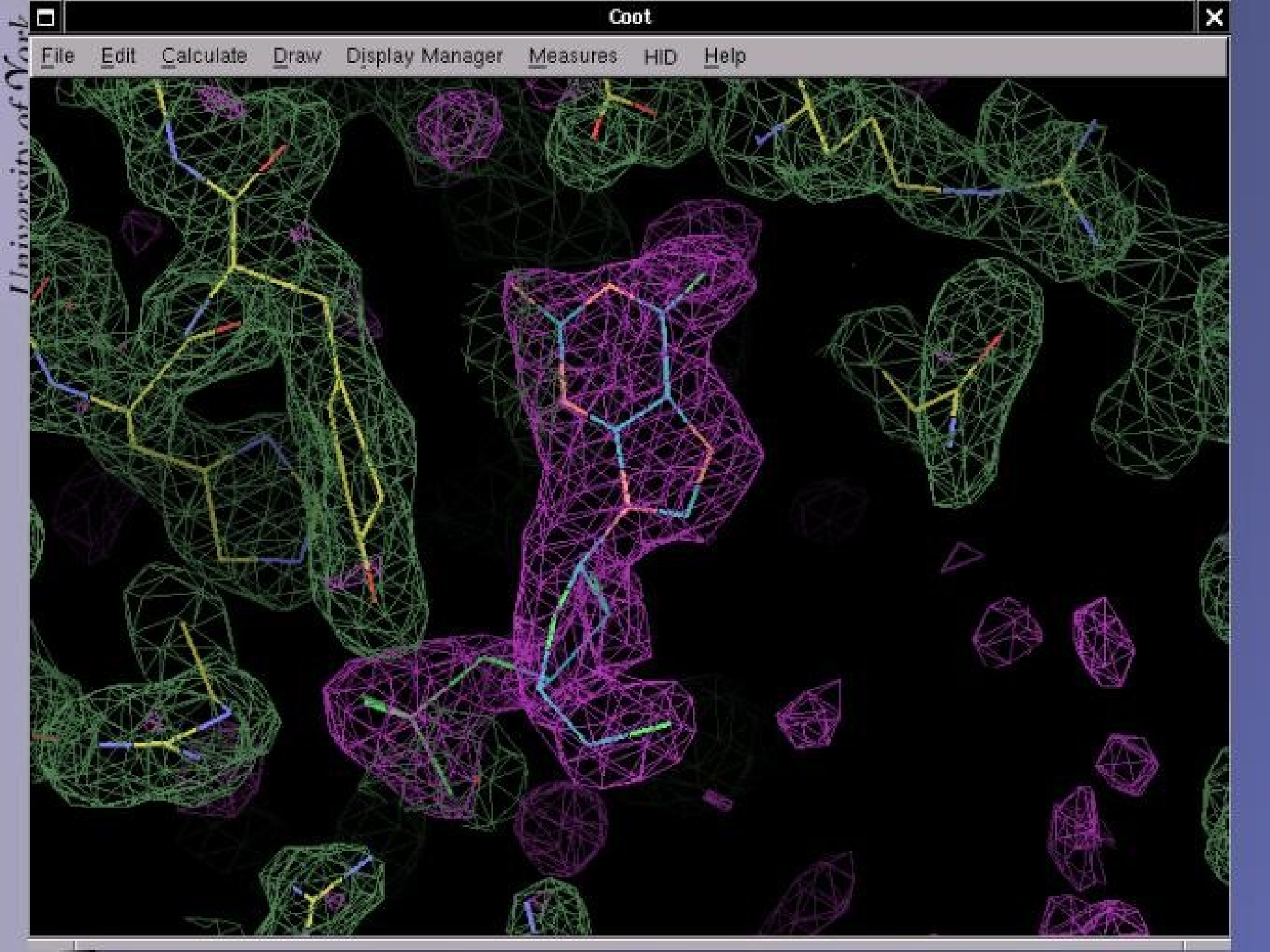

1
0
0
1


1
0
0
1


1
0
0
1


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

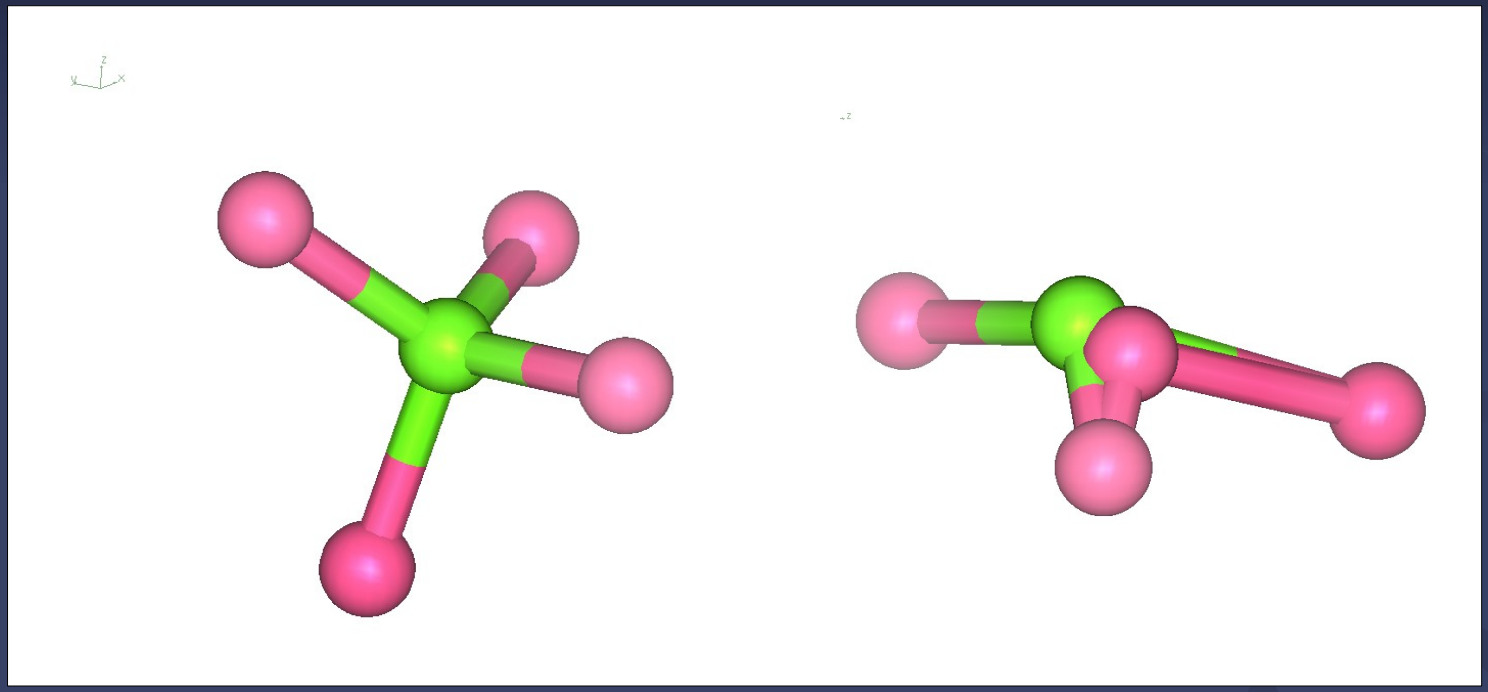
Conformation Idealization

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

Ligand validation

Why validate?

- 70K structures in the PDB
- 11K chemical compounds in 50K structures

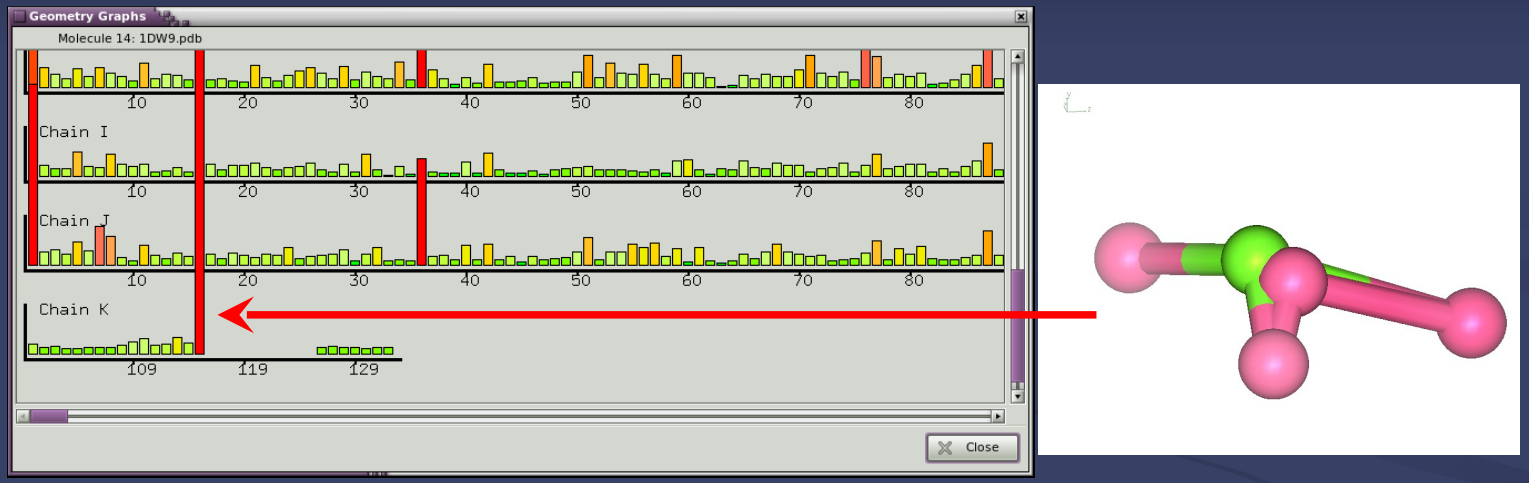


  - sulfate ions in 1DW9
  - 1.65Å resolution
  - R/Rfree: 0.15/0.19

Ligand validation

- How to validate ligand geometry?
  - Compare the observed structure to the restraints

