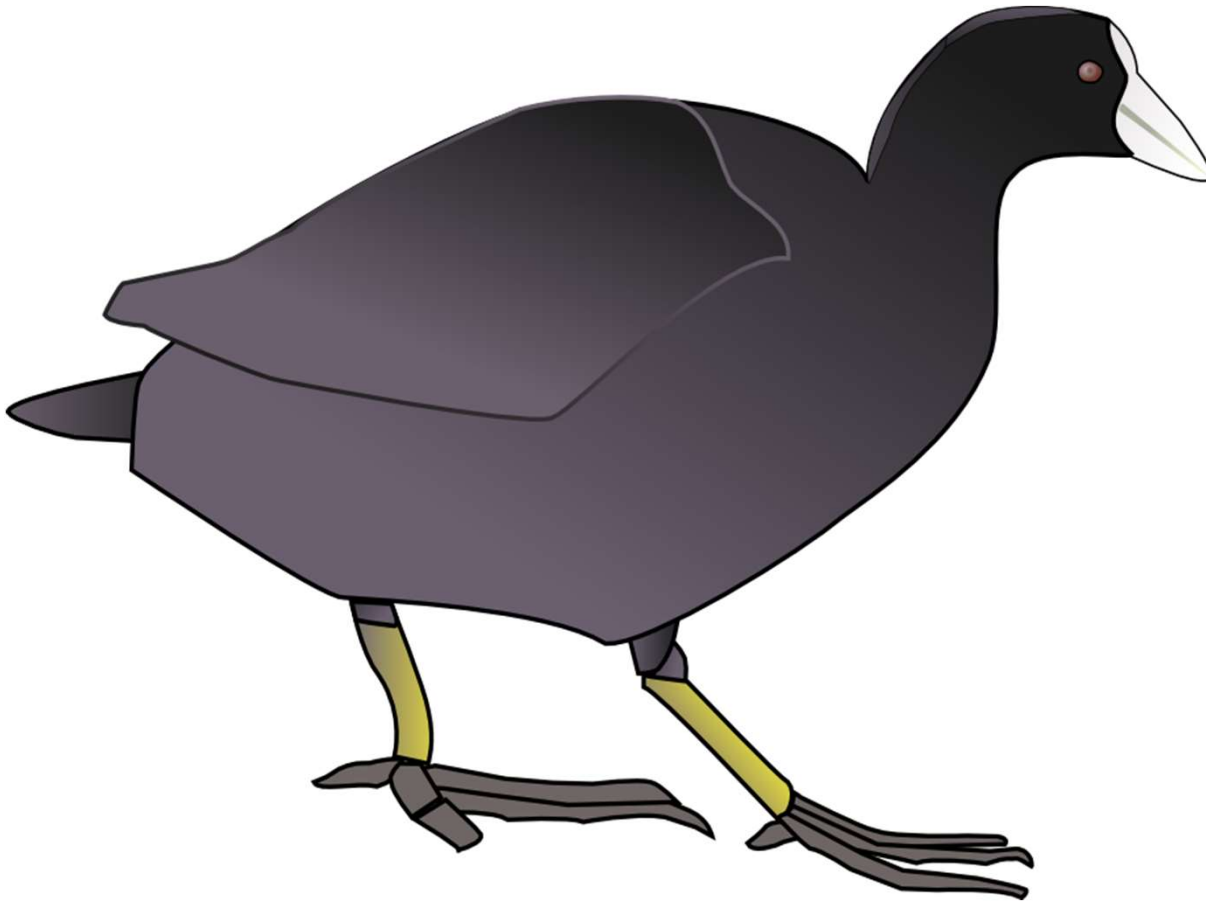


Coot



Judit Debreczeni (AZ, UK)
2021 December, DLS

Acknowledgements

Credit where it's due:

- **Developers:**
 - Paul Emsley
 - Bernhard Lohkamp
 - Kevin Cowtan
- **Libraries:**
 - Alexei Vagin
 - Garib Murshudov
 - Eugene Krissinel
 - Greg Landrum

Useful links

Downloads:

- **Coot:** <https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/binaries/>
- **wincoot:** <https://bernhardcl.github.io/coot/>
- **CCP4:** <http://www.ccp4.ac.uk/download>
- **CCPEM:** <https://www.ccpem.ac.uk/download.php>

Manual: <https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/web/docs/>

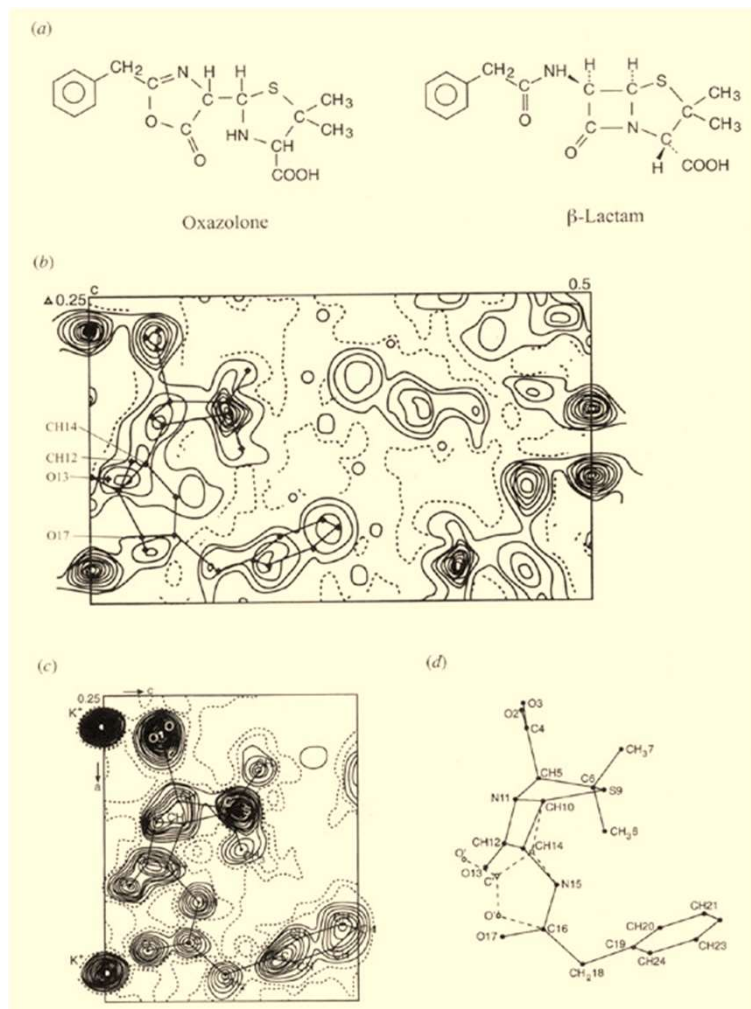
Blog: <https://pemsley.github.io/coot/>

Moorhen

Coot



Once upon a time...



Is there a 4-membered ring?

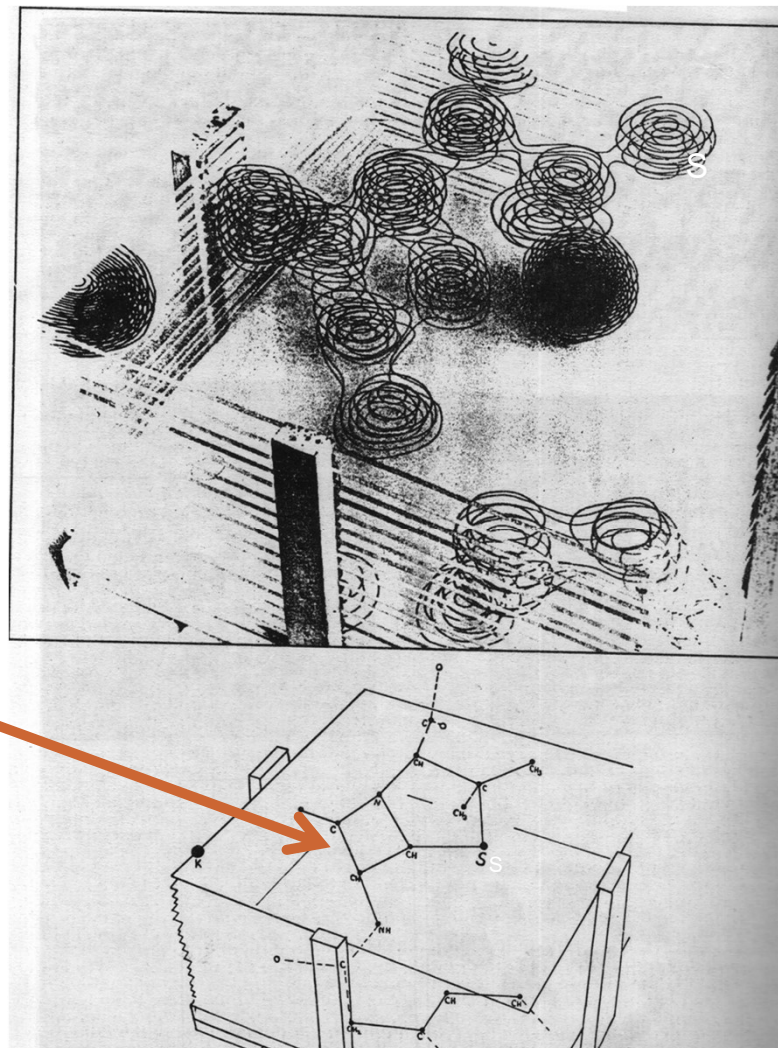
- First projection of Fourier map
- Final map
(yes, there is a 4-membered ring)

Once upon a time...

The 3-dimensional penicillin G map calculated in 1944. This unequivocally determined its chemical structure.

(The hand however is wrong.)

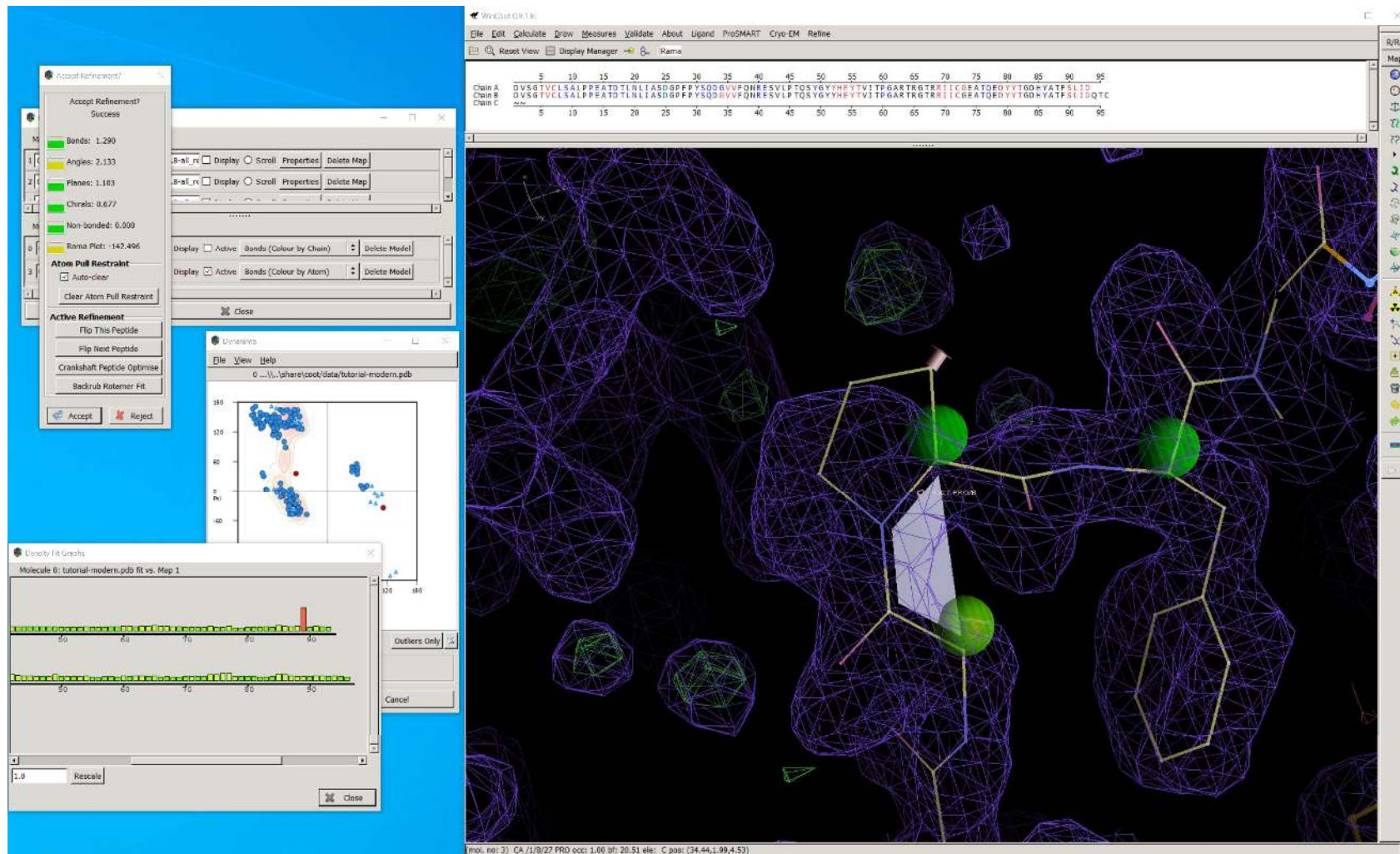
The chemical interpretation of the electron density map. The four membered beta-lactam ring, the centre of the controversy.

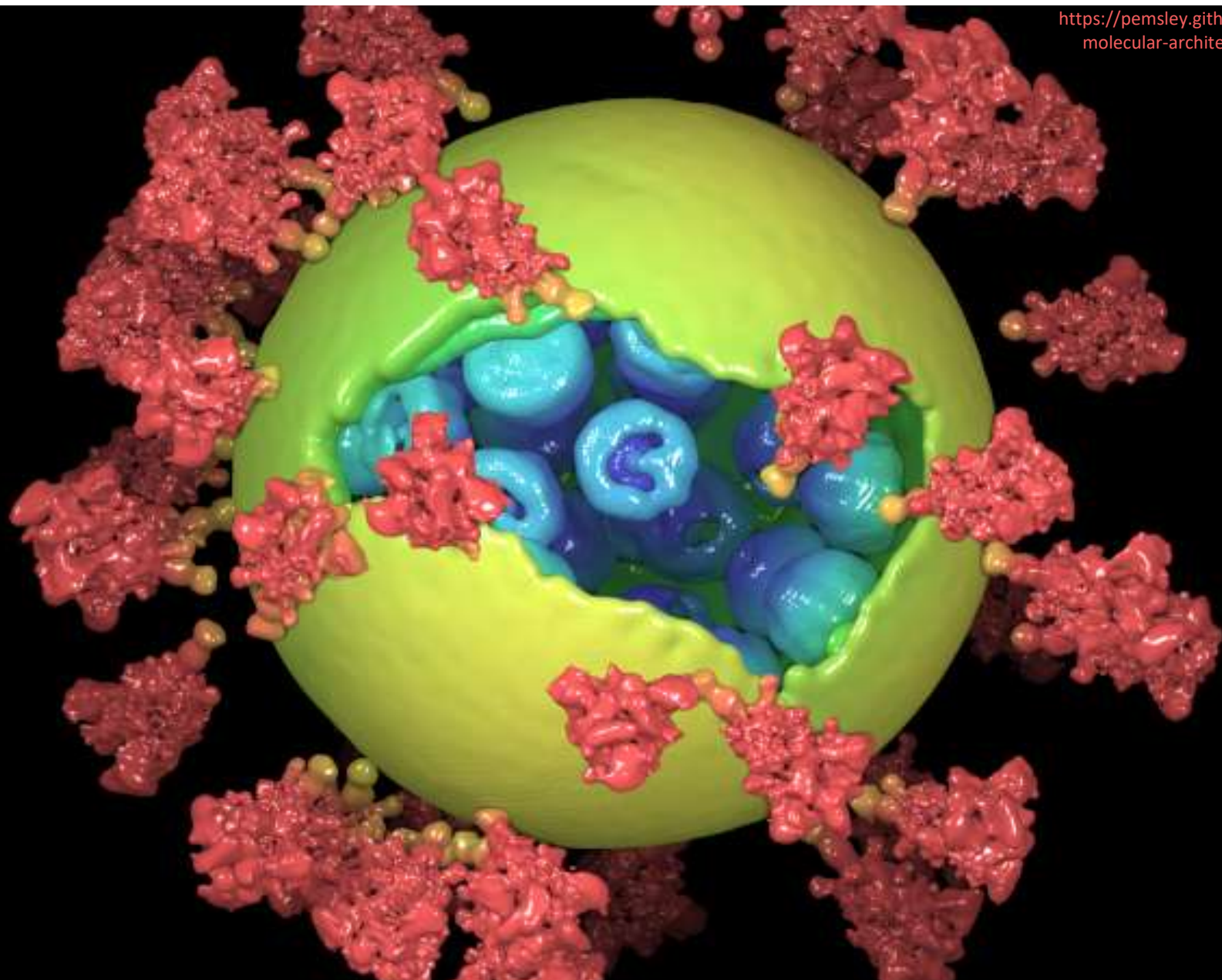


First model of insulin...



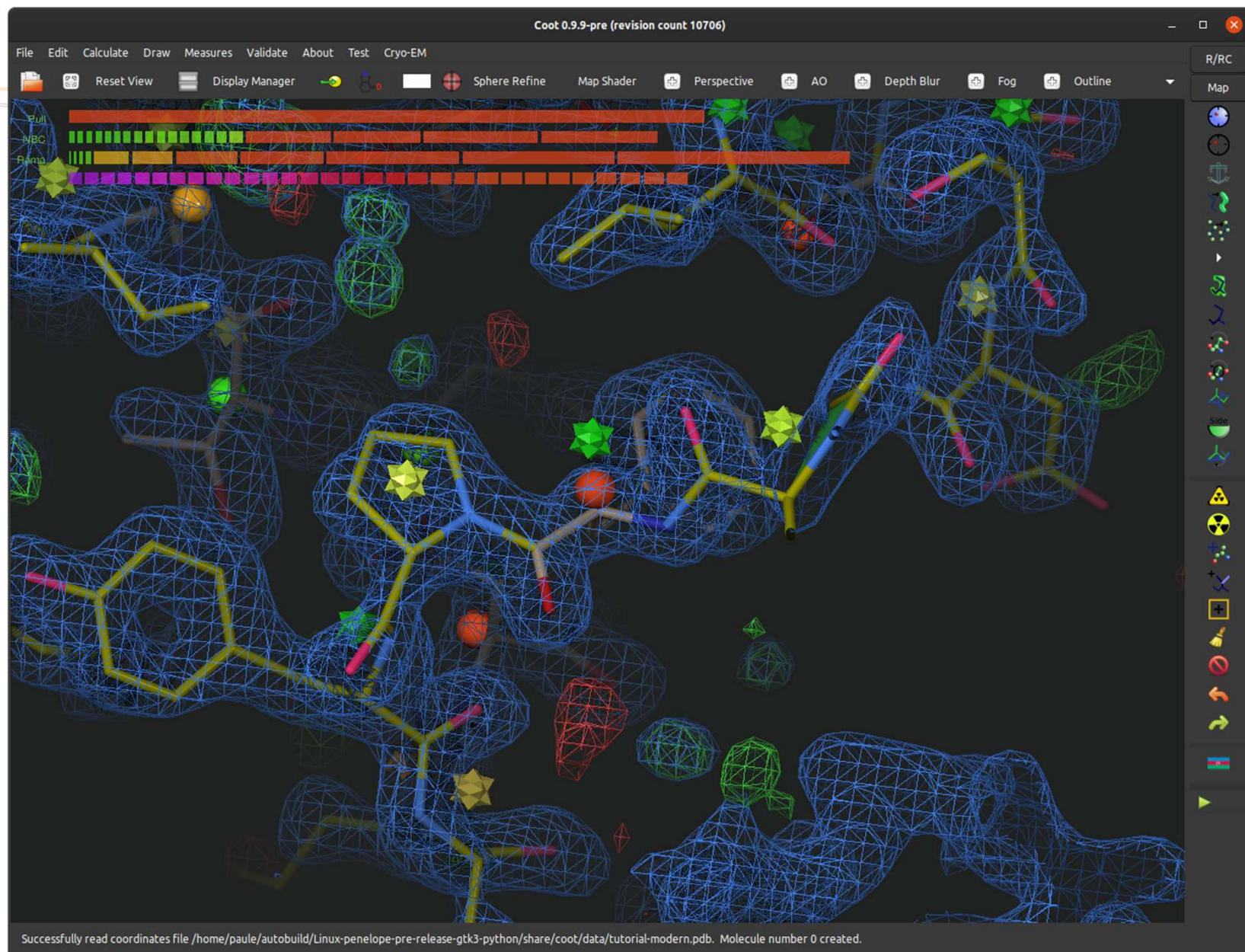
Right now...



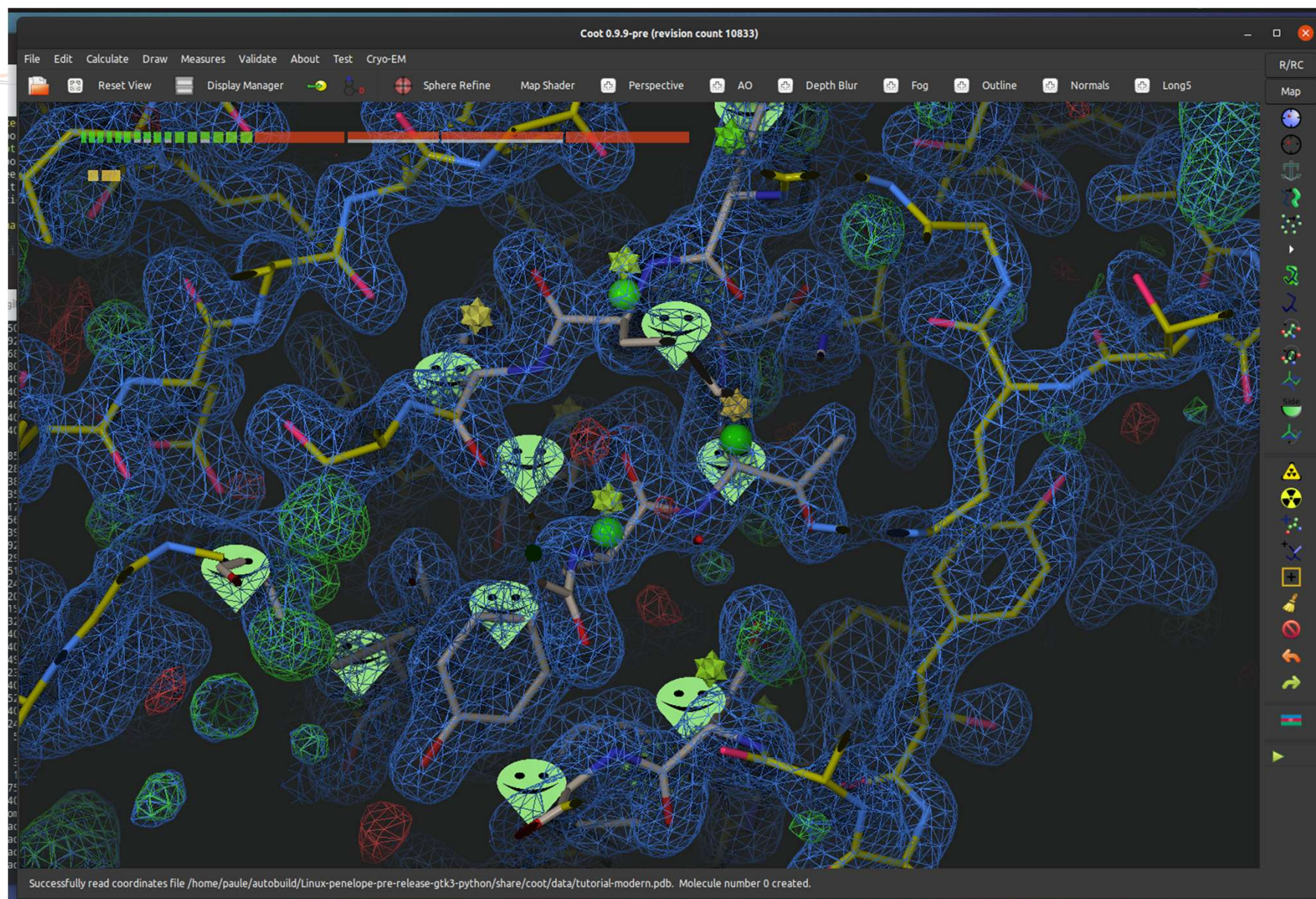


The future...?

- instead of
 - multiple windows
 - flat view and little depth
 - obsolete components (GTK1/2, Python2.x, OpenGL v1)
- the future is:
 - single full screen view
 - HUD representation of metrics
 - more depth using shading, occlusion, shadows, blur, fog depth
 - modern technology (Python3, Gtk+3, Gtk4, Vulkan)







Outline

- Look and feel
- Speed: preferences and key bindings
- Extra resources: Curlew
- Real space refinement – 0.8.x vs 0.9.x
- Low resolution tools
- Validation
- Handling NCS

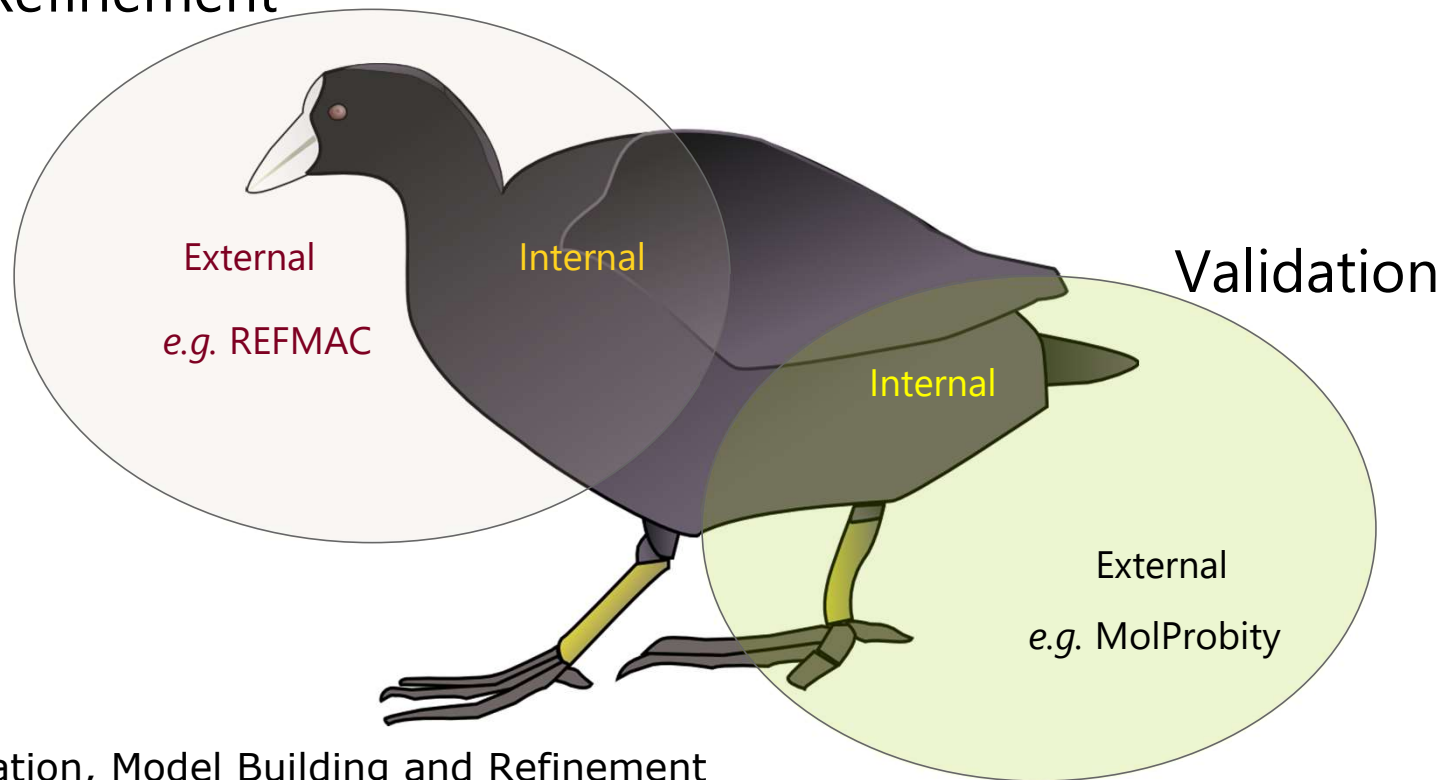
Coot

- Crystallographic Object-Oriented Tool-kit
- Primarily a tool for the interpretation of electron density generated from X-ray data (now CryoEM as well)
- with tools for modelling:
 - rotate/translate, rotamers,
 - refinement & regularization
 - add, delete
 - ligand fitting
- Interface to other programs: SHELXL, Refmac, Libcheck, Probe&Reduce (Molprobit), EBI, EDS, Povray, Raster3D, PHENIX, ...

