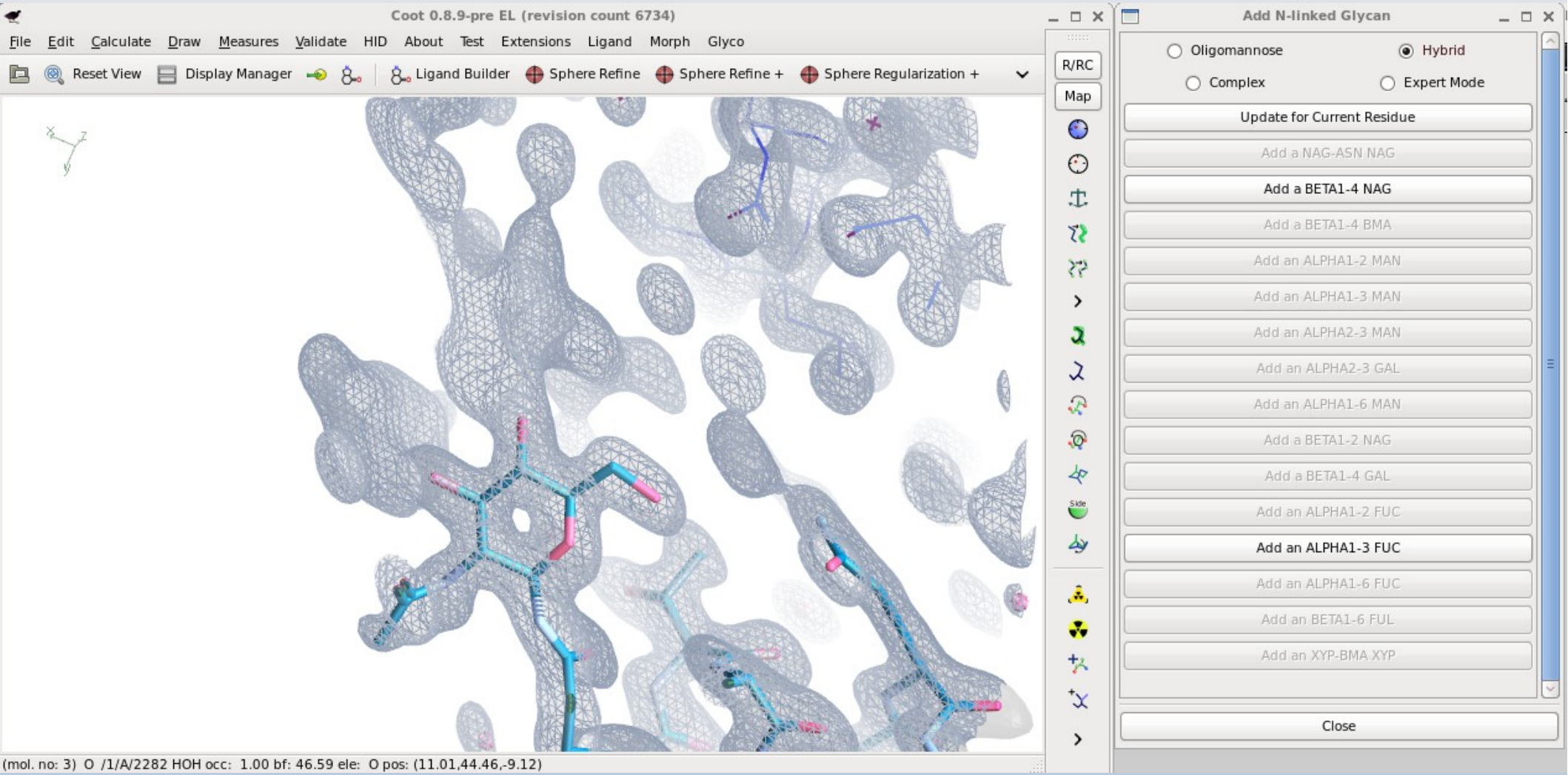
Coot Updates

Paul Emsley
Sept 2017

N-linked Carbohydrates

- Improved algorithm and re-worked GUI



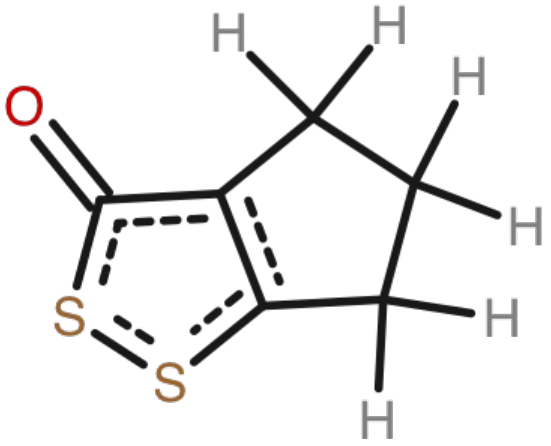
Improved Parsing of PDB CCD

- i.e. Improved construction of RDKit "in memory" representations
- Using sdf, mol2, SMILES for tools that use protein-ligand structures is now an unnecessary inconvenience
- Do everything from CCD entry (or restraints)
  - CCD Information-rich storage format and information there has been (or is being) embedded into RDKit representations of molecules
  - 2 ways to create molecules
    - Direct from the mmcif dictionary file
    - Using PDB file, residue specifier and mmcif dictionary file

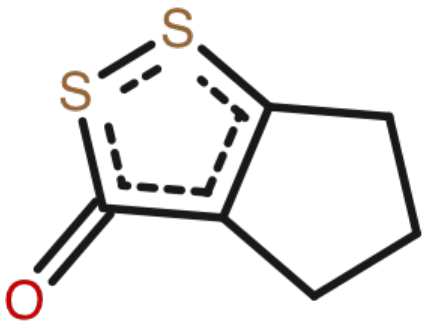
Better Use of Coot from Python

- Reworking tools so that they can be used from Python
  - rather than scripted in Python in the *Coot* application
- e.g. output to PNG files
  - rather than a representation in Lidia
  - direct coding of Cairo API
  - improved Rendering of Chemical Diagrams
    - (aka structure depiction)

