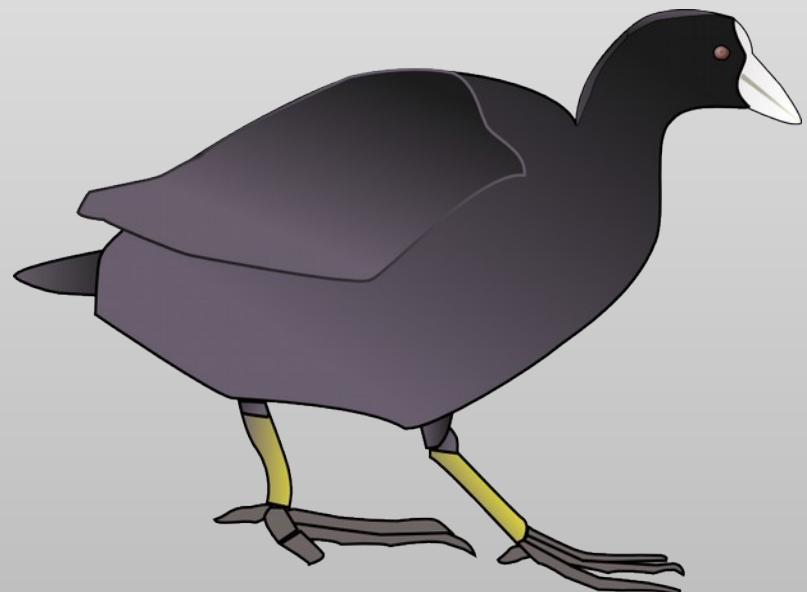




Model-Building of Cryo-EM Reconstructions With *Coot*



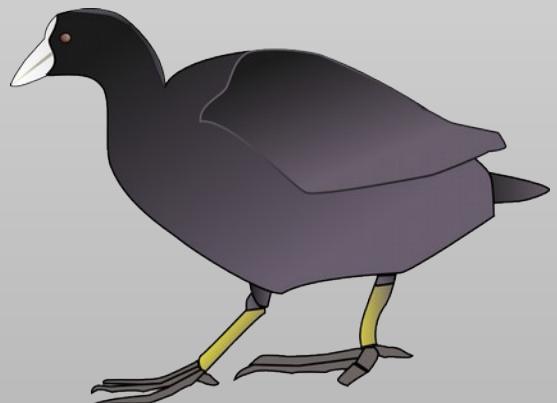
Paul Emsley
Sept 2017

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
- Infrastructure needed for access to other applications via scripting API, allows
 - Test-Driven Development
 - Easy GUI access
 - Server-Client configuration
 - Coot as a Python module

Cryo-EM model-building

- Typically need to move more atoms than one does for crystallography
- the maps are lower resolution and the starting model is further from where you want them to be – usually systematically so
- addressing these needs has been the focus of my work extending/changing coot for cryo-EM
 - adjusting and scripting the tools to require less user intervention



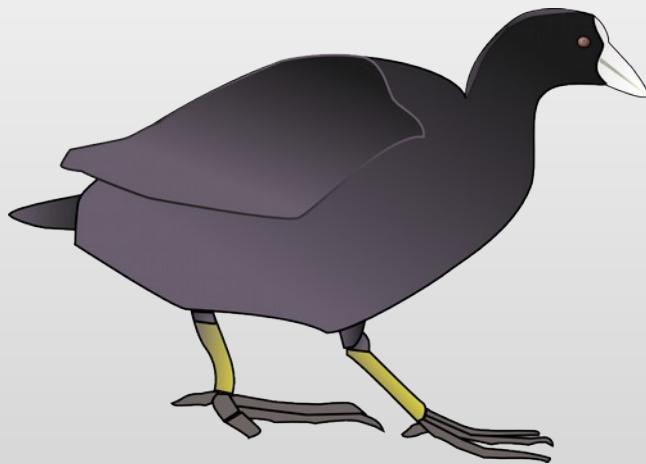
For What is Coot Useful?

- What resolution ranges for cryo-EM?
 - “Resolution Revolution” quality map
 - 2 – 4Å range of Coot's strengths
- Seeing side-chains and purines vs pyrimidines
- Local good fit of model to density

Cryo-EM Model-building

- autozone multi-residue
- sphere refine, sphere refine+
- sphere regularize, sphere regularize +
- Geman-McClure distance restraints
- Multi-threading/parallel processing
 - (WIP)
- backrub rotamers

Acknowledgments, Collaborators



Bernhard
Lohkamp



Kevin
Cowtan



Eugene
Krissinel



Stuart
McNicholas



Martin
Noble



Alexei
Vagin

Yeast Mitochondrial Large Ribosomal Subunit

Science AAAS.ORG | FEEDBACK | HELP | LIBRARIANS All Science Journals ▾ Enter Search Term BATES COLLEGE ALERTS | ACCESS RIGHTS

AAAS NEWS SCIENCE JOURNALS CAREERS MULTIMEDIA COLLECTIONS

▶ Science The World's Leading Journal of Original Scientific Research, Global News, and Commentary.

Science Home Current Issue Previous Issues Science Express Science Products My Science About the Journal

Home > [Science Magazine](#) > [28 March 2014](#) > Amunts *et al.*, [343](#) (6178): 1485-1489

Article Views
Science 28 March 2014:
Vol. 343 no. 6178 pp. 1485-1489
DOI: 10.1126/science.1249410

Full Text
RESEARCH ARTICLE
Structure of the Yeast Mitochondrial Large Ribosomal Subunit
Alexey Amunts[†], Alan Brown[†], Xiao-chen Bai[†], Jose L. Llacer[†], Tanweer Hussain, Paul Emsley, Fei Long, Garib Murshudov, Sjors H. W. Scheres[‡], V. Ramakrishnan[‡]

Author Affiliations
[†]Corresponding author. E-mail: scheres@mrc-lmb.cam.ac.uk (S.H.W.S.); ramak@mrc-lmb.cam.ac.uk (V.R.)
[‡]These authors contributed equally to this work.

Article Tools
Leave a comment (0)
Save to My Folders
Download Citation
Alert Me When Article is Cited
Post to CiteULike
Article Usage Statistics
E-mail This Page
Rights & Permissions
Commercial Reprints and E-Prints
View PubMed Citation

ABSTRACT **EDITOR'S SUMMARY**

Mitochondria have specialized ribosomes that have diverged from their bacterial and cytoplasmic counterparts. We have solved the structure of the yeast mitoribosomal large subunit using single-particle cryo-electron microscopy. The resolution of 3.2 angstroms enabled a nearly complete atomic model to be built *de novo* and refined, including 39 proteins, 13 of which are unique to mitochondria, as well as expansion segments of mitoribosomal RNA. The structure reveals a new exit tunnel path and architecture, unique elements of the E site, and a putative membrane docking site.

Mitochondria are organelles in eukaryotic cells that play a major role in metabolism, especially the synthesis of adenosine triphosphate (ATP).

Related Resources

Tools for macromolecular model building and refinement into electron cryo-microscopy reconstructions

Alan Brown, Fei Long, Robert A. Nicholls, Jaan Toots, Paul Emsley and Garib Murshudov*

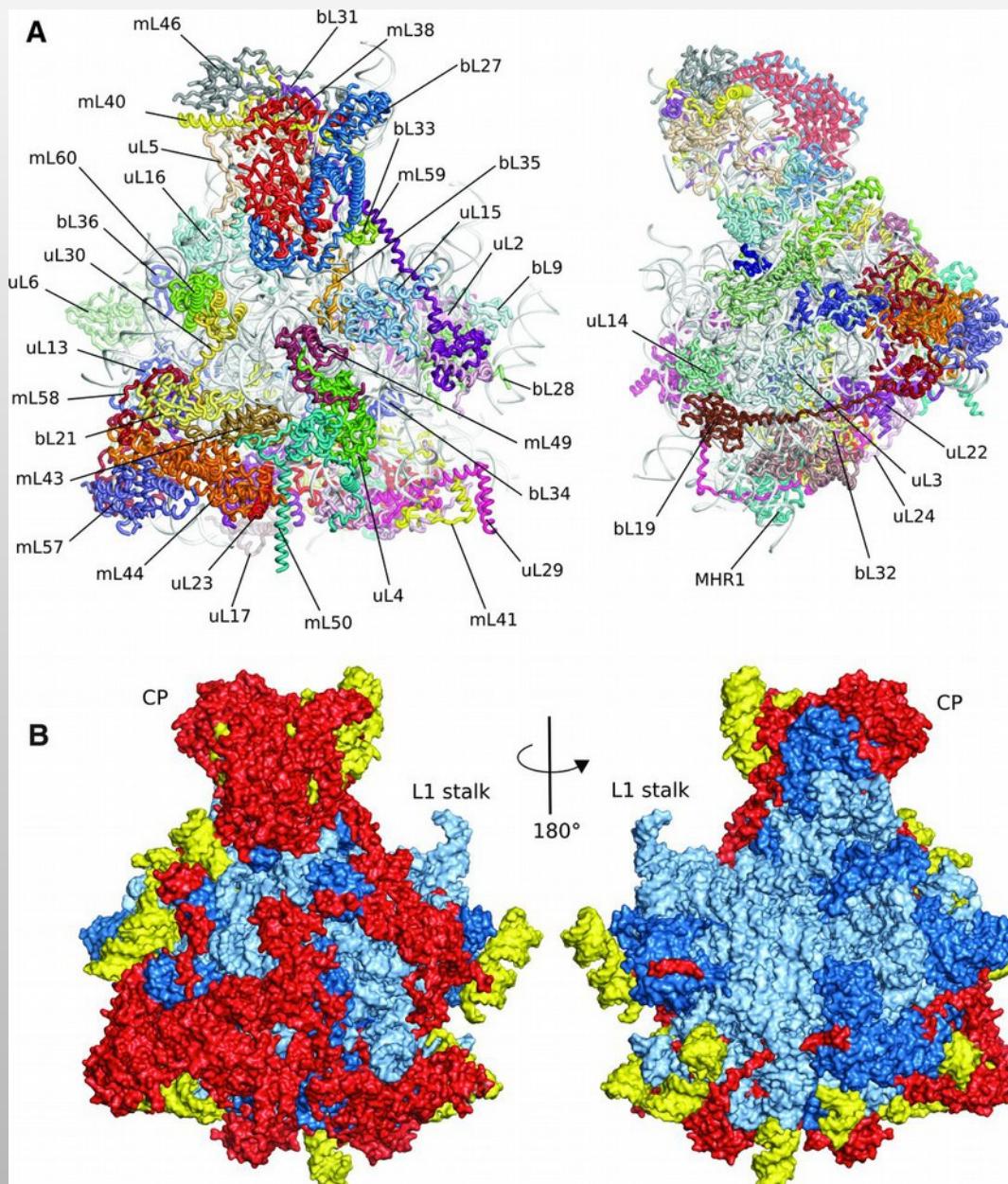
MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge CB2 0QH, England

Correspondence e-mail:
garib@mrc-lmb.cam.ac.uk

Received 3 June 2014
Accepted 1 October 2014

The recent rapid development of single-particle electron cryo-microscopy (cryo-EM) now allows structures to be solved by this method at resolutions close to 3 Å. Here, a number of tools to facilitate the interpretation of EM reconstructions with stereochemically reasonable all-atom models are described. The *BALBES* database has been repurposed as a tool for identifying protein folds from density maps. Modifications to *Coot*, including new Jiggle Fit and morphing tools and improved handling of nucleic acids, enhance its functionality for interpreting EM maps. *REFMAC* has been modified for optimal fitting of atomic models into EM maps. As external structural information can enhance the reliability of the derived atomic models, stabilize refinement and reduce overfitting, *ProSMART* has been extended to generate interatomic distance restraints from nucleic acid reference structures, and a new tool, *LIBG*, has been developed to generate nucleic acid base-pair and parallel-plane restraints. Furthermore, restraint generation has been integrated with visualization and editing in *Coot*, and these restraints have been applied to both real-space refinement in *Coot* and reciprocal-space refinement in *REFMAC*.

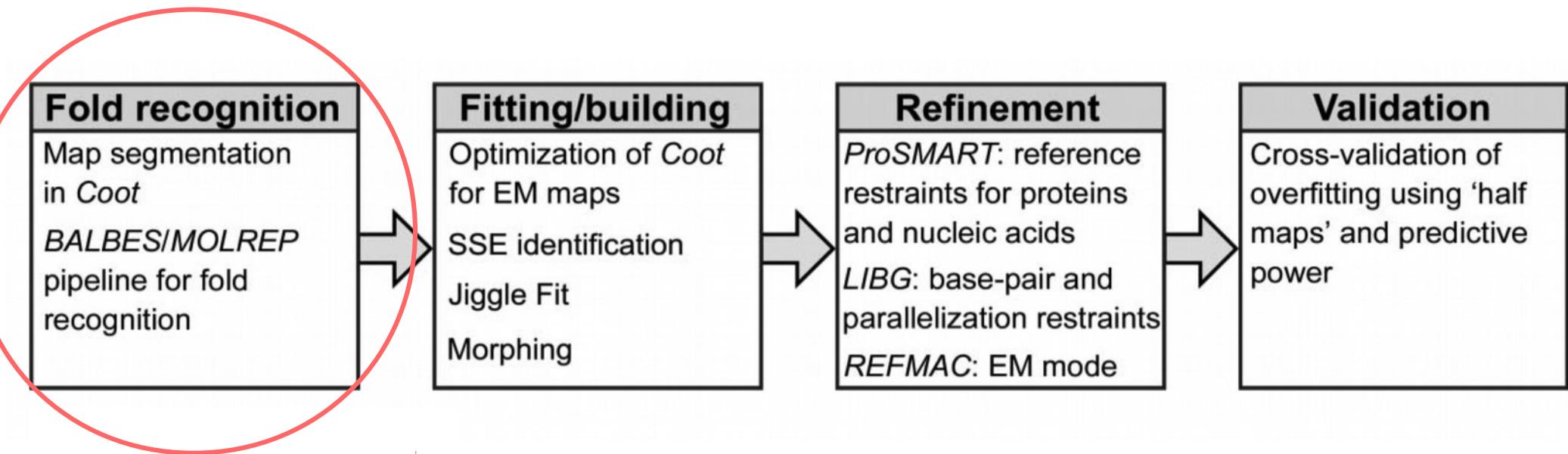
New Components



blue: conserved
red: found in both
yellow: yeast only

Model-Building Tools

Recent Developments

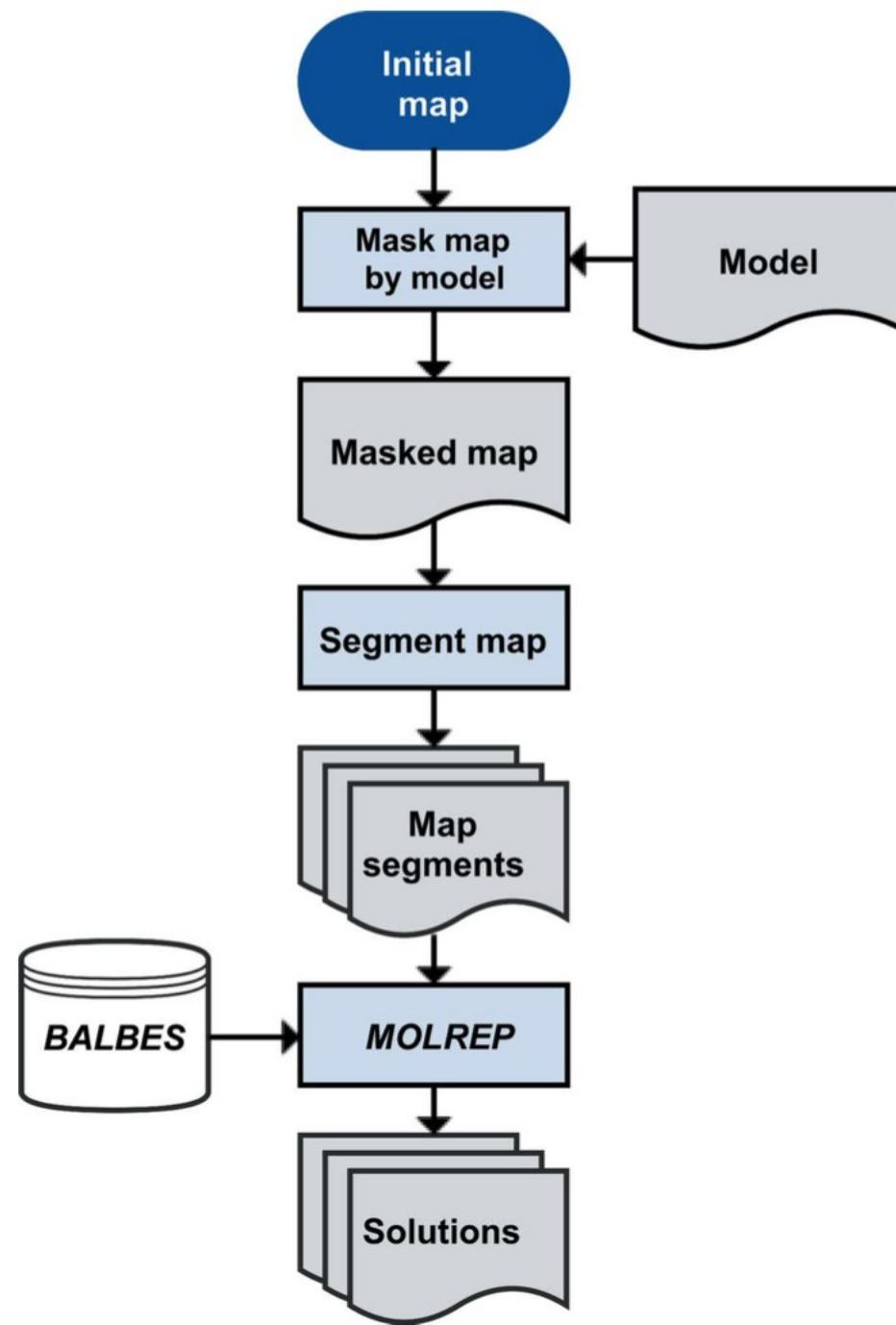


Cryo-EM Data

- Ability to collect data from native sources
 - in such cases, the composition of the complex may not be known
 - Mere “docking” of high resolution structures/fragments cannot work
 - At 4Å, it may be possible to trace the backbone
 - At better than 4Å it may be possible to assign the amino-acid sequence
 - thus search a sequence database for possible matches

Cryo-EM data

- Alternatively, use fold recognition
- Using the BALBES database
 - (originally design for molecular replacement)
 - screen unknown density against domains



Fold Recognition

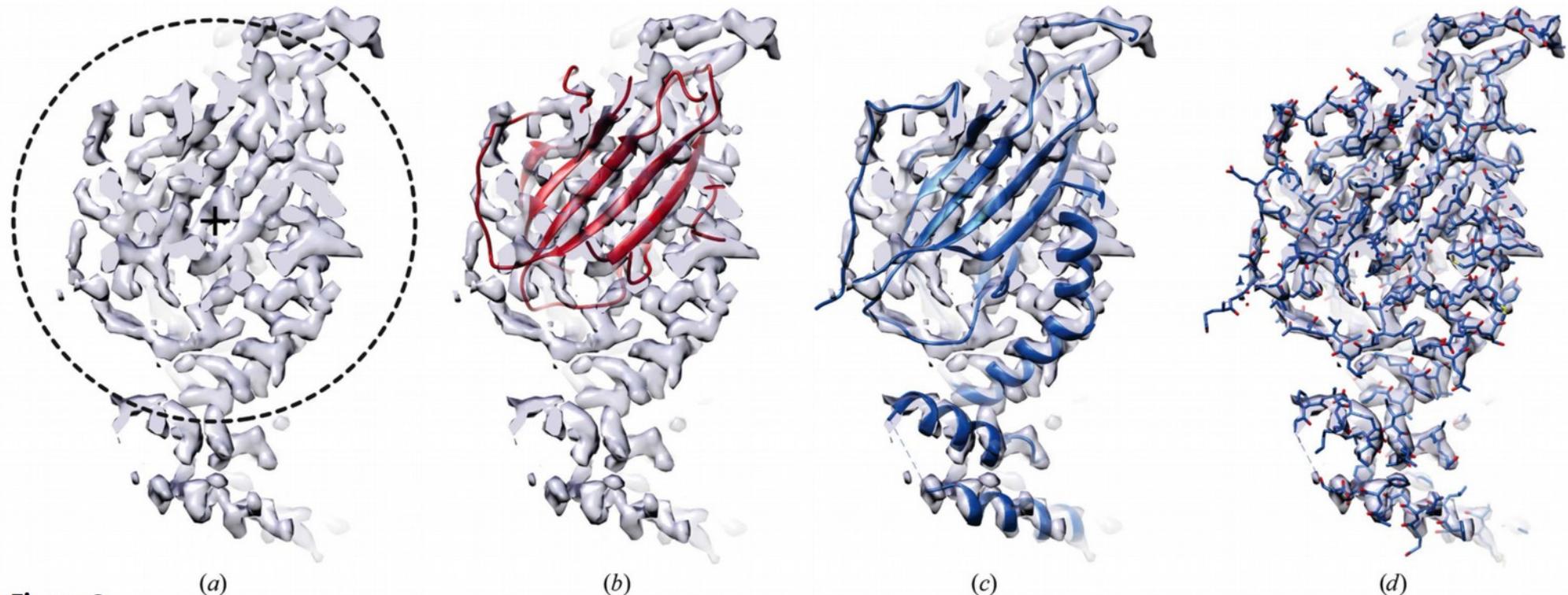
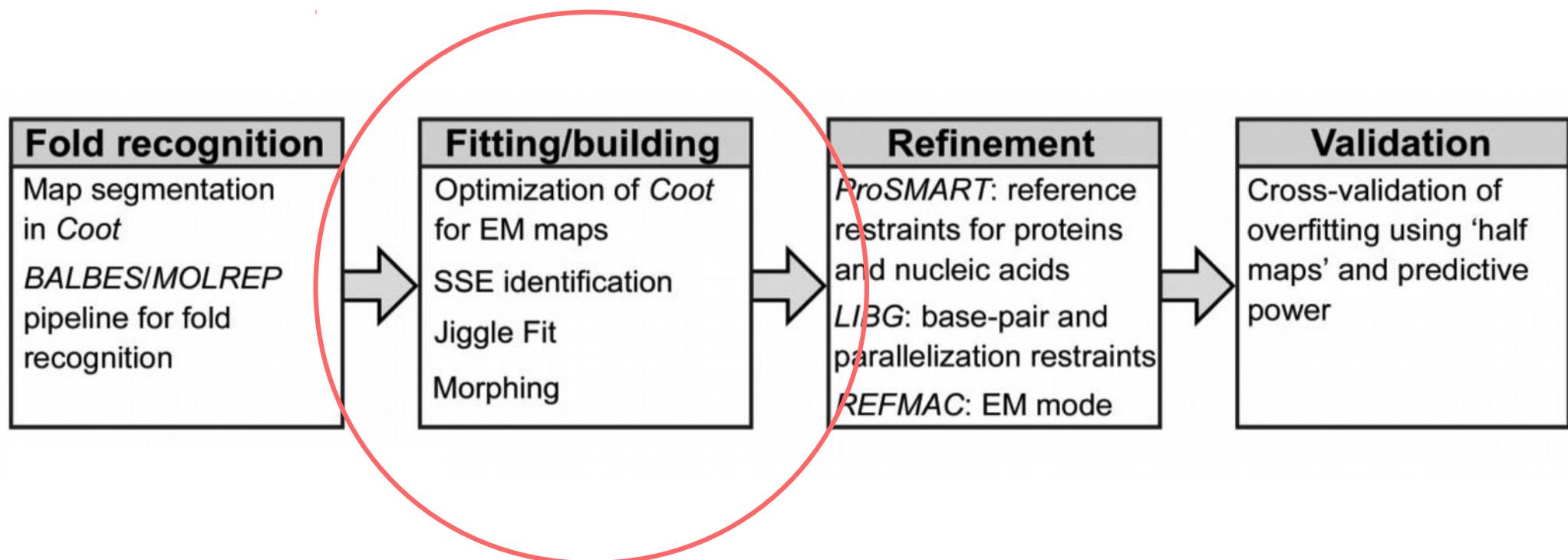


Figure 3

Fold recognition can identify template molecules for model building. (a) Density map corresponding to the final model of the mitoribosomal protein mL38 with the segmented search map indicated. (b) Top solution from the *BALBES-MOLREP* pipeline. (c, d) Final refined model of mL38 in (c) cartoon and (d) full-atom representation.

Model-Building Tools

Recent Developments



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
 - Including links and modifications
- Provides “interactive” refinement
- Subject to substantial extension

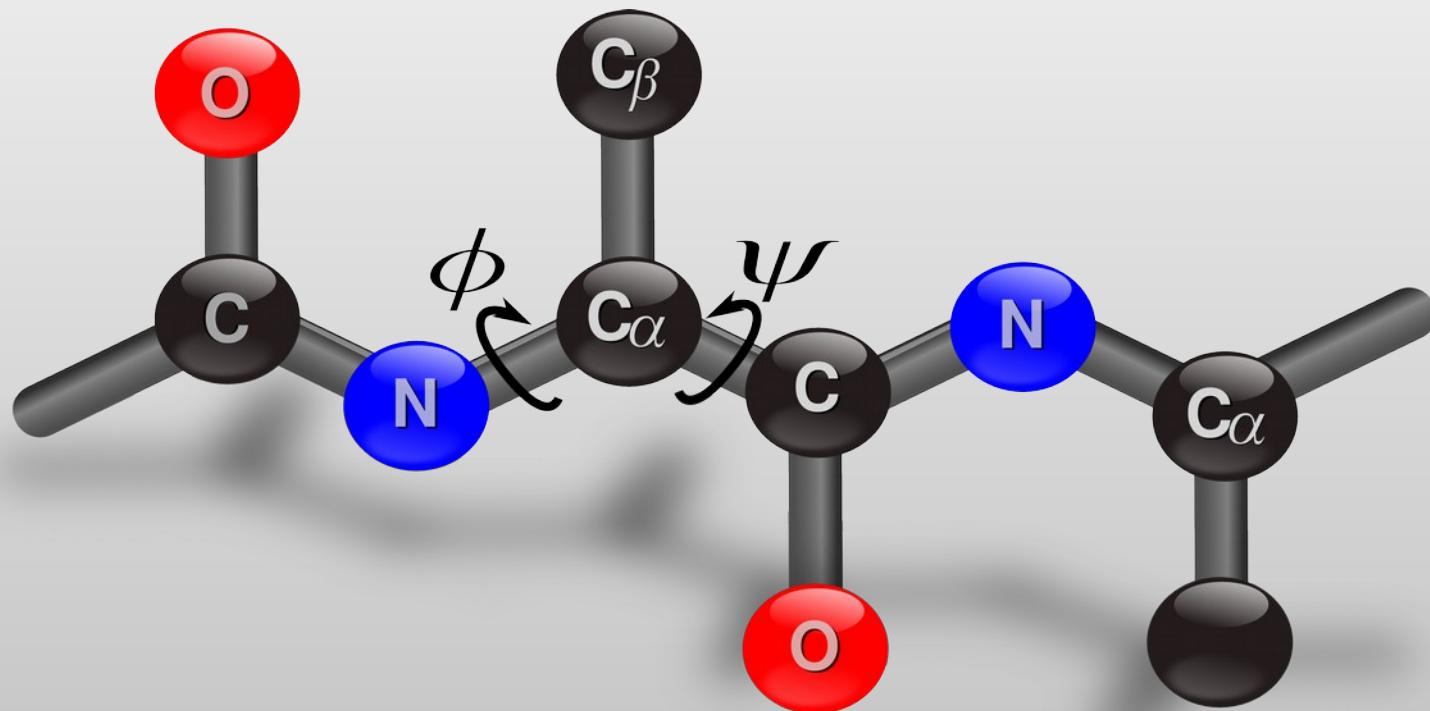
Refinement Techniques

- Single-Atom Drag
 - Over-dragging
- Key-bindings:
 - Triple Refine
 - Single Residue Refine with Auto-accept

Low Resolution Model-Building

- “Backrub” rotamers

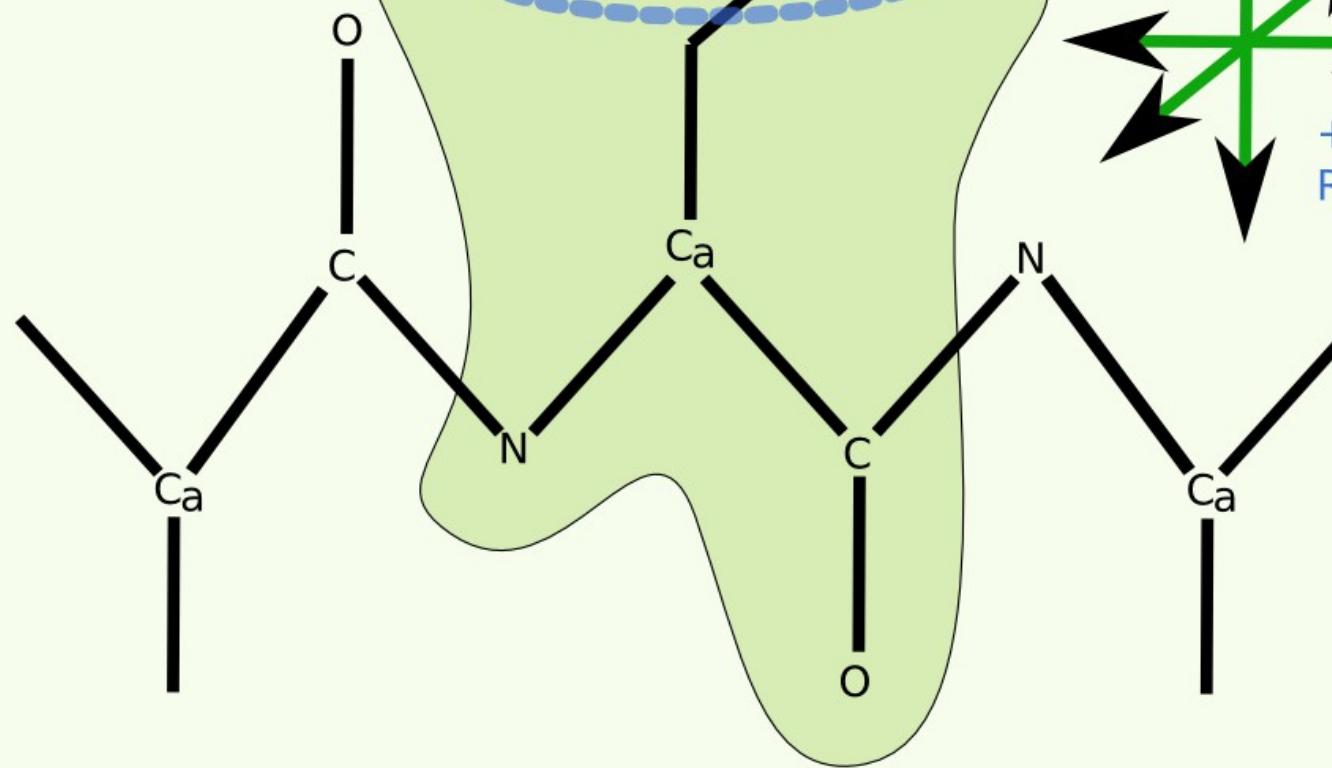
Peptide Backbone Geometry



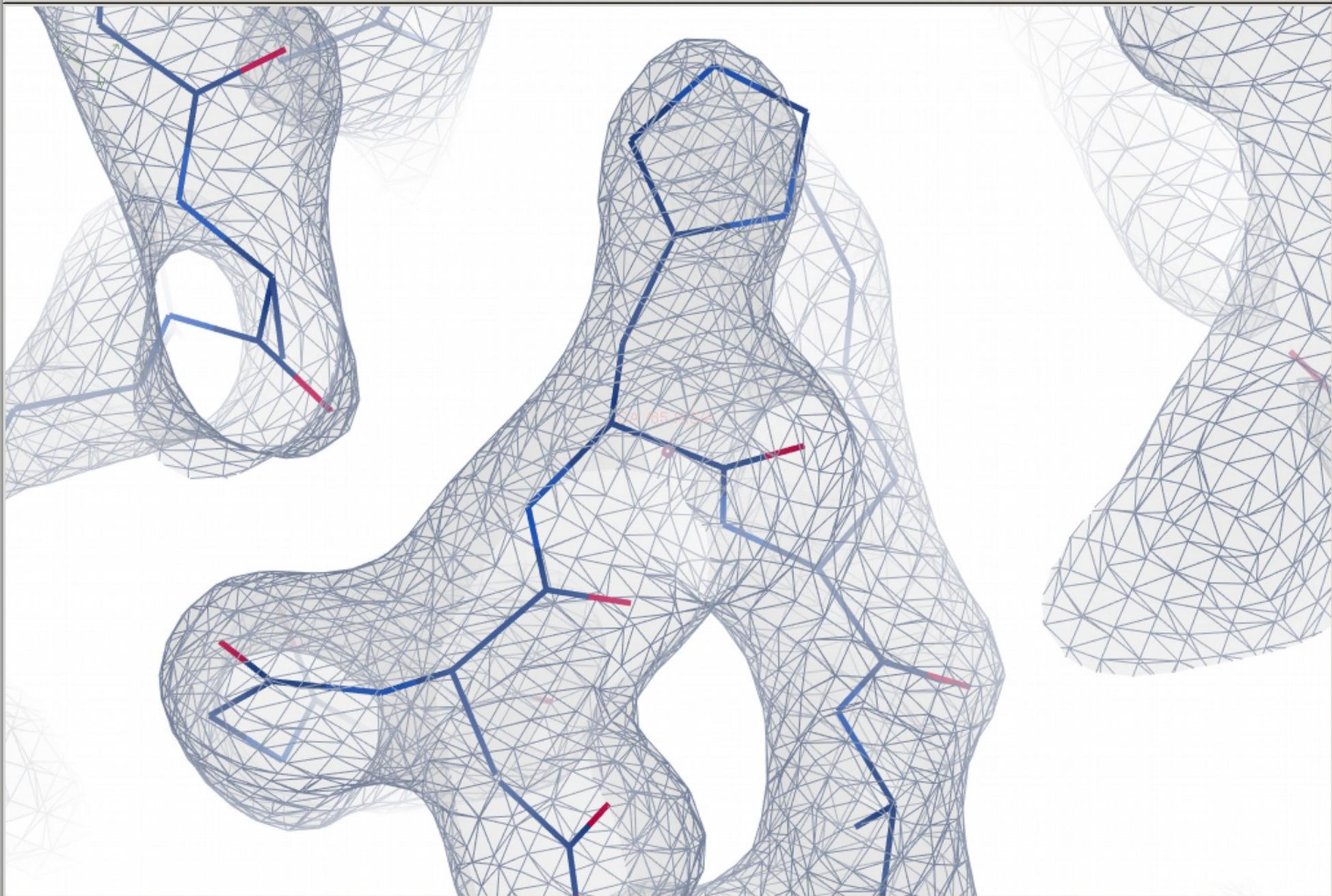
~~Current~~ Low Resolution Rotamer Search