A “workhorse”, not a show-pony

Feature Integration

Refinement



Validation, Model Building and Refinement
should be used together

Coot

- Interactive model-building, refinement and validation
- (Fast) automation over interaction
- Functionality over speed
- Speed over beauty
- Beauty over ugly (as long as it doesn't cost too much time)

A workhorse not a show pony

– Optimised for the world of chicken-wire

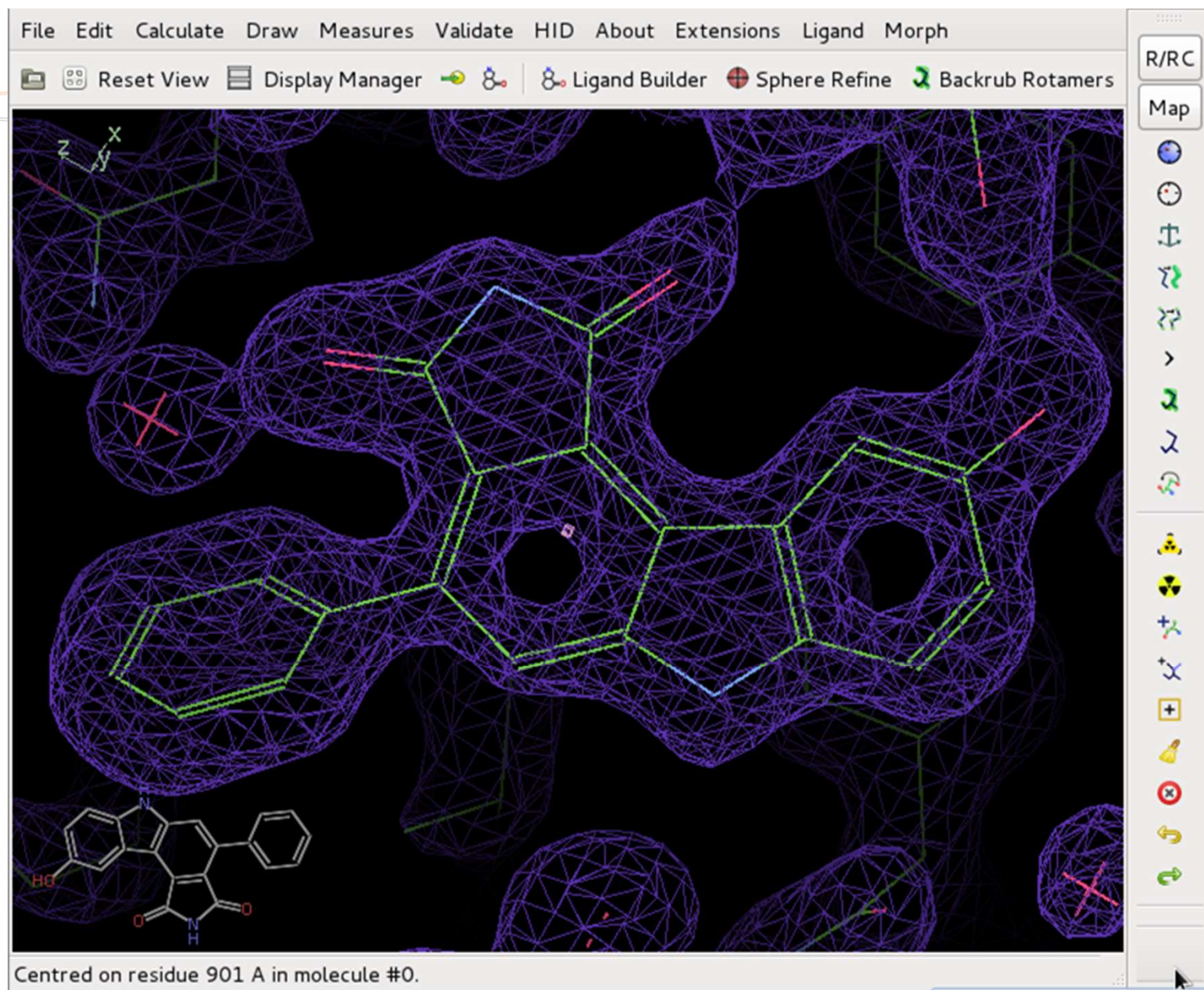
Presentation quality screenshots with Coot

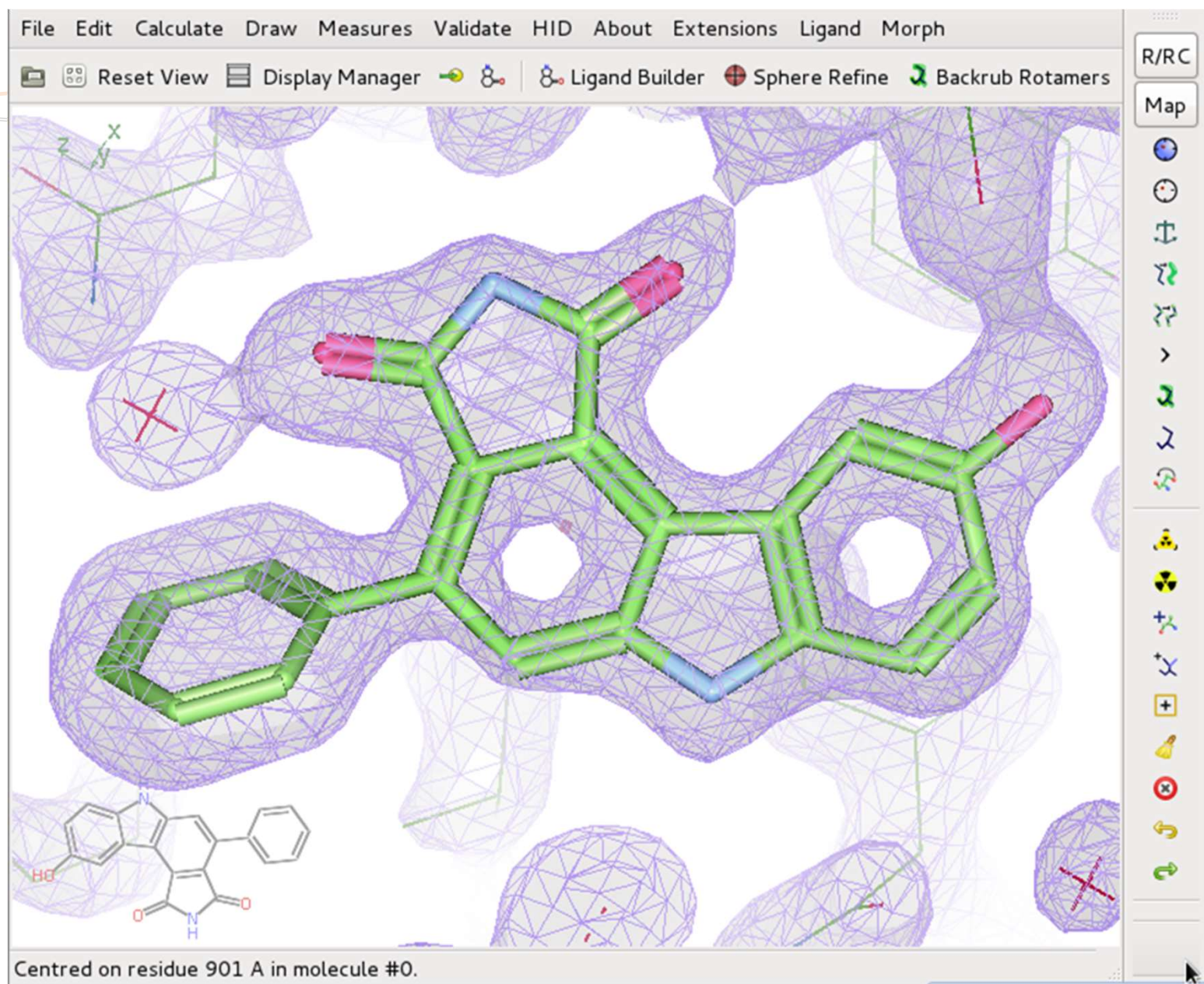
Standard Screenshot

- Density is too dark for high illumination seminar rooms

Better option:

- Edit → Background Colour → White
- Draw → Additional Representations → Ball & Stick → Add Representation
- Resample the map (1.8 → 2.2)
- Map Colour → “Light Blue/Grey”

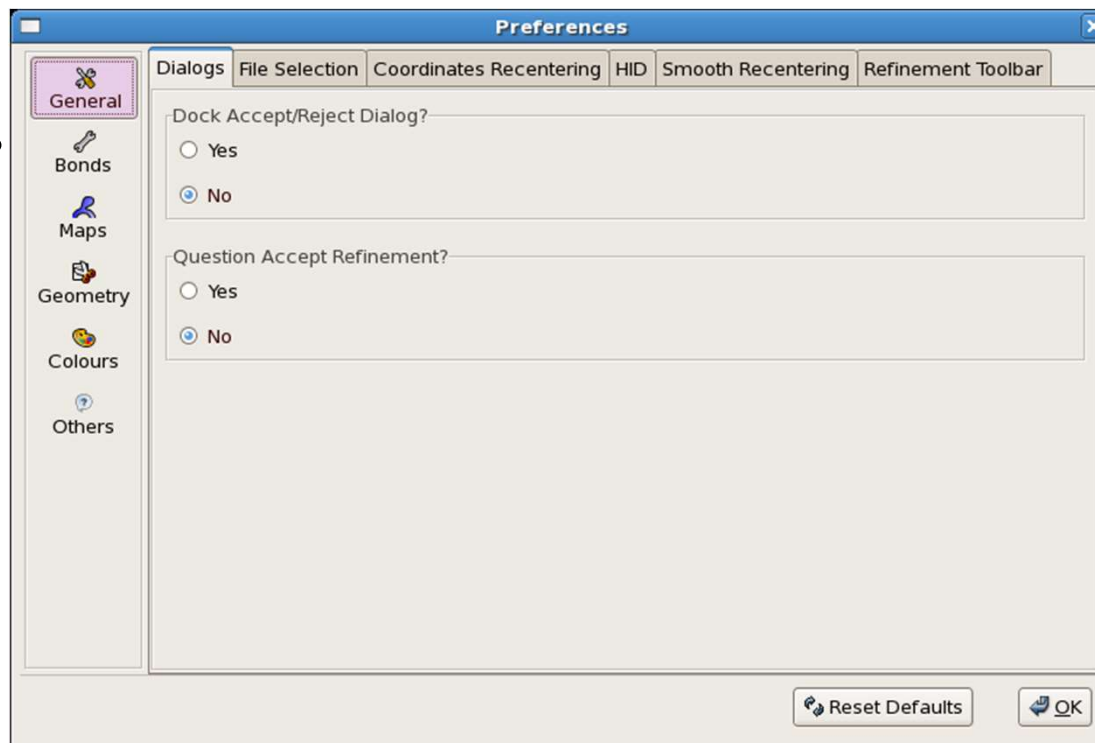




Speed: user preferences

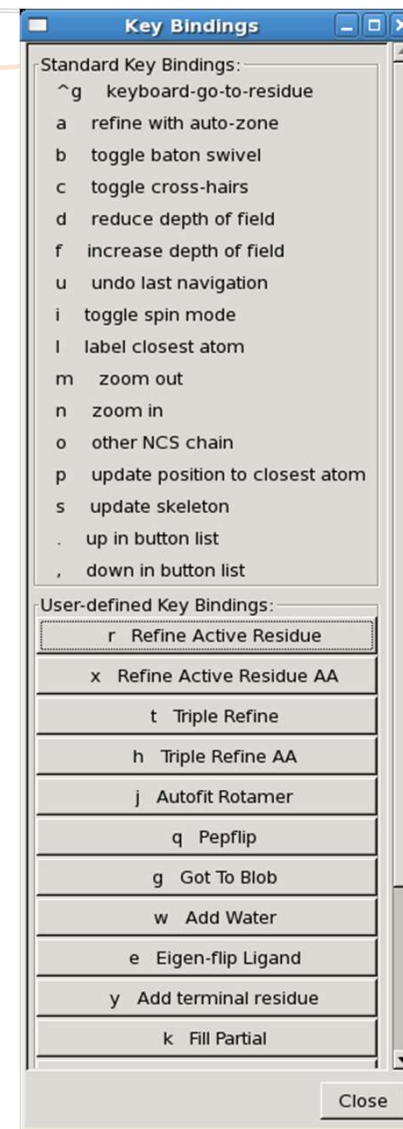
Preferences tabs

Preferences groups



Speed: key bindings

- Short-cuts to speed up common tasks
 - Some build-in
 - User defined ones
- examples on Coot WIKI pages:
<https://strucbio.biologie.uni-konstanz.de/ccp4wiki/index.php/Coot>



Speed: key bindings

- example: active residue refinement → key “x”

in scheme:

```
(add-key-binding "Refine Active Residue" "x"  
  (lambda () (refine-active-residue)))
```

in python

```
add_key_binding("Refine Active Residue", "x",  
  lambda: refine_active_residue())
```

- Makes Coot easy to use (but harder to learn)
- Crib sheet available here:

<https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/web/docs/crib-sheet.pdf>

demo?

Curlew

- Coot Utility Refinement Library Extension Wrangler mechanism to install extensions, tweaks etc
- new in 0.9.x
- File → Curlew



Real Space Refinement



Real Space Refinement

- The adjustment of model parameters (co-ordinates) so that the calculated structure factors match the observations as nearly as possible
- In “one-shot” real-space refinement, such as in Coot, this translates to:
 - move the atoms into as high density as possible while minimizing geometrical distortions

Real Space Refinement

- Major Feature of Coot
 - Gradient-based minimiser (BFGS derivative)
- Geometry library is the standard CIF-based Refmac dictionary
 - Minimise deviations in bond length, angles, torsions, planes, chiral volume, non-bonded contacts, Ramachandran
 - Including links and modifications
- Provides “interactive” refinement – fast and animated
- Nice and tight geometry (can set X-ray/geometry weight)
- Subject to substantial extension

Real Space Refinement

- reftmac monomer library (e.g.: bonds in Ala)

loop_

_chem_comp_bond.comp_id

_chem_comp_bond.atom_id_1

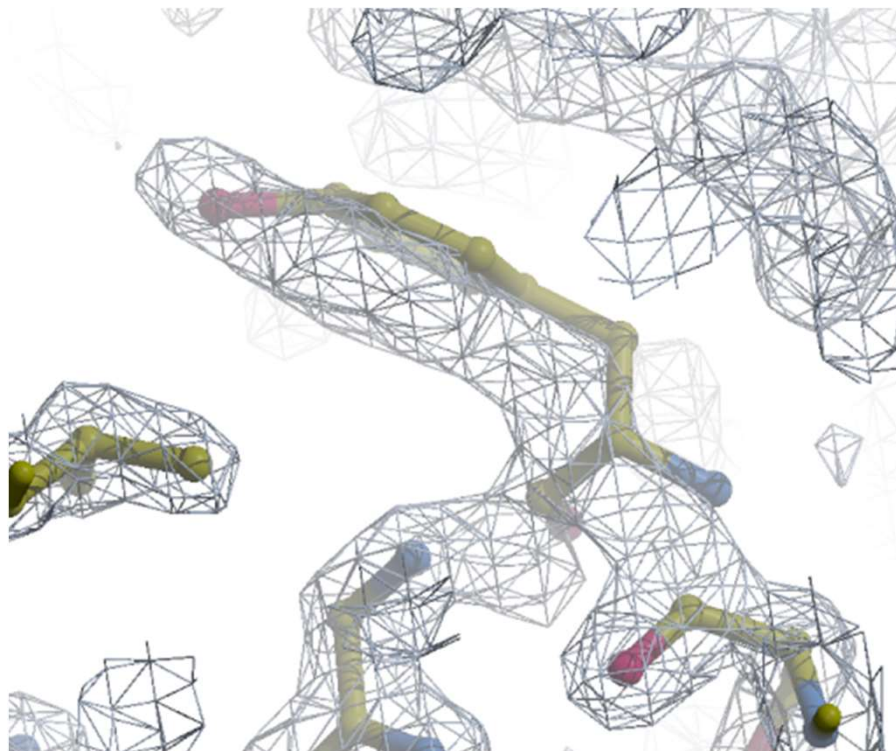
_chem_comp_bond.atom_id_2

_chem_comp_bond.type

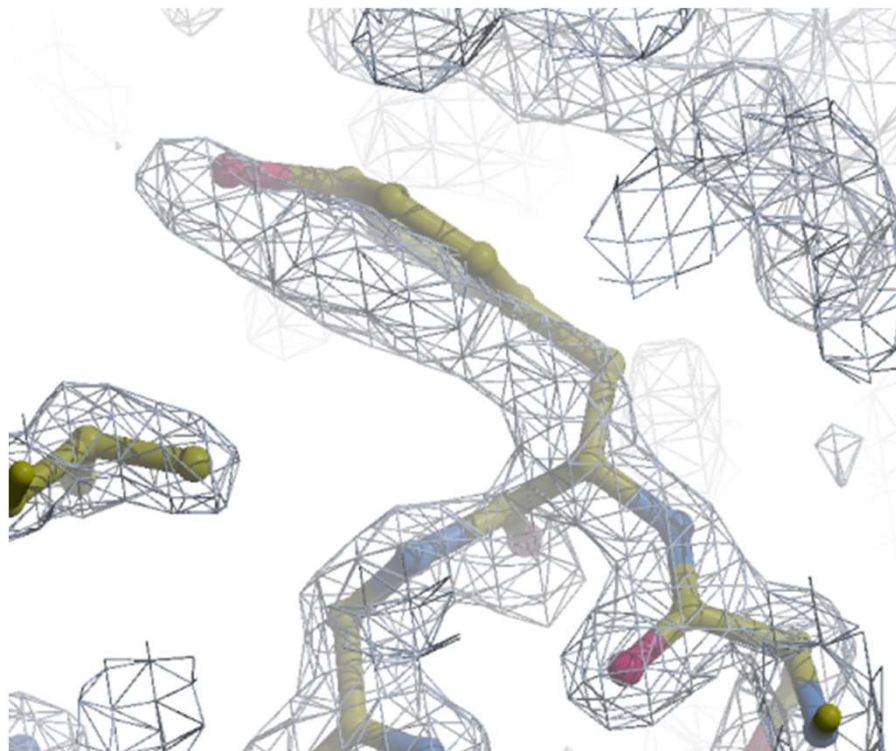
_chem_comp_bond.value_dist

_chem_comp_bond.value_dist_esd

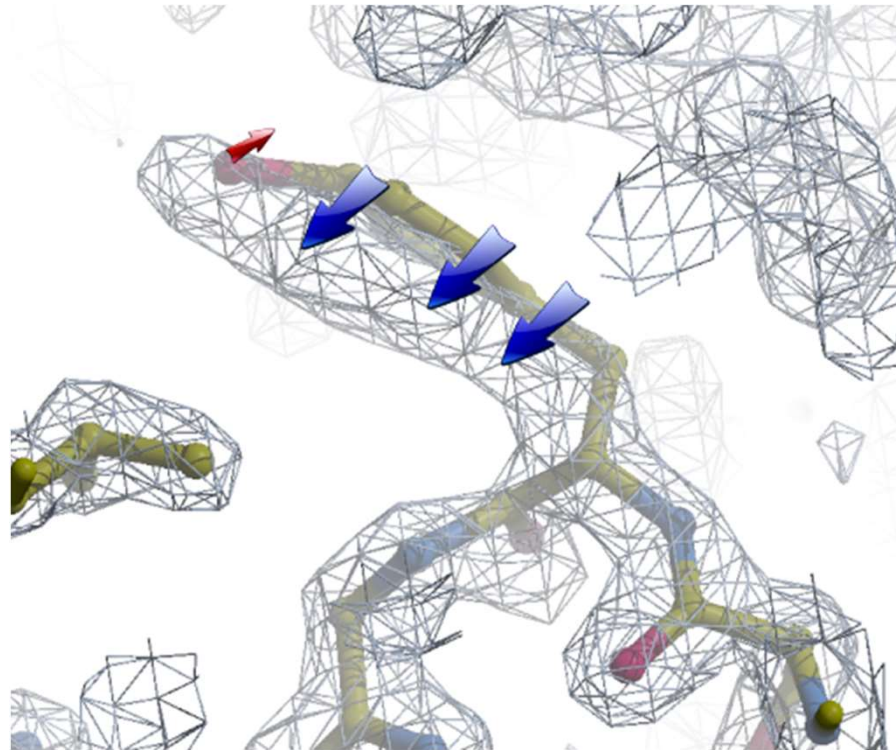
ALA	N	H	single	0.860	0.020
ALA	N	CA	single	1.458	0.019
ALA	CA	HA	single	0.980	0.020
ALA	CA	CB	single	1.521	0.033
ALA	CA	C	single	1.525	0.021
ALA	C	O	double	1.231	0.020



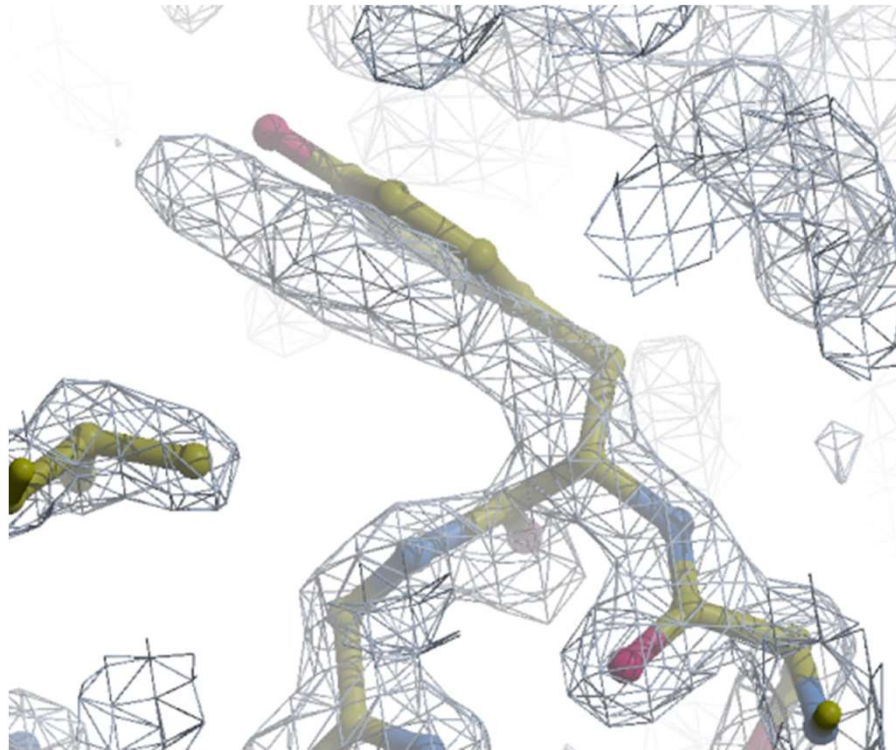
distorted geometry



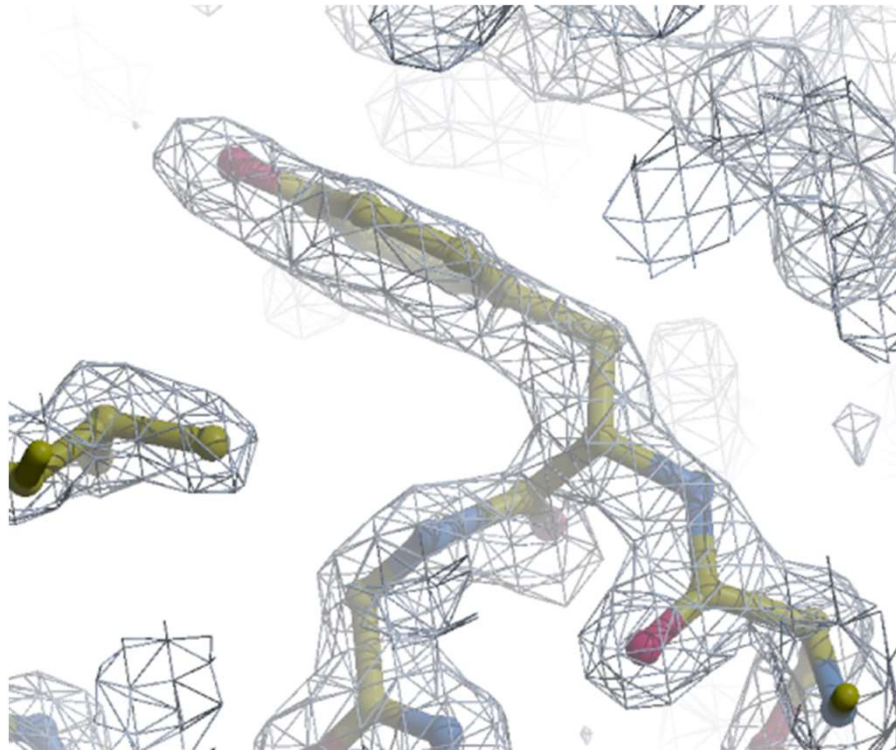
mainchain atoms connect



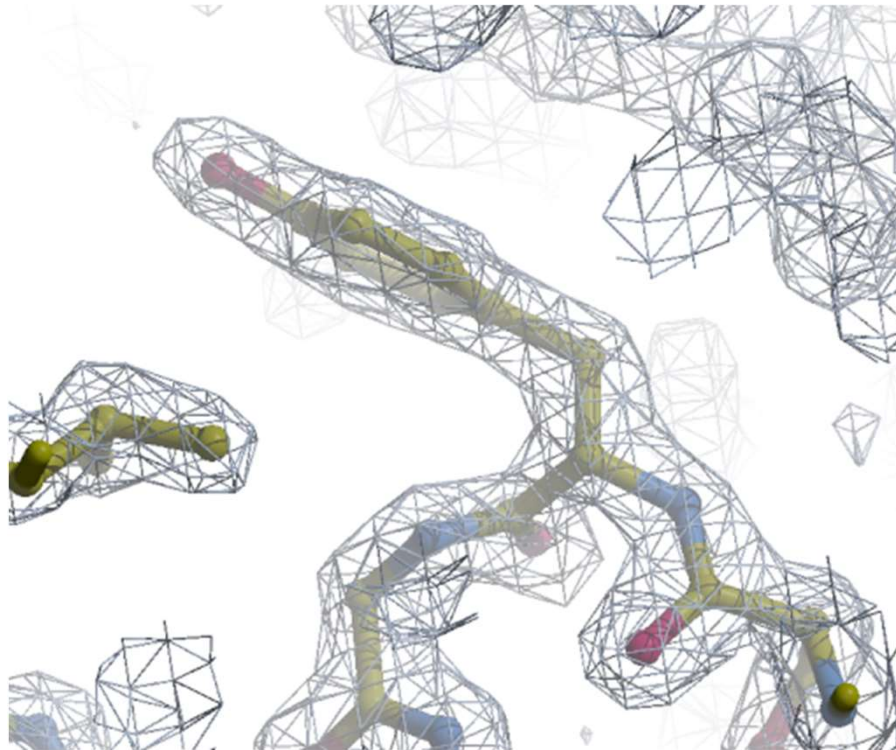
refinement
gradients



after a few cycles of refinement



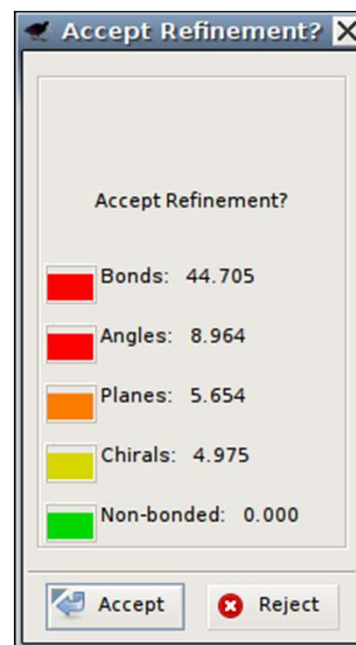
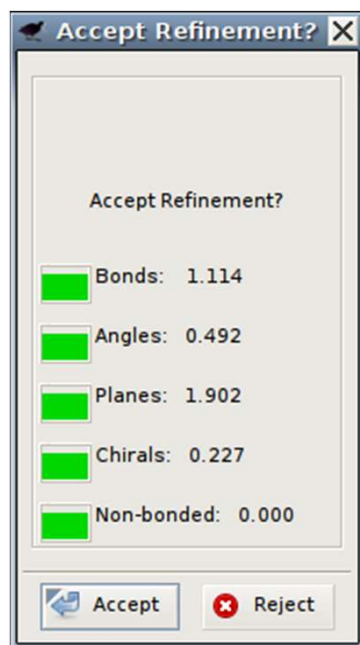
more cycles



convergence

Results: traffic lights

“Traffic Lights” represent the RMSd values for each of the refined geometry types



Refinement techniques

- Auto-zone
- Single-Atom Drag
- ~~Over-dragging~~ Obsolete! Instead: pink stick, green peas
- Key-bindings:
 - Triple Refine
 - Single Residue Refine with Auto-accept
 - Sphere refinement
- Ramachandran Refinement

Sphere Refinement

- Given an “Active” Residue
- Define a sphere of residues around it and use them all for refinement
- NOT just a linear selection: residues from different chains (or different parts of the same chain) interact
- Make CYS-CYS or glycosylation links as you find them
- Use the group and link_list chem_link in the dictionary
- Most powerful keybinding!

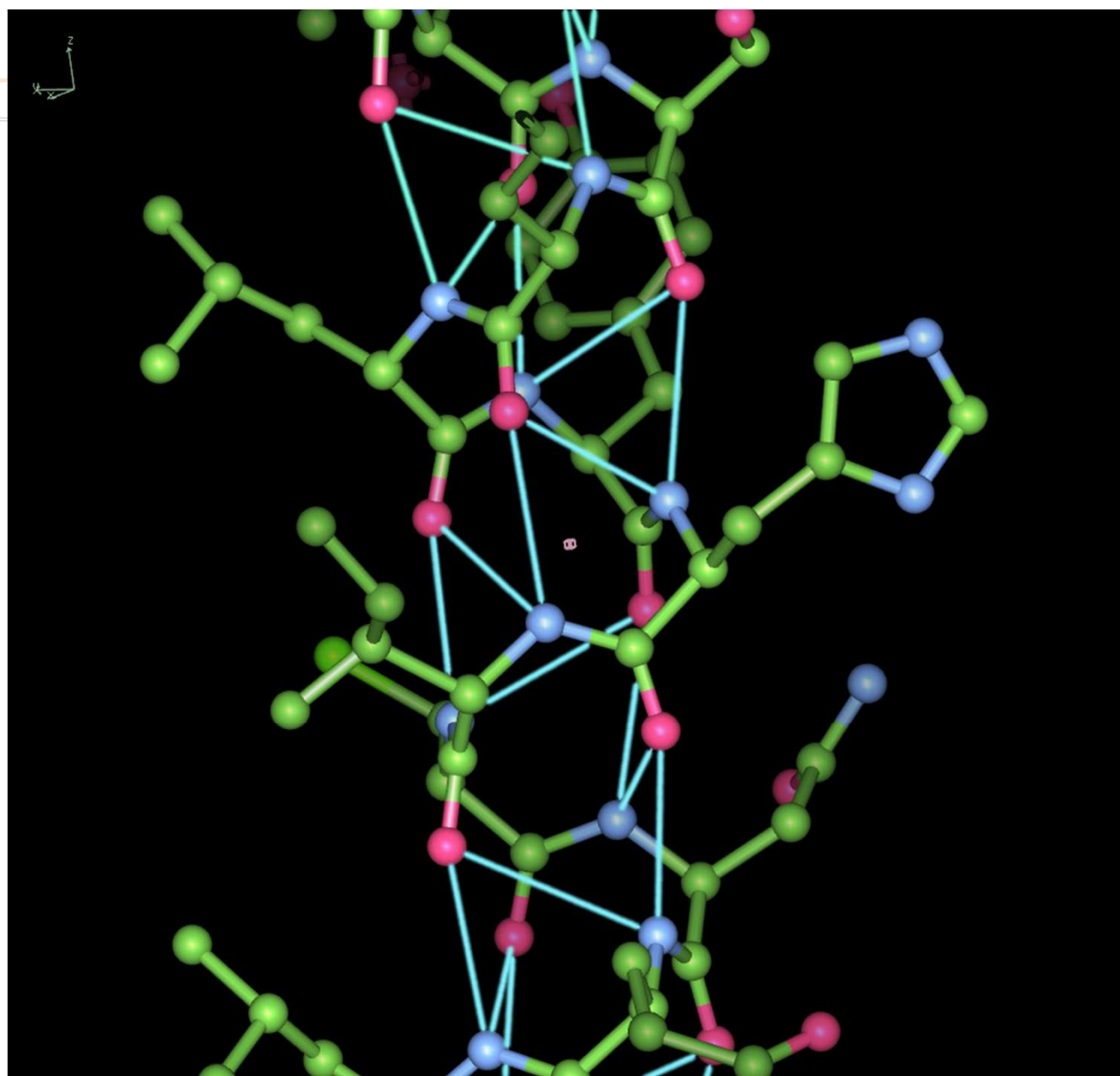
Pink Stick, Green Peas

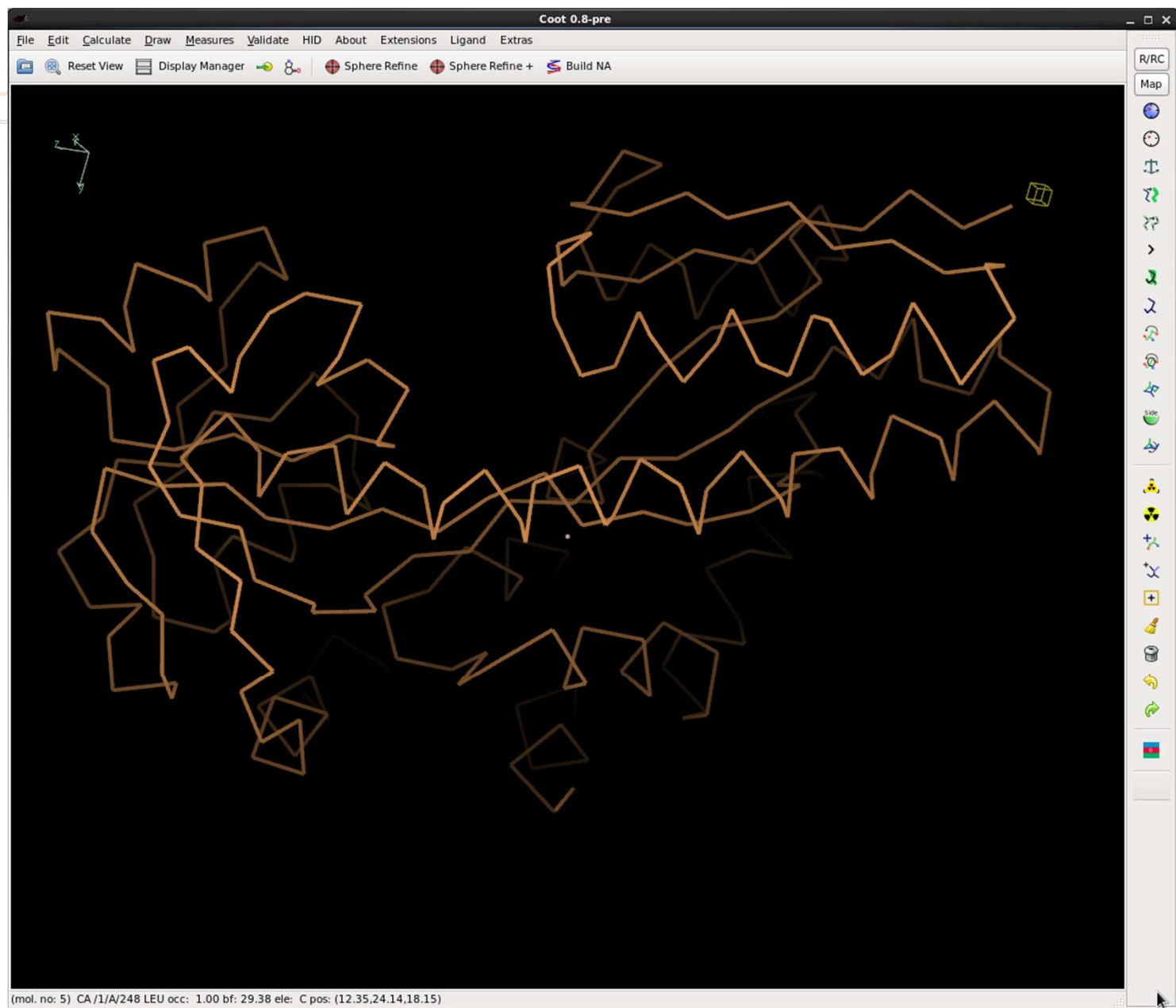
- new real space refinement mode
- new feel: instead of rubber sheet movement, atoms are refined dynamically
- gentle tugs instead of yanking of atoms
- Pink Stick:
 - shows the difference between where an atom is and where you want it to go.
 - should be short and often difficult to see. If you see several long pink sticks, then that's usually a sign that something has gone wrong.
- Green Peas: markers for Ramachandran probabilities

