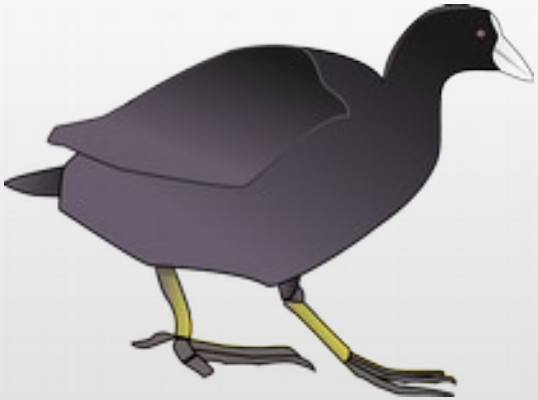
Oct 2018 CSHL

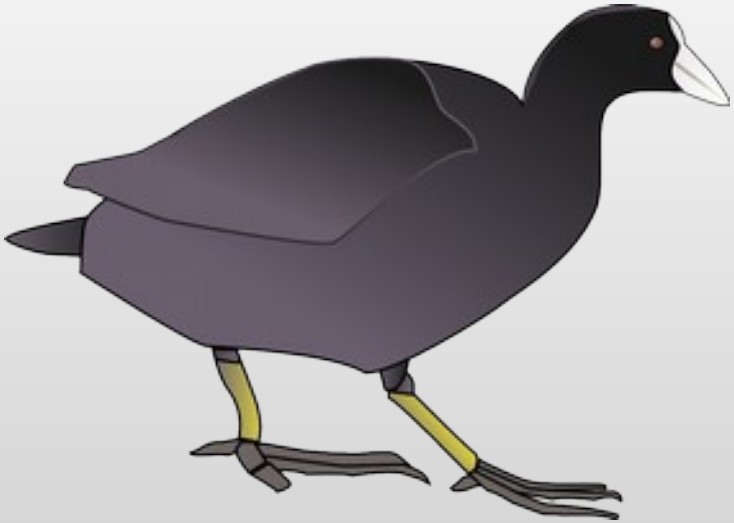


Model-Building using X-ray data

(with *Coot*)

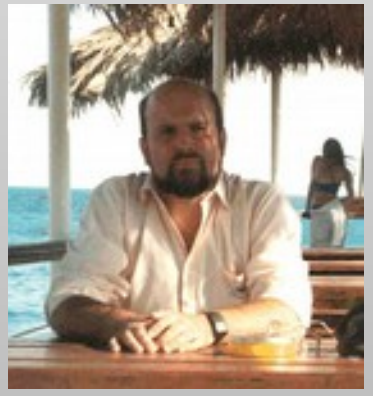
Paul Emsley,
MRC Laboratory of Molecular Biology
Cambridge, UK

Coot Collaborators



Bernhard Lohkamp

Kevin Cowtan



Eugene Krissinel

Stuart McNicholas

Martin Noble

Alexei Vagin

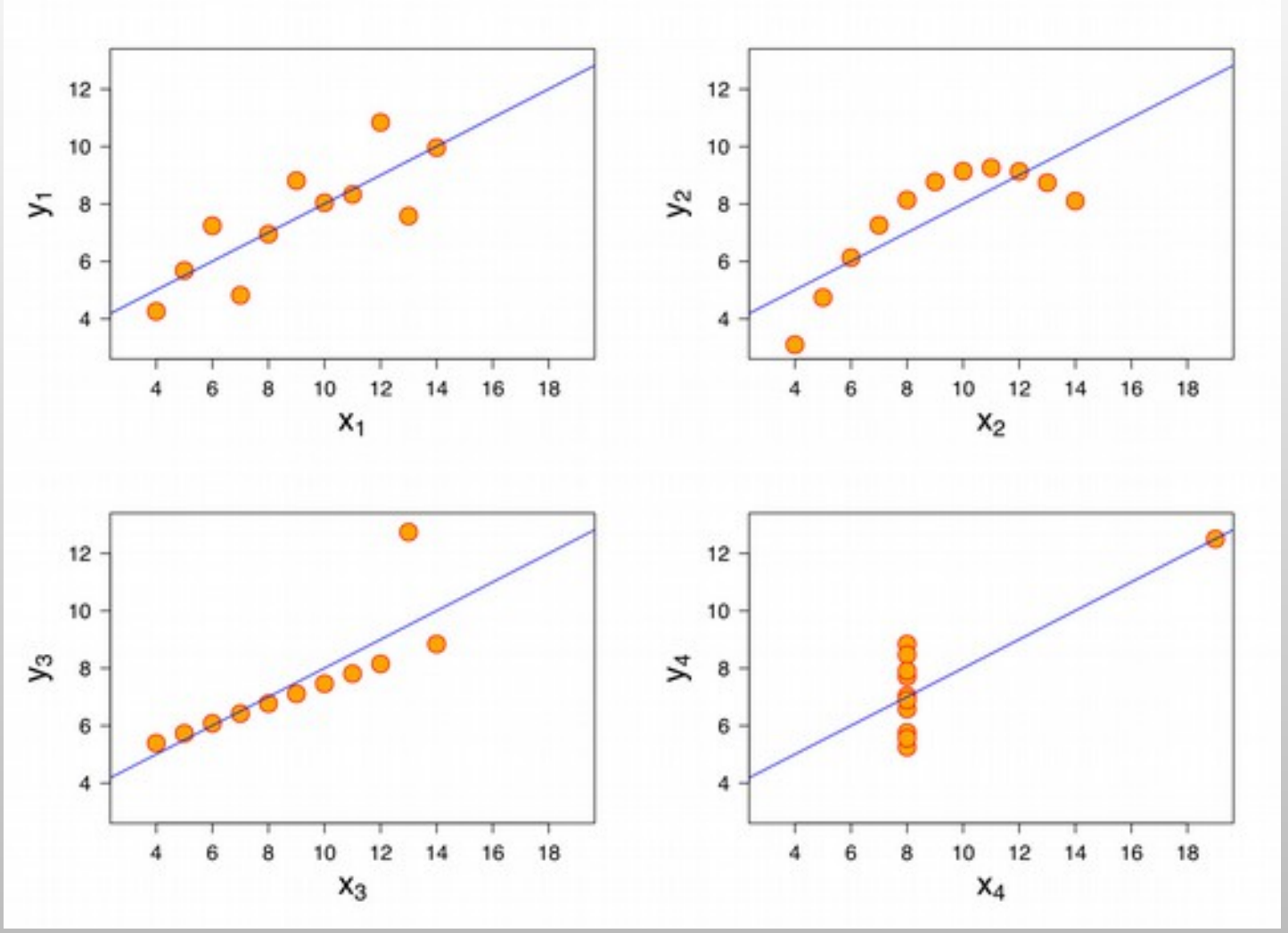
A bit of context

- Why use 3D graphics?

Summary Statistics

- Are useful, but don't tell the whole story
- Let's say we have 10 data points
  - X mean 9
  - Y mean 7.5
  - correlation 0.816
  - regression $y = 3 + \frac{1}{2} x$

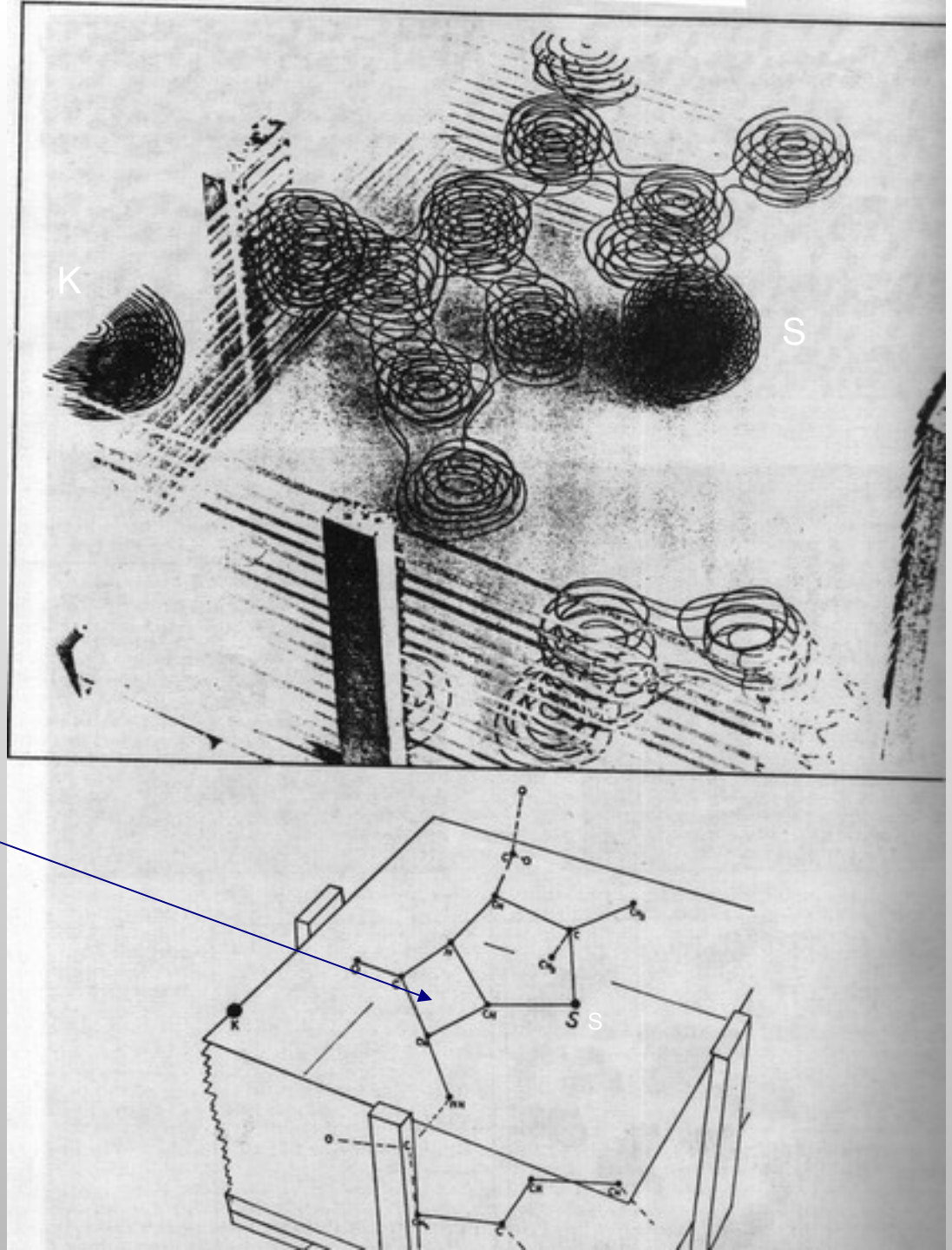
View Your Data and Model



Anscombe's Quartet

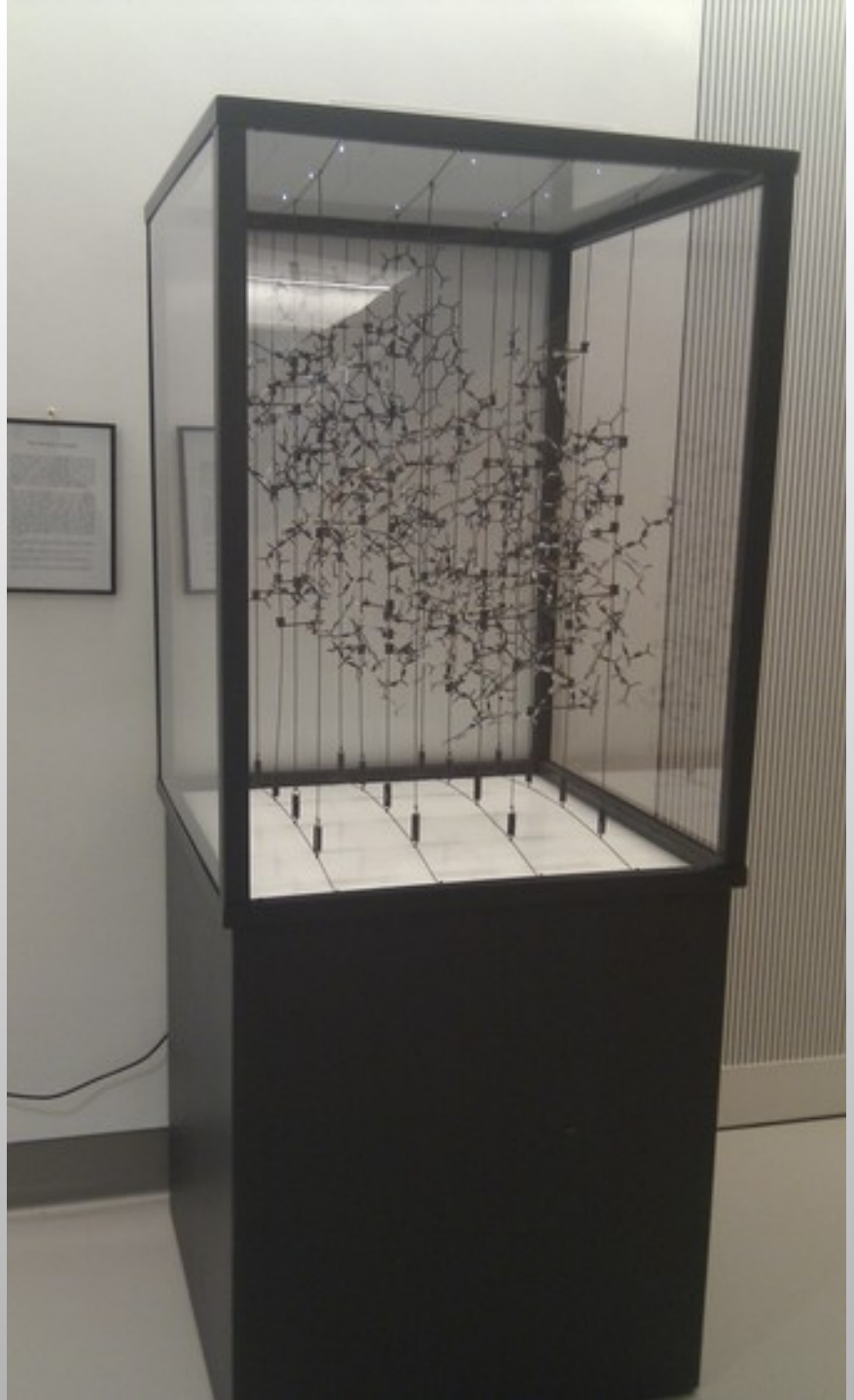
The 3-dimensional penicillin G map calculated in 1944. This unequivocally determined its chemical structure. (The hand however is wrong.)

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

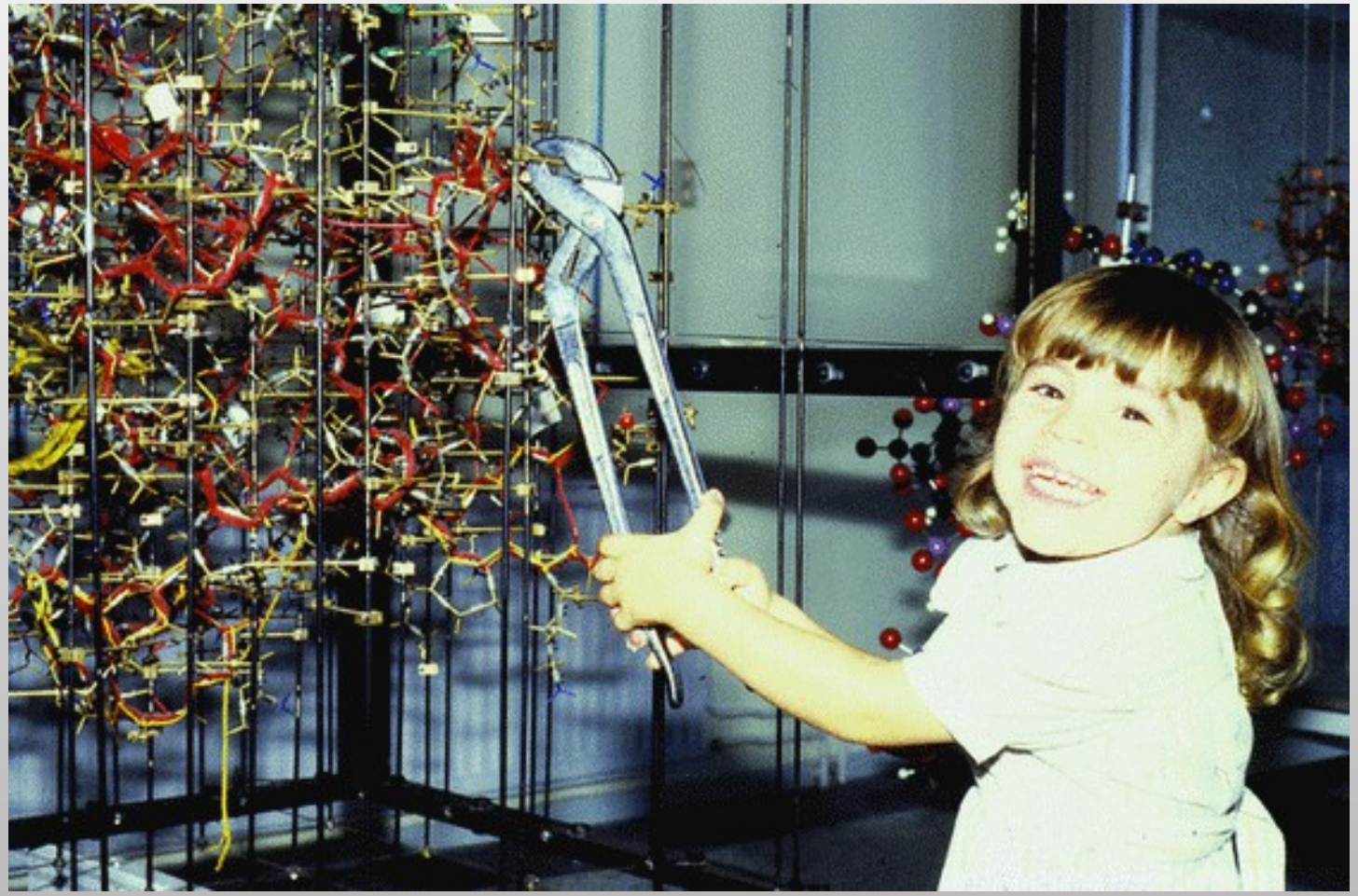


DCH building first insulin model

Lysozyme 1965



Kendrew wire model of alcohol dehydrogenase that is about to undergo a round of rebuilding by Maelle Cambillau



T. Alwyn Jones (2004)

History moves on...

- INTER, BUILDER, CHAIN, FRODO, O...

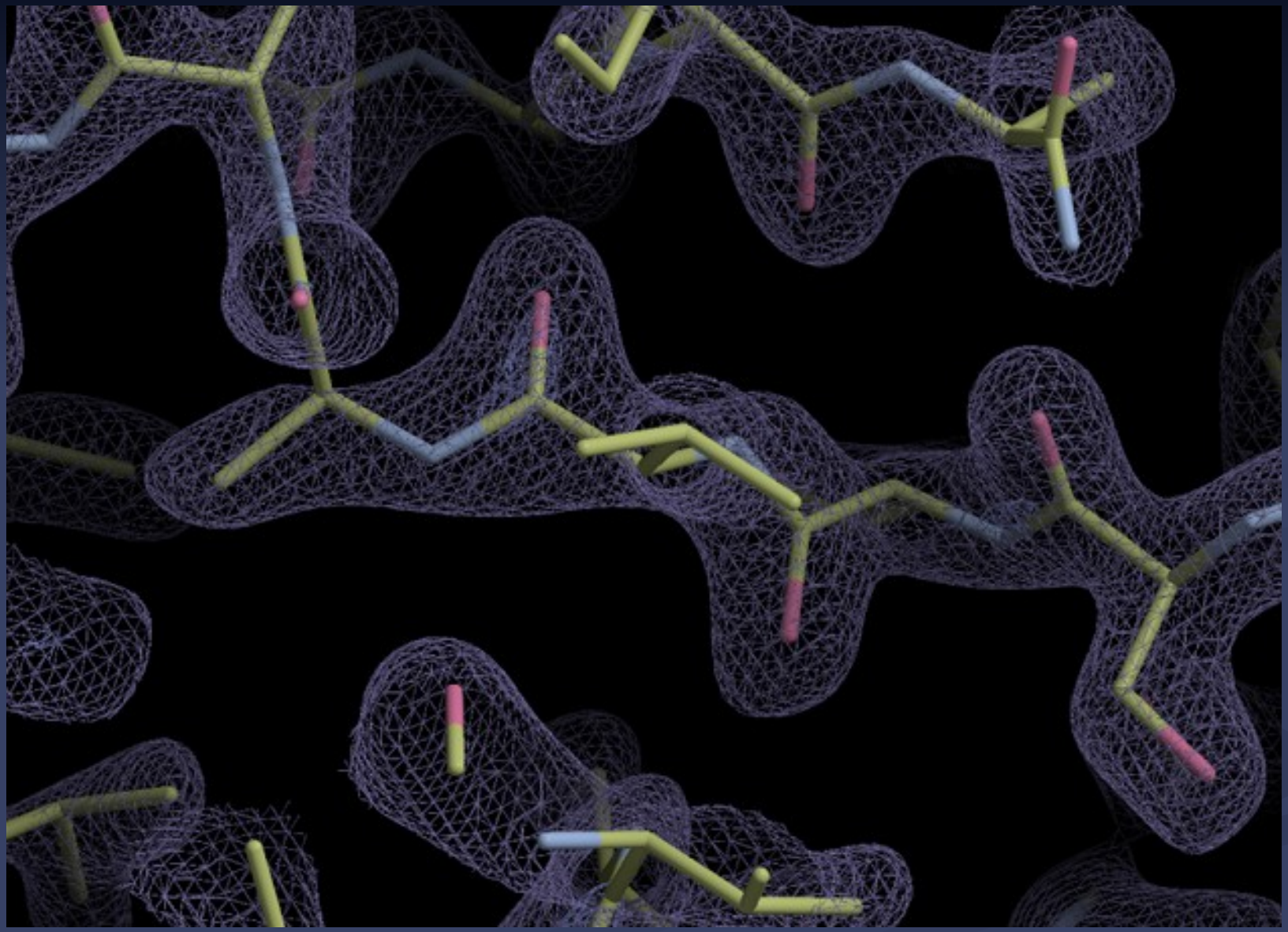
Coot

- Molecular Graphics application
  - Protein Crystallographic model-building tools
  - Designed to “fill the gap” where automatic methods fail
    - (generally, we don't use molecular graphics programs to do what automatic methods can do)
- Interface to other programs: SHELXL, Refmac, Libcheck, Probe&Reduce (Molprobity), EBI, EDS, Povray… and others

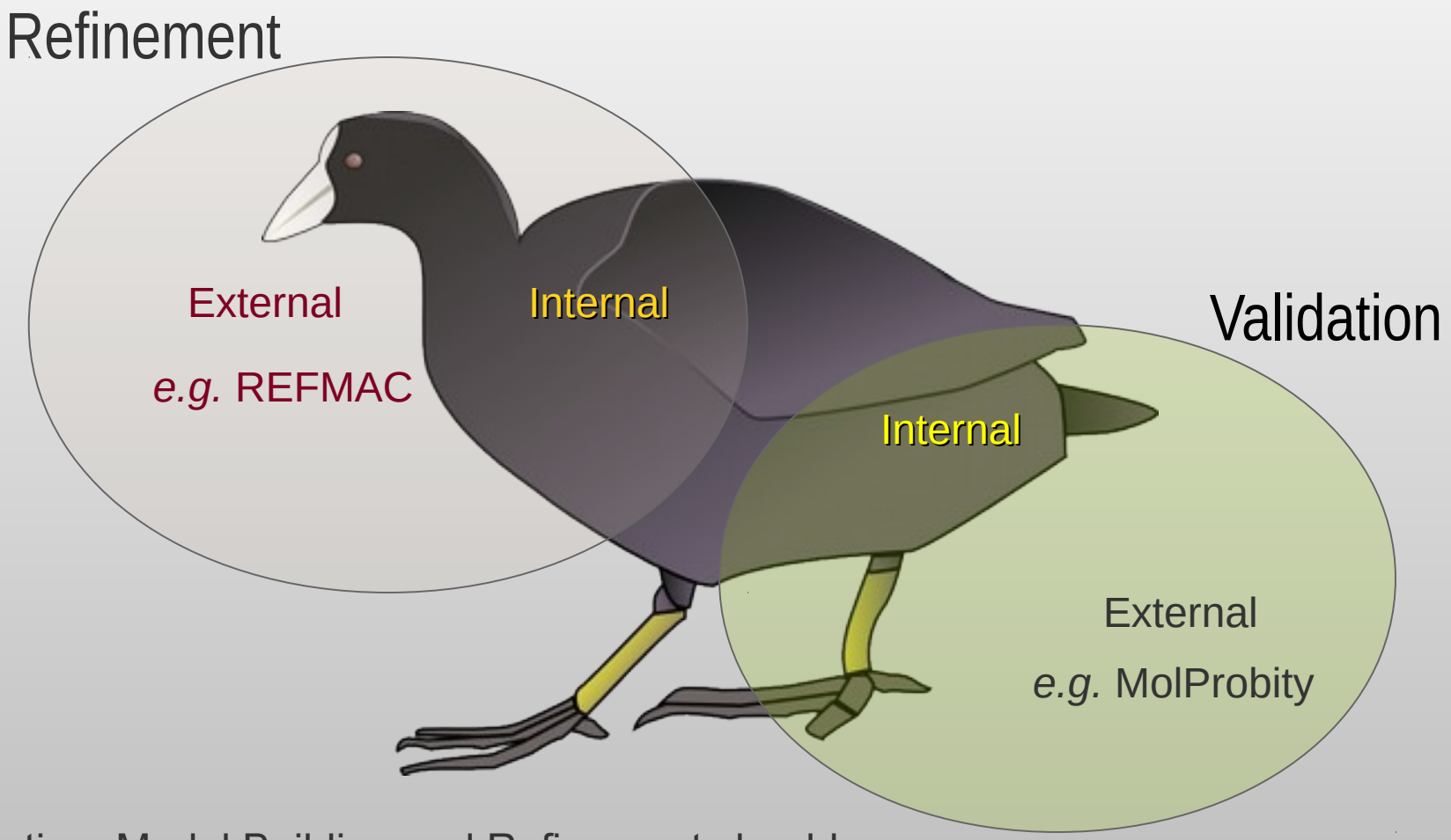
But Why Bother?

- Automated model-building for complete models is still impossible
  - It takes a brain to validate
- Concerted correction/improvement of a model is difficult on the larger scale

2.2Å



Feature Integration



Validation, Model Building and Refinement should be used together

What is “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 minimizer (BFGS derivative)
  - Based on mmCIF standard dictionary
  - Minimizing bonds, angles, planes, non-bonded contacts, torsions, chiral volumes
  - Additional user-defined restraints,
    - •secondary structure restraints
    - •homologous protein local environment restraints
- Provides “interactive refinement”

Refinement in *Coot* has been extended in several ways...

What prior geometric information do we have?

- We know chemistry....
  - We know bond lengths and uncertainties
  - We know bond angles and uncertainties
  - We know the chiral centres
  - We know which atoms should lie in a plane
  - We know (more or less) about torsions
- We combine the gradients from the data with those from molecular mechanics in the minimisation