    Coot: Validate/Geometry analysis



  - but what if the parametrisation is wrong?

    Perfect refinement with wrong restraints

    → distorted geometry

Ligand validation

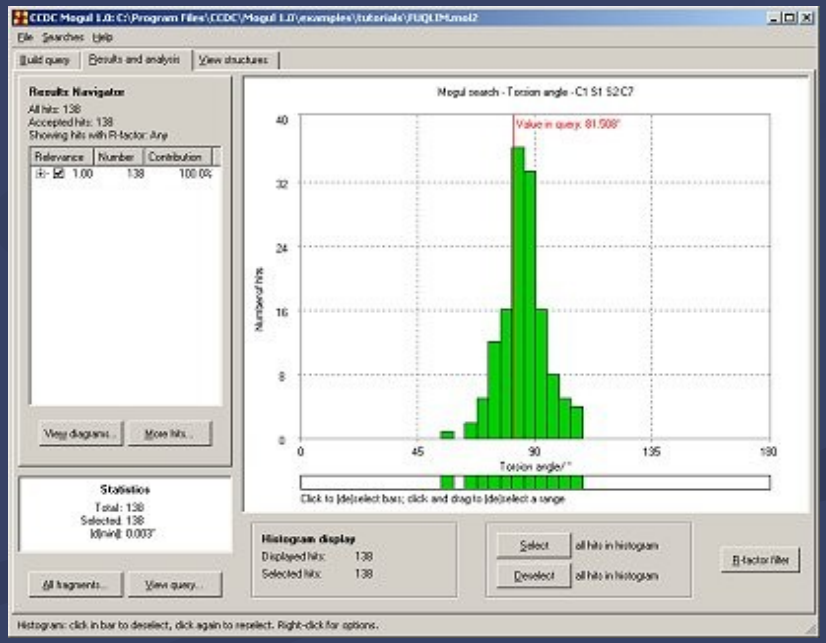
How to validate ligand geometry?

- QM (minimised structures, forces on atoms)
  - CPU hungry
  - in vacuo
  - low energy conformation does not necessarily correspond to the ligand-bound conformation
- PDB (e.g. ValLigURL)
  - good source of cofactor structures,
  - less useful for novel ligands
  - occasionally questionable quality

Ligand validation

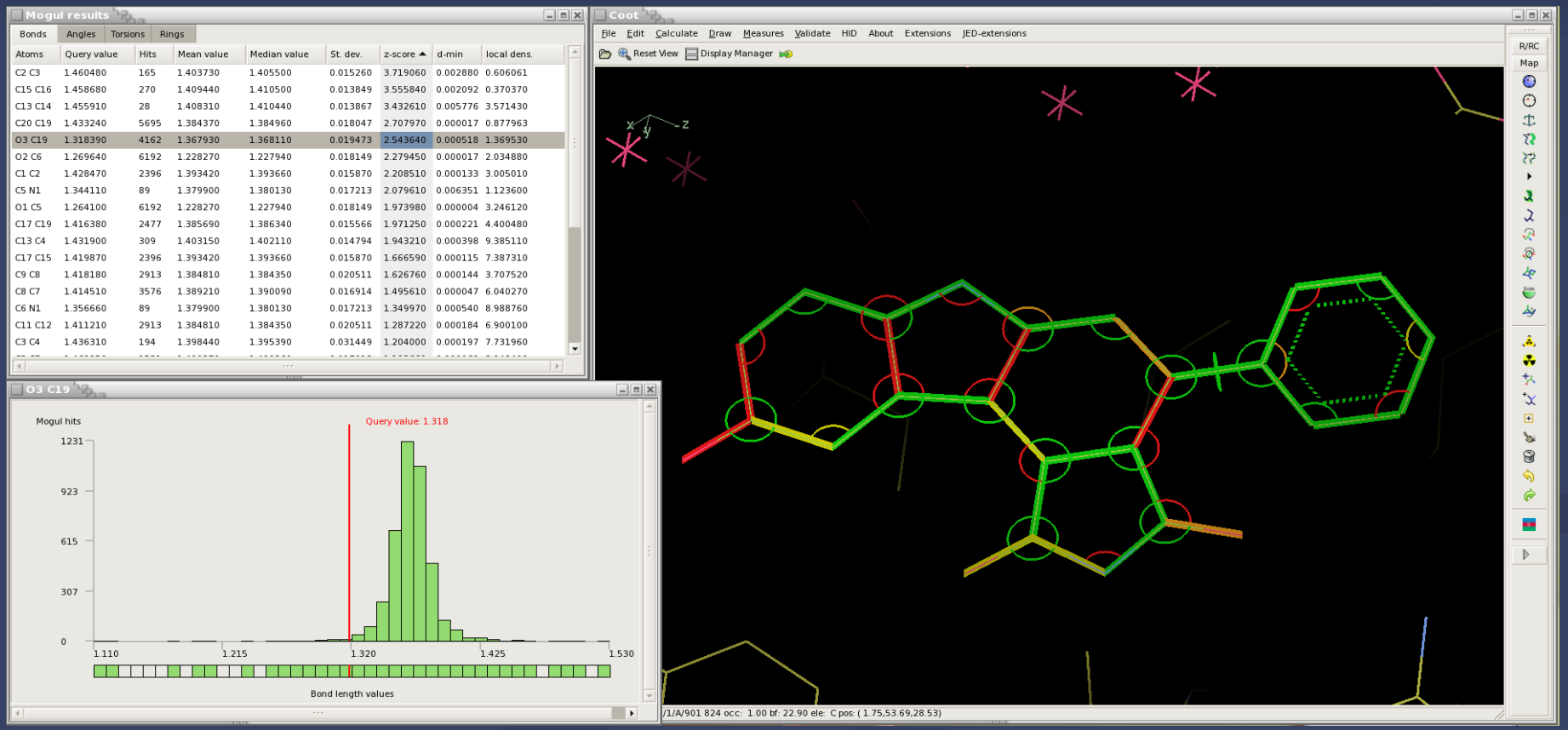
How to validate ligand geometry?

- CSD (are small molecule geometries relevant in protein structures?)
  Mogul:
  - knowledgebase of geometric parameters based on CSD
  - can be run as batch job on pdb files (automatic atom typing)
  - mean, Z-score, # of hits etc.



Ligand validation

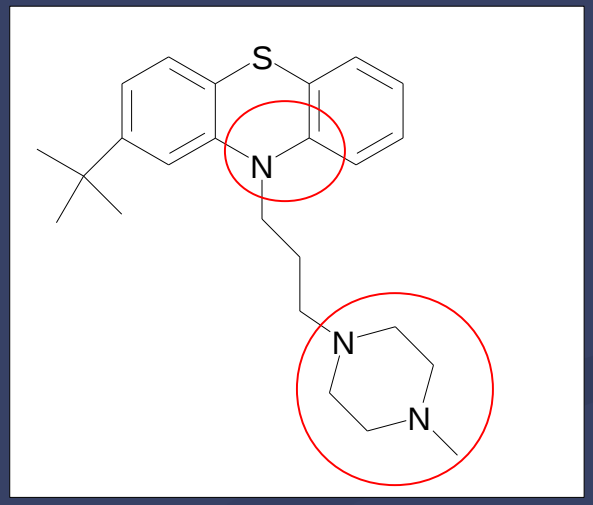
- Mogul plugin in Coot:
  - run Mogul from coot
  - update/correct restraints (target and esd for bonds, angles)