Previous

Rotamer Search



+ Rigid Body
Refinement



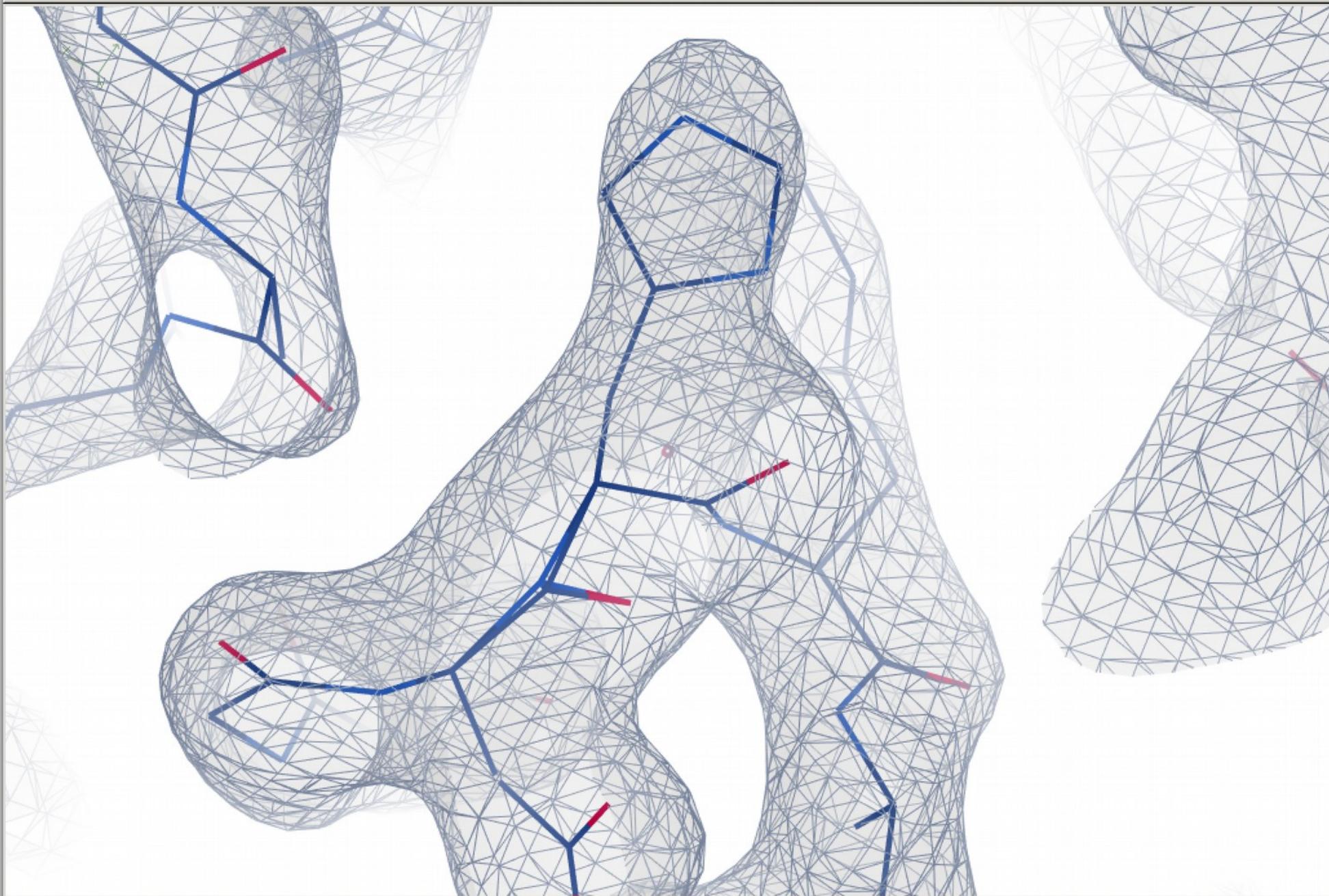
(mol. no: 3) CA /1/A/85 HIS occ: 1.00 bf: 19.16 ele: C pos: (57.45,15.65,14.20)



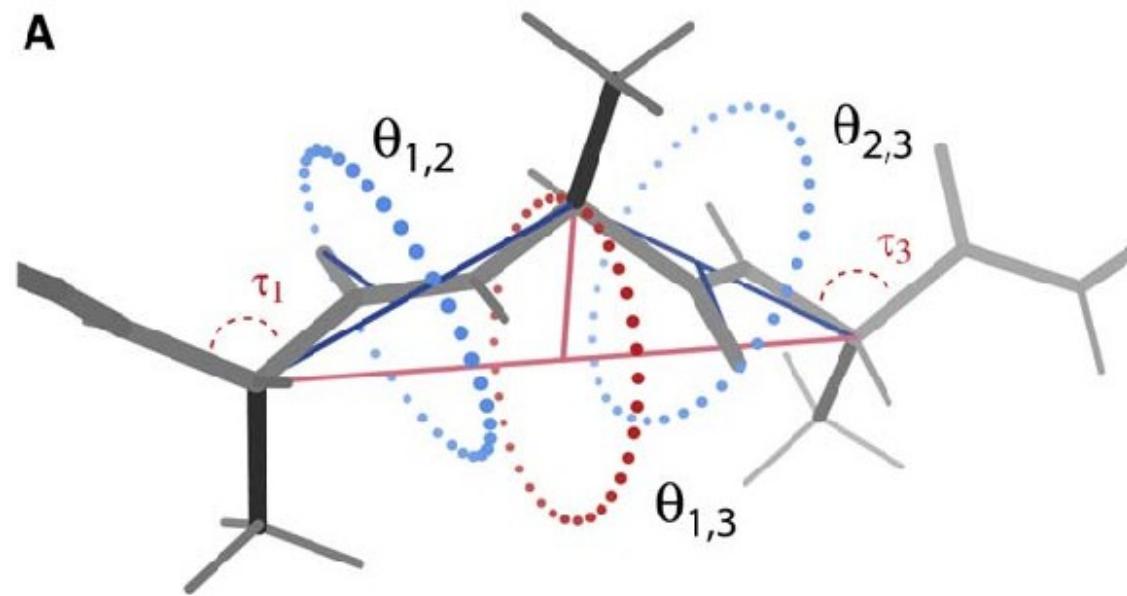
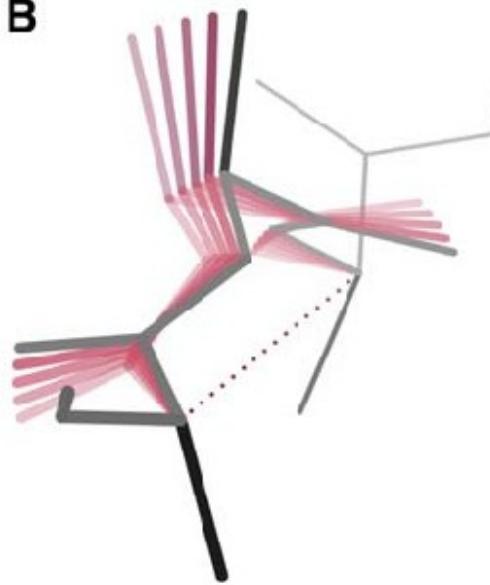
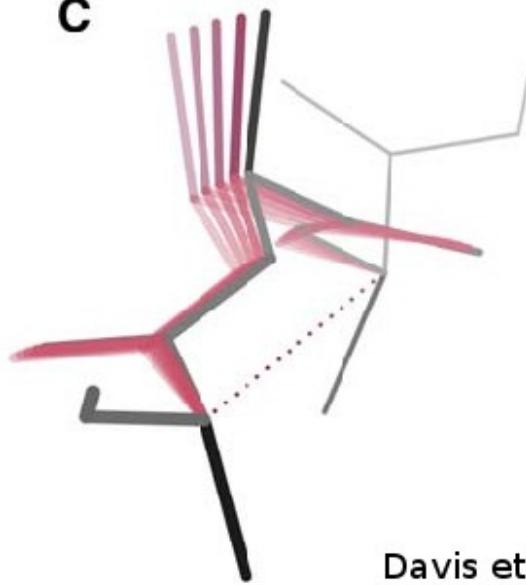
Coot 0.8.7-pre EL (revision count 6456)

File Edit Calculate Draw Measures Validate HID About Ligand Extensions Debug

Reset View Display Manager Ligand Builder Sphere Refine Sphere Refine + Backrub Rotamers

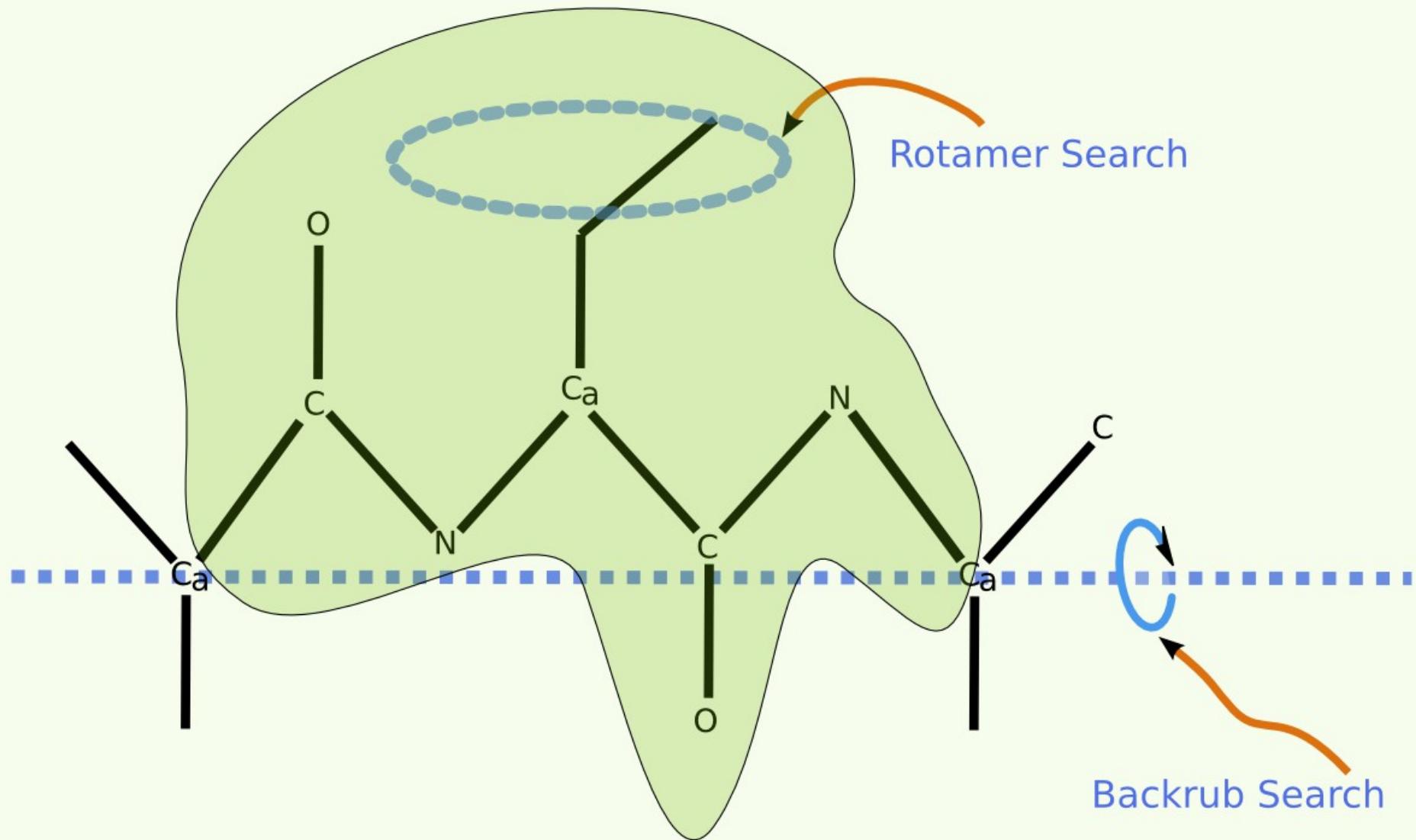


(mol. no: 3) CA /1/A/85 HIS occ: 1.00 bf: 19.16 ele: C pos: (57.45,15.65,14.20)

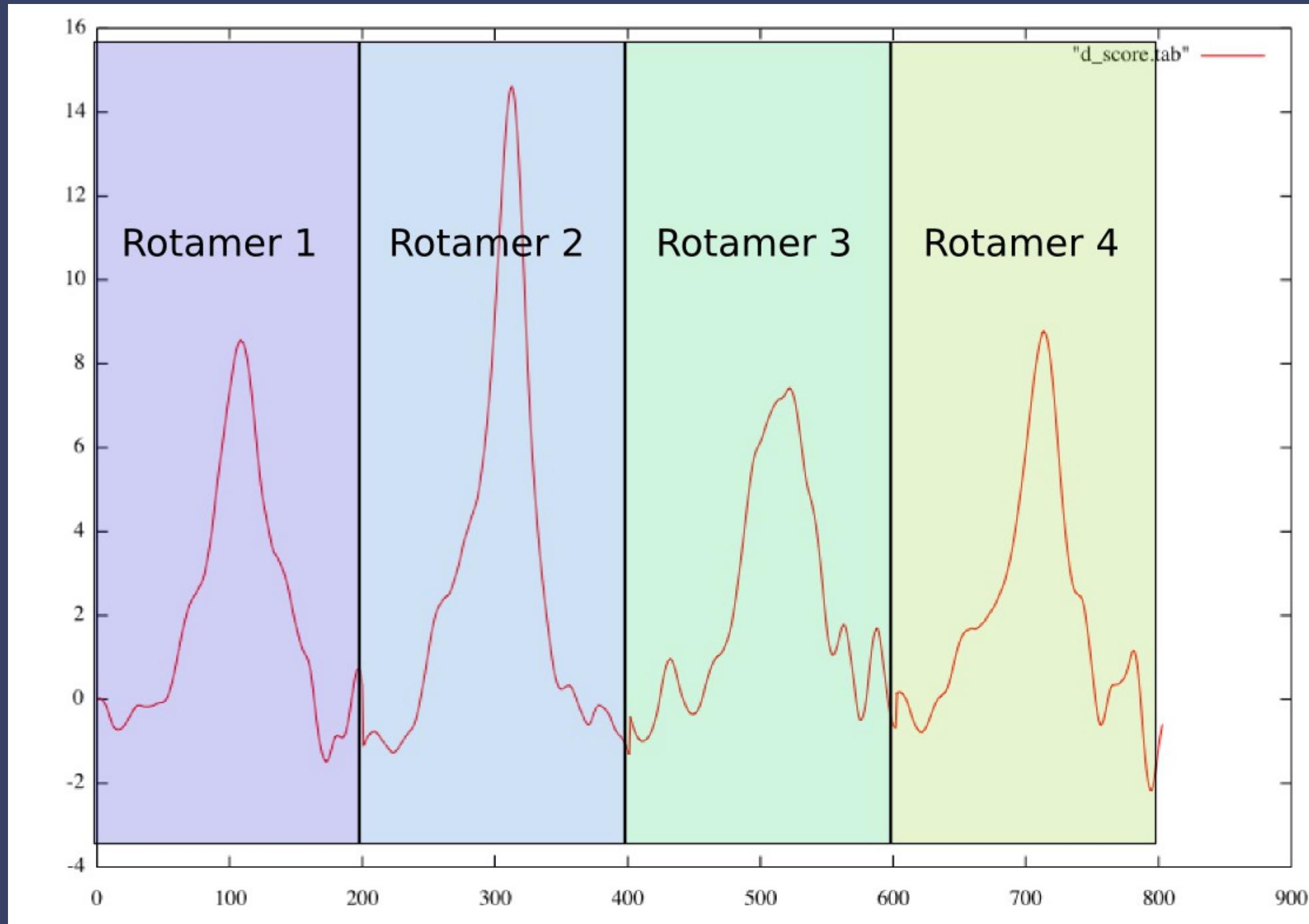
A**B****C**

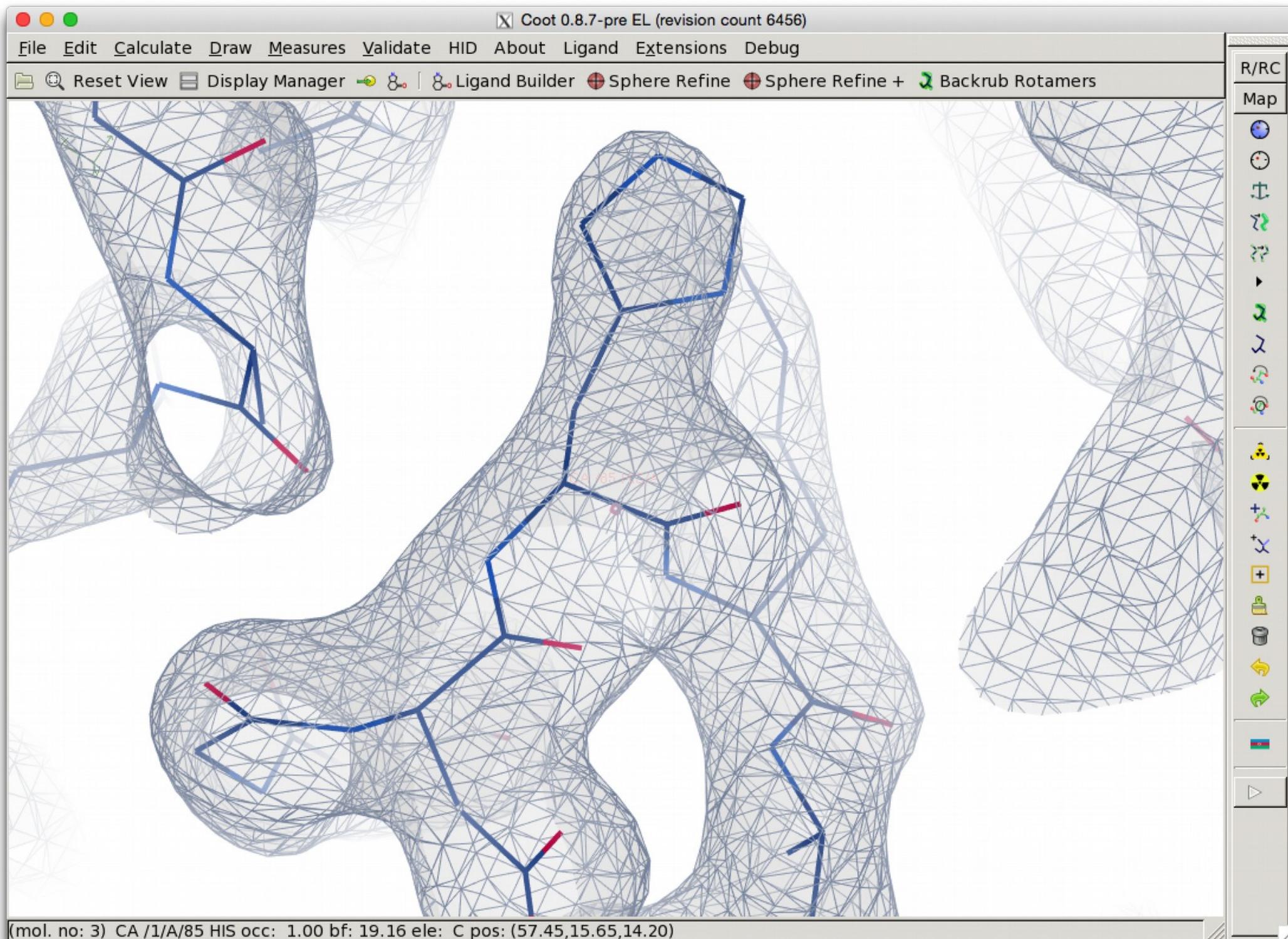
Davis et al. (2006) Structure

New Low Resolution Rotamer Search



After Fitting Tools in KING/Molprobity



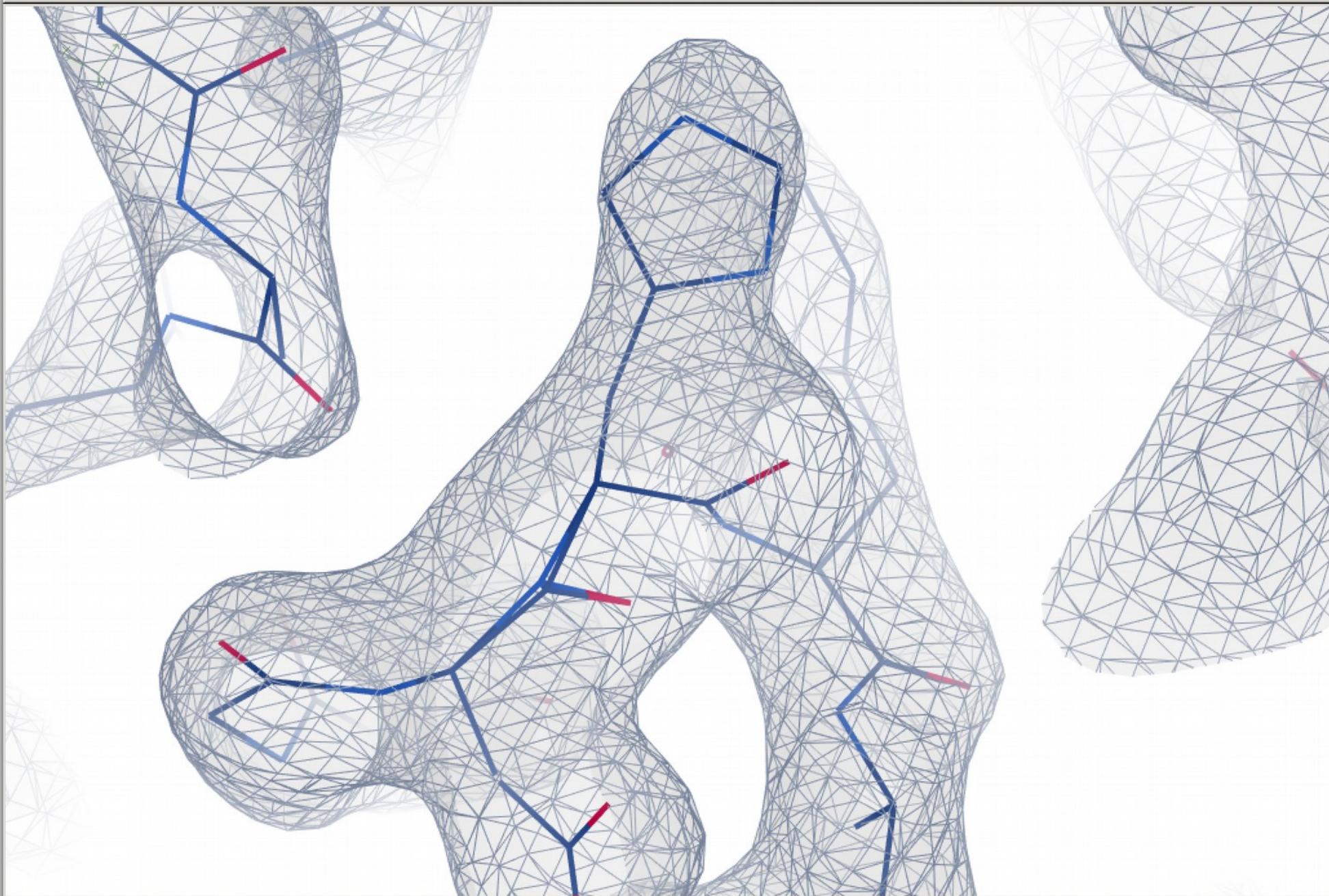




Coot 0.8.7-pre EL (revision count 6456)

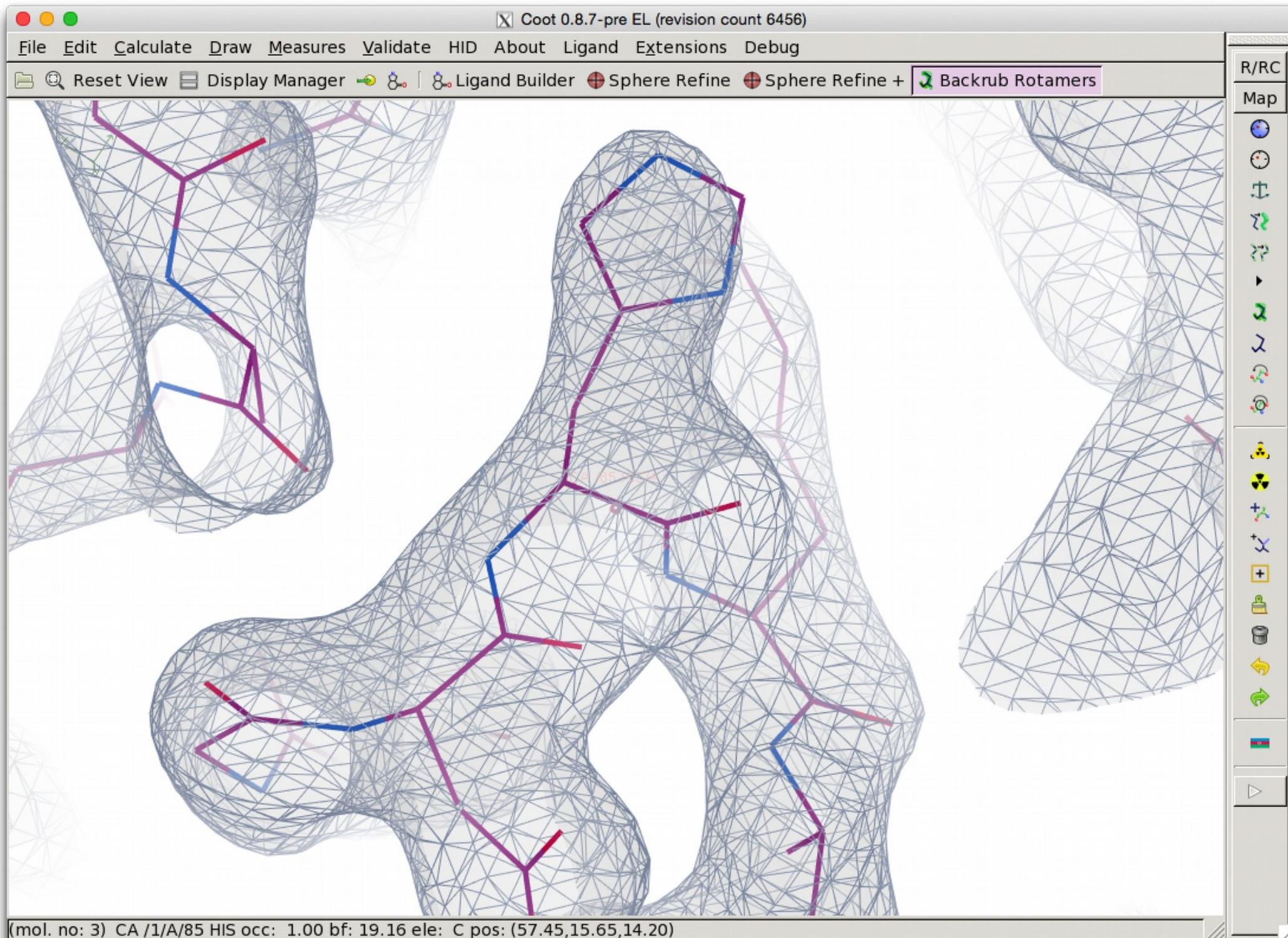
File Edit Calculate Draw Measures Validate HID About Ligand Extensions Debug

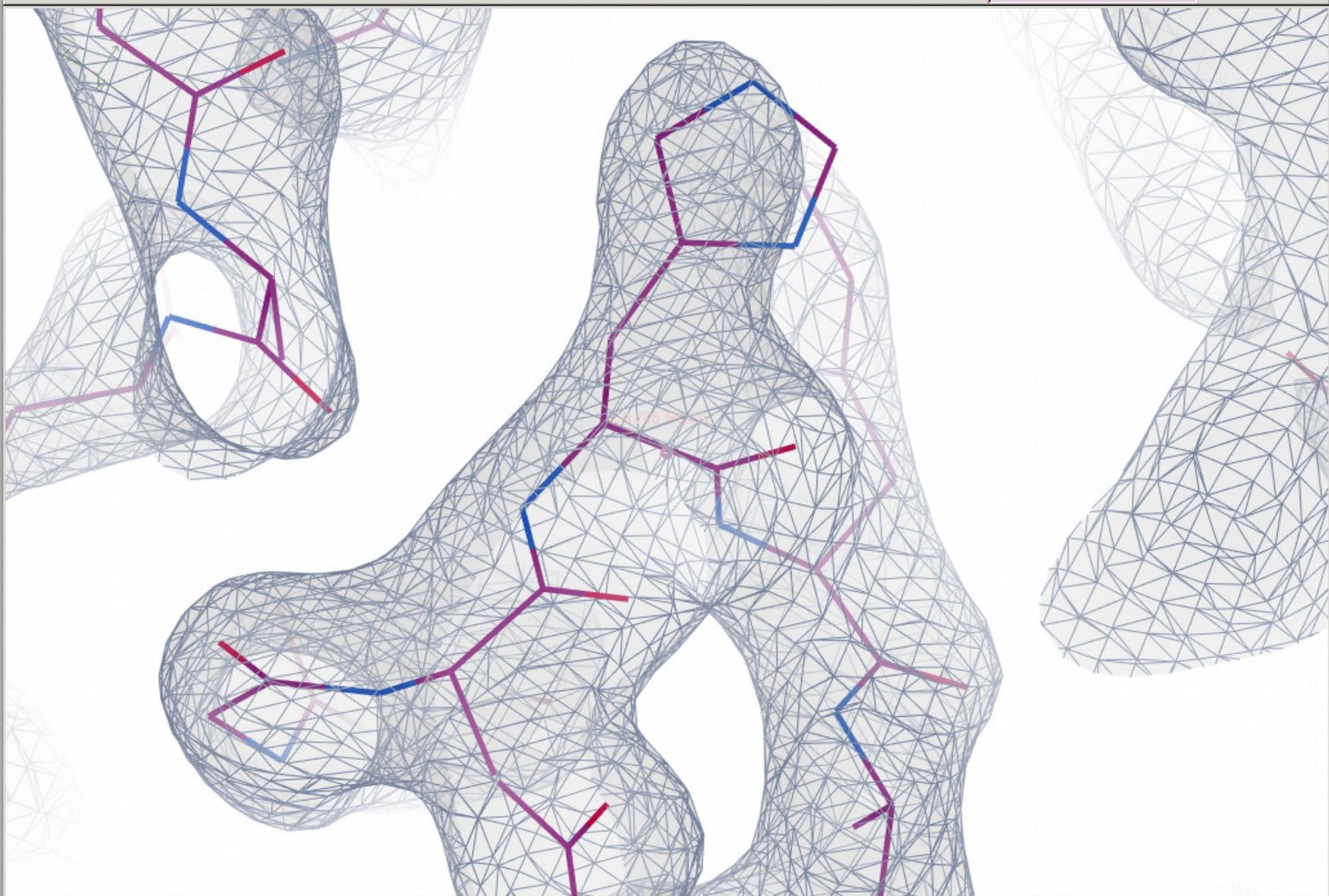
Reset View Display Manager Ligand Builder Sphere Refine Sphere Refine + Backrub Rotamers