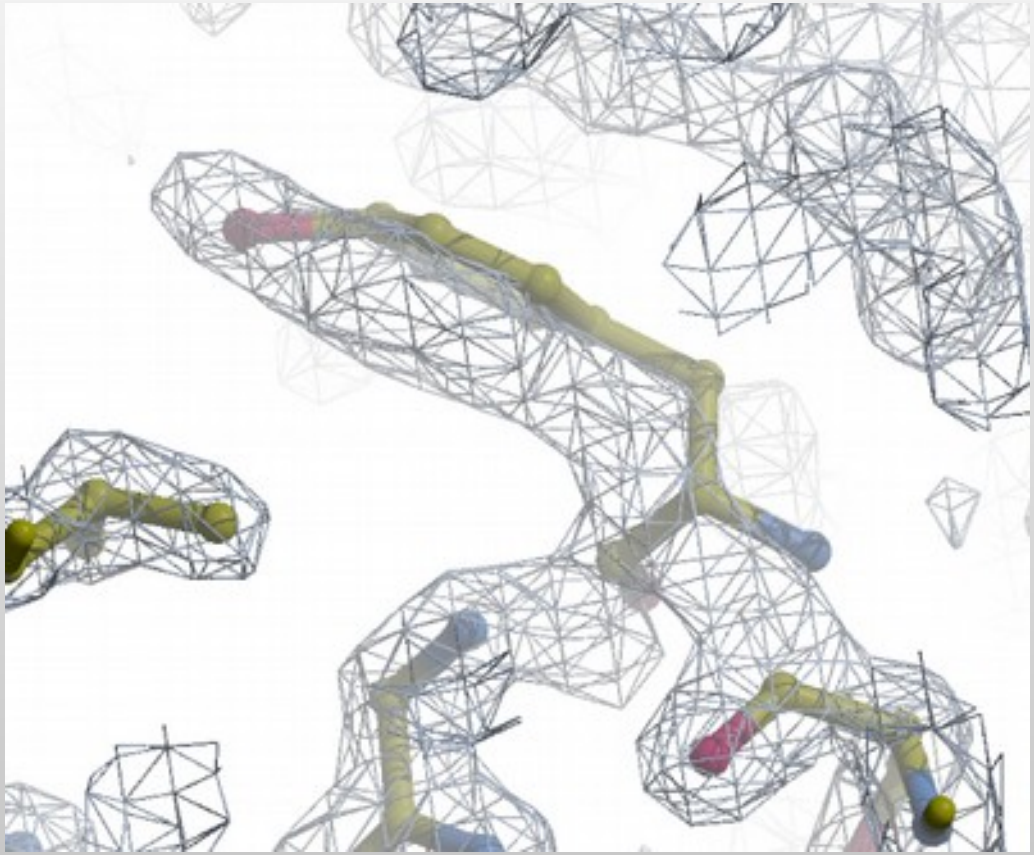
REFMAC Monomer Library
chem_comp_bond

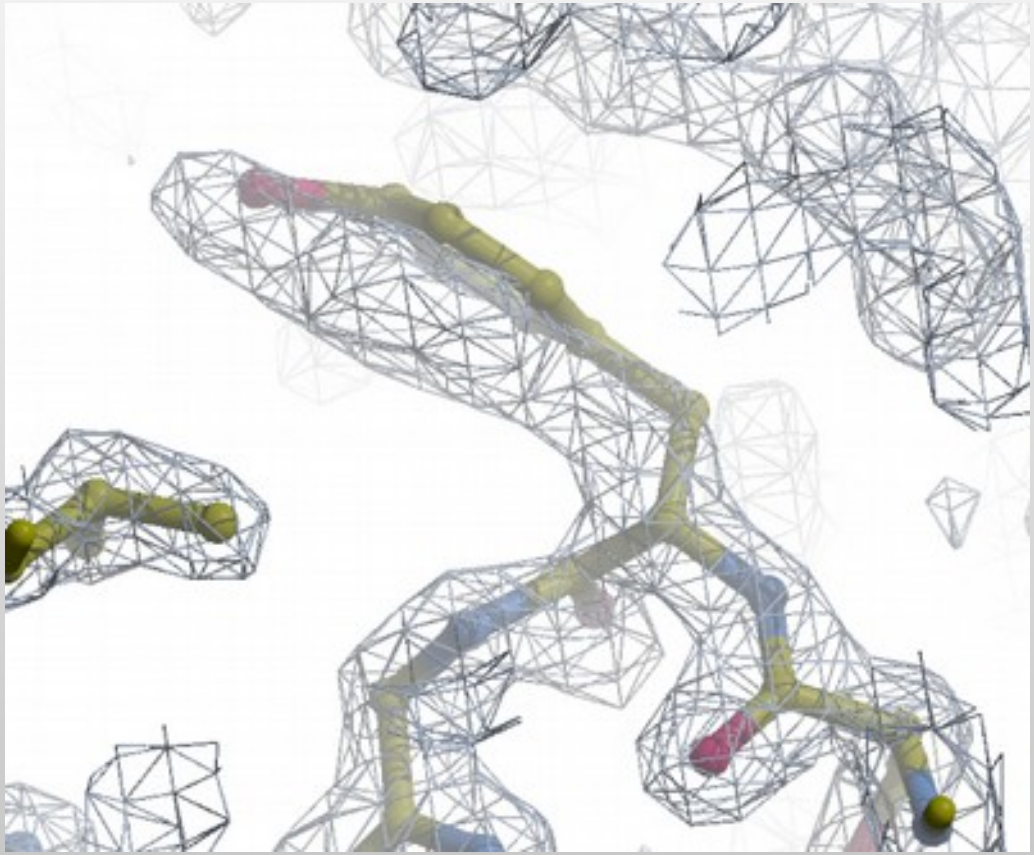
```
loop_
_chem_comp_bond.comp_id
_chem_comp_bond.atom_id_1
_chem_comp_bond.atom_id_2
_chem_comp_bond.type
_chem_comp_bond.value_dist
_chem_comp_bond.value_dist_esd
 ALA      N    H       single     0.860    0.020
 ALA      N    CA      single     1.458    0.019
 ALA      CA   HA      single     0.980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

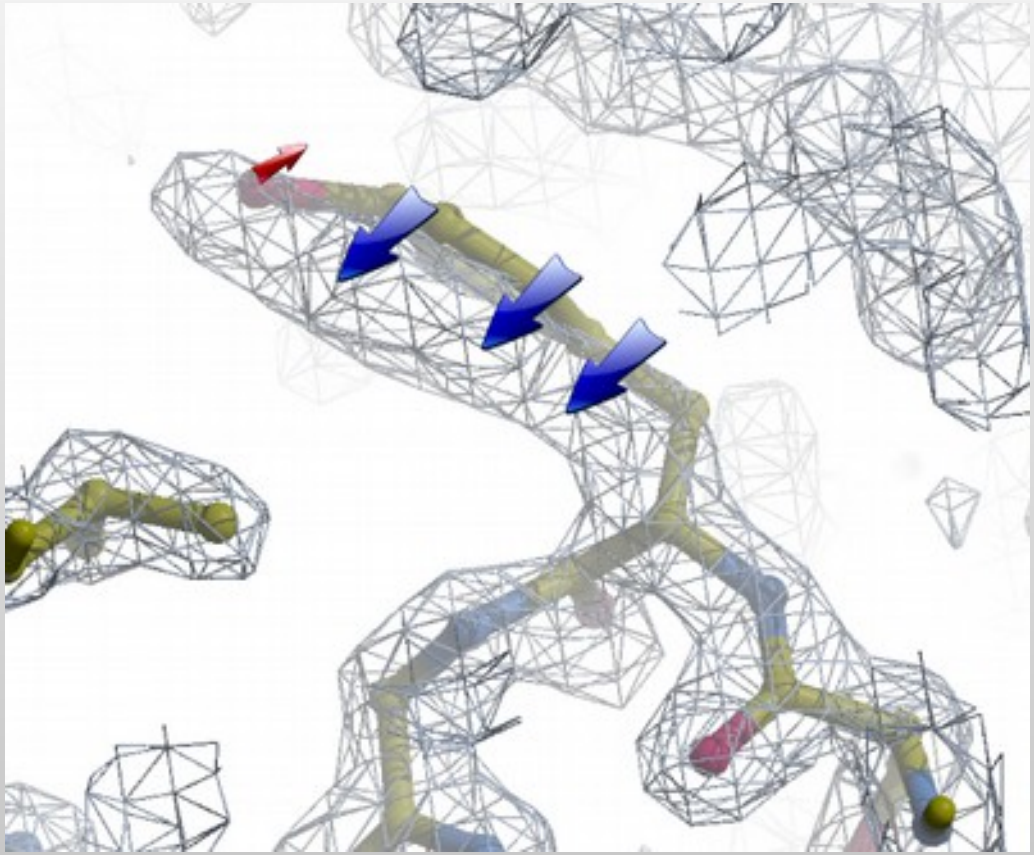
APPENDIX *A*
Regularization and refinement derivatives

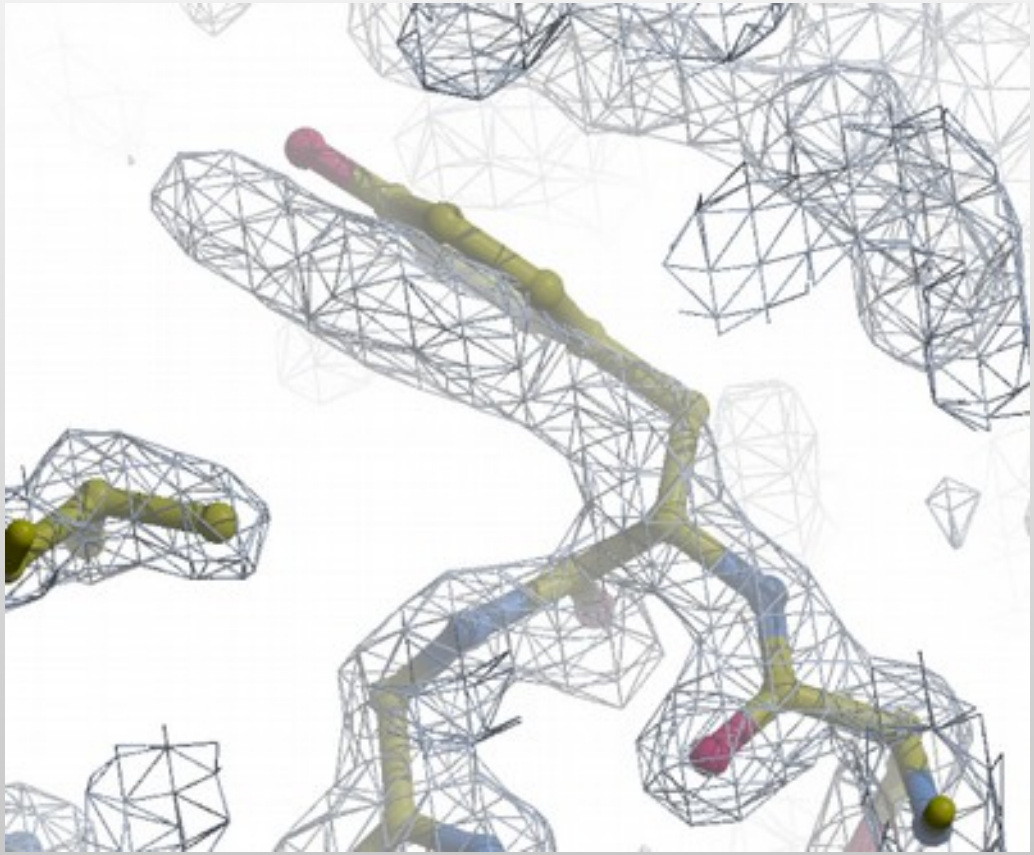
The function that we are trying to minimize is S, where

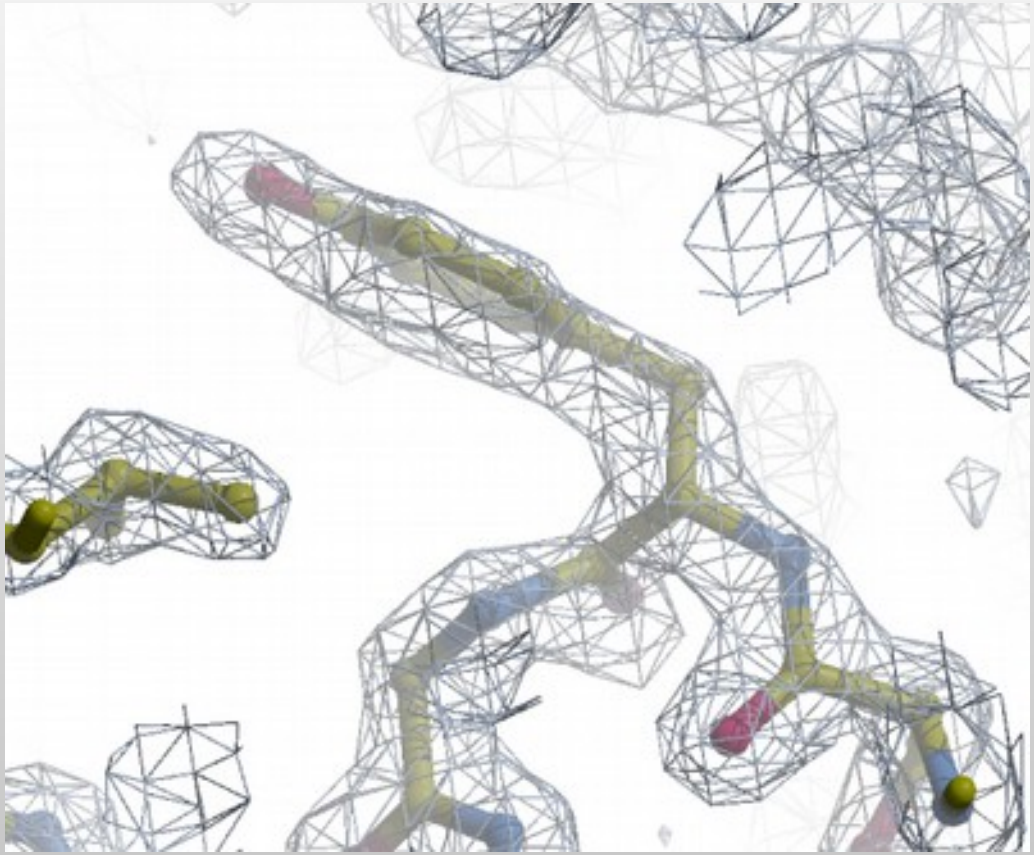
$$S = S_{\text{bond}} + S_{\text{angle}} + S_{\text{torsion}} + S_{\text{plane}} + S_{\text{nbc}} + S_{\text{chiral}}$$

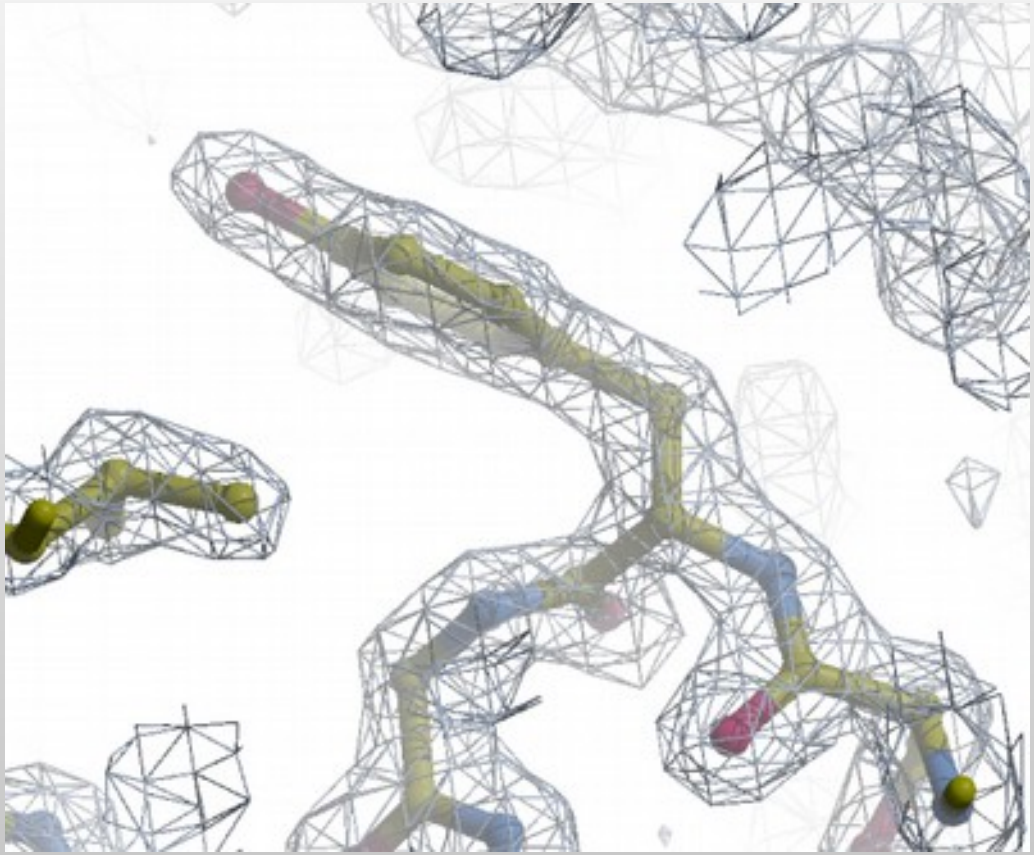








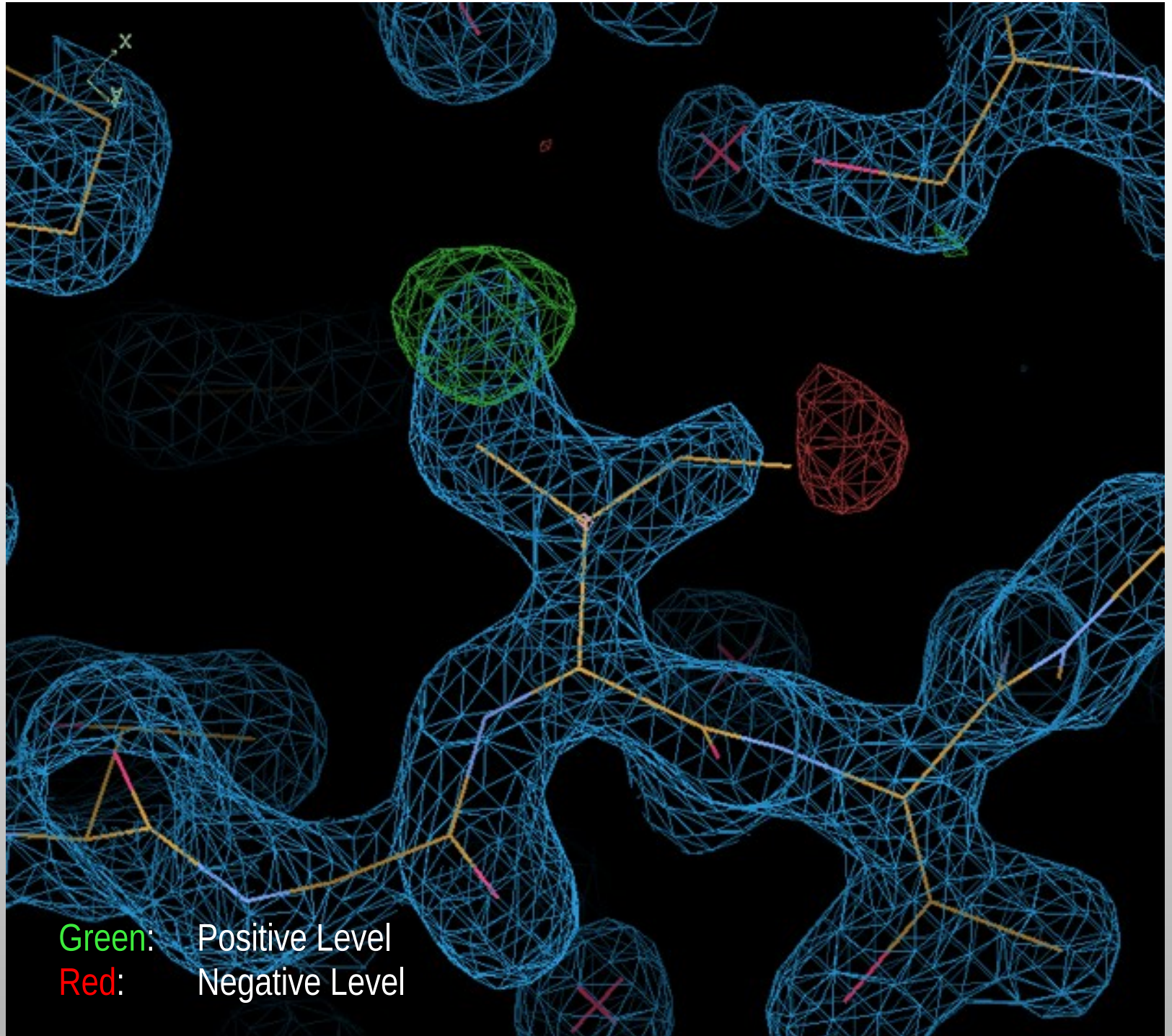




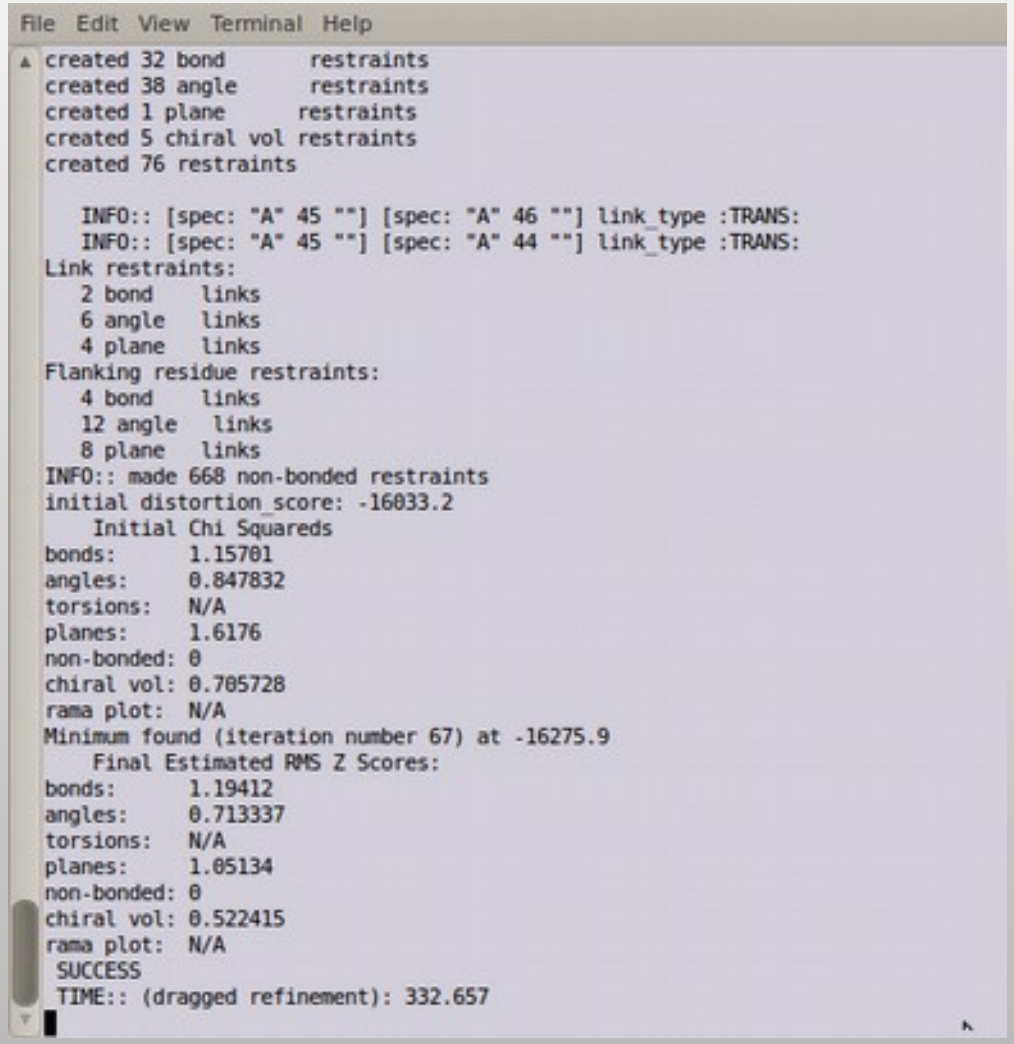
Different types of electron density maps

- “Experimental” maps
  - maps that result directly from the crystallographic data analysis: MIR, MAD, SAD
- Direct Maps:
  - where the atoms are
- Coefficients F_o-F_c (“difference map”)
  - Identifies errors in the model. Locations in space where there should be atoms show positive peaks, while locations where the model contains atoms that should not be there show negative peaks.

Morten Kjeldgaard


Green: Positive Level
Red: Negative Level

Representation of Results:

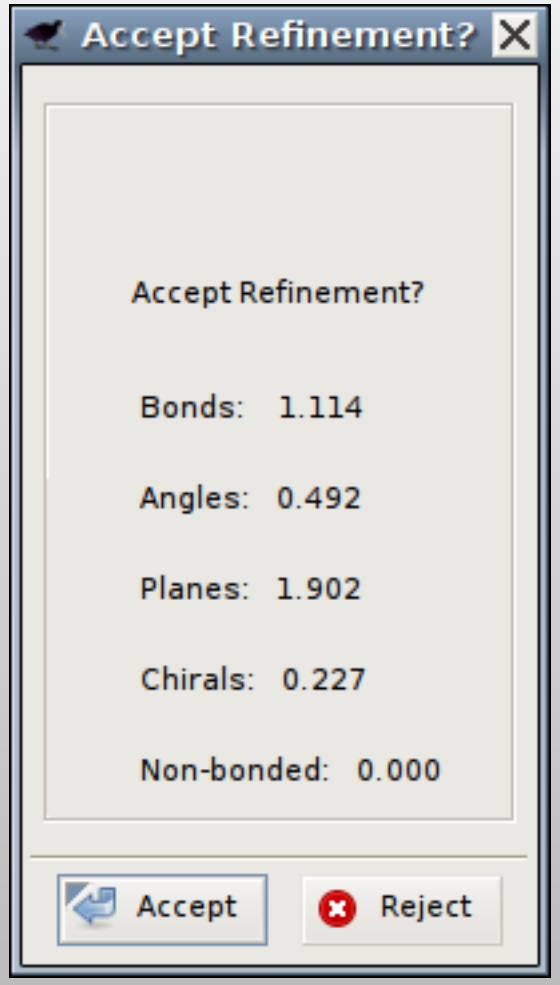


The first attempt

Student Reaction:

“Oh, I don't look at that window...”

Representation of Results:



Second attempt...

Student Reaction:

“Oh, box of meaningless numbers.

Go away”

Representation of Results: “Traffic Lights”

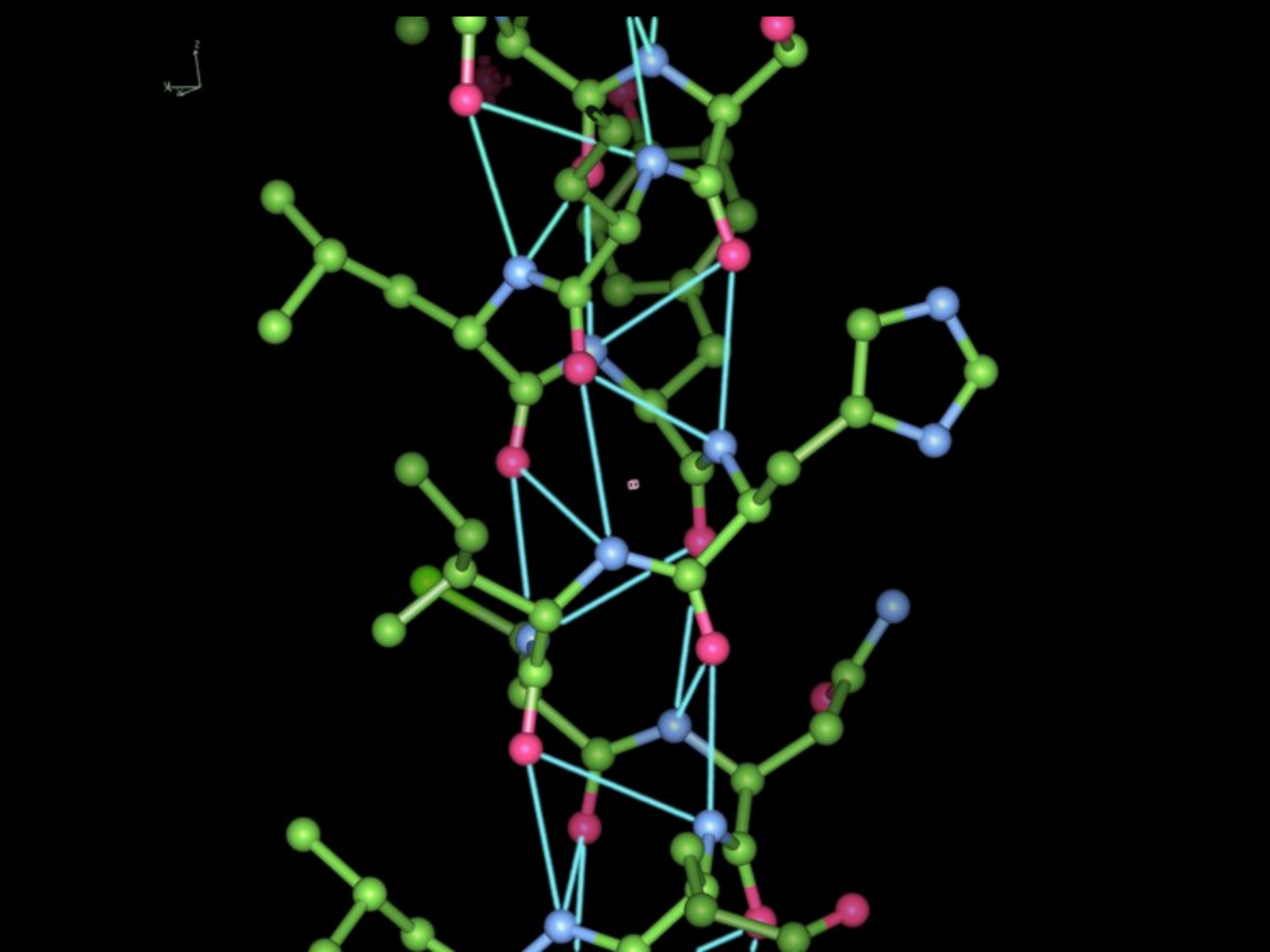
“Traffic Lights” represent the RMSd values for each of the refined geometry types

Good refinement

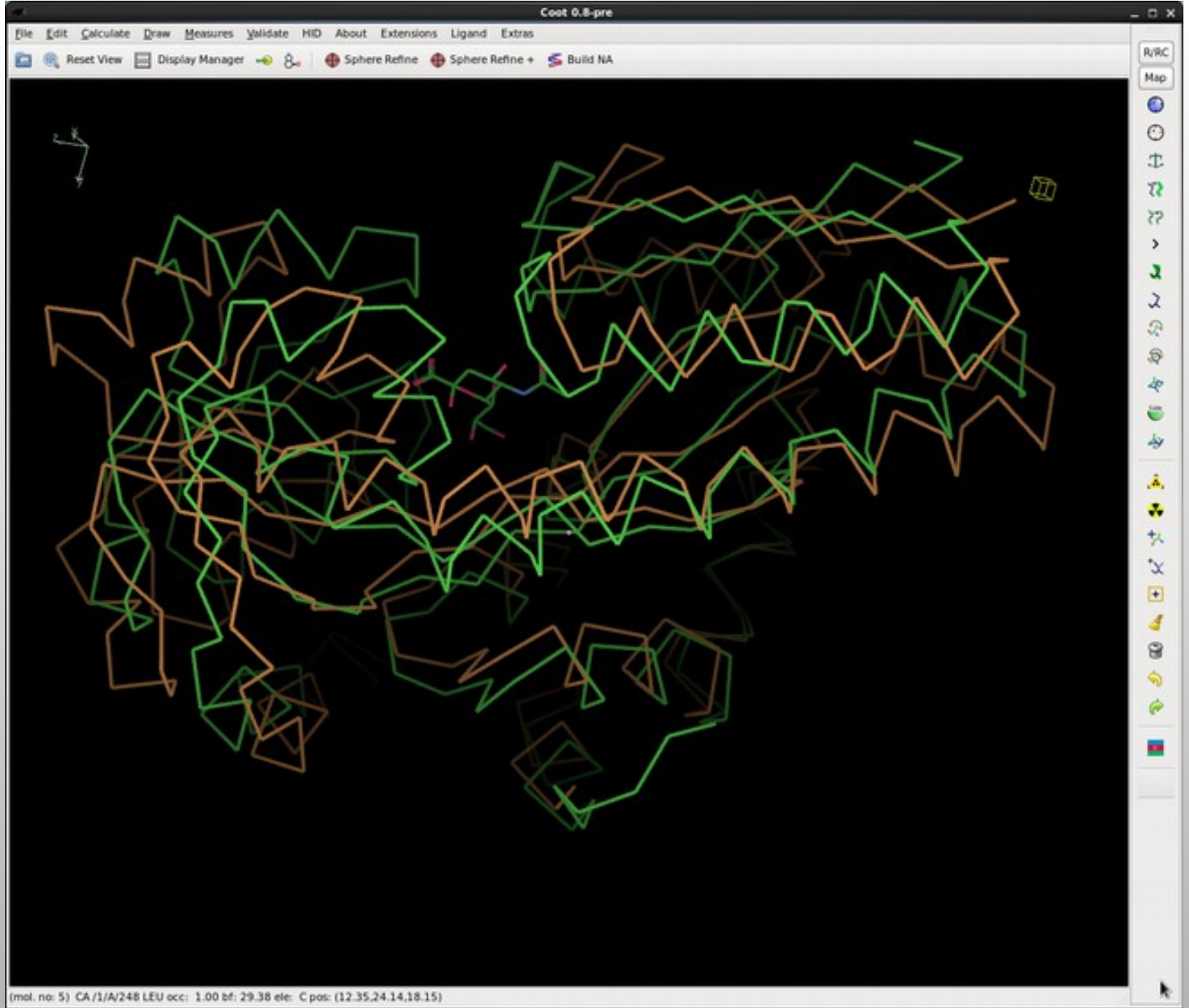
Bad refinement

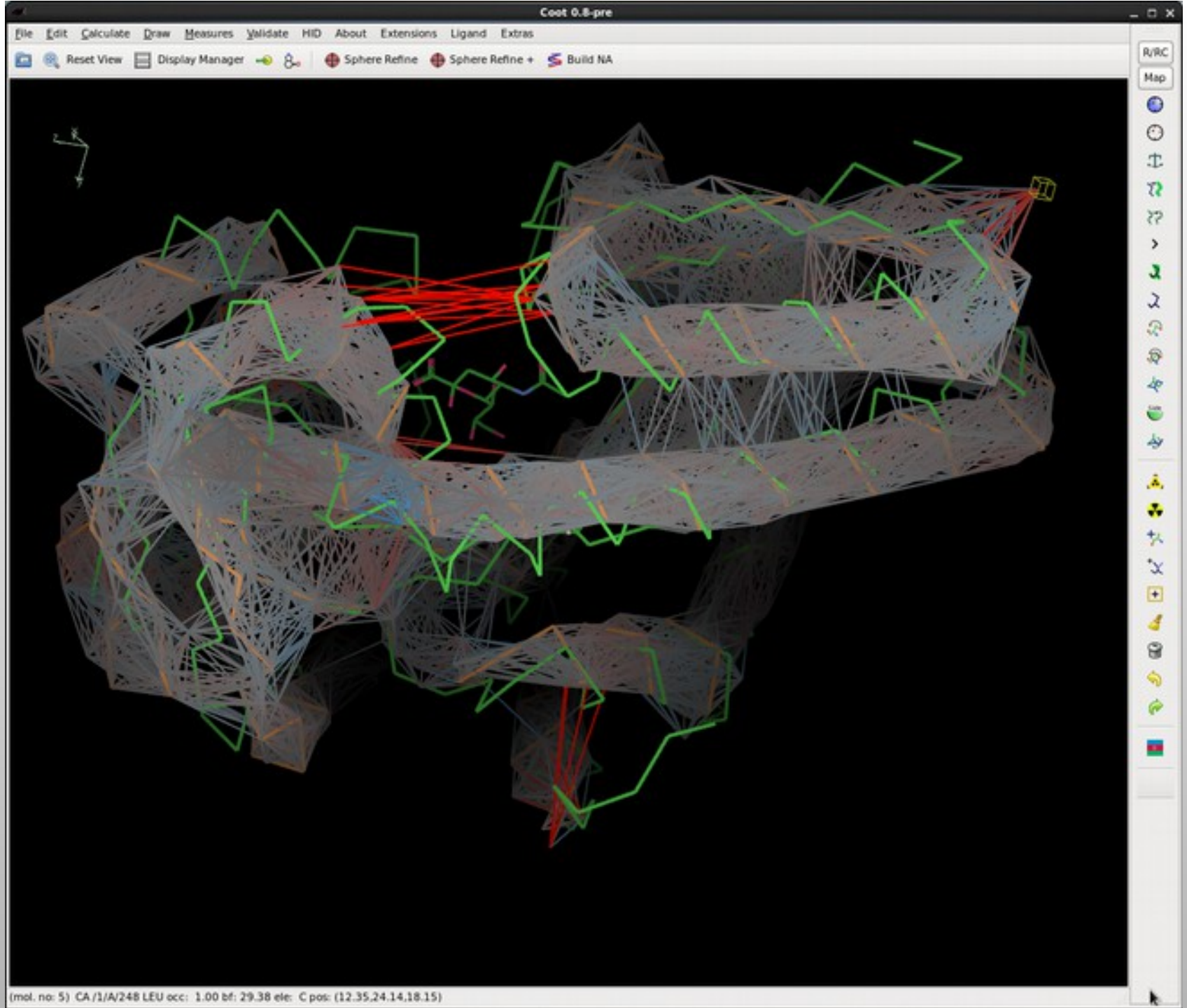
ProSMART Interface

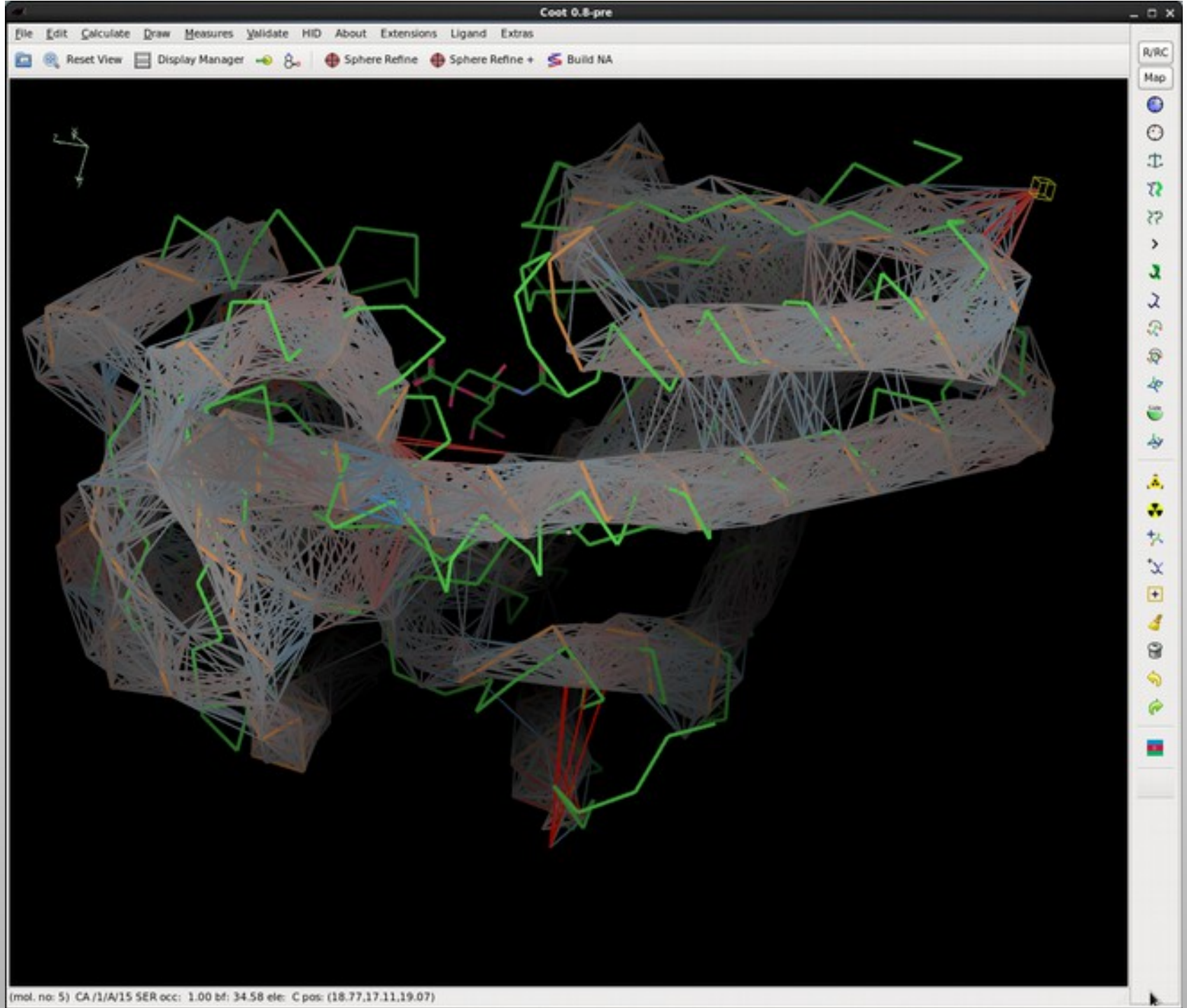
- Use previous-solved “template” structures to inform the refinement of the (low resolution) target protein
- Conformation-independent structural comparison/superposition
- and restraint generation




Coot 0.8-pre
File Edit Calculate Draw Measures Validate HID About Extensions Ligand Extras
Reset View
Display Manager
Sphere Refine
Sphere Refine +
Build NA
R/RC
Map
(mol. no: 5) CA /1/A/248 LEU occ: 1.00 bf: 29.38 ele: C pos: (12.35,24.14,18.15)


Coot 0.8-pre
File Edit Calculate Draw Measures Validate HID About Extensions Ligand Extras
Reset View