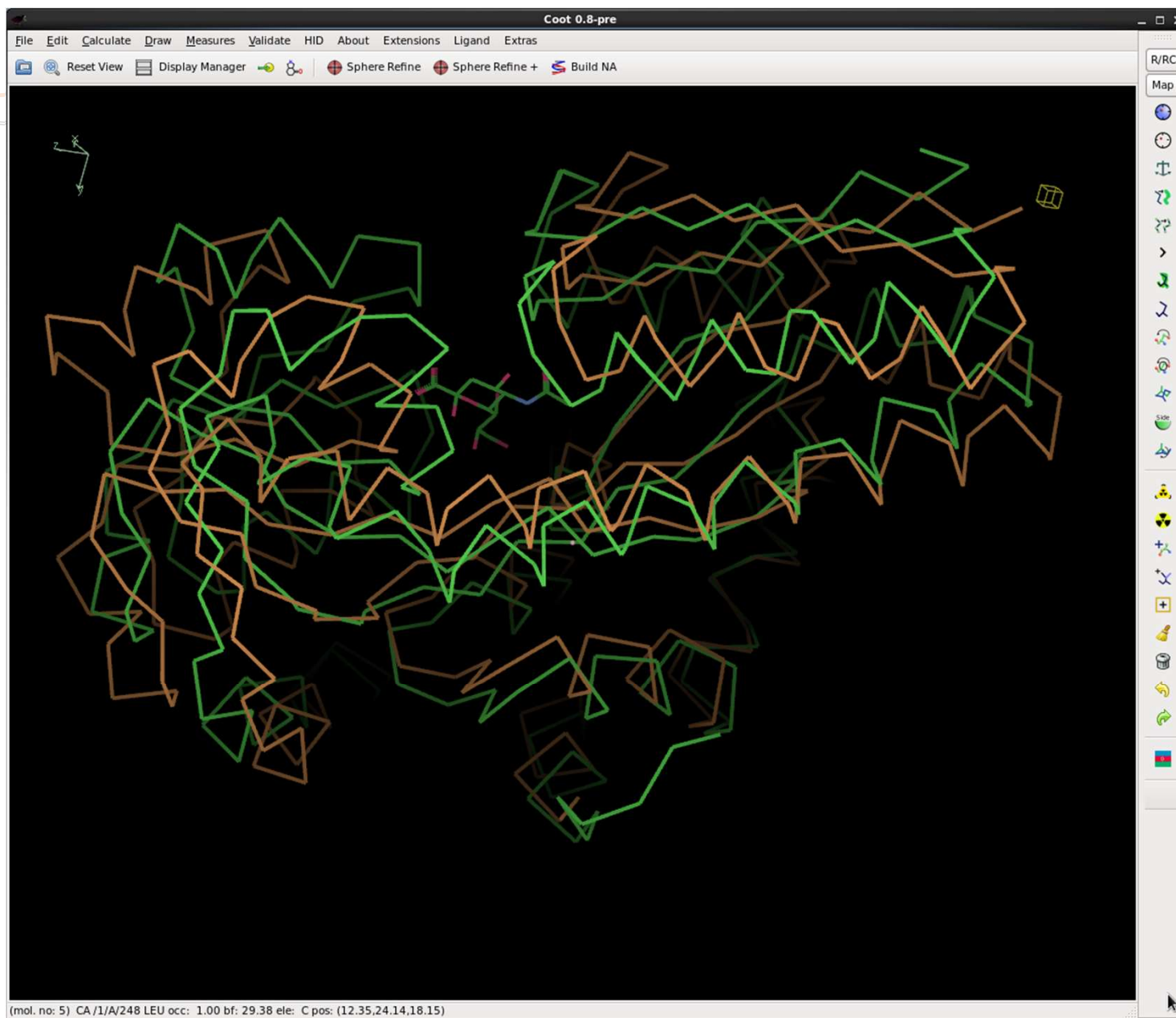
demo?

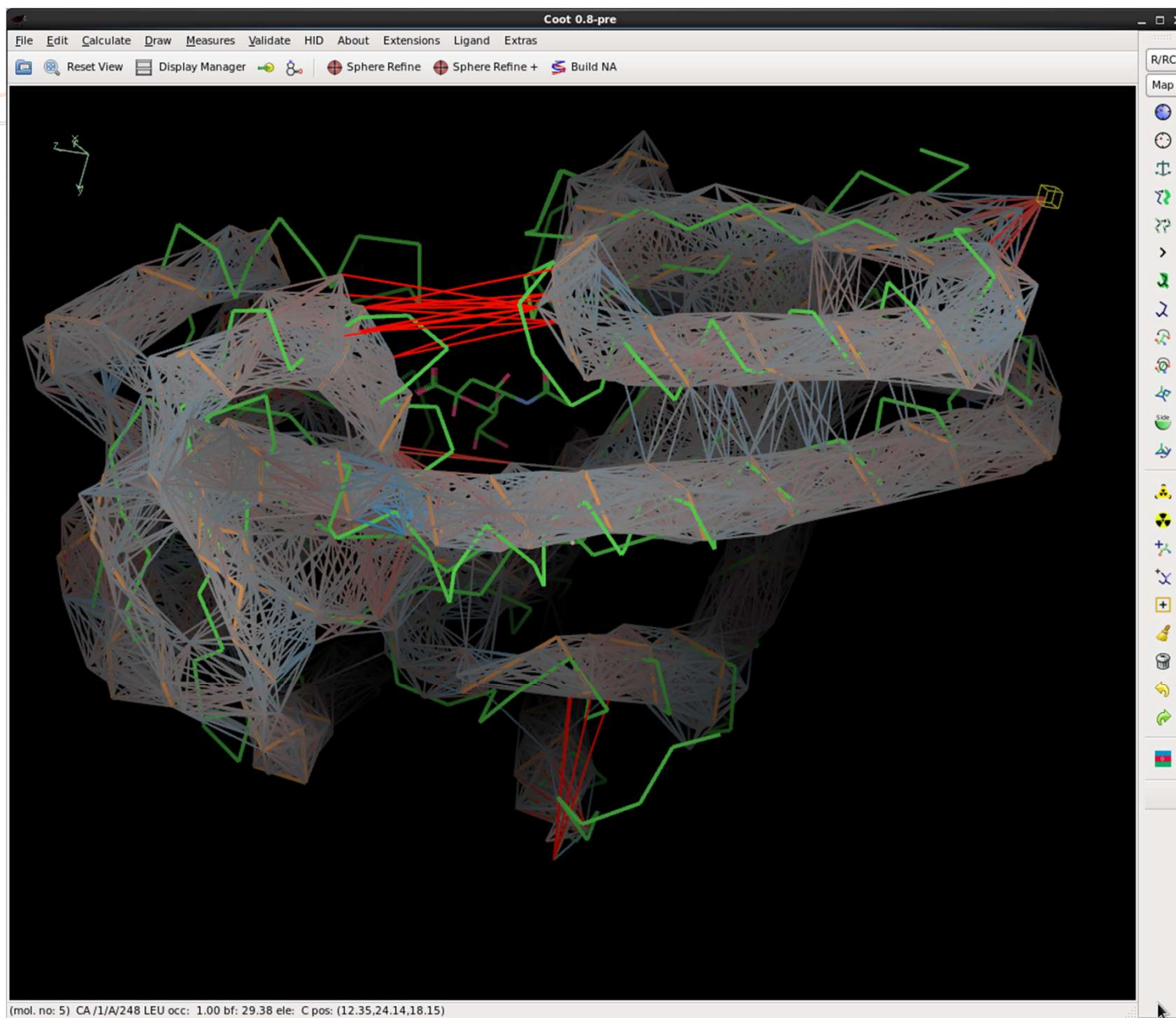
ProSMART interface

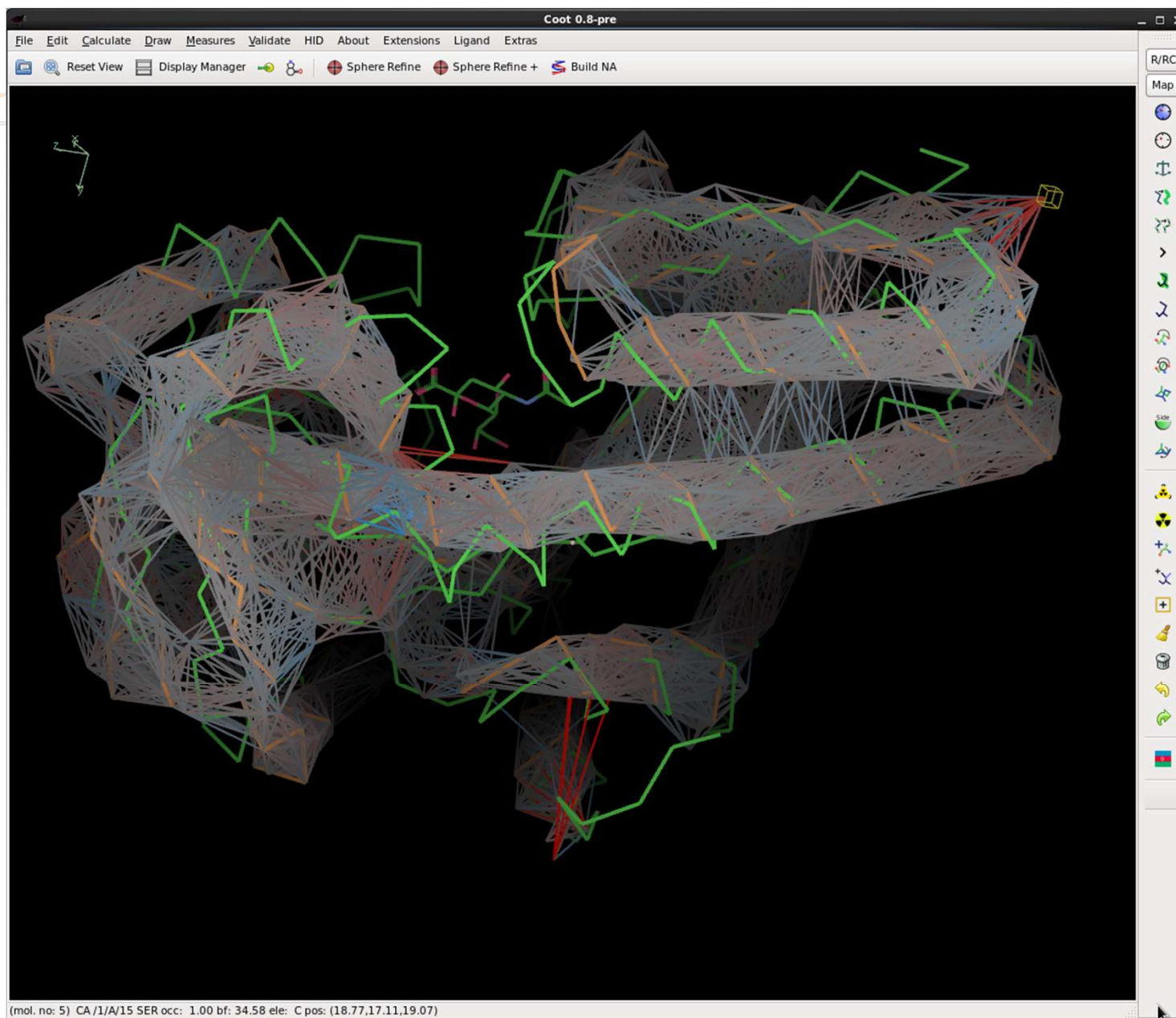
- Use previous-solved “template” structures to inform the refinement of the (low resolution) target protein
- Conformation-independent structural comparison/superposition and restraint generation



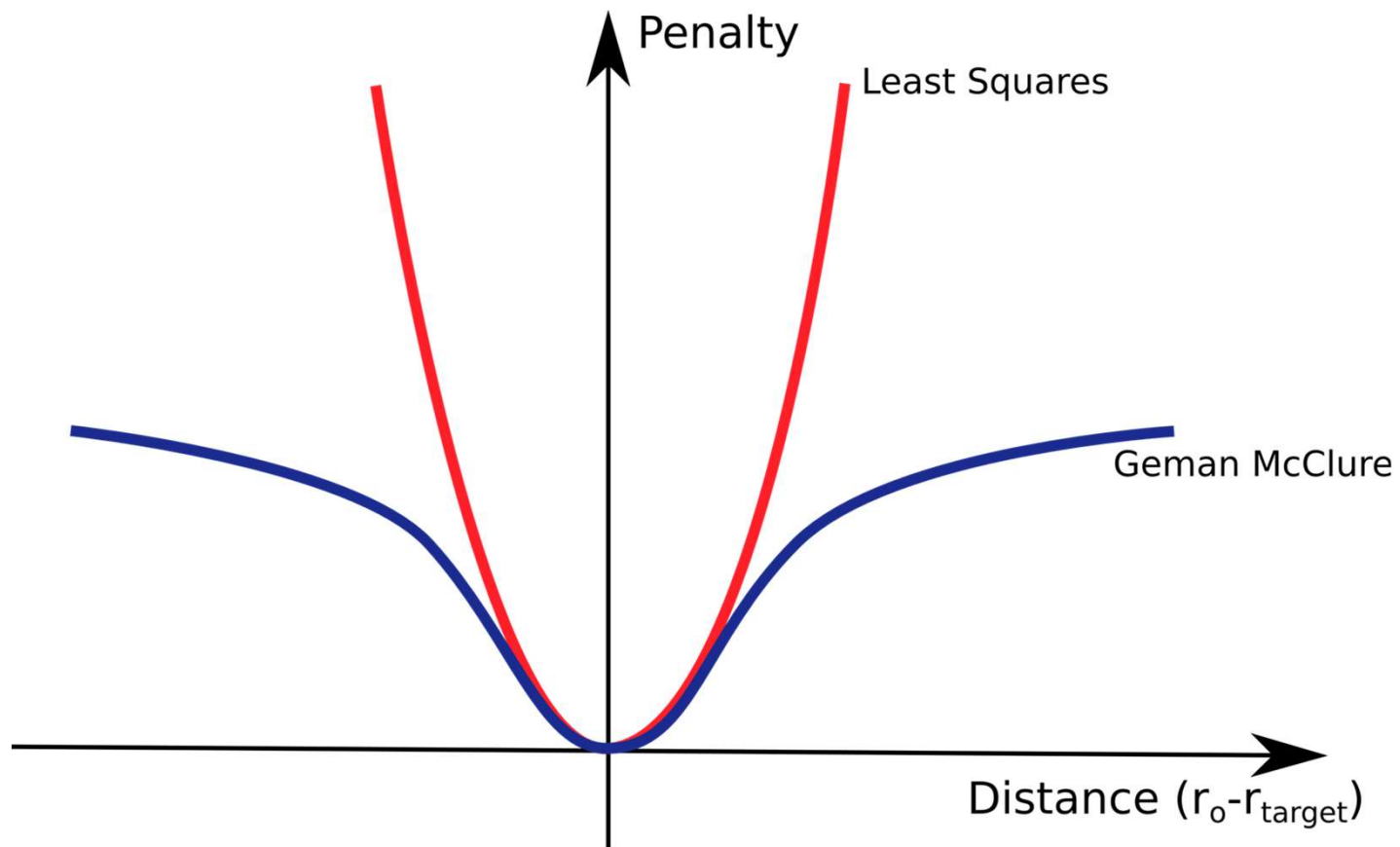








Modified Target Function



Chain Refinement

- a powerful low resolution fitting tool
- refines all atoms of a chain
- ProSMART restraints are recommended (self or external)
- particularly useful with the new **proportional editing** function
- Ctrl+scroll to adjust the radius

demo?

Tools for low resolution




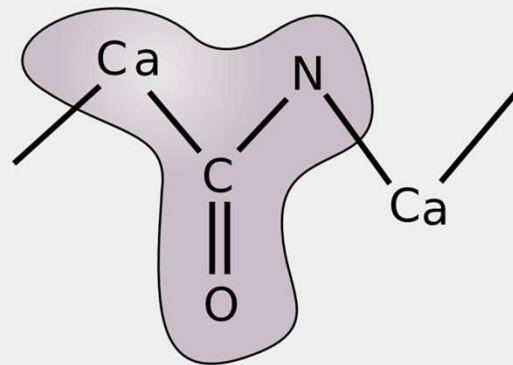
Tools for low resolution

- Additional restraints
 - Peptide plane restraints
 - Ramachandran restraints
 - Secondary structure restraints
 - ProSmart interface
- Backrub rotamers
- Jiggle Fit
- Map sharpening

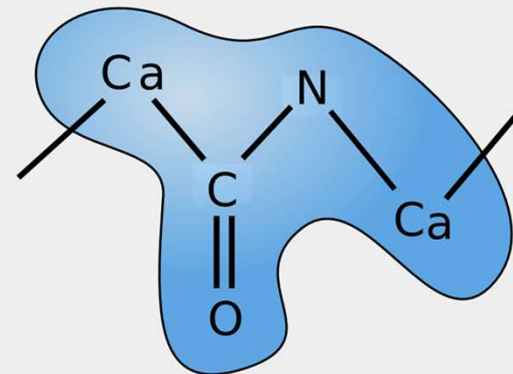
Additional Restraints

Coot's Extra Peptide Plane Restraint

 Default Refmac Peptide Plane



 Extended Plane in Coot



(add-planar-peptide-restraints)

Ramachandran Restraints

- Scenario:

I have a loop, with poor density, I know the atoms are there somewhere and I want to provide a “reasonable” model

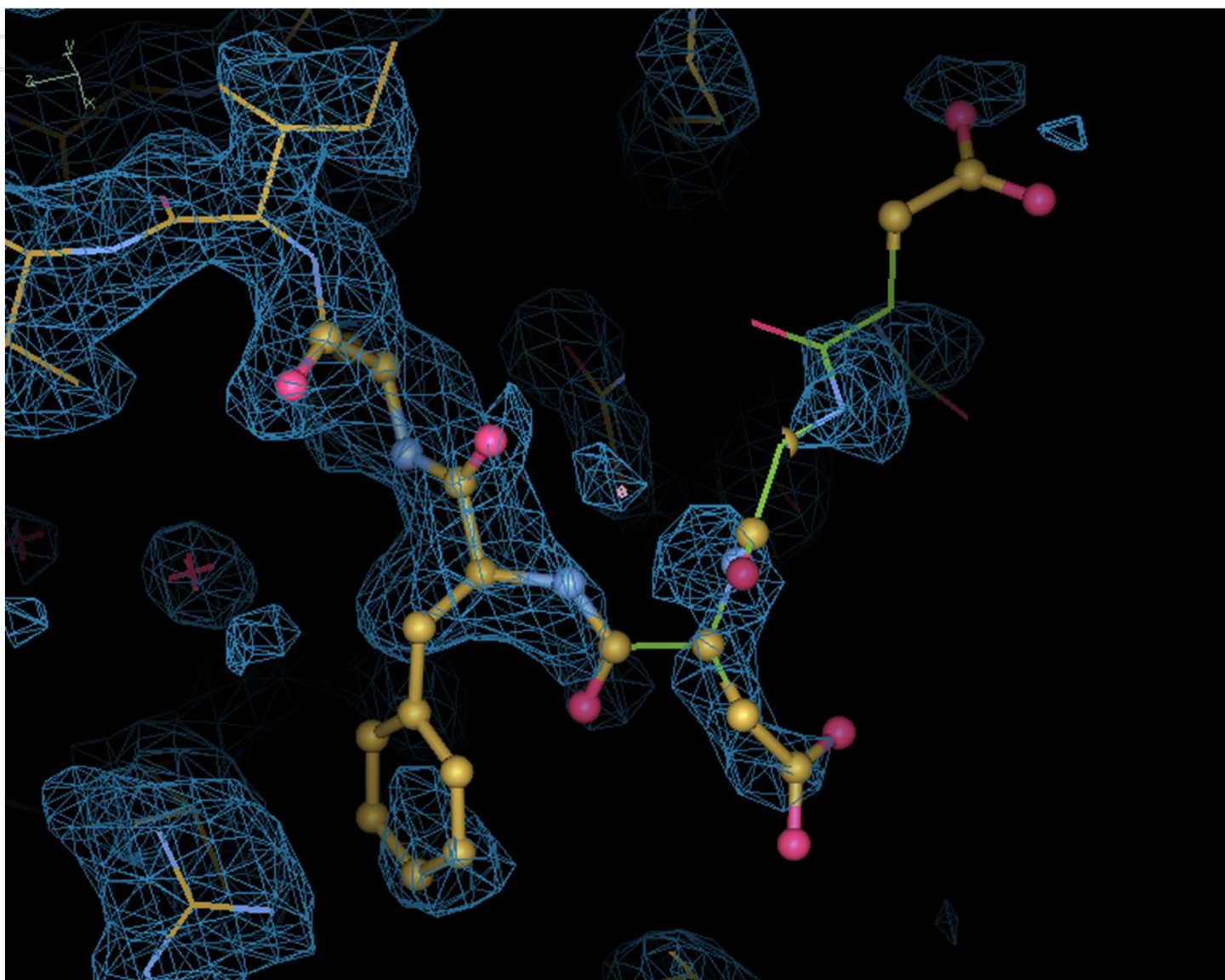
- Controversial Feature?

Ramachandran Plots have been used for “validation” - but here we are deliberately optimizing them

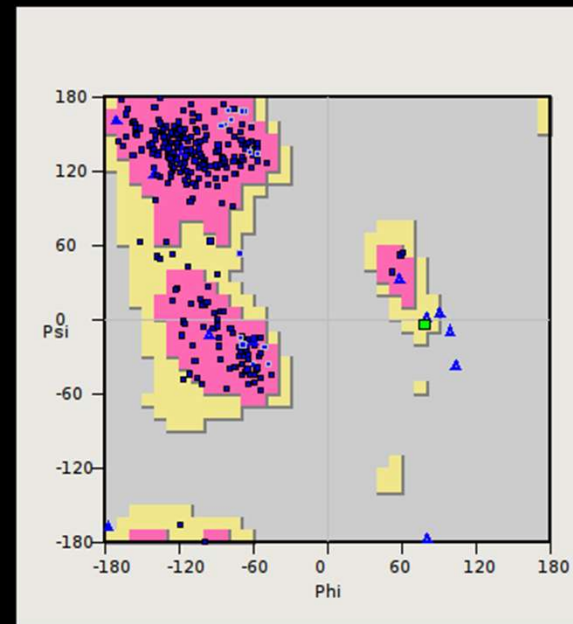
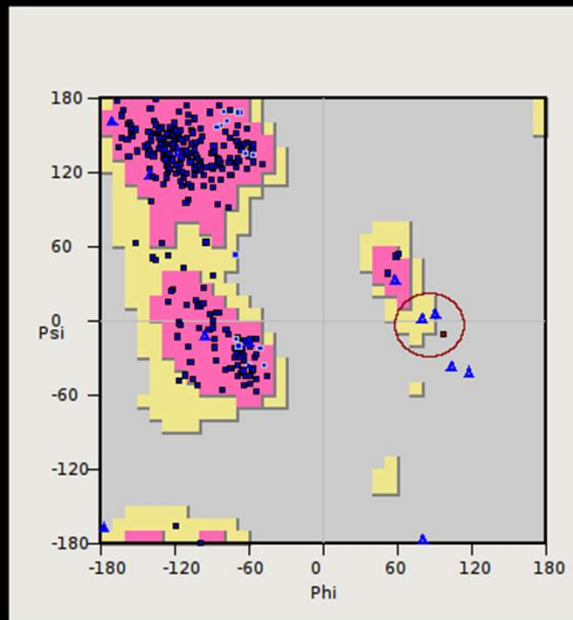
- But to quote Jane Richardson:

Do you want a better structure – or a better idea of the quality of your structure?

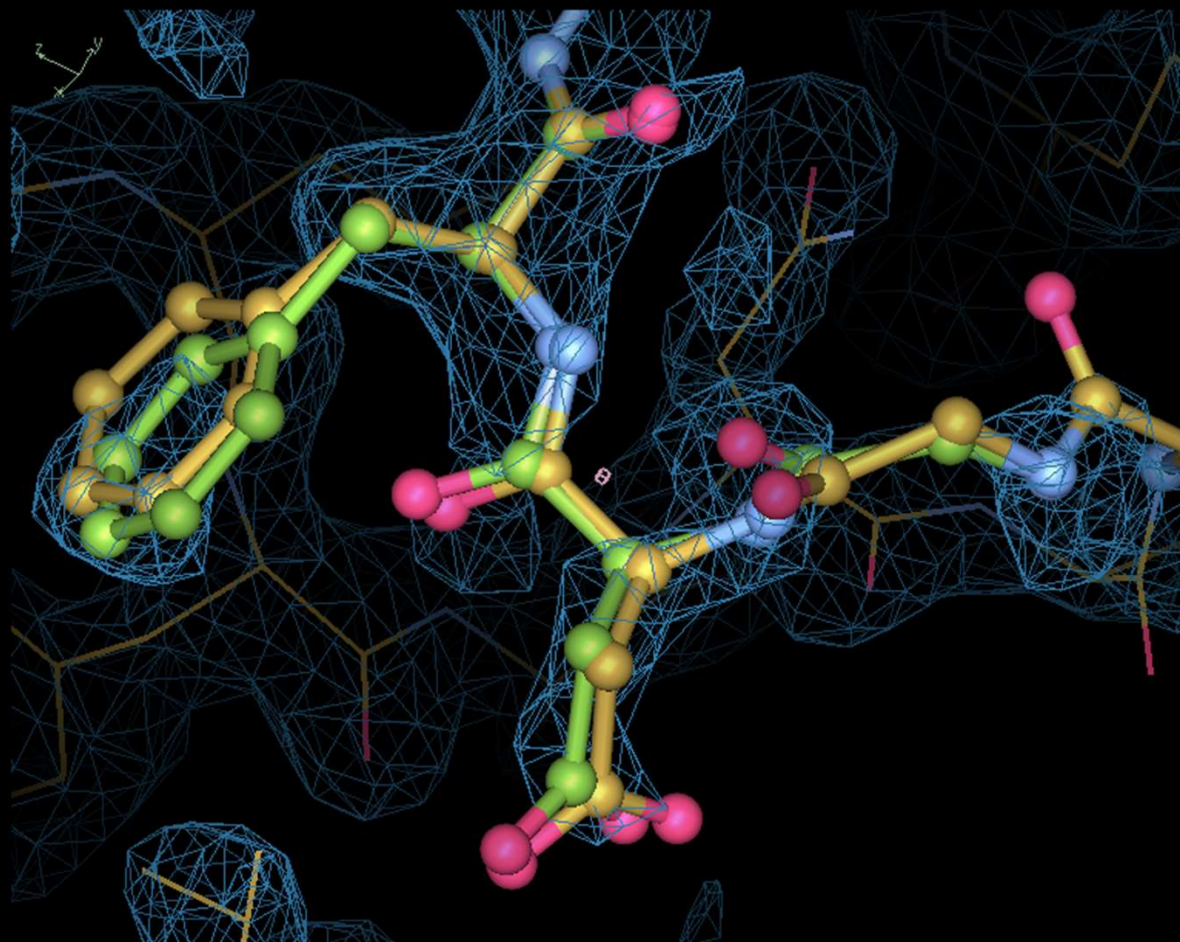
- Ramachandran Plots can be added to the geometry target function