(mol. no: 3) CA /1/A/85 HIS occ: 1.00 bf: 19.16 ele: C pos: (57.45,15.65,14.20)









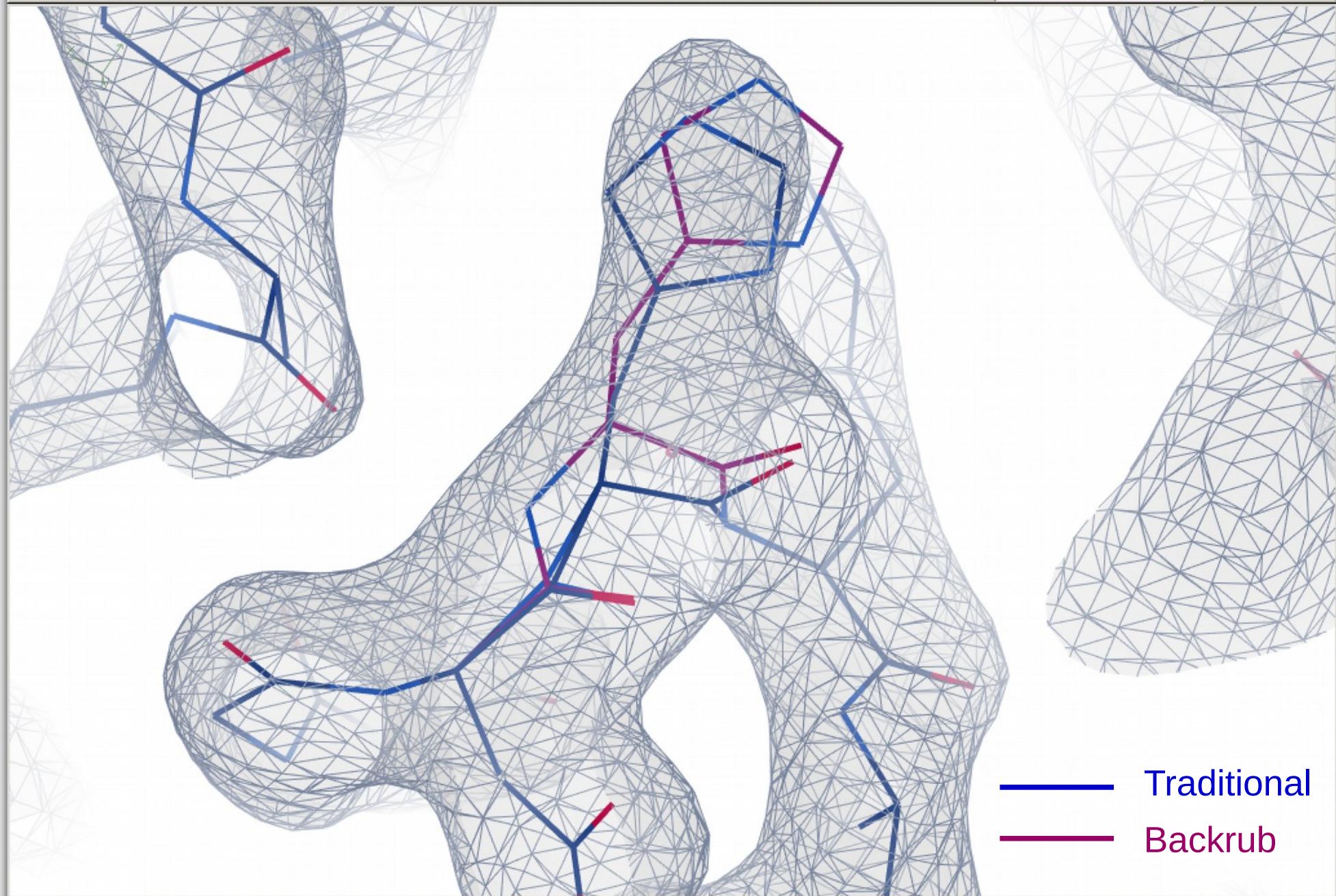
Coot 0.8.7-pre EL (revision count 6456)

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

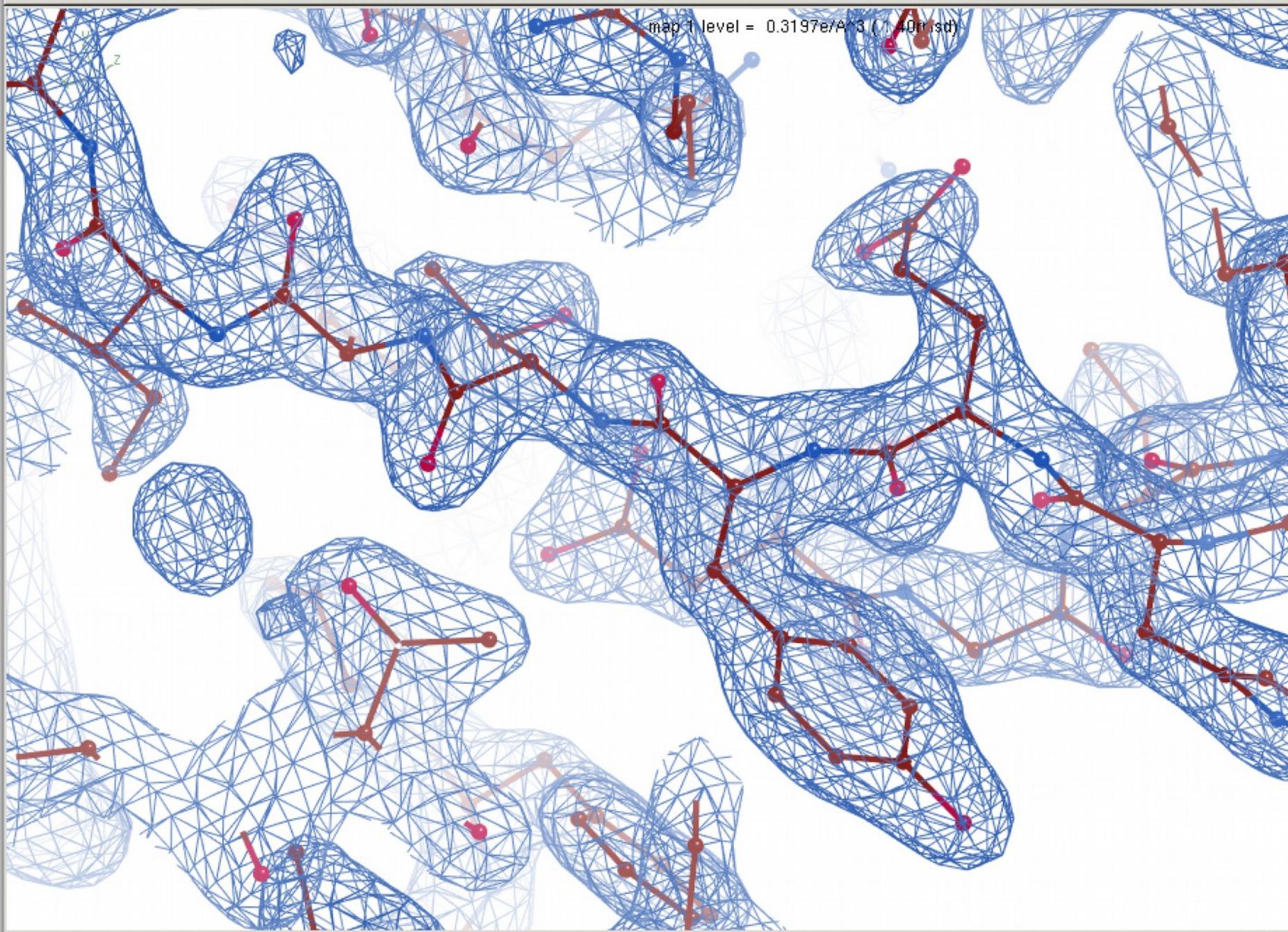


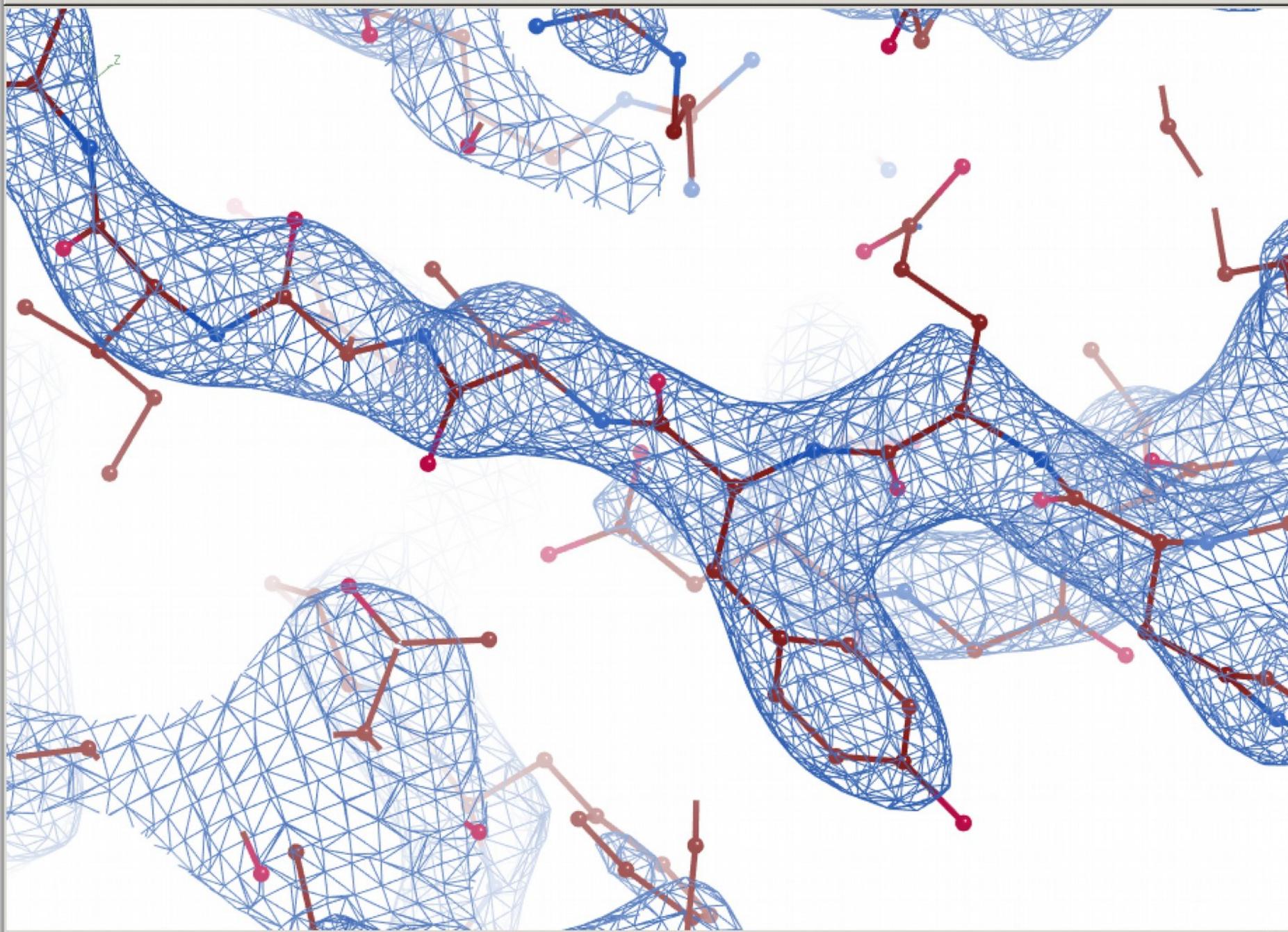
(mol. no: 3) CA /1/A/89 PHE occ: 1.00 bf: 20.19 ele: C pos: (53.21,12.14, 7.56)

X Coot 0.8.9-pre EL (revision count 6954)

File Edit Calculate Draw Measures Validate HID About Extensions Debug Glyco ProSMART Ligand Morph

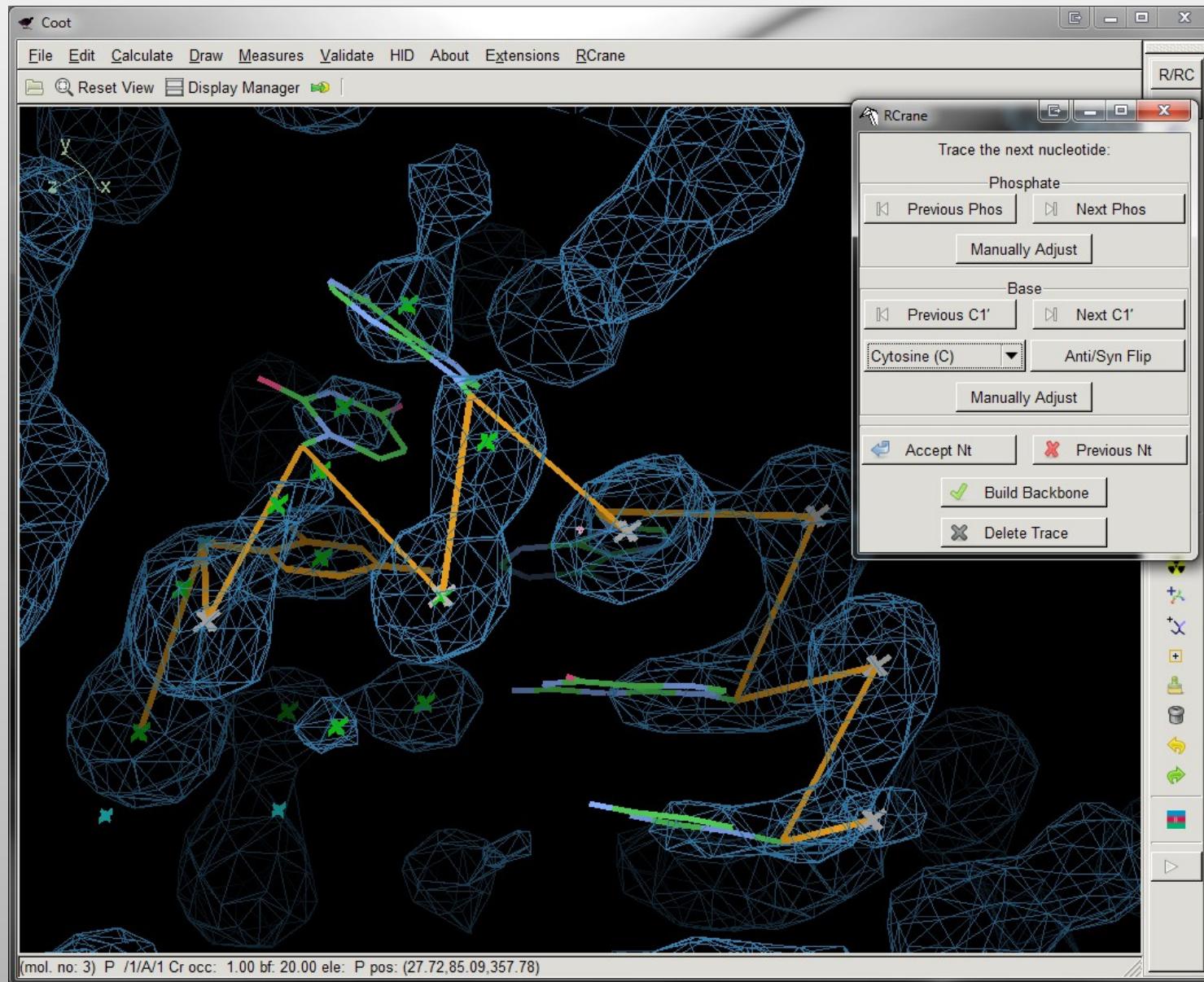
Reset View Display Manager Ligand Builder Sphere Refine Sphere Refine + Sphere Regularization +



R/RC
Map

Building RNA...

RCrane: Semi-Automated Building of RNA



Ana Marie Pyle
Kevin Keating
Yale

Handling EM maps

Coot

File Edit Calculate Draw Measures Validate HID About Extensions

Reset View

Display Manager

R/RC

Map

○

○

+

○

○

○

○

○

○

○

○

○

○

○

○

○

○

○

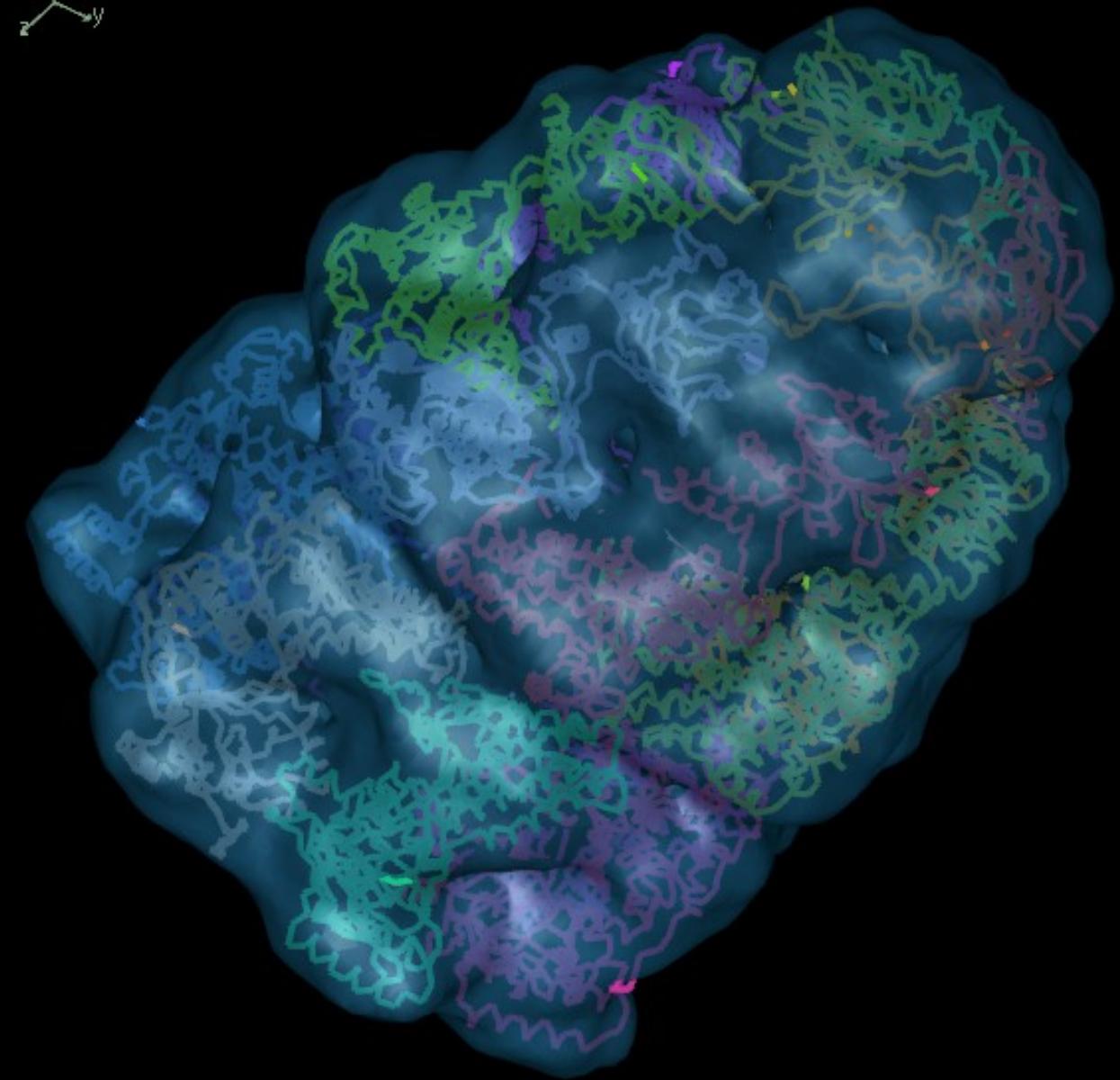
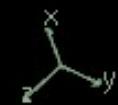
○

○

○

○

○



...coordinates file /home/paule/em-challenge/groEL/1GRU.pdb.gz. Molecule number 1 created.

Coot

File Edit Calculate Draw Measures Validate HID About Extensions

Reset View

Display Manager

R/RC

Map

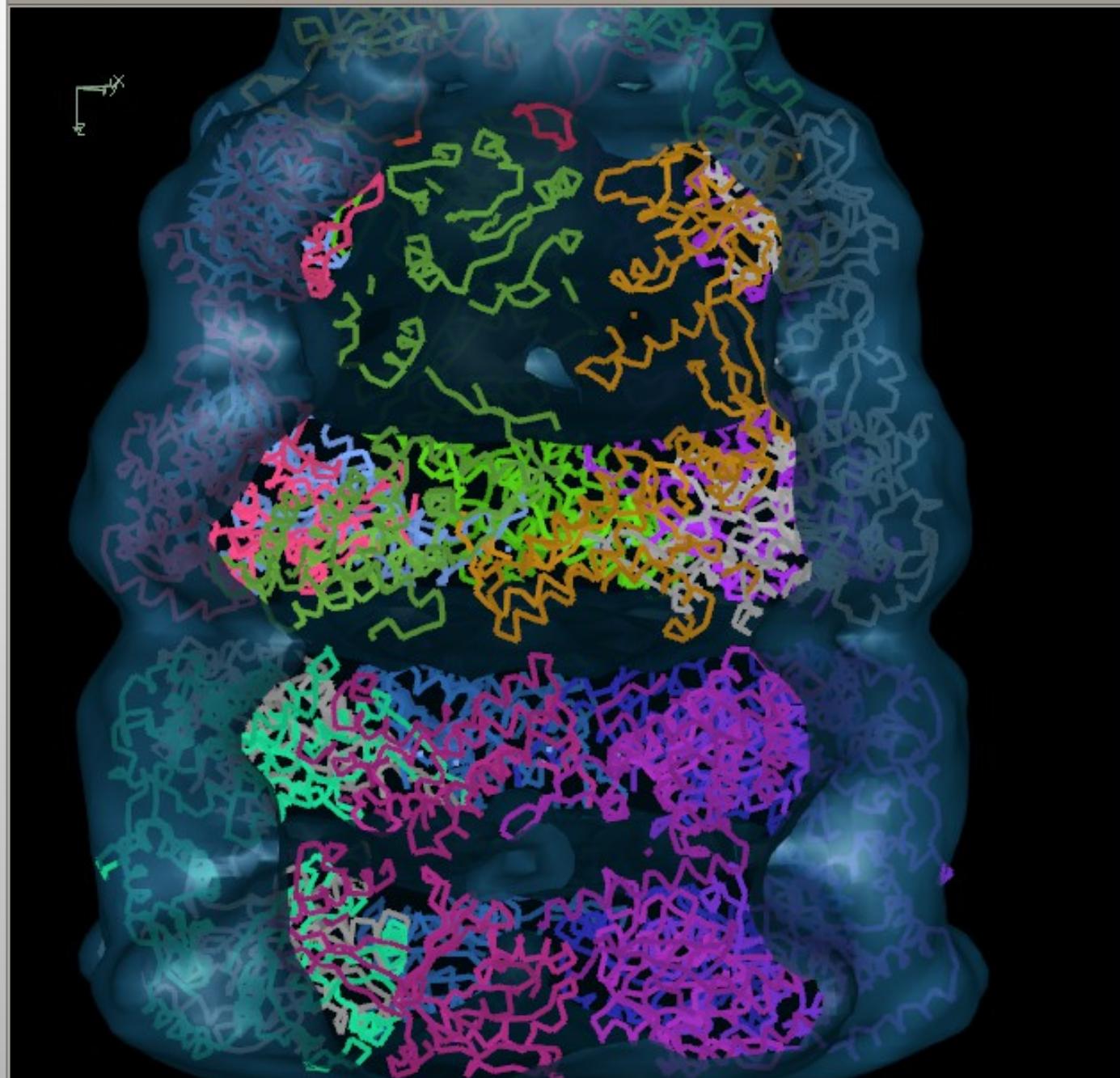
Atom

Site

Vector

Surface

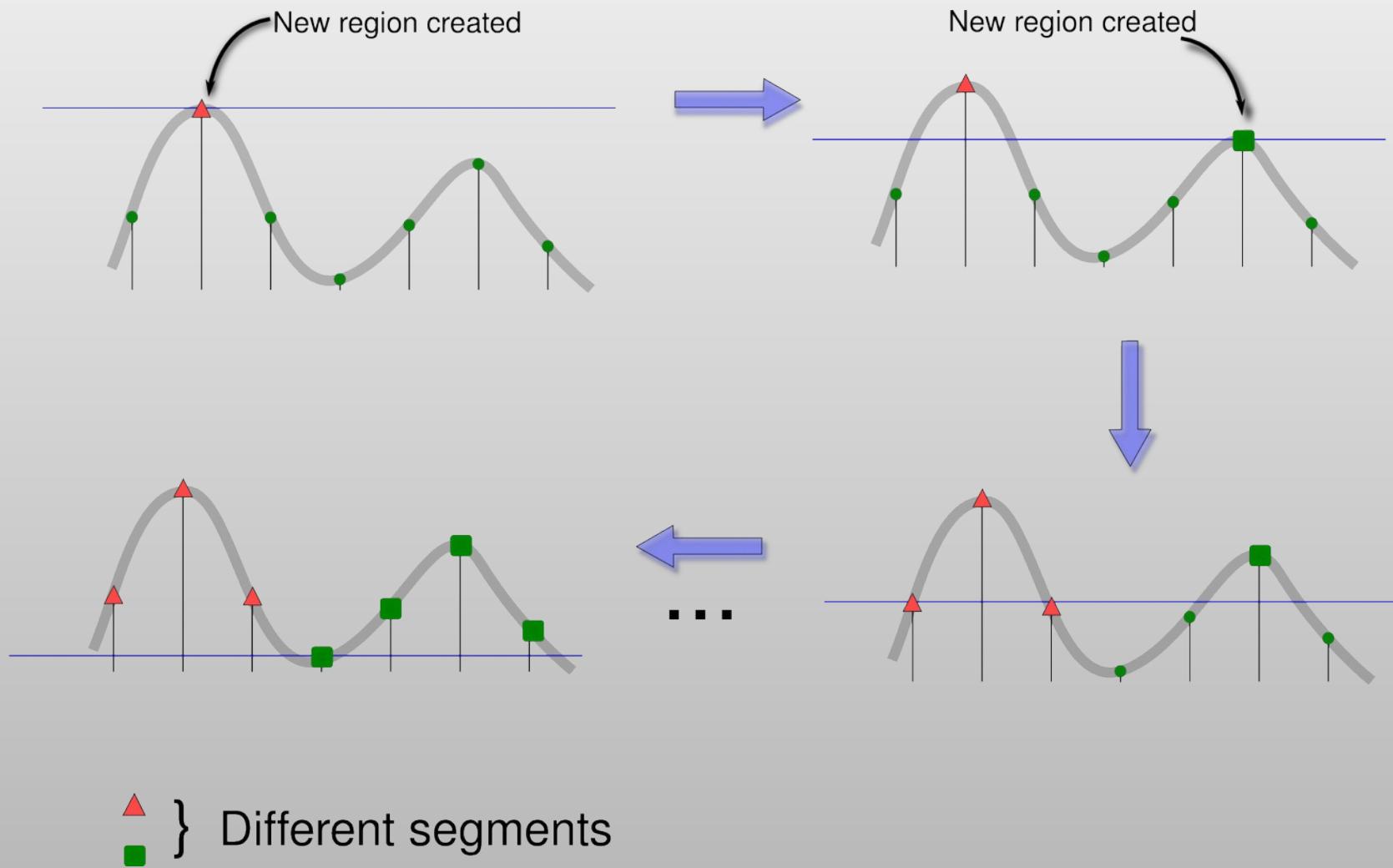
Map

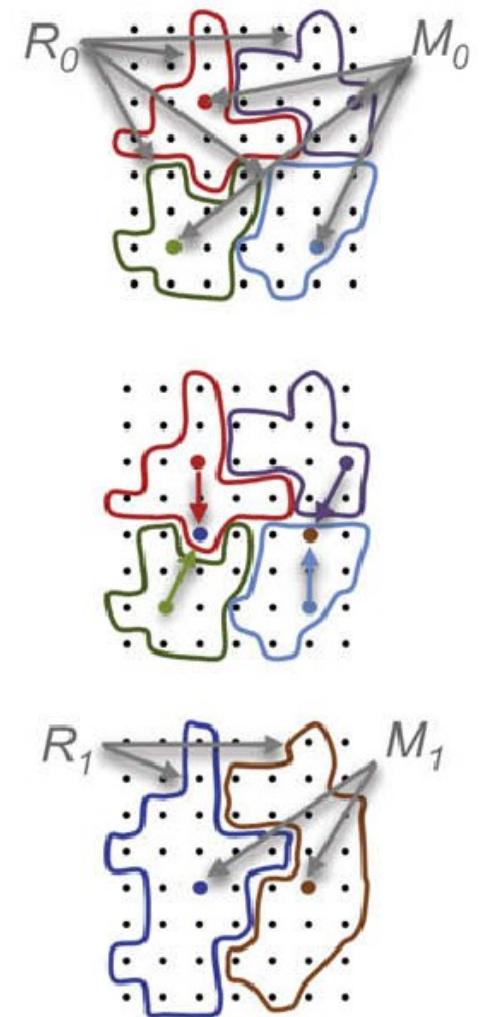
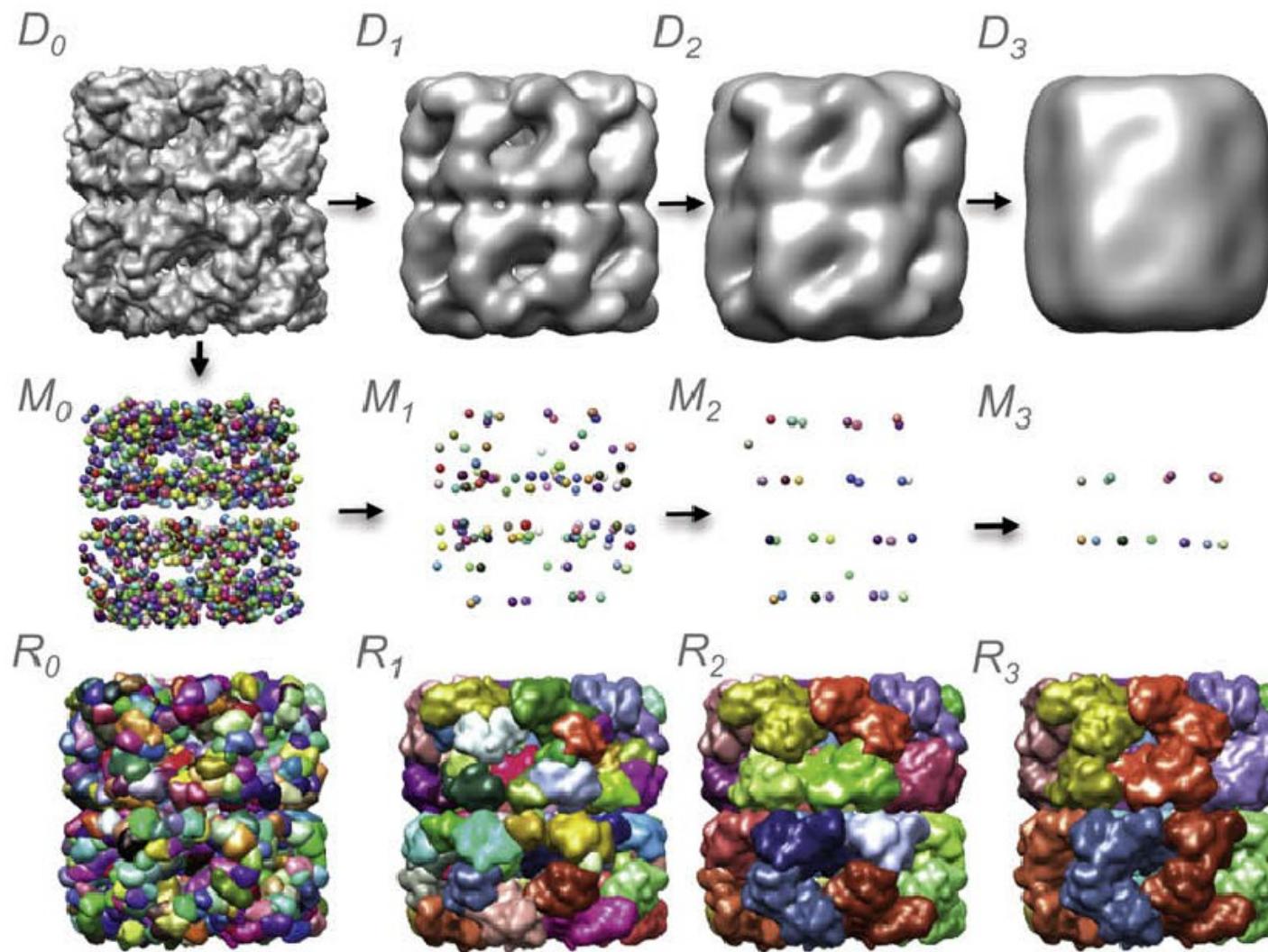


(mol. no: 1) CD /1/M/209 GLU occ: 1.00 bf: 100.09 ele: C pos: (167.80,210.31,219.52)

Partitioning Maps: Watershed Algorithm

1D-analog





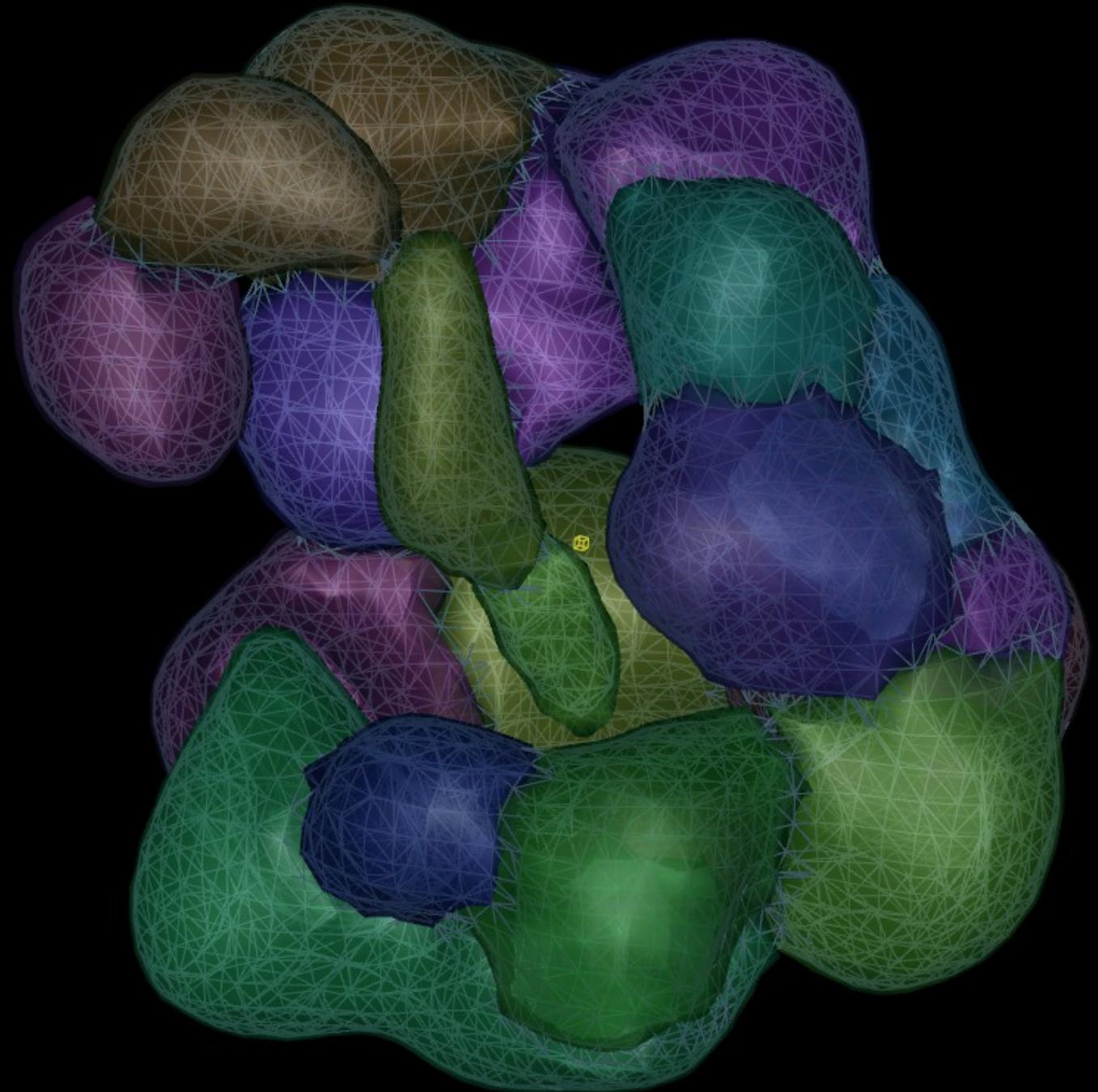
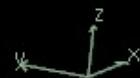
Coot

File Edit Calculate Draw Measures Validate HID About Extensions Solid

Reset View Display Manager

R/RC

Map



Successfully read coordinates file test.pdb. Molecule number 1 created.