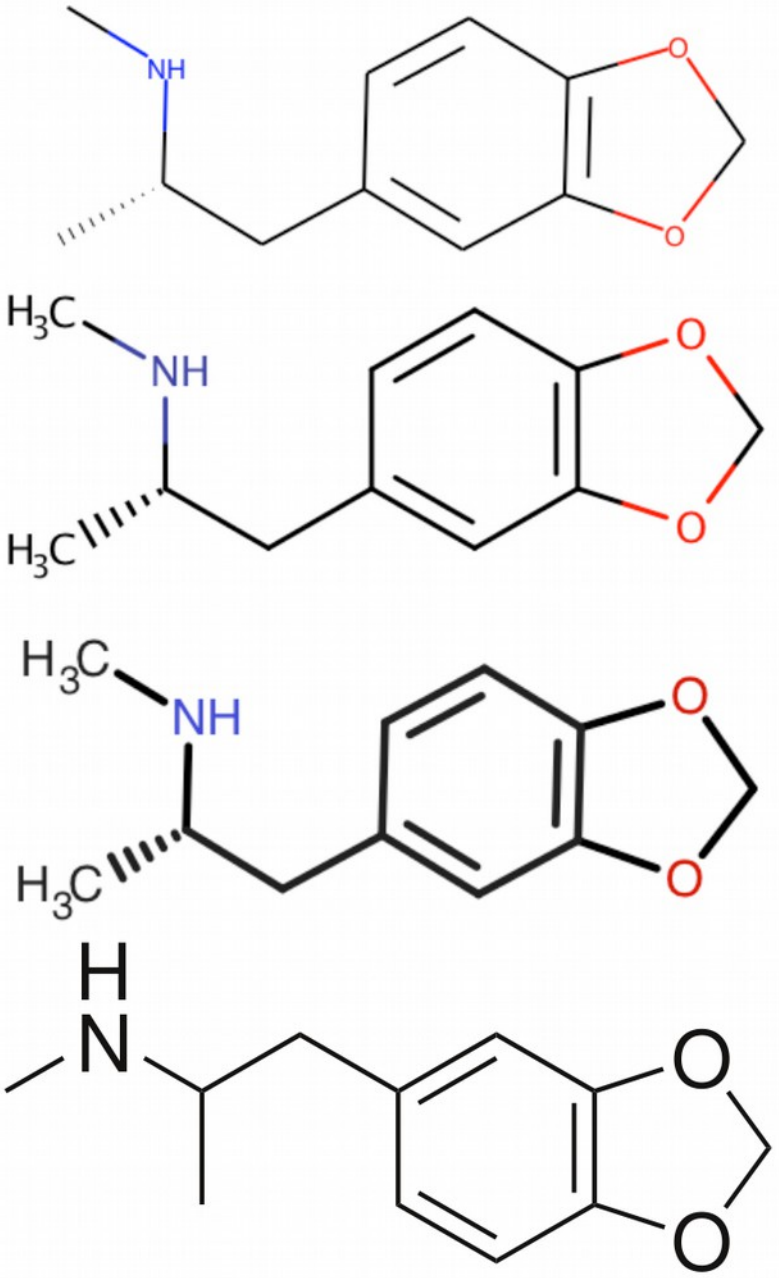
Depictions



before



after



Extended Atom Typing

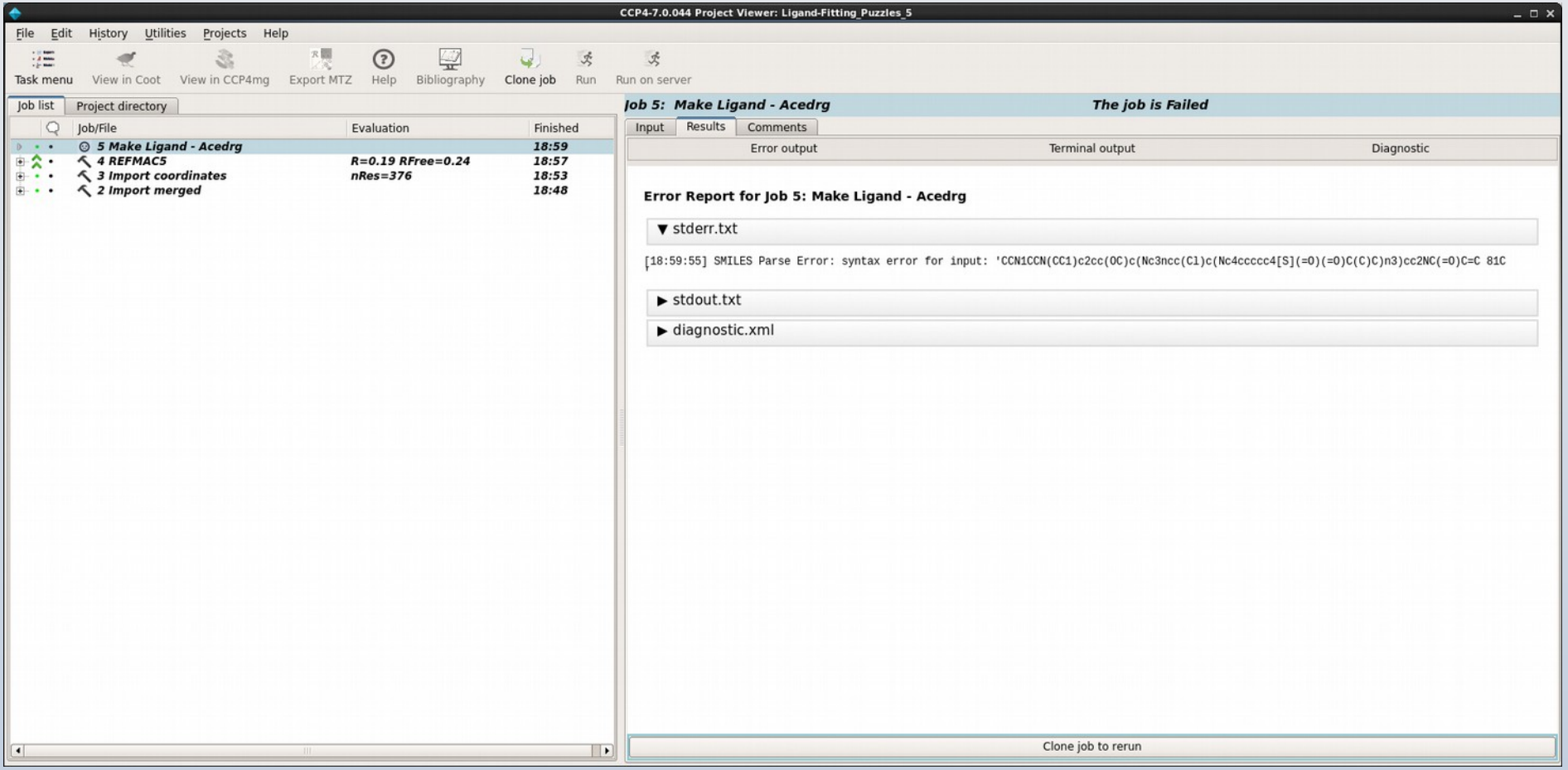
- Bringing parm@Frosst atom types to pyrogen
- More problematic because of Amber has tricky/confused atom types for imidazole and fused ring systems
  - Reference PATTY file uses Kekulized representation
  - SMARTS use aromatization
    - and RDKit's aromatization model
- On the plus side, there is an extensive validation suite
  - but not all of these molecules can be parsed with sanity!