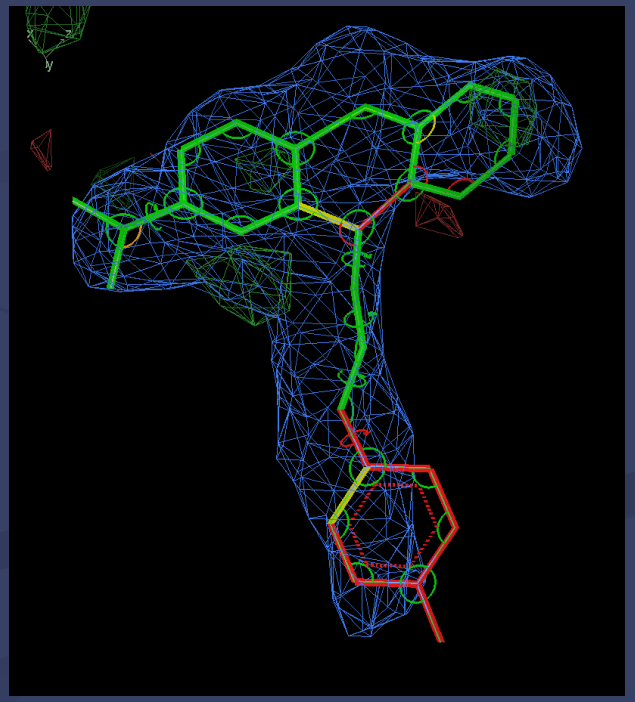


Ligand validation

- Sources of errors
  - Incorrect parametrisation
  - Incorrect fitting: wrong ligand or density not supporting conformation or binding mode

    PDB-code: 1CTR, TFP residue
    - alternate conformation not modelled
    - methyl-piperazin group planar





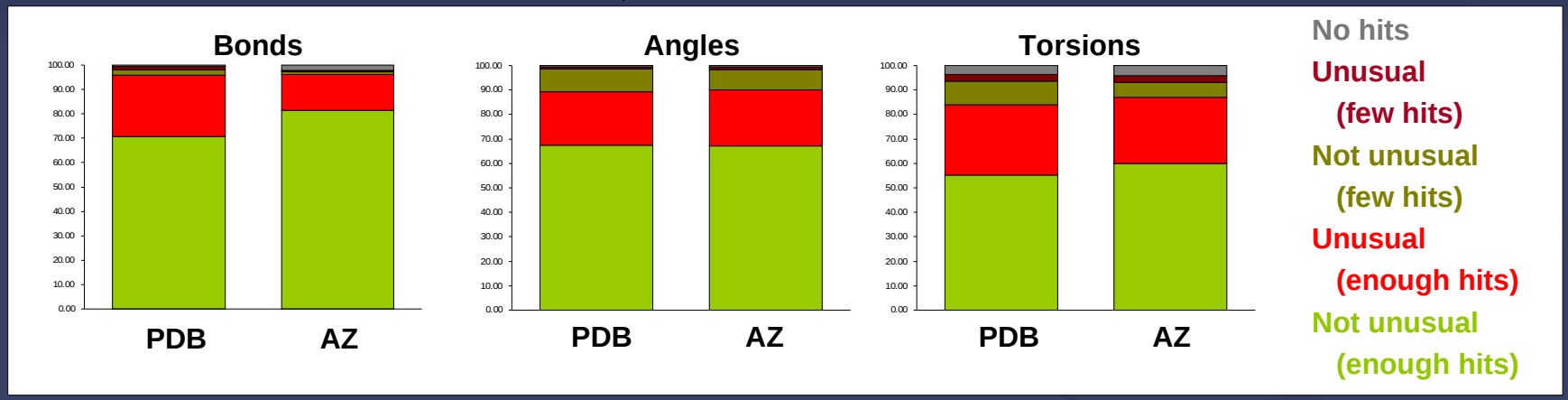
Ligand validation

- AZ vs PDB: similar error rates

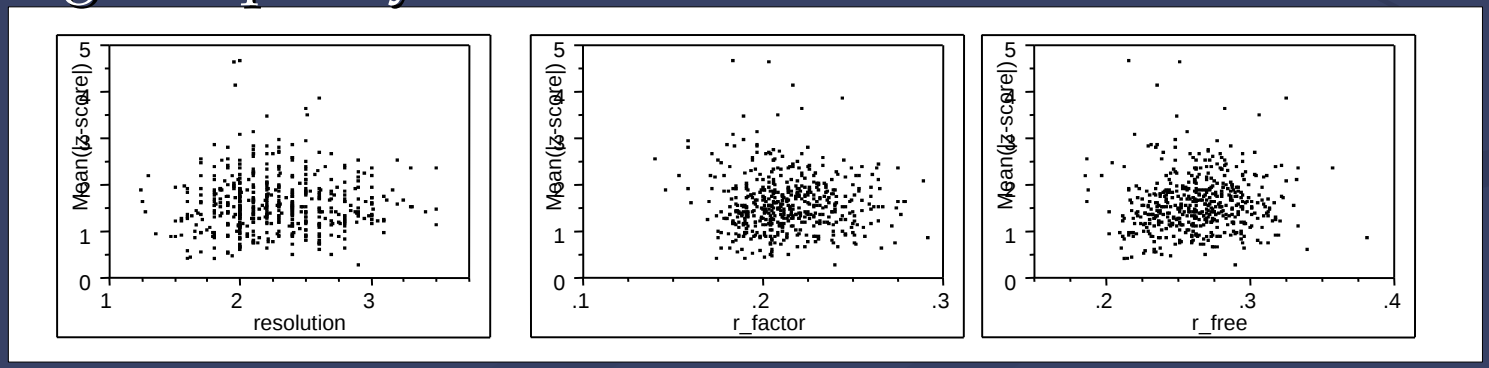
  bonds: better, corrected in-house libraries

  angles: not tightly restrained (+ errors in restraints)

  torsions: not restrained, different from small mol. values



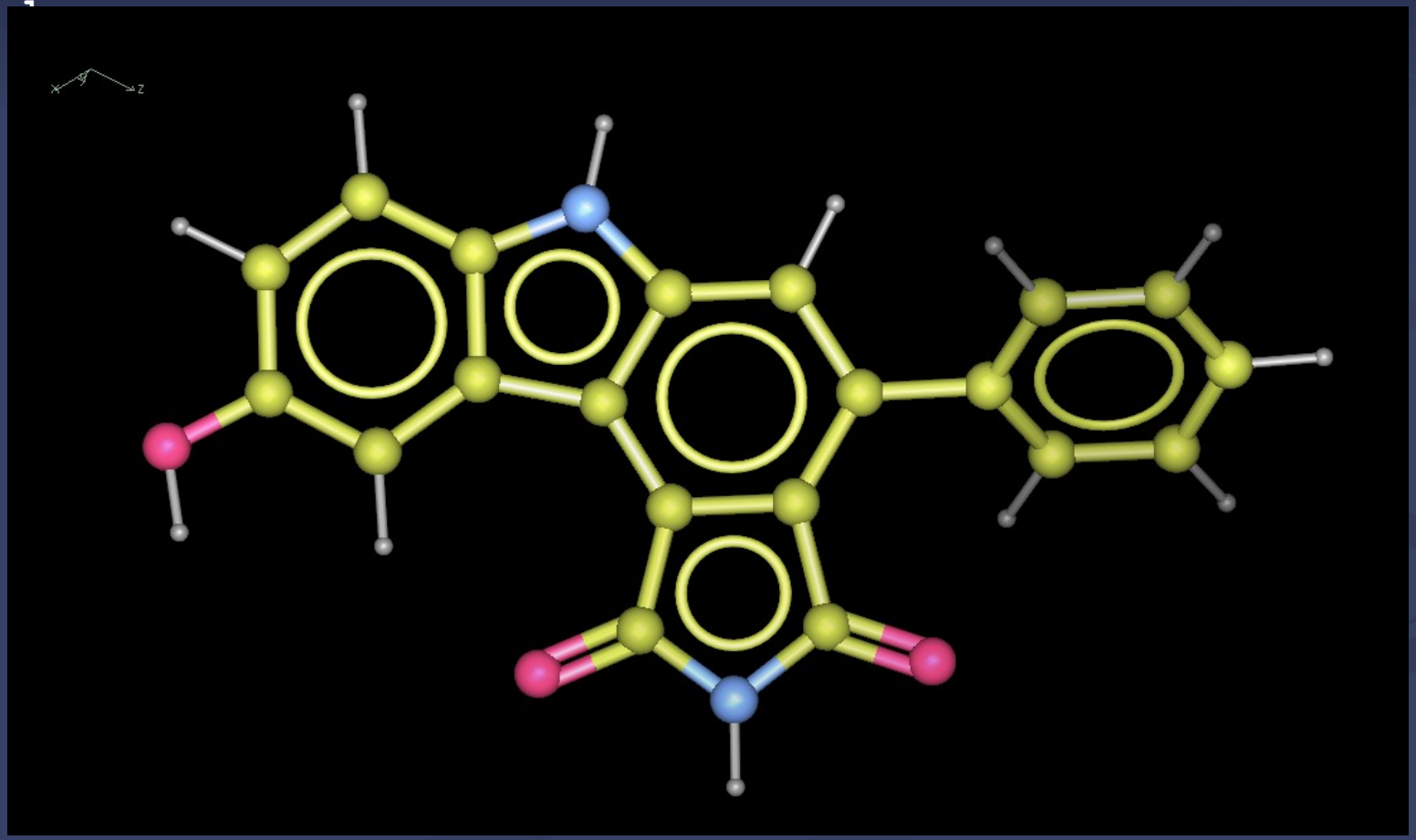
- Ligand quality not correlated with overall structure quality