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
- Now extended to fit arbitrary atom selections
 - e.g. by Chain

Jiggle Fit: How it Works

- Loop n 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
 - (non-oversharpened map)

Coot 0.8-pre (revision 4826)

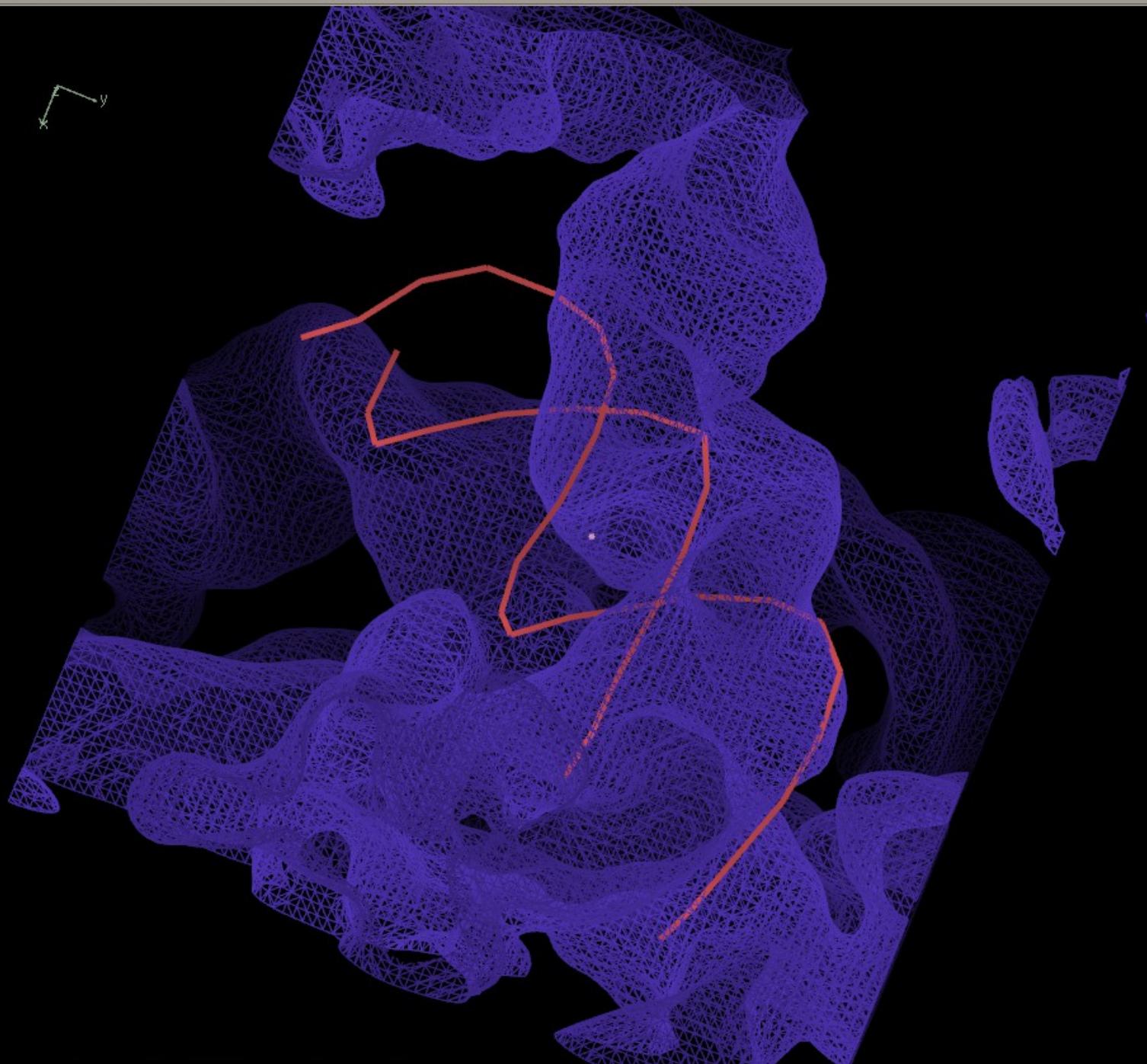
File Edit Calculate Draw Measures Validate HID About Extensions Ligand Morph

Reset View Display Manager Full screen Sphere Refine

R/RC

Map

x
y



(mol. no: 4) P /1/A/2099 U occ: 1.00 bf: 121.66 ele: P pos: (194.71,174.88,278.64)

Coot 0.8-pre (revision 4826)

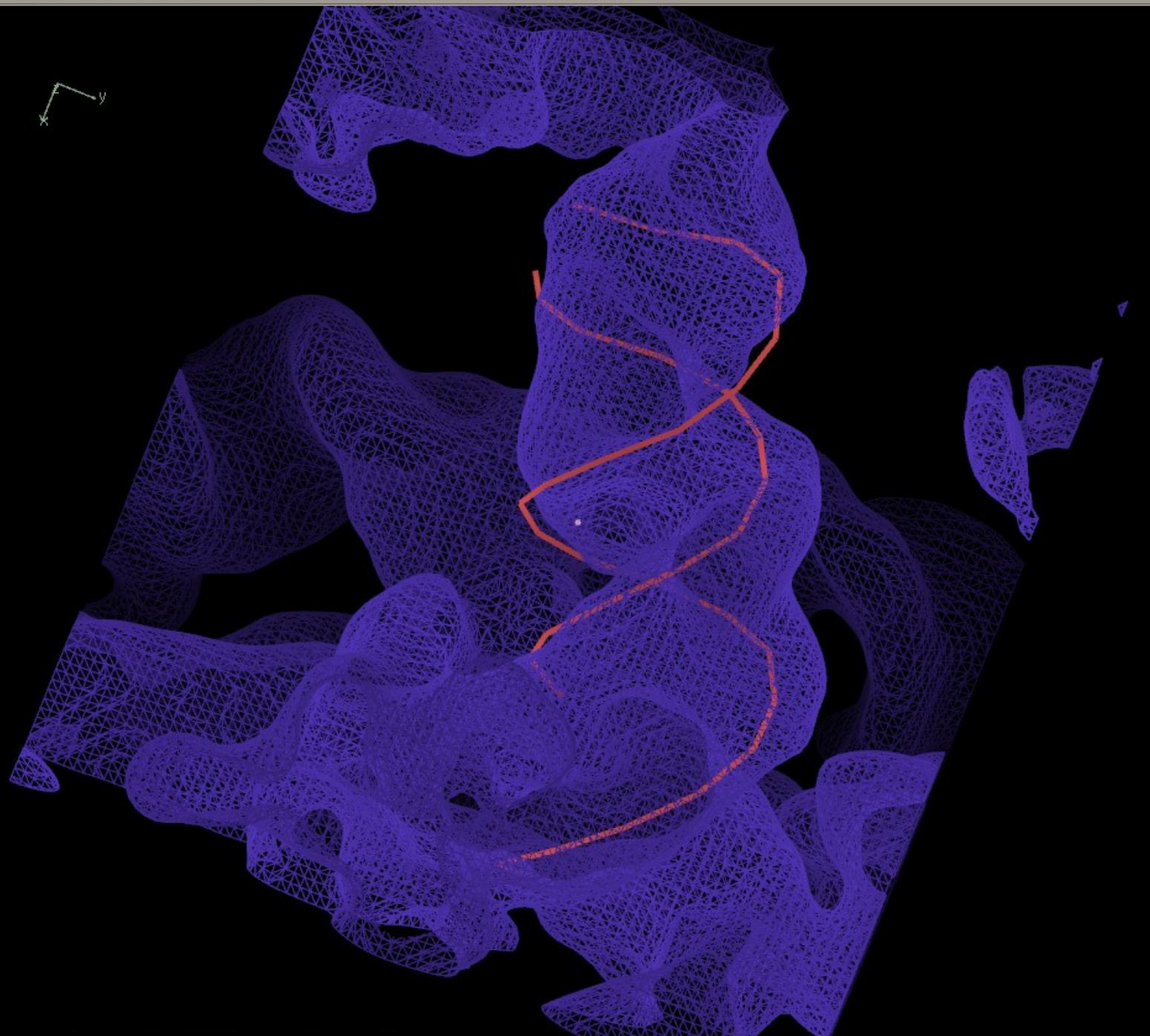
File Edit Calculate Draw Measures Validate HID About Extensions Ligand Morph

Reset View Display Manager Full screen Sphere Refine

R/RC

Map

x
y



(mol. no: 4) P /1/A/2099 U occ: 1.00 bf: 121.66 ele: P pos: (194.71,174.88,278.64)

Coot 0.8-pre (revision 4826)

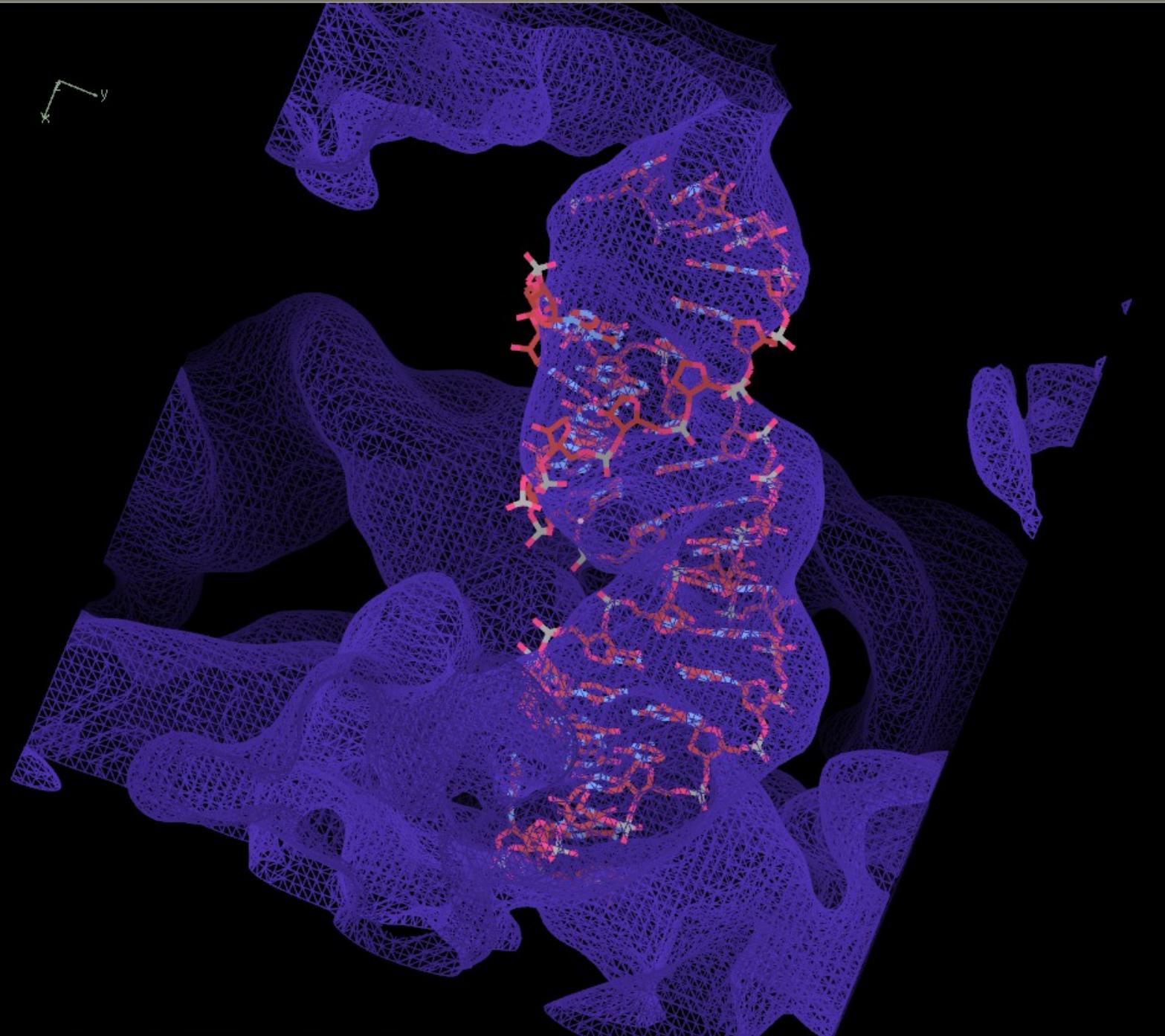
File Edit Calculate Draw Measures Validate HID About Extensions Ligand Morph

Reset View Display Manager Full screen Sphere Refine

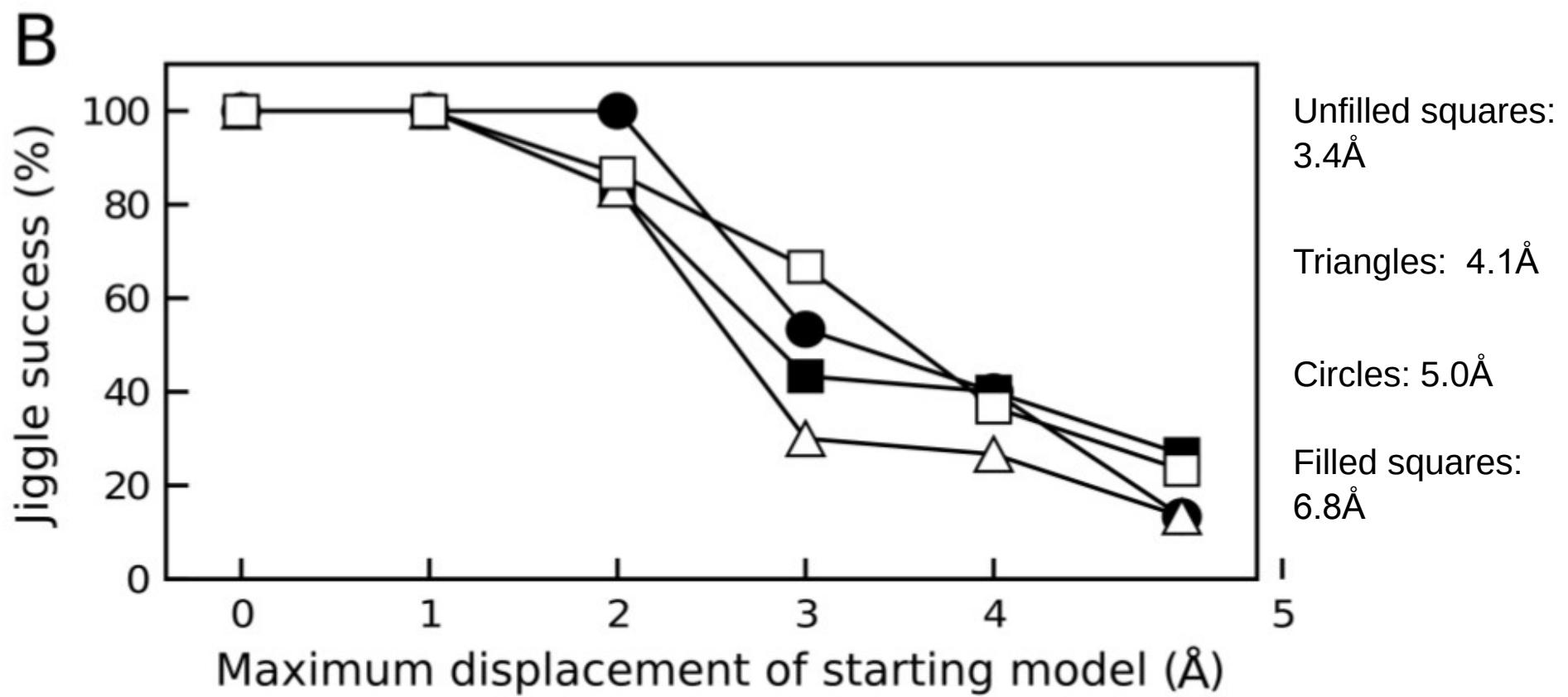
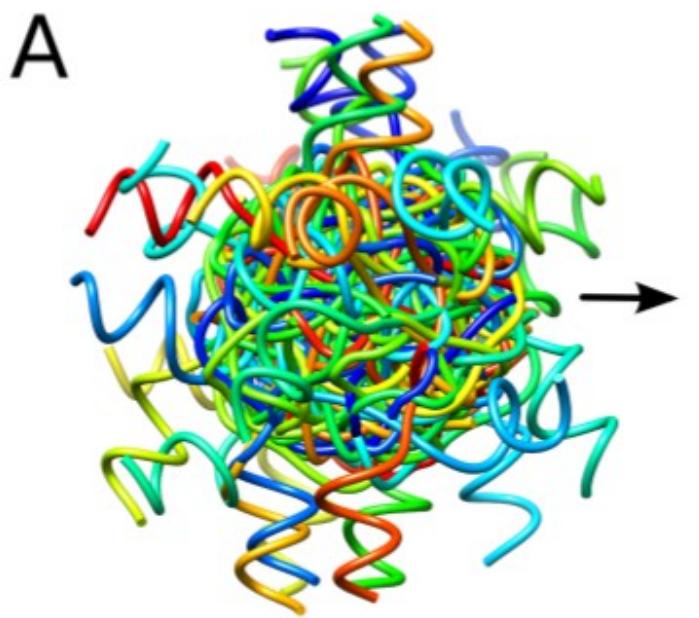
R/RC

Map

x
y



(mol. no: 4) P /1/A/2099 U occ: 1.00 bf: 121.66 ele: P pos: (194.71,174.88,278.64)



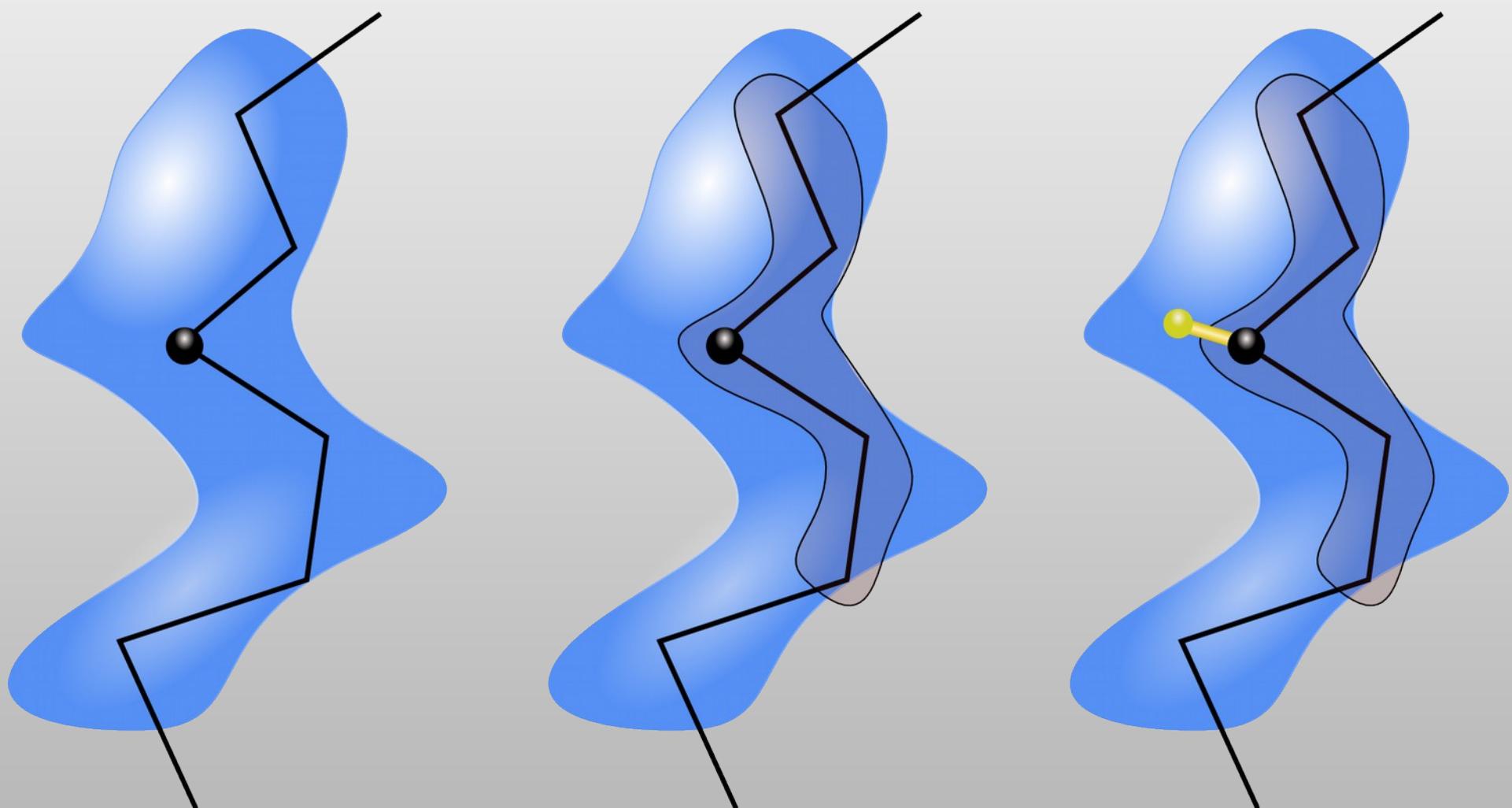
So we have our ideal RNA or homologous protein sitting roughly in the density

(not a great fit)

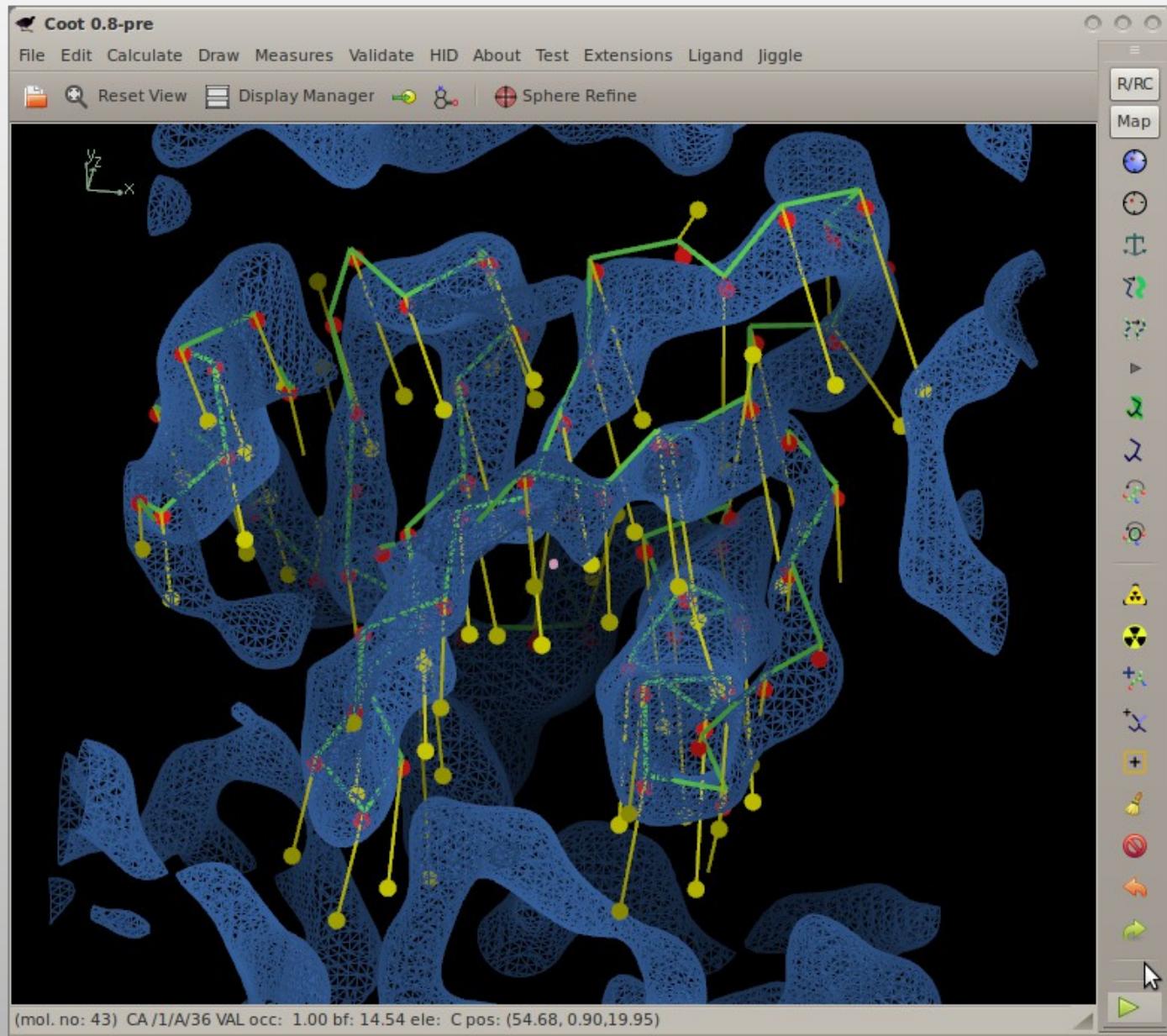
Model Morphing: How it Works

- For each residue in a chain, we ask:
 - where does a small fragment centred on this residue want to go?
 - (Robust) average the transformations and apply them on a per-residue basis
- Repeat

Model Morphing: Generating the Raw RTs

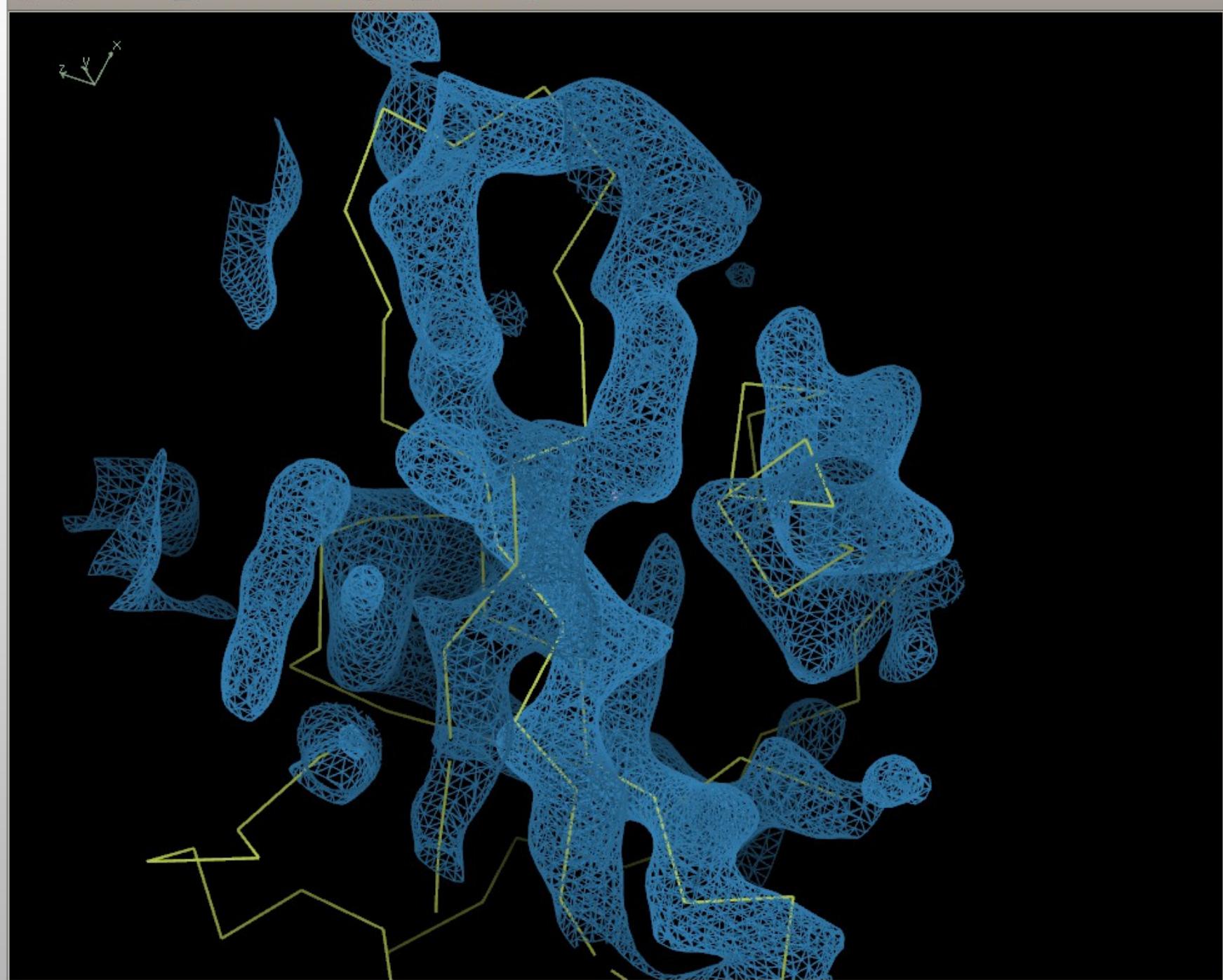


Model Morphing: Example



Model Morphing: Robust Averaging

- What are the residues in the environment of a residue?
 - What are their RTs?
 - Create a metric 'distance', sort on that
 - Discard the top and bottom 25%
 - Use remaining RTs to generate average
 - ...which is then applied to central residue
- Repeat for all residues
- Larger environment radii make the shifts smaller/more conservative
 - More cycles needed



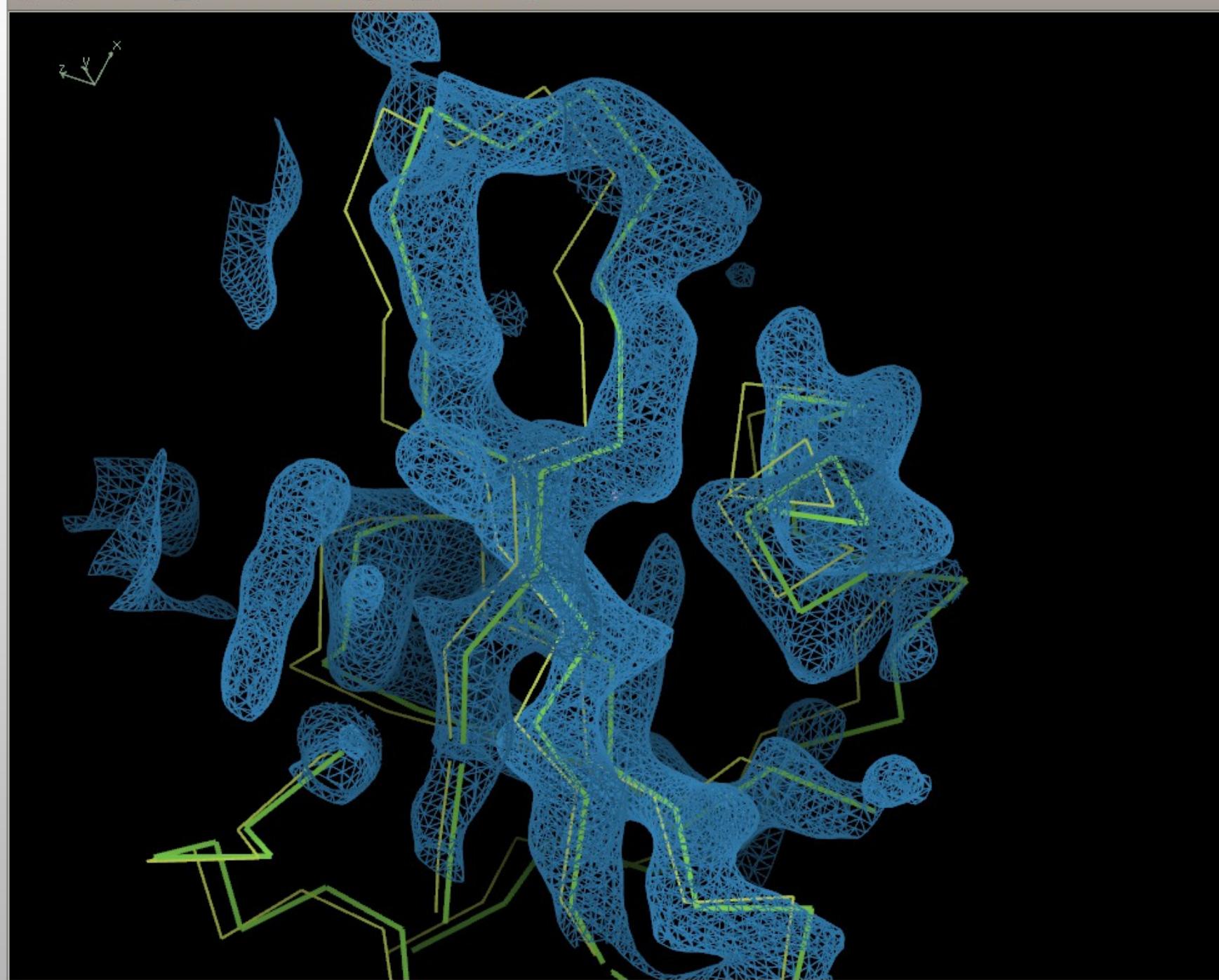
Coot 0.8-pre (revision 4826)

File Edit Calculate Draw Measures Validate HID About Extensions Ligand Morph

Reset View Display Manager

Full screen

Sphere Refine



Successfully read coordinates file /home/paule/autobuild/build-coot+rdkit-pre-release-gtk2-python/share/coot/data/tutorial-modern.pdb. Molecule number 2 created.