Scriptable PLIFs

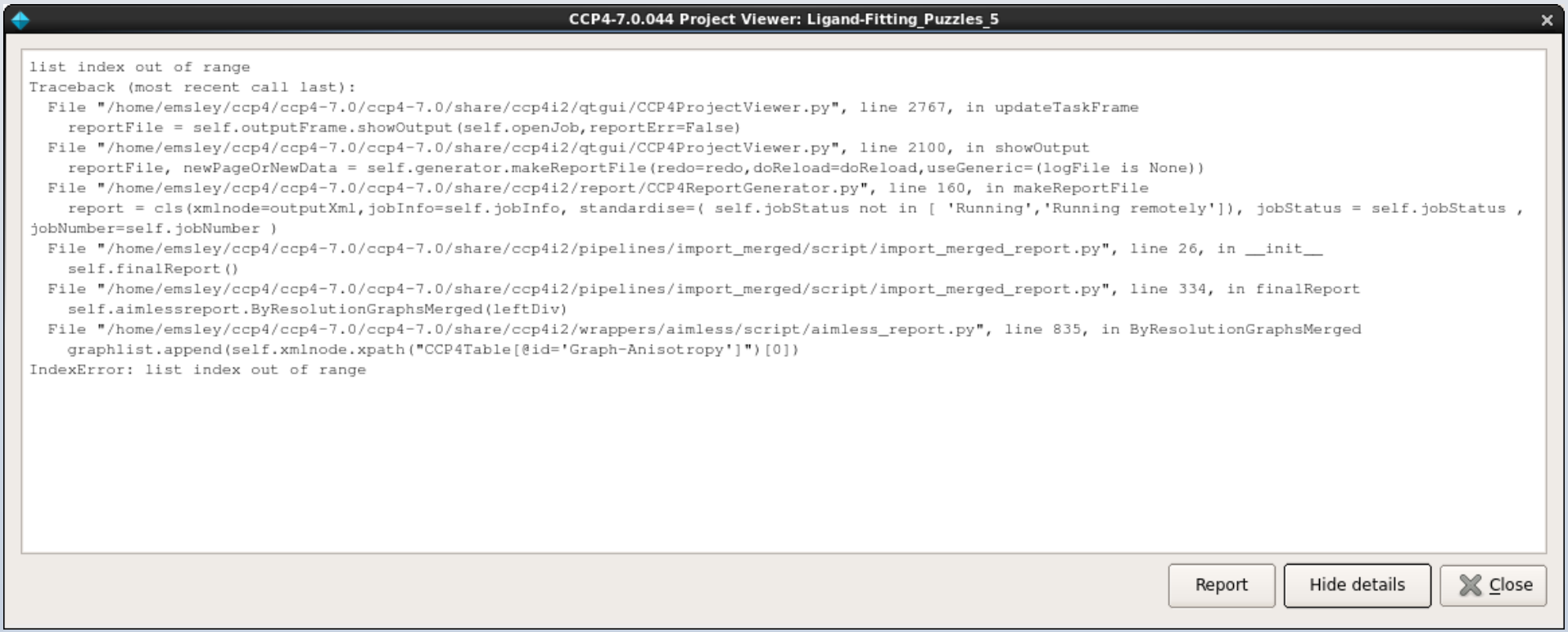
- Lidia/FLEV had been calculating and displaying protein ligand interactions
  - output to user via a (X11) canvas widget
  - a non-GUI API means access needs to be promoted to the Python API
  - Use SGC Fragment Screening structures for PLIF fingerprints and analysis