Ligand representations

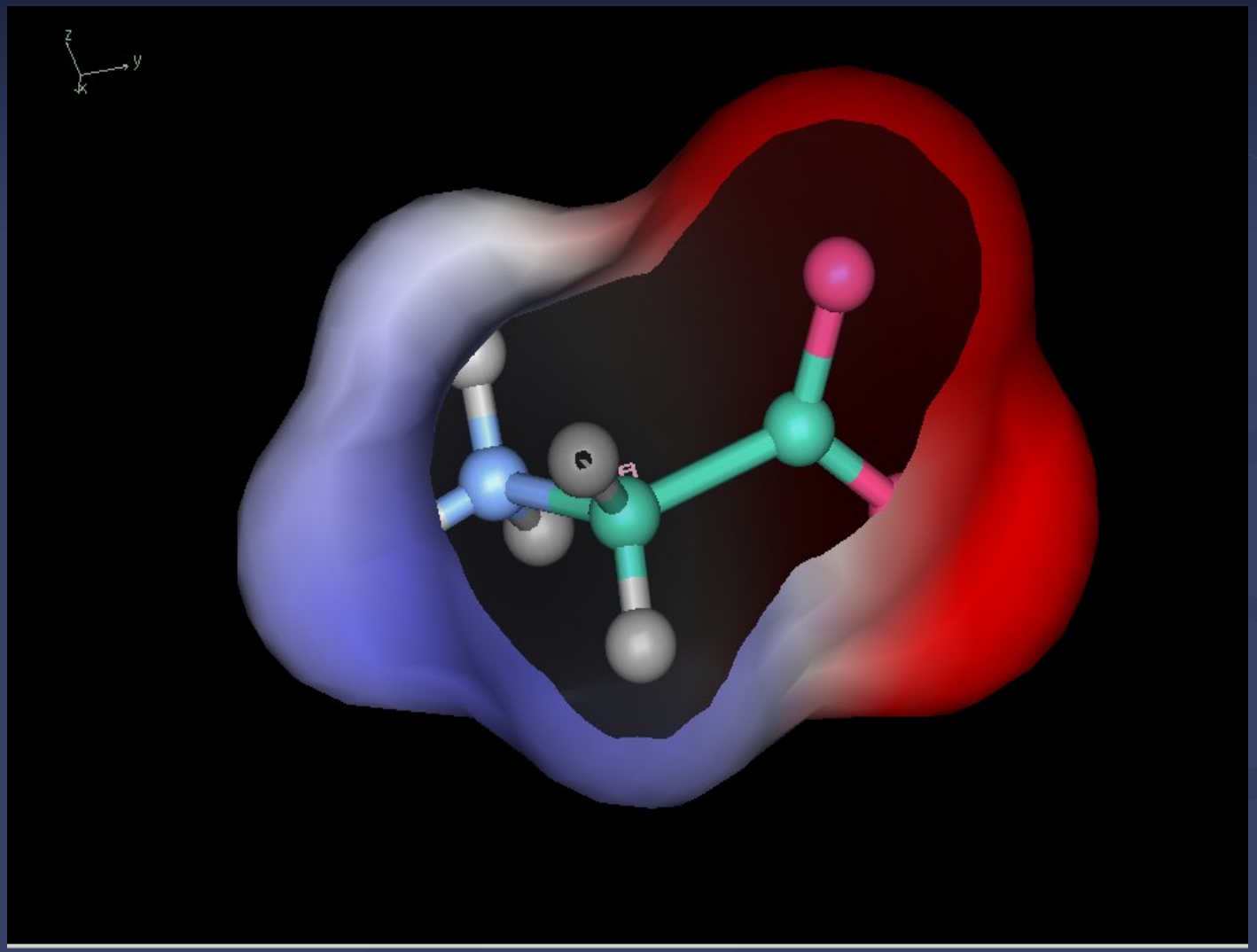
Ligand representations

- Bond order representation
  from dictionary definitions



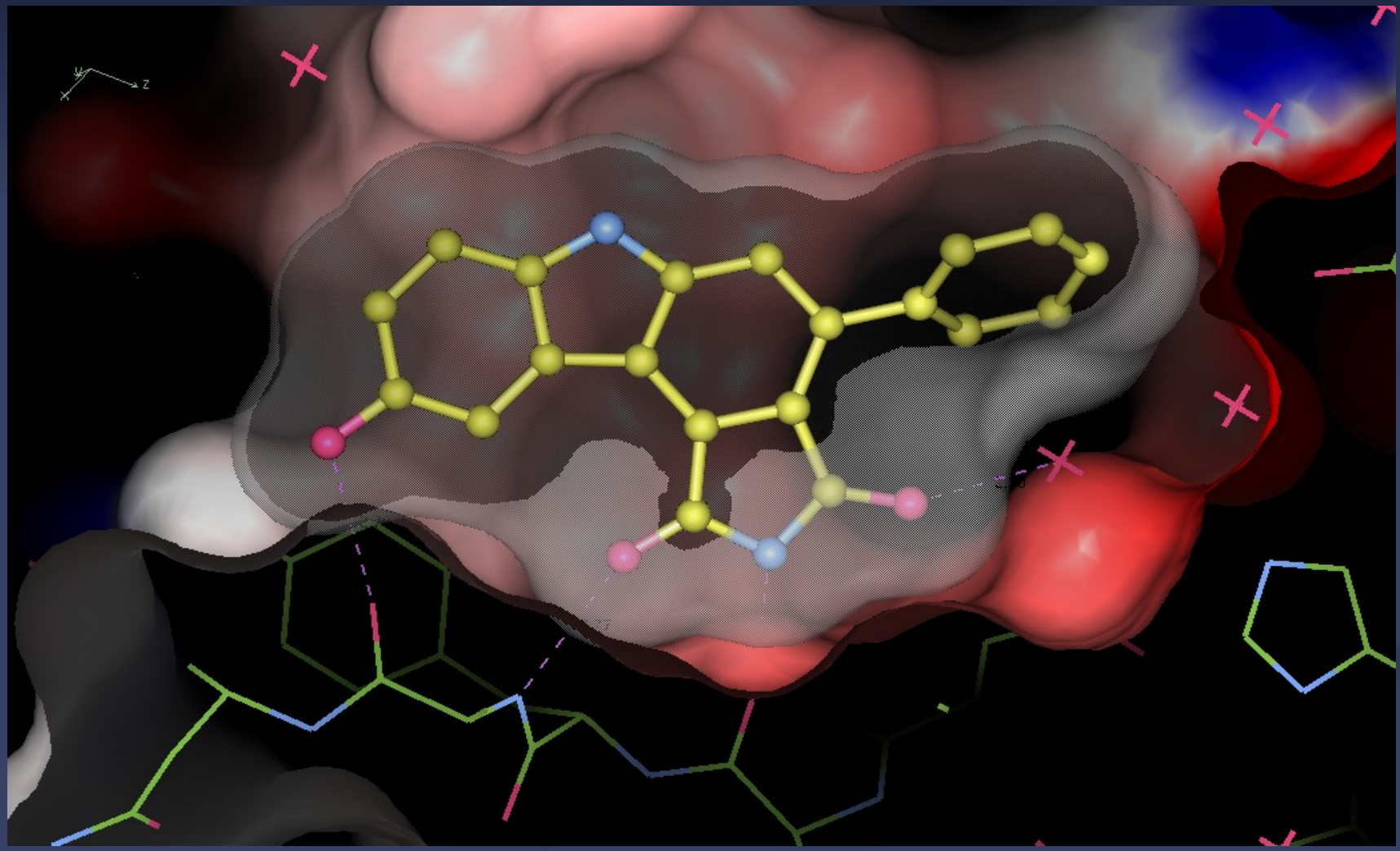
Surfaces

- Partial charges



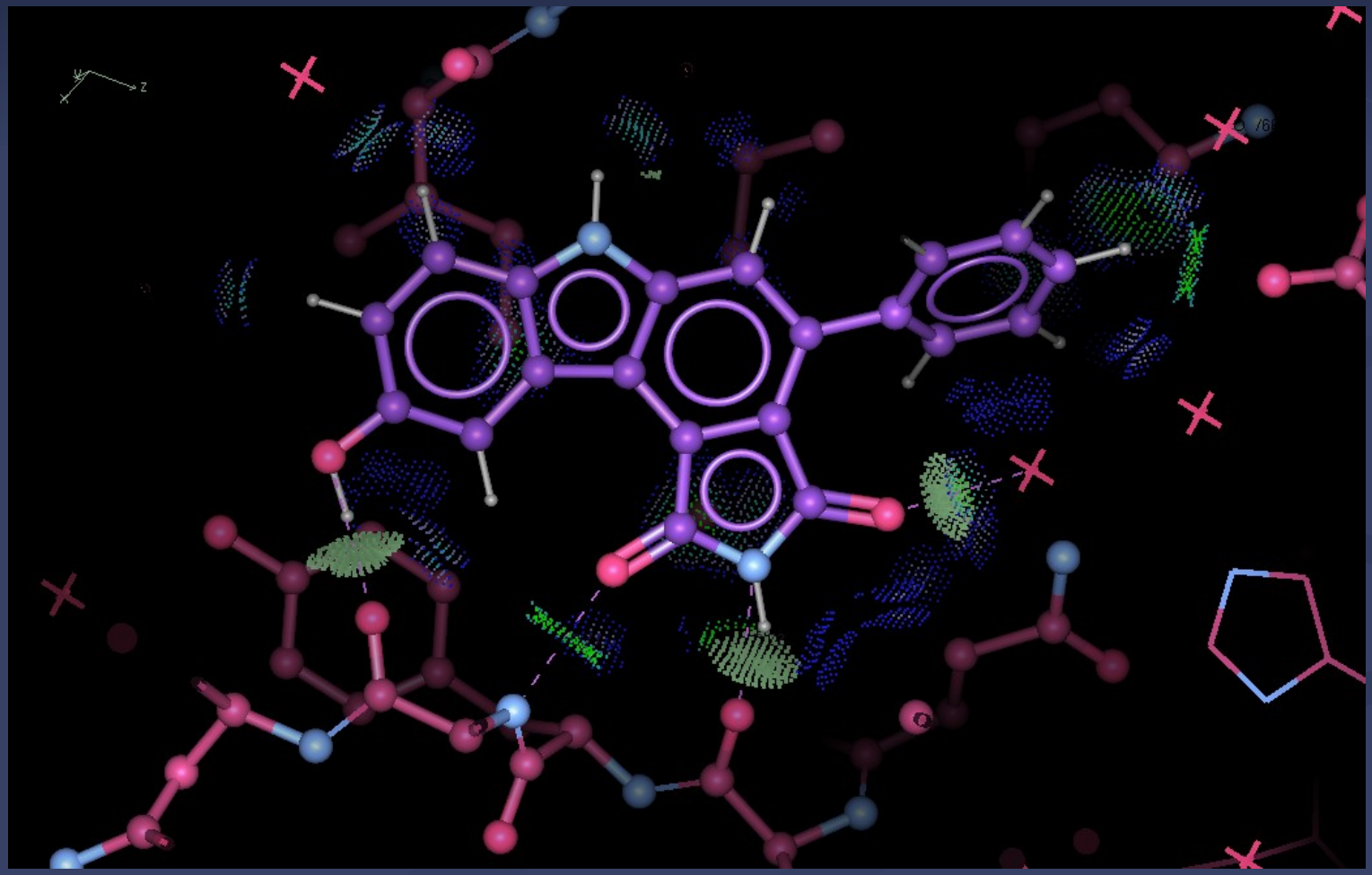
Surfaces

- Transparent surfaces - surface complementarity



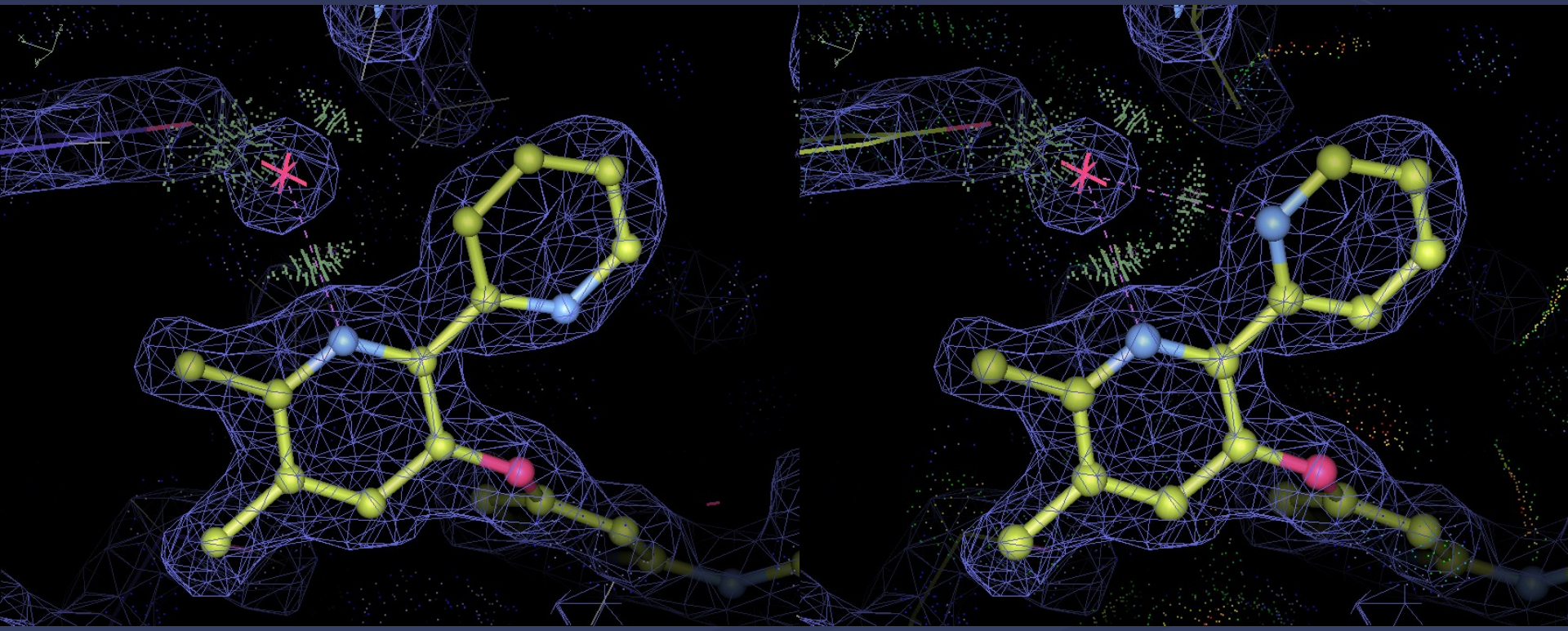
Binding mode analysis

- Binding site highlighting