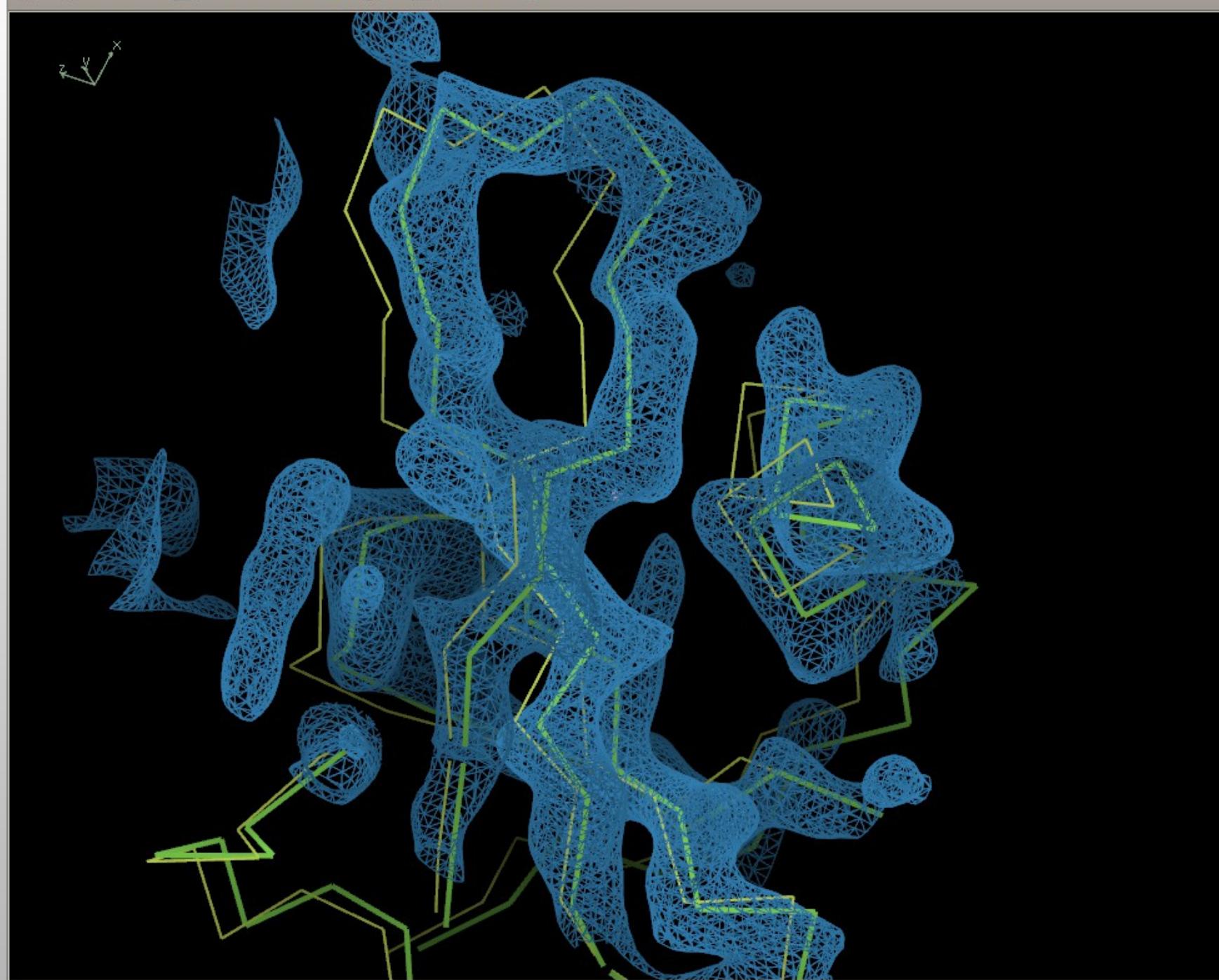
Coot 0.8-pre (revision 4826)

File Edit Calculate Draw Measures Validate HID About Extensions Ligand Morph

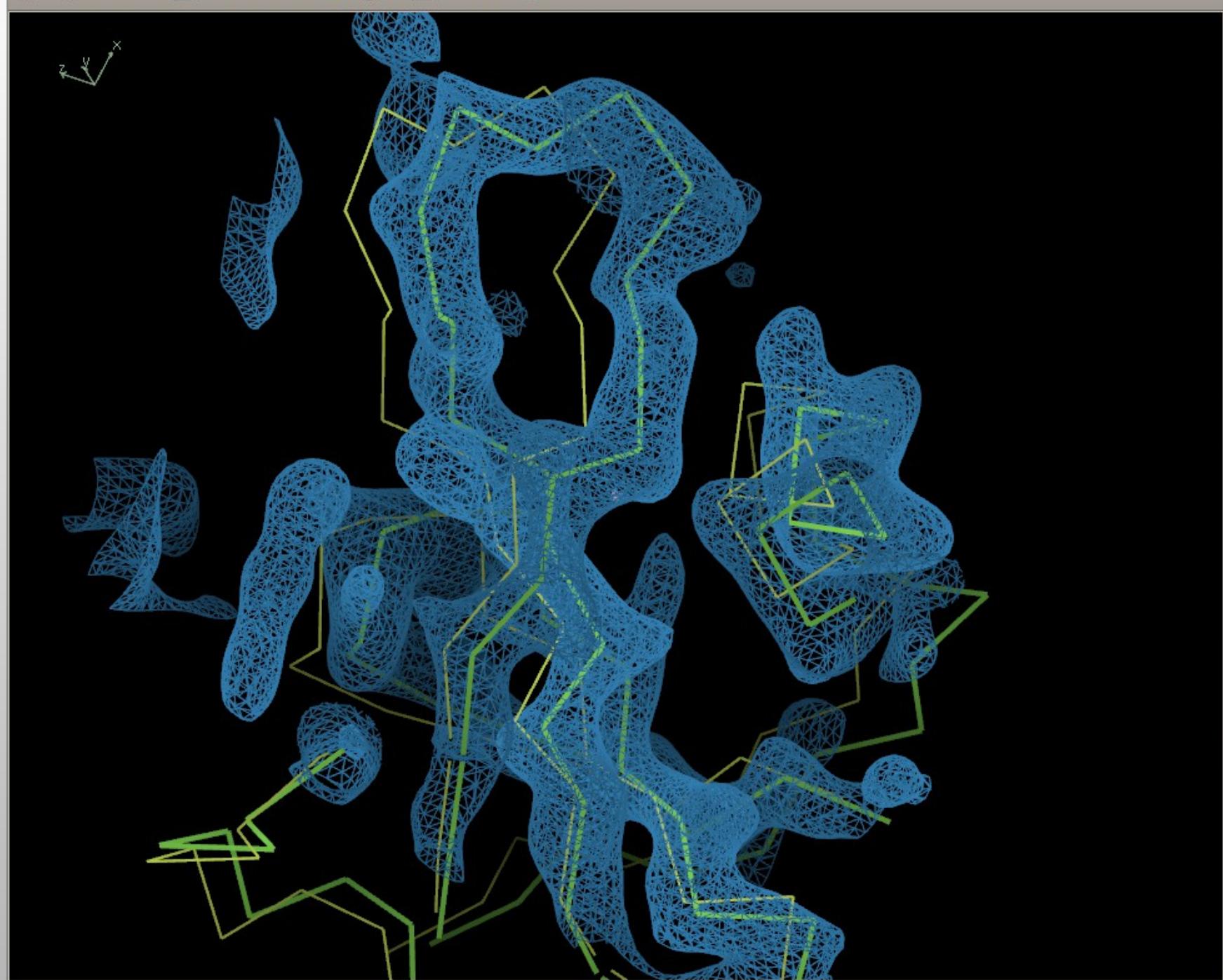
Reset View Display Manager

Full screen

Sphere Refine

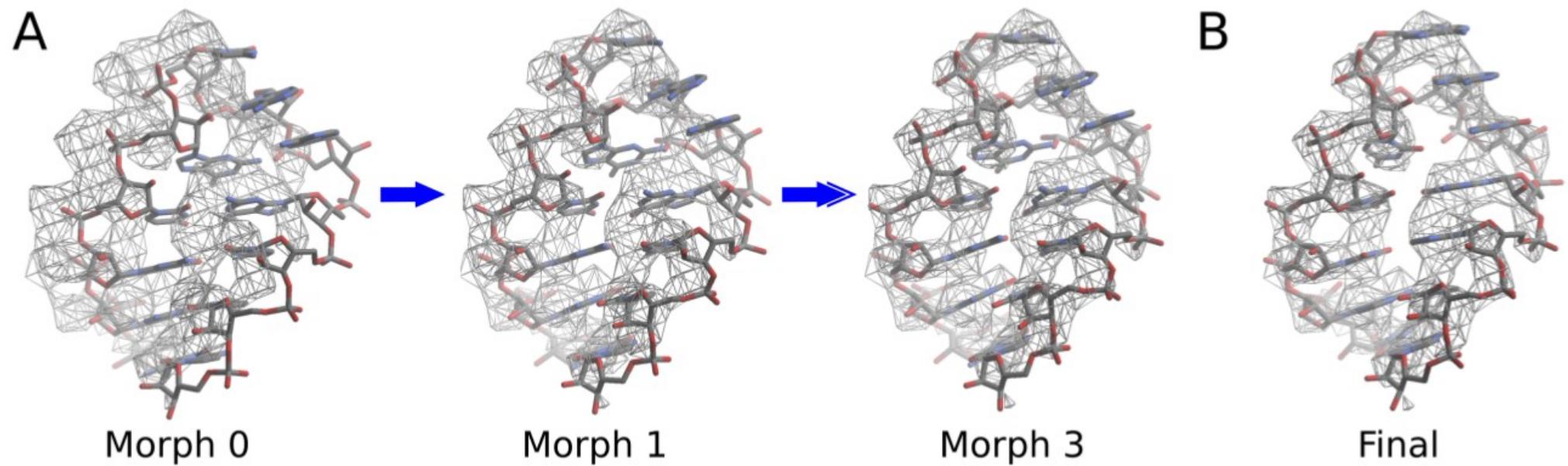


Successfully read coordinates file /home/paule/autobuild/build-coot+rdkit-pre-release-gtk2-python/share/coot/data/tutorial-modern.pdb. Molecule number 2 created.





Model Morphing



Morphing Caveat

- Because the transformations are applied on a per-residue basis
 - peptide N-C geometry is distorted
 - Use hand-in-hand with REFMAC refinement
 - and ProSMART restraints
 - (work in progress)

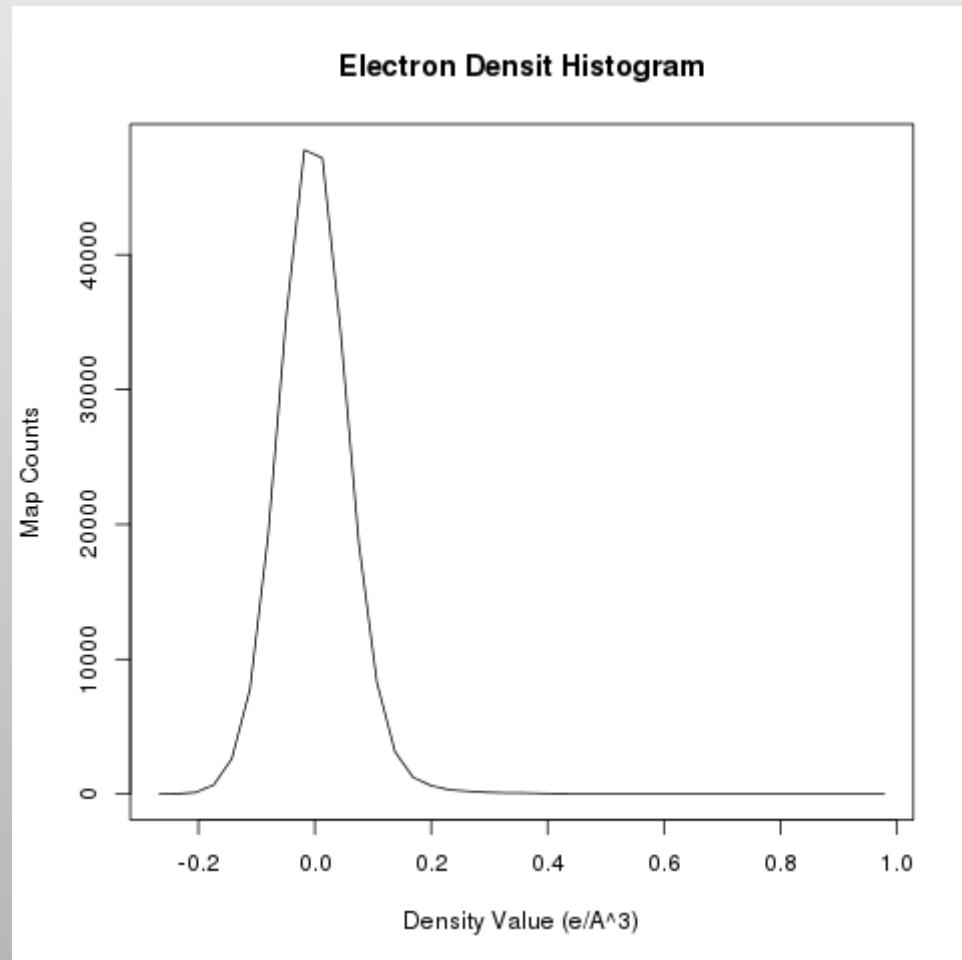
Helix Fitting

- The distribution of electron density is quite unlike that of x-ray maps
 - e.g. You don't see main-chain atoms at 4 rmsd in x-ray maps
 - regions of dense electron density contribute negatively to helix score

Helix Fitting

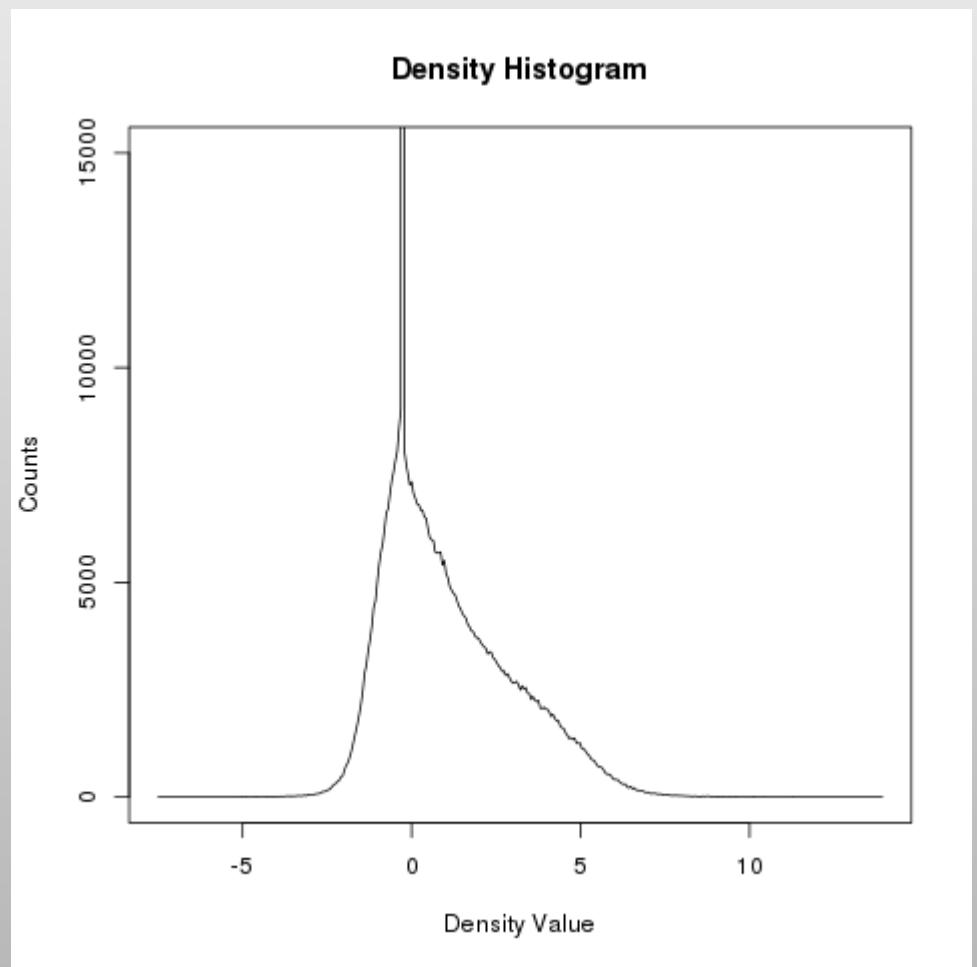
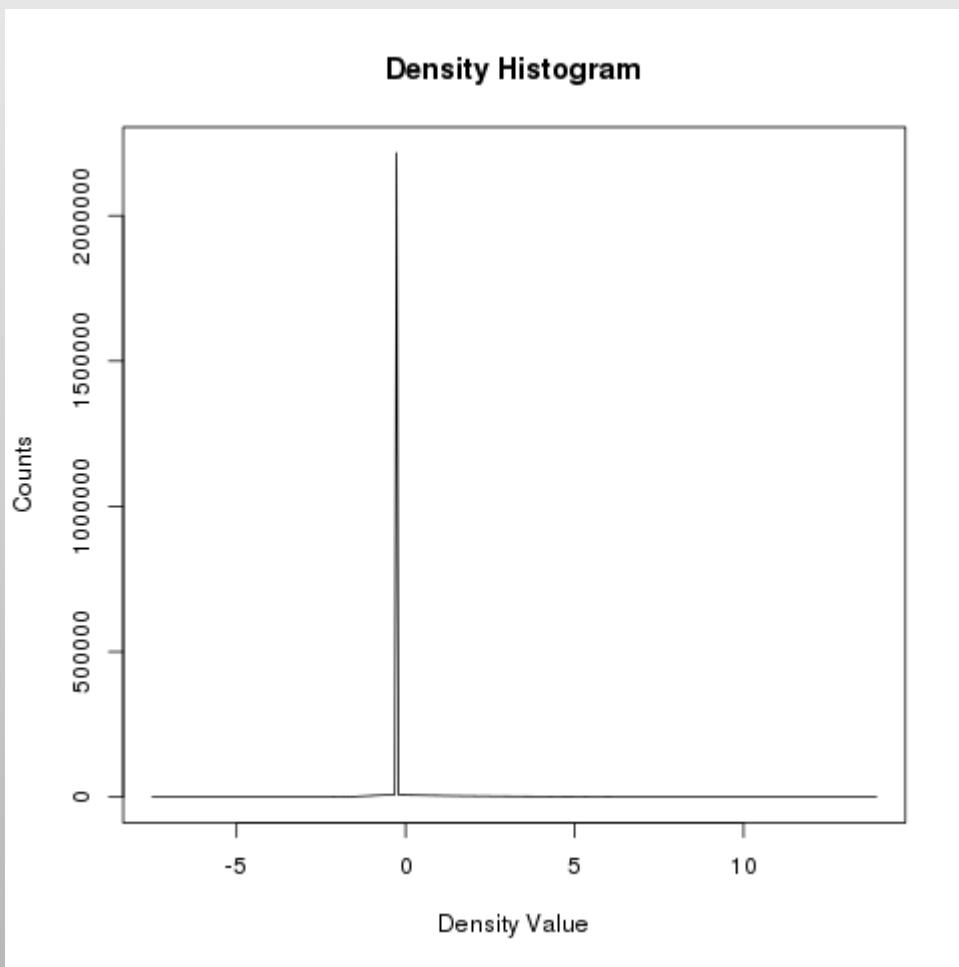
- The distribution of electron density is quite unlike that of x-ray maps

Typical Density Histogram
from an
X-ray map



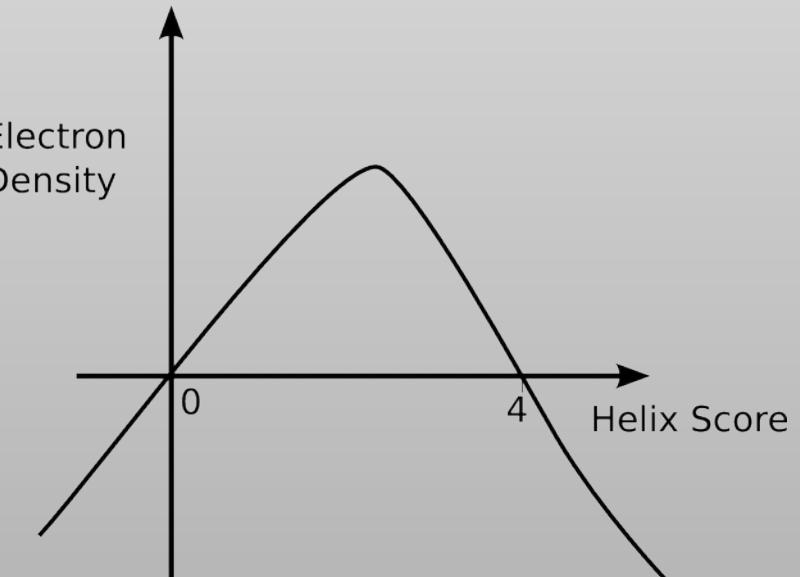
Helix Fitting

- The distribution of electron density is quite unlike that of x-ray maps



Helix Fitting

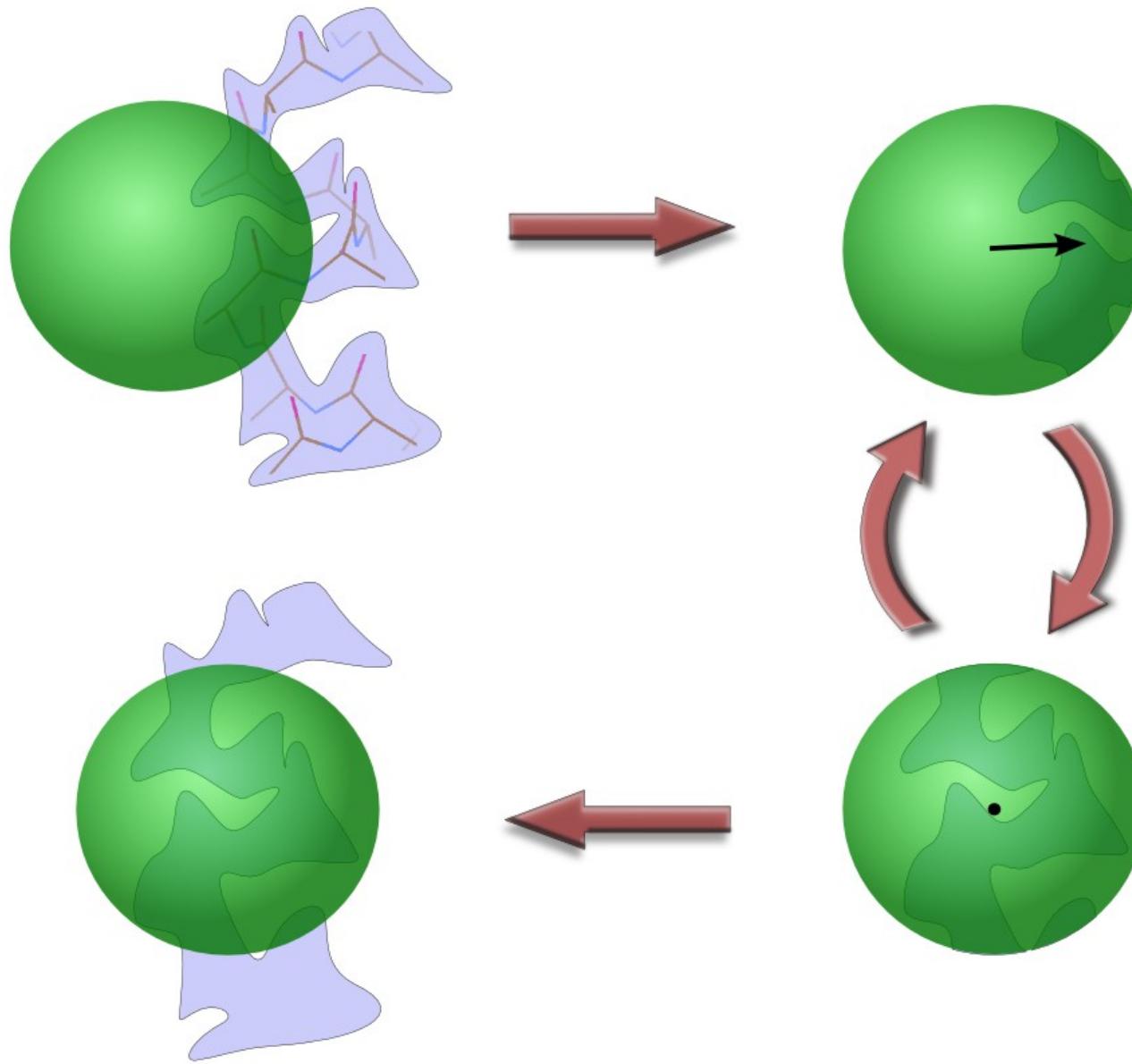
- The distribution of electron density is quite unlike that of x-ray maps
 - e.g. You don't see main-chain atoms at 4 rmsd in x-ray maps
 - regions of dense electron density contribute negatively to helix score
 - These EM maps were sharpened and in a big box of mostly nothing
 - Lots to see at 4 rmsd



Alpha Helix Placement

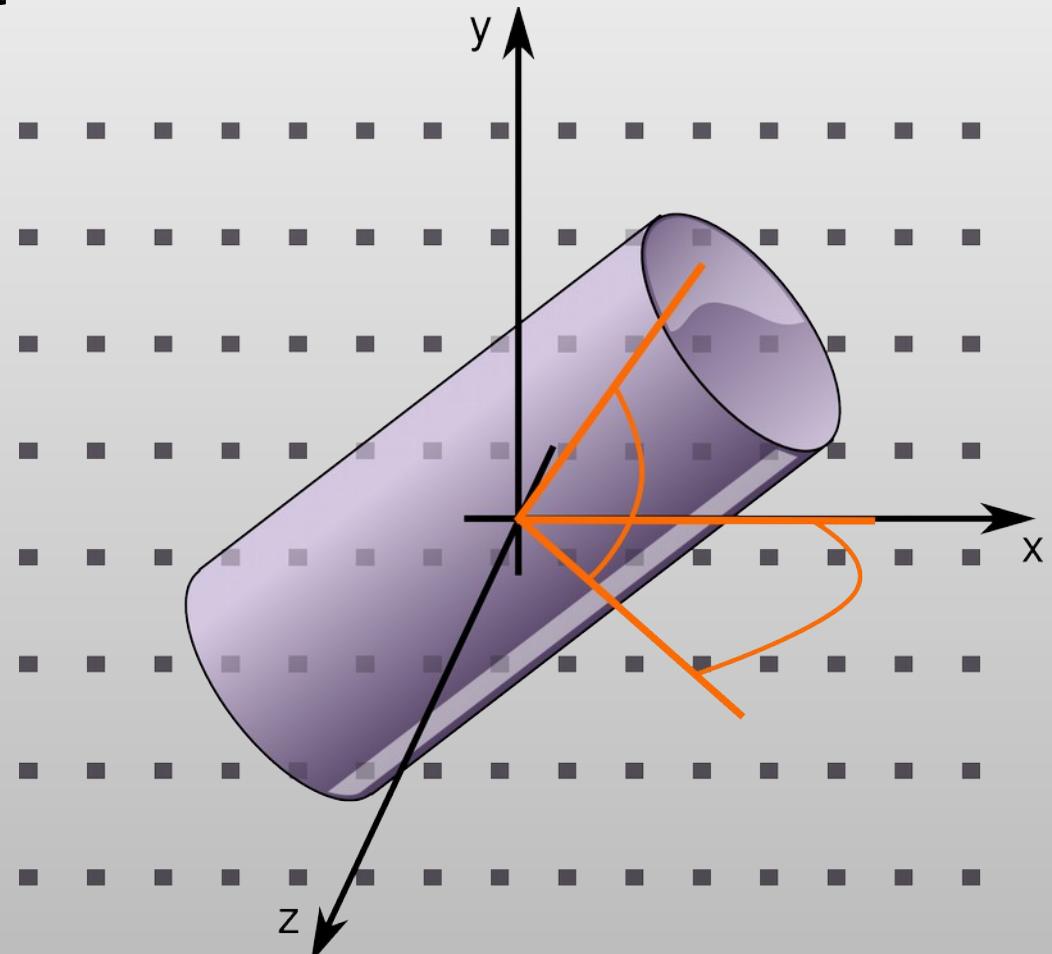
- Scenario: Looking at a new map, not built with automatic tools:
 - “I can see that there’s a helix here - build it for me!”
- From a given point:
 - Move to local averaged maximum
 - Do a 2D MR-style orientation search on a cylinder of electron density
 - Build a helix (both directions)
 - 1D Rotation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

Centering the Rotation point



Real Space Molecular Replacement Cylinder Search

- Pick the orientation that encapsulates the most electron density



Using 2 rotation axes



2 x 1-D Helix orientation searches

Top



Bottom

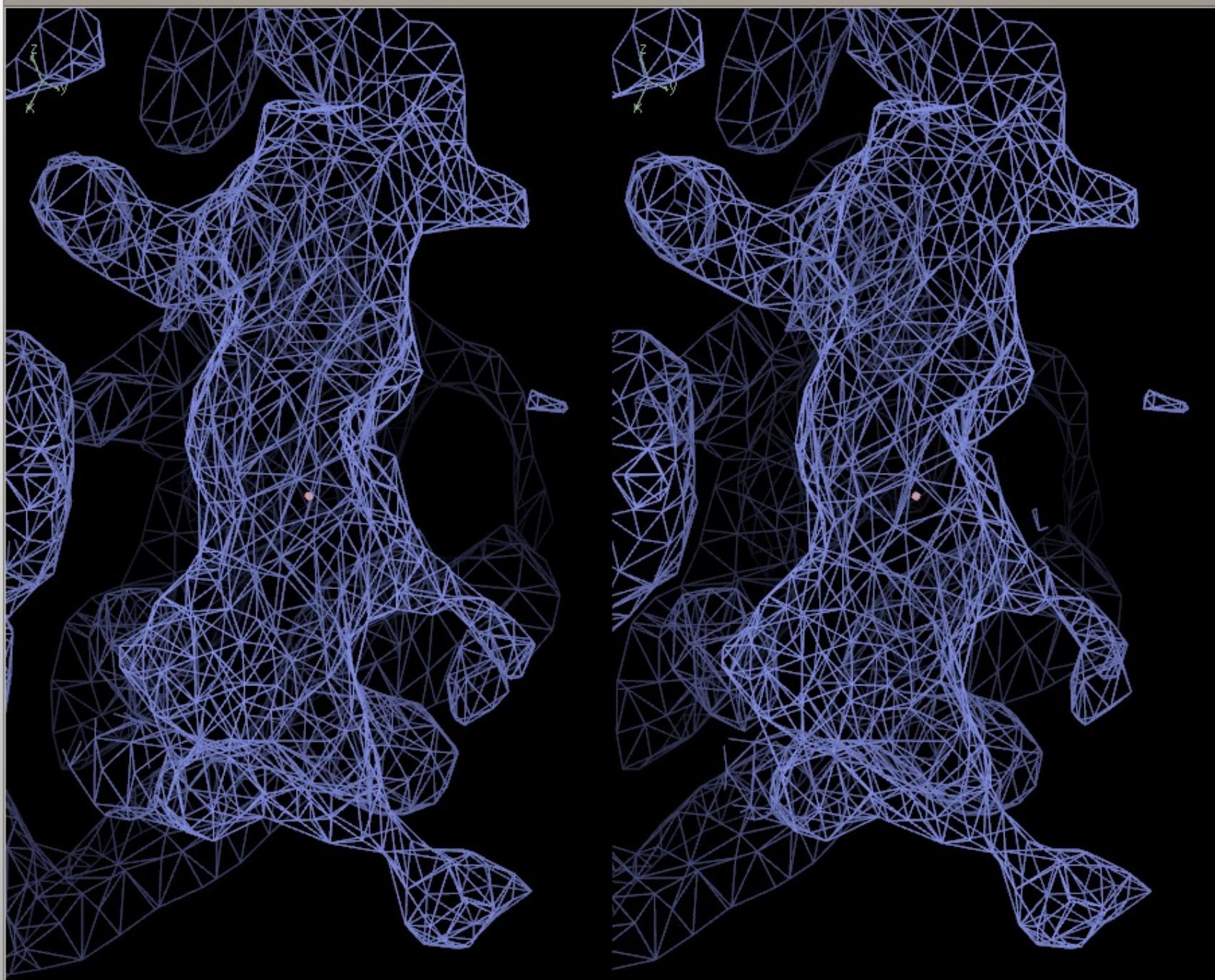
Coot

File Edit Calculate Draw Measures Validate HID About Extensions Lidia

Reset View Display Manager

R/RC

Map



...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

...