Display Manager
Sphere Refine
Sphere Refine +
Build NA
R/RC
Map
(mol. no: 5) CA /1/A/248 LEU occ: 1.00 bf: 29.38 ele: C pos: (12.35,24.14,18.15)

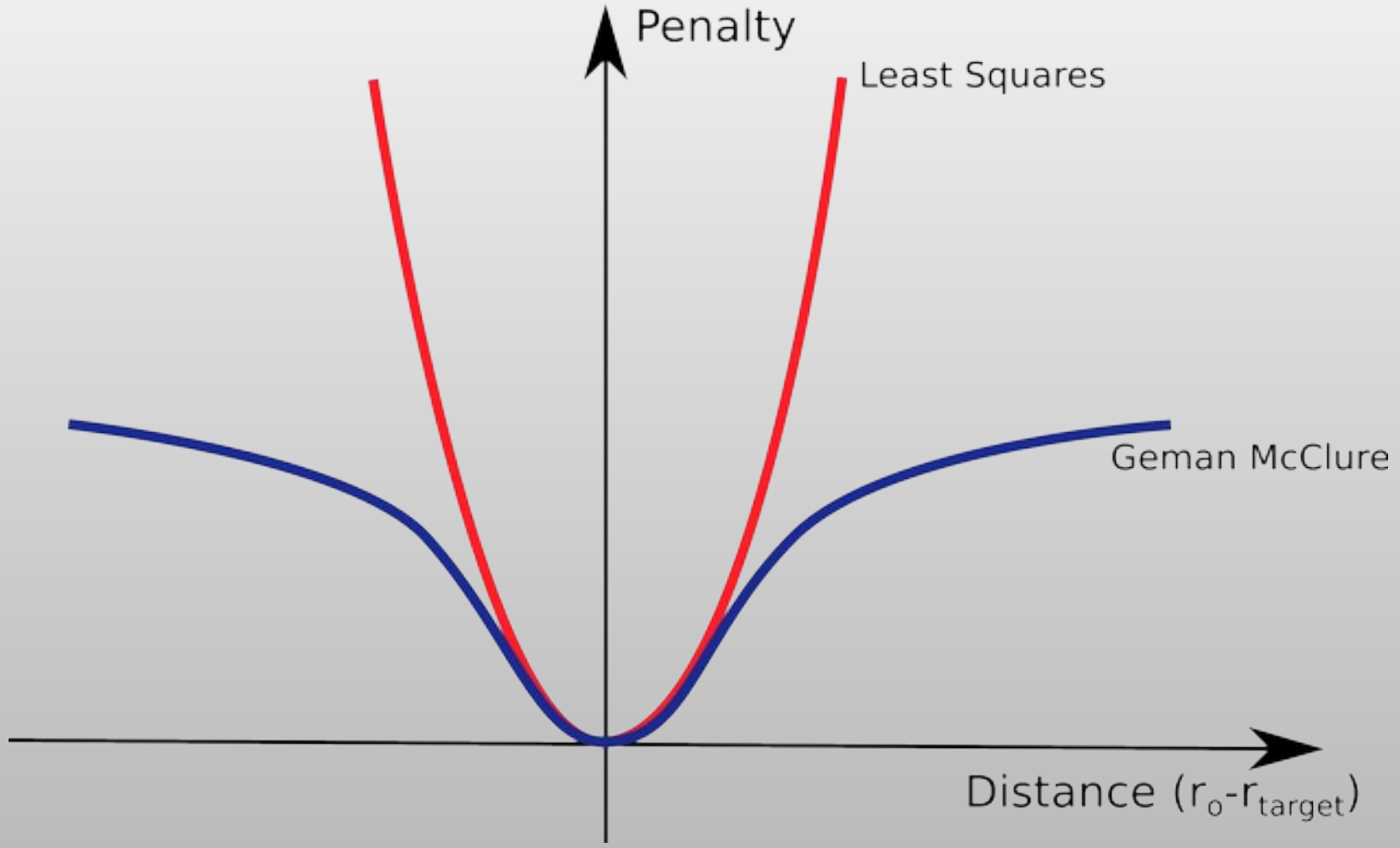

Coot 0.8-pre
File Edit Calculate Draw Measures Validate HID About Extensions Ligand Extras
Reset View
Display Manager
Sphere Refine
Sphere Refine +
Build NA
R/RC
Map
(mol. no: 5) CA /1/A/248 LEU occ: 1.00 bf: 29.38 ele: C pos: (12.35,24.14,18.15)


Coot 0.8-pre
File Edit Calculate Draw Measures Validate HID About Extensions Ligand Extras
Reset View
Display Manager
Sphere Refine
Sphere Refine +
Build NA
R/RC
Map
(mol. no: 5) CA /1/A/15 SER occ: 1.00 bf: 34.58 ele: C pos: (18.77,17.11,19.07)

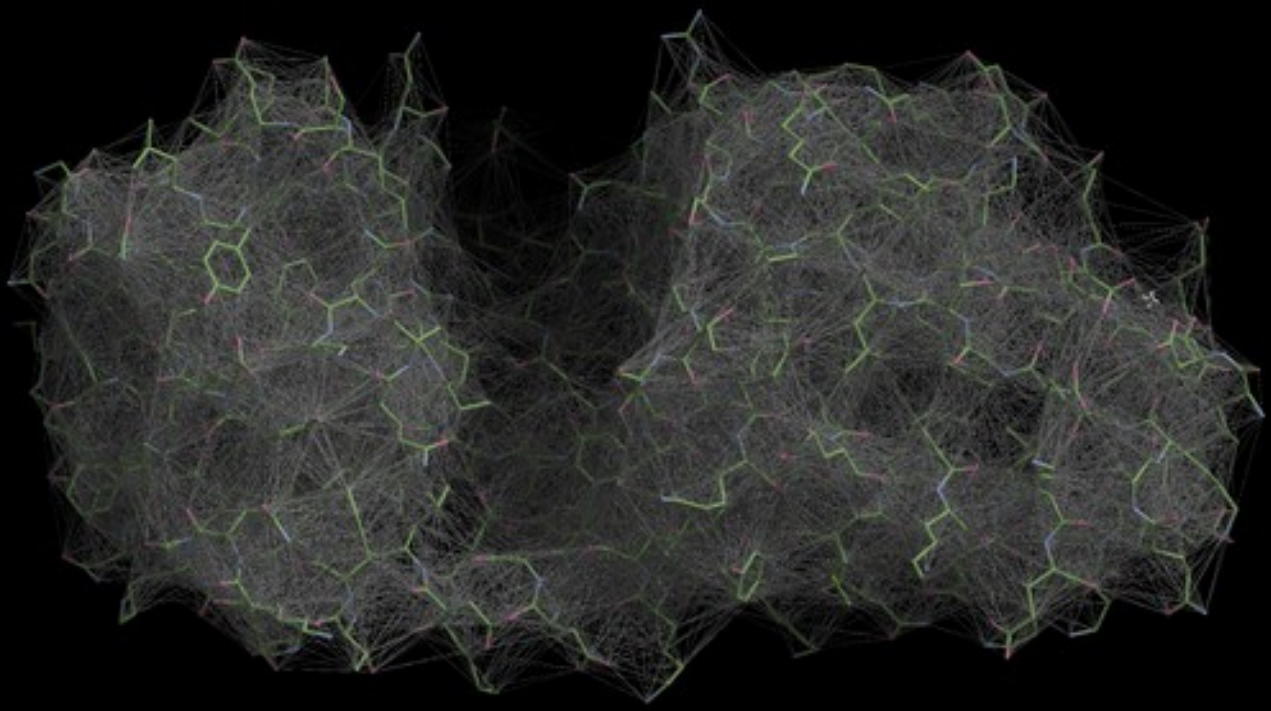
ProSMART integration

- ProSMART generates distance restraints from homologous structures
  - to be applied to current model for refinement
  - now available in *Coot*

Modified Target Function



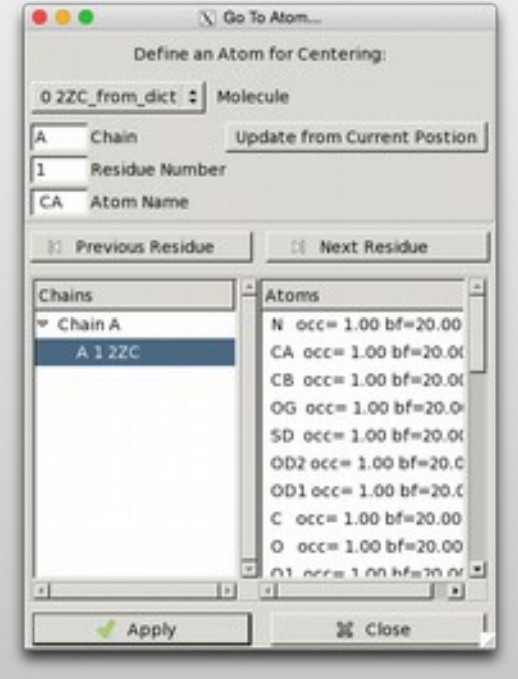
ProSMART Restraints



A note on *Coot*'s GUI

- It used to be clean
- Now lots of features have been added without much thought
- “Somewhat difficult to navigate”
- “Hidden” hot-keys

IISTDTIDIW



- If I See This Dialog Then I'm Doing It Wrong

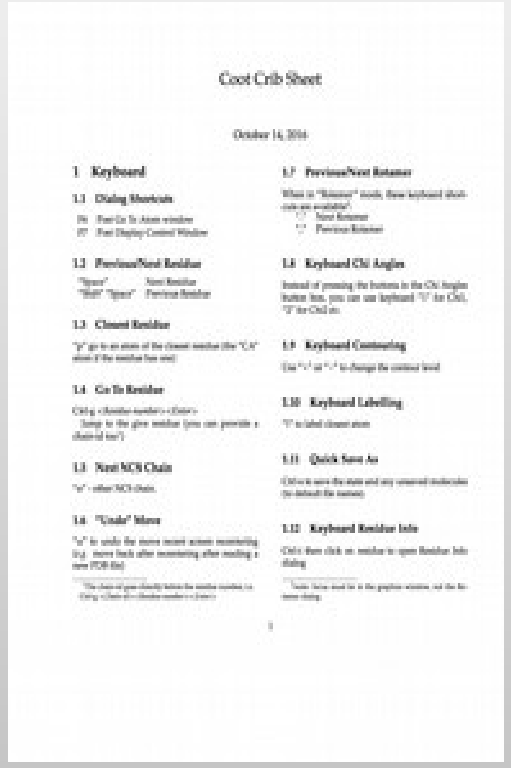
CURLEW:
*C*oot *U*tilites and *R*efinement *L*ibrary *E*xtention *W*ranger

- Easy access to ”interesting” *Coot* scripts

Refinement Techniques

- Single-Atom Drag
  - Over-dragging
- Key-bindings:
  - Triple Refine “T”, with auto-accept: “H”
  - Single Residue Refine: “R” with Auto-accept: “X”
  - Add Residue: “Y”
  - Autofit rotamer” “J”
    - Residue Flip: E, Shift: Opt-Alt- → Rotate: Ctl Shft - →
  - Hybridization-aware residue fragment rotation: “Shift F”

Coot Key-binding Crib-Sheet



Coot Crib Sheet

October 14, 2016

1 Keyboard

1.1 Dialog Shortcuts

[b] Pore Go To Atom window
[F] Pore Display Control Window

1.2 Previous/Next Residue

"Space" Next Residue
"Shift" "Space" Previous Residue

1.3 Closest Residue

"g" go to an atom of the closest residue (for "C,A" atom if the residue has one)

1.4 Go To Residue

Ctrl-g <Residue number><Enter>
Jump to the give residue (you can provide a chain-id too[1])

1.5 Next NCS Chain

"n" - other NCS chain.

1.6 "Undo" Move

"u" to undo the move (scott scores monitoring (e.g. move back after monitoring after reading a new PDB file)

[1] The chain-id (pair (chain)) before the residue number (i.e. Ctrl-g <Chain ID><Residue number><Enter>)

1.7 Previous/Next Rotamer

When in "Rotamer" mode, these keyboard shortcuts are available[2]:

'.' Next Rotamer
',' Previous Rotamer

1.8 Keyboard Chi Angles