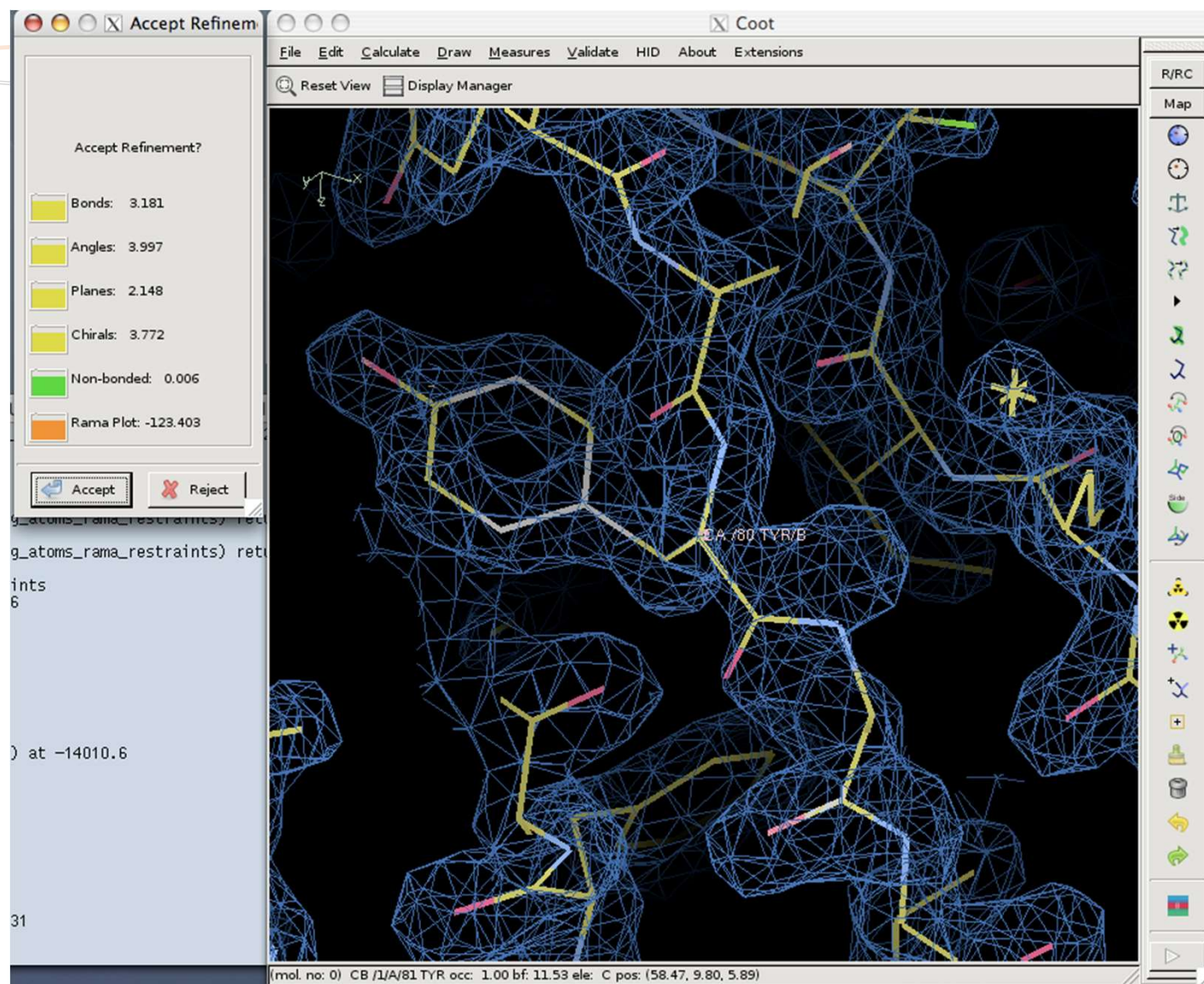


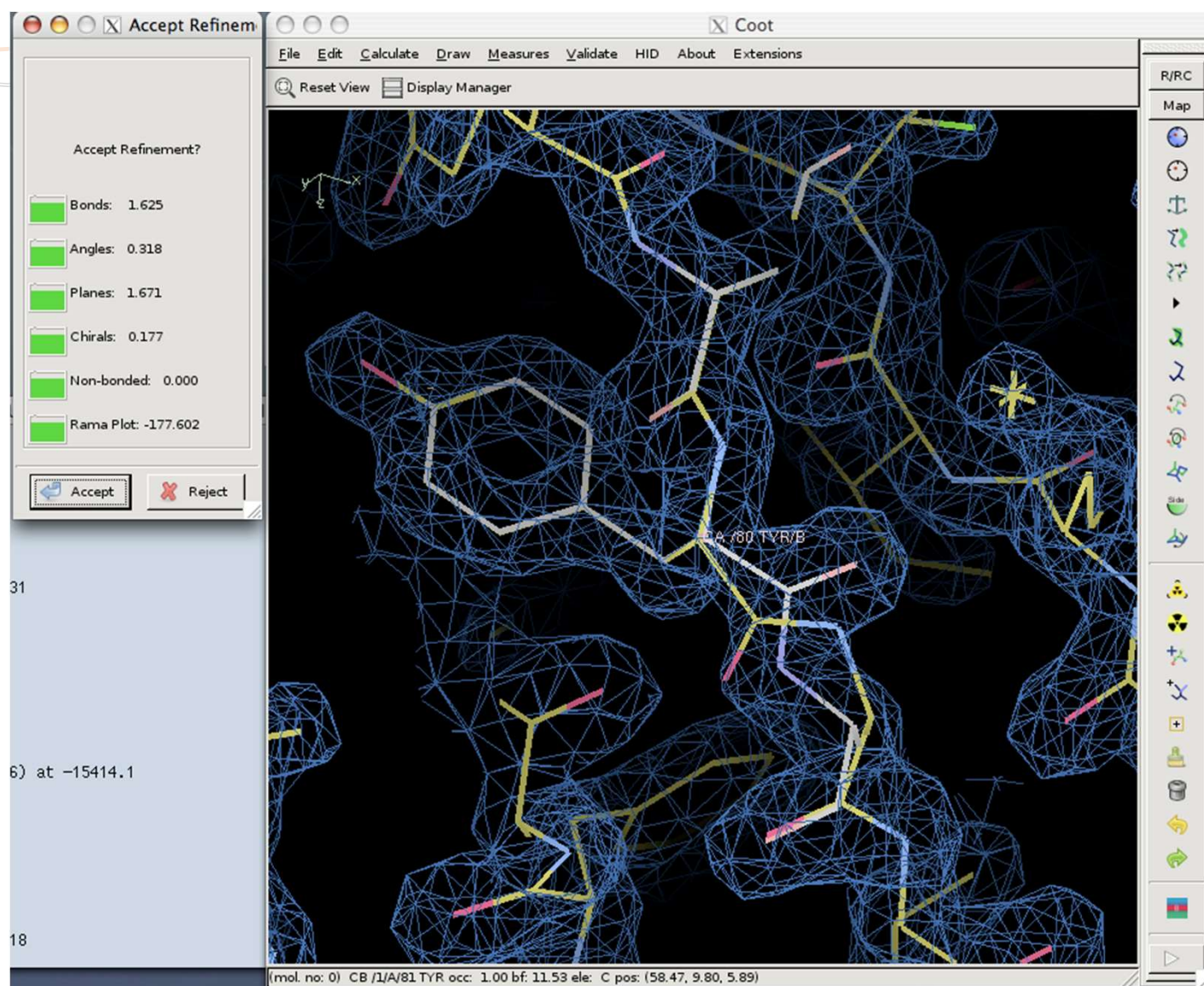
Tweaking a Ramachandran Outlier



Tweaking Phi and Psi





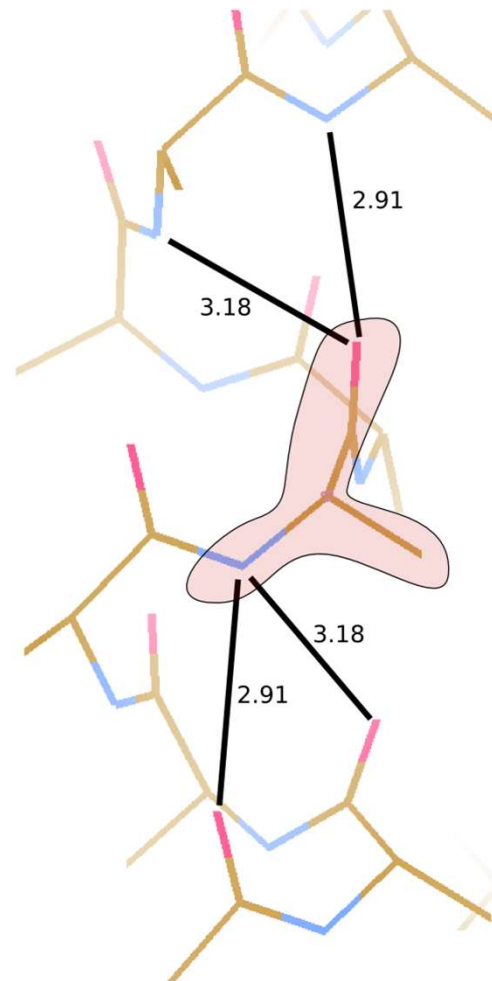


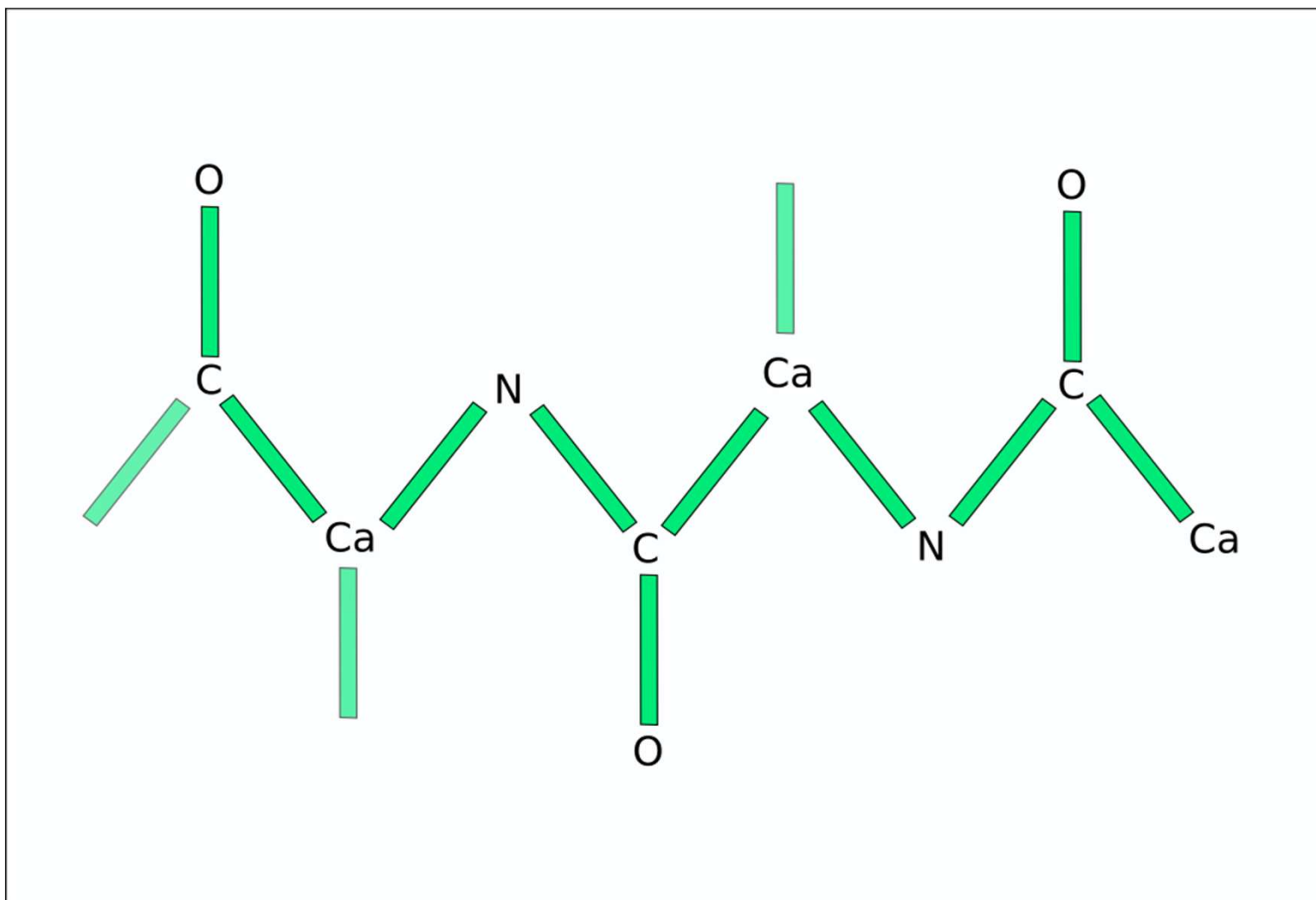
Torsion angle restraints

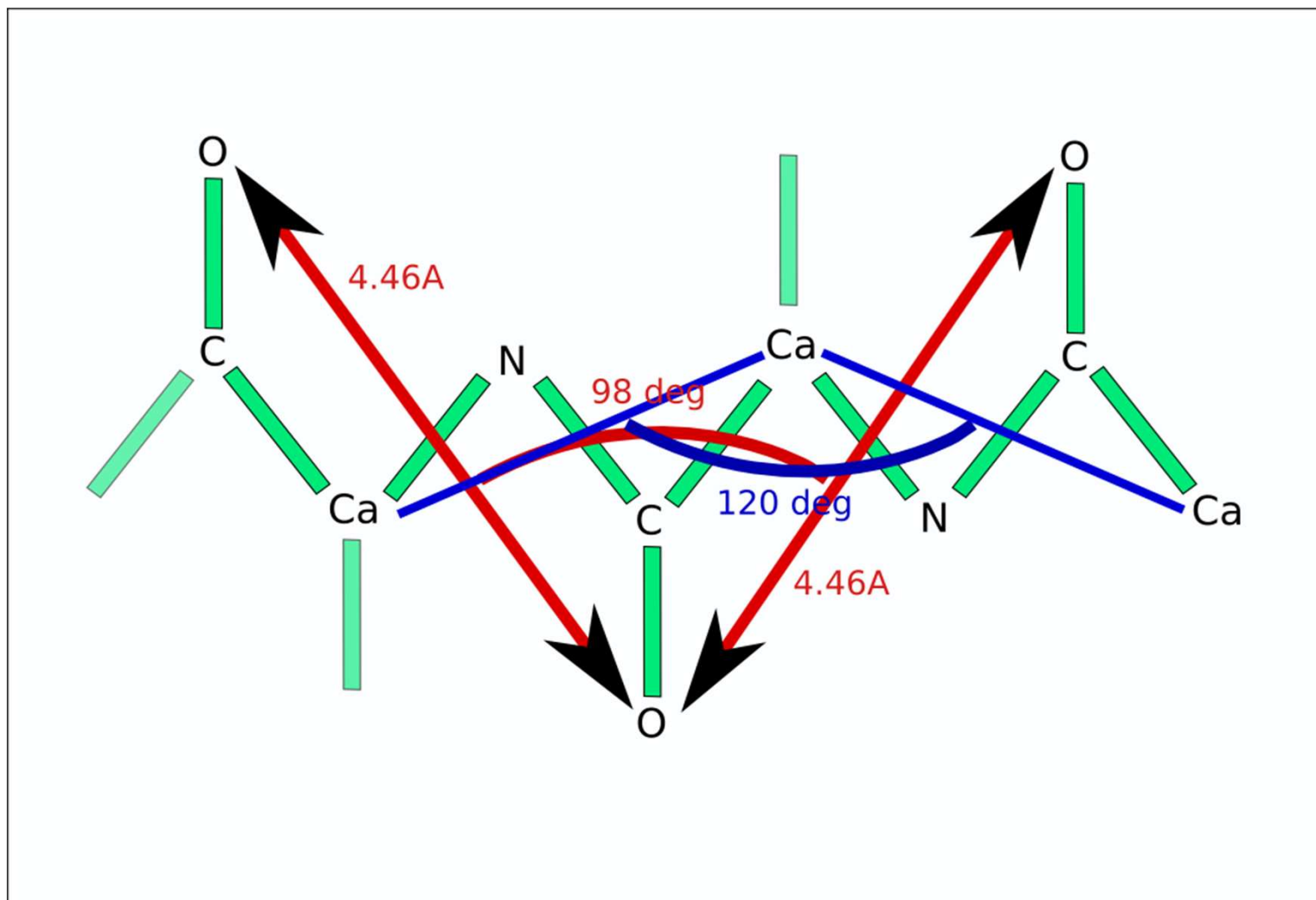
- Torsion angle refinement is slow (relatively)
- Simple addition of these restraints to the geometry target function often makes the region “stuck and unsatisfied”
 - i.e. trapped in local minimum
 - apply pseudo-bond restraints instead

Alpha Helix pseudo-bond
restraints

Restrain the Hydrogen-bonding
atom distances



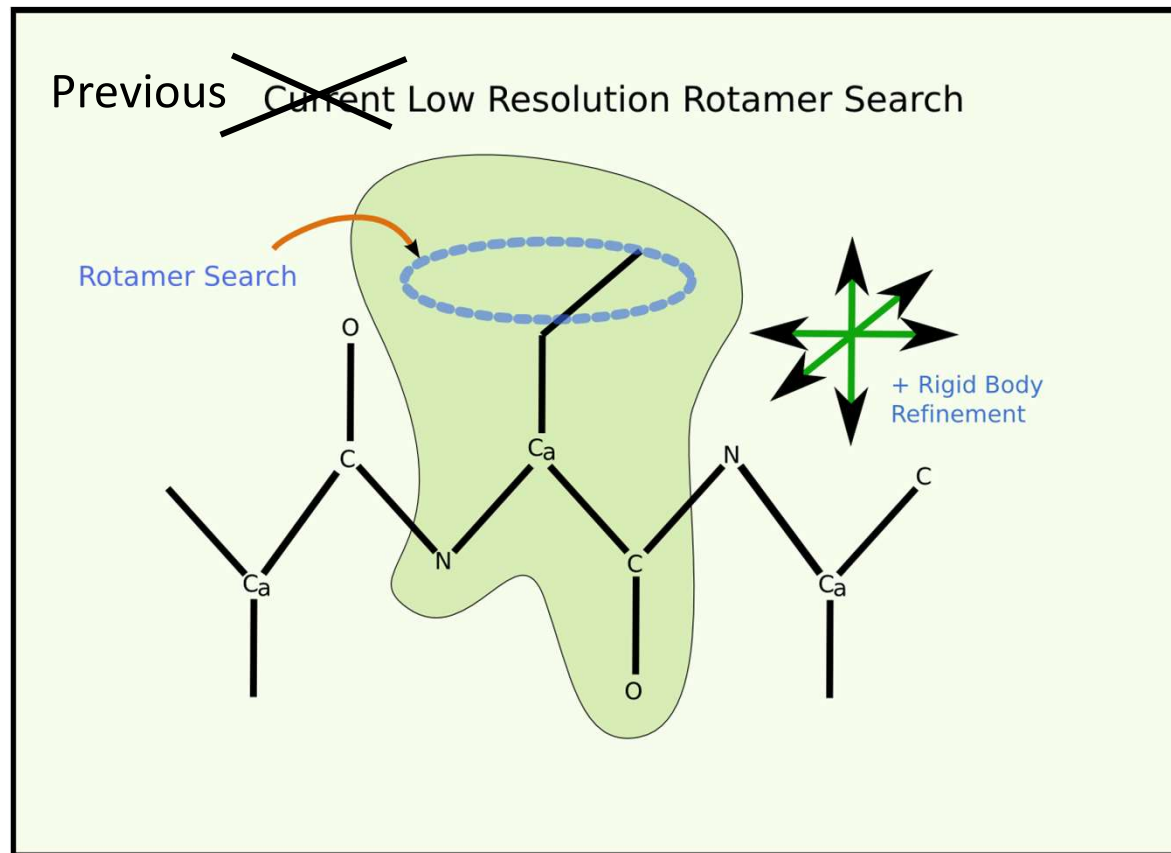


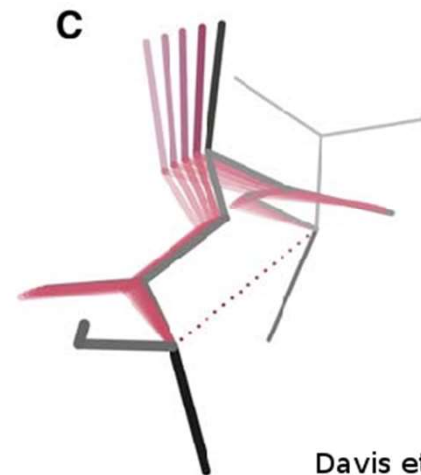
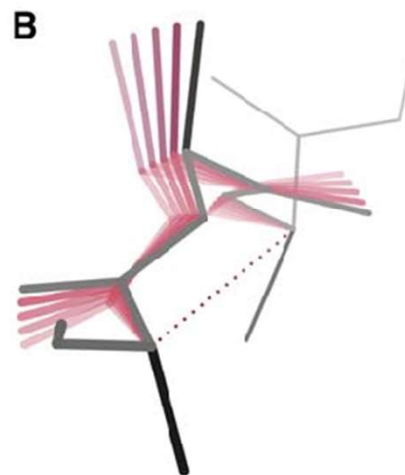
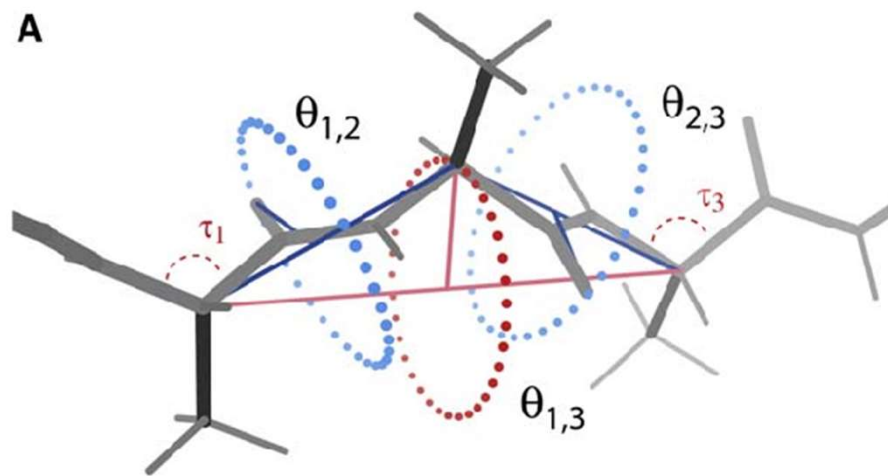


Backrub Rotamers

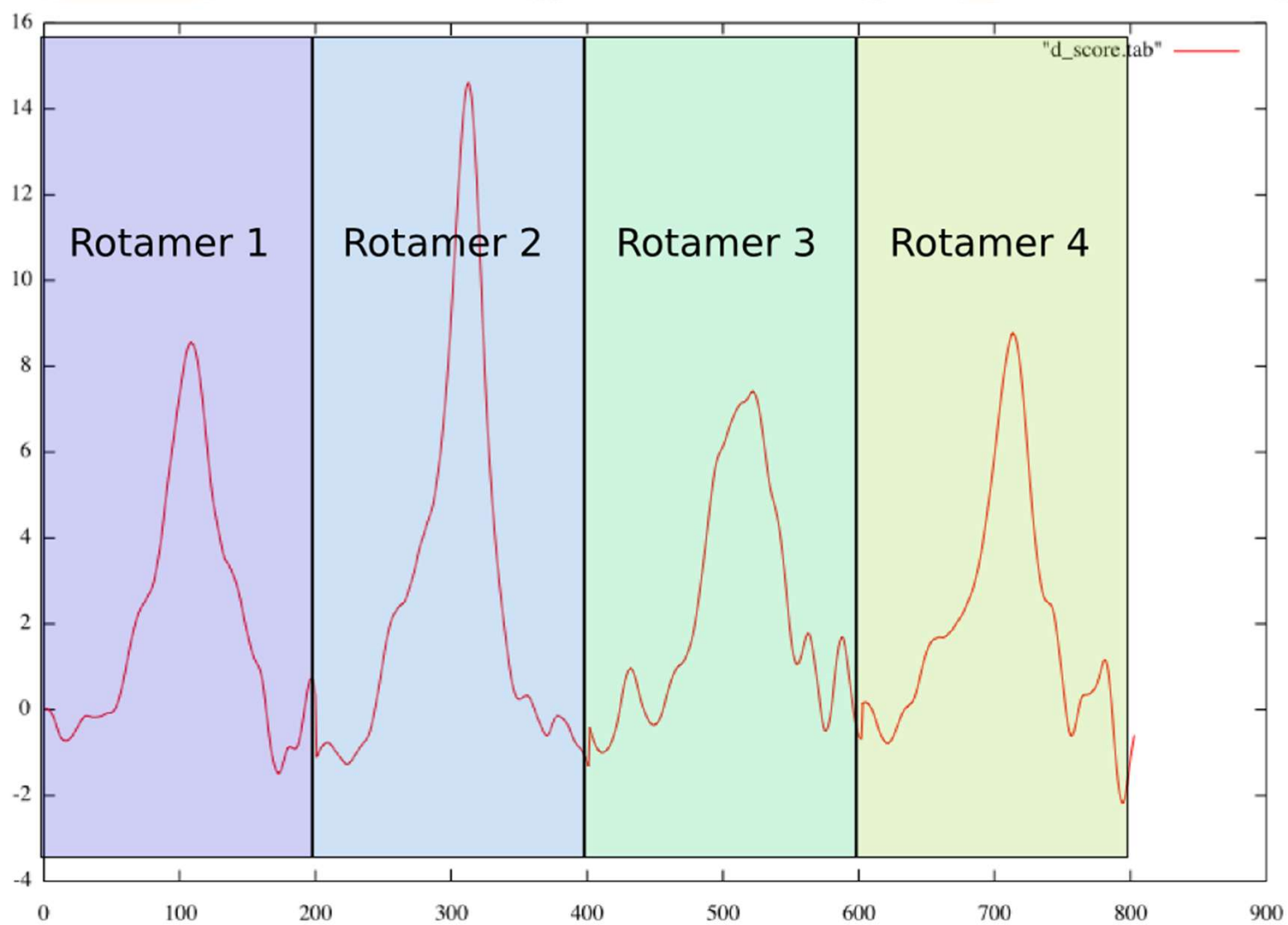
- High probability models with low resolution data
- Rotamers are preferred configurations of a side-chains rotatable bonds, where “preferred” means these configurations occur more frequently in a set of reference protein structures
- “preferred” because they are low-energy conformations
- Several Rotamer “databases” exist
 - best: (Son of) Penultimate Rotamer Library
- To activate it:
 - (ROTAMERSEARCHLOWRES)
 - Default for resolutions worse than 2.6Å
 - Add a toolbar button (right click on bar...)