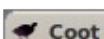
...

...

...

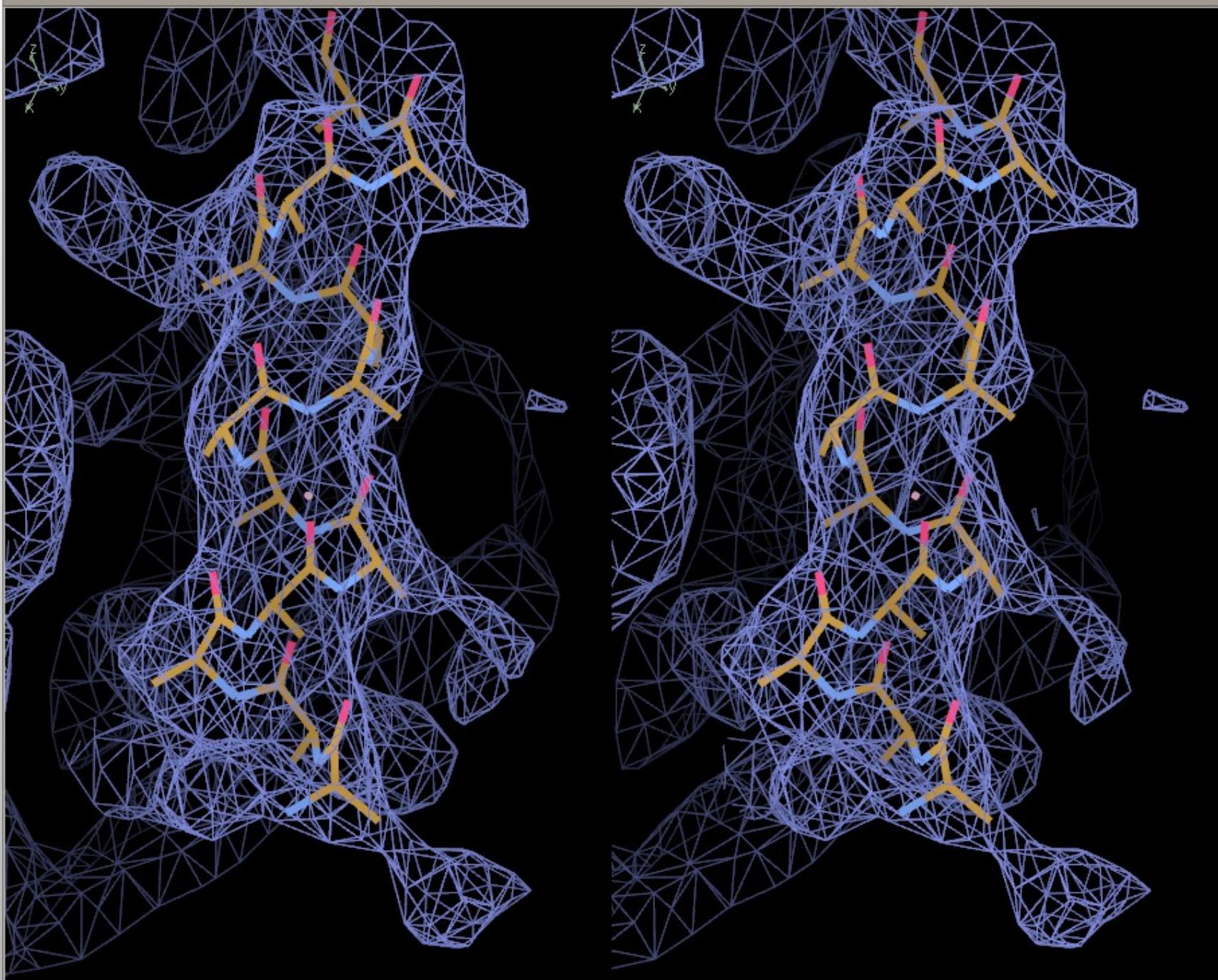
...

...



File Edit Calculate Draw Measures Validate HID About Extensions Lidia

Reset View



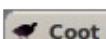
Helix added



R/RC

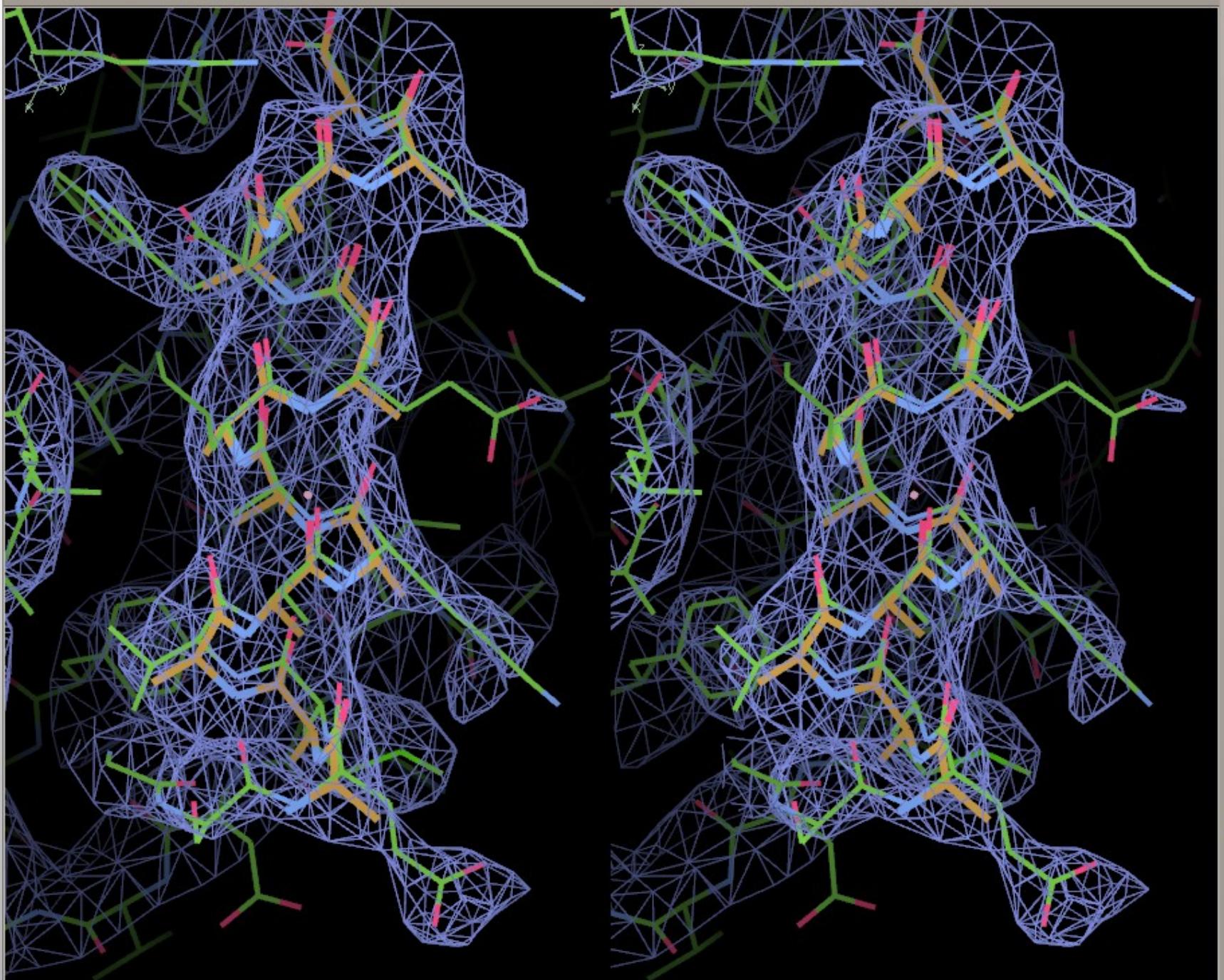
Map





File Edit Calculate Draw Measures Validate HID About Extensions Lidia

Reset View Display Manager

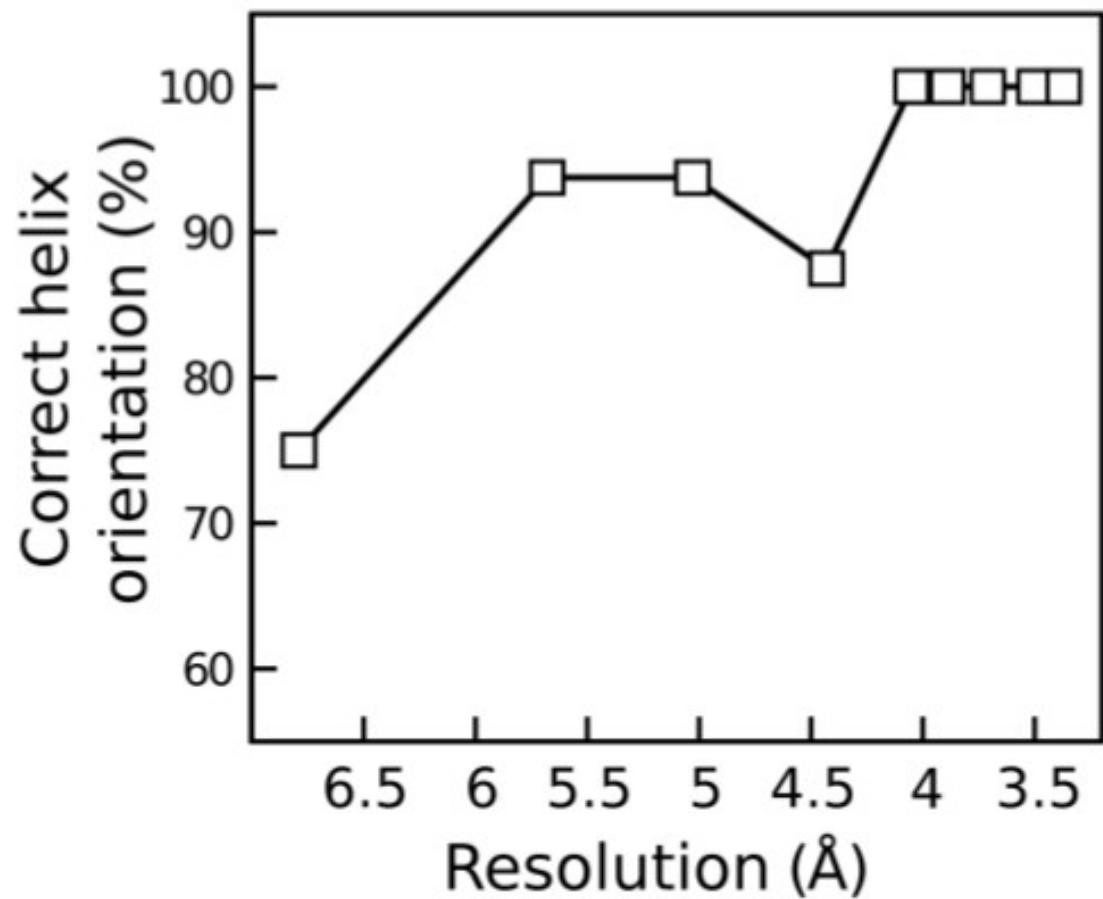
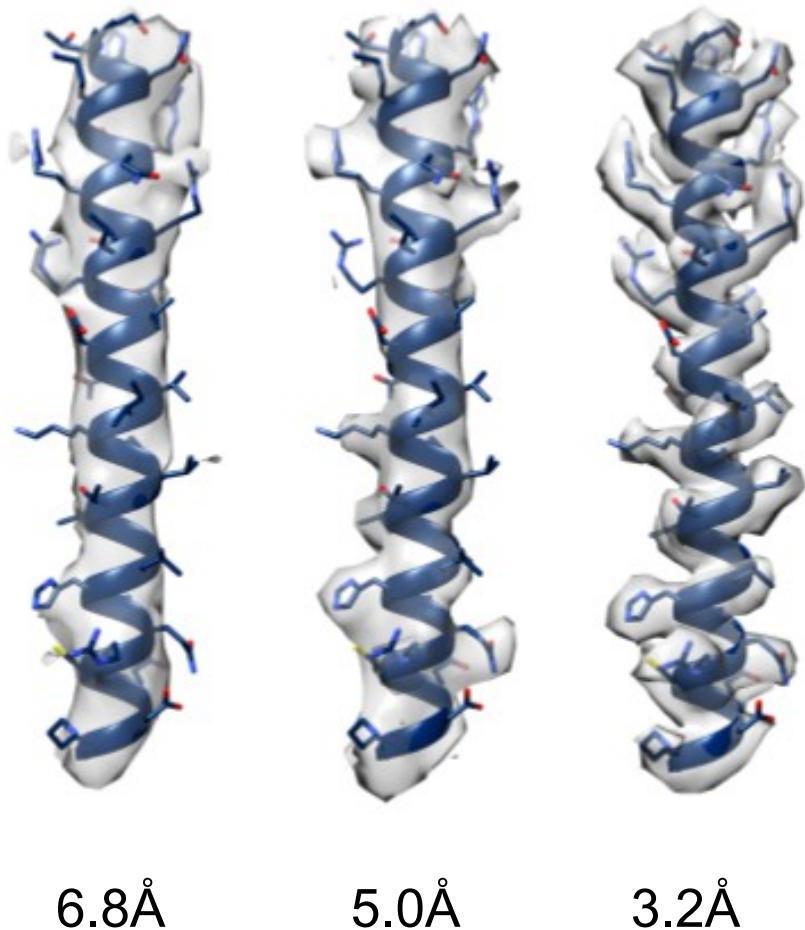


R/RC

Map



Helix Fit Orientation Tests



Post-Poly-ALA Helix: Cootaneer

Use a likelihood function based on conserved density features.

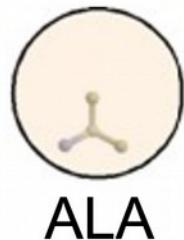
The same likelihood function is used several times. This makes the program very simple (<3000 lines), and the whole calculation works over a range of resolutions.

Finding, growing: Look for C-alpha environment

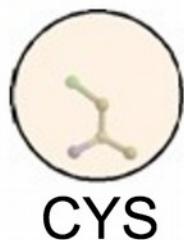


(4.0Å sphere about C α)

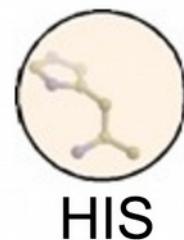
Sequencing: Look for C-beta environment



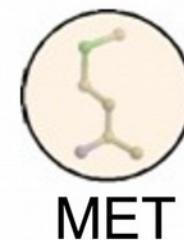
ALA



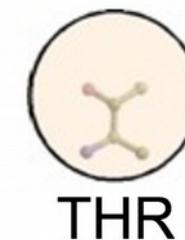
CYS



HIS



MET



THR

(5.5Å sphere about C β)

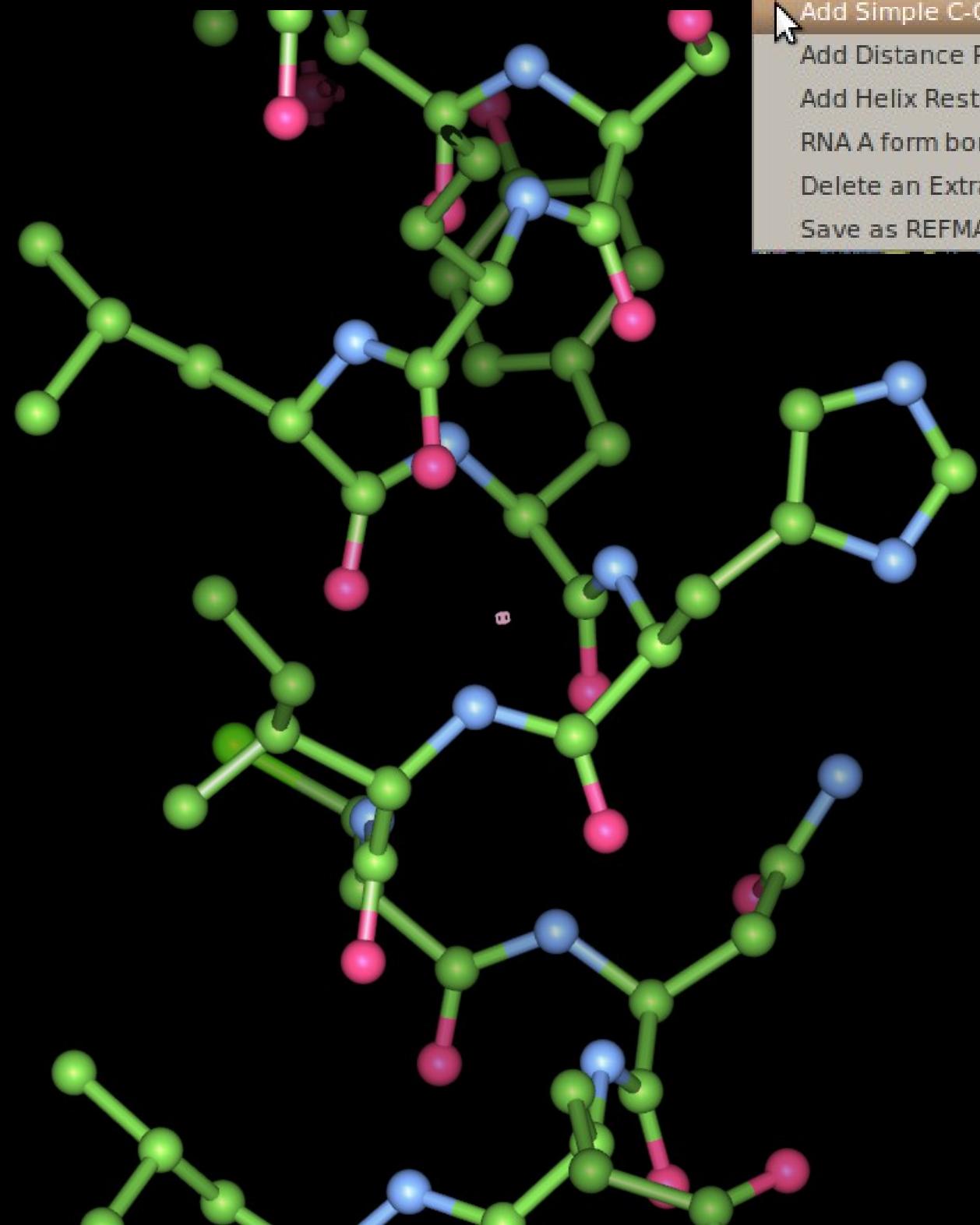
... x20

Additional Restraints

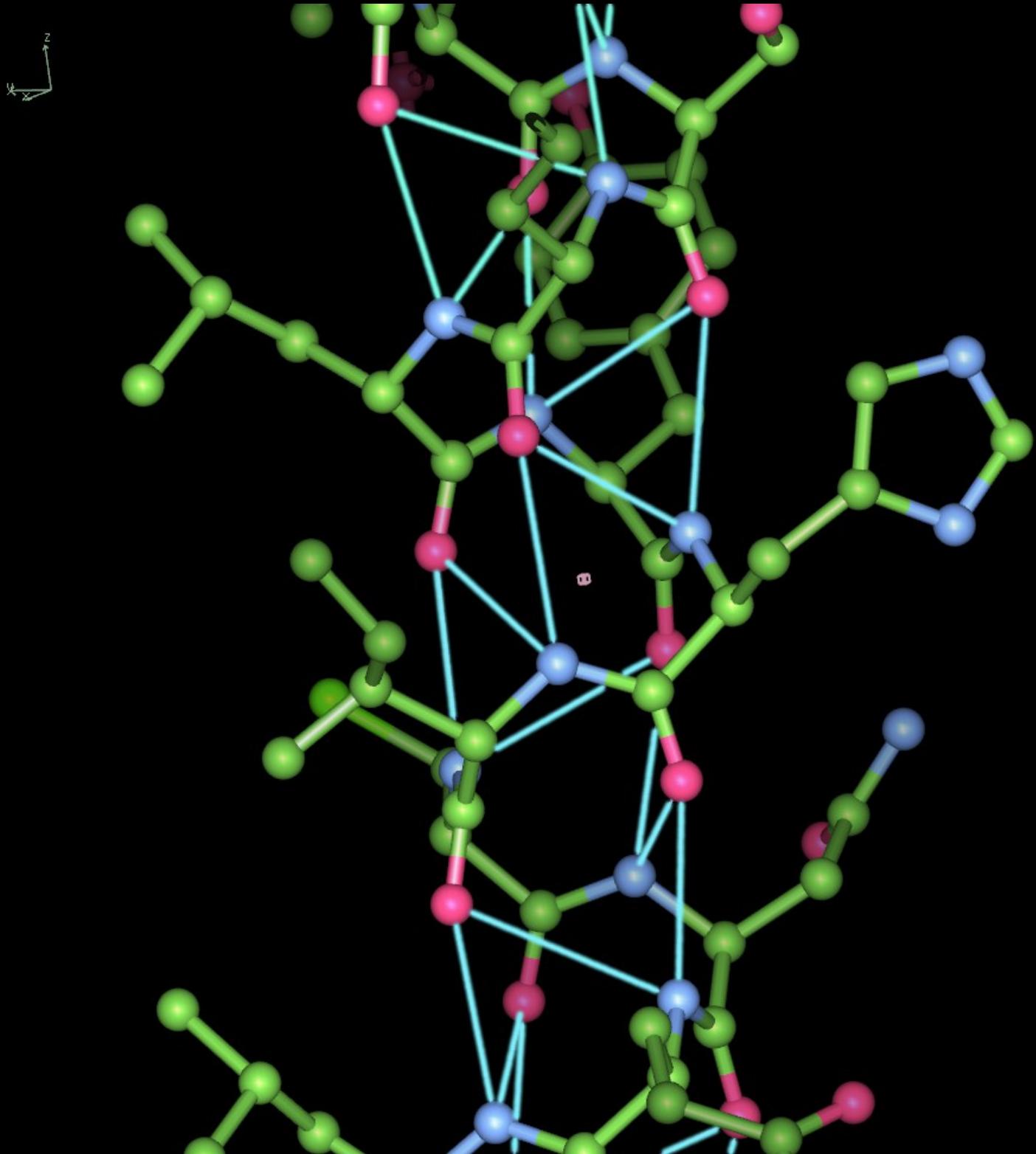
Restraints Editing in Coot

- Distance Restraints:
 - Alpha helices, A-form RNA, B-form DNA
- Add and delete individual restraints
 - User-selectable sigma
- Select 2 residues for range
- User-defined torsion restraints
- Input from ProSMART
- Output to Refmac

x
y
z

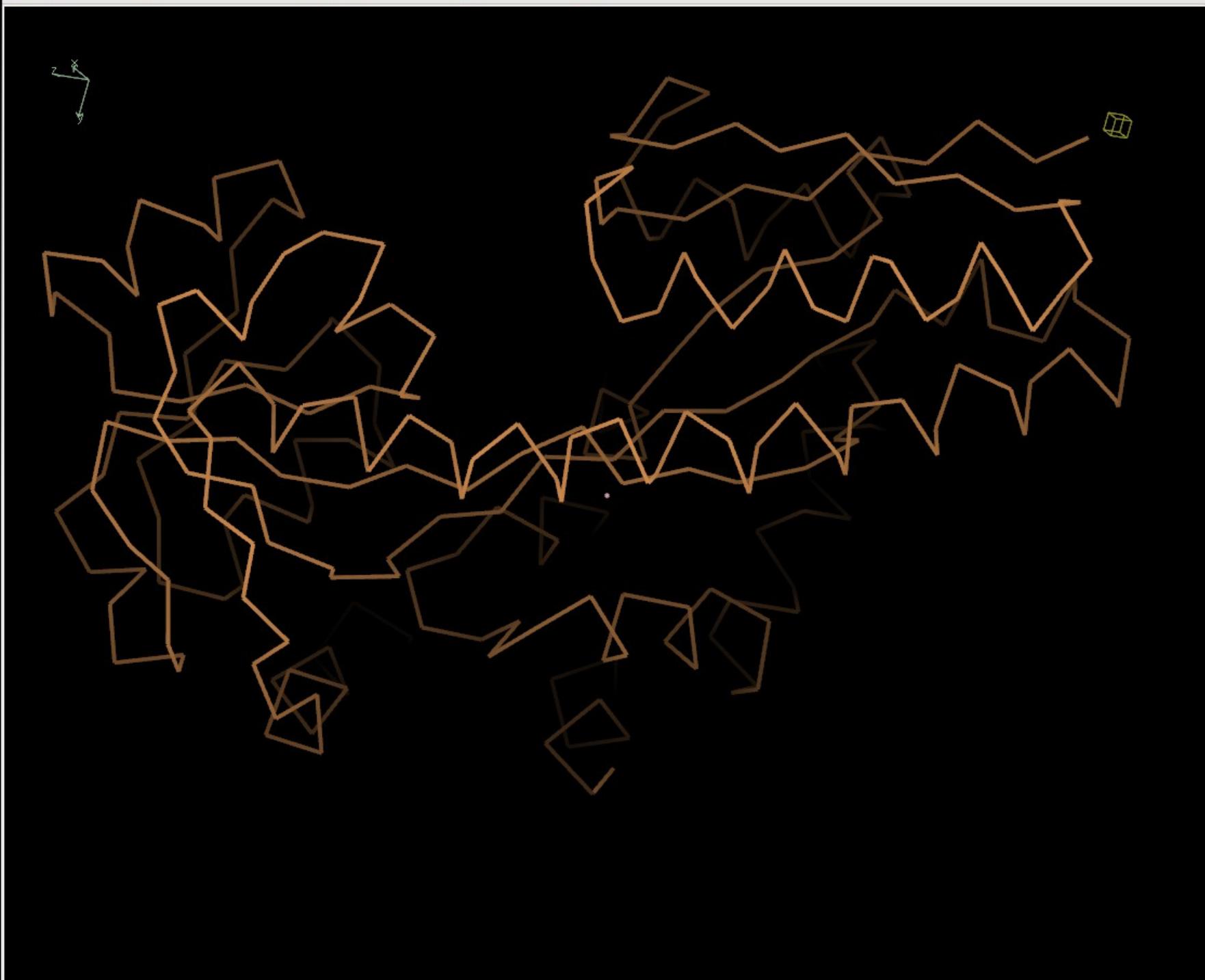


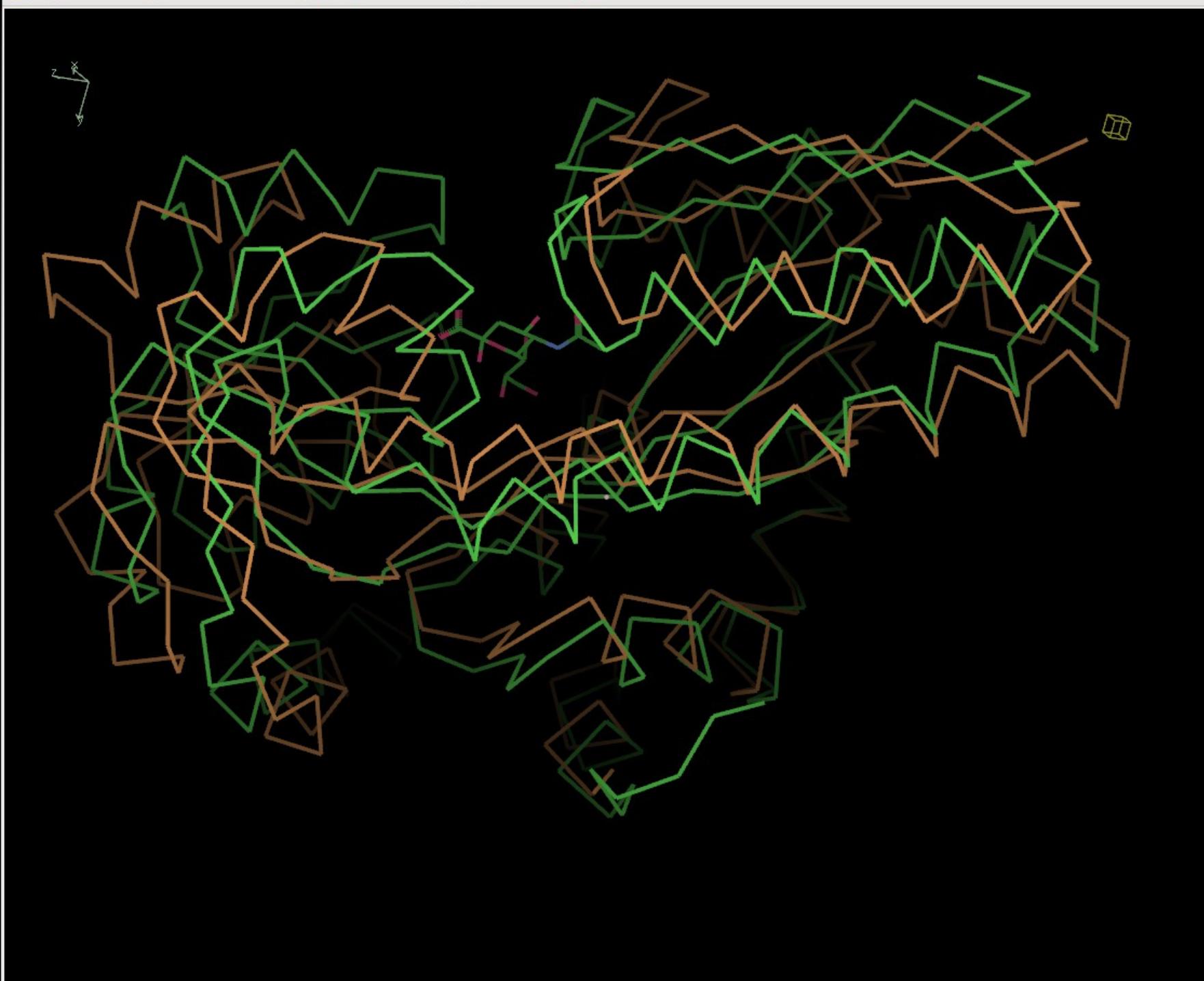
- [Add Simple C-C Single Bond Restraint...](#)
- [Add Distance Restraint...](#)
- [Add Helix Restraints...](#)
- [RNA A form bond restraints...](#)
- [Delete an Extra Restraint...](#)
- [Save as REFMAC restraints...](#)