- Isolated molprobity dots



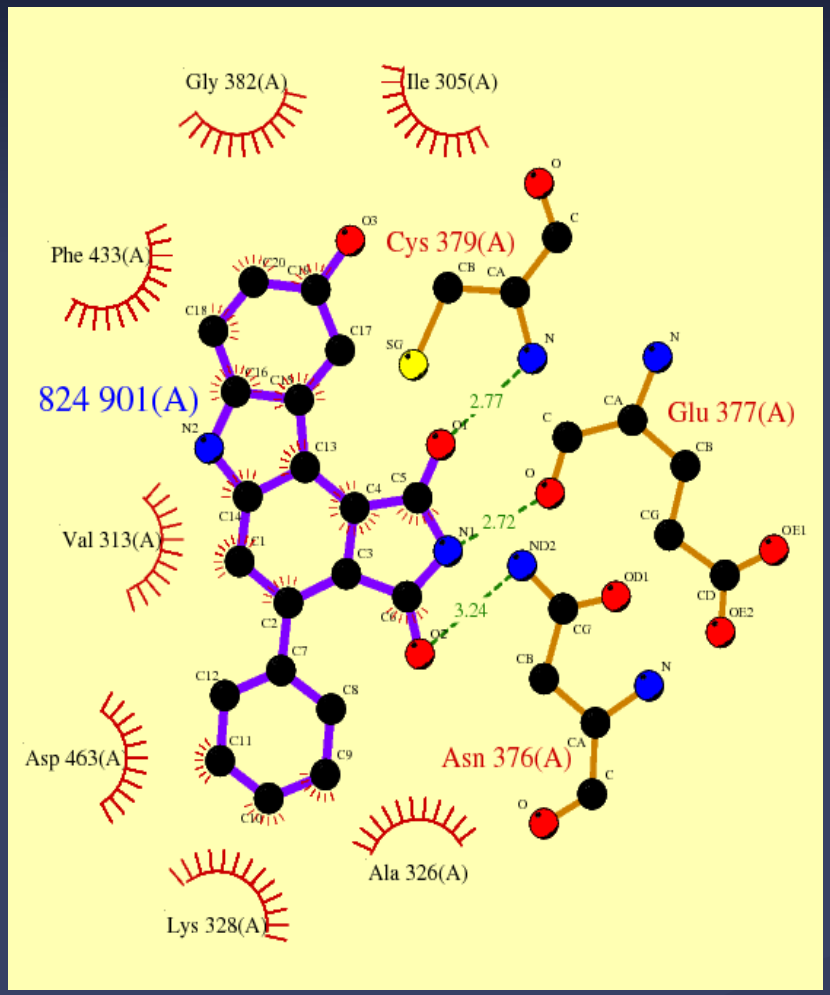
Binding mode analysis

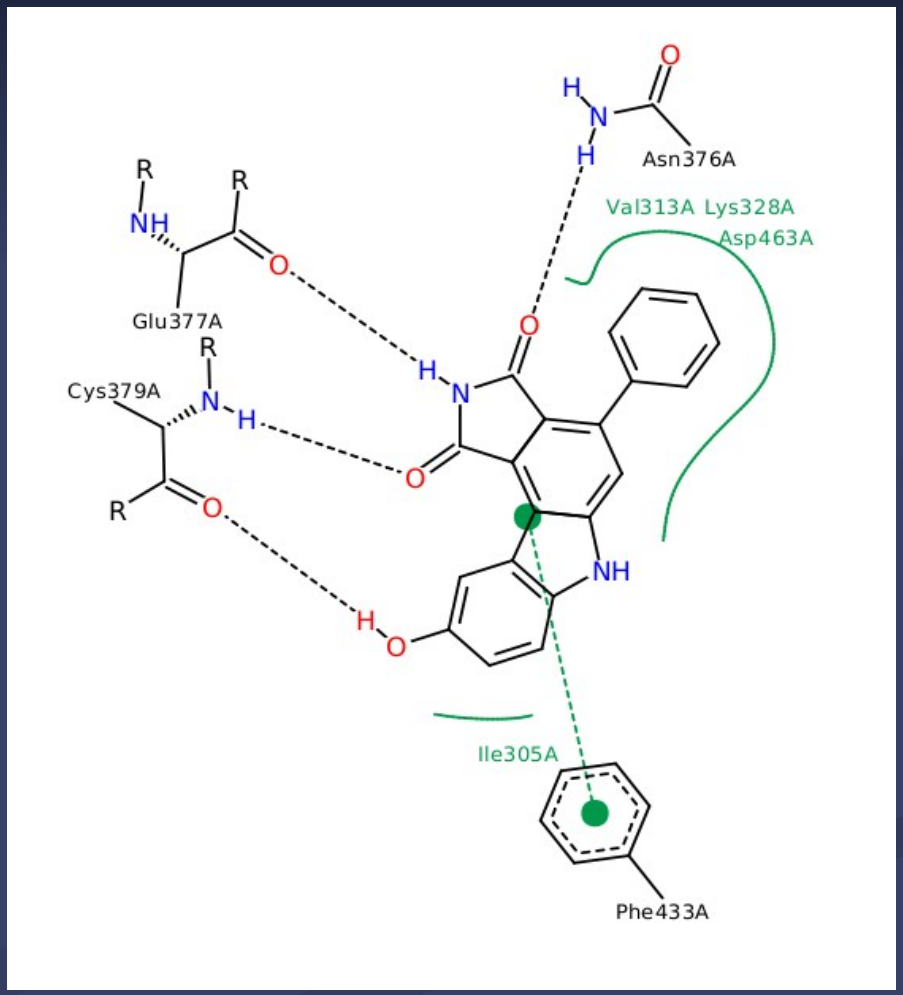
- Binding site highlighting
- Isolated molprobity dots



Ligand environment layout

- 2D ligand pocket layout (ligplot, poseview)





Can we do better?

Ligand environment layout

- Binding pocket residues
- Interactions
- Substitution contour
- Solvent accessibility halos