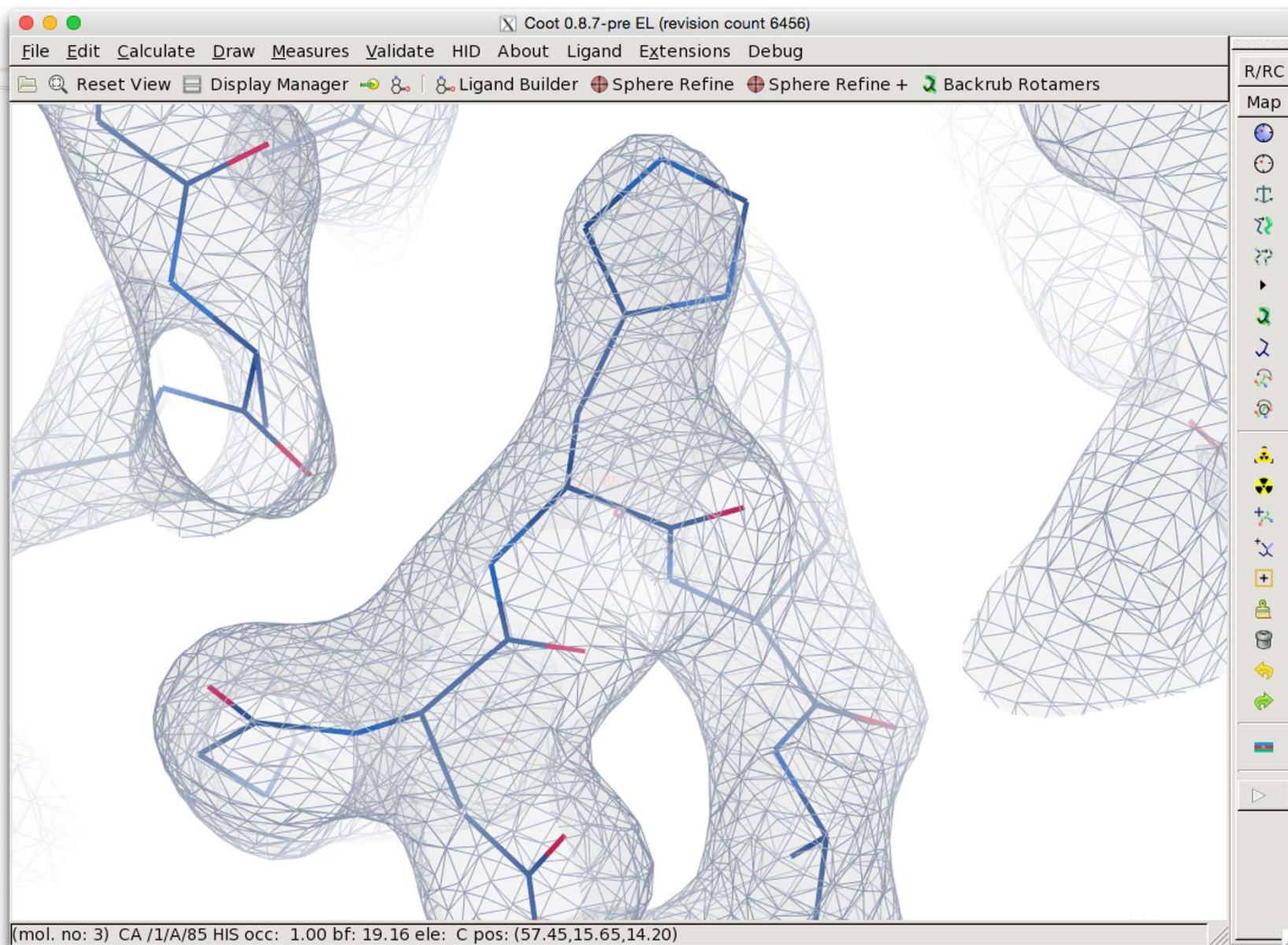
Backrub Rotamers

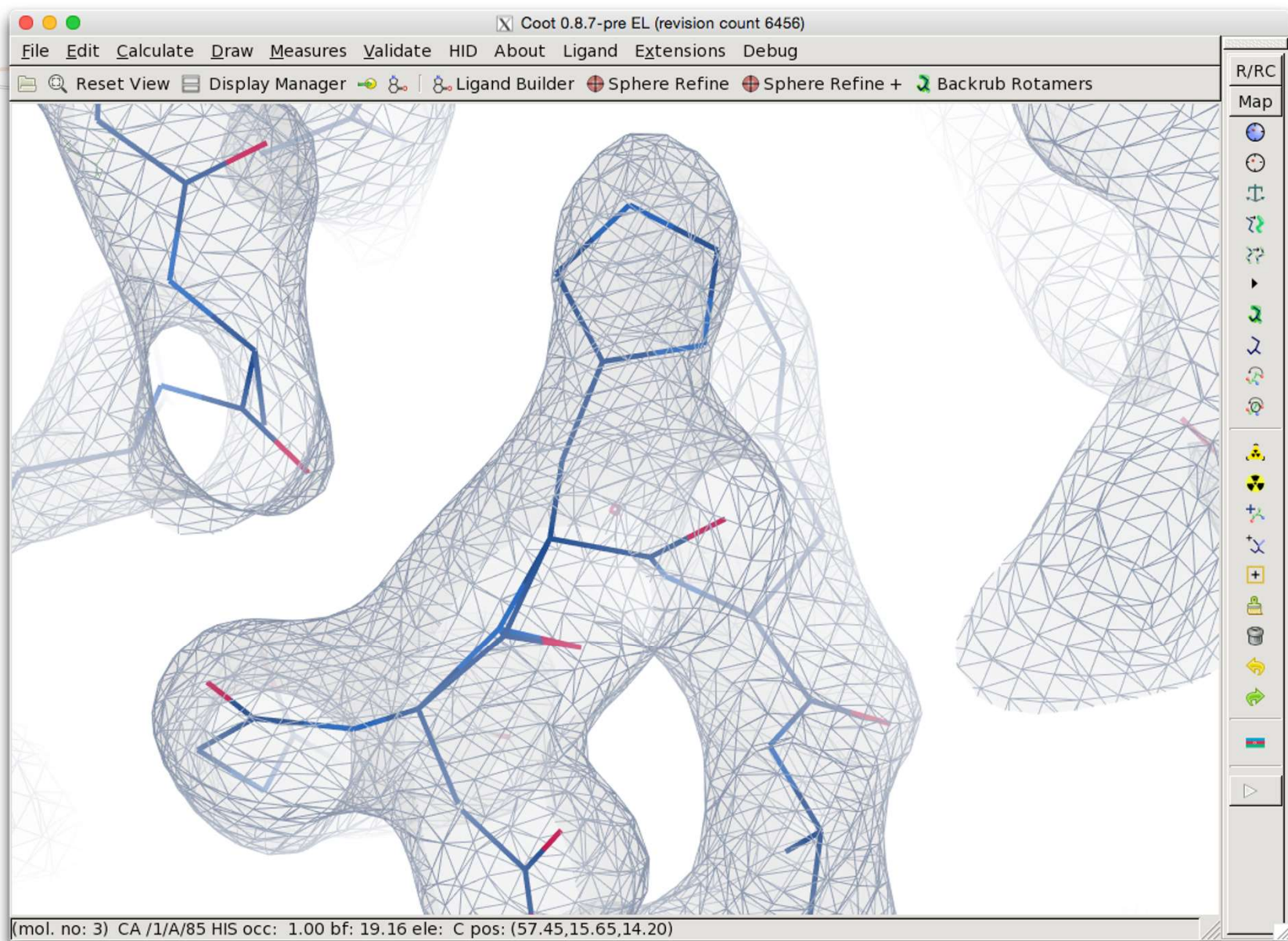


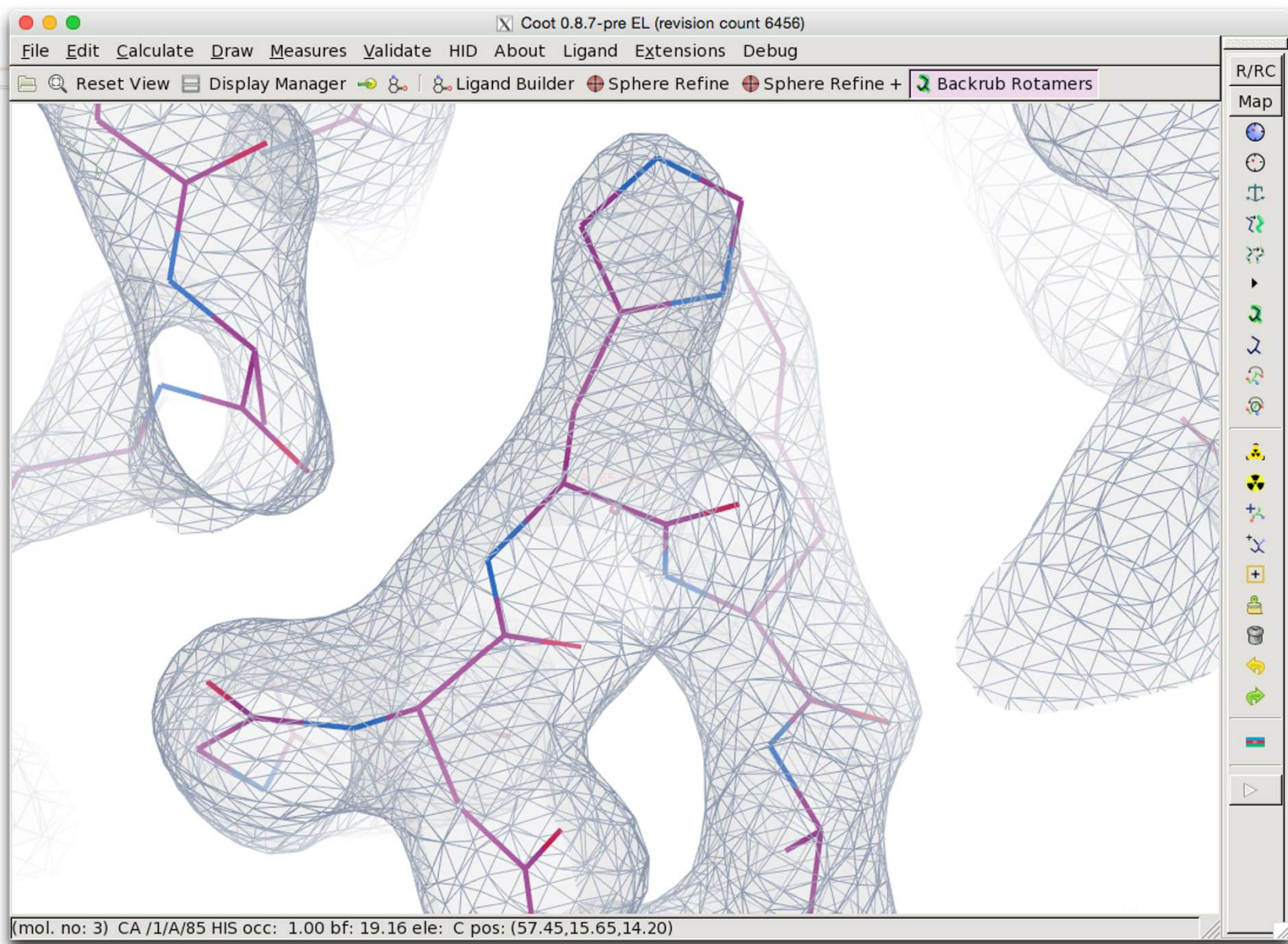


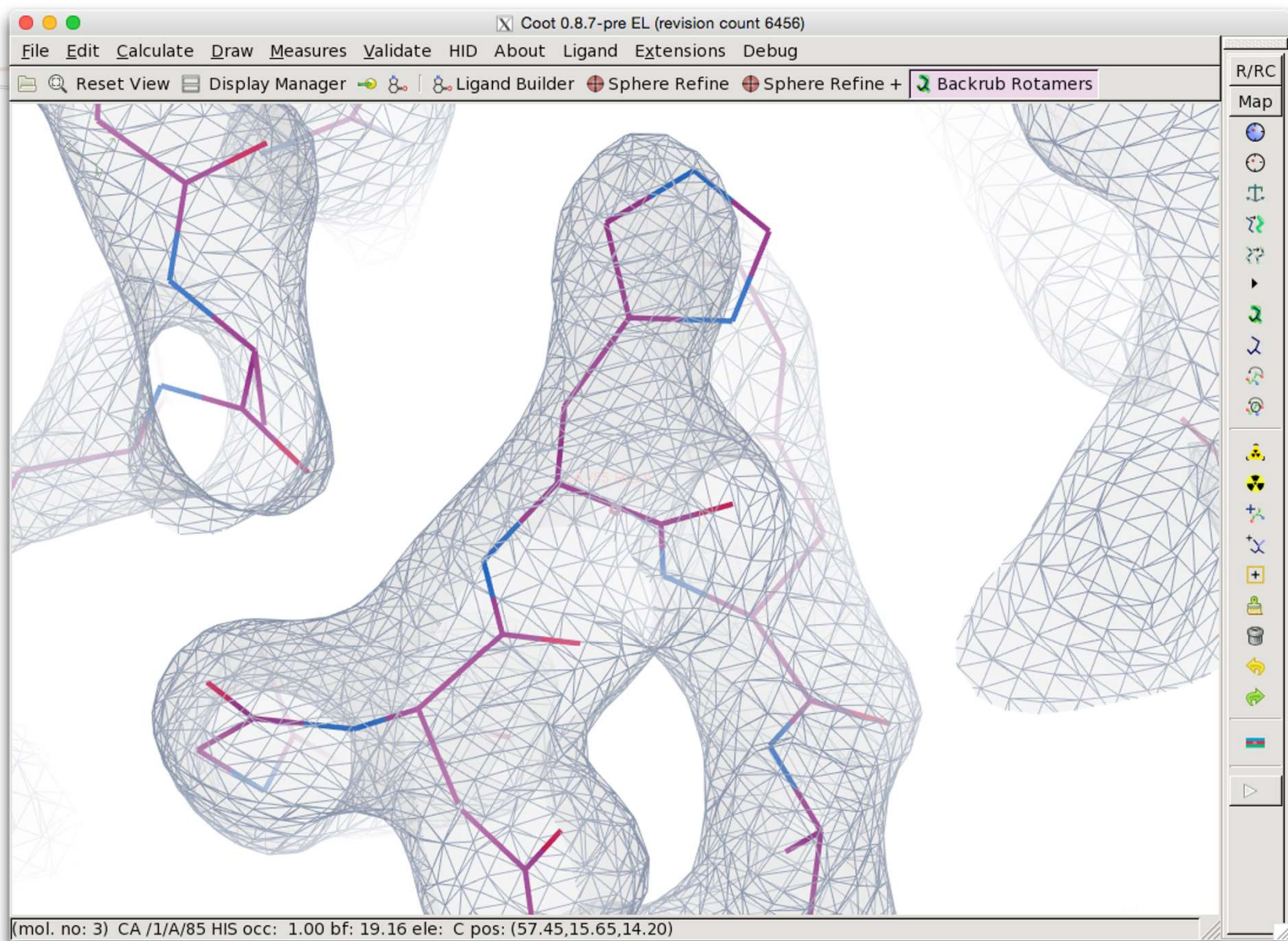
Davis et al. (2006) Structure

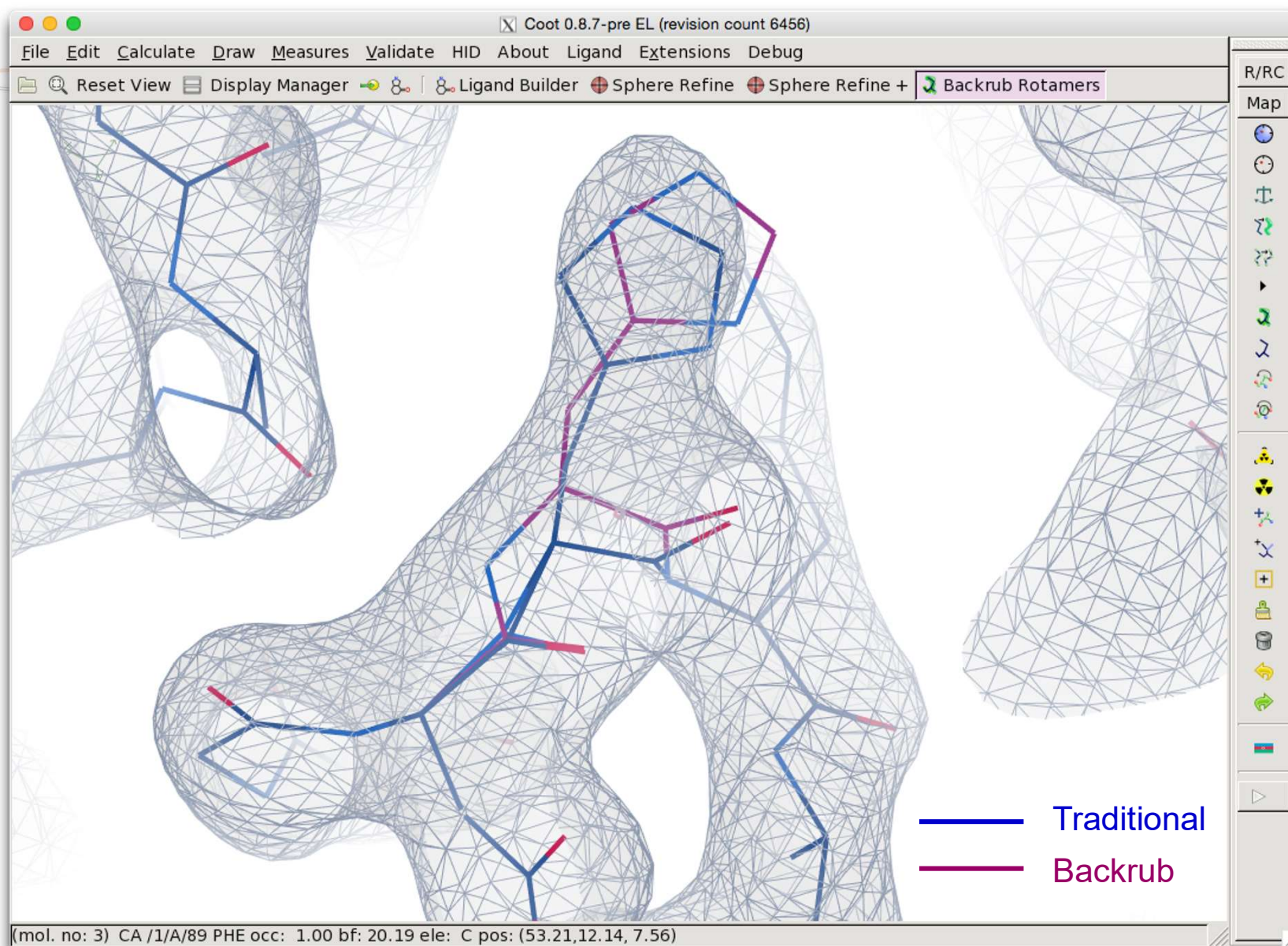










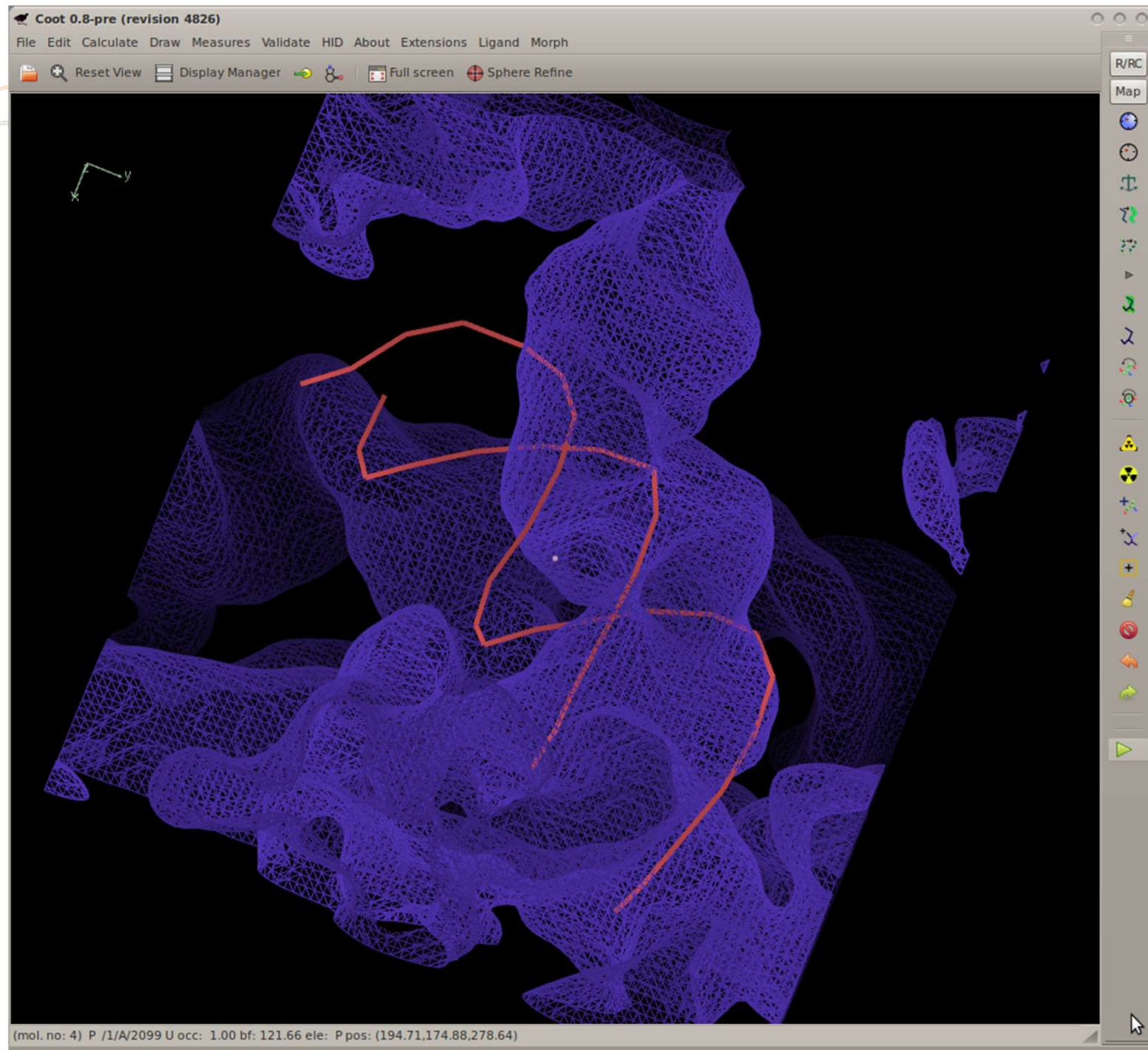


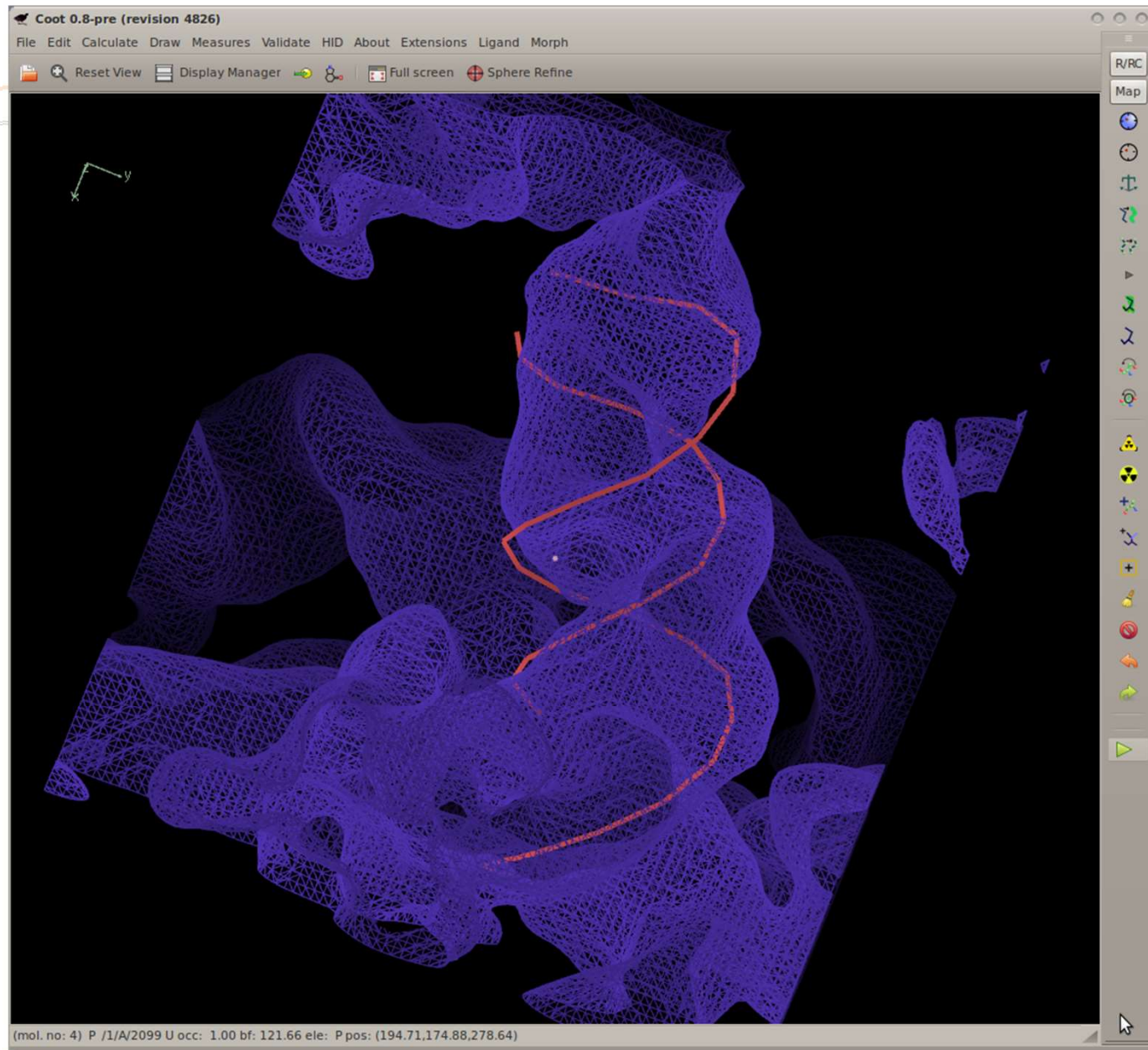
Jiggle Fit

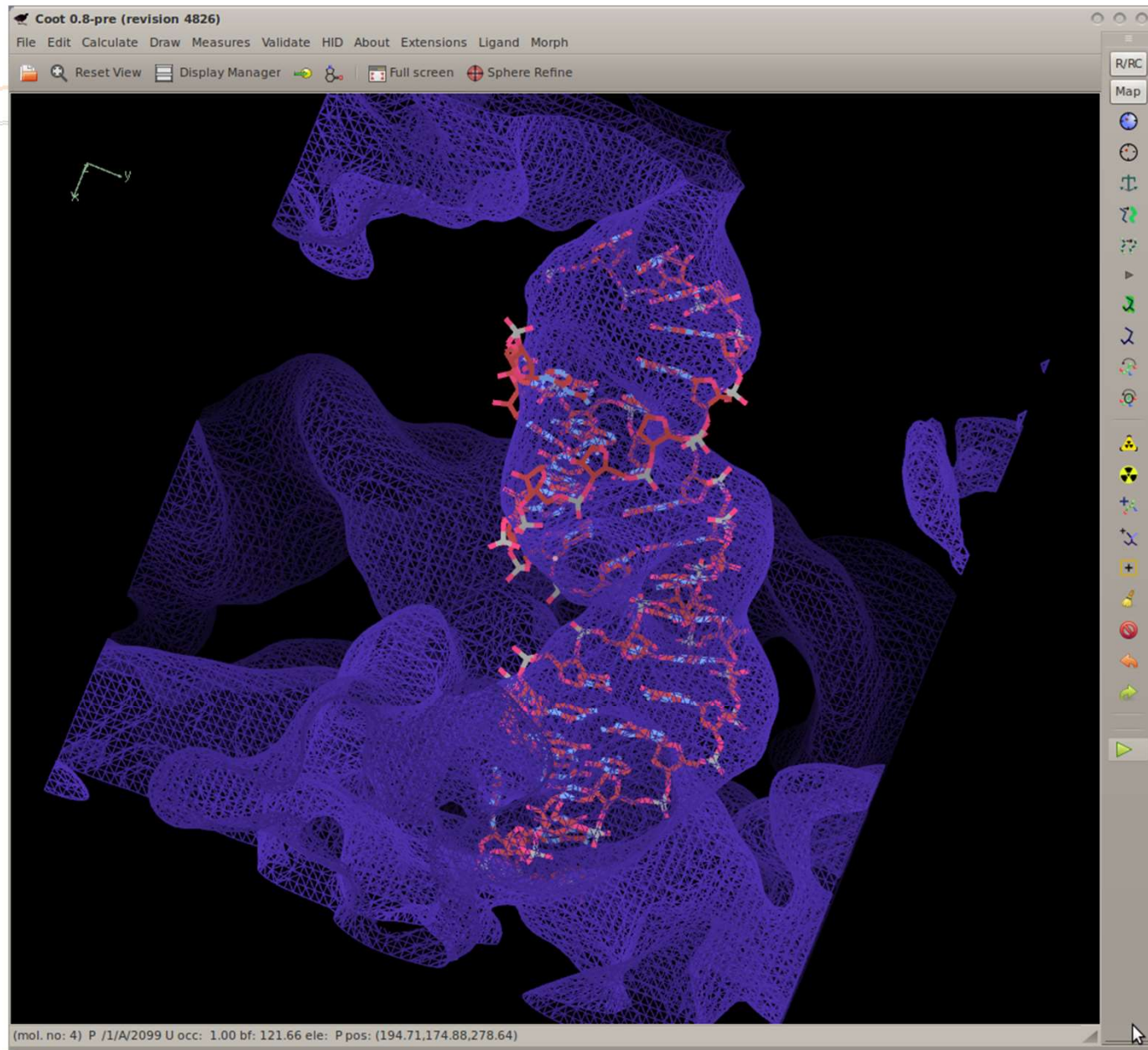
- How do I rotate and translate these atoms to fit the density?
6-dimensional problem
- Originally used to fit simple ligands/solvent molecules to blobs of density (with more or less success)
- Now extended to fit arbitrary atom selections
e.g. by Chain

Jiggle Fit

- Loop 1000 times:
 - Generate random angles and translations
 - Transform atom selection by these rotations and translation
 - Score and store the fit to density
 - Rank density fit scores,
 - Pick top 20 solution, for each of them
 - Rigid body fit and score solutions
 - Pick the highest scoring solution if it's better than the starting model)
- Radius of Convergence is larger when using a low-pass map