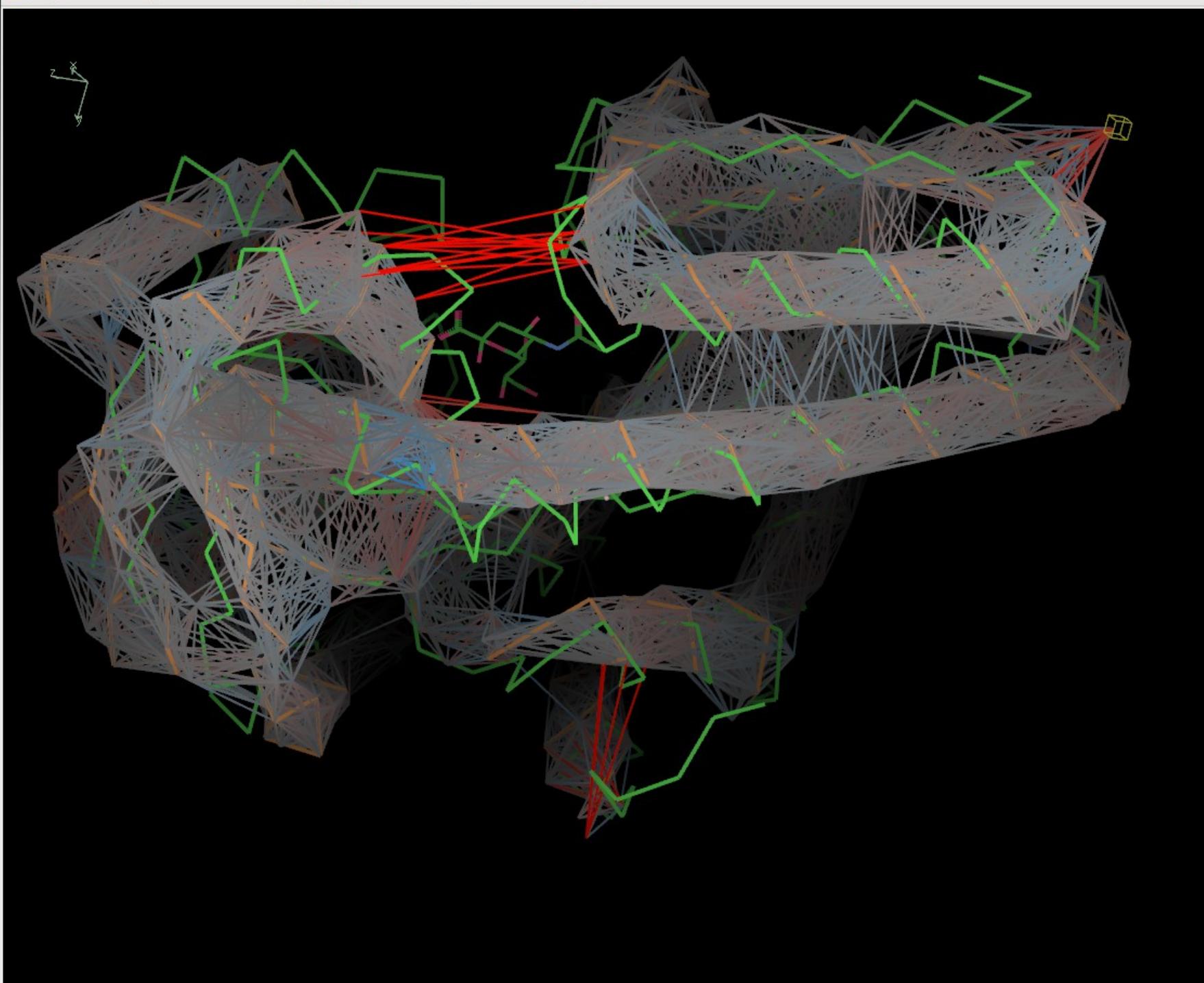


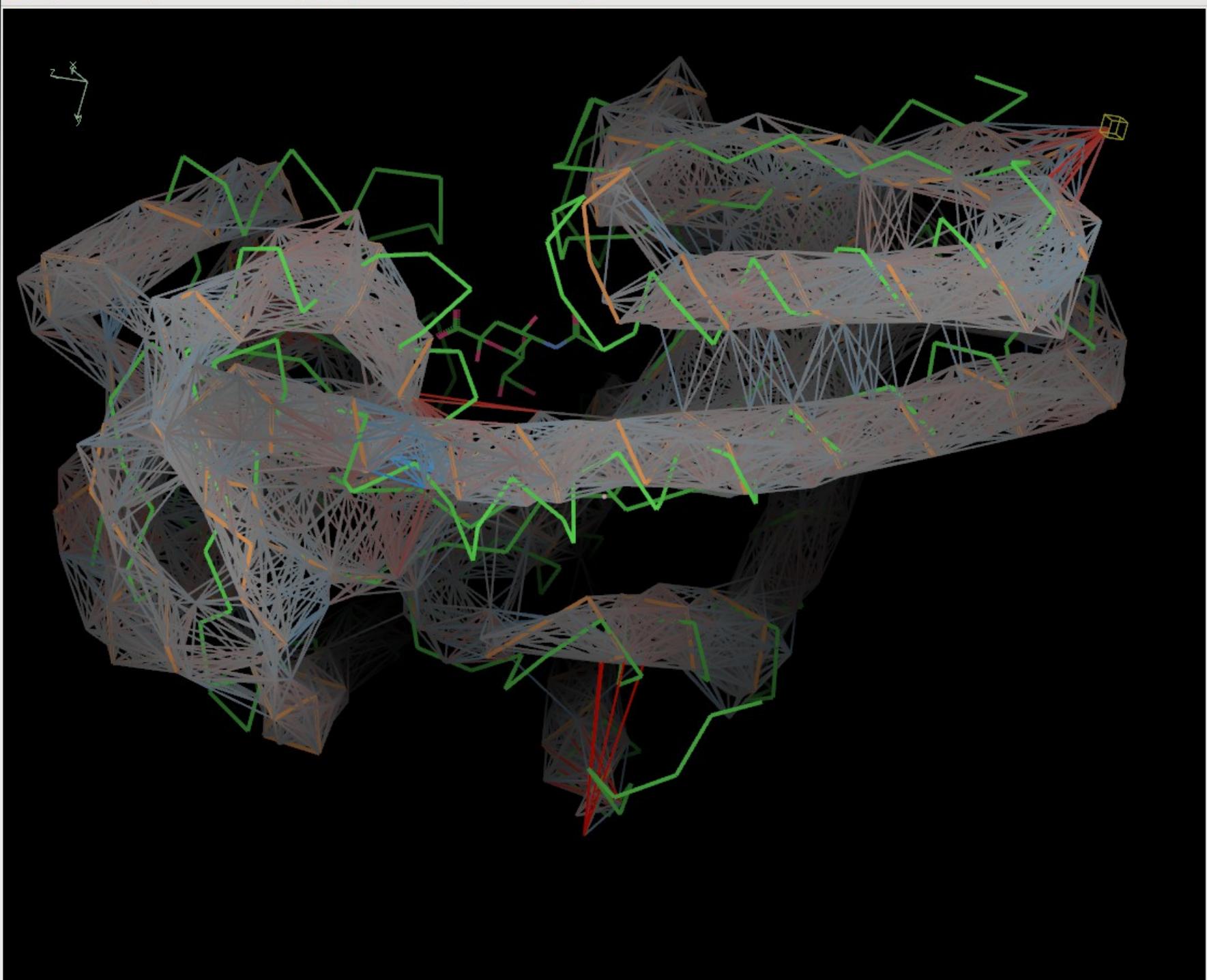
ProSMART Interface

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

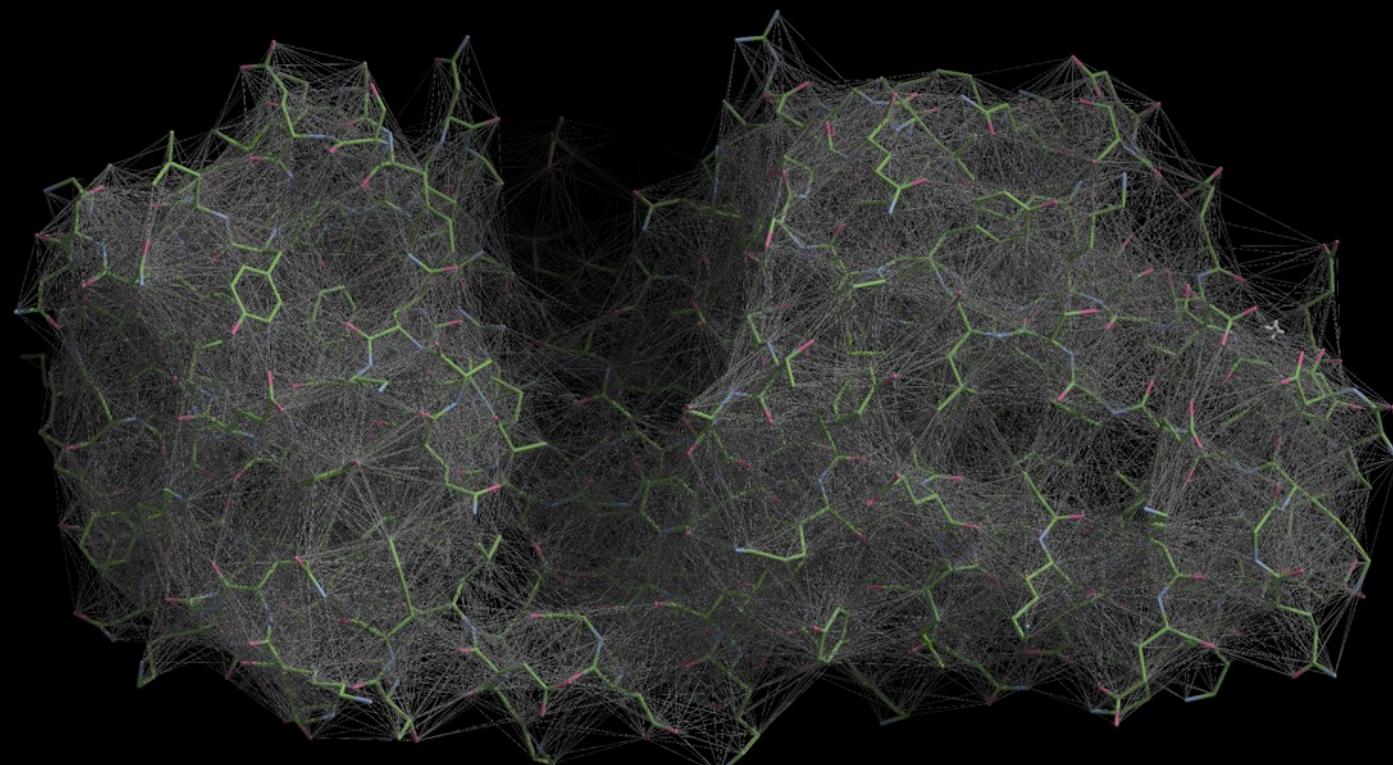








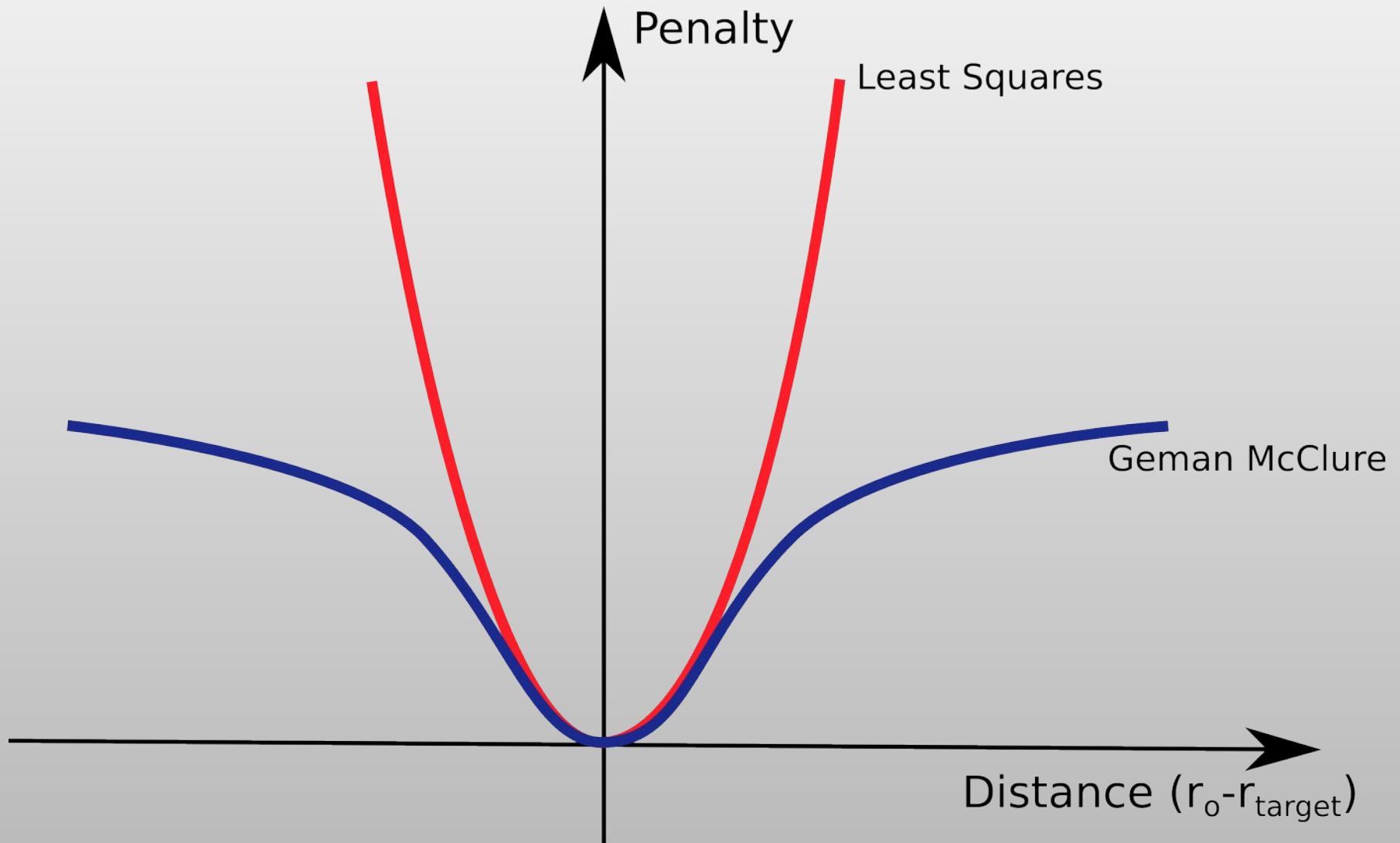
ProSMART Restraints



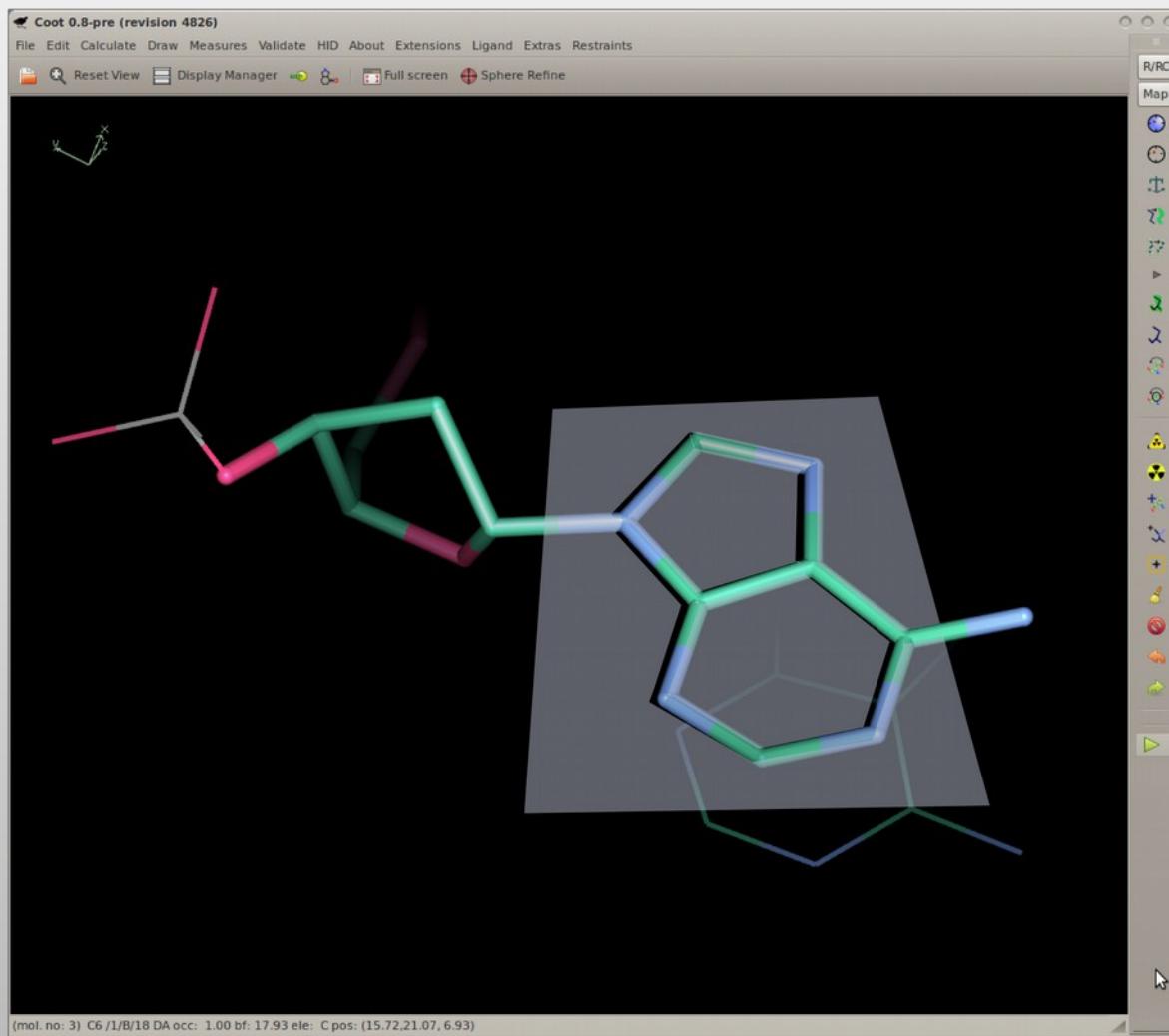
ProSMART integration

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

Modified Target Function

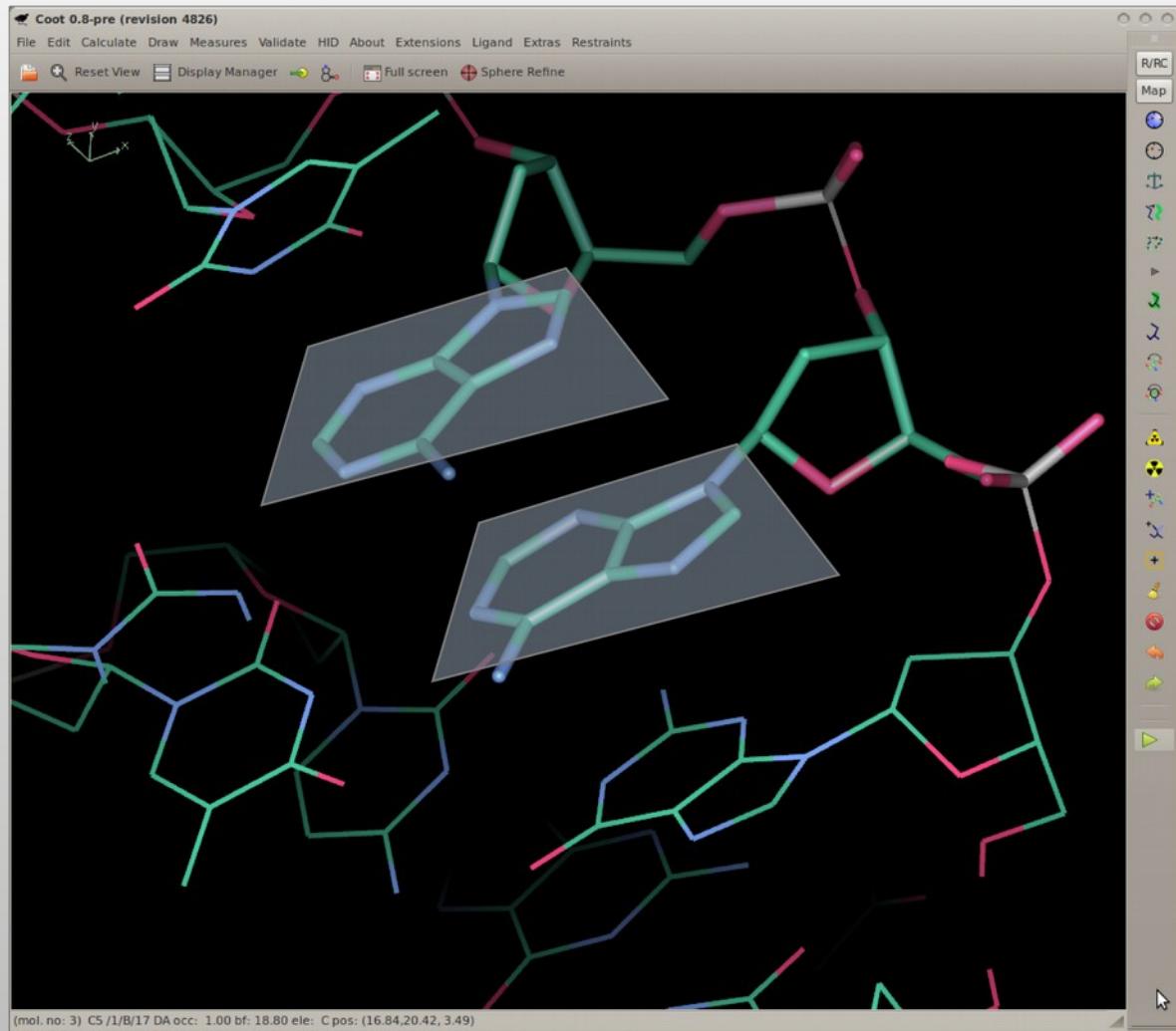


Plane Restraints



Derivativatives are an eigenvector scaled by out-of-plane distance

Parallel Planes Restraints



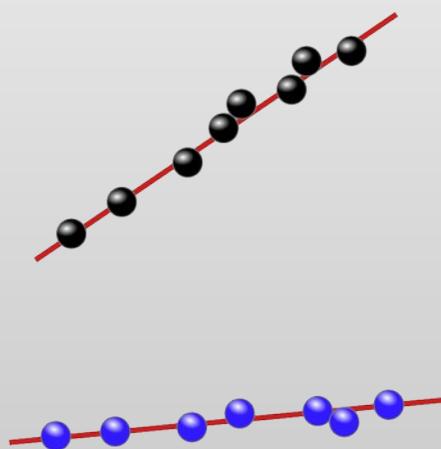
$$S = (a_1 - a_2)^2 + (b_1 - b_2)^2 + (c_1 - c_2)^2$$

Not easy to use in Coot

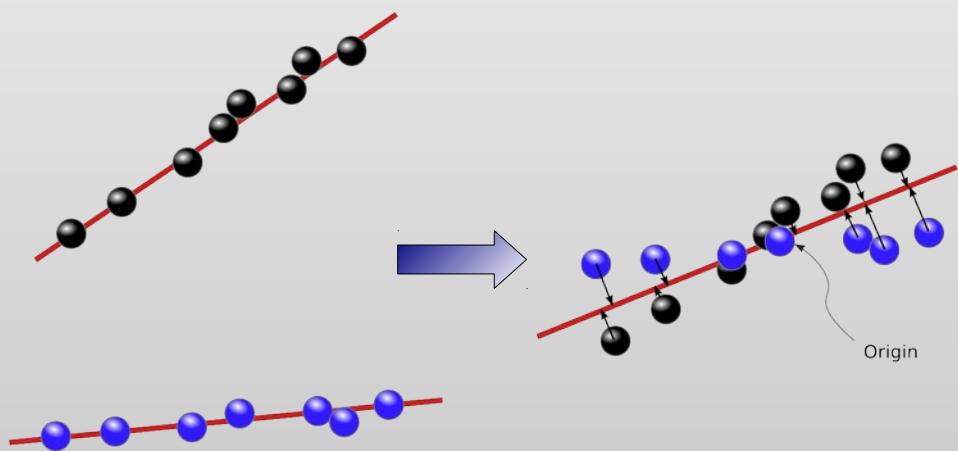
Parallel Planes Restraints

- Also, we have considered parallel-planes distance restraints
 - More tricky still to implement
 - Not implemented yet (not in *Coot*, anyway)