Instead of pressing the buttons in the Chi Angles button box, you can use keyboard "1" for Chi1, "2" for Chi2 etc.

1.9 Keyboard Contouring

Use "+" or "-" to change the contour level

1.10 Keyboard Labelling

"l" to label closest atom

1.11 Quick Save As

Ctrl-w to save the main and any connected molecules (to default file names).

1.12 Keyboard Residue Info

Ctrl-i then click on residue to open Residue Info dialog

[2] Under Linux mode for in the graphics window, not the Rotamer dialog.

1.13 Keyboard Translation

Keypad 5 Push View (-Z translation)
Keypad . Pull View (Z translation)

1.14 Keyboard Undo/Redo

Ctrl-z Undo last modification
Ctrl-y Redo last modification
z Undo last move/navigation

1.15 Editing

Ctrl-c Copy active molecule
Ctrl-x Delete active residue

1.16 Keyboard Zoom and Clip

a Zoom out
m Zoom in
J Slim clip
f Fatten clip

1.17 Crosshairs

c cross hairs

1.18 Skeleton

s Generate skeleton around current point[3]

1.19 Continuous Rotate

t Toggle continuous spin

1.20 Baton Mode

b toggle into baton rotate mode[4]

[3] a skeleton is being displayed
[4] enter baton rotate mode

2 Mouse

Mouse actions are occasionally augmented with keyboard modifiers

Left-mouse Drag	Rotate view
Ctrl Left-Mouse Drag	Translate view
Shift Left-Mouse Click	Label Atom
Right-Mouse Drag	Zoom in and out
Shift Right-Mouse Drag	Change clipping and Front Screen Z The movement is along or near up–right/down–left shift up–left/down–right shift
Ctrl Shift Right-Mouse Drag	Rotate View about Screen
Middle-mouse Click	Centre on atom
Scroll wheel Forward	Increase map contour level
Scroll wheel Backward	Decrease map contour level

Intermediate (white) atoms can be dragged around by clicking on them

Left-mouse Drag	Move all intermediate atoms by lesser about an atom with
Left-mouse Drag with "A" key	non-linear about
Left-mouse Drag with "Ctrl"	Move a single atom

3 Refinement Extras

Use "A" to define a residue range[5] with a single-click. Useful in Refinement and Regularization.

- Click "Real Space Refine Zone"
- Click on an atom
- Press the "A" key

[5] a residues from the current residue

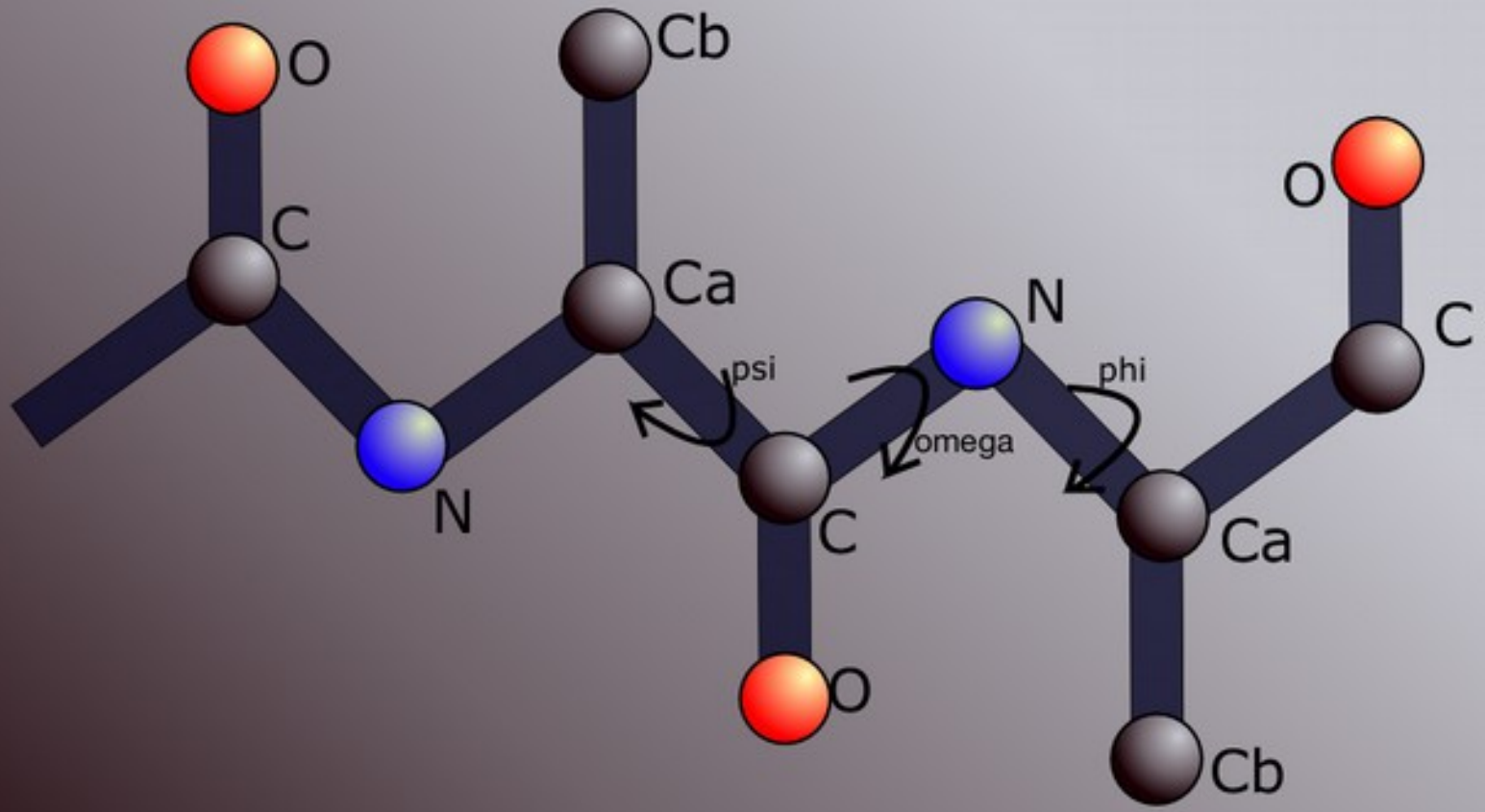
4 Template Key-bindings

E	Flip Ligand
G	Go To Blob (under cursor)
H	Neighbour refine
I	Toggle Fit This Residue
K	Fill Partial Side-chain
R	Refine Active Residue
T	Triple Residue Refine
X	Refine Active Residue and Auto-accept
W	Add Water
Y	Add Terminal Residue
Shift-Q	Rotamer Dialog for Residue
Shift-R	Sphere Refine
Shift-A	Accept Baton Position
Shift-B	Sphere Regularize
Shift-F	Delete Residue Hydrogens
Shift-Y	Undo Symmetry View
Shift-N	Edit Chi Angles
Shift-W	Add Water to Blob