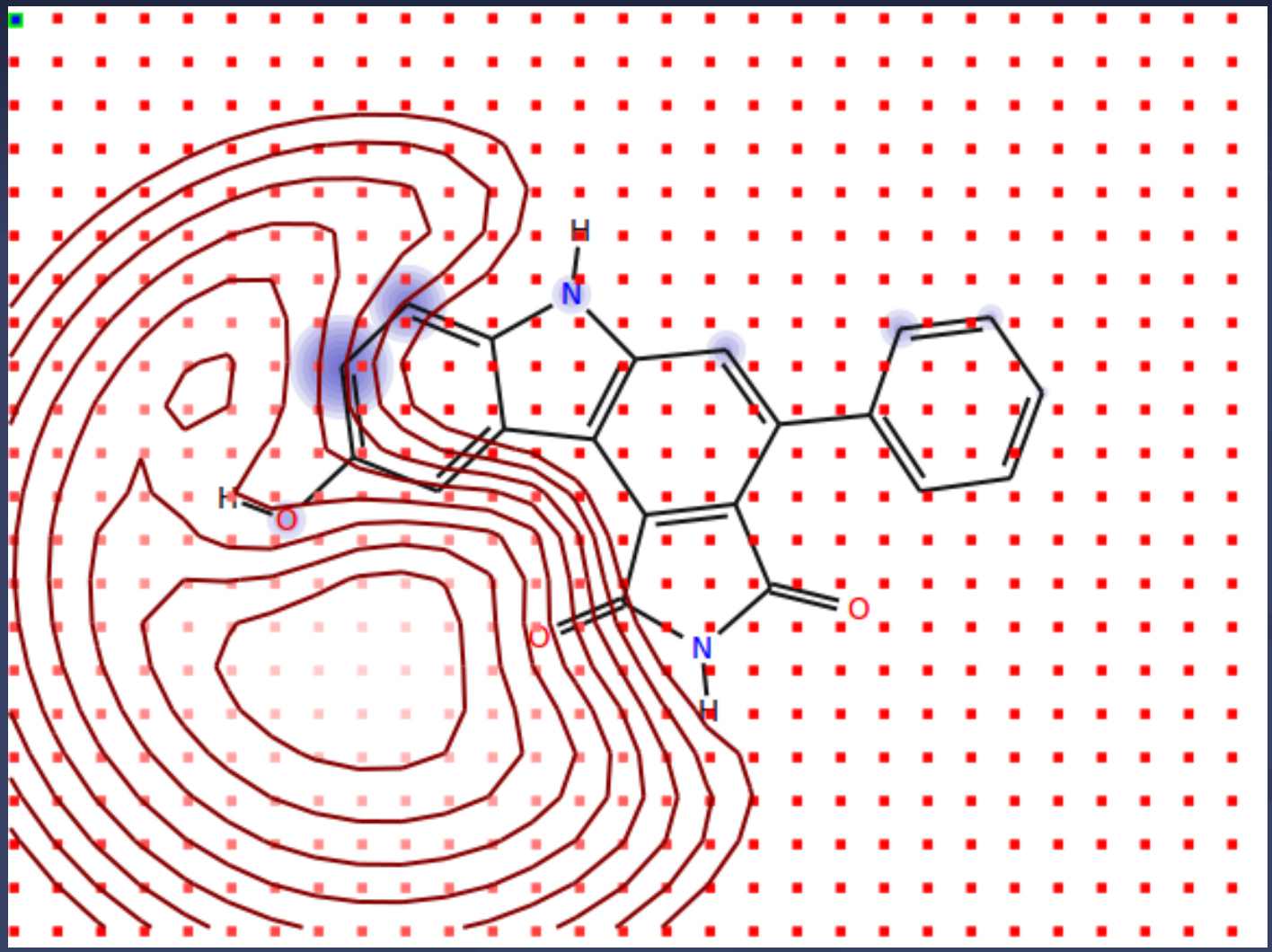
Ligand environment layout

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

Work in progress

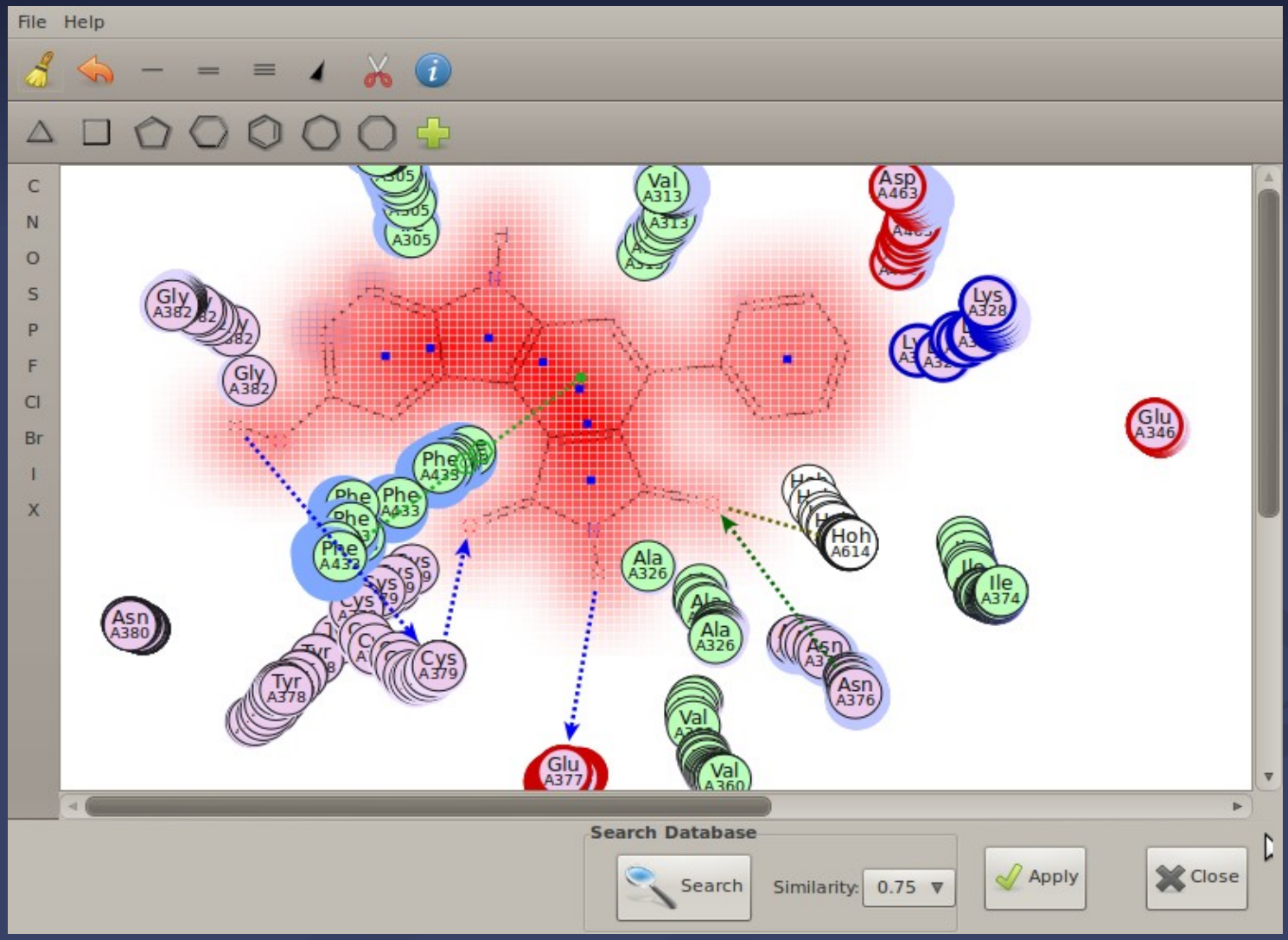
Ligand environment layout

- Initial residue placement



Ligand environment layout

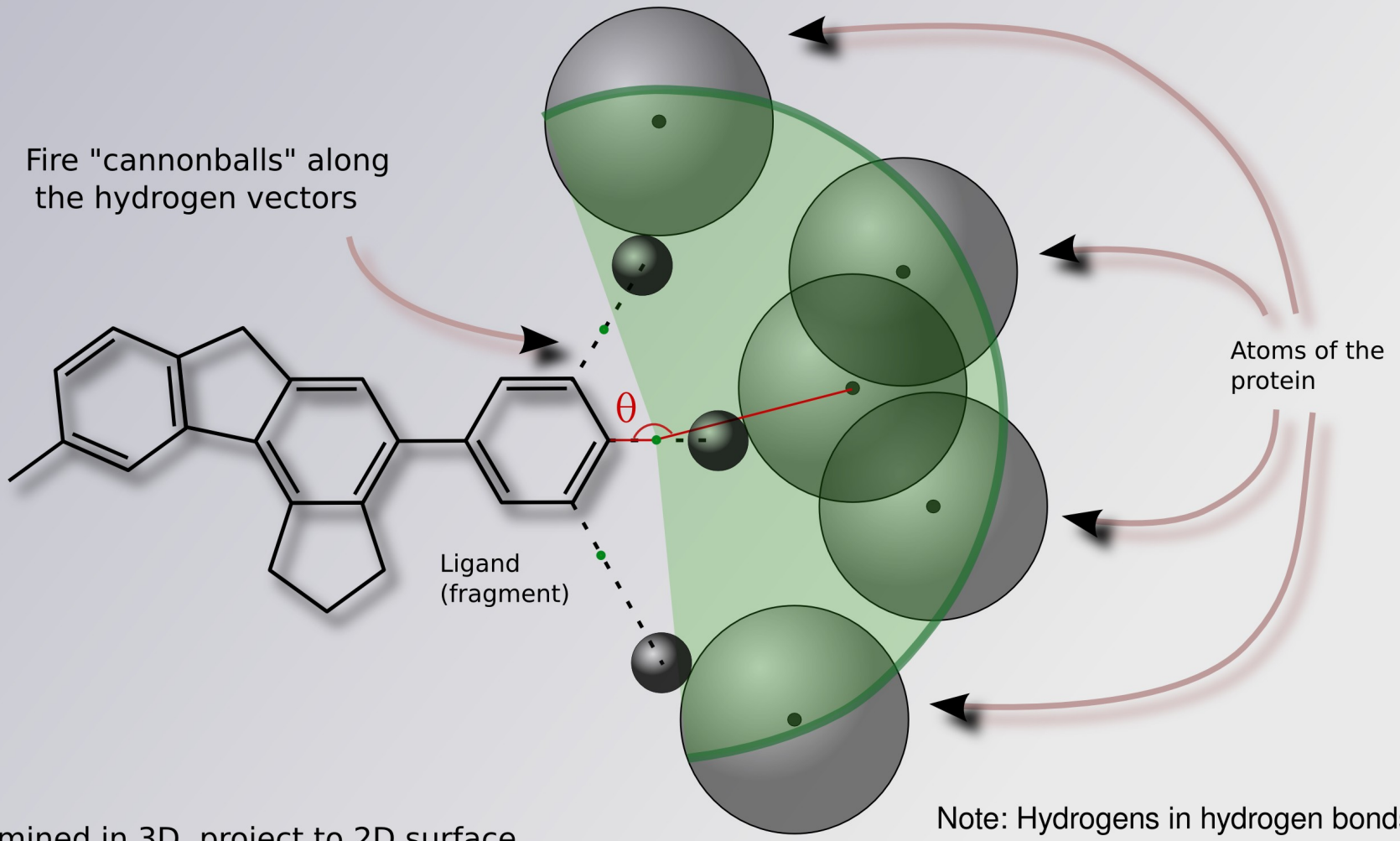
- Residue position minimisation



Determination of the Substition Contour