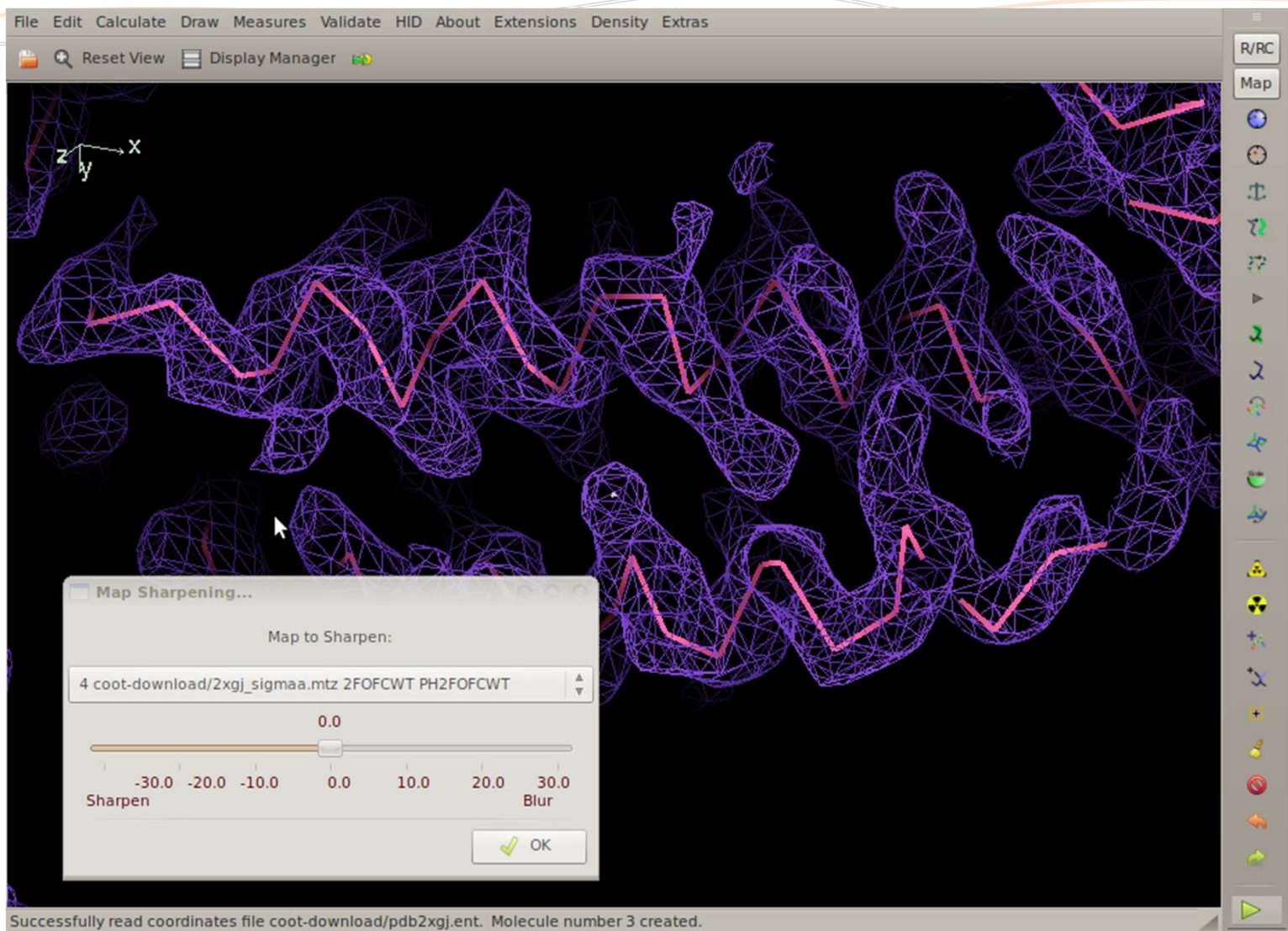


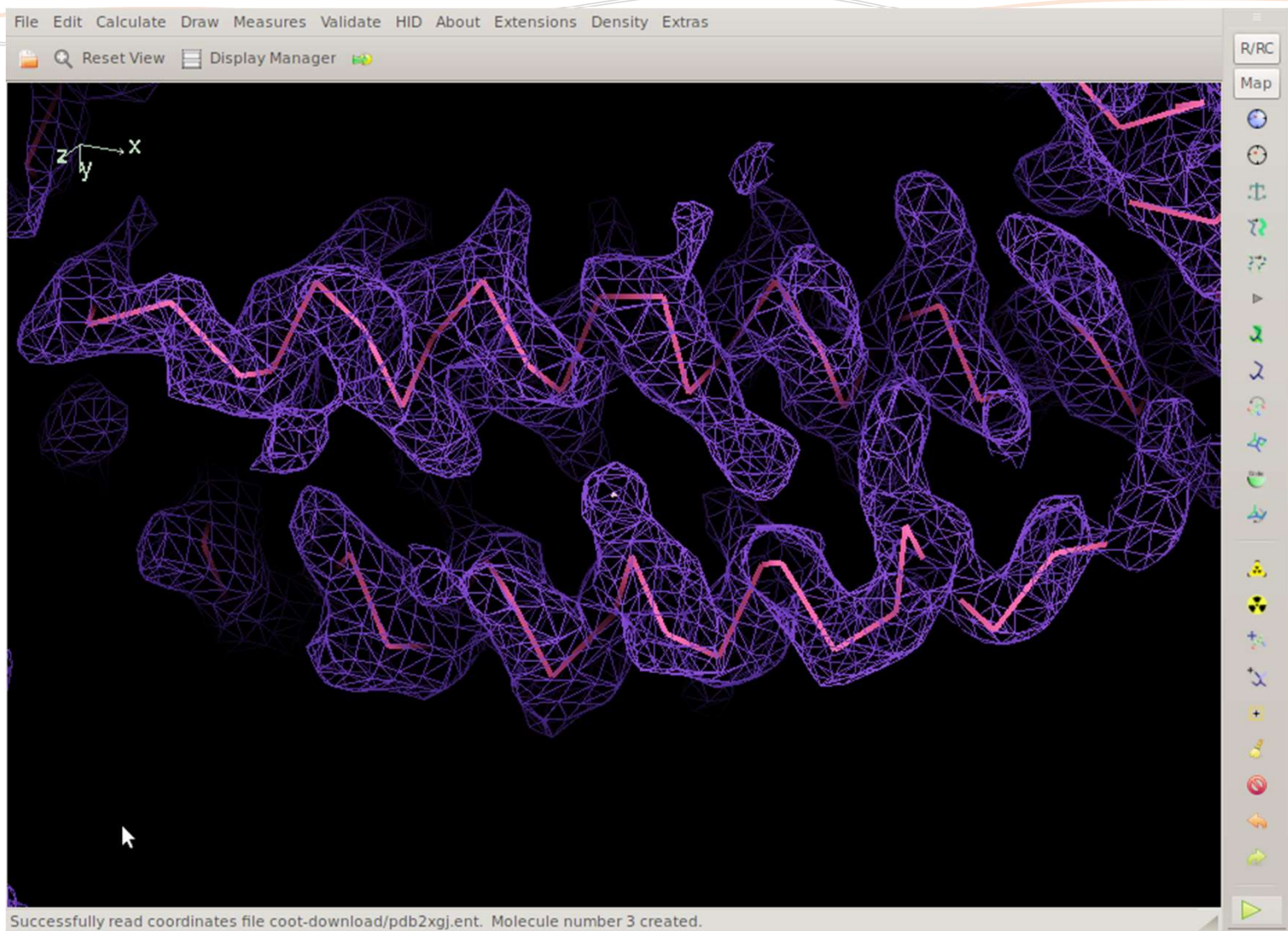


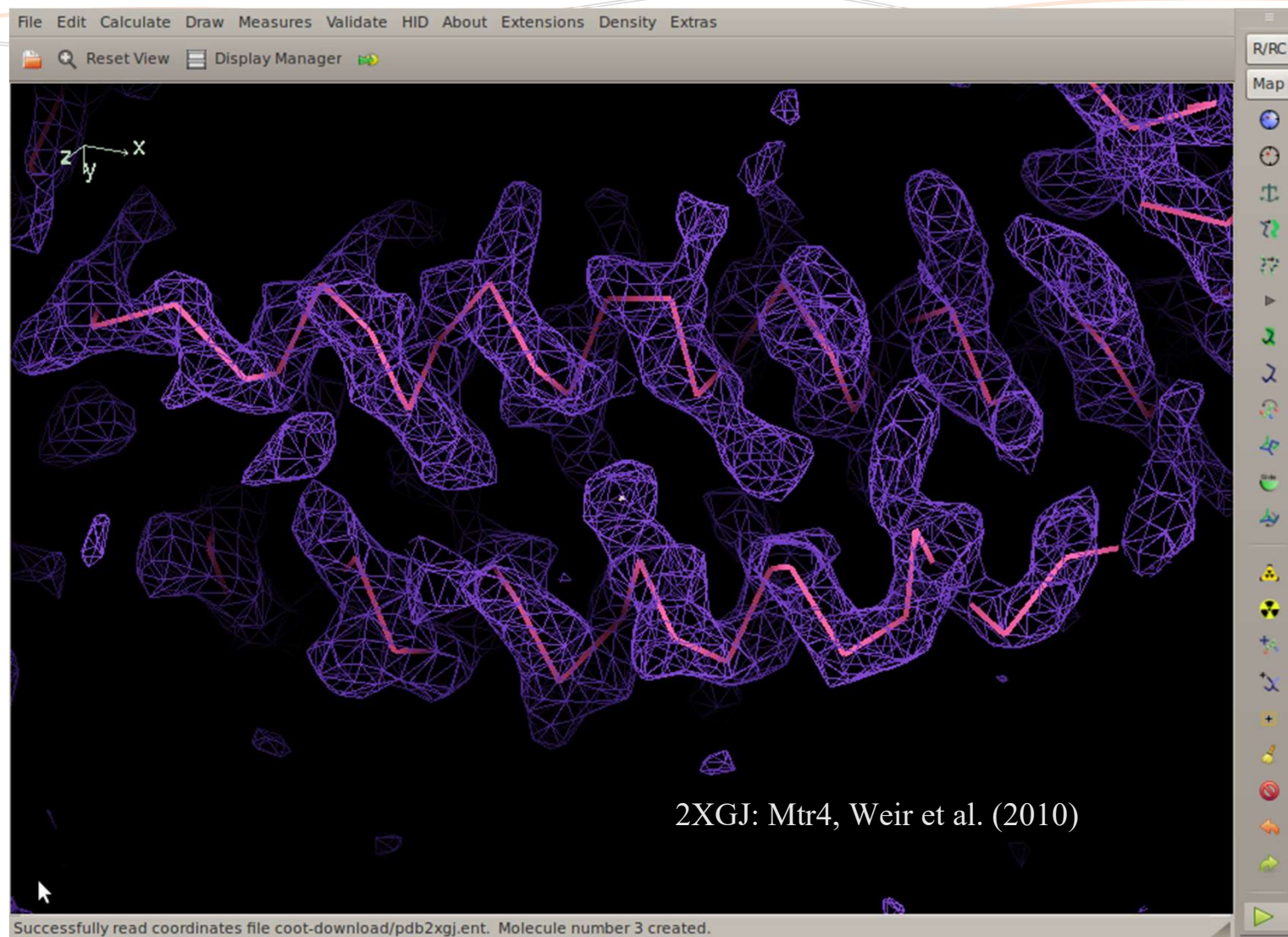


Map Sharpening

- Fine tuning the B-factor: which B-factor shall I use to get the most interpretable map?
- Interactively adjust the structure factor amplitudes and re-generate the map with FFT and recontouring...







More *Coot* tools

- Rotamer search
- Chi angle editing
- Alternate Conformations
- Add terminal residue
- Ligand fitting/search
- Rigid-body Fitting
- Steepest Descent
- Simplex (slower but better)
- “Move Molecule Here”
- Water Search
- Fill-partial-residues (post-MR)
- “All model” tools (post-MR) eg stepped refine
- Additional representations:
dots, ball&stick, CA trace etc representation
- ... and many more

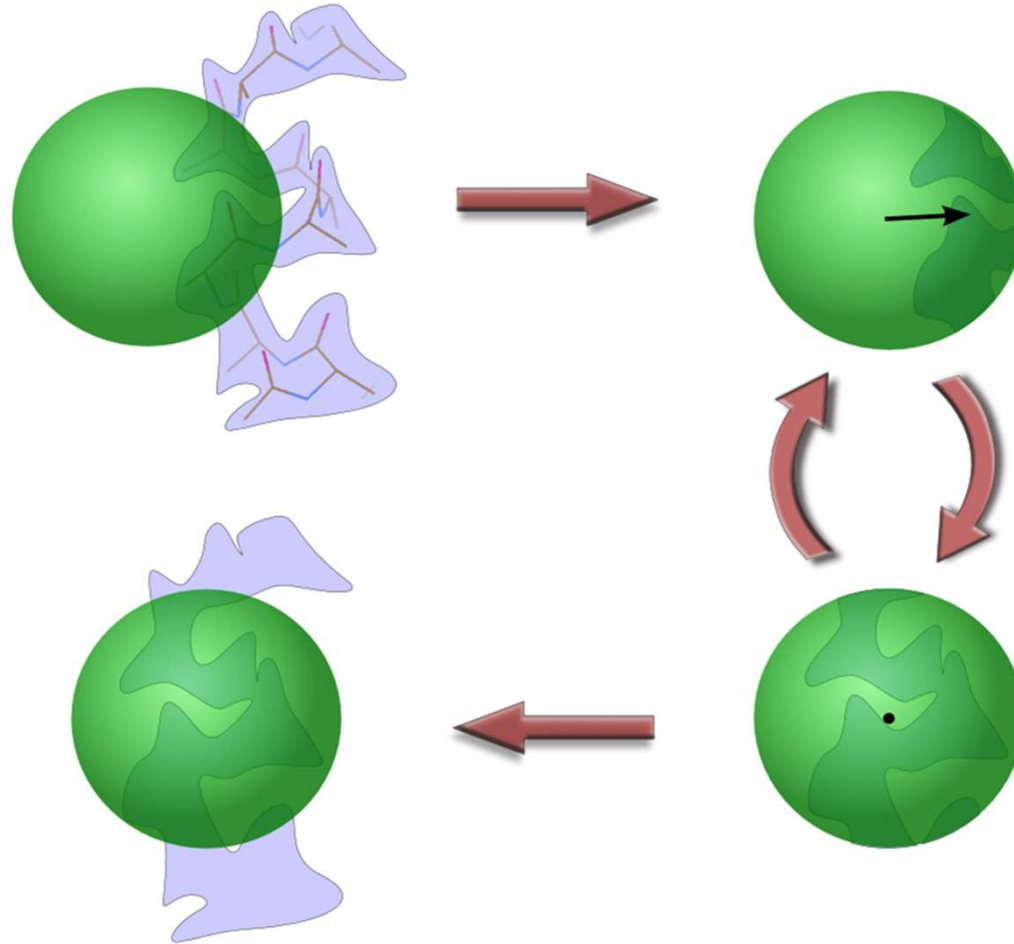


Building from scratch

Alpha helix placement

- Scenario: Looking at a new map, not built with automatic tools:
“I can see that there’s a helix here - build it for me!”
- From a given point:
 - Move to local averaged maximum
 - Do a 2D MR-style orientation search on a cylinder of electron density
 - Build a helix (both directions)
 - 1D Rotation search to find best fit
 - Score based on density at CB positions
 - Trim ‘n Grow

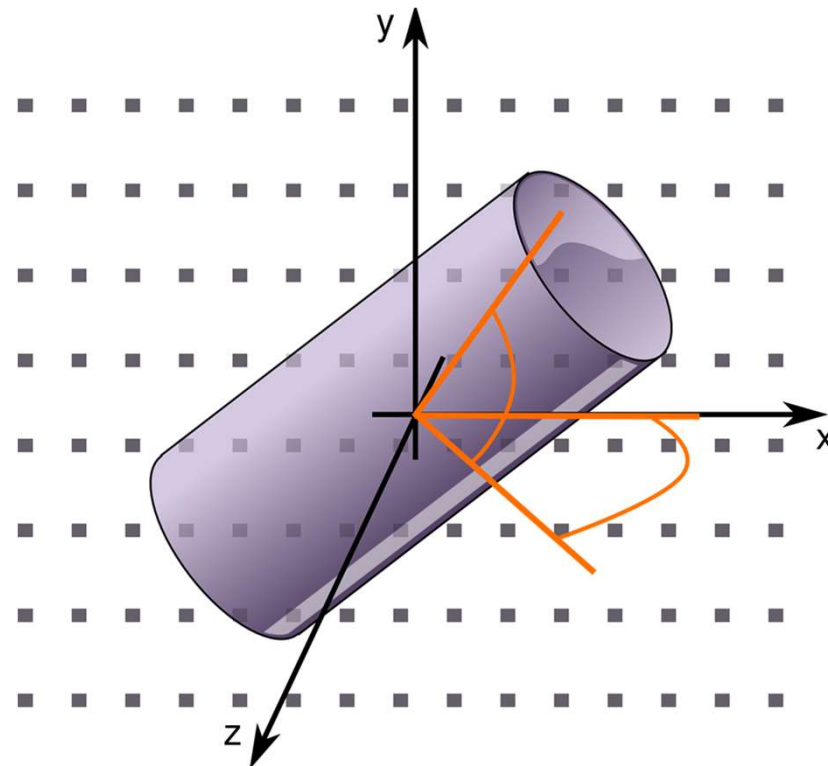
Centering the point of rotation

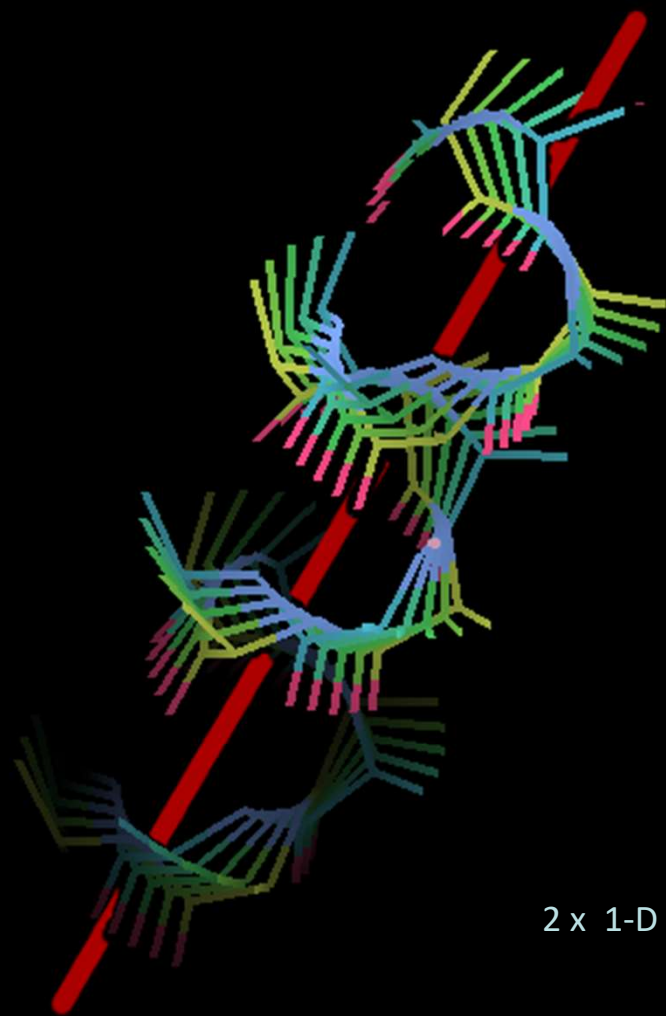


Cylinder search

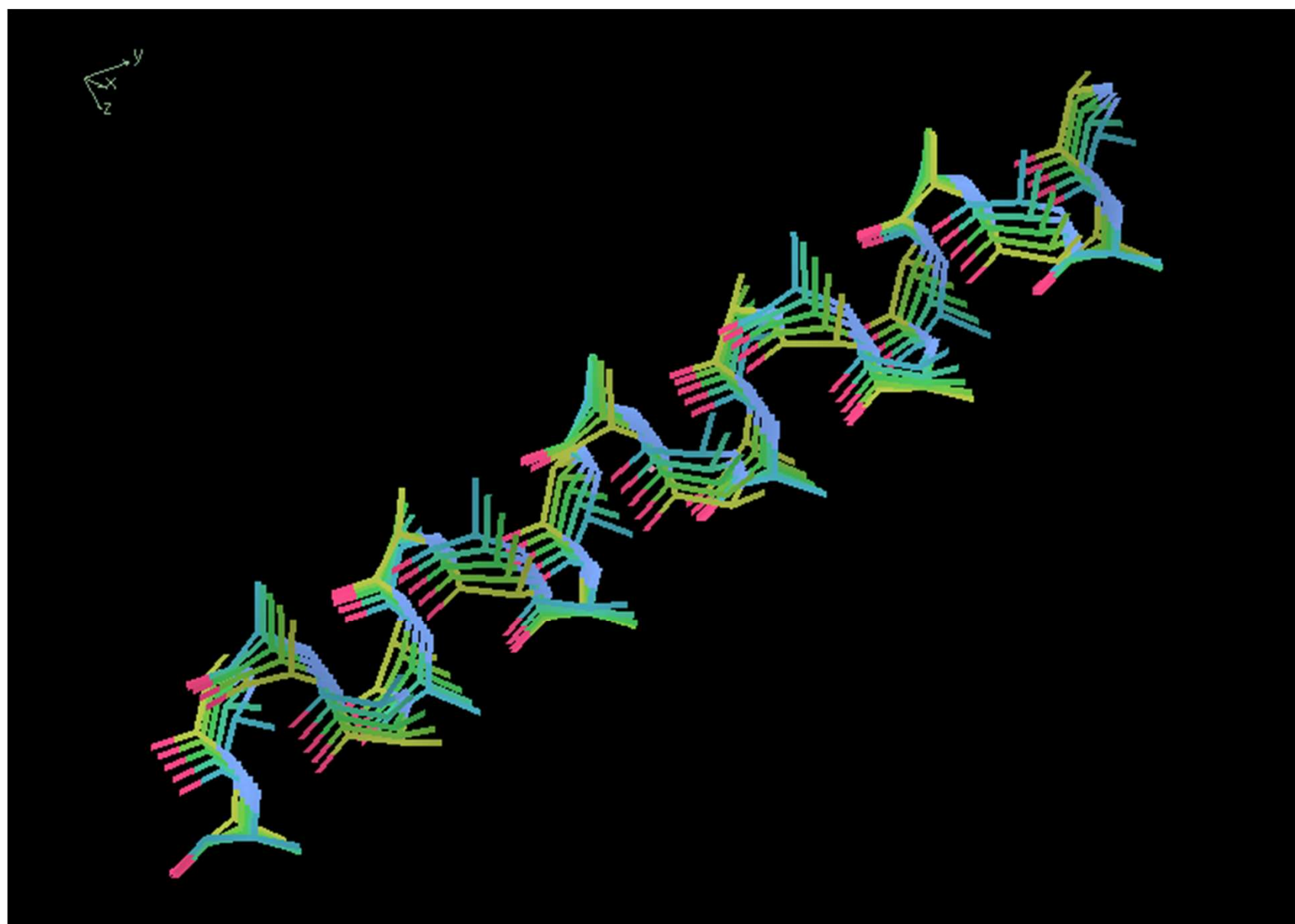
Pick the orientation that encapsulates the most electron density