Shift-A	Bid and Stick for Ligand

Rotamer Searching

- Two methods
  - Traditional
  - Backrub

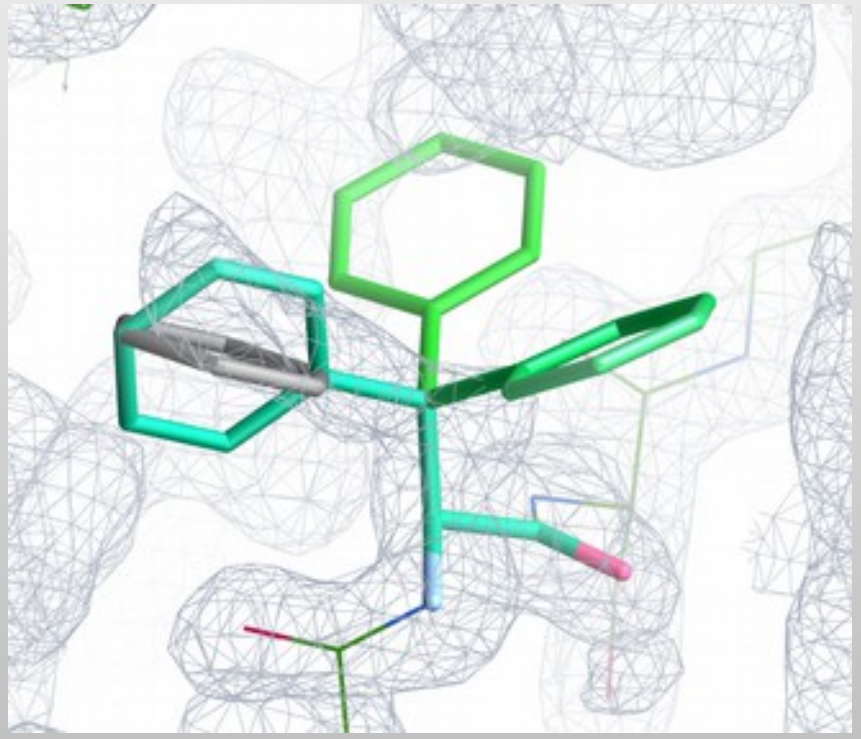
Peptide Torsion Angles

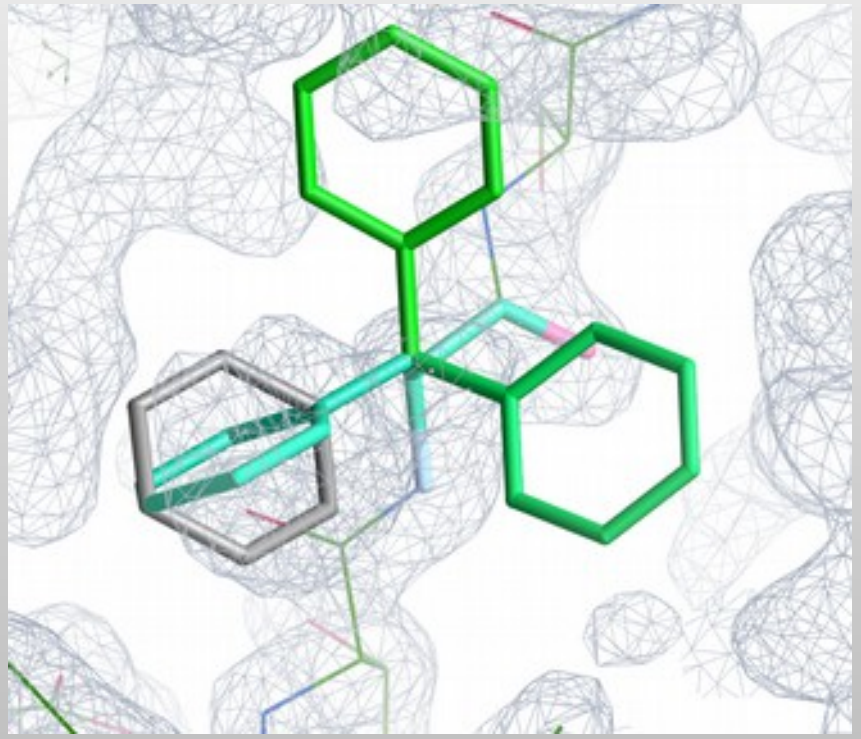


Rotamers

- 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

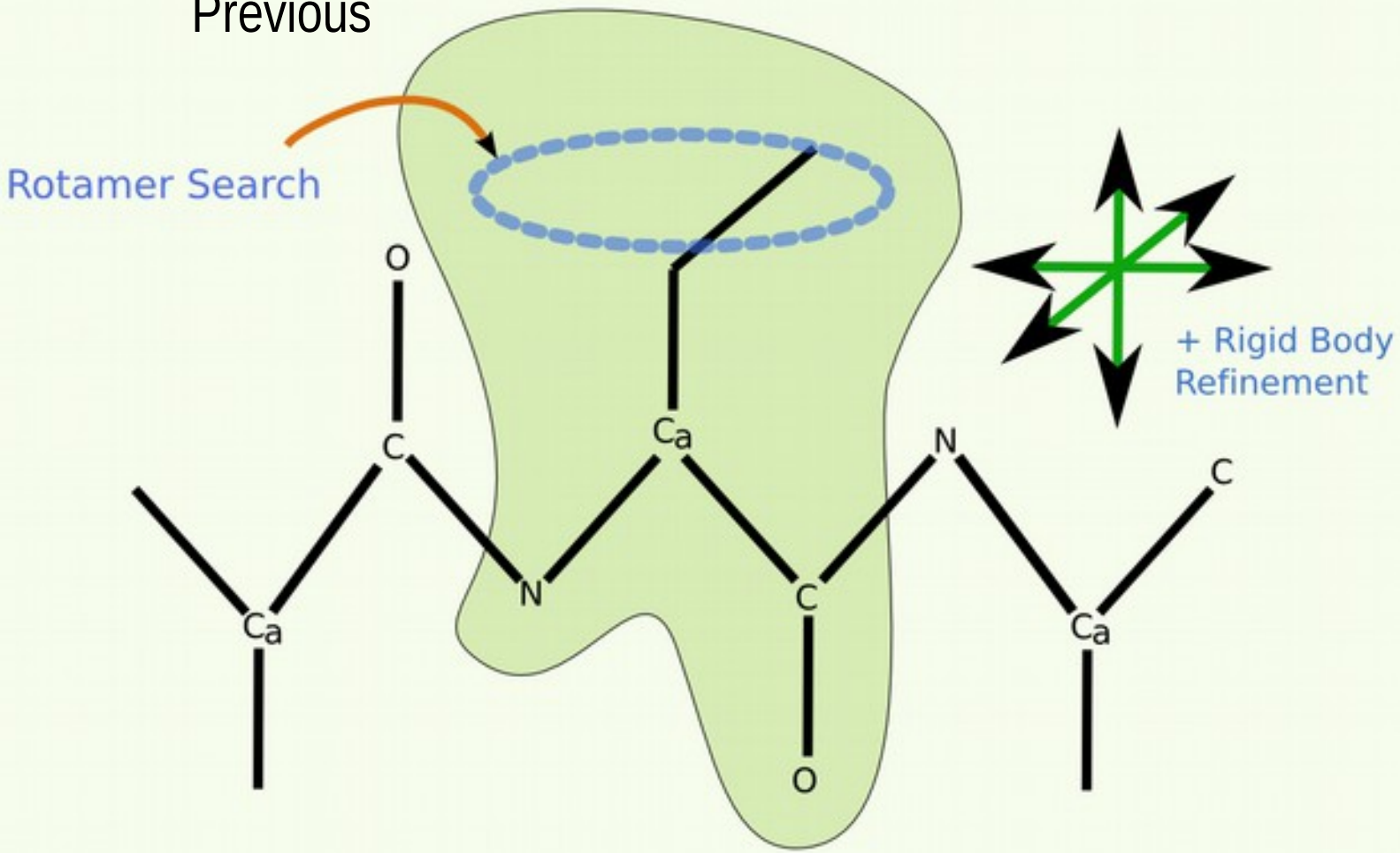
4 PHE Rotamers

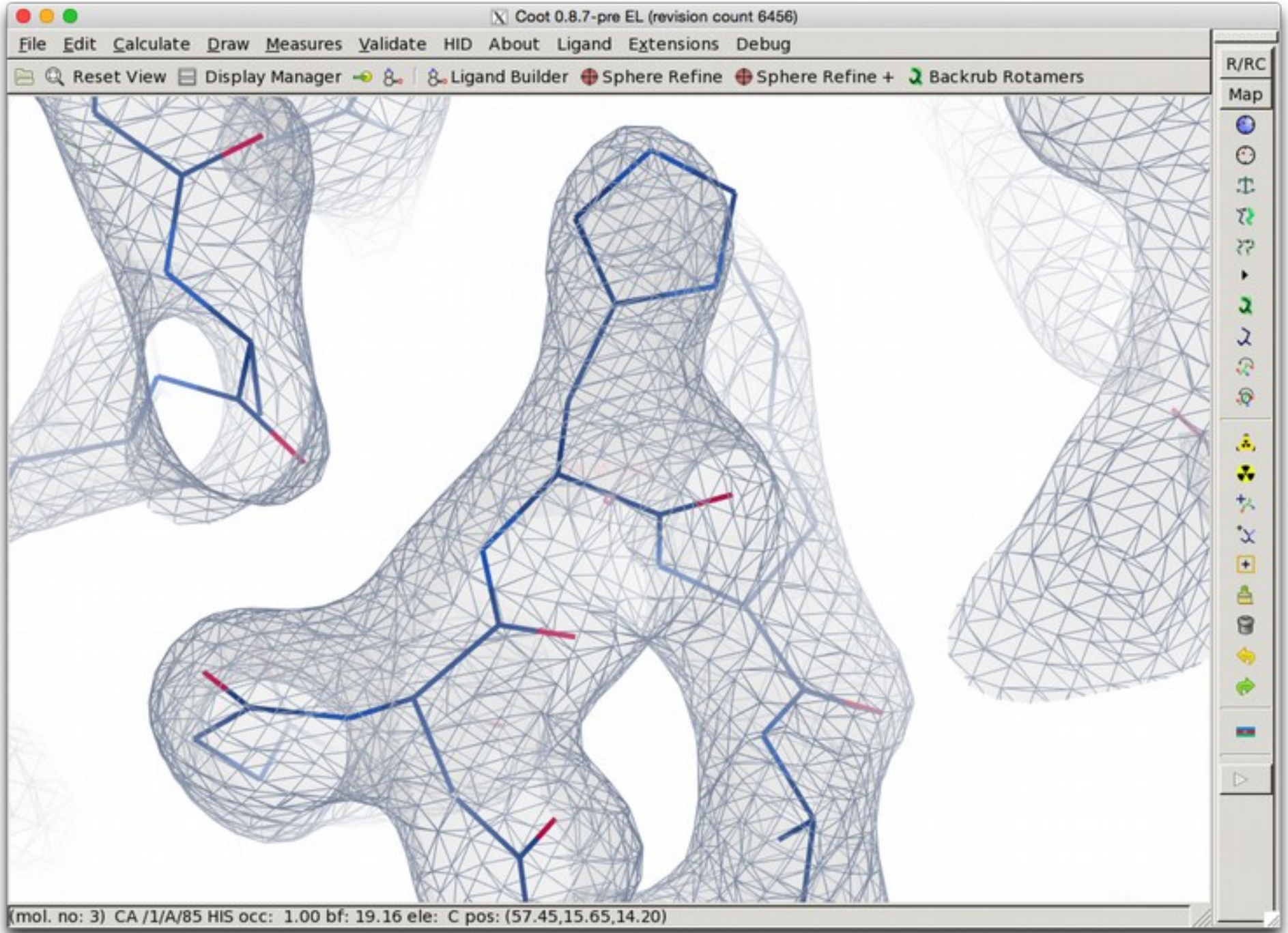


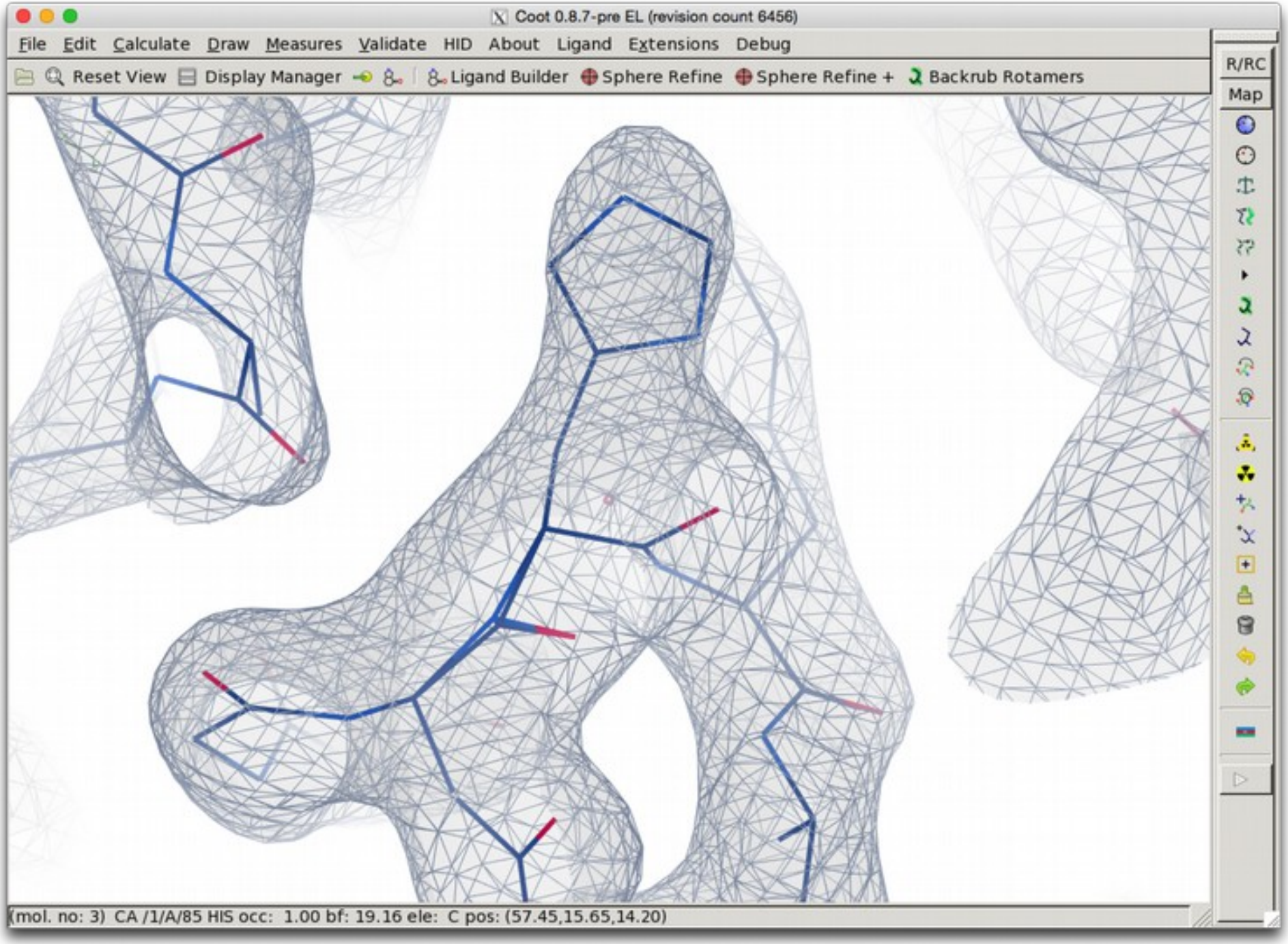


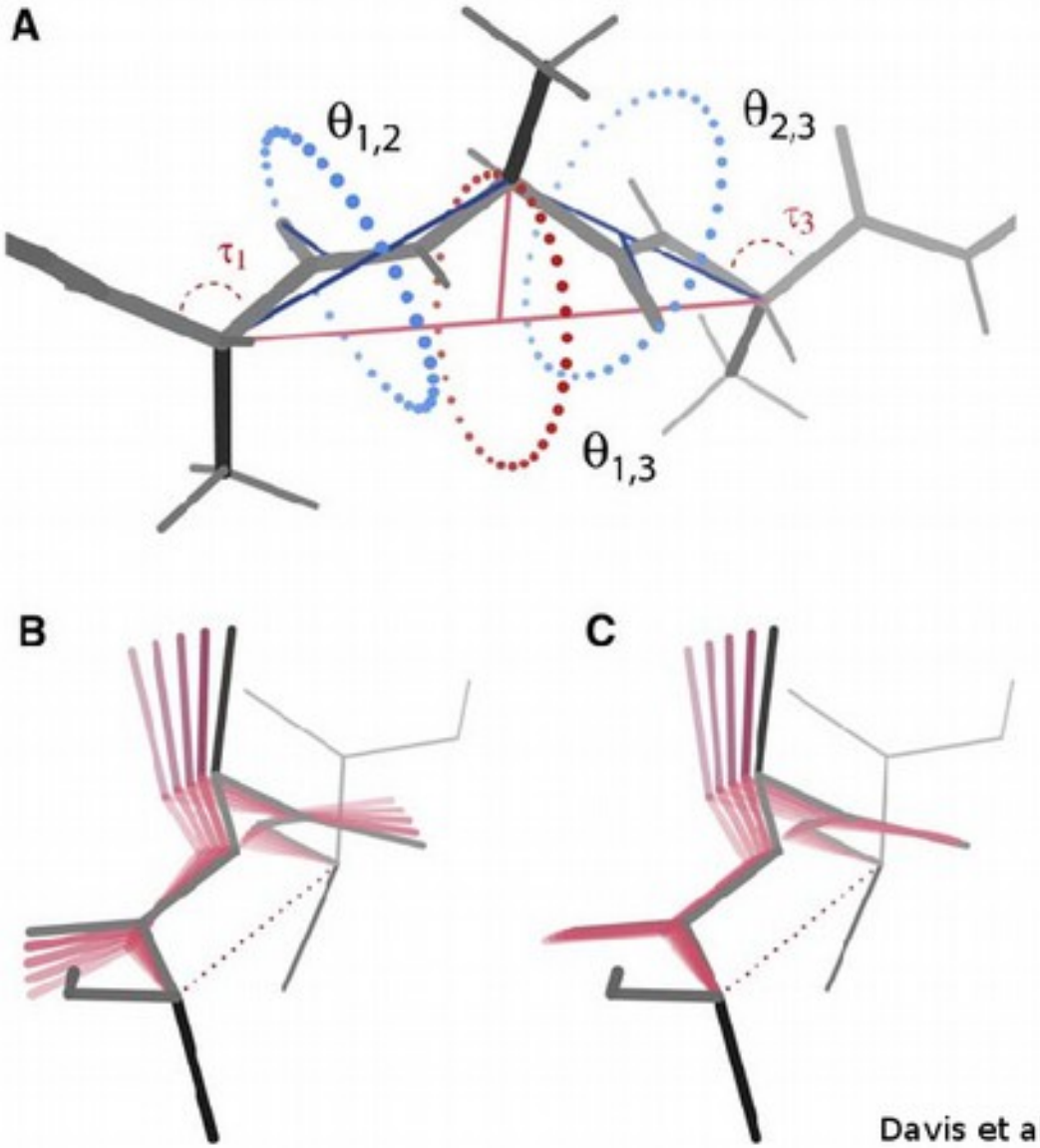
~~Current~~ Low Resolution Rotamer Search

Previous



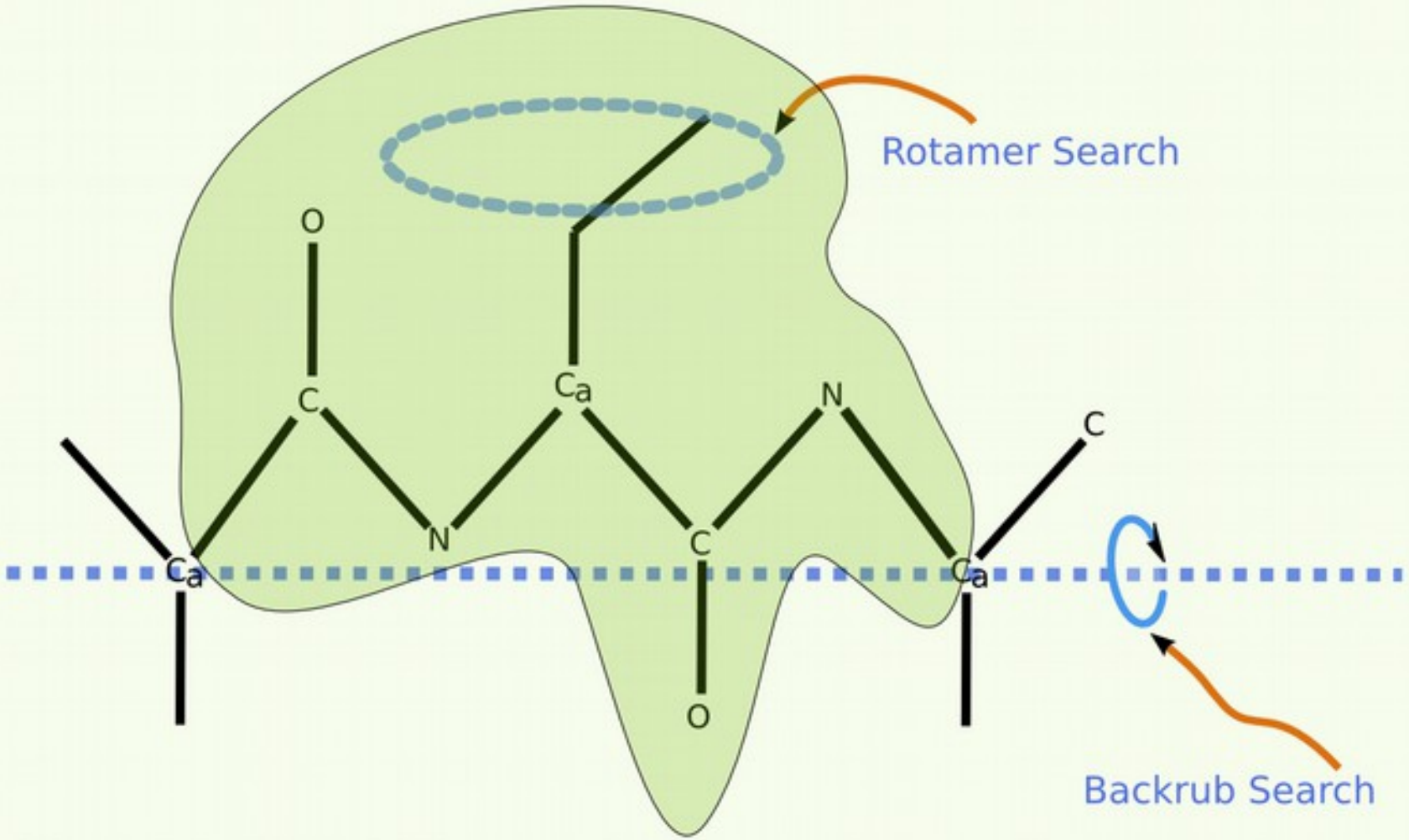

Coot 0.8.7-pre EL (revision count 6456)
File Edit Calculate Draw Measures Validate HID About Ligand Extensions Debug
Reset View
Display Manager
Ligand Builder
Sphere Refine
Sphere Refine +
Backrub Rotamers
R/RC
Map
(mol. no: 3) CA /1/A/85 HIS occ: 1.00 bf: 19.16 ele: C pos: (57.45,15.65,14.20)


Coot 0.8.7-pre EL (revision count 6456)
File Edit Calculate Draw Measures Validate HID About Ligand Extensions Debug
Reset View
Display Manager
Ligand Builder
Sphere Refine
Sphere Refine +
Backrub Rotamers
R/RC
Map
(mol. no: 3) CA /1/A/85 HIS occ: 1.00 bf: 19.16 ele: C pos: (57.45,15.65,14.20)

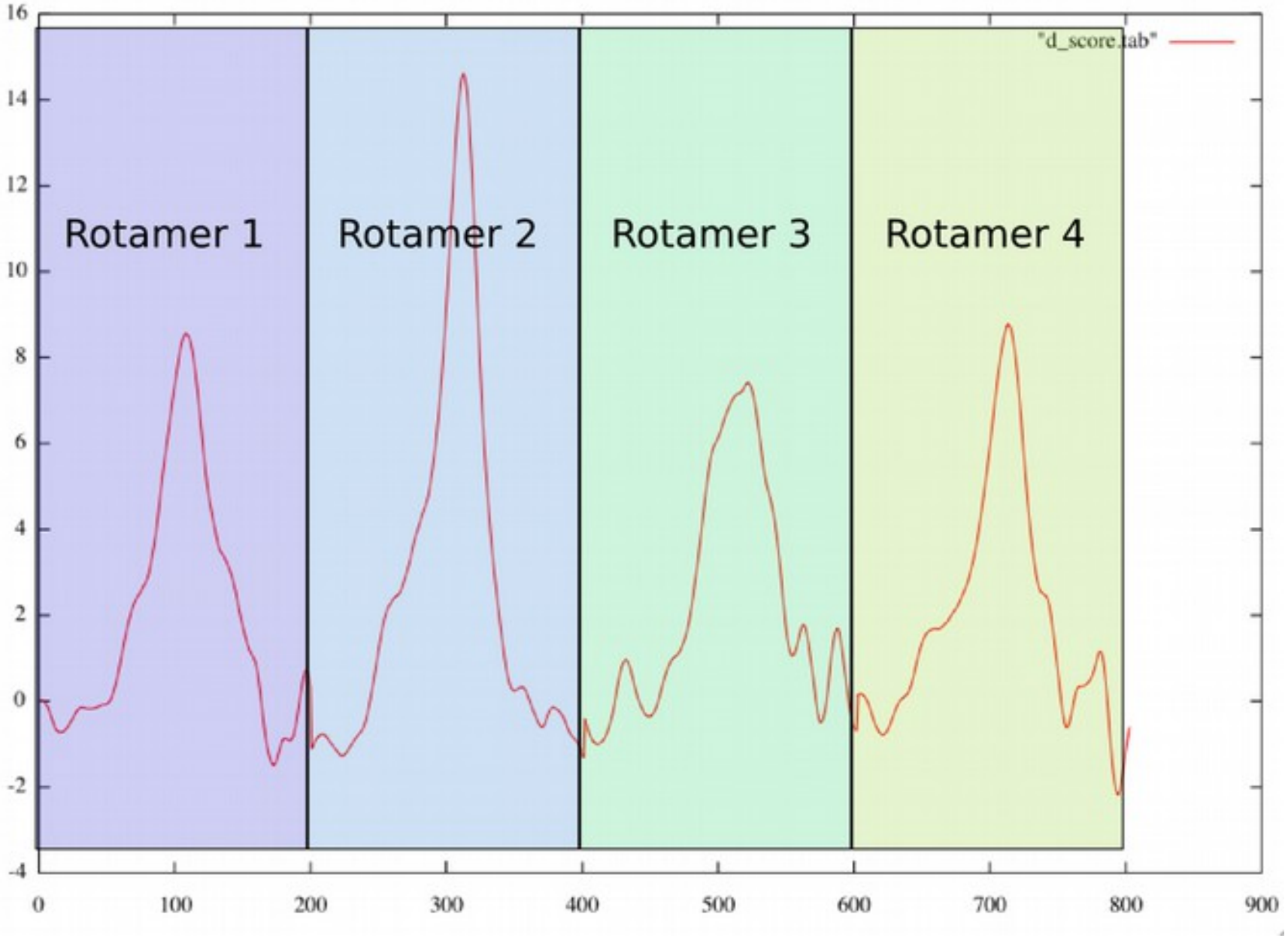


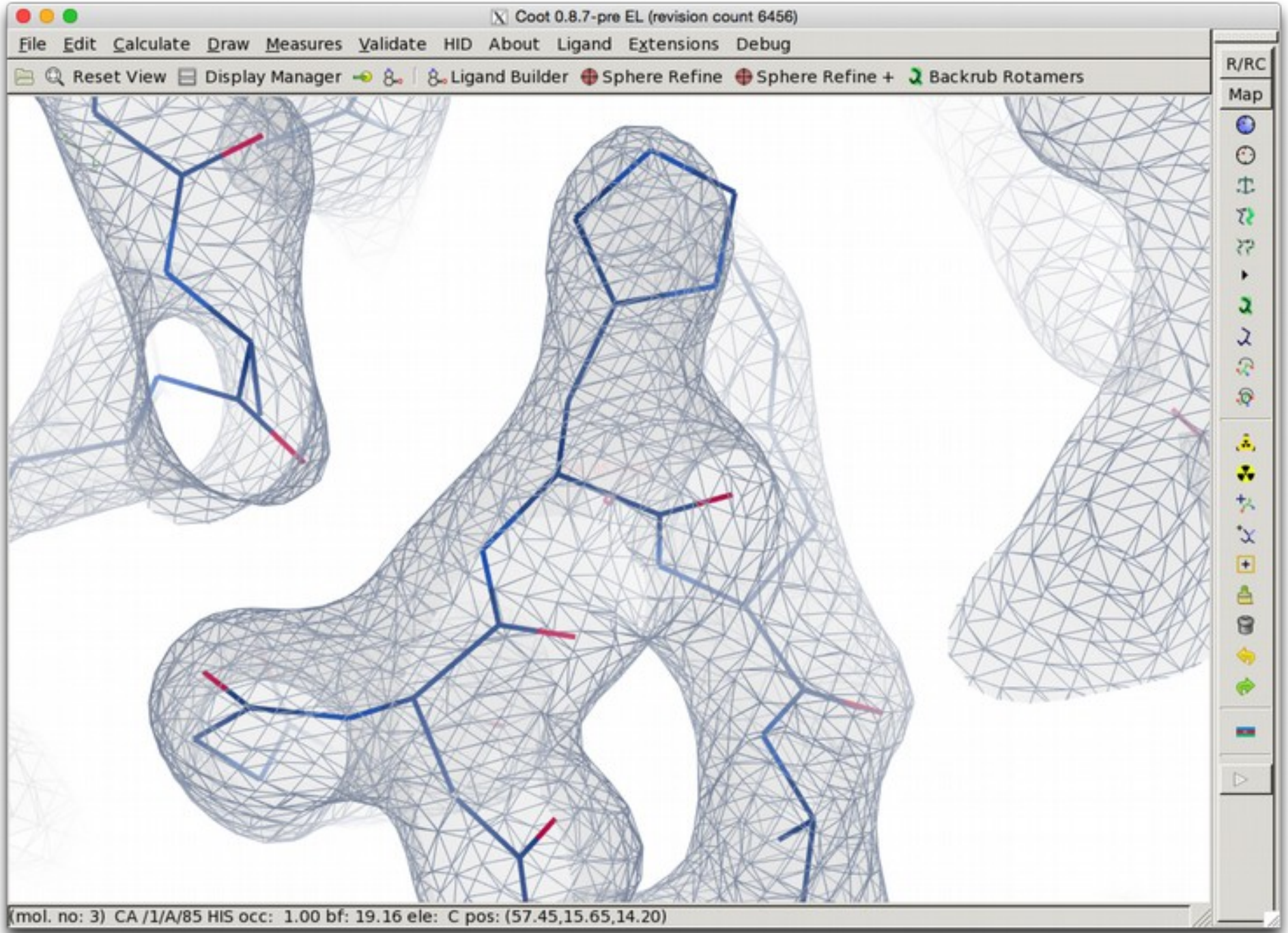
Davis et al. (2006) Structure

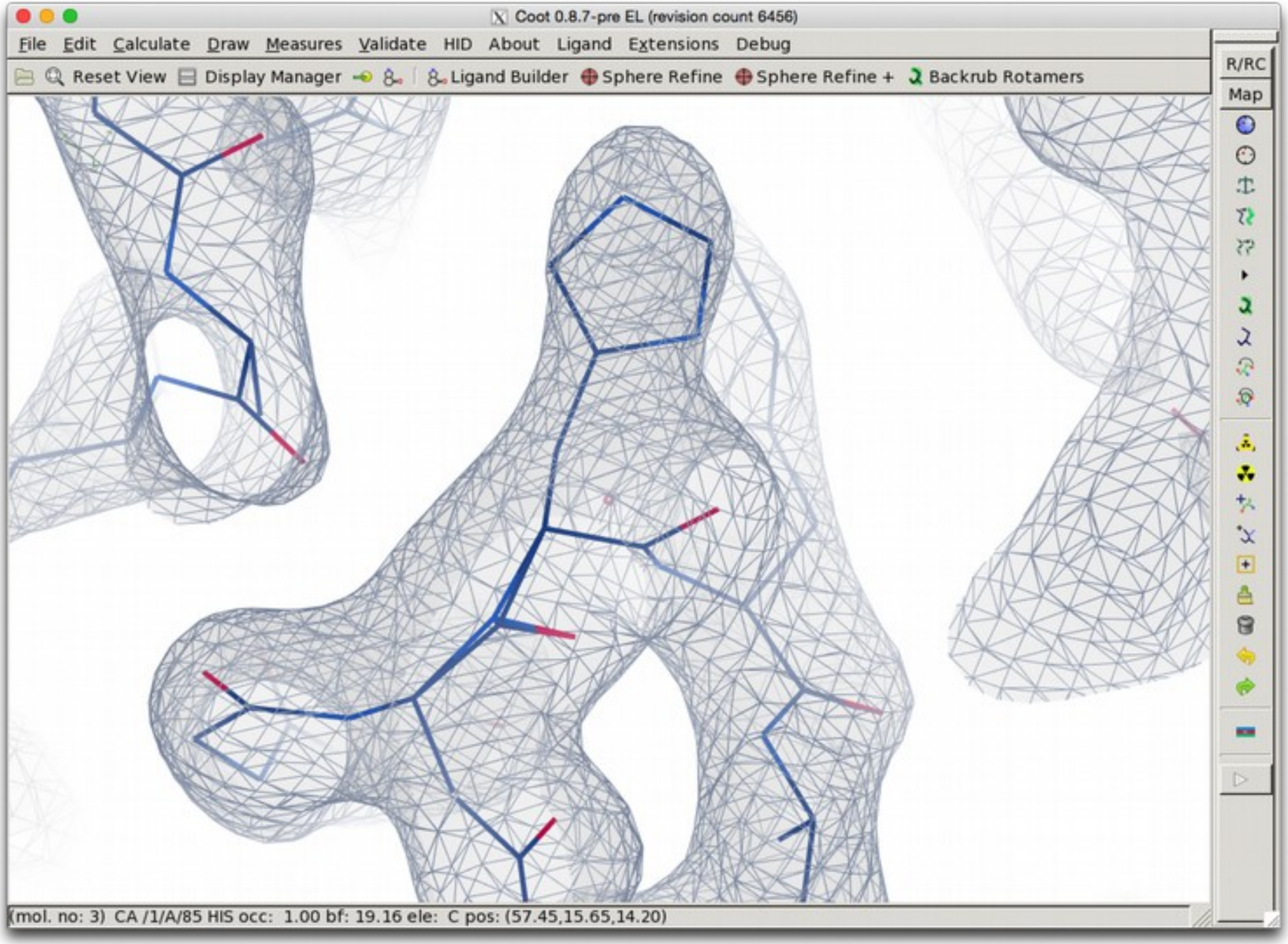
New Low Resolution Rotamer Search

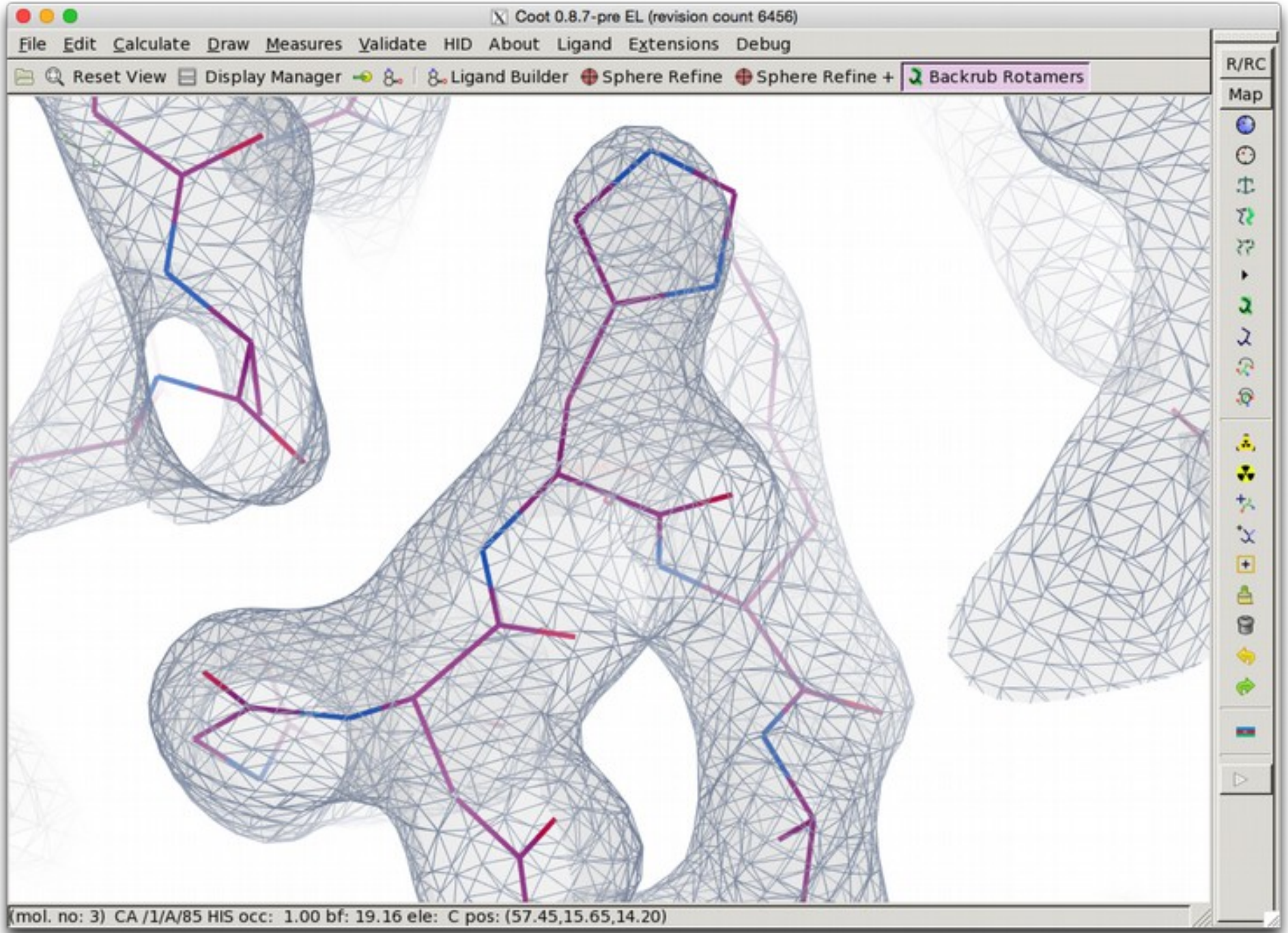


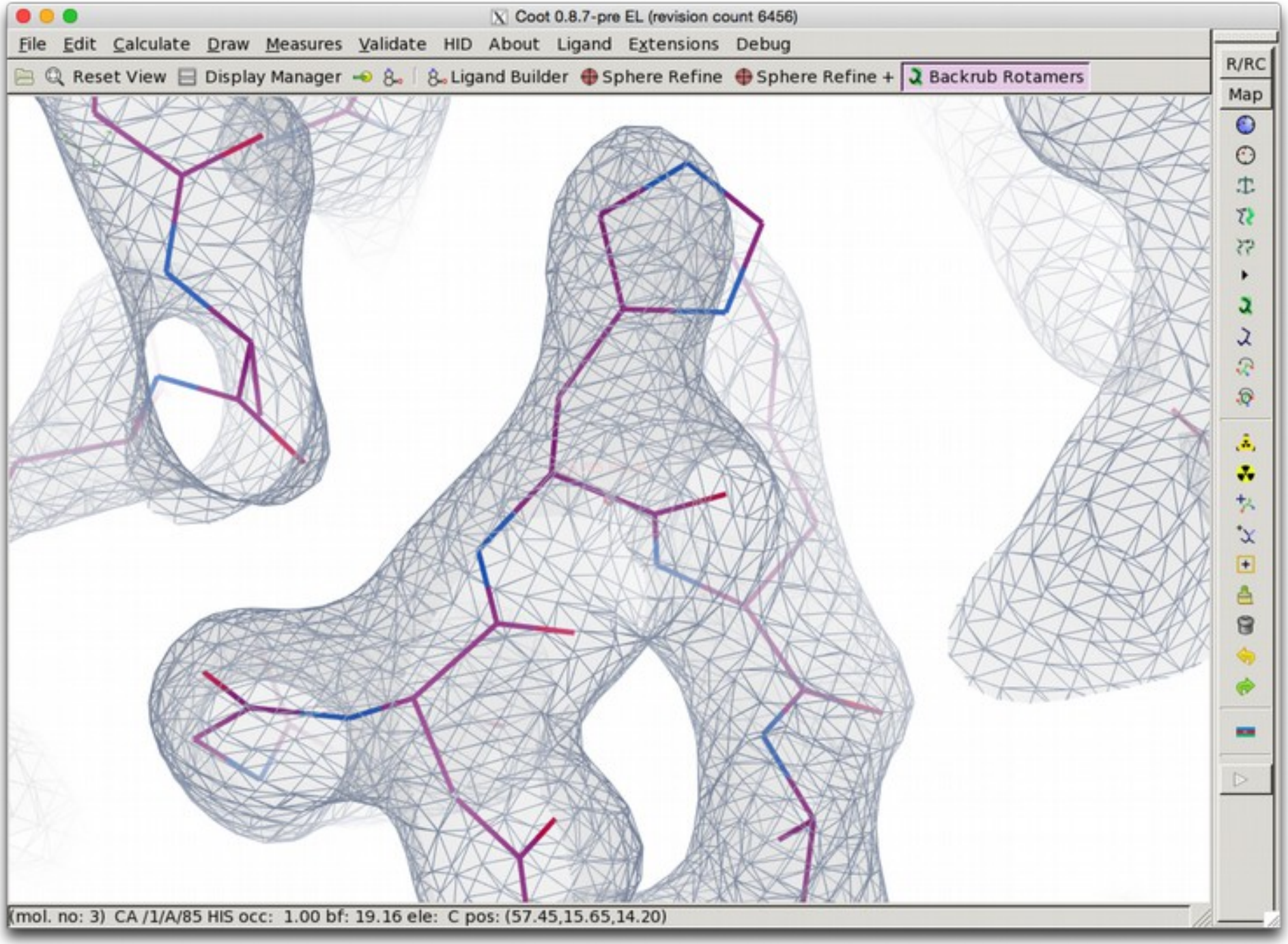
After Fitting Tools in KING/Molprobity

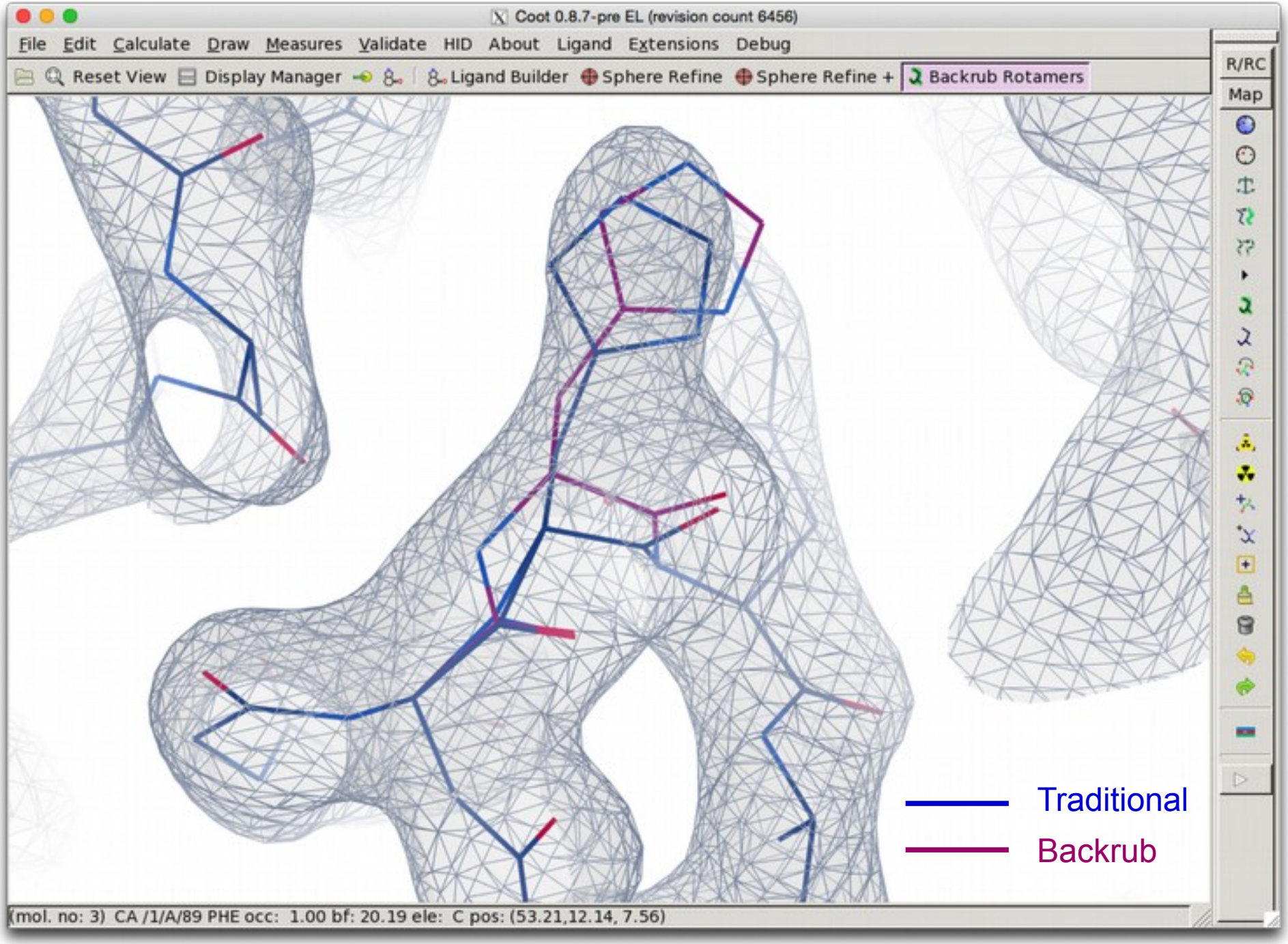

"d_score.tab"
Rotamer 1
Rotamer 2
Rotamer 3
Rotamer 4
16
14
12
10
8
6
4
2
0
-2
-4
0
100
200
300
400
500
600
700
800
900


Coot 0.8.7-pre EL (revision count 6456)
File Edit Calculate Draw Measures Validate HID About Ligand Extensions Debug
Reset View
Display Manager
Ligand Builder
Sphere Refine
Sphere Refine +
Backrub Rotamers
R/RC
Map
(mol. no: 3) CA /1/A/85 HIS occ: 1.00 bf: 19.16 ele: C pos: (57.45,15.65,14.20)

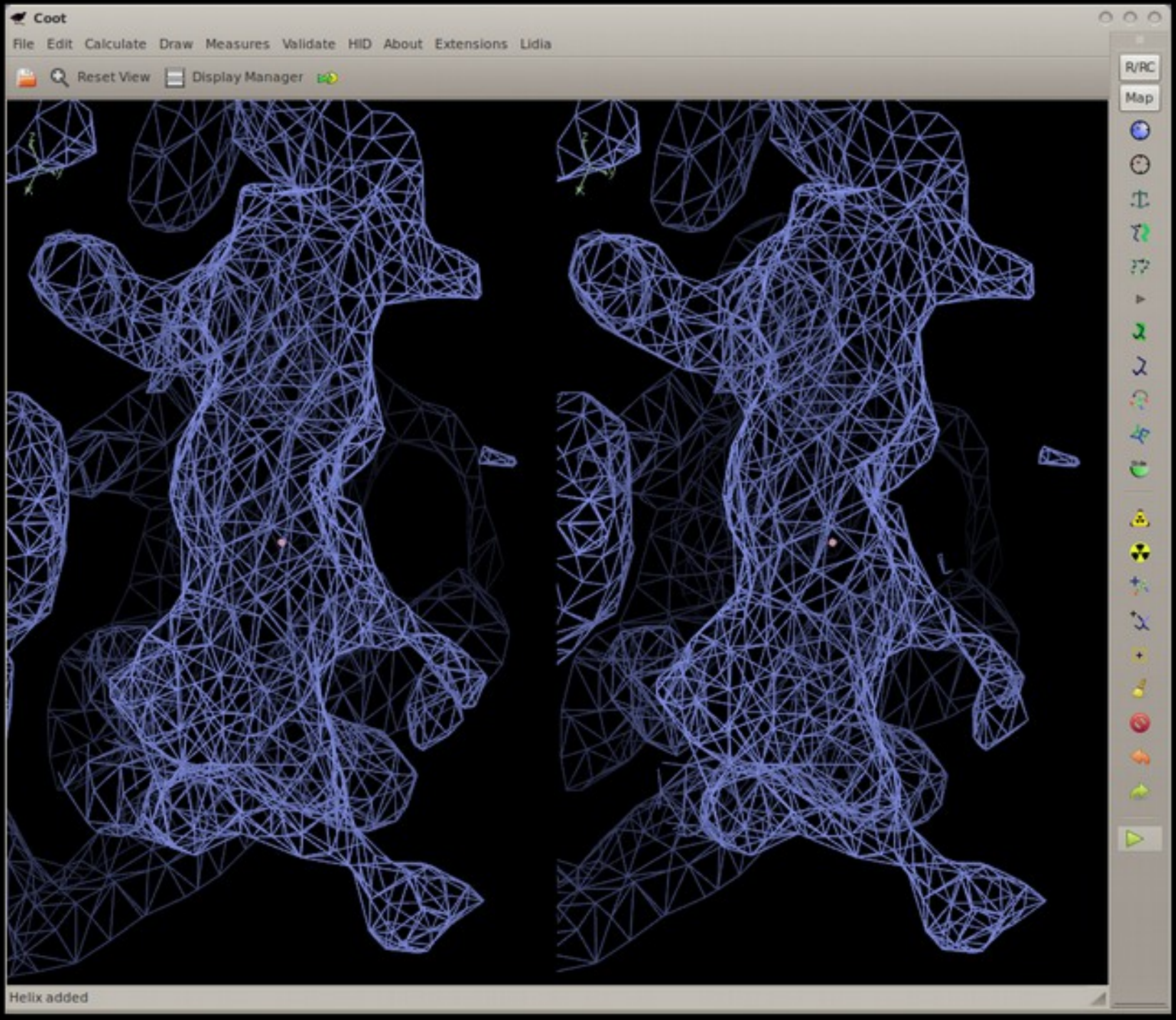

Coot 0.8.7-pre EL (revision count 6456)
File Edit Calculate Draw Measures Validate HID About Ligand Extensions Debug
Reset View
Display Manager
Ligand Builder
Sphere Refine
Sphere Refine +
Backrub Rotamers
R/RC
Map
(mol. no: 3) CA /1/A/85 HIS occ: 1.00 bf: 19.16 ele: C pos: (57.45,15.65,14.20)


Coot 0.8.7-pre EL (revision count 6456)
File Edit Calculate Draw Measures Validate HID About Ligand Extensions Debug
Reset View
Display Manager
Ligand Builder
Sphere Refine
Sphere Refine +
Backrub Rotamers
R/RC
Map
(mol. no: 3) CA /1/A/85 HIS occ: 1.00 bf: 19.16 ele: C pos: (57.45,15.65,14.20)


Coot 0.8.7-pre EL (revision count 6456)