c.f. Clarke & Labute (2007)

How far can we go (in the direction of the hydrogens) before hitting atoms of the protein?

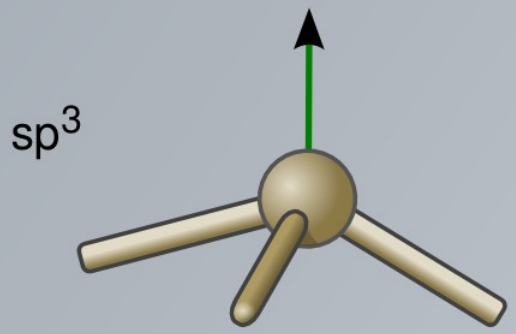


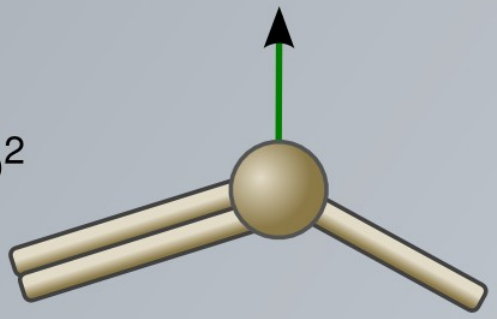
Determined in 3D, project to 2D surface

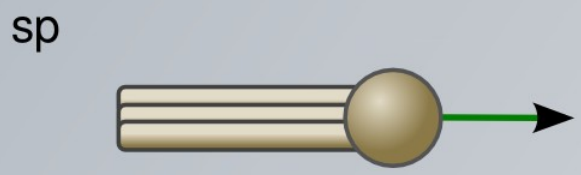
Note: Hydrogens in hydrogen bonds are a confounding factor