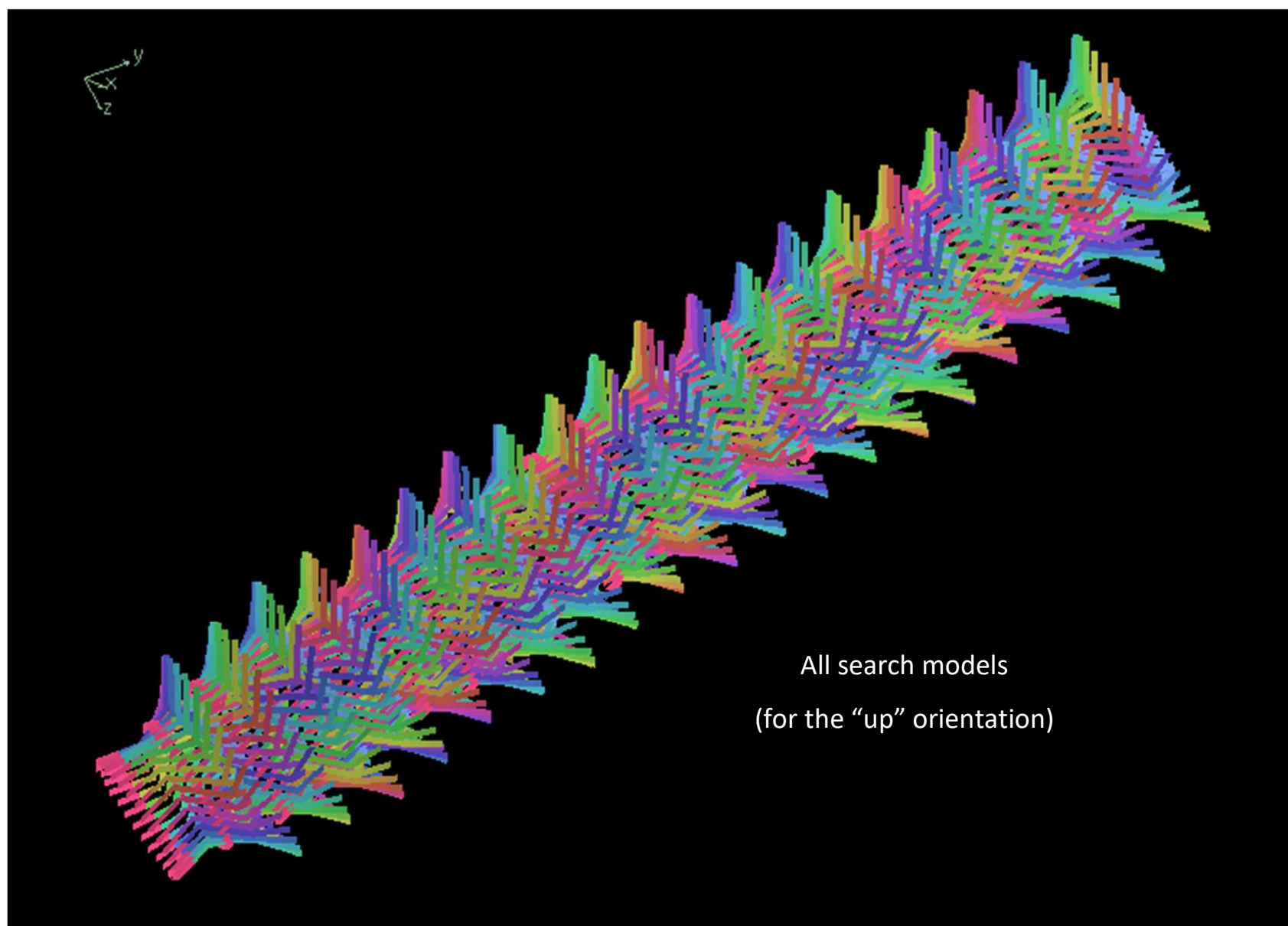
Using 2 rotation axes



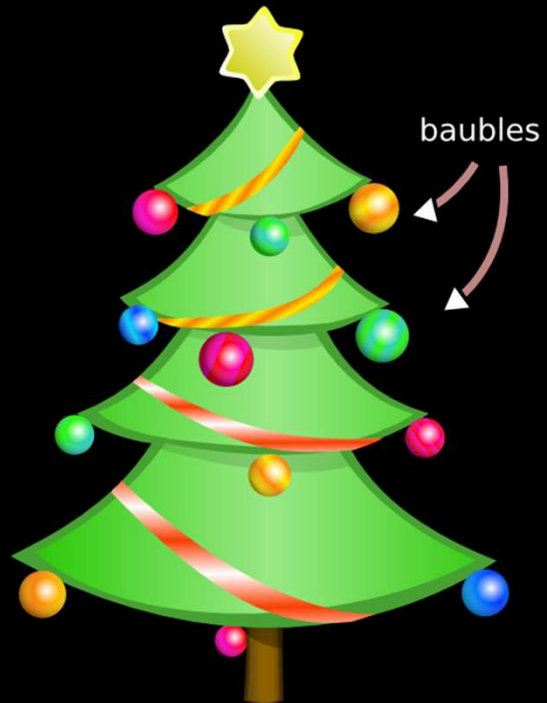


2 x 1-D Helix orientation searches



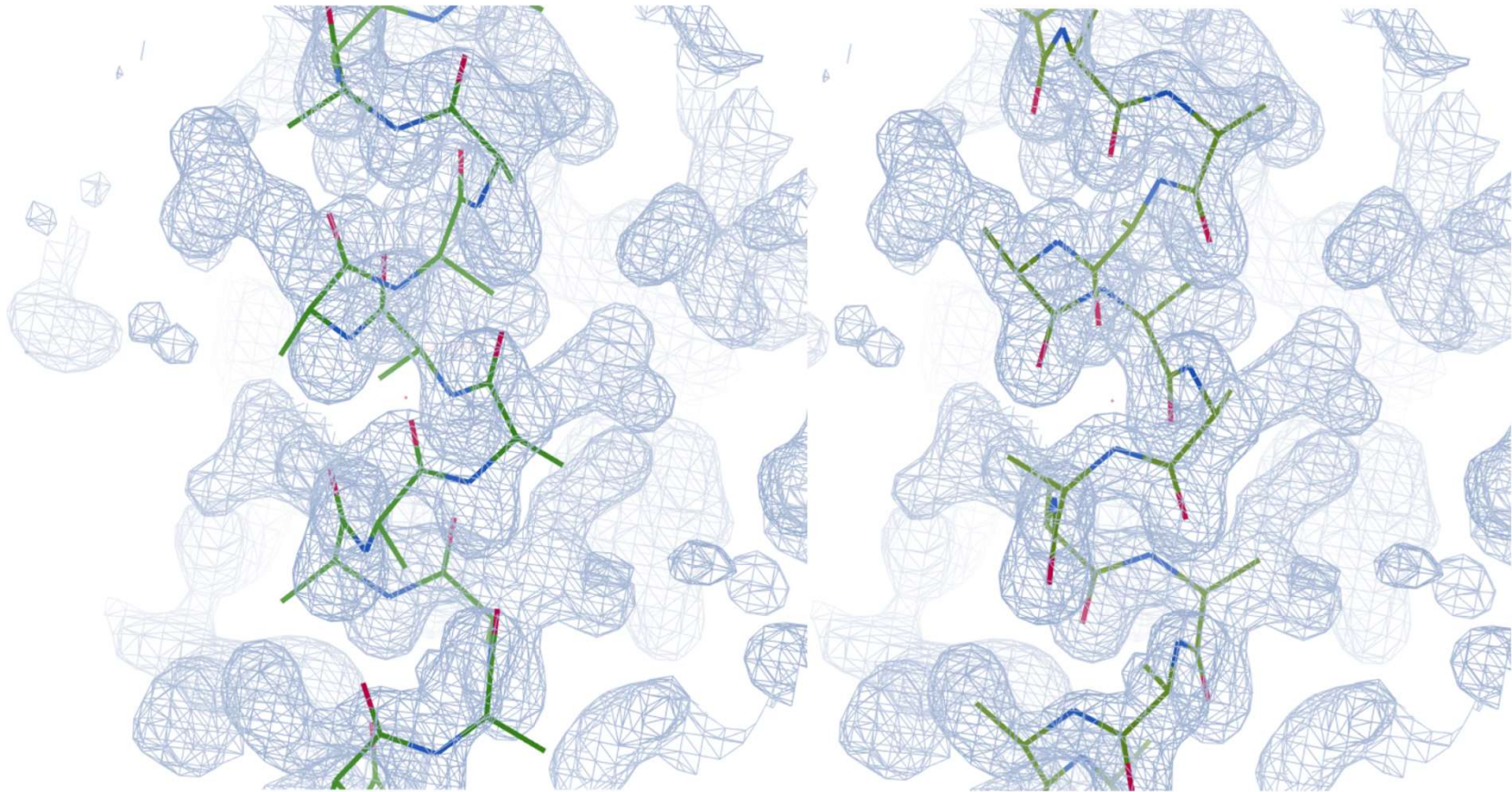


Top

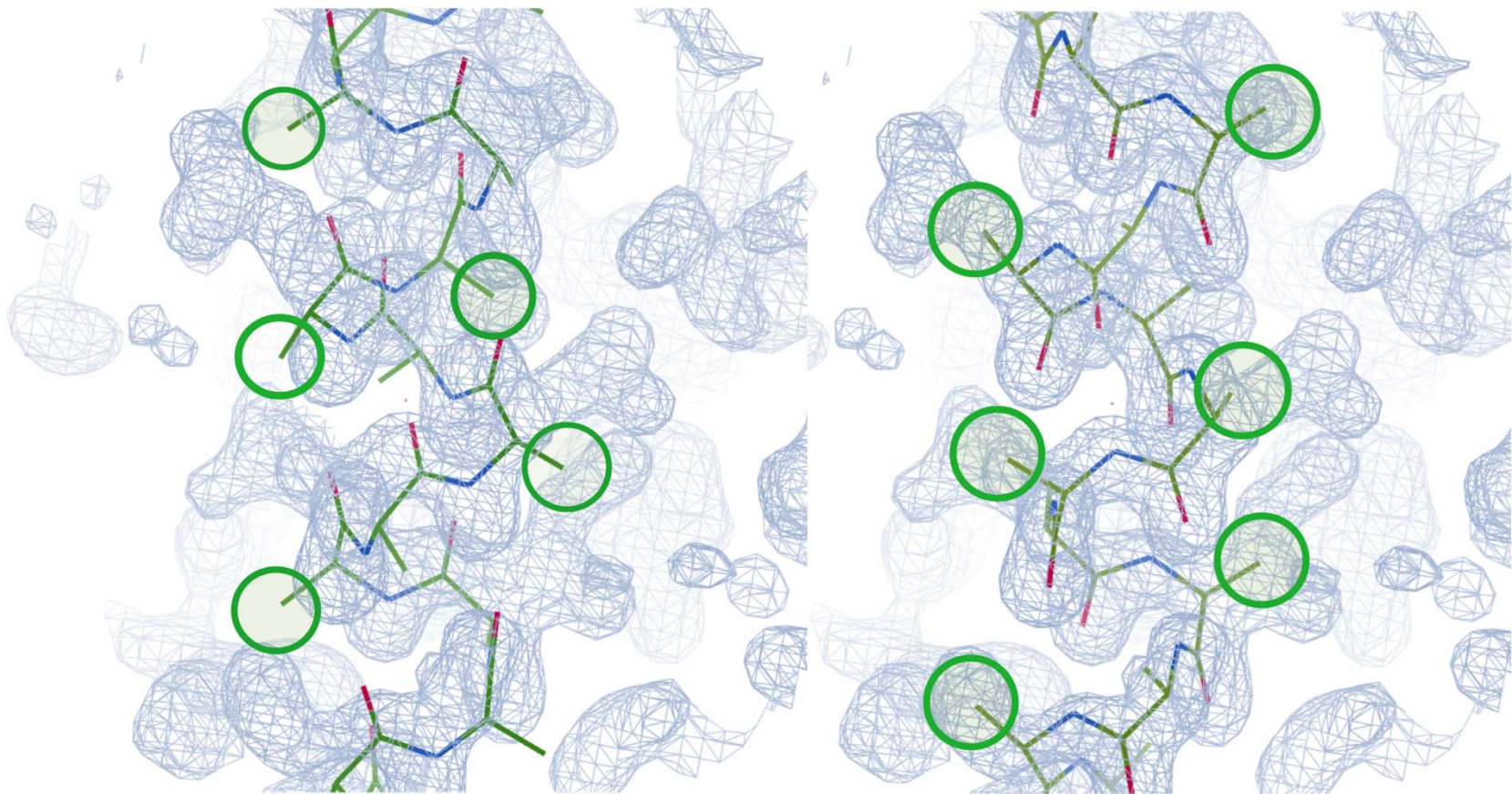


Bottom

Helix fitting in two orientations

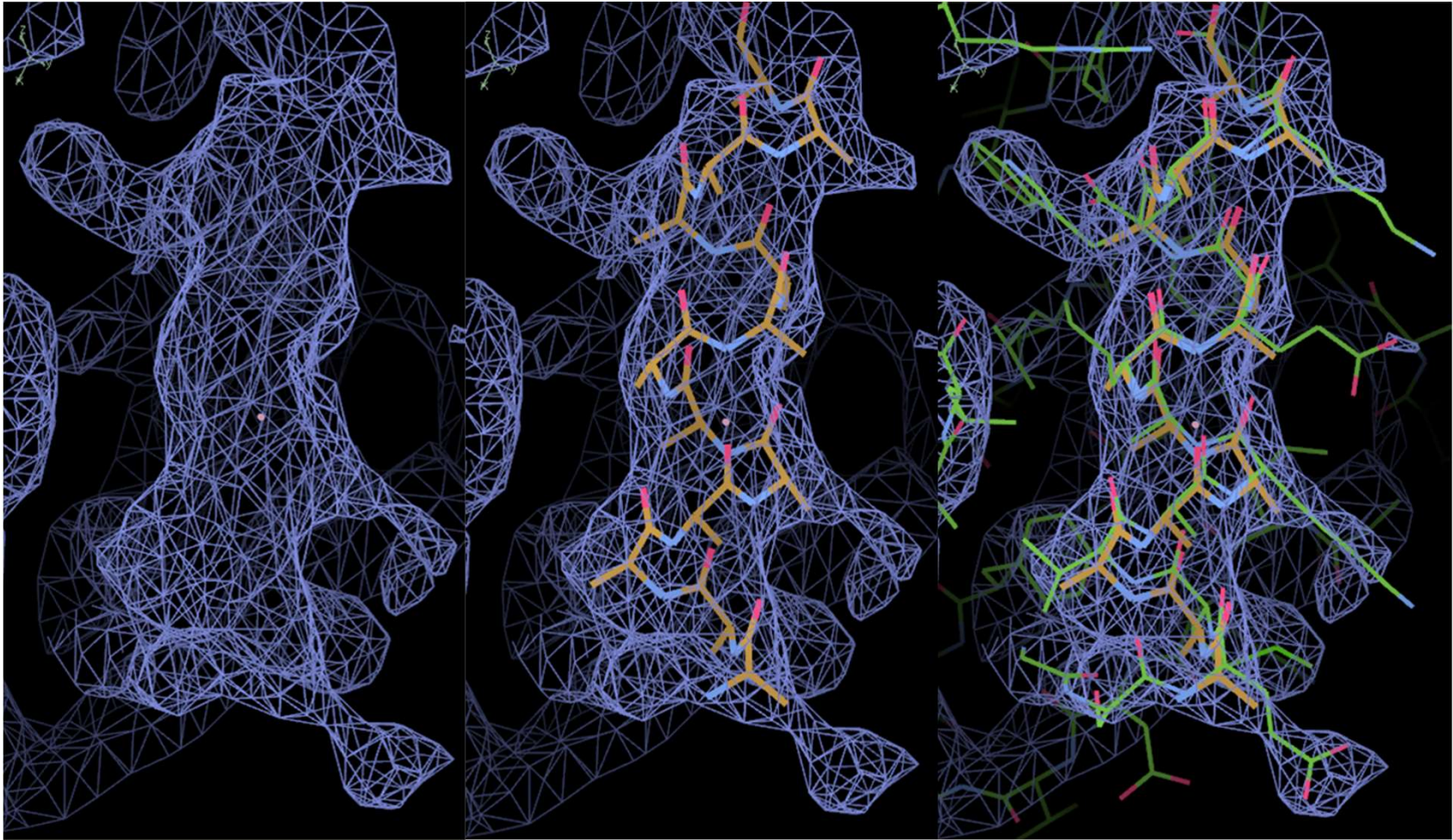


Scoring



c-betas are not fitted and are used for scoring

Helix fitting in action



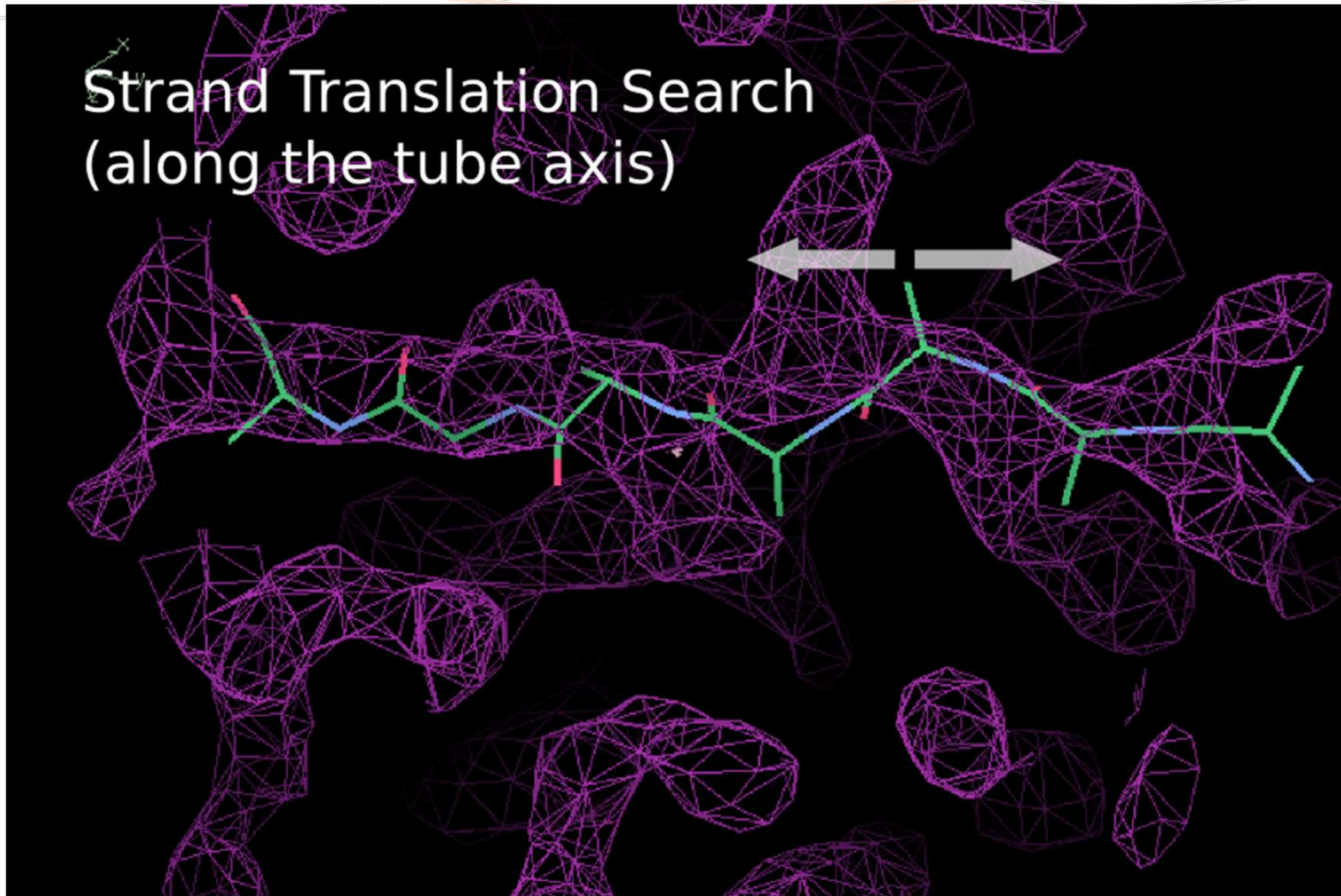
Placing strands

- Unlike Helices, Strands have to be treated as non-idealized
- Repeating a single phi/psi value doesn't make a structure that fits “real-world” density
- Curvature of strands should be taken into account
- Use selections from a “database” of good structures

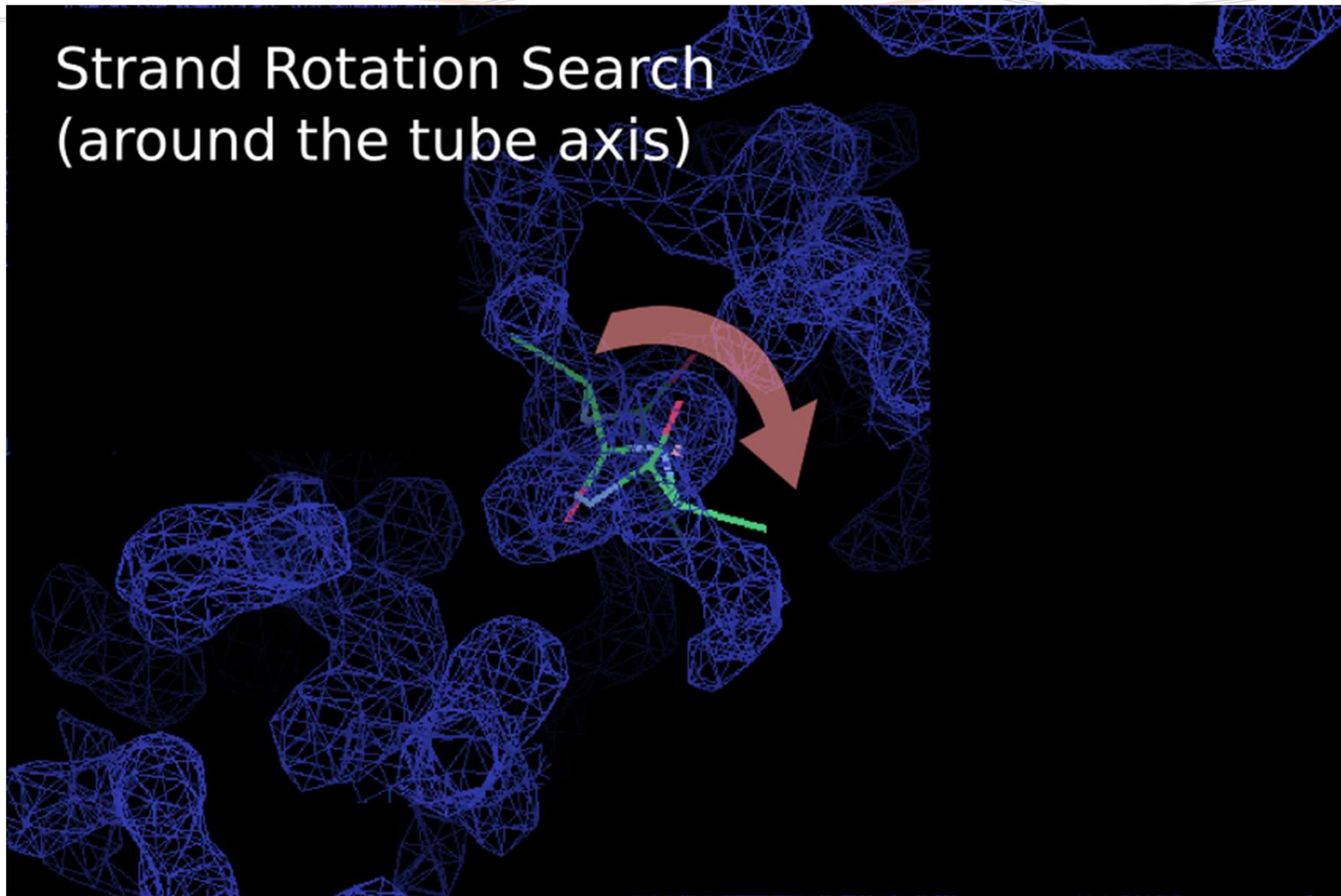
Strand fitting algorithm

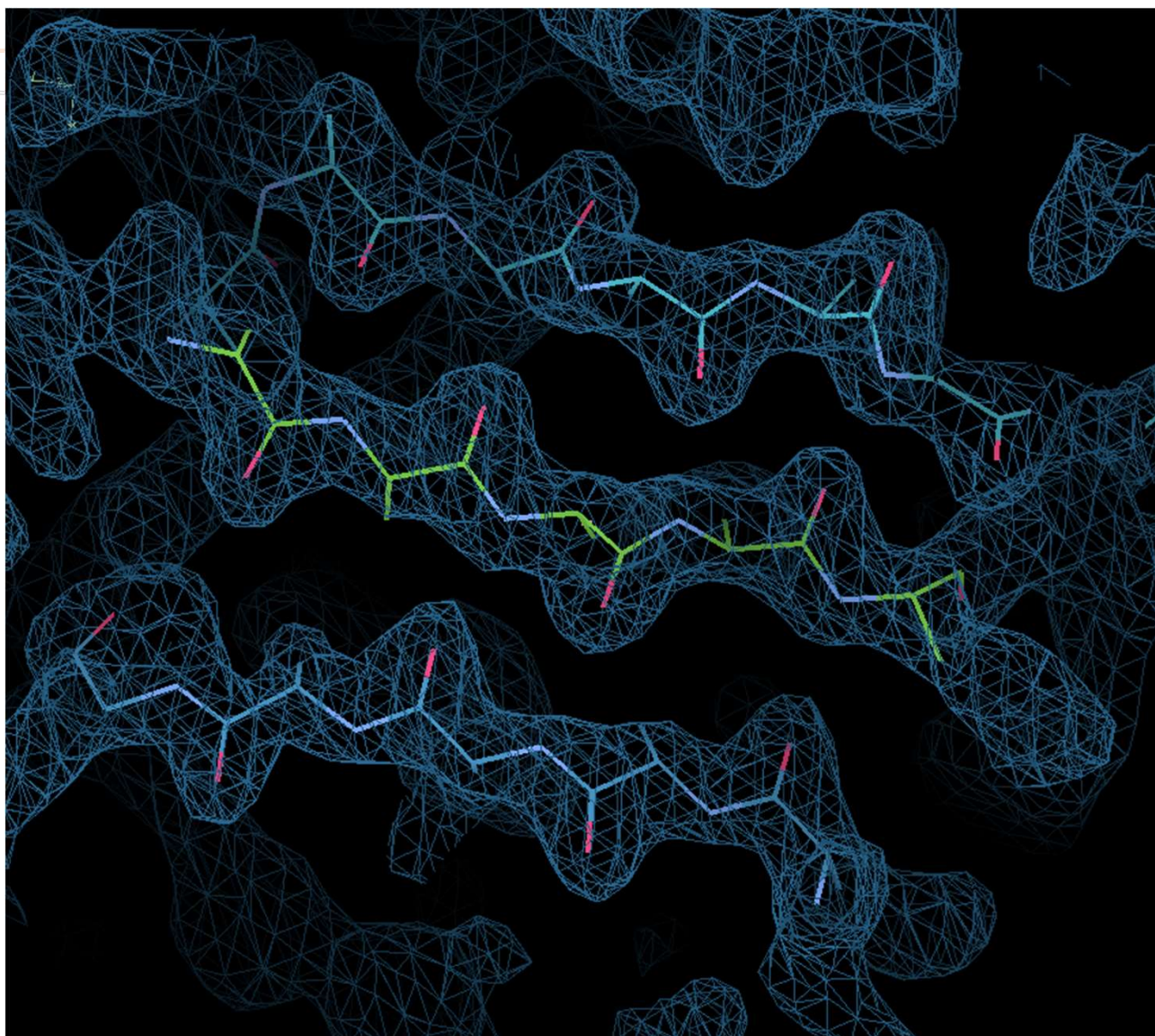
1. Cylinder search
2. Get N fragments of length l from database
3. 1-D Translation search along the tube
4. 1-D Rotation search around the tube
5. Direction flip search
6. Rigid body refine best solutions
7. Real-space refine best solution

Strand Translation Search (along the tube axis)

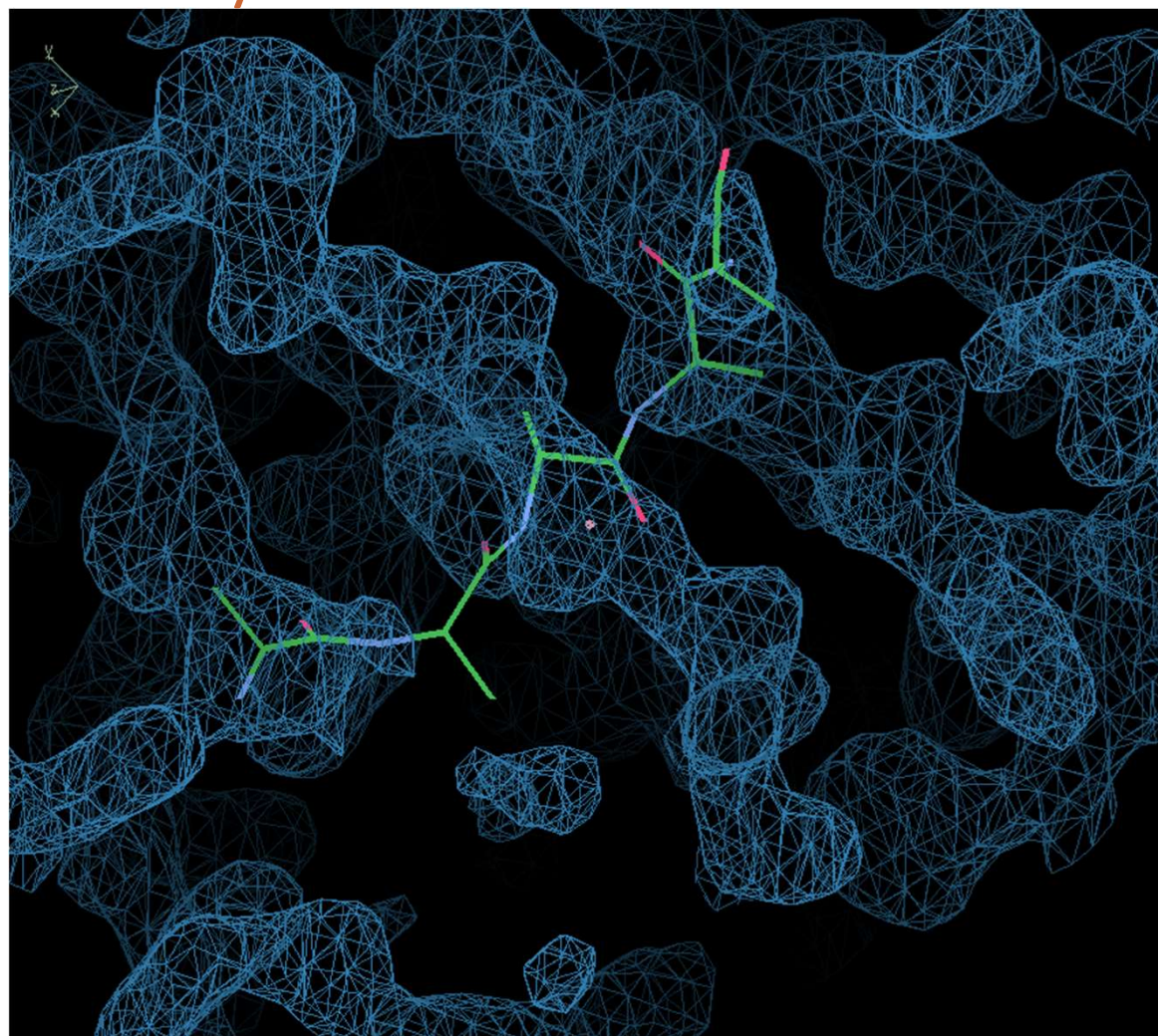


Strand Rotation Search (around the tube axis)





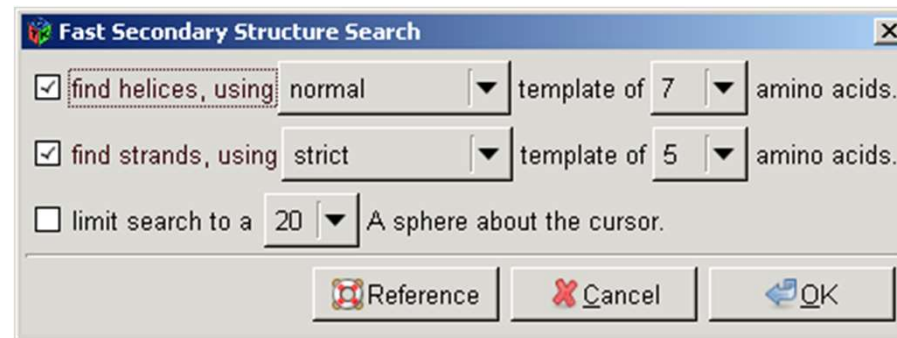
Not all is rosey...



Strand fitting caveat

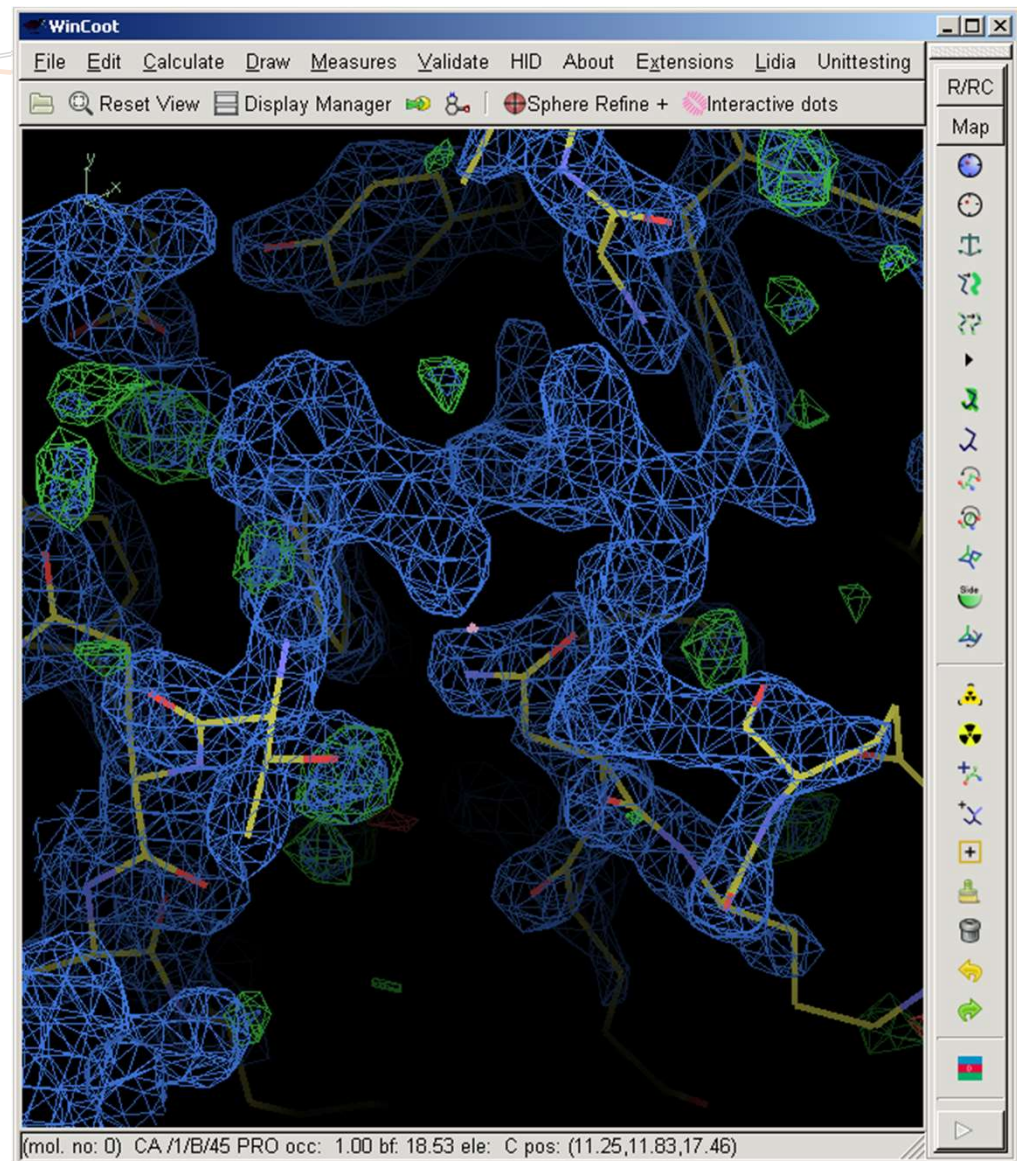
- In the case of strand-fitting, the initial translation search centring the cylinder is not performed (the search cylinder is too thin)
- The user is responsible for centring the search point “in the middle of the tube”
- Not at a C-alpha position

Automated secondary structure search



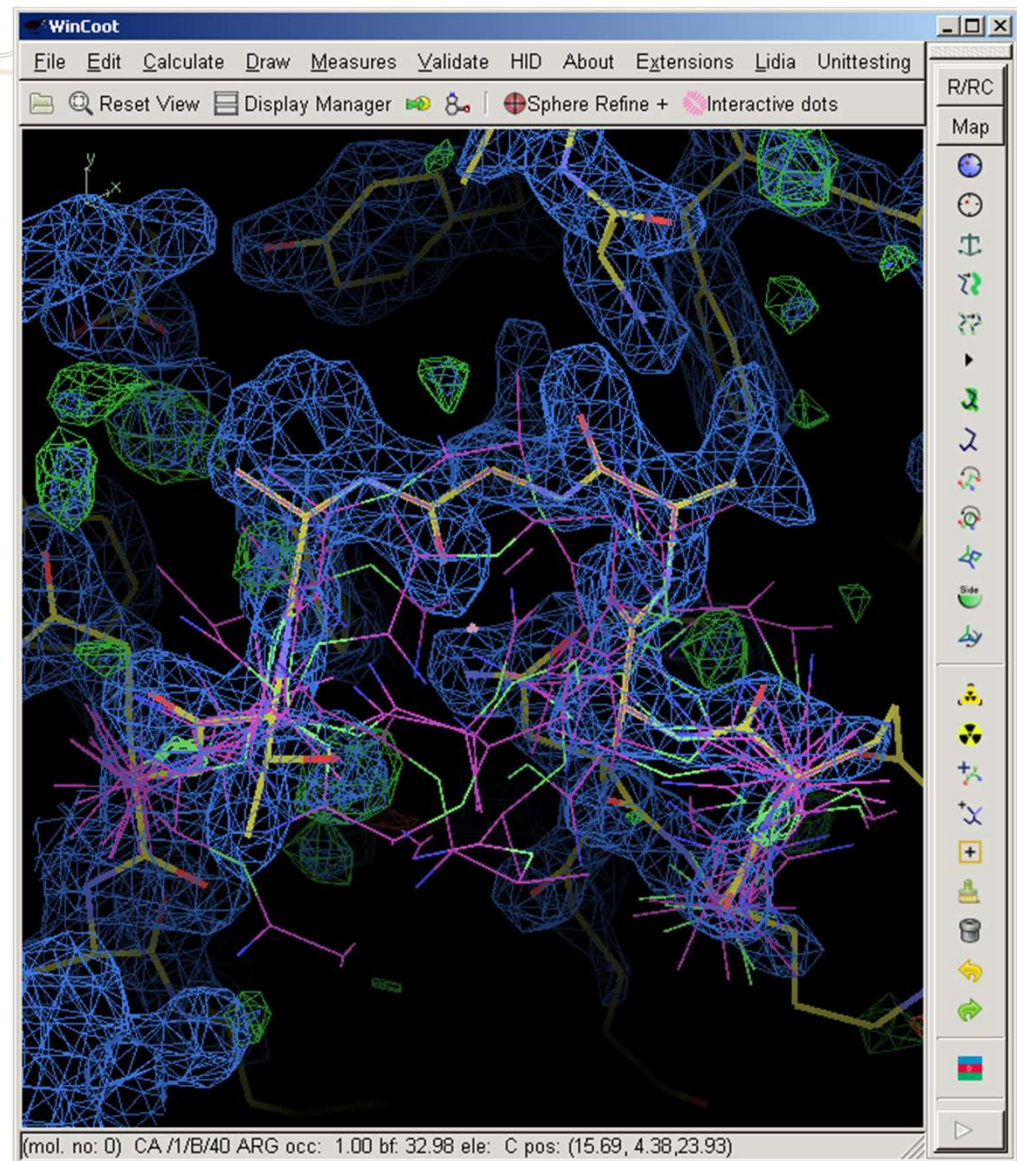
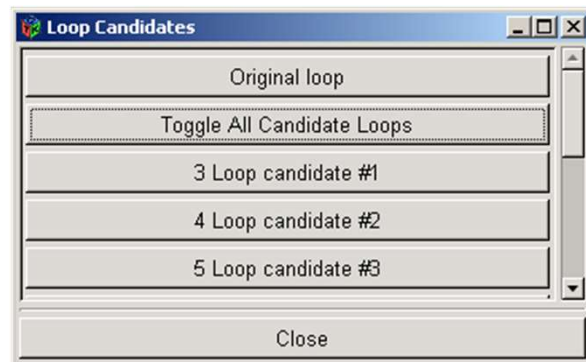
Loop fitting

- Simple loop fitting
Add residue by residue
(from both termini)
- DB loop
Fitting fragments from
database



Loop fitting

- Simple loop fitting
Add residue by residue
(from both termini)
- DB loop
Fitting fragments from
database