File Edit Calculate Draw Measures Validate HID About Ligand Extensions Debug
Reset View
Display Manager
Ligand Builder
Sphere Refine
Sphere Refine +
Backrub Rotamers
R/RC
Map
(mol. no: 3) CA /1/A/85 HIS occ: 1.00 bf: 19.16 ele: C pos: (57.45,15.65,14.20)


Coot 0.8.7-pre EL (revision count 6456)
File Edit Calculate Draw Measures Validate HID About Ligand Extensions Debug
Reset View
Display Manager
Ligand Builder
Sphere Refine
Sphere Refine +
Backrub Rotamers
R/RC
Map
Traditional
Backrub
(mol. no: 3) CA /1/A/89 PHE occ: 1.00 bf: 20.19 ele: C pos: (53.21,12.14, 7.56)

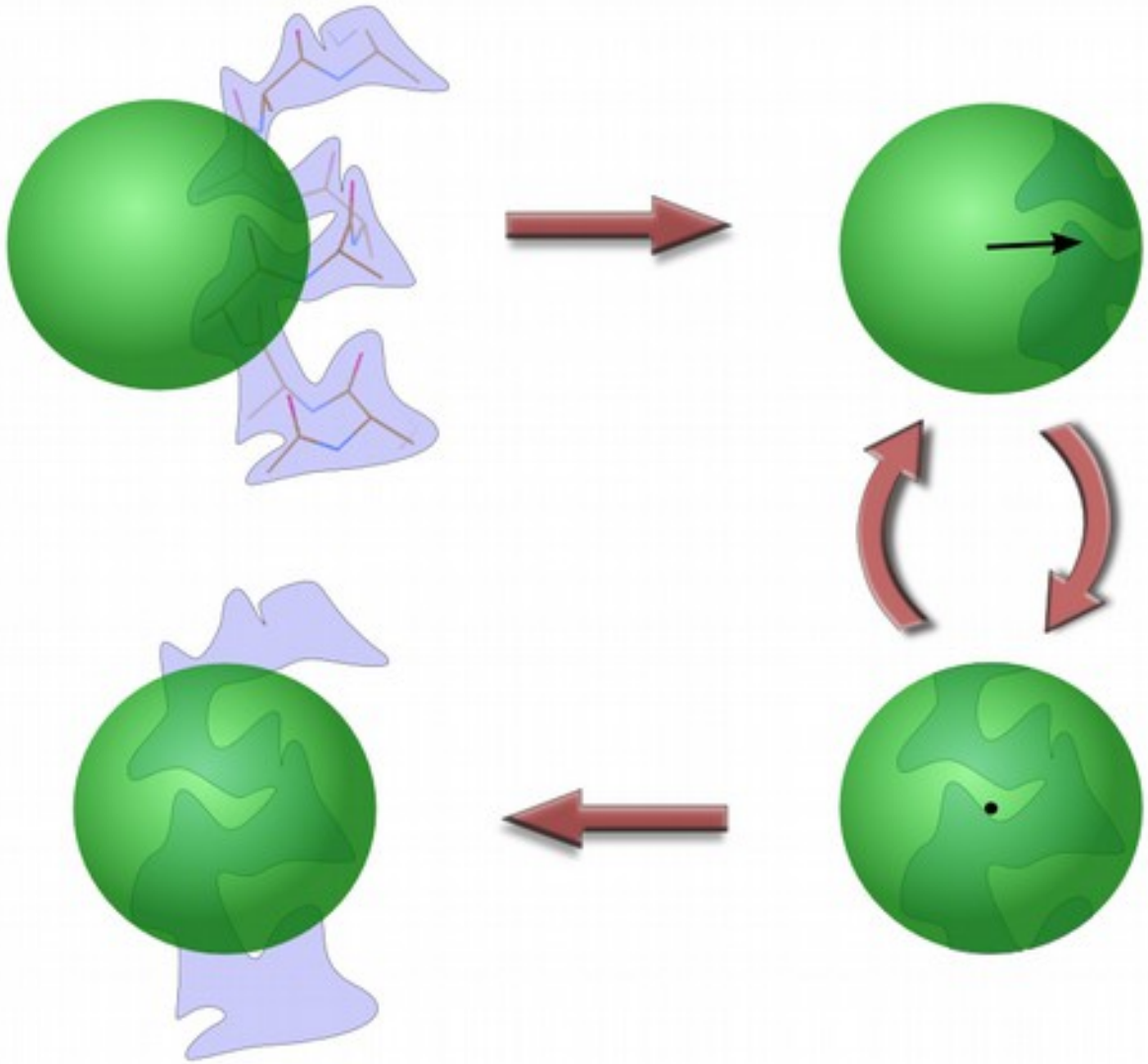
Helix-Building


Coot
File Edit Calculate Draw Measures Validate HID About Extensions Lidia
Reset View
Display Manager
R/RC
Map
Helix added

Alpha Helix Placement

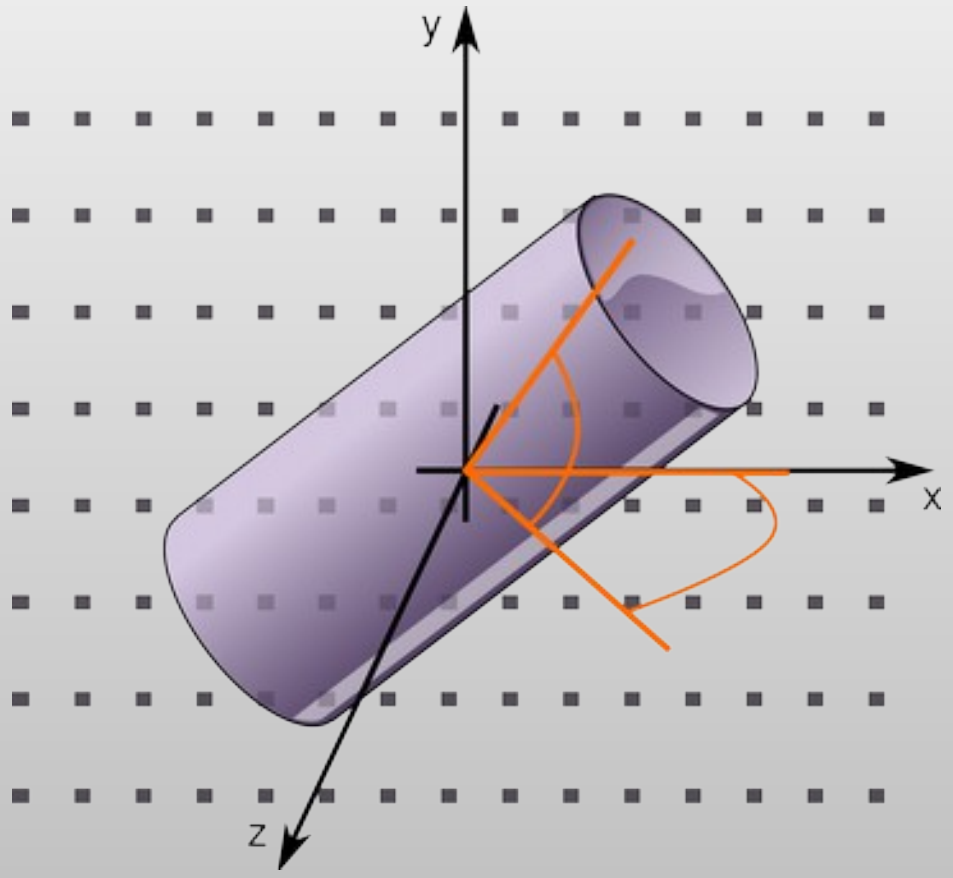
- Scenario: Looking at a new map, not built with automatic tools:
  - "I can see that there's a helix here - build it for me!"
- From a given point:
  - Move to local averaged maximum
  - Do a 2D MR-style orientation search on a cylinder of electron density
  - Build a helix (both directions)
  - 1D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

Centering the Rotation point

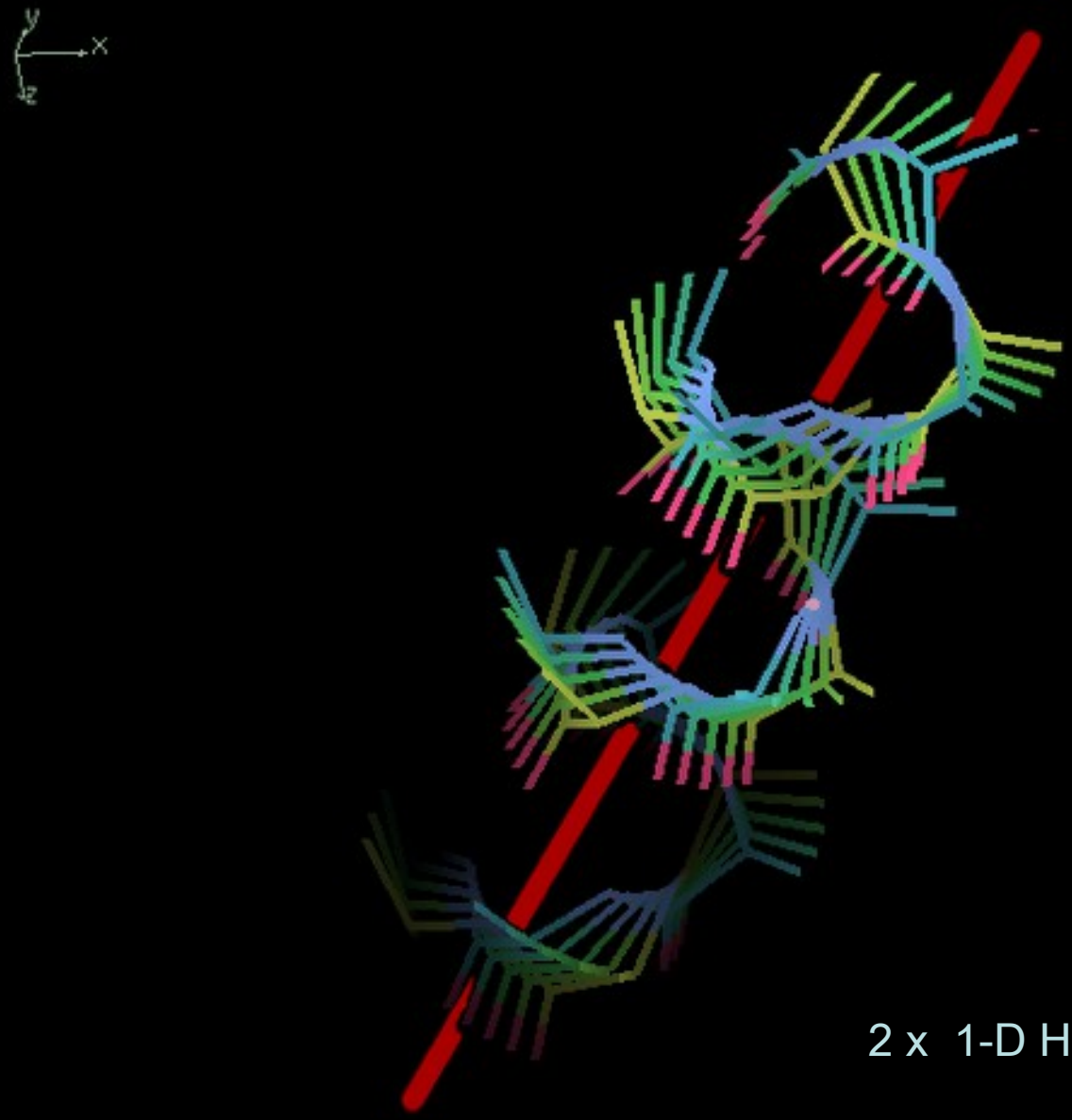


Helix Fitting: Cylinder Search

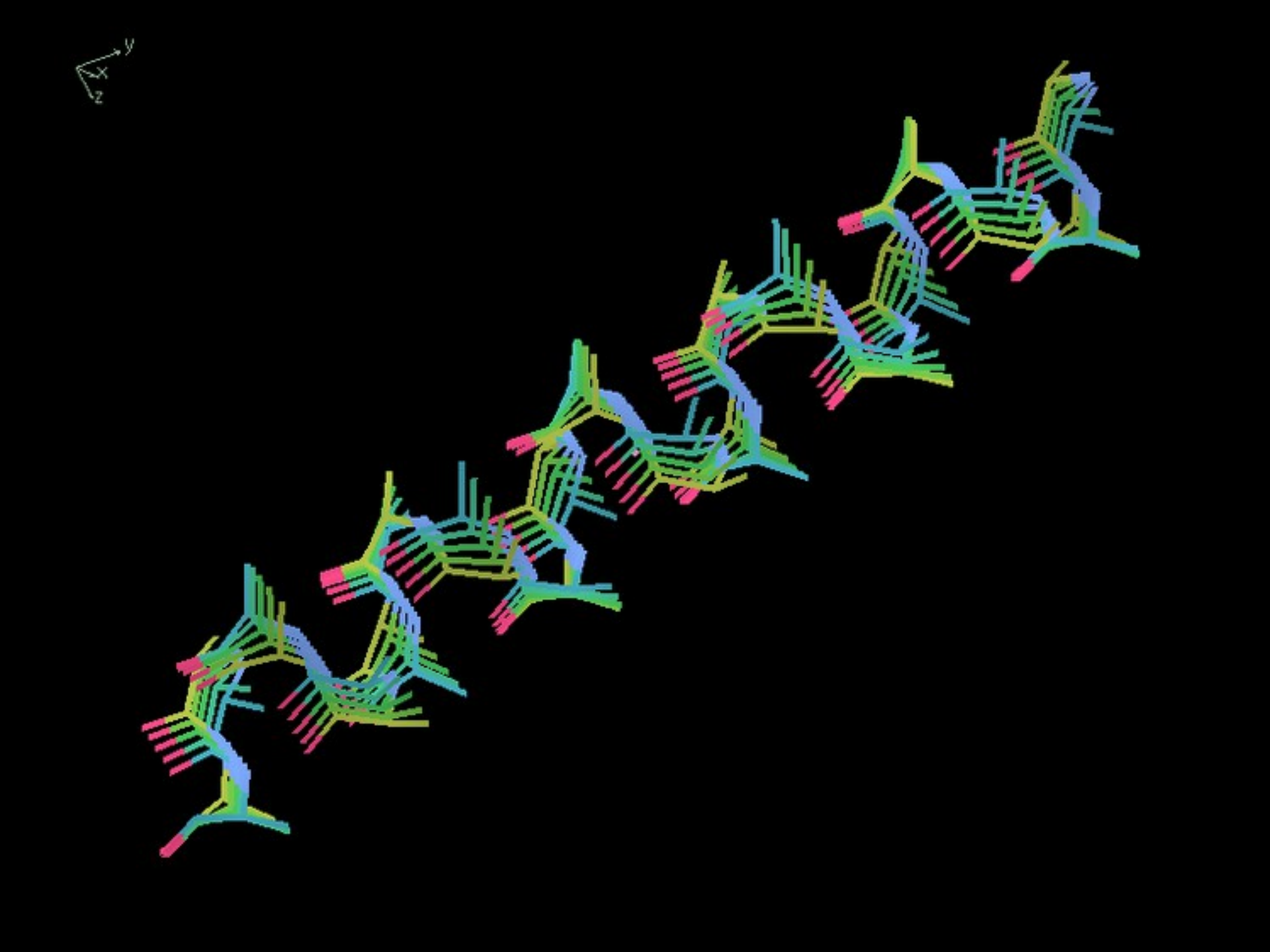
- Pick the orientation that encapsulates the most electron density

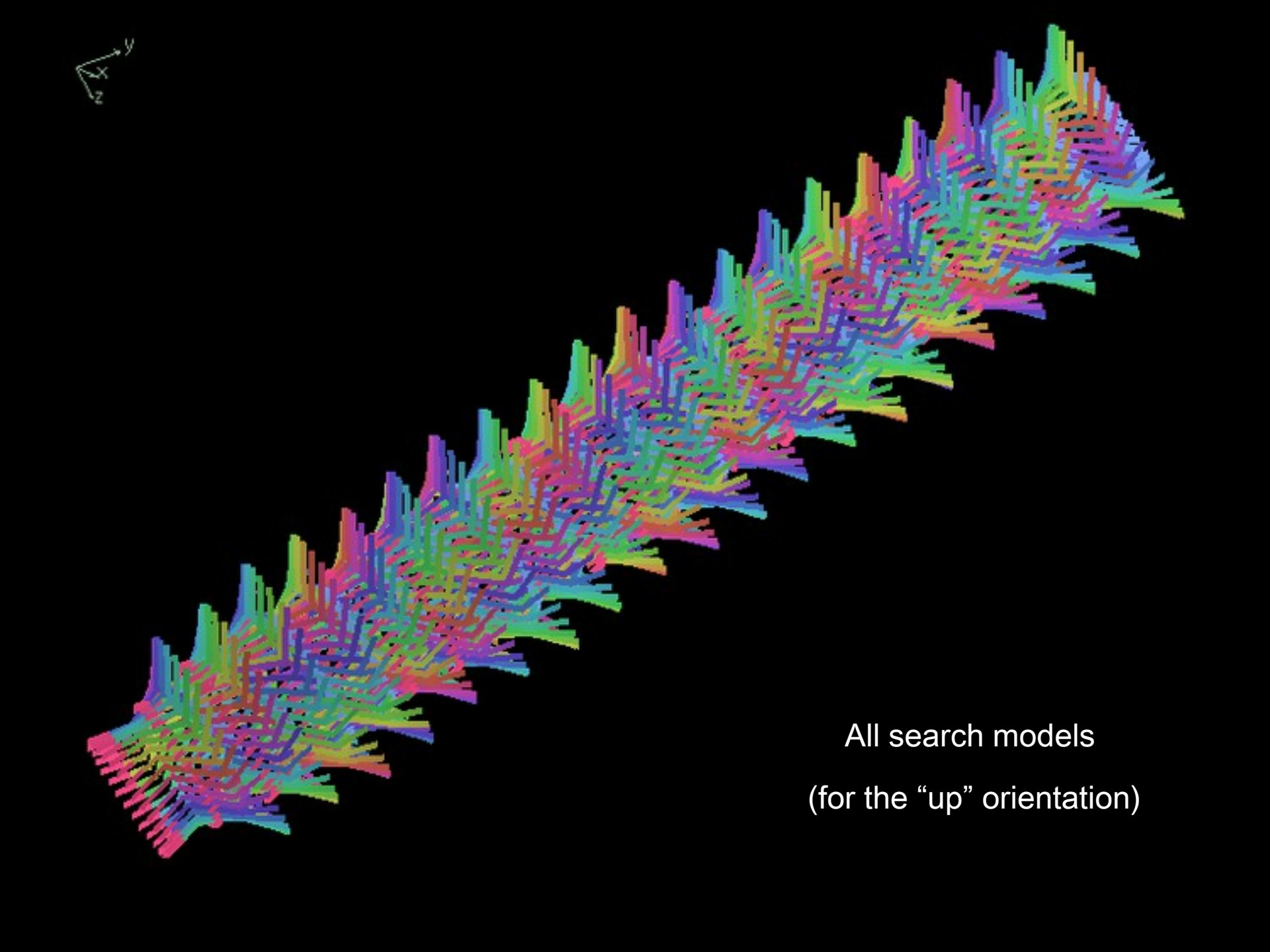


Using 2 rotation axes



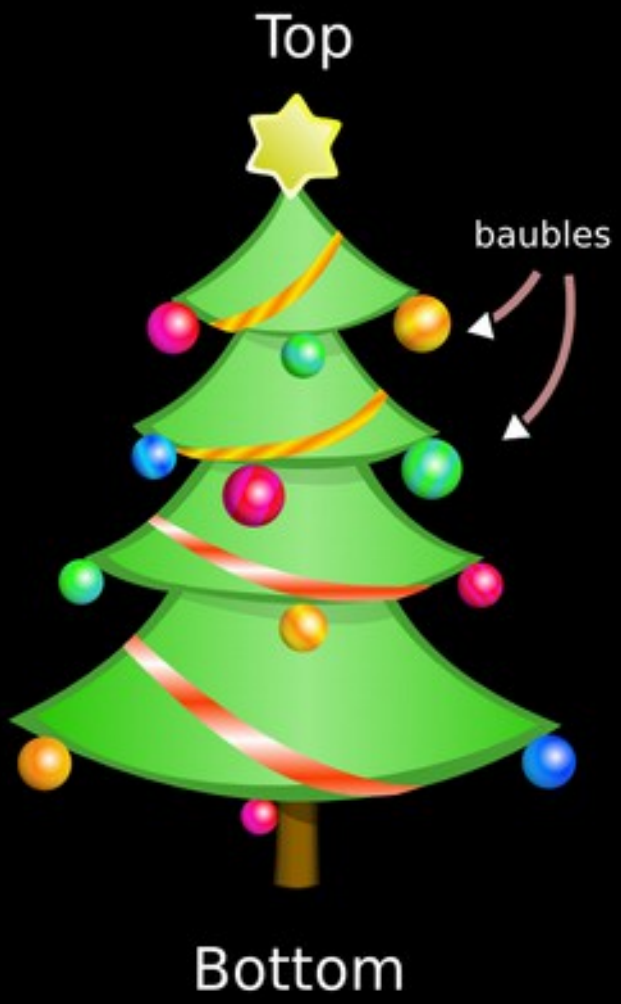
2 x 1-D Helix orientation searches


y
x
z

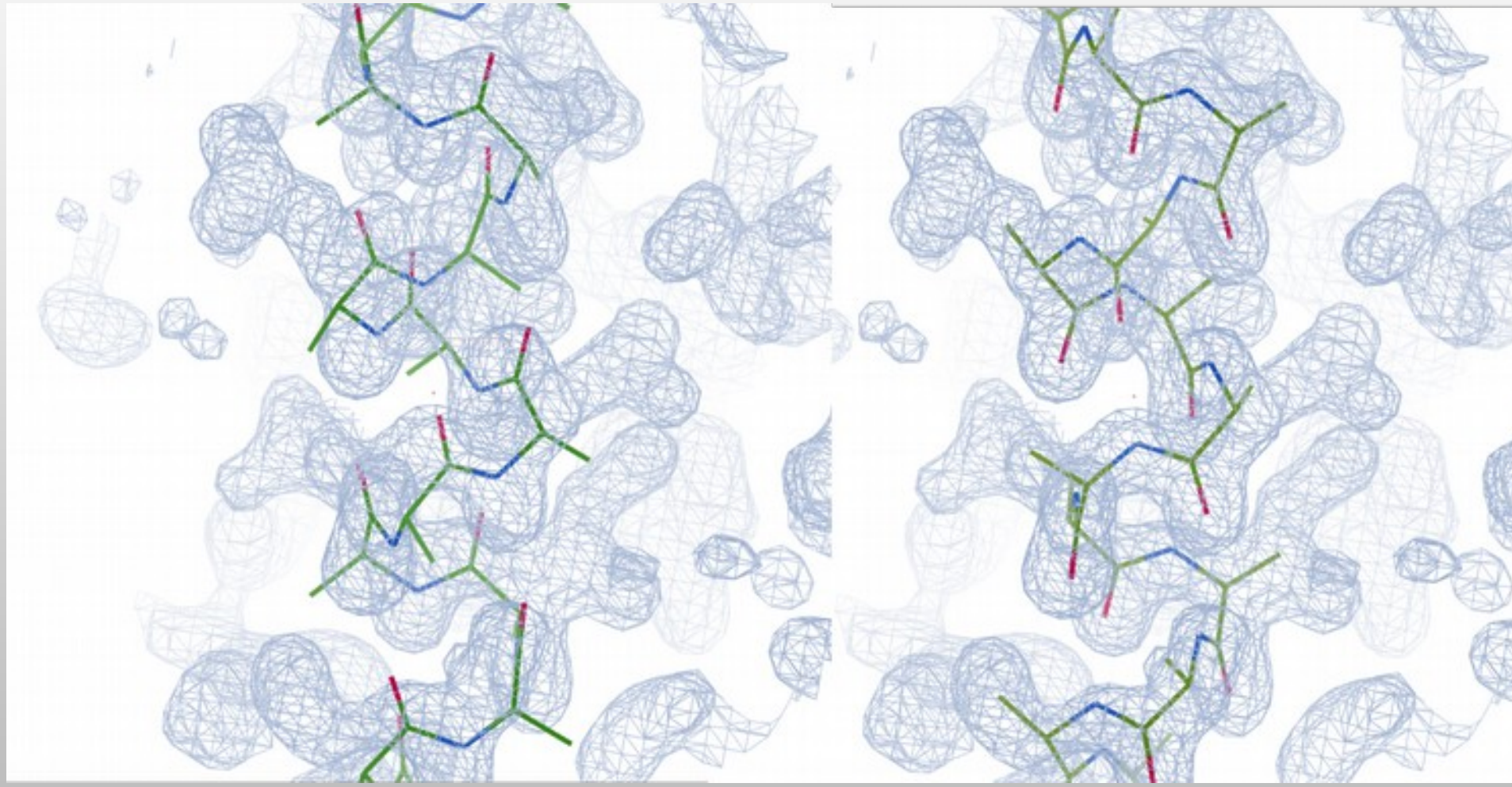


All search models

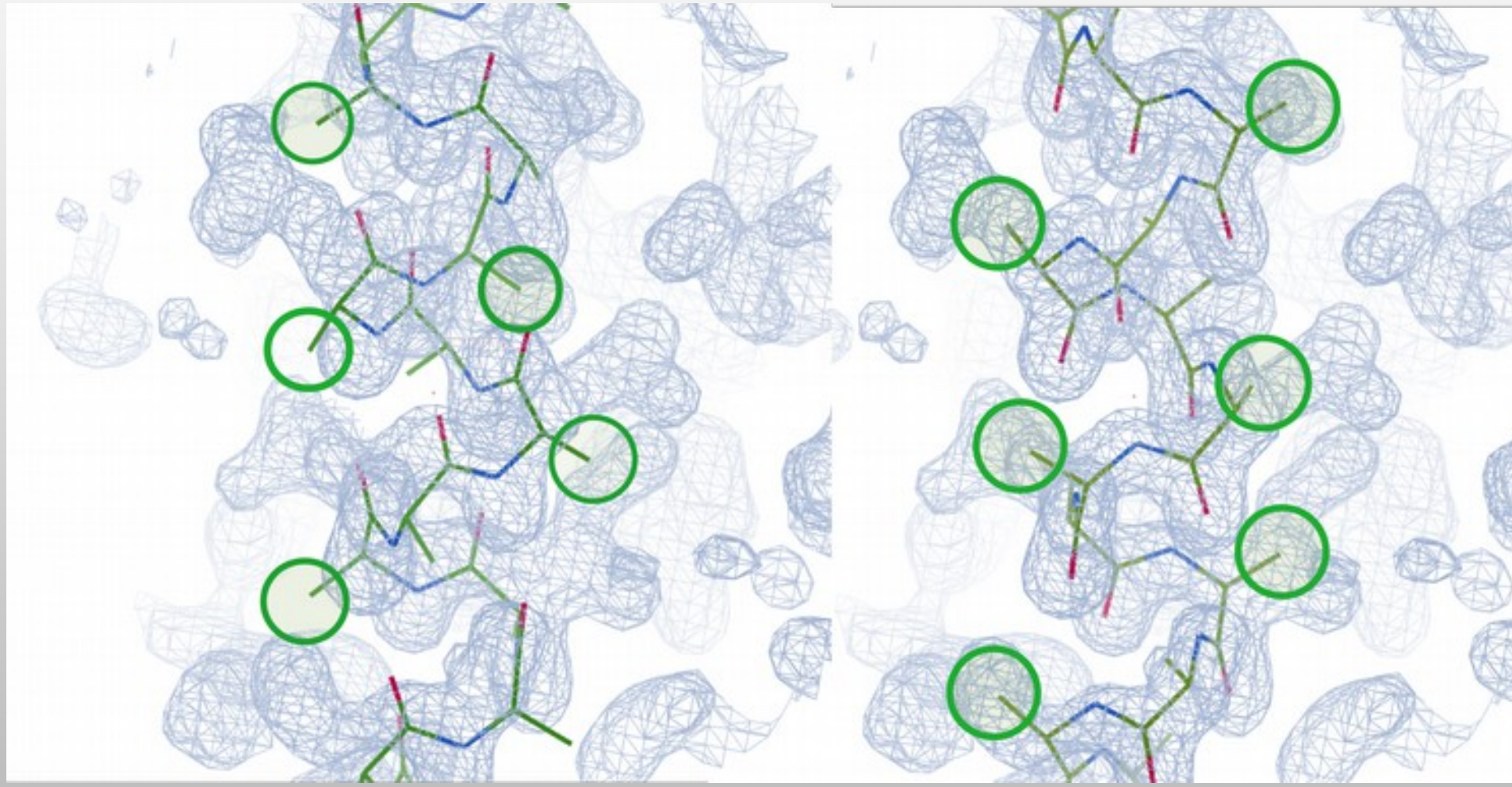
(for the “up” orientation)


Top
baubles
Bottom

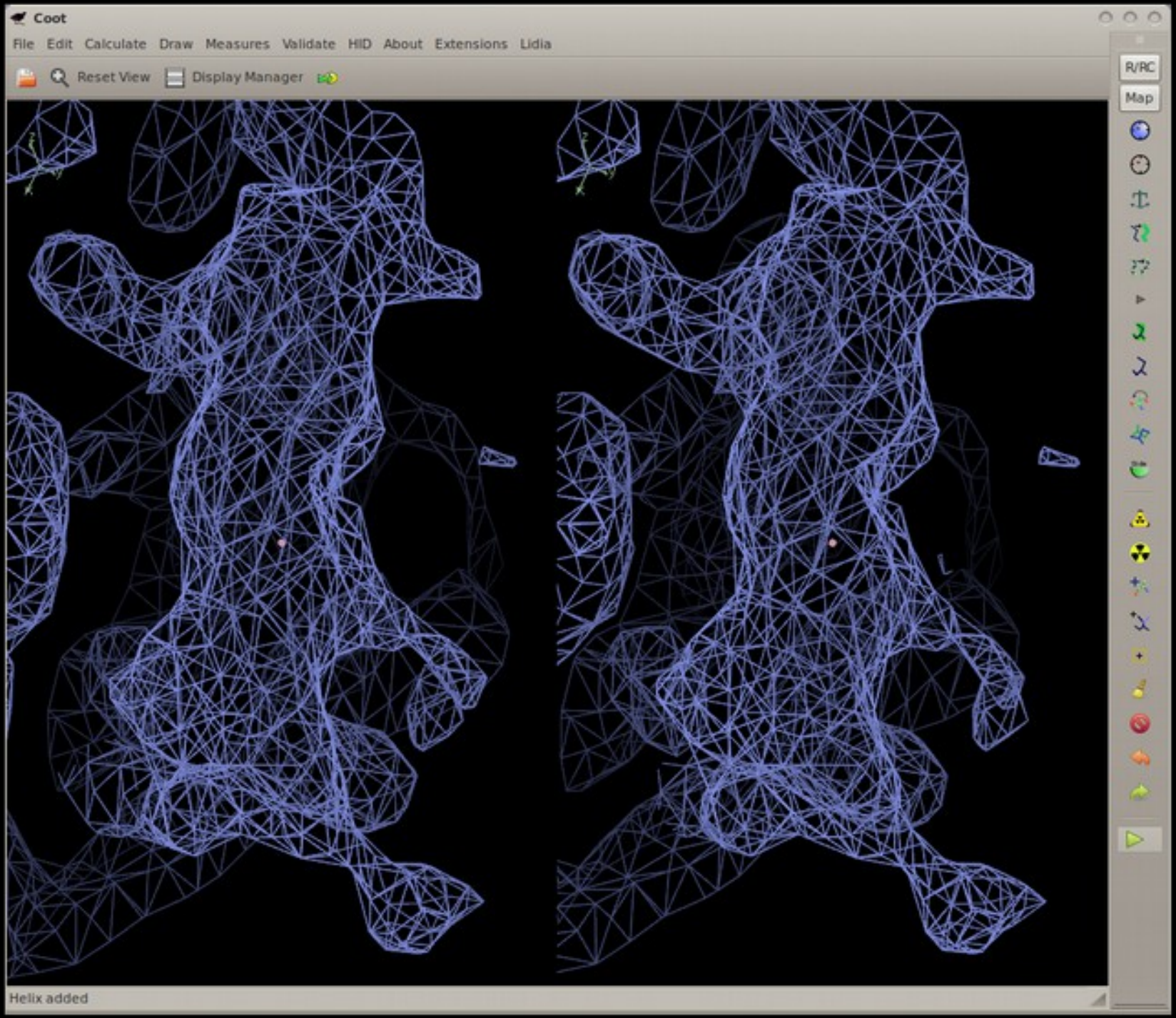
Helix Fitting
Comparing orientation hypotheses

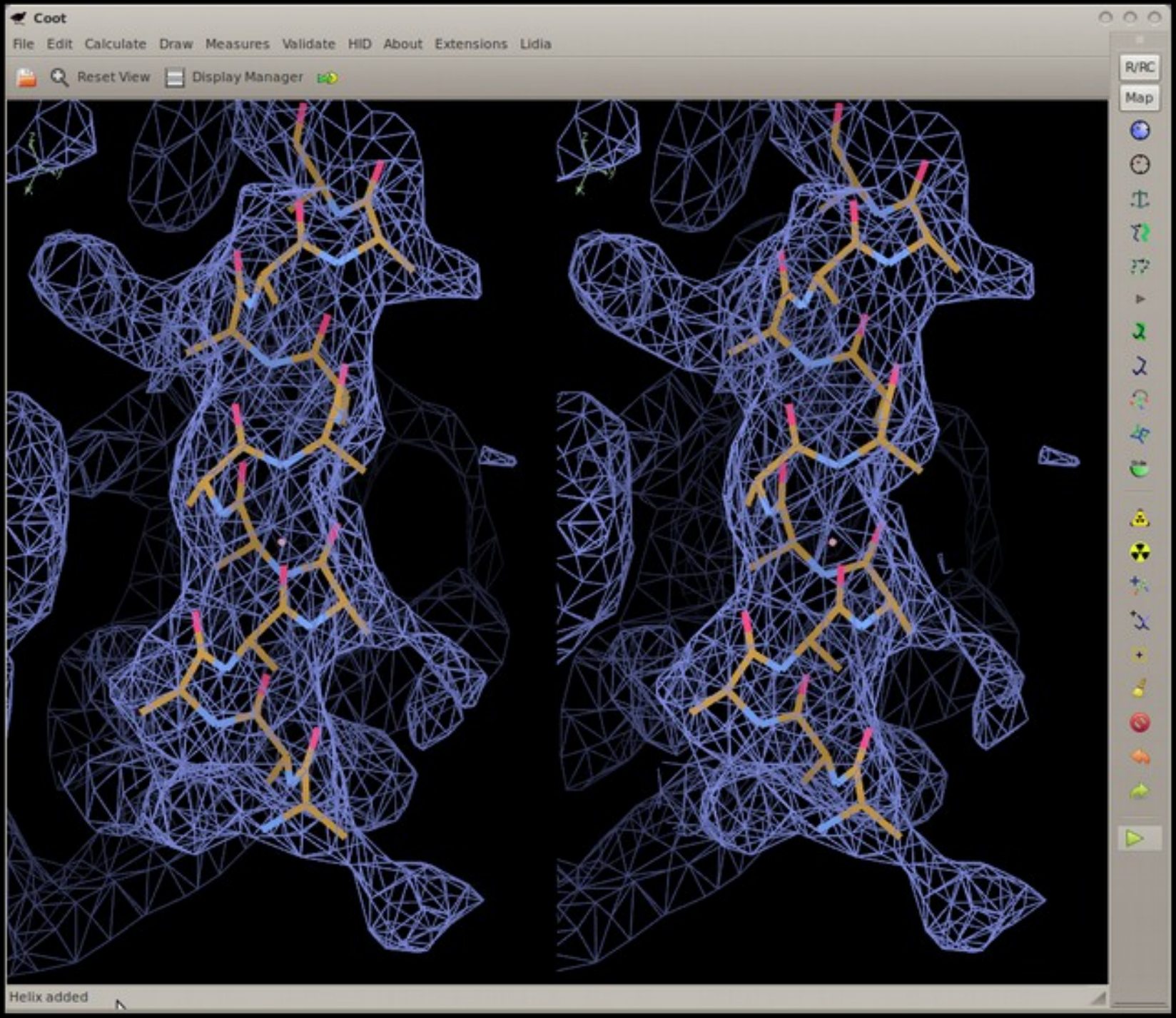


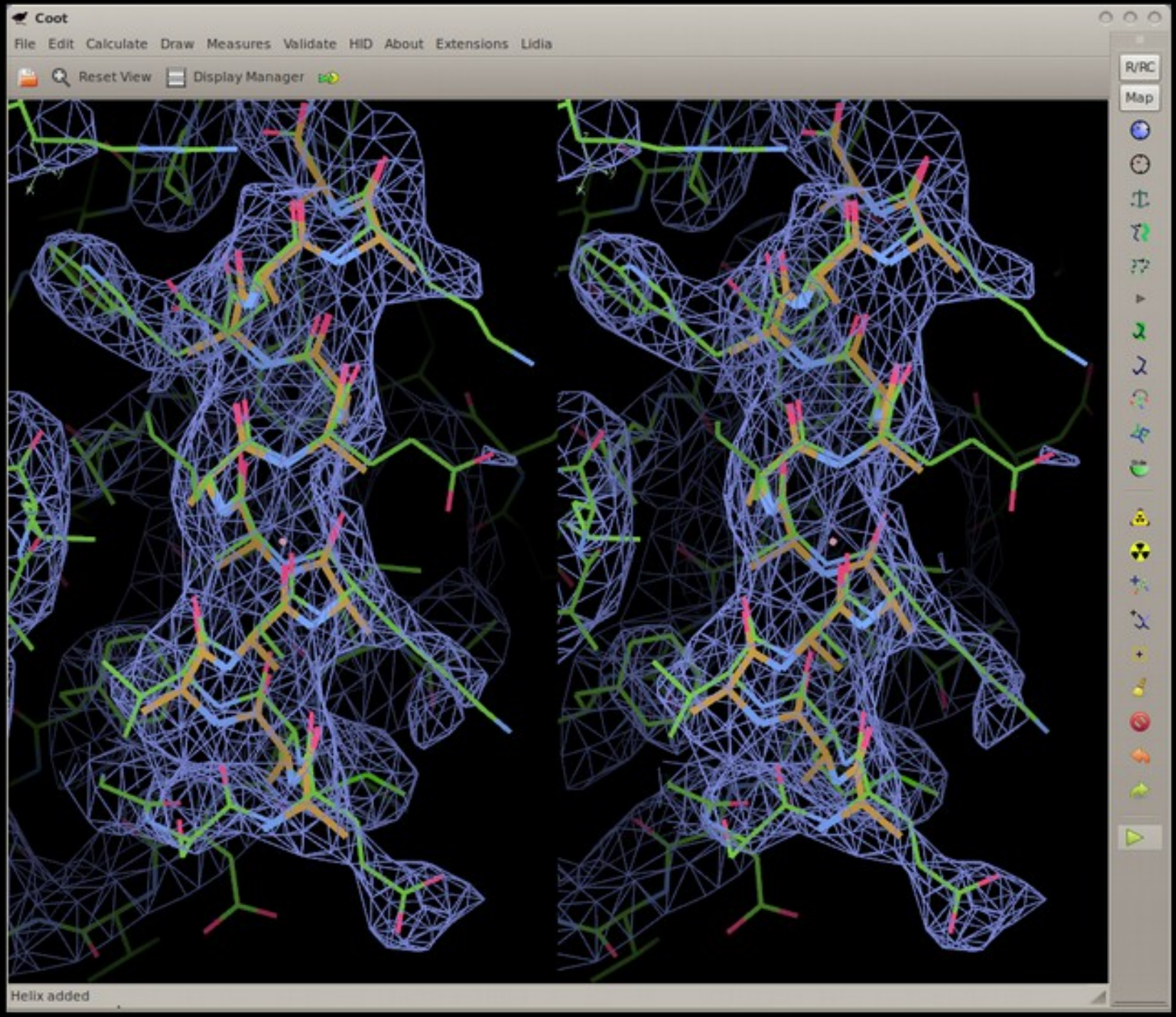
Helix Fitting
Comparing orientation hypotheses



c-betas not fitting and are used for scoring


Coot
File Edit Calculate Draw Measures Validate HID About Extensions Lidia
Reset View
Display Manager
R/RC
Map
Helix added


Coot
File Edit Calculate Draw Measures Validate HID About Extensions Lidia
Reset View
Display Manager
R/RC
Map
Helix added


Coot
File Edit Calculate Draw Measures Validate HID About Extensions Lidia
Reset View
Display Manager
R/RC
Map
Helix added

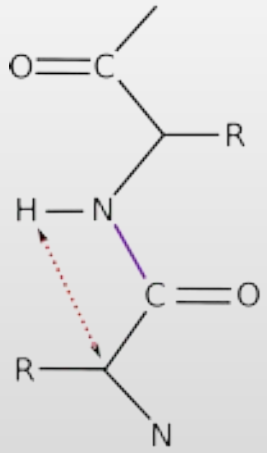
cis-Peptides

- What is a cis-peptide?
- Peptide restraints in Coot 2004-2015

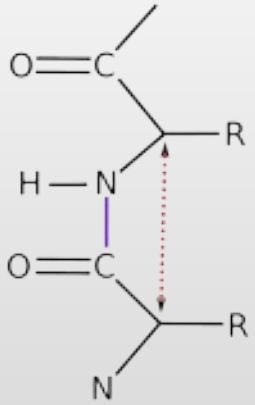
cis-Peptides

- A number of paper have been published recently highlighting the unusually large number of cis-peptides in some structures:
  - Croll: The rate of cis-trans conformation errors is increasing in low-resolution crystal structures *Acta Cryst.* (2015). **D**71, 706-709
  - Touw *et al*.: Detection of trans–cis flips and peptide-plane flips in protein structures *Acta Cryst.* (2015). **D**71, 1604-71614

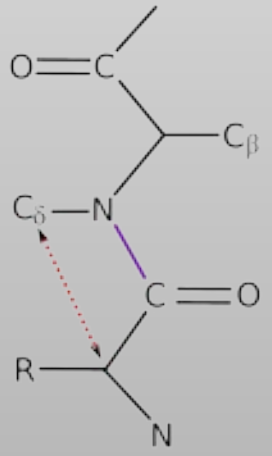
cis-Peptides



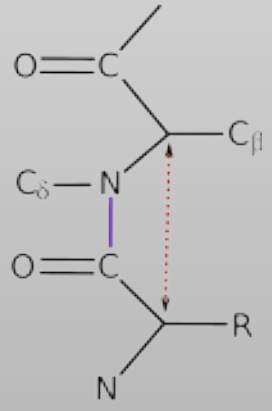
trans-peptide



cis-peptide

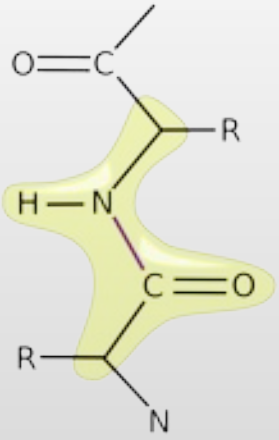


PRO *trans-peptide*

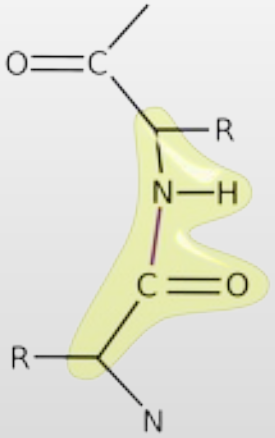


PRO *cis-peptide*

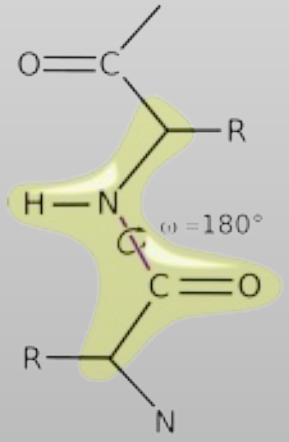
cis-Peptides



trans-peptide
with plane restraints

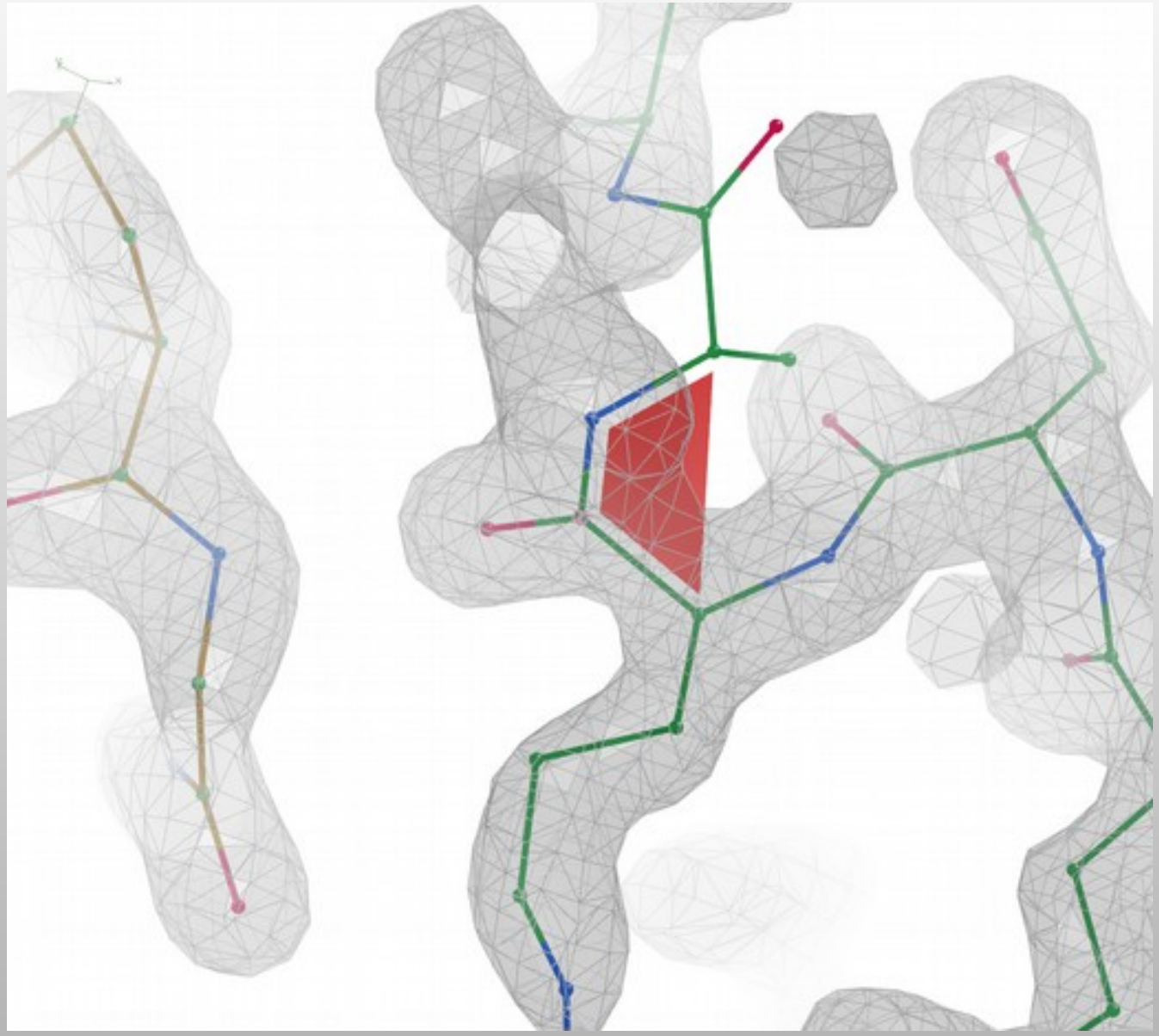


cis-peptide
with plane restraints



trans-peptide
with plane and trans restraints

cis-peptide Representation



Pre-PRO

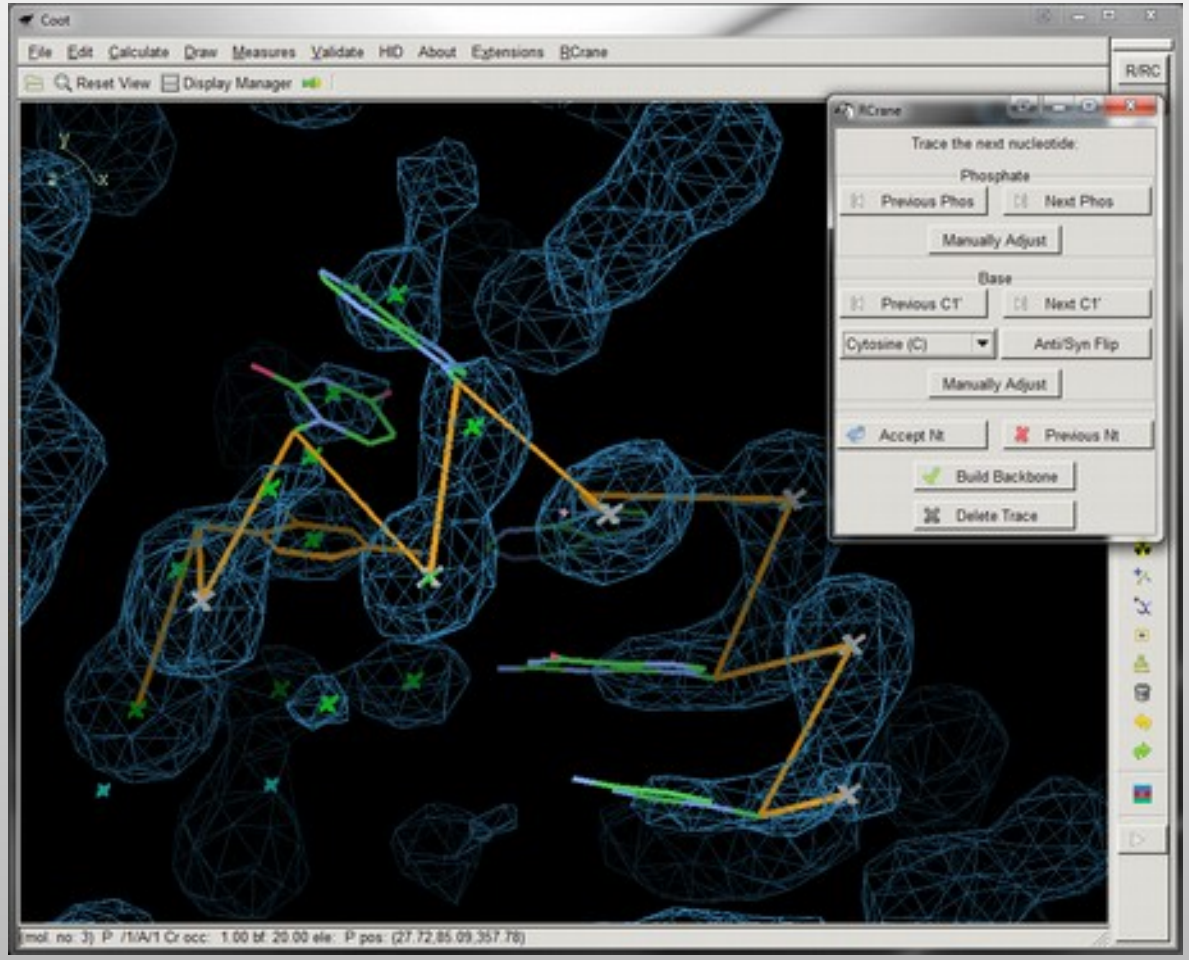
Twisted-trans

Non-pre-PRO

A Sample of Tools

- A few extra tools...

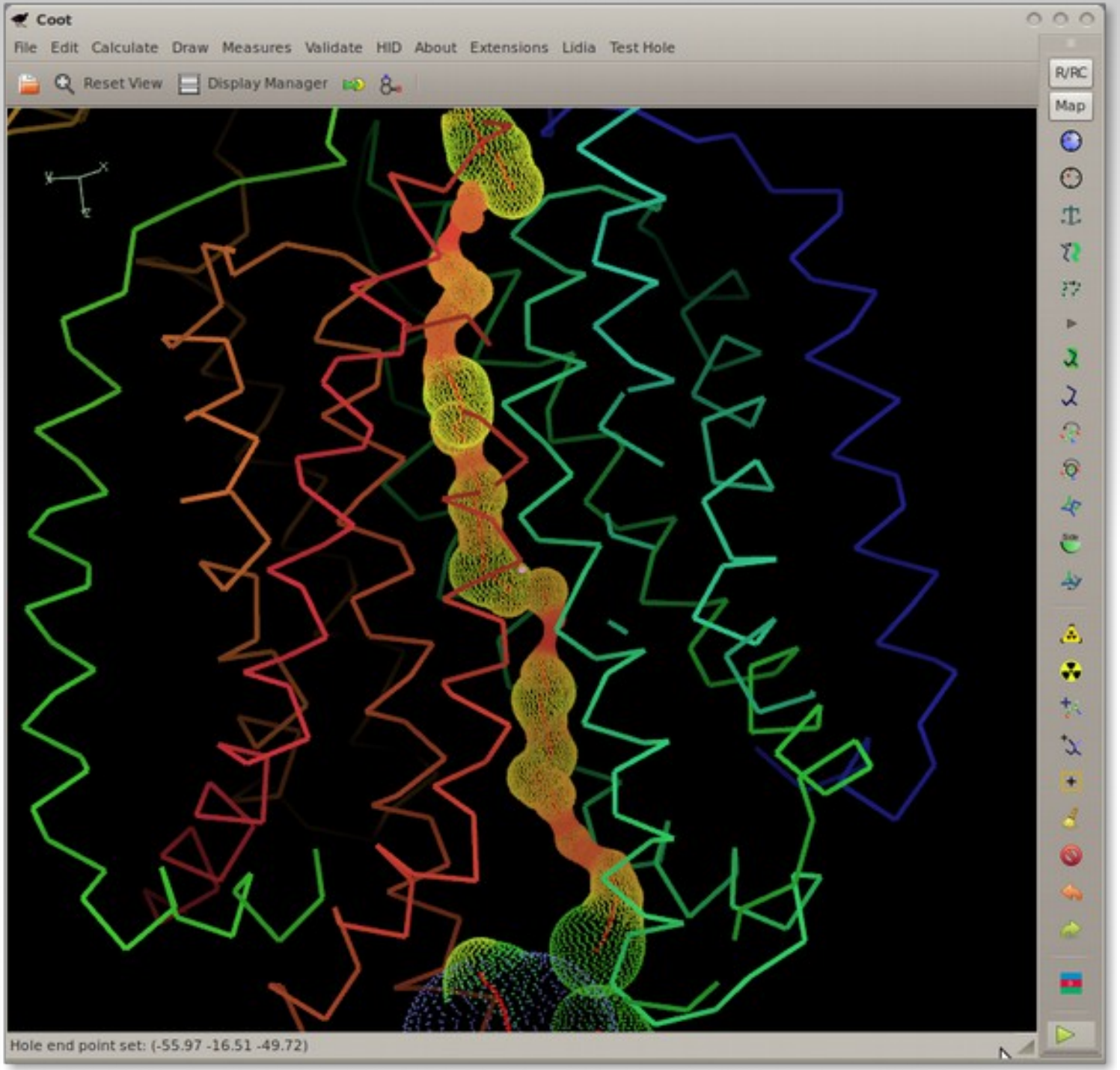
RCrane: Semi-automated RNA building



Kevin Keating

Finding Holes

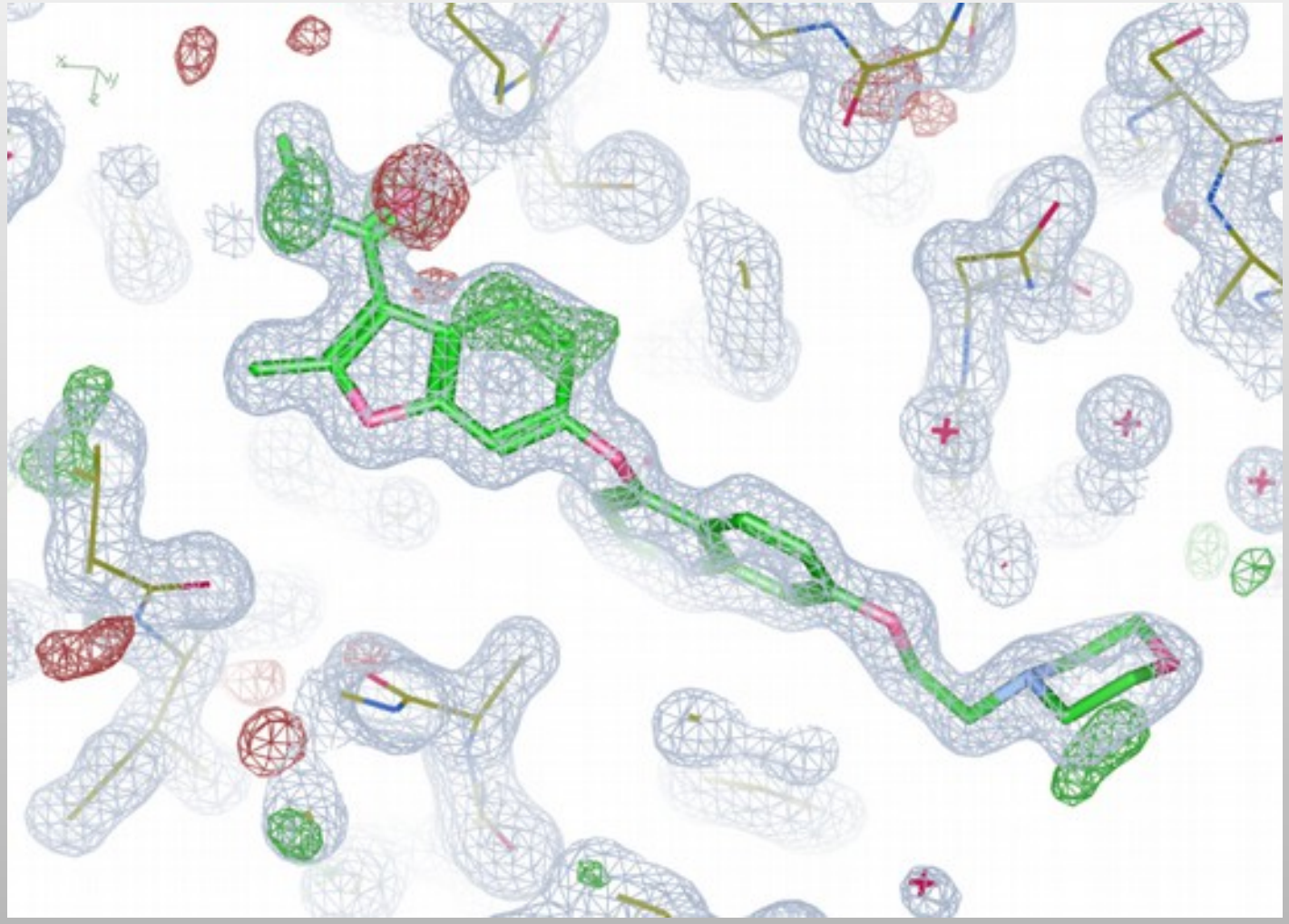
- An implementation of
  - Smart, Goodfellow & Wallace (1993) Biophysics Journal **65**, 2455
  - Atomic radii from AMBER
  - I used
    - radii from CCP4 monomer library
    - sans simulated annealing


Coot
File Edit Calculate Draw Measures Validate HID About Extensions Lidia Test Hole
Reset View
Display Manager
R/RC
Map
Hole end point set: (-55.97 -16.51 -49.72)

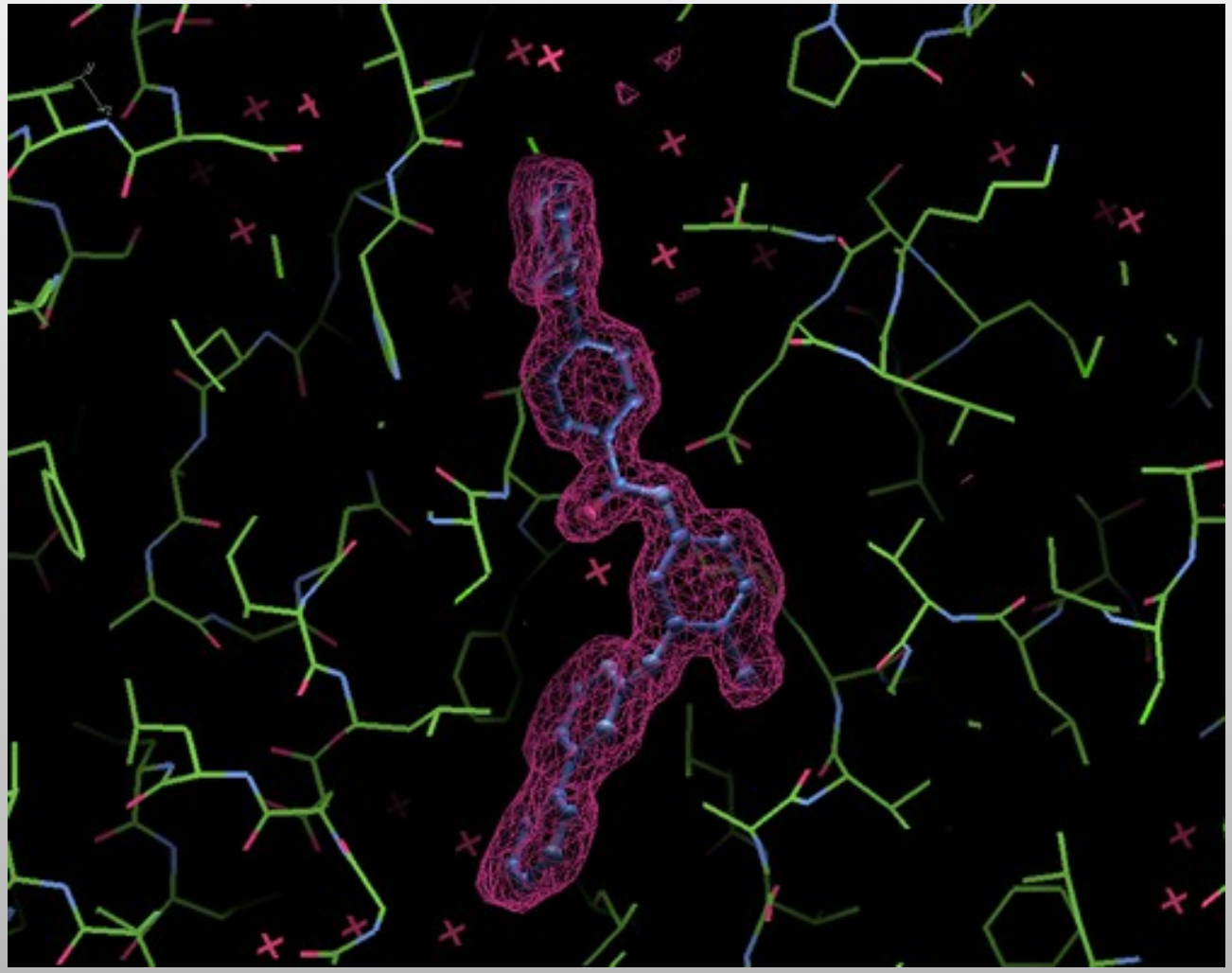
Making Density Slides with *Coot*

- White background
- “High” Oversampling (2.3x)
- Pale gray (or very pastel) density colour
- Enable Cut-glass mode 5-10%
- Anti-aliased Coot
  - `$ setenv __GL_FSAA_MODE 5`
  - 0.8.3 will do a better job of anti-aliasing out the box
  - (transfer to CCP4-built binaries)

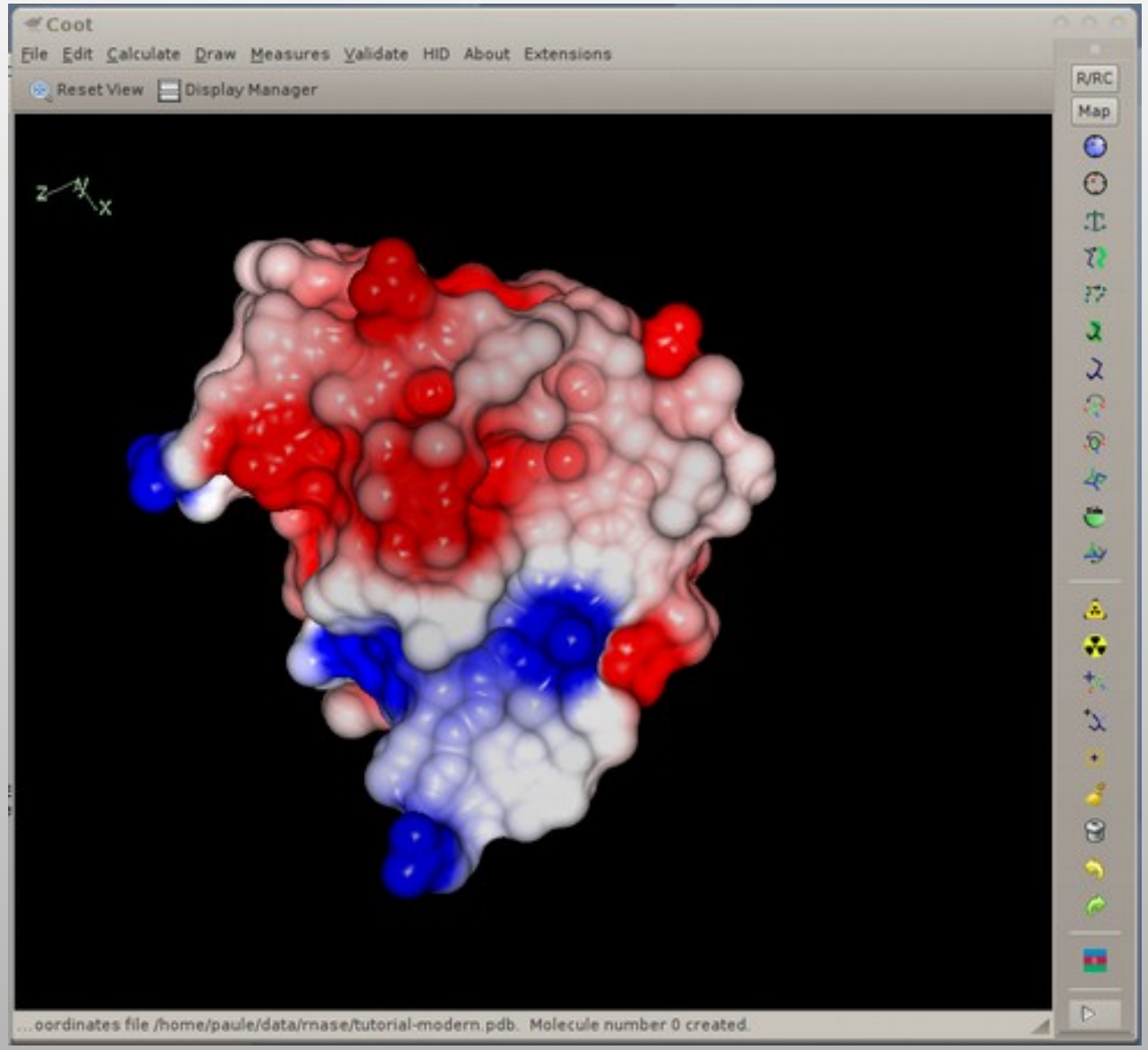
Example Density Slide



Some Representation Tools



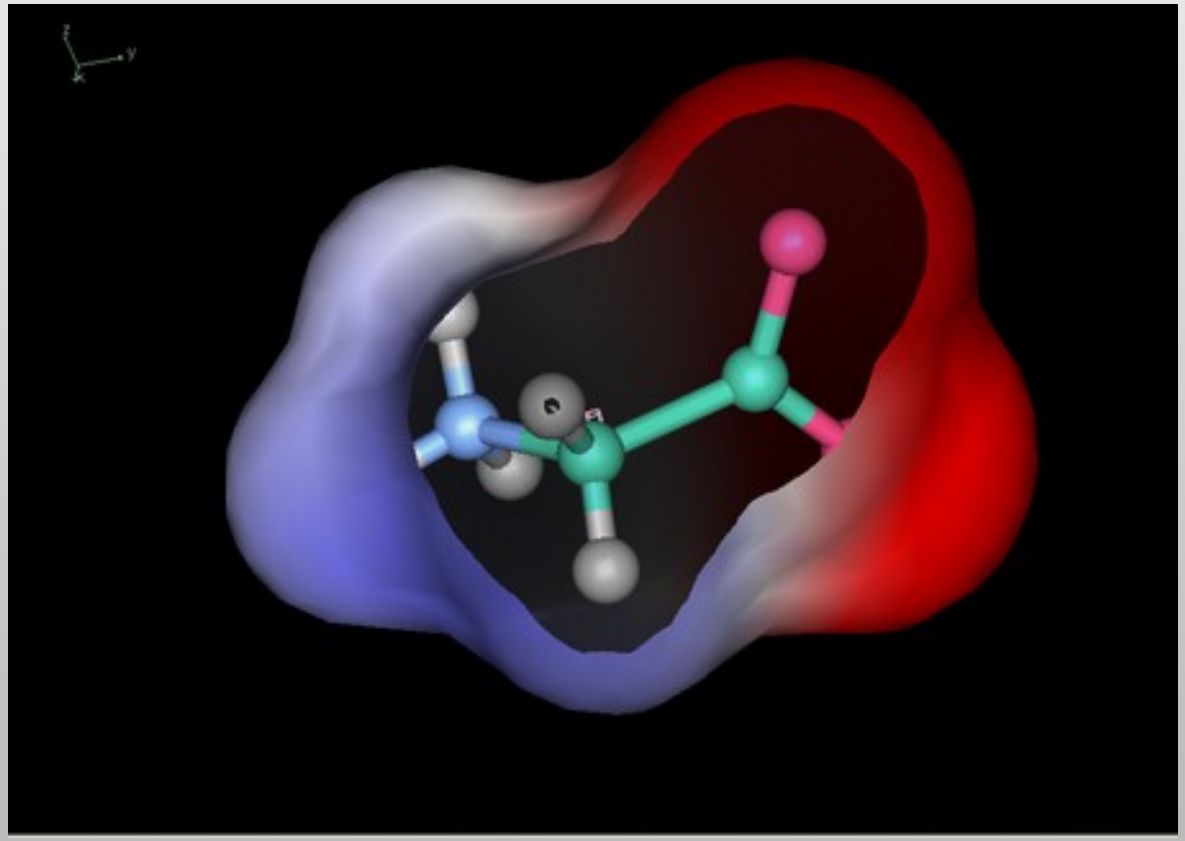
Some Representation Tools



Gruber & Noble (2007)

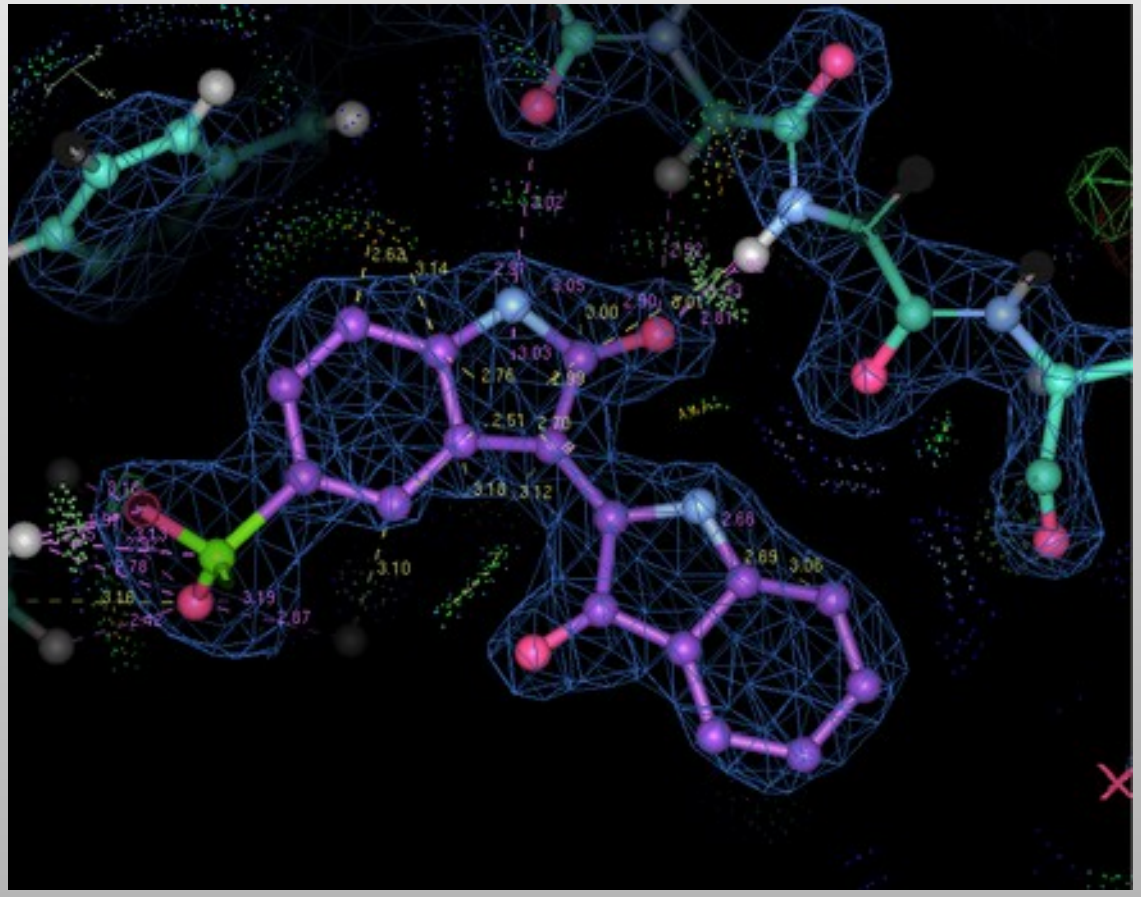
Other Things

- Surfaces that use dictionary partial charges

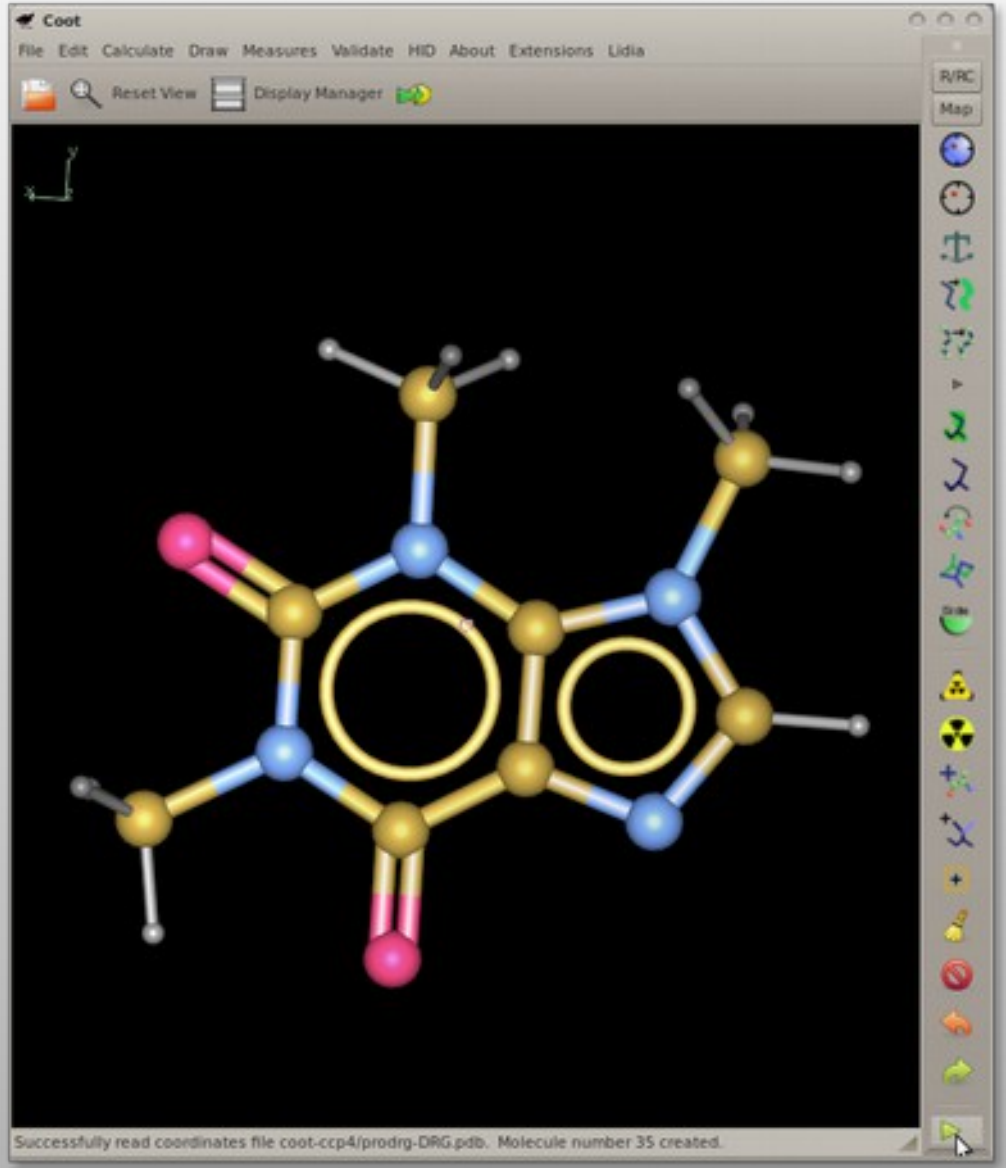


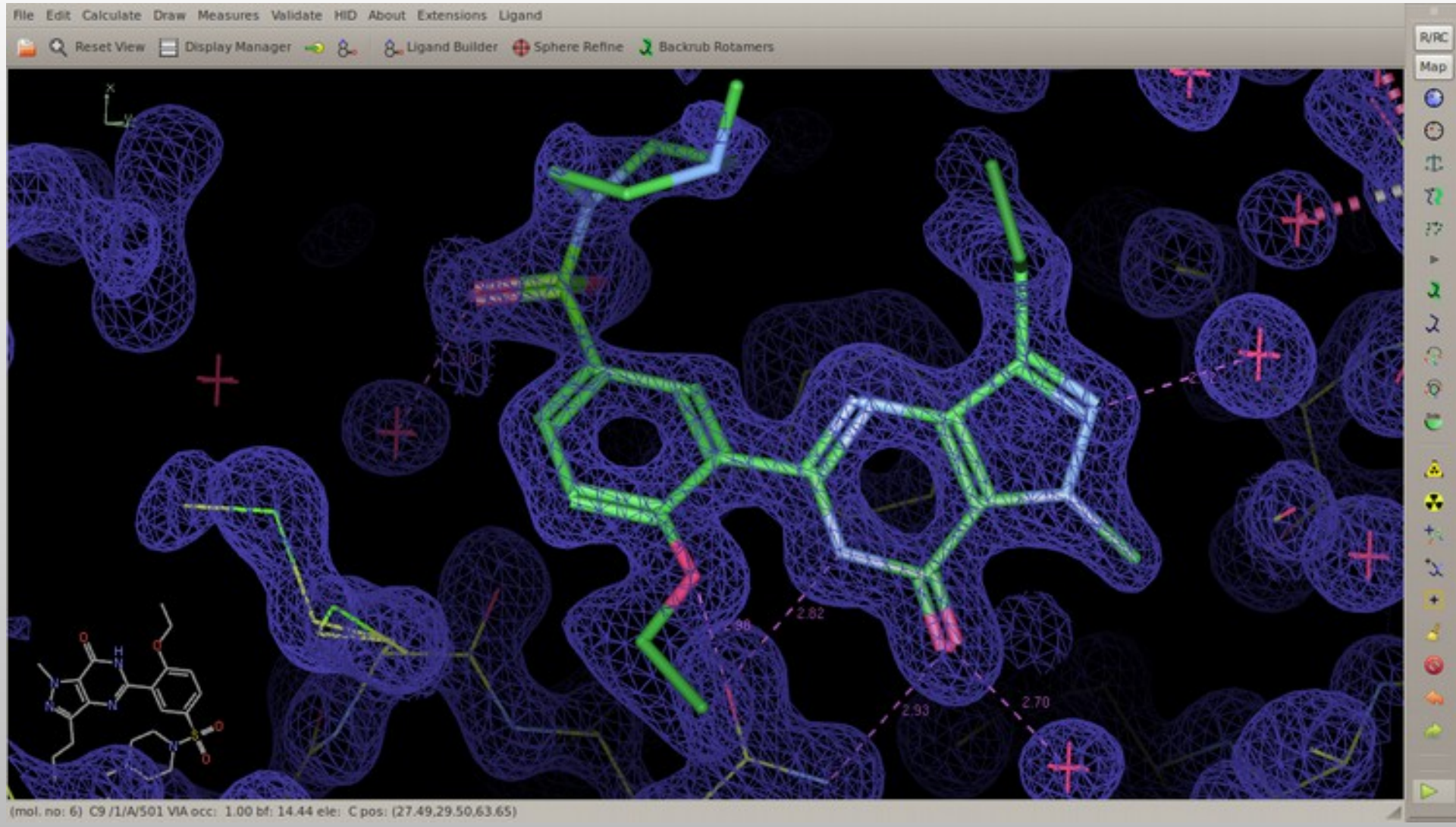
Other Tools

- Molprobity dots for ligands
  - Highlight interesting site



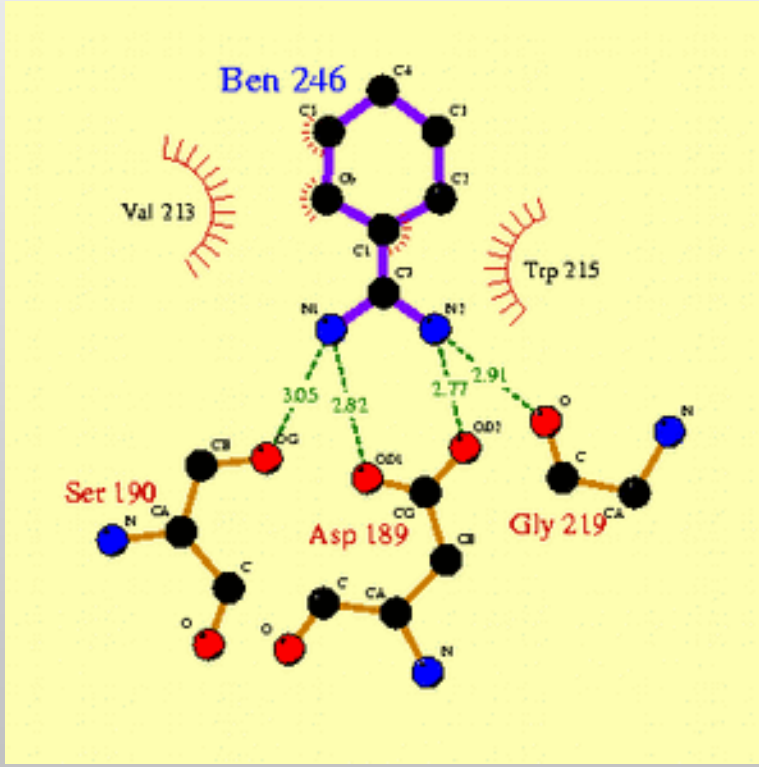
Representing Bond Orders

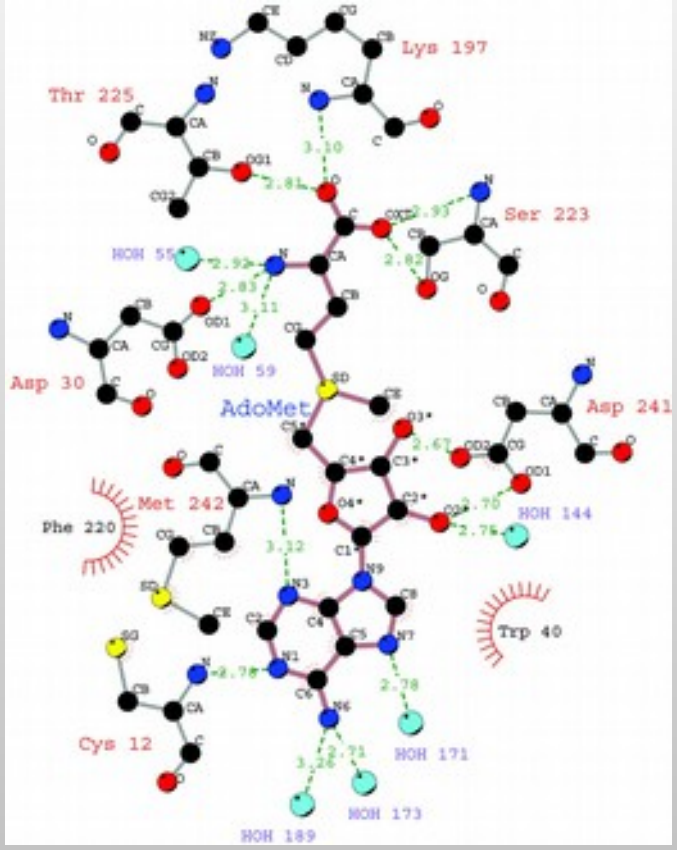



File Edit Calculate Draw Measures Validate HID About Extensions Ligand
Reset View
Display Manager
Ligand Builder
Sphere Refine
Backrub Rotamers
R/RC
Map
2.82
2.93
2.70
(mol. no: 6) C9 /1/A/501 VIA occ: 1.00 bf: 14.44 ele: C pos: (27.49,29.50,63.65)

Residue Environment Layout

Ligplot