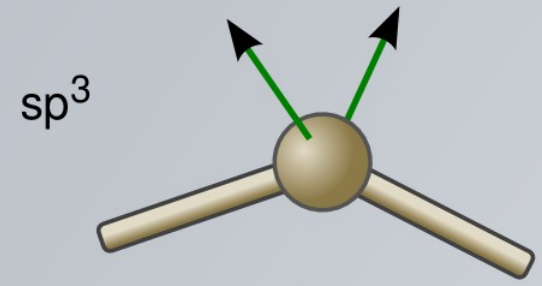
Substitution Contour: Extending Along Hydrogens

Riding Hydrogens



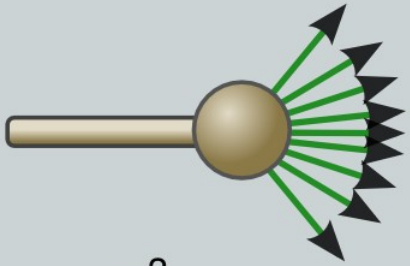


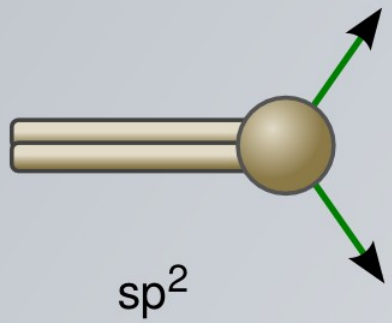




Torsionable Hydrogens

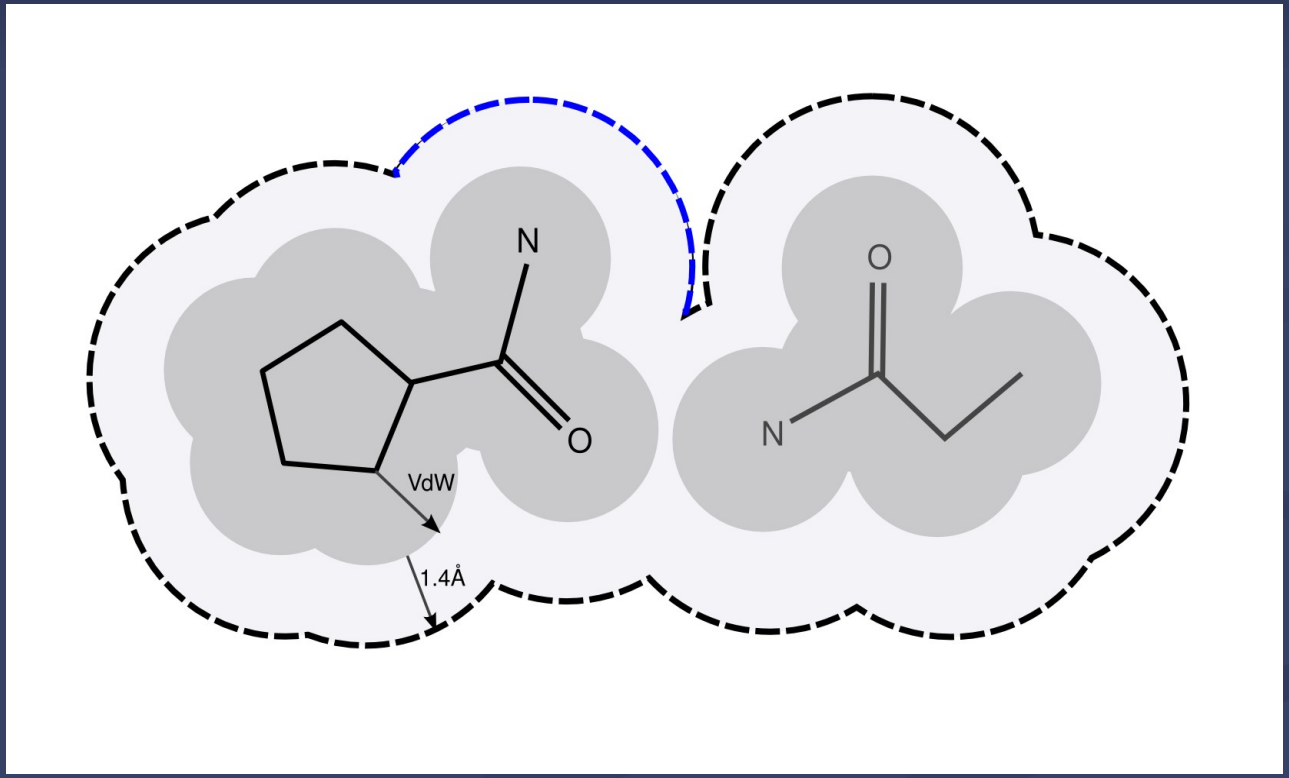
(test multiple directions)

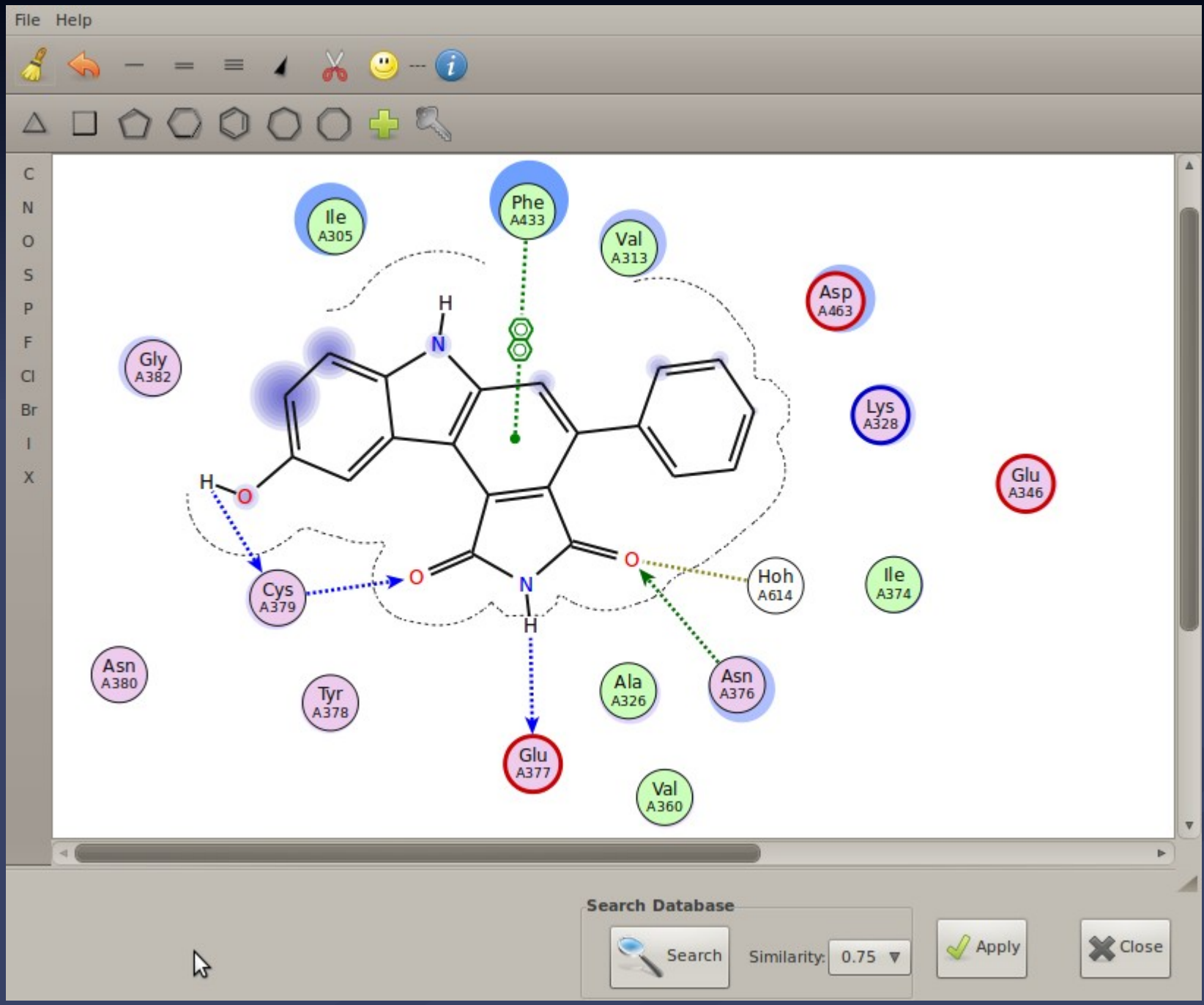




Solvent exposure calculation

- Identification of solvent accessible atoms
- Different from substitution contour




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

Acknowledgements

- Paul Emsley
- Bernhard Lohkamp
- Kevin Cowtan

http://www.biop.ox.ac.uk/coot/

or

Google: Coot

- Libraries, dictionaries
  - Alexei Vagin, Eugene Krissinel, Stuart McNicholas
  - Richardsons (Duke)
  - Martin Noble (electrostatic surfaces)