Parallel Planes Restraints

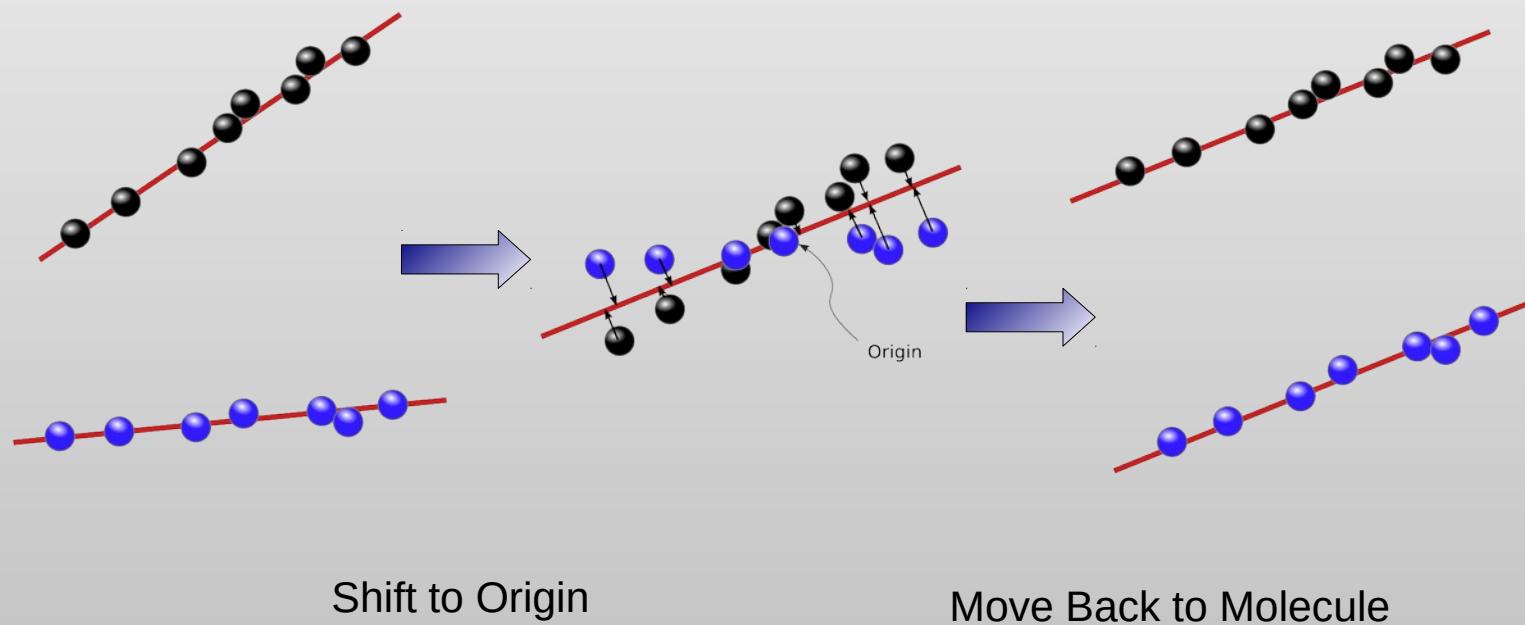


Parallel Plane Restraints



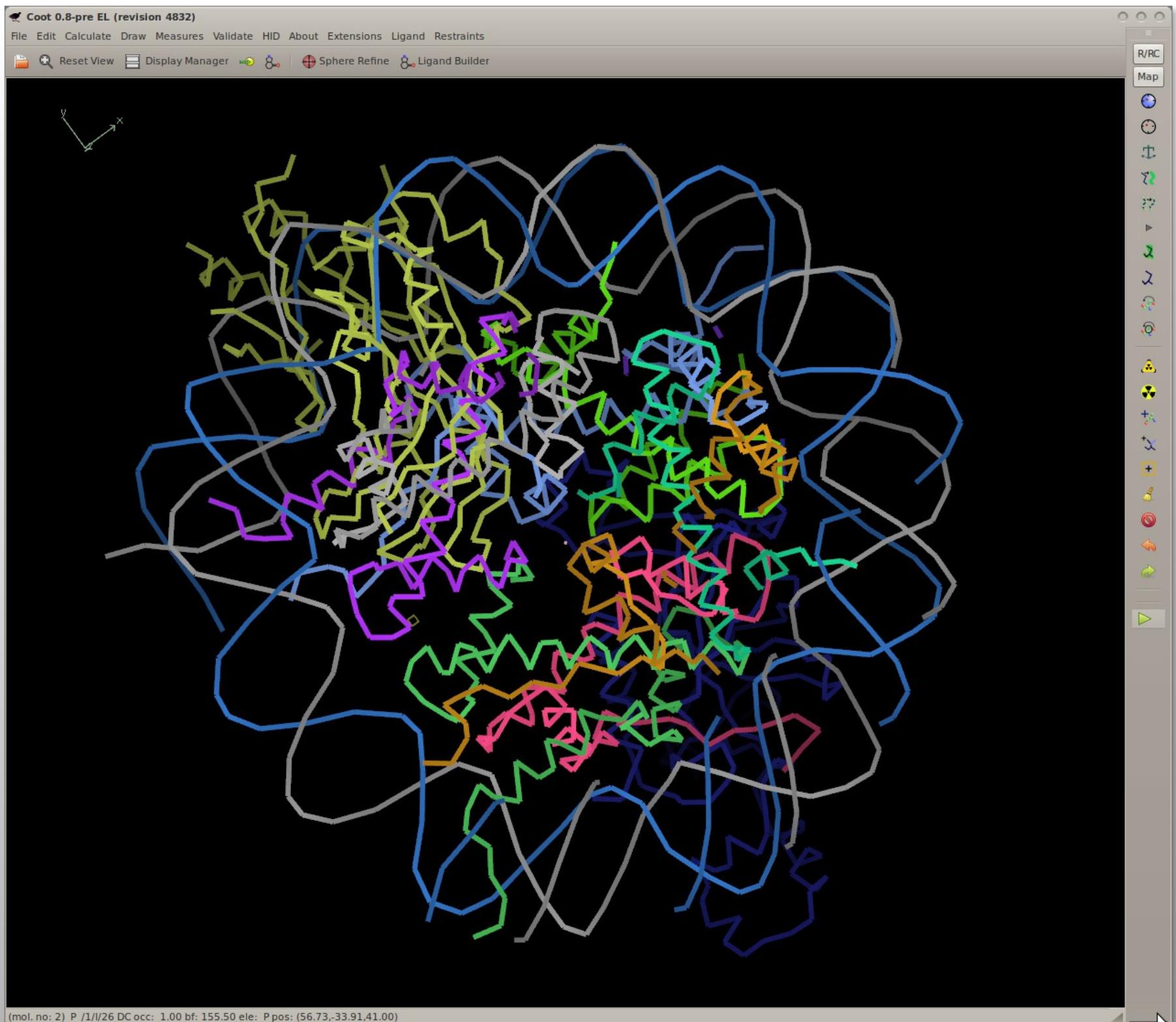
Shift to Origin

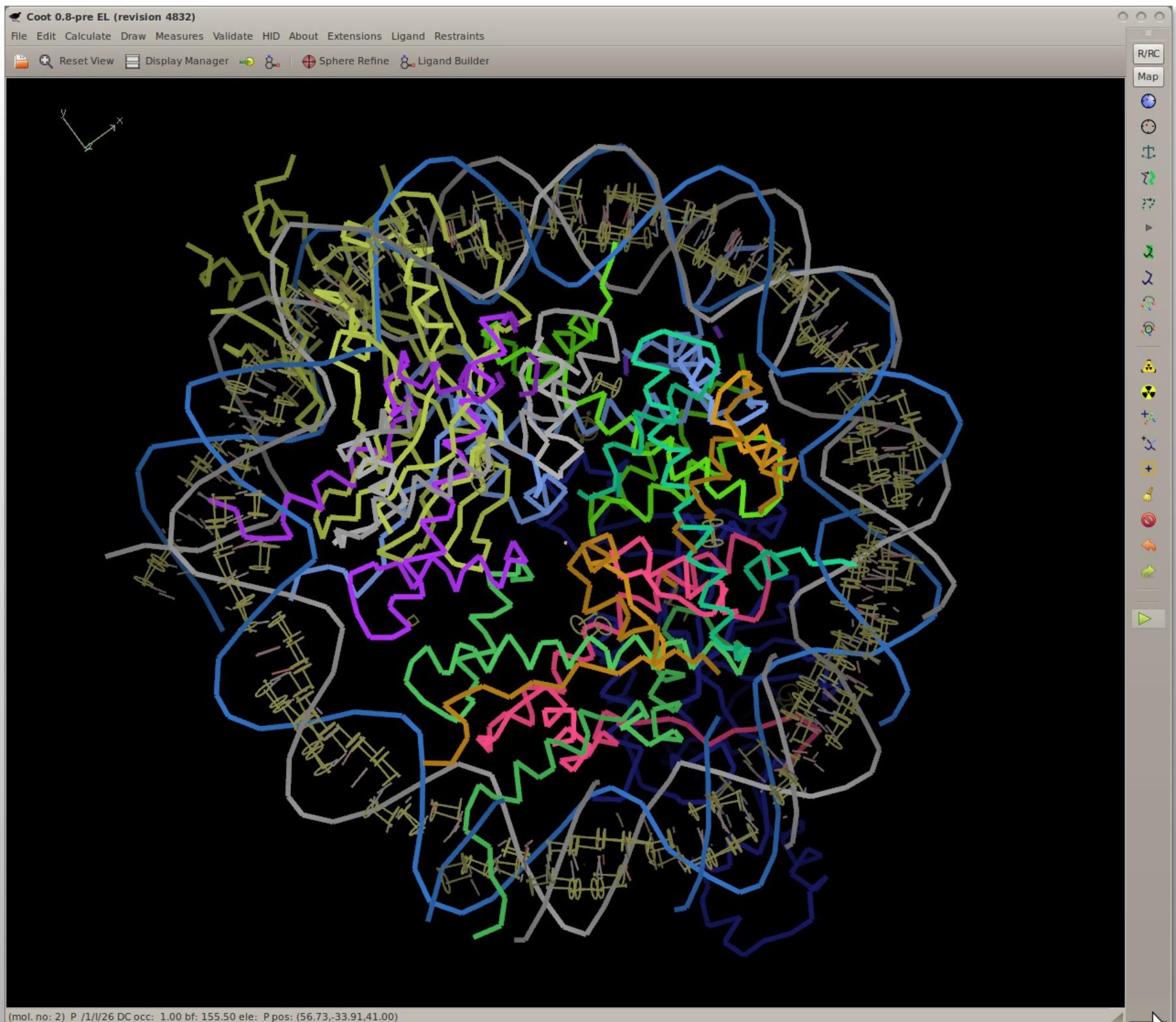
Parallel Planes Restraints



Automatic Generation of Base Pairing and Stacking Restraints

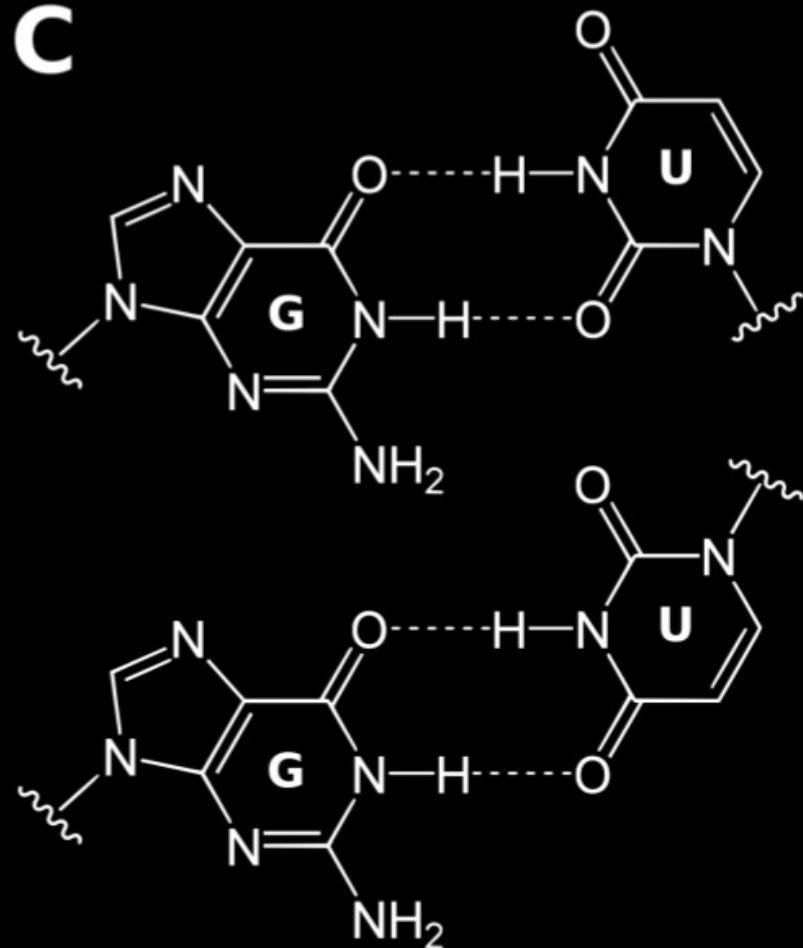
- Fei Long's `libg_d`
 - Provide it with a model and it writes out Refmac restraint descriptions
 - ... which *Coot* can also read
 - *Coot* can also create user-defined base-pairing and stacking restraints



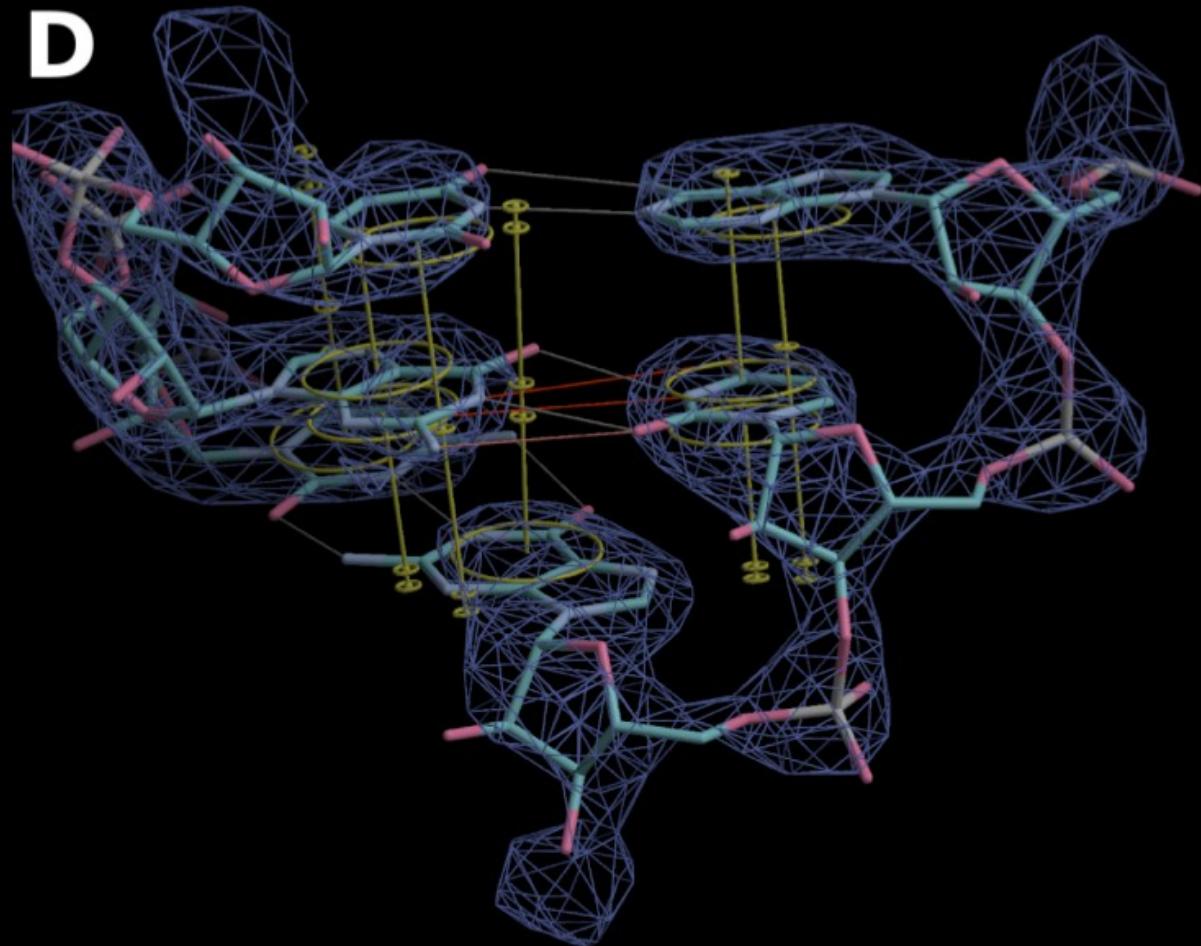


Libg restraints

C



D



(Watson Crick and)
Wobble, Reverse Wobble

Representation in Coot

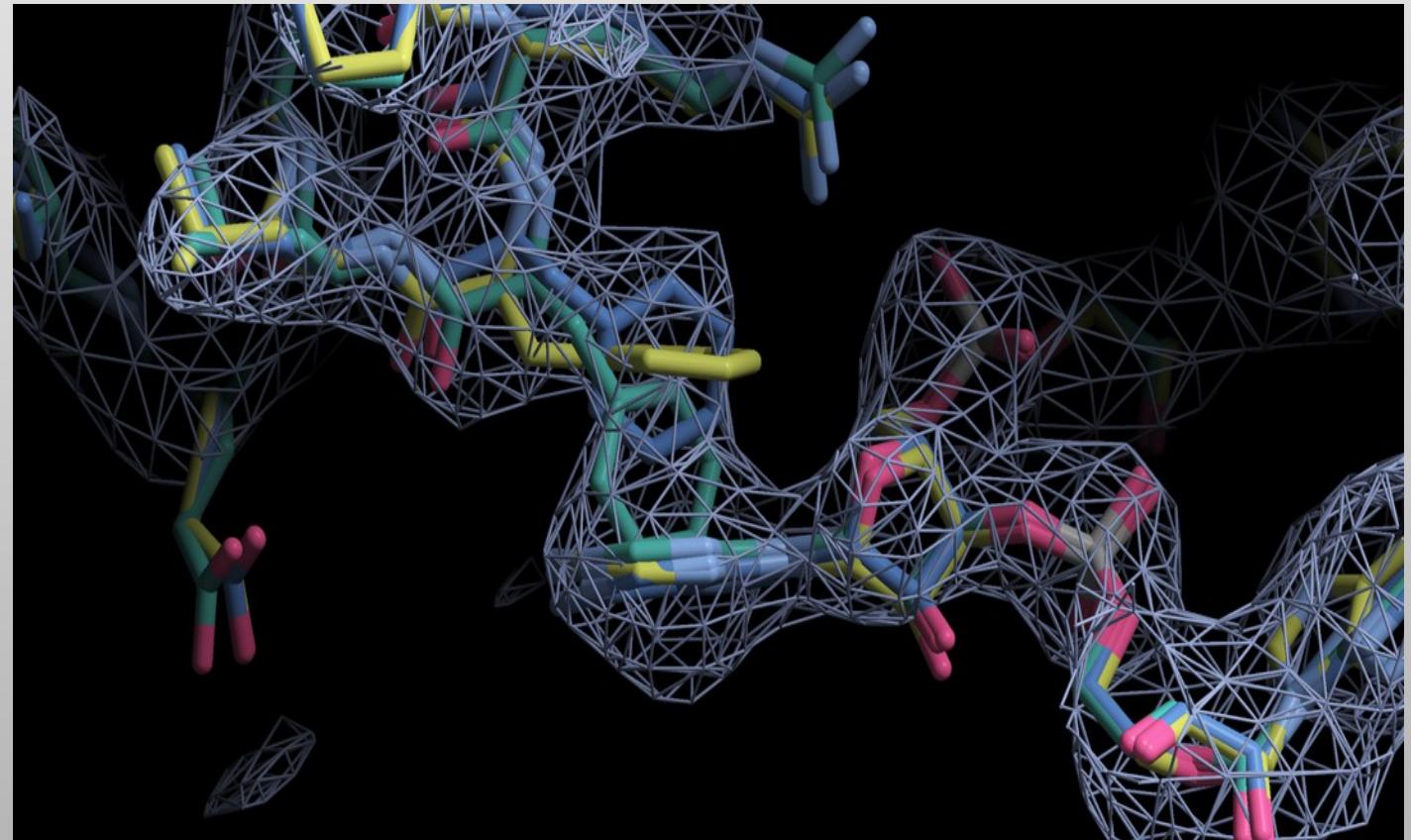
LIBG Restraints for DNA/RNA

LIBG – for the generation of nucleic acid restraints

Base-stacking restraints:
(parallel plane restraints)

Example:
3.2 \AA cryo-EM

Green – before refinement
Blue – refined without LIBG
Yellow – refined with LIBG



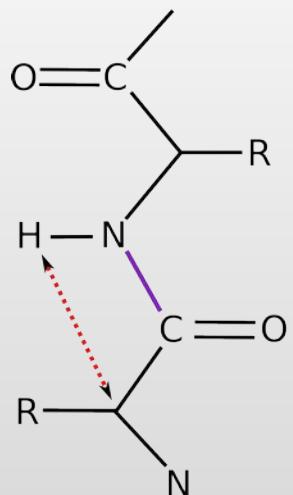
cis-Peptides

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

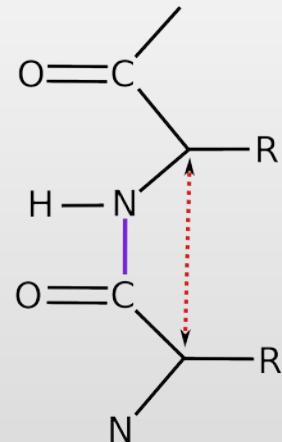
cis-Peptides

- 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). **D71**, 706-709
 - Touw *et al.*: Detection of trans–cis flips and peptide-plane flips in protein structures *Acta Cryst.* (2015). **D71**, 1604-71614

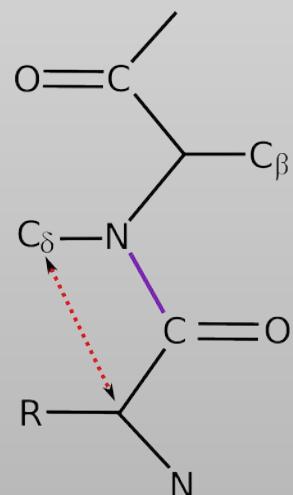
cis-Peptides



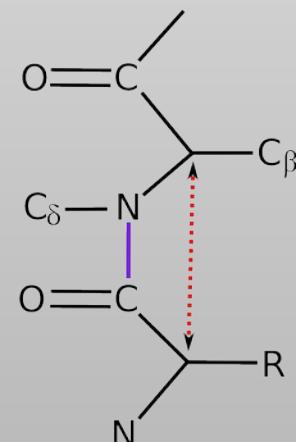
trans-peptide



cis-peptide

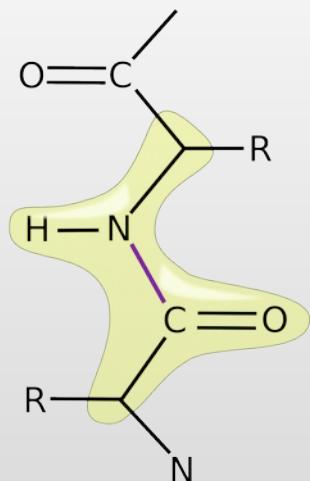


PRO *trans-peptide*

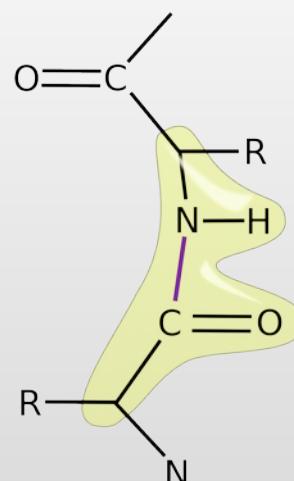


PRO *cis-peptide*

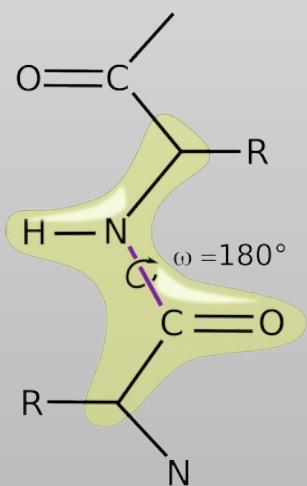
cis-Peptides



trans-peptide
with plane restraints

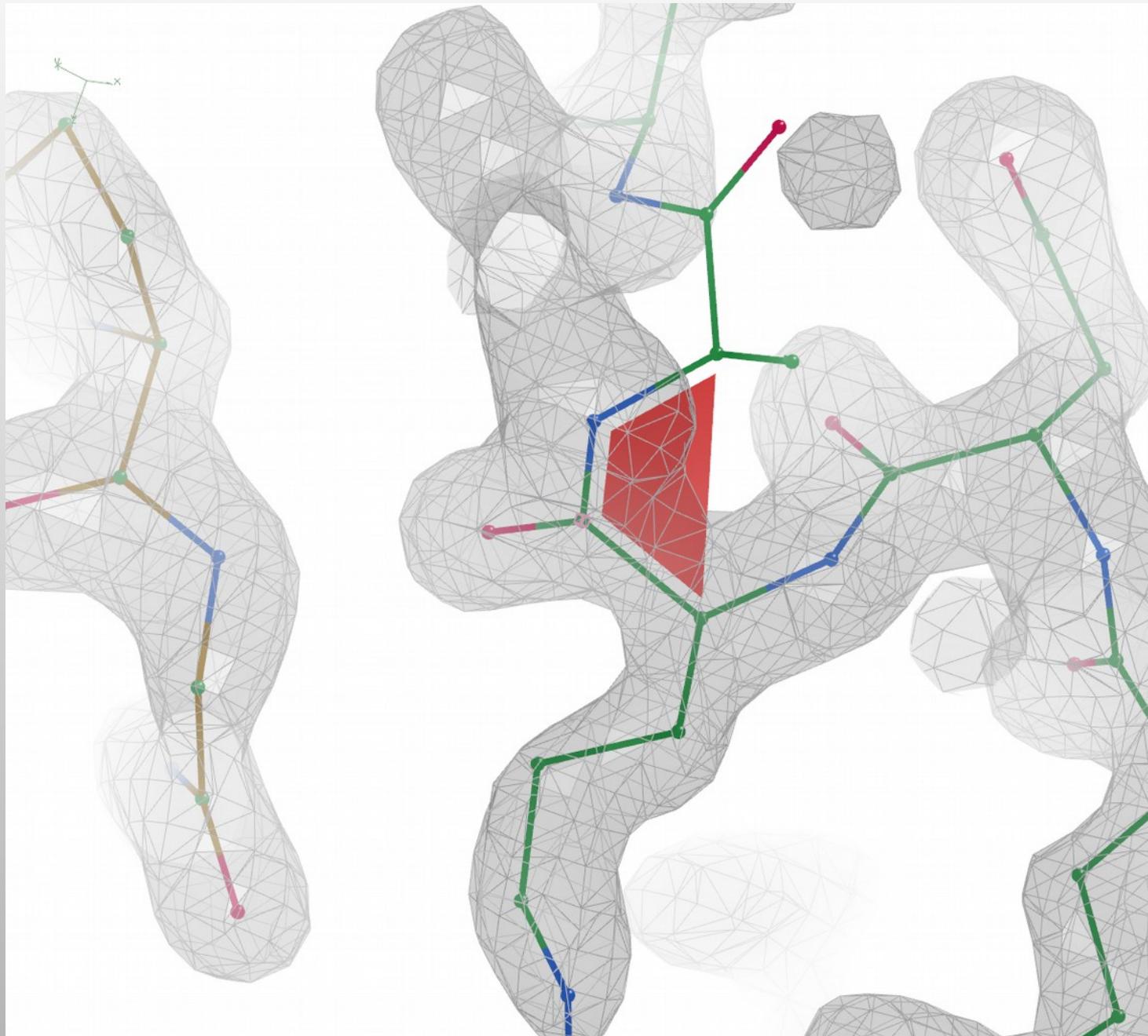


cis-peptide
with plane restraints

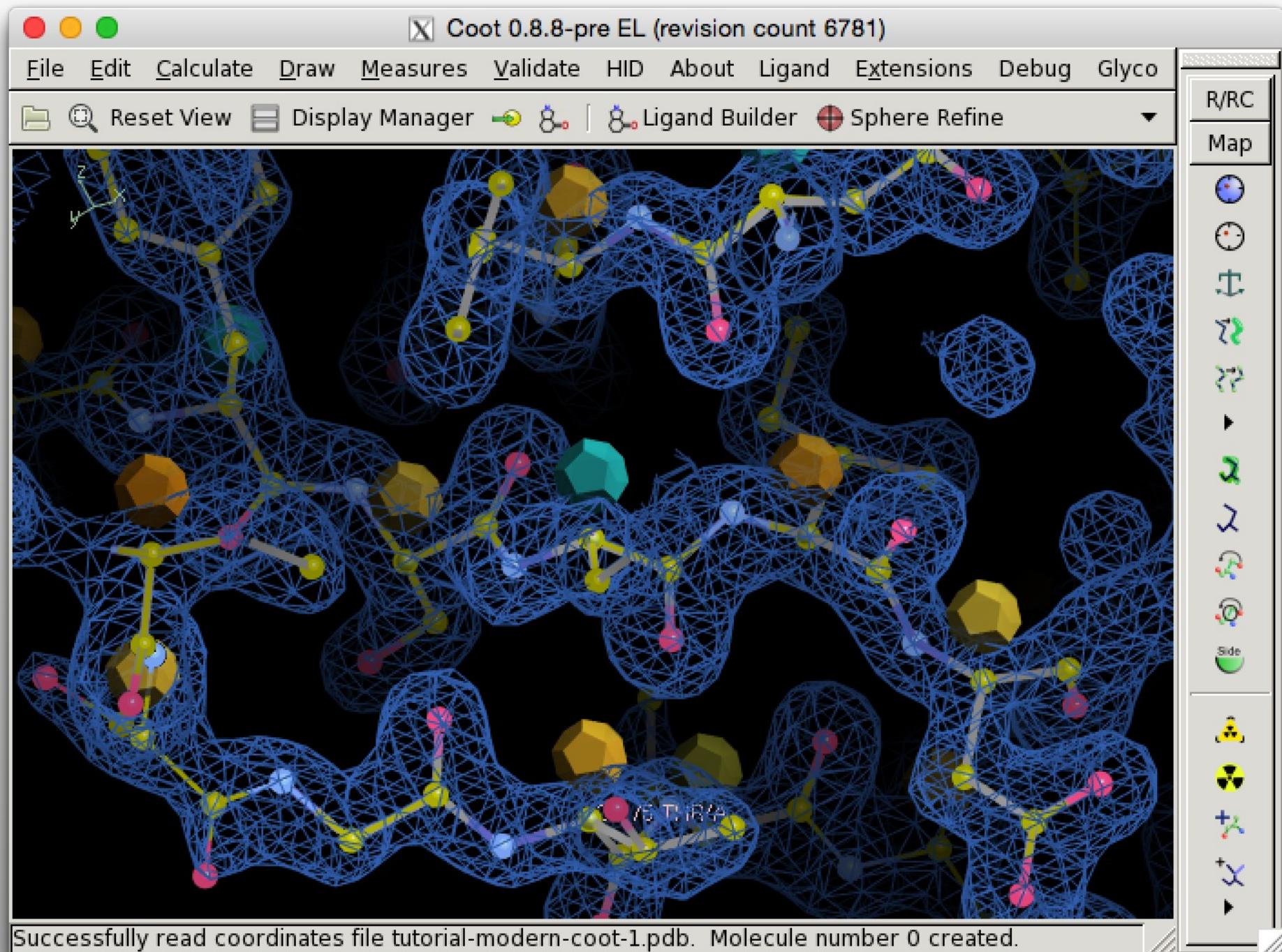


trans-peptide
with plane and trans restraints

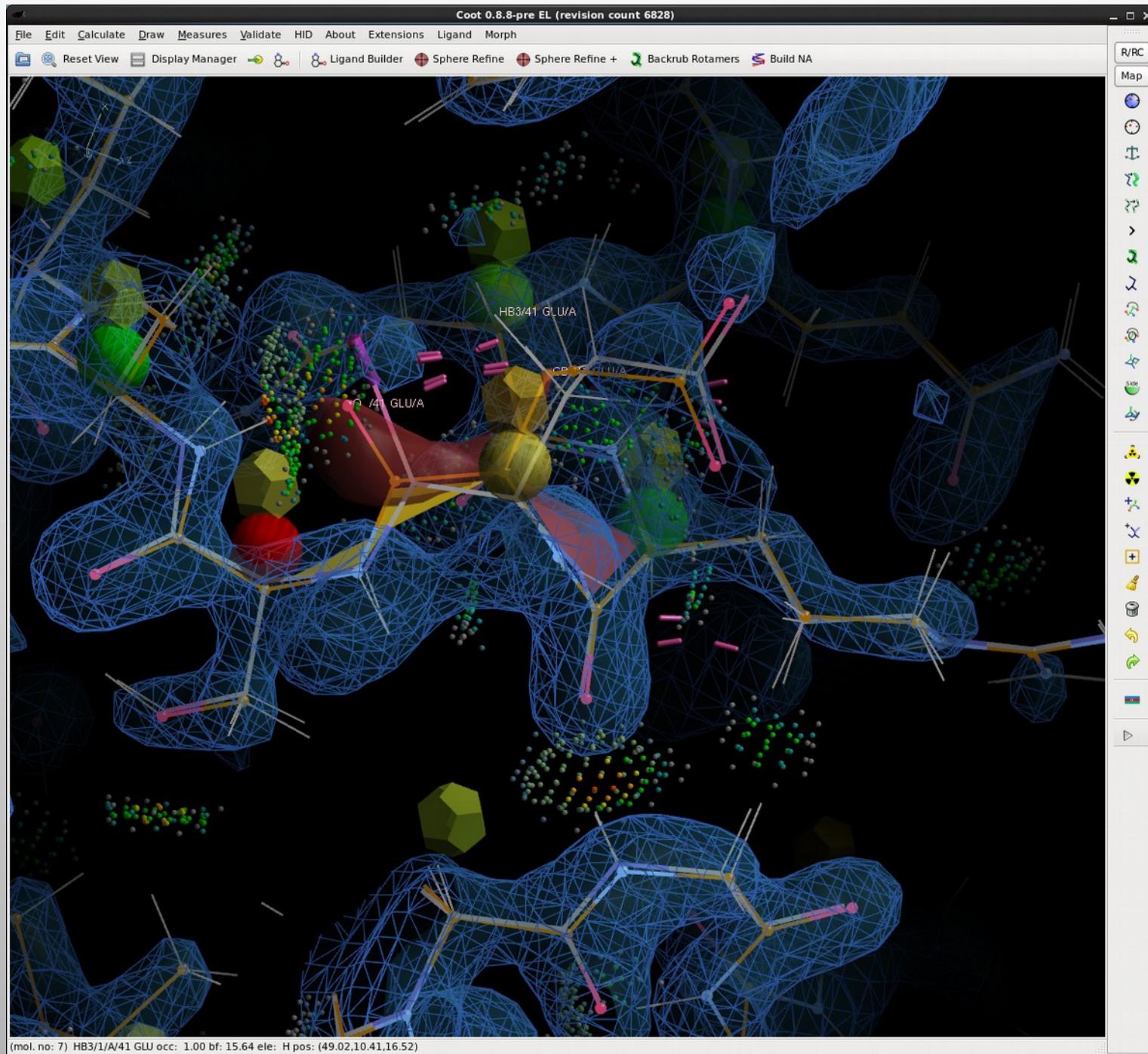
cis-peptide Representation



Interactive Rotamer Goodness



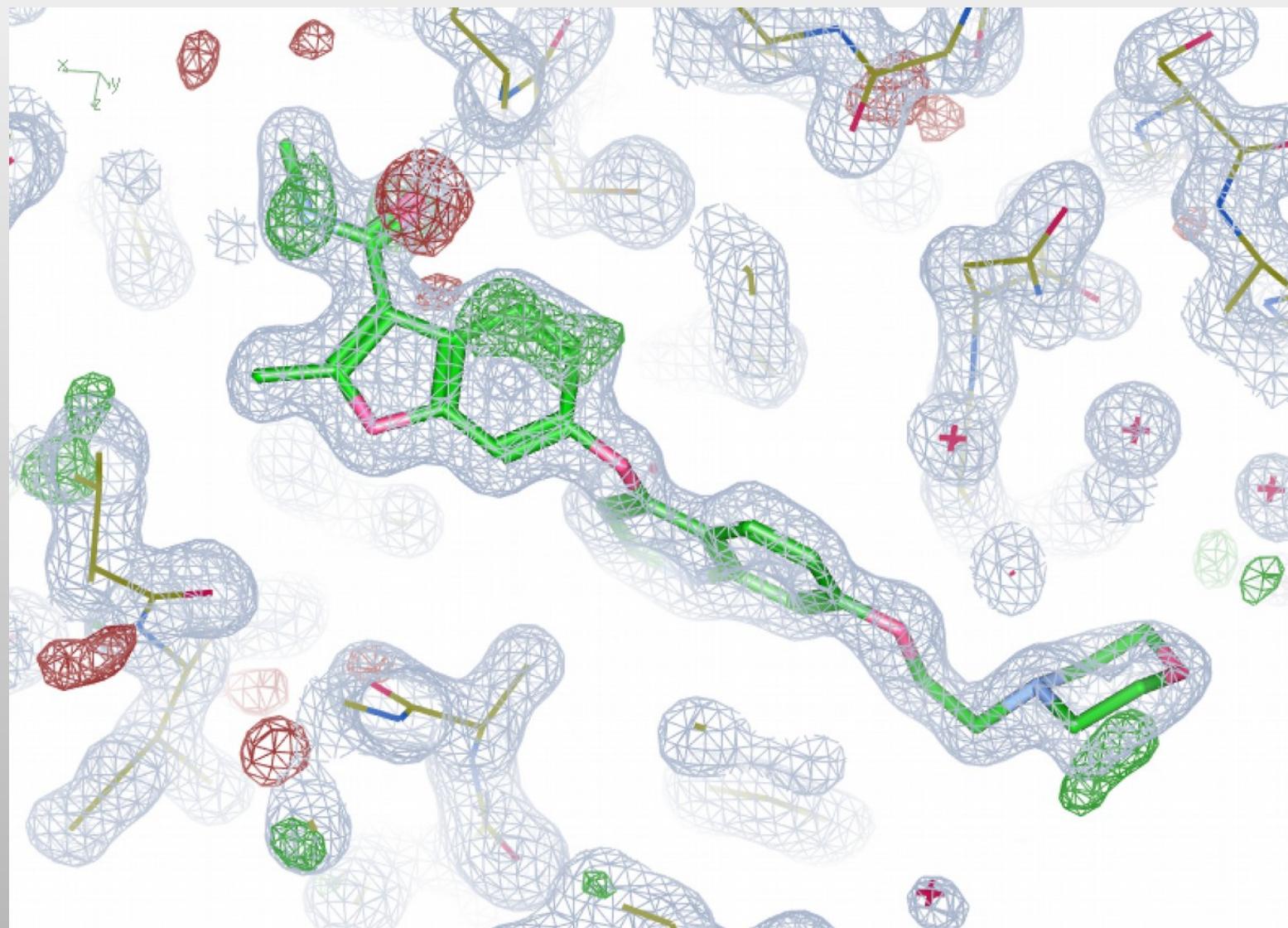
Multi-Criteria Markup



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



Acknowledgements

- LMB:
 - Garib Murshudov, Rob Nicholls, Fei Long,
 - Alexey Amunts, Alan Brown
- Kevin Cowan, Bernhard Lohkamp
- Libraries & Dictionaries:
 - Jane & Dave Richardson
 - Alexei Vagin
 - Eugene Krissinel