Can we do better?

Residue Environment Layout

- This can't be solved by an “algorithmic/one-pass” procedure
  - Not in the general case
- Introduce “energy penalty terms” for displeasing interactions
- And use 2D energy minimisation to solve

Layout Energy Terms

$$E = \sum\sum w_{ij}(d_{ij}^2 - D_{ij}^2) +$$

Residues match 3D Distances

$$\sum\sum \exp(-\frac{1}{2}d_{ij}^2) +$$

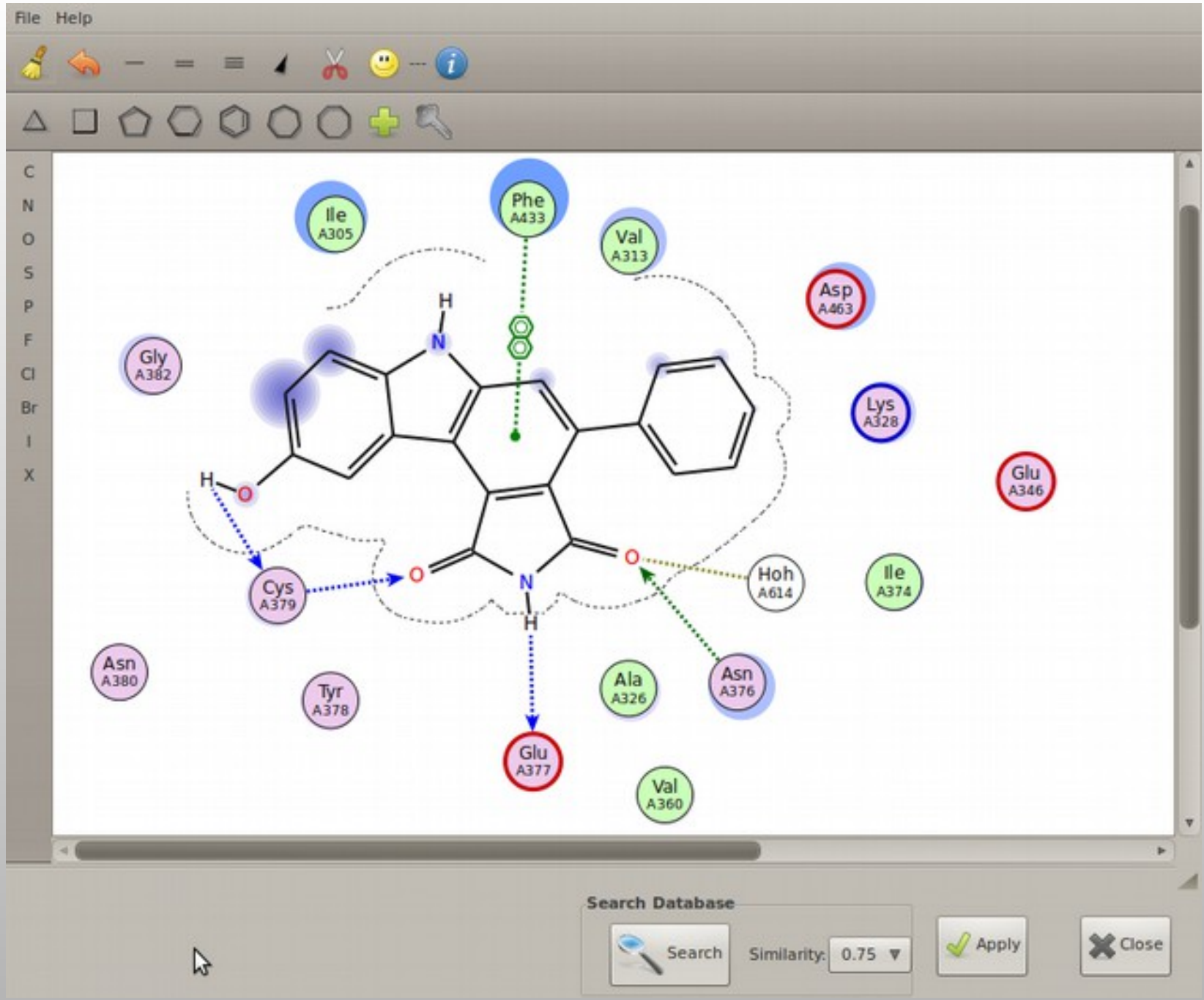
Residues don't overlay each other

$$\sum\sum (d_{ik}^2 - D_{ik}^2) +$$

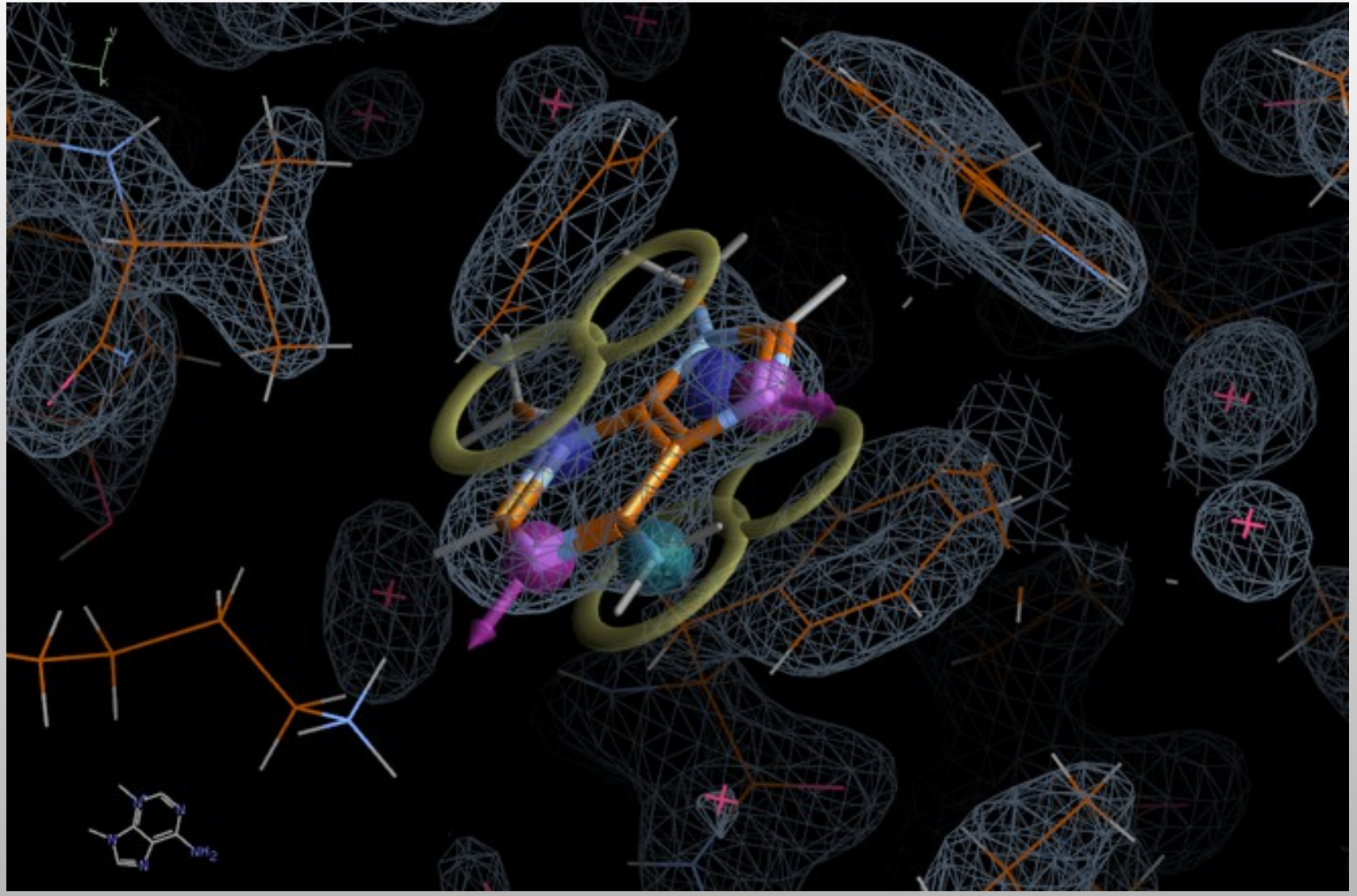
Residues are close to H-bonding ligand atoms

$$\sum\sum \exp(-\frac{1}{2}d_{ik}^2)$$

Residues don't overlap ligand


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

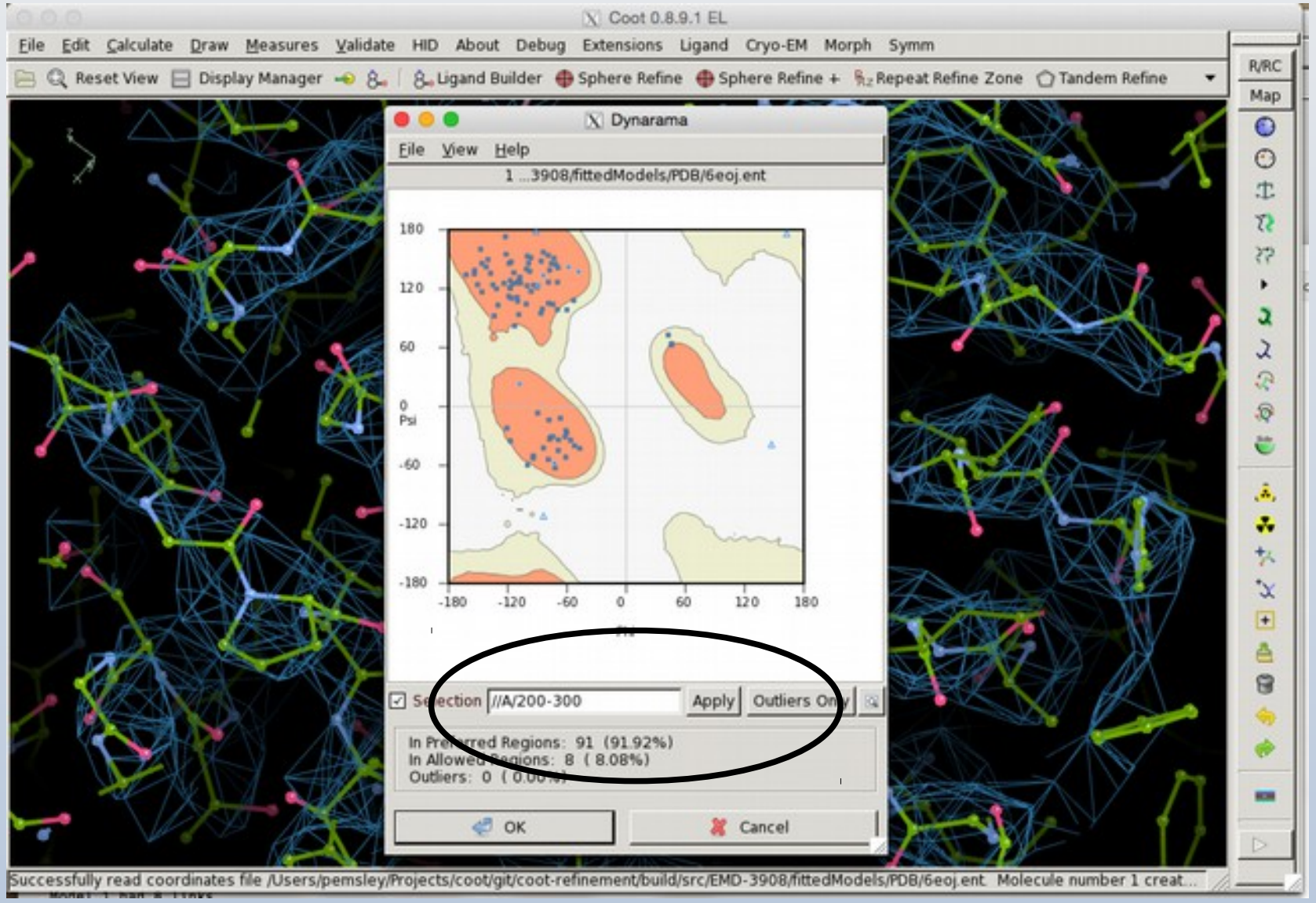
Chemical Features



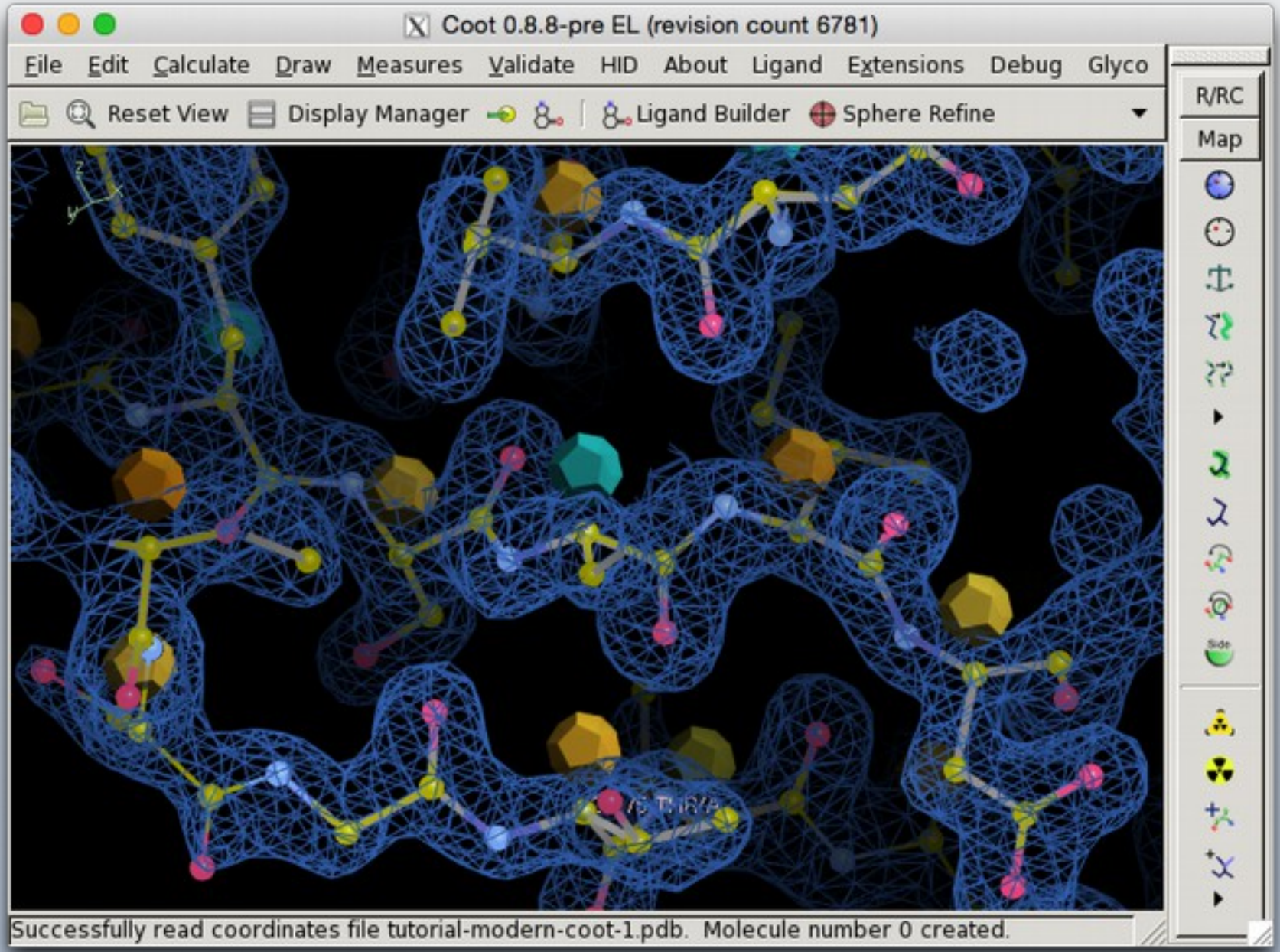
GUI Updates

Map Properties

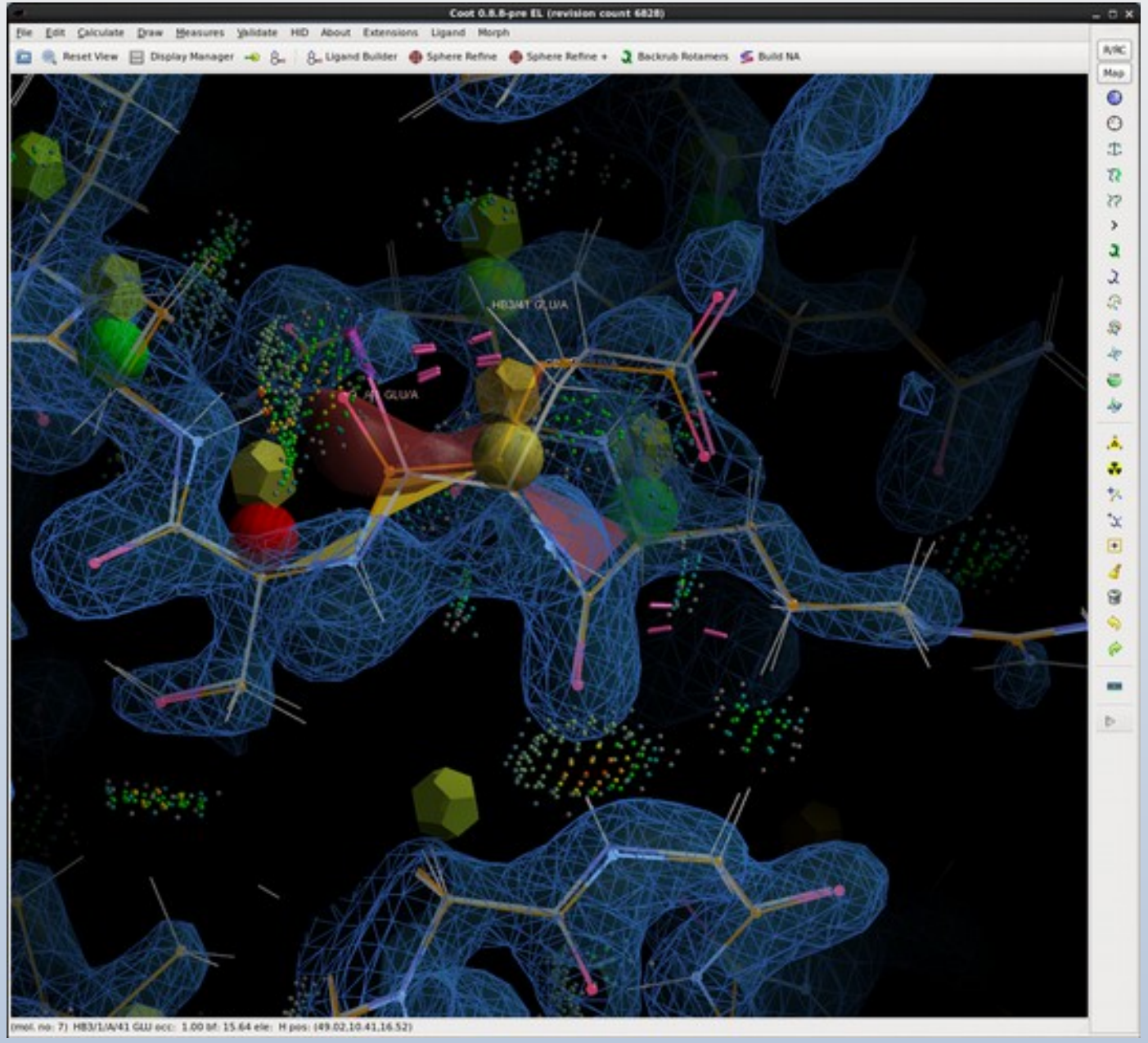
The New Ramachandran Plot



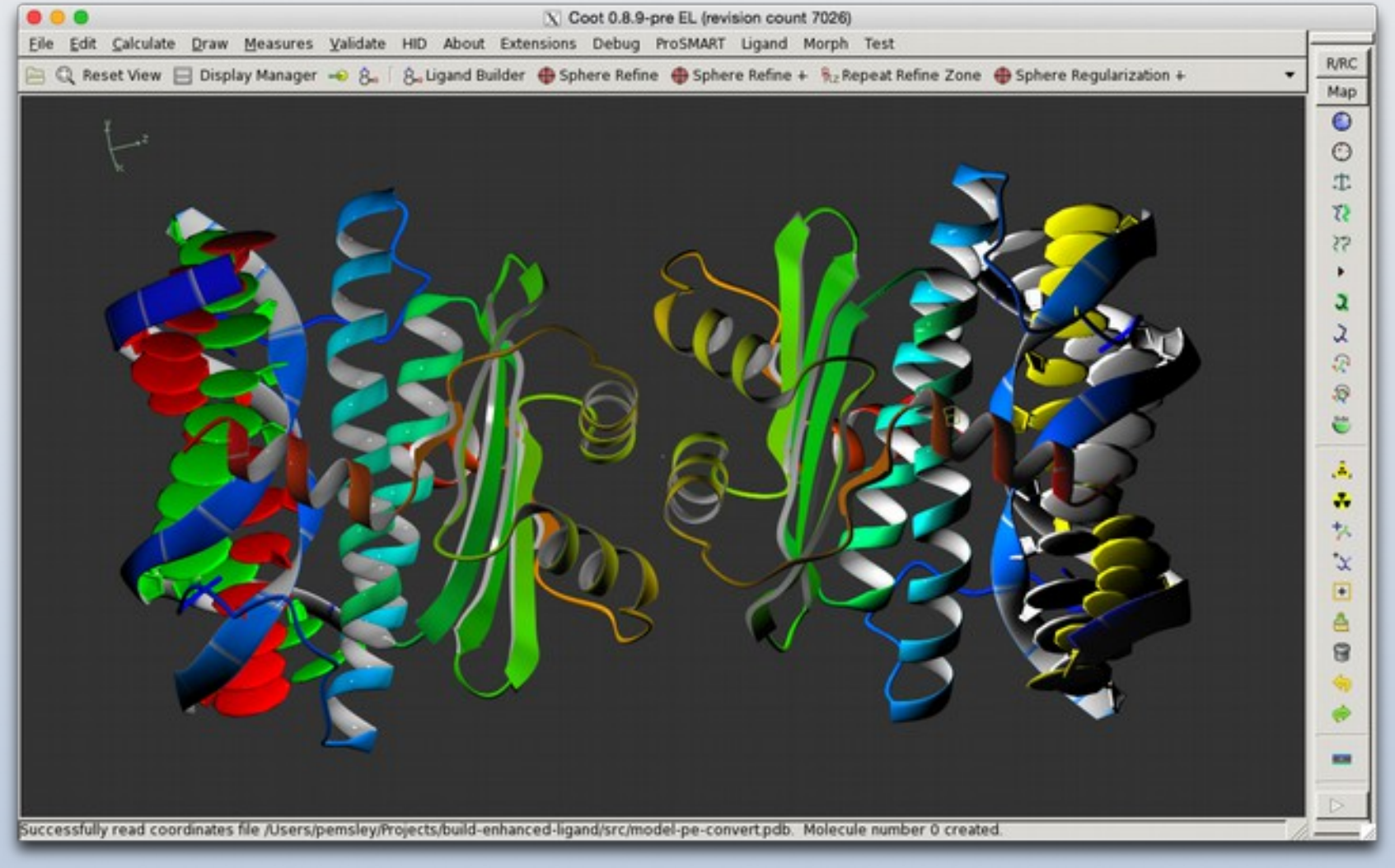
Interactive Rotamer Goodness



Multi-Criteria Markup



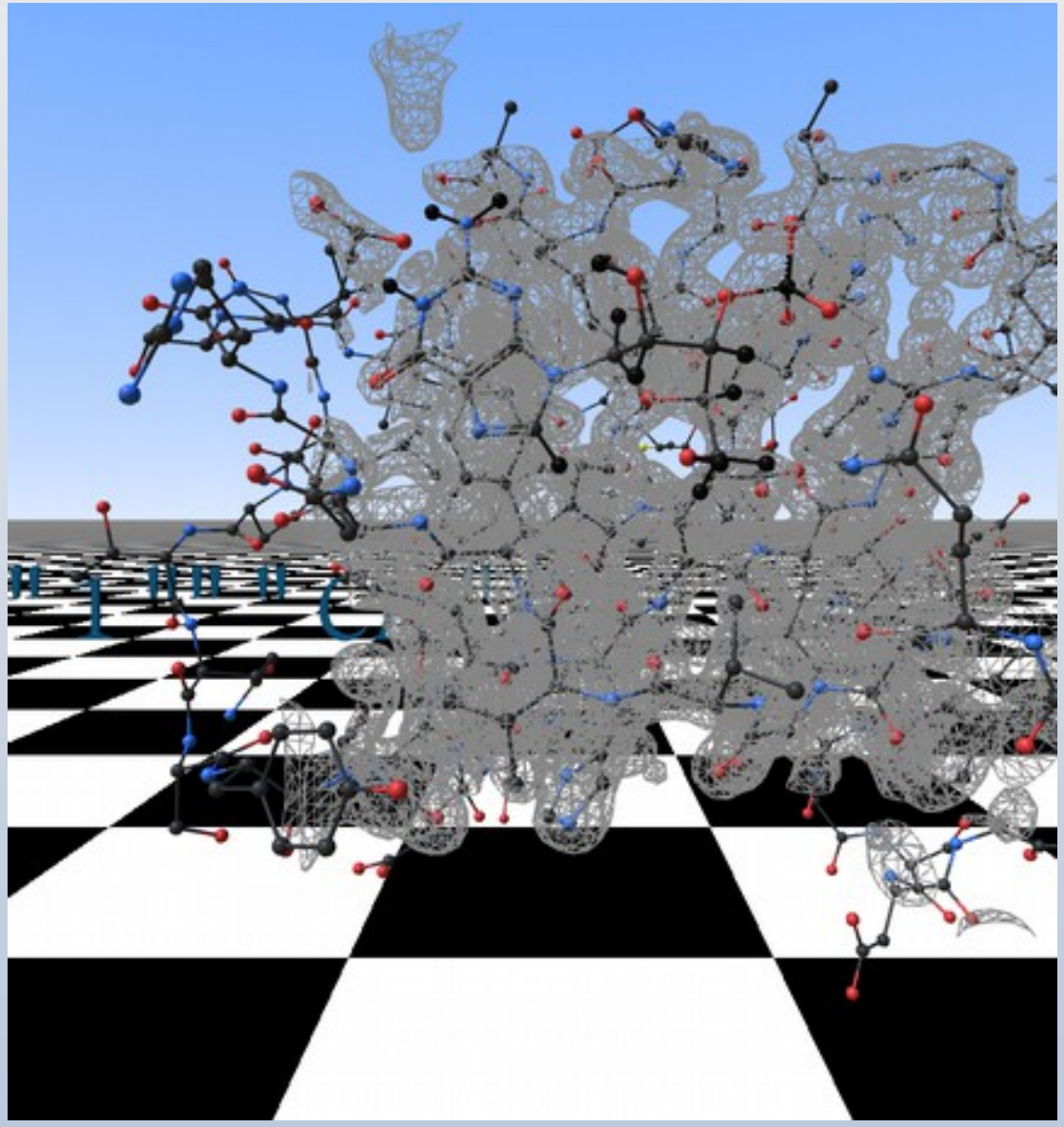
Coot Futures: GPU Ribbons



with Martin Noble

Coot Futures: Virtual Reality

Hamish Todd



- **An Intuitive Interface:**
- Stereoscopic Representation
- Greater Field of View
- 2 Hands with Articulation
- However:
  - current tools are not immediately transferable
  - because: nausea

CootVR

- Demonstrated at CCP-EM Meeting in Keele in April

A Few Tools More...

- Fitting Low-Resolution/EM maps
- Ligands:
  - dictionaries
  - ligand-fitting
  - analysis
- Carbohydrate-fitting
  - N-linked glycosylation

Acknowledgements

- Kevin Cowtan
- Bernhard Lohkamp
- Eleanor Dodson
- Keith Wilson

- Libraries, dictionaries
  - Alexei Vagin, Eugene Krissinel
  - Richardsons (Duke)

- Funding
  - BBSRC, CCP4 & MRC

Non-Crystallographic Symmetry

What is 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

Handling NCS

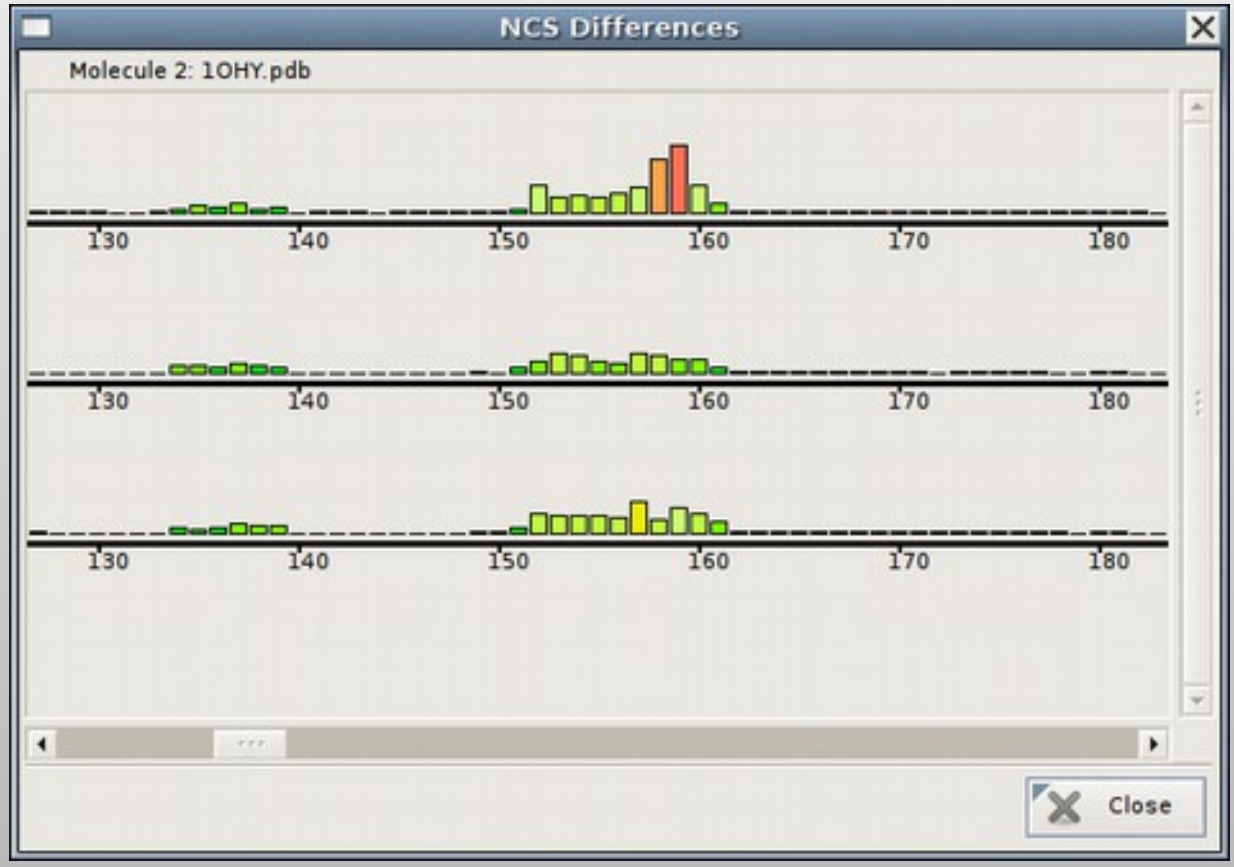
- What are the Problems?
- Strict NCS:
  - NCS should appear like crystallographic symmetry does [exact copies]
- Non-Strict NCS:
  - Molecules are different
  - How to cope with differences, but minimize unnecessary rebuilding?

Handling NCS

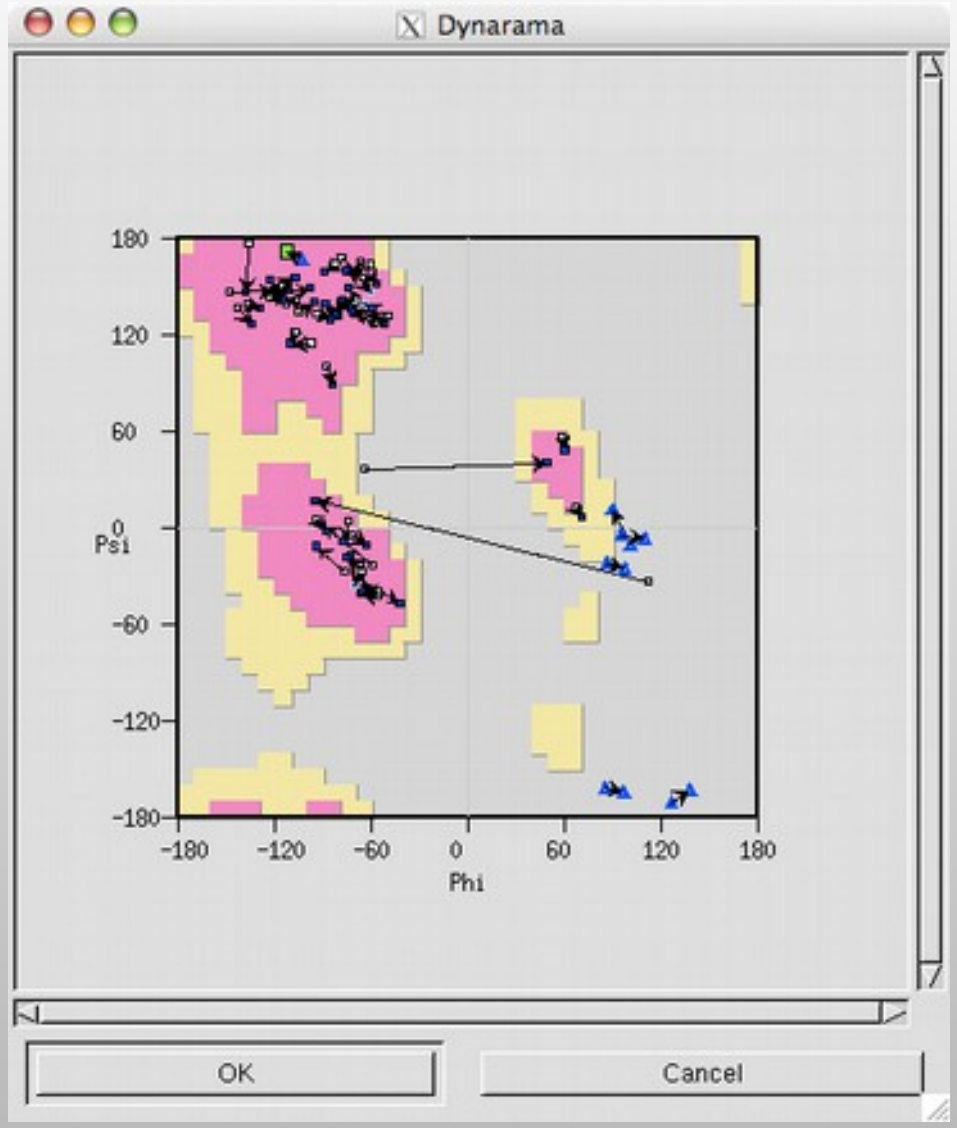
Typical Scenario:

- 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?

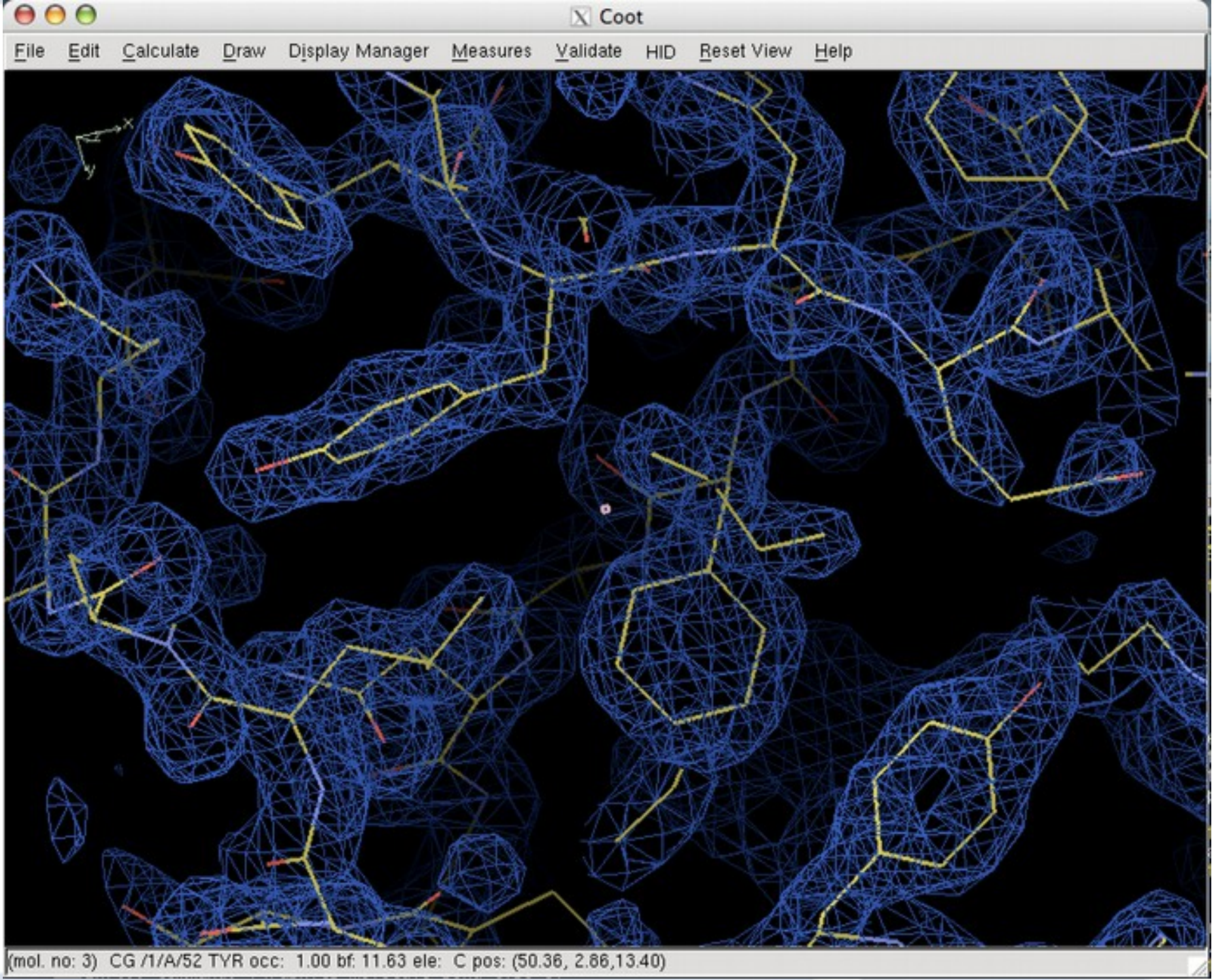
...or new NCS Differences graph



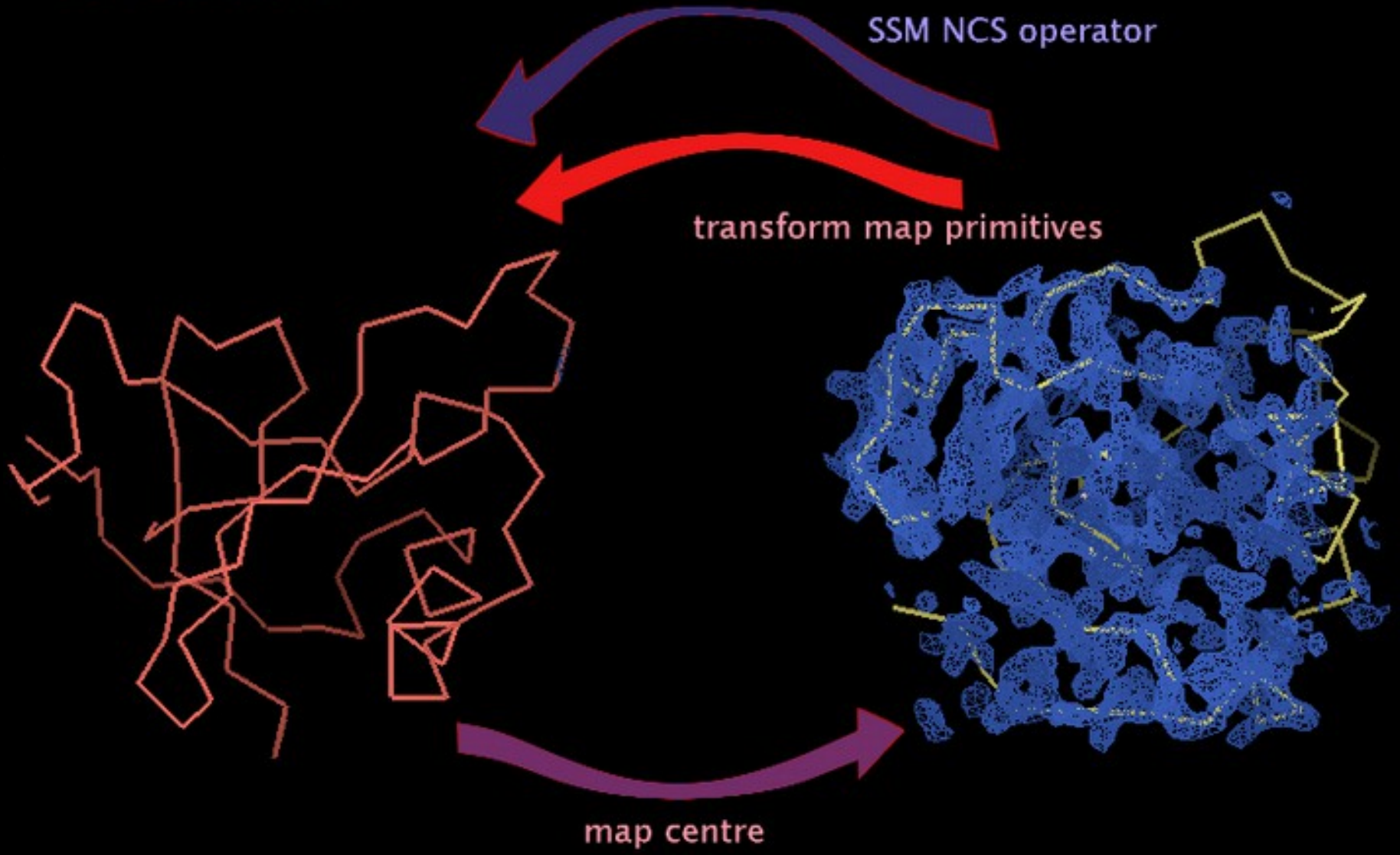
…or Kleywegt Plots[*]

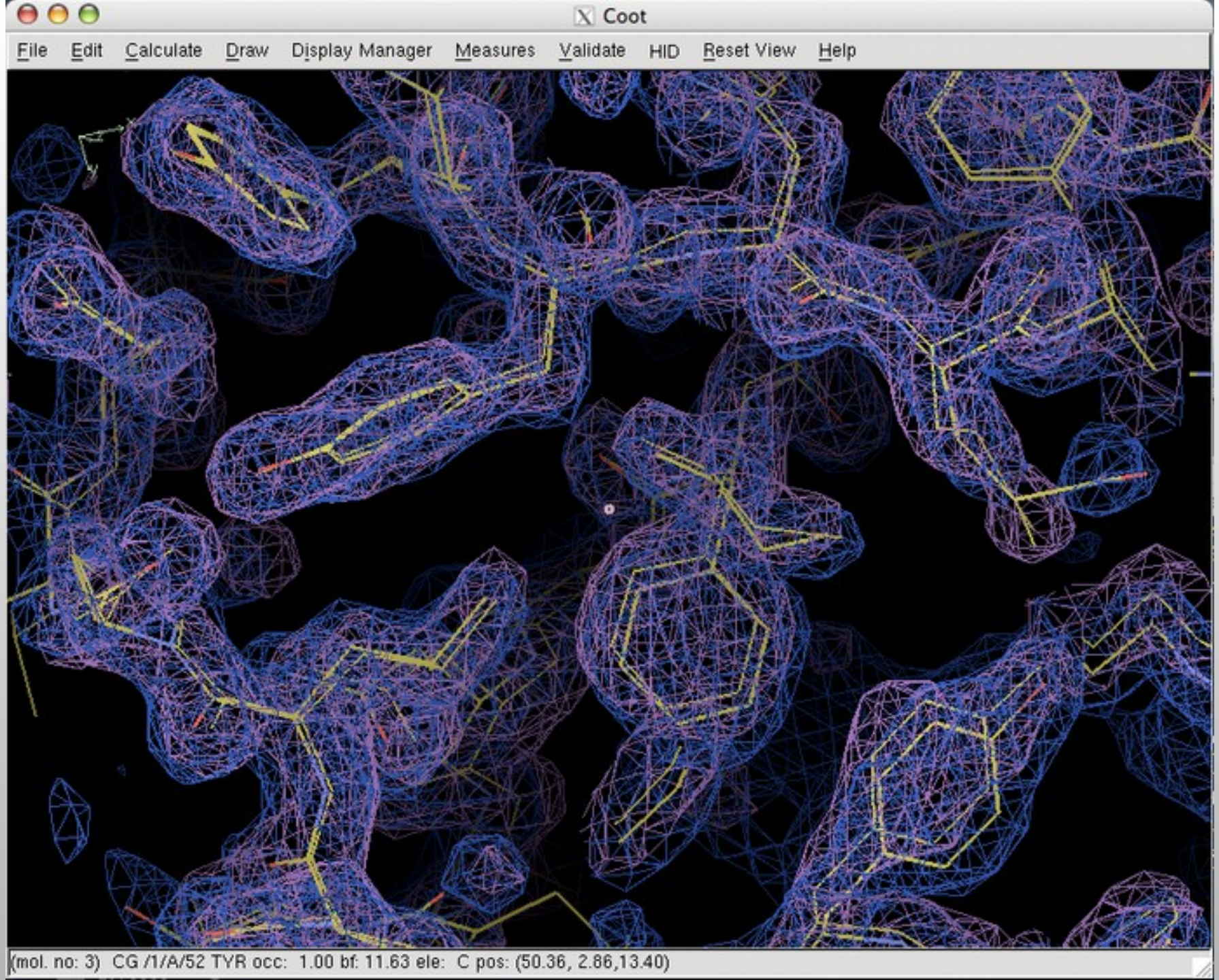


[*] Named by George Sheldrick


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, 2.66,13.40)

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)