Ligand Puzzles 5

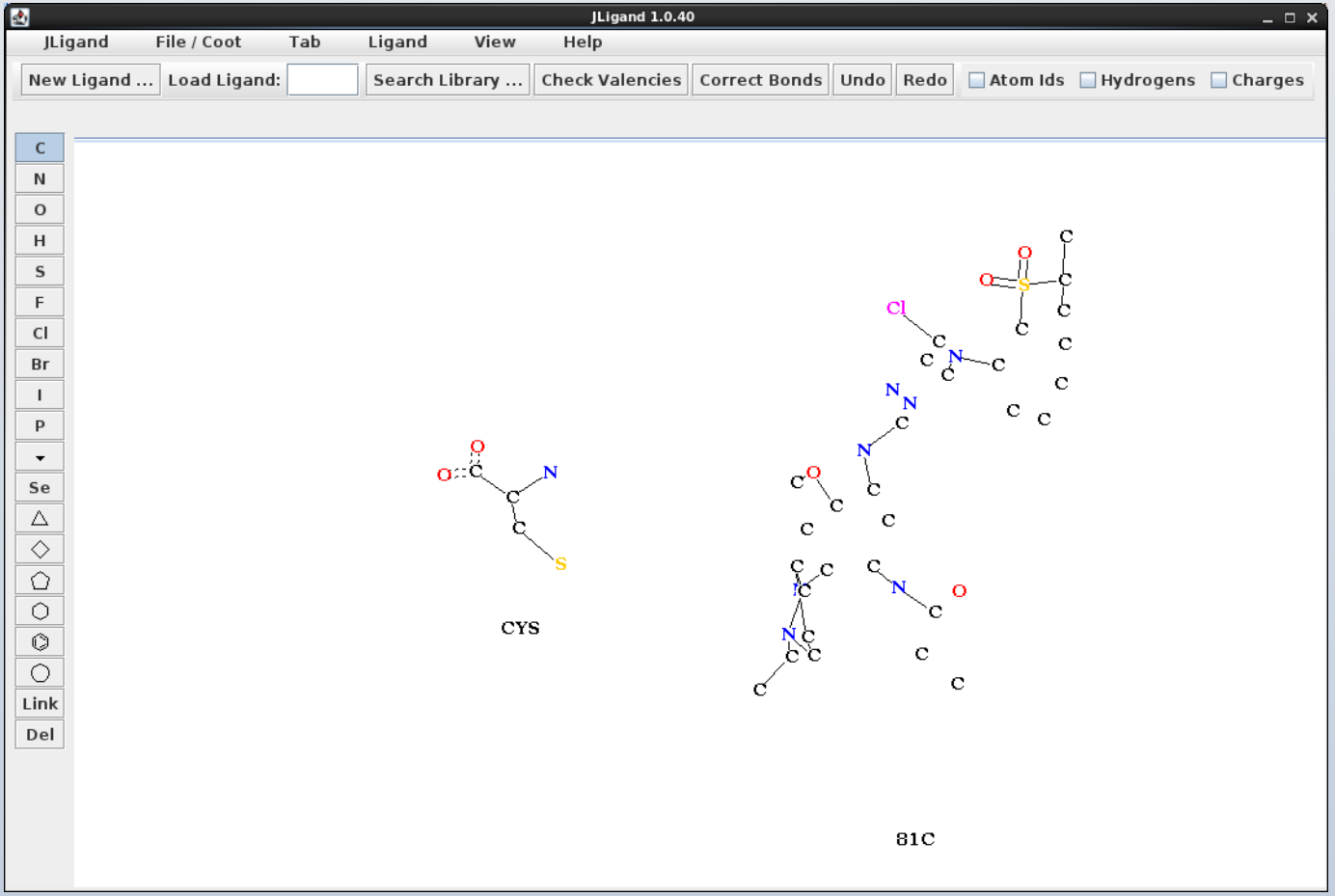
Puzzles 5-1


CCP4-7.0.044 Project Viewer: Ligand-Fitting_Puzzles_5
File Edit History Utilities Projects Help
Task menu
View in Coot
View in CCP4mg
Export MTZ
Help
Bibliography
Clone job
Run
Run on server
Job list
Project directory
Job/File
Evaluation
Finished
5 Make Ligand - Acedrg 18:59
4 REFMAC5 R=0.19 RFree=0.24 18:57
3 Import coordinates nRes=376 18:53
2 Import merged 18:48
Job 5: Make Ligand - Acedrg
The job is Failed
Input
Results
Comments
Error output
Terminal output
Diagnostic
Error Report for Job 5: Make Ligand - Acedrg
▼ stderr.txt
[18:59:55] SMILES Parse Error: syntax error for input: 'CCN1CCN(CC1)c2cc(OC)c(Nc3ncc(Cl)c(Nc4ccccc4[S](=O)(=O)C(C)C)n3)cc2NC(=O)C=C 81C
▶ stdout.txt
▶ diagnostic.xml
Clone job to rerun