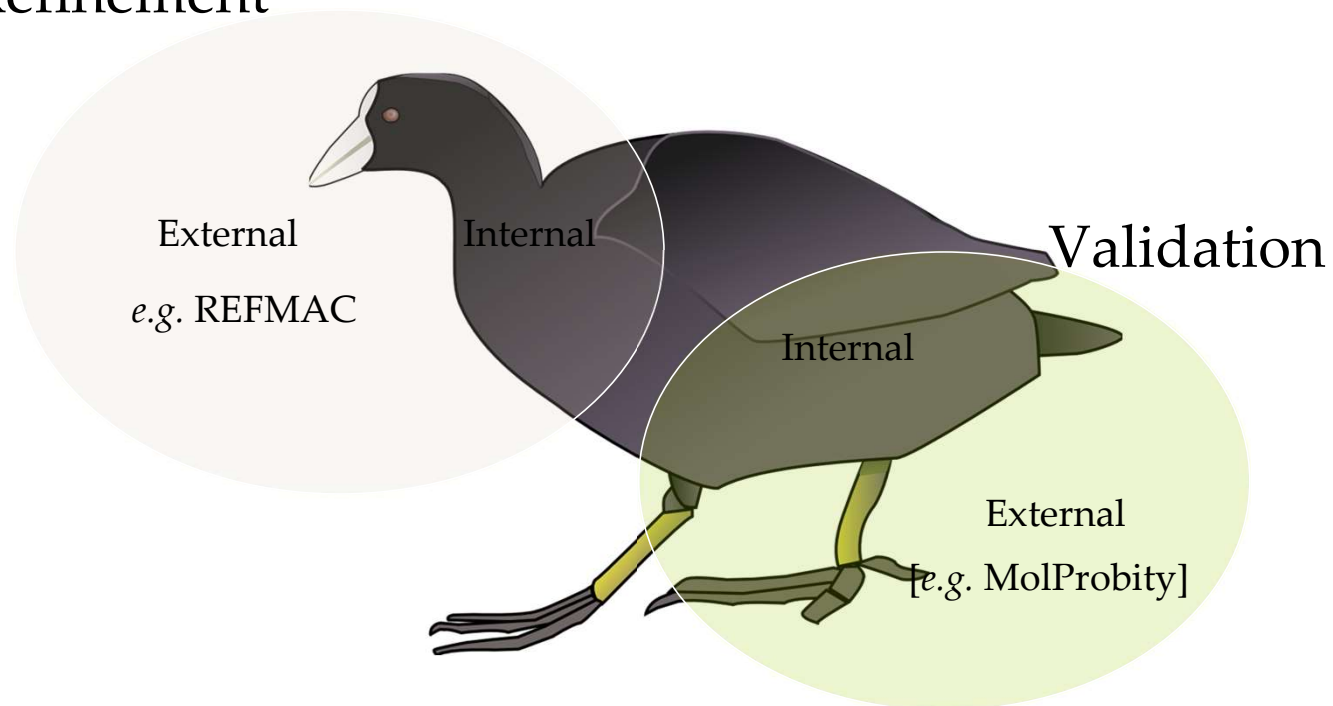


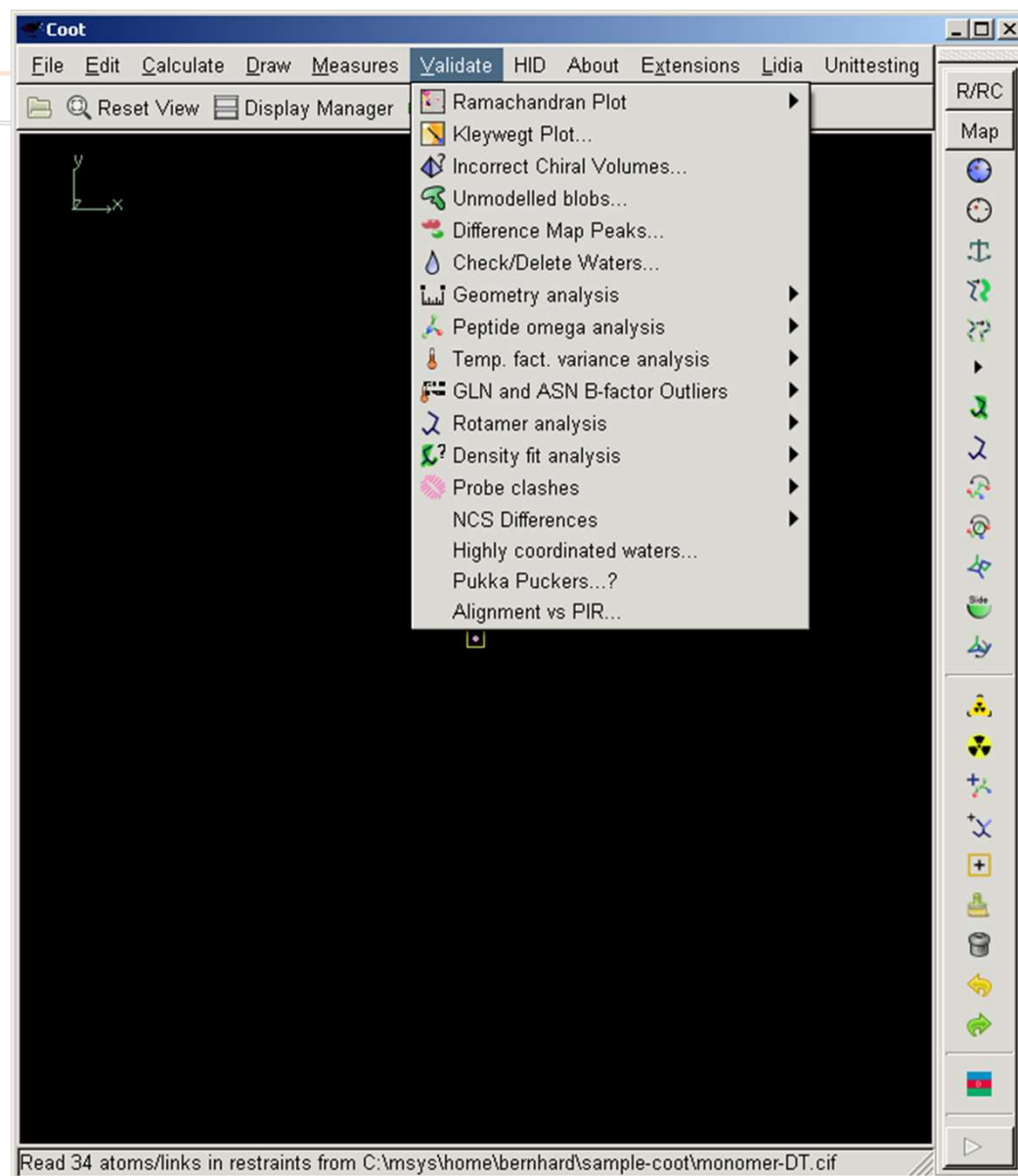


Validation

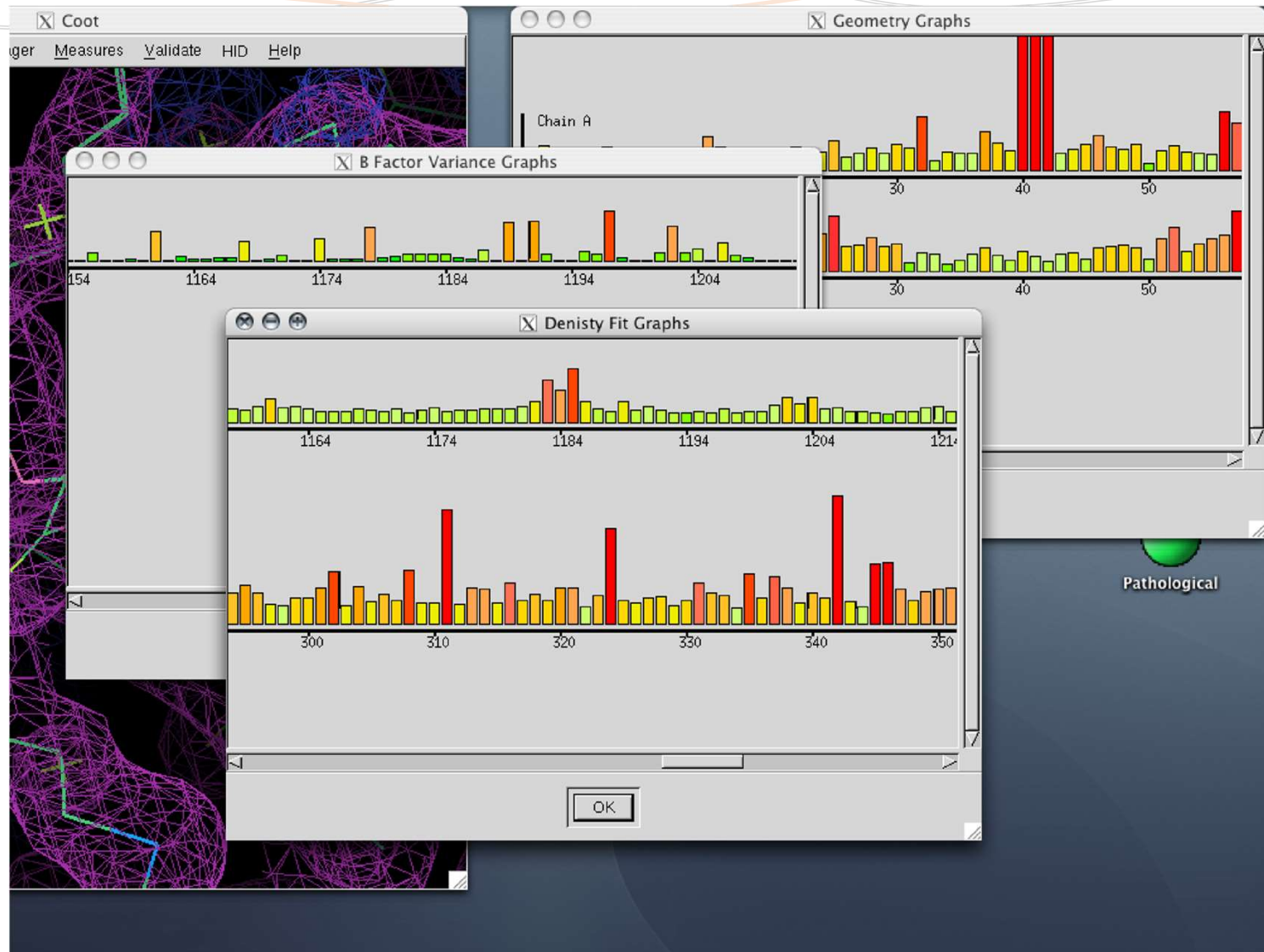
Feature integration

Refinement

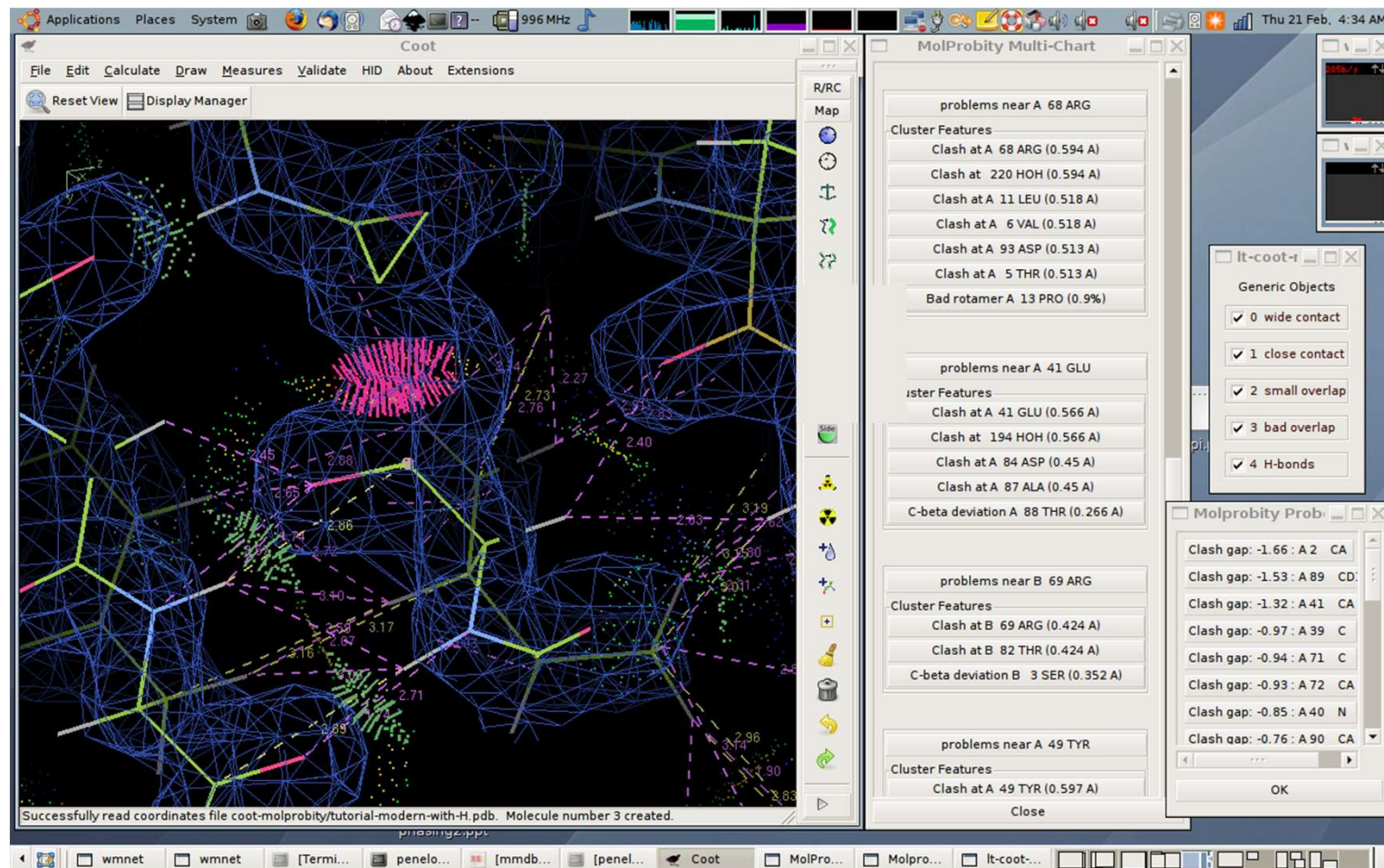




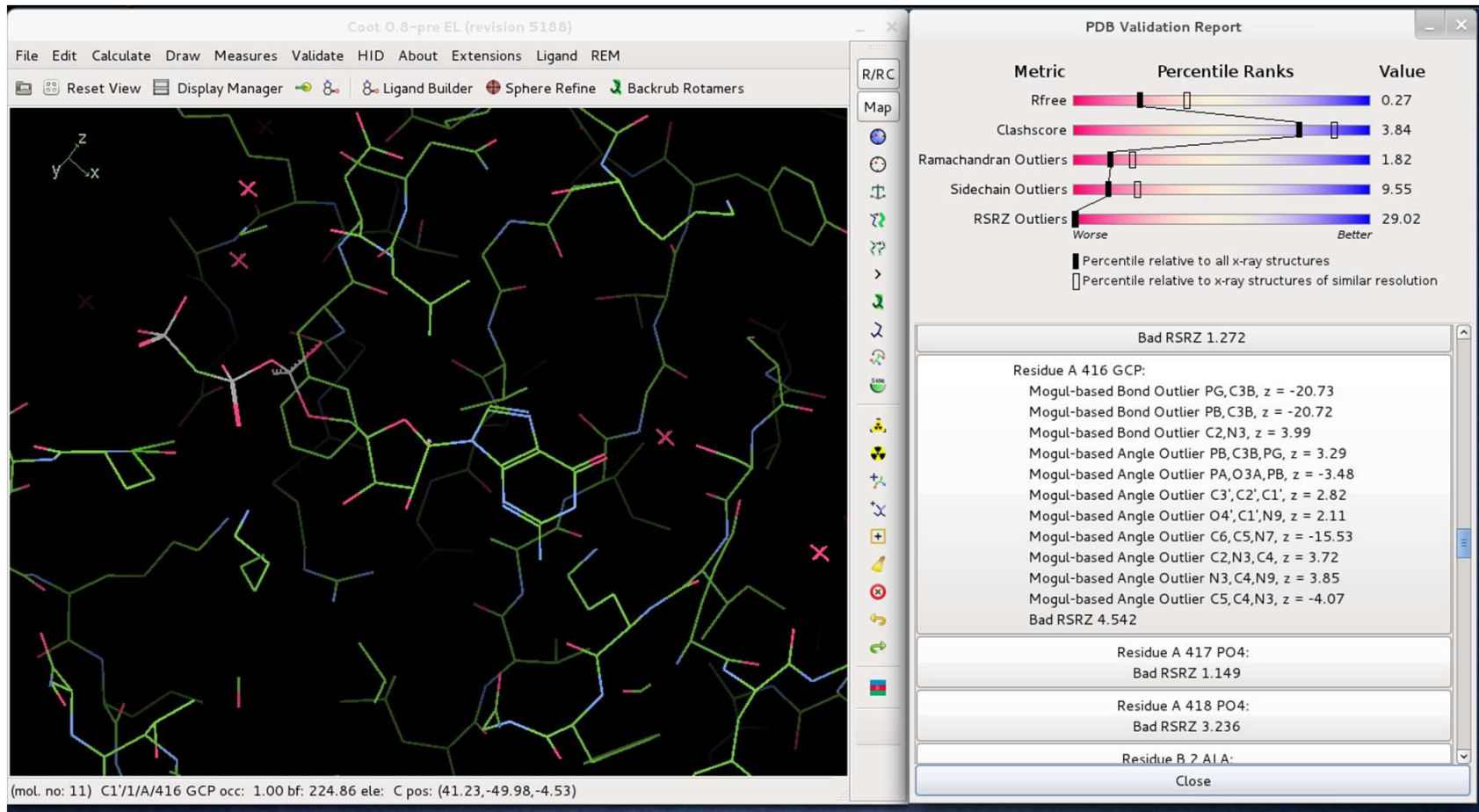
Charts



Molprobity integration



PDB validation report parsing





NCS: non-crystallographic symmetry

NCS: non-crystallographic symmetry

- 2 or more copies of a molecule in the unit cell not related by crystallographic symmetry
- Crystallographic copies of molecules are (of course) treated as if they were exactly the same across the unit cell – and indeed across the whole crystal
- Non-crystallographically related molecules provide different representations of the same molecule
 - This can be useful for model-building
 - But difficult to use in practice

NCS: non-crystallographic symmetry

Problems?

- Strict NCS:
NCS should appear like crystallographic symmetry does
[exact copies]
- Non-Strict NCS:
Molecules are different

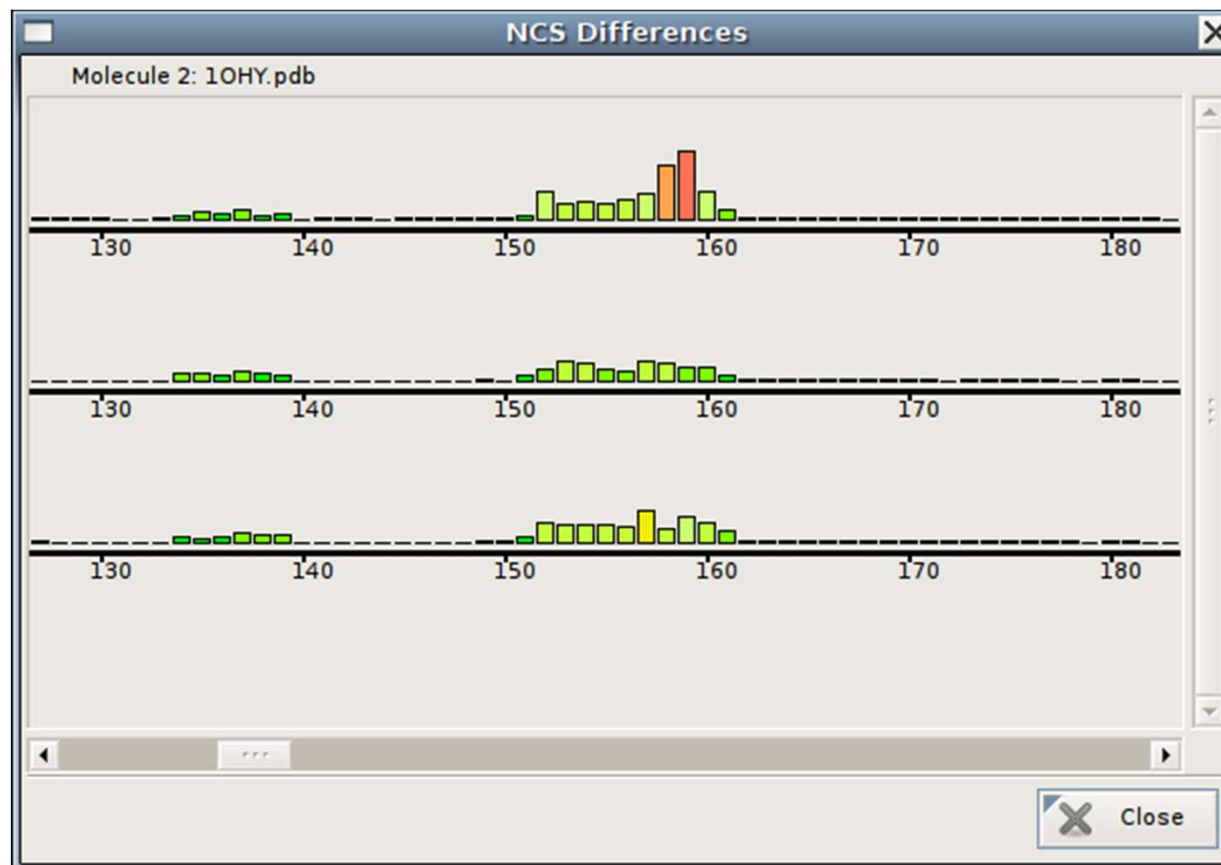
How to cope with differences, but minimize unnecessary rebuilding?

NCS: non-crystallographic symmetry

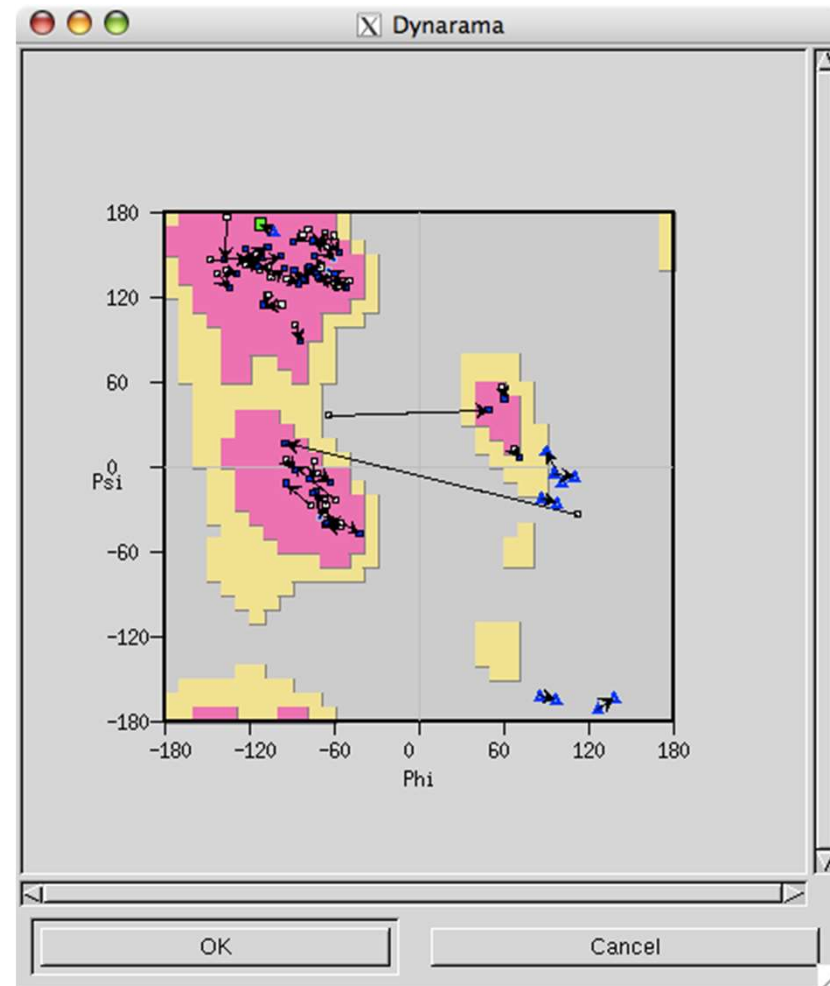
Typical Scenario:

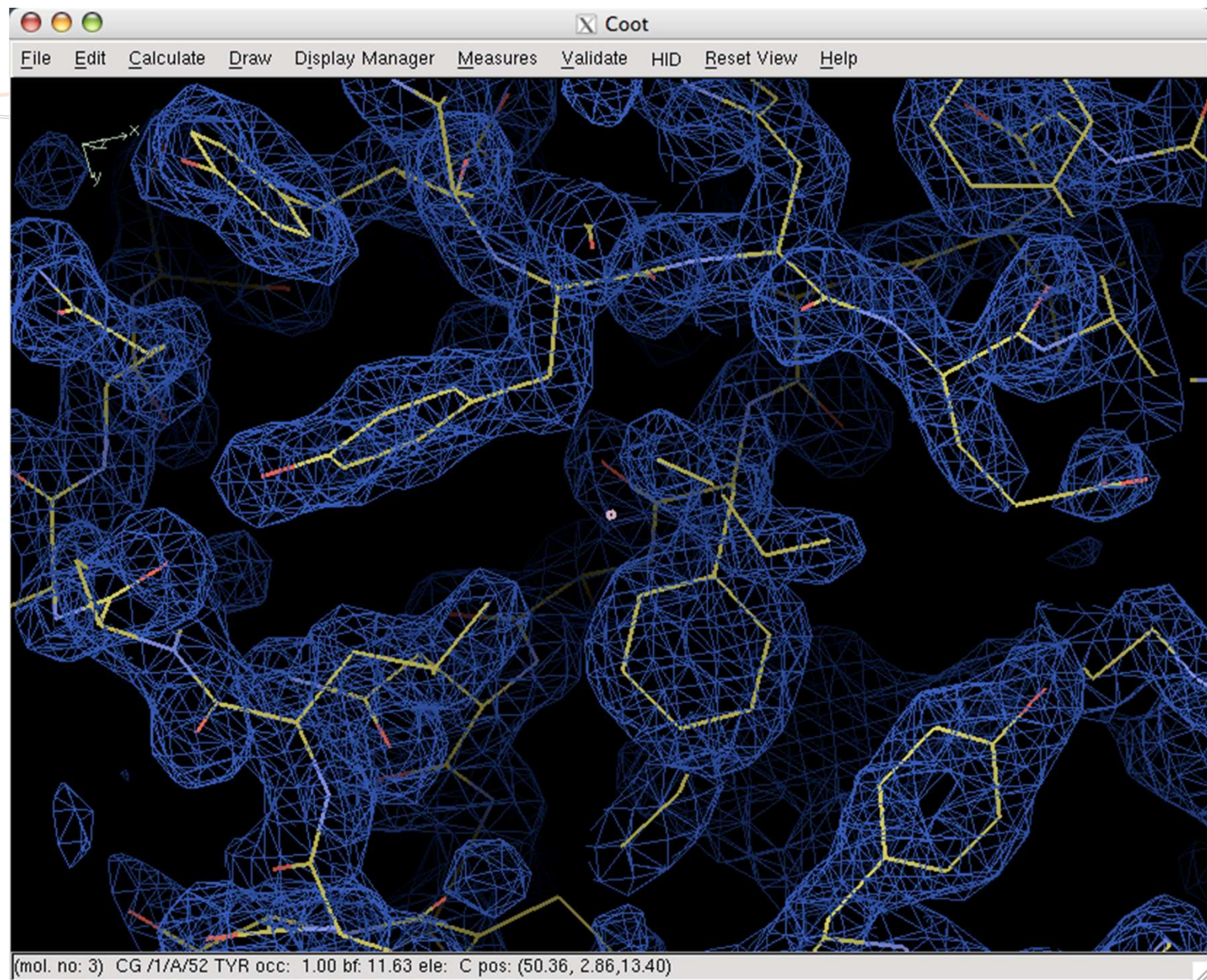
- I have done an LSQ overlap of my NCS-related molecules and from the graph, have seen significant deviations in the positions of some side-chains.
- Why are they different?

NCS differences graph

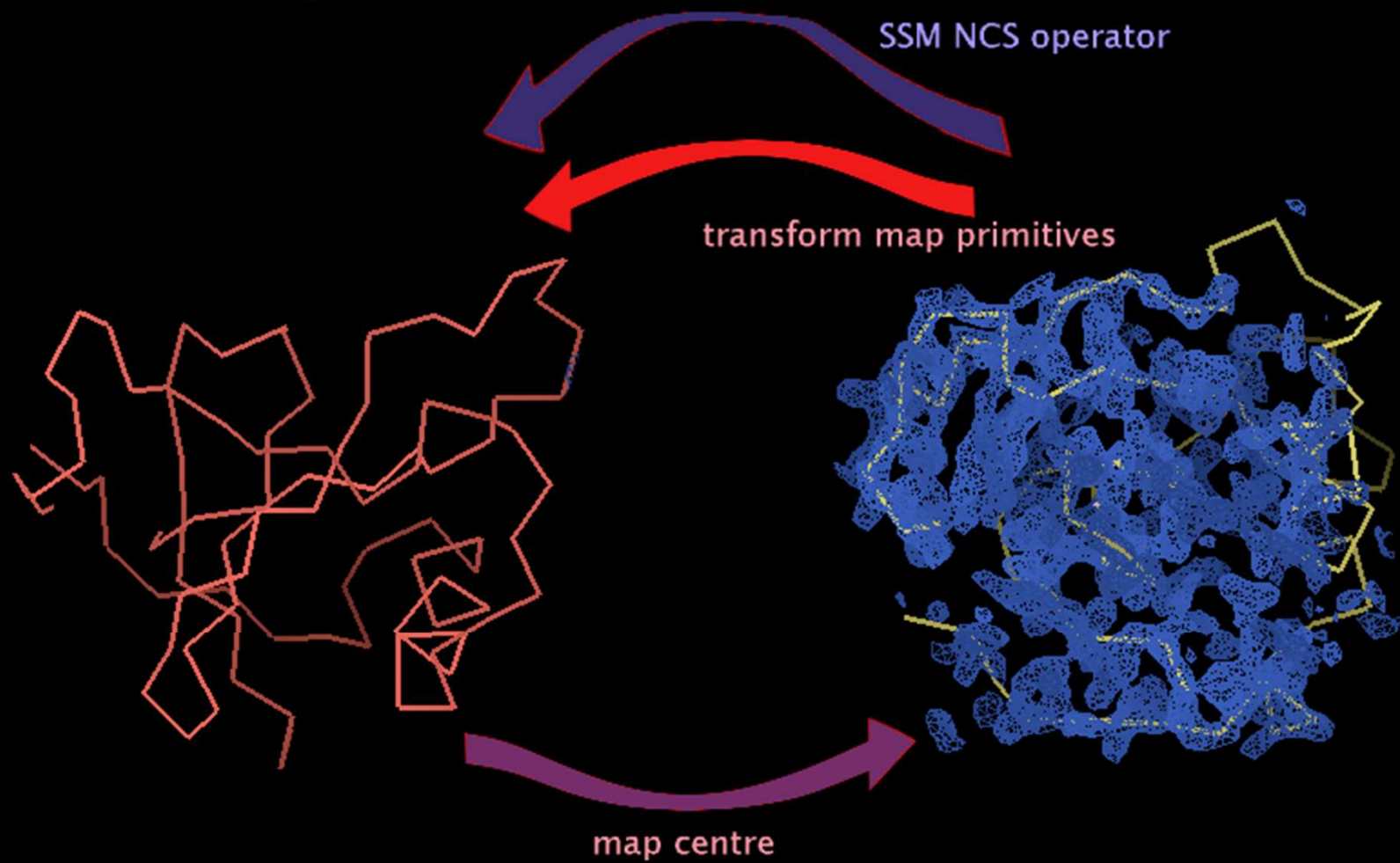


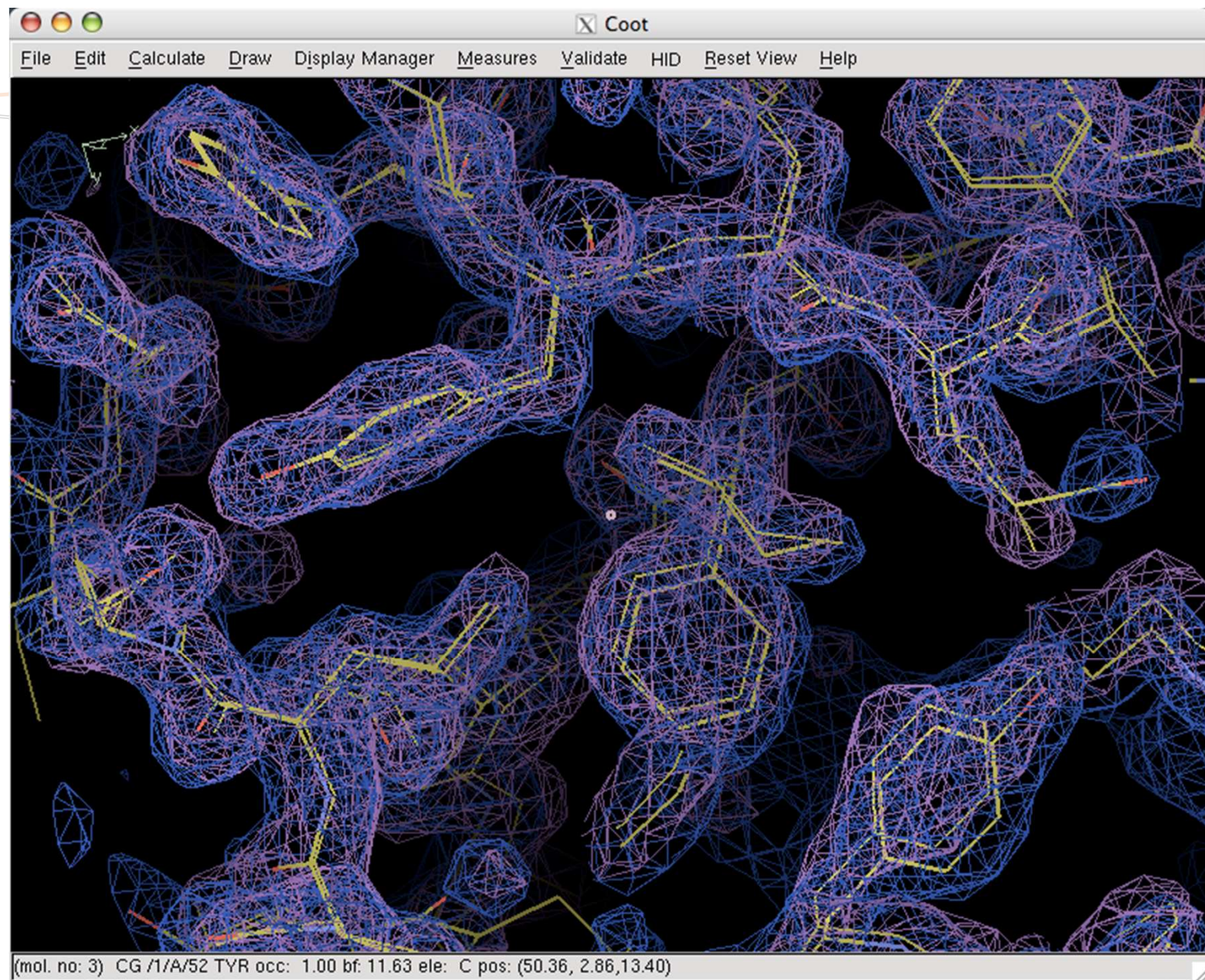
Kleywegt plot





NCS Overlays





NCS 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)
- NCS ligands