Puzzles 5-1

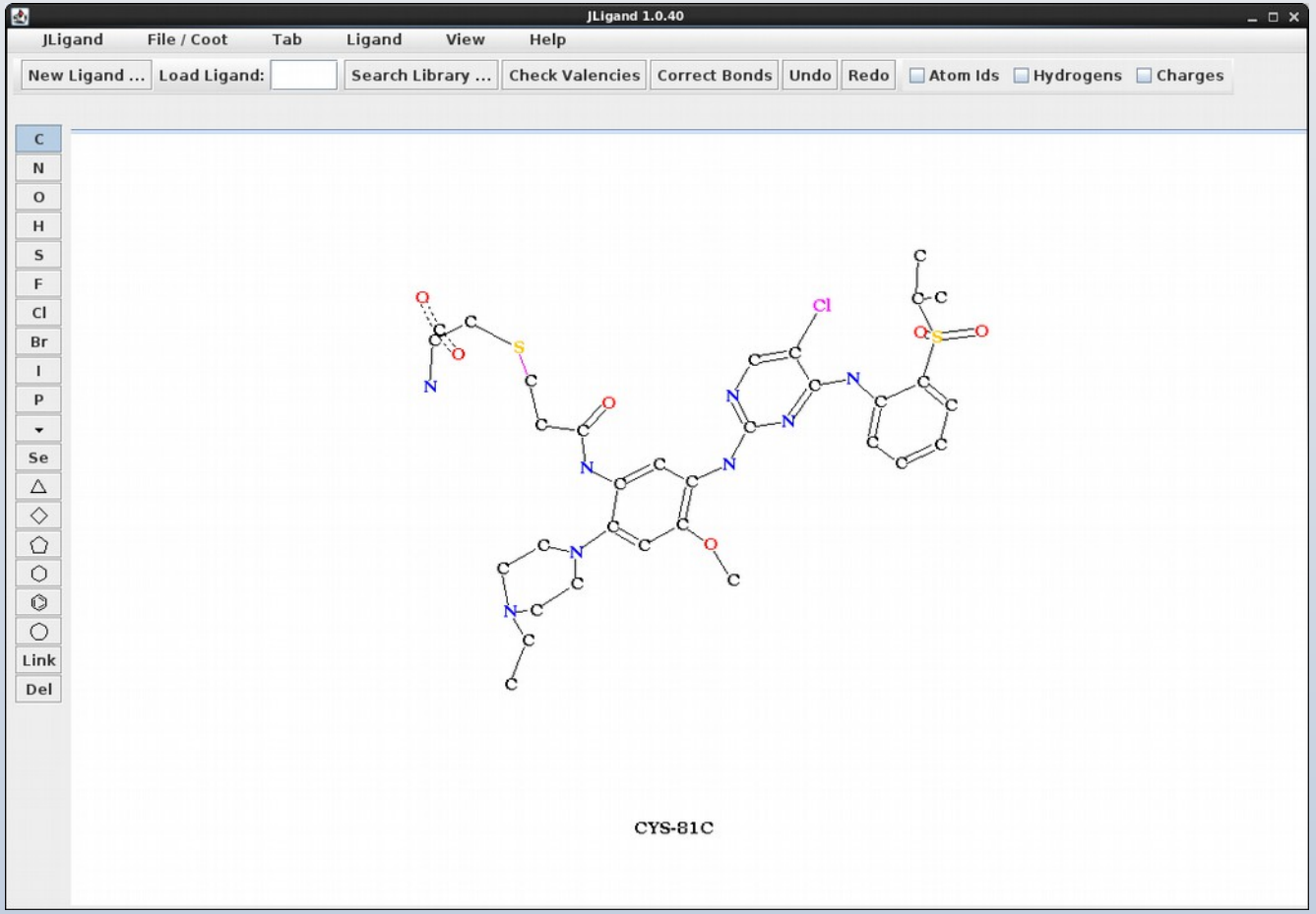

CCP4-7.0.044 Project Viewer: Ligand-Fitting_Puzzles_5
list index out of range
Traceback (most recent call last):
File "/home/emsley/ccp4/ccp4-7.0/ccp4-7.0/share/ccp4i2/qtgui/CCP4ProjectViewer.py", line 2767, in updateTaskFrame
reportFile = self.outputFrame.showOutput(self.openJob,reportErr=False)
File "/home/emsley/ccp4/ccp4-7.0/ccp4-7.0/share/ccp4i2/qtgui/CCP4ProjectViewer.py", line 2100, in showOutput
reportFile, newPageOrNewData = self.generator.makeReportFile(redo=redo,doReload=doReload,useGeneric=(logFile is None))
File "/home/emsley/ccp4/ccp4-7.0/ccp4-7.0/share/ccp4i2/report/CCP4ReportGenerator.py", line 160, in makeReportFile
report = cls(xmlnode=outputXml,jobInfo=self.jobInfo, standardise=(self.jobStatus not in ['Running','Running remotely']), jobStatus = self.jobStatus , jobNumber=self.jobNumber)
File "/home/emsley/ccp4/ccp4-7.0/ccp4-7.0/share/ccp4i2/pipelines/import_merged/script/import_merged_report.py", line 26, in __init__
self.finalReport()
File "/home/emsley/ccp4/ccp4-7.0/ccp4-7.0/share/ccp4i2/pipelines/import_merged/script/import_merged_report.py", line 334, in finalReport
self.aimlessreport.ByResolutionGraphsMerged(leftDiv)
File "/home/emsley/ccp4/ccp4-7.0/ccp4-7.0/share/ccp4i2/wrappers/aimless/script/aimless_report.py", line 835, in ByResolutionGraphsMerged
graphlist.append(self.xmlnode.xpath("CCP4Table[@id='Graph-Anisotropy']")[0])
IndexError: list index out of range
Report
Hide details
Close

JLigand: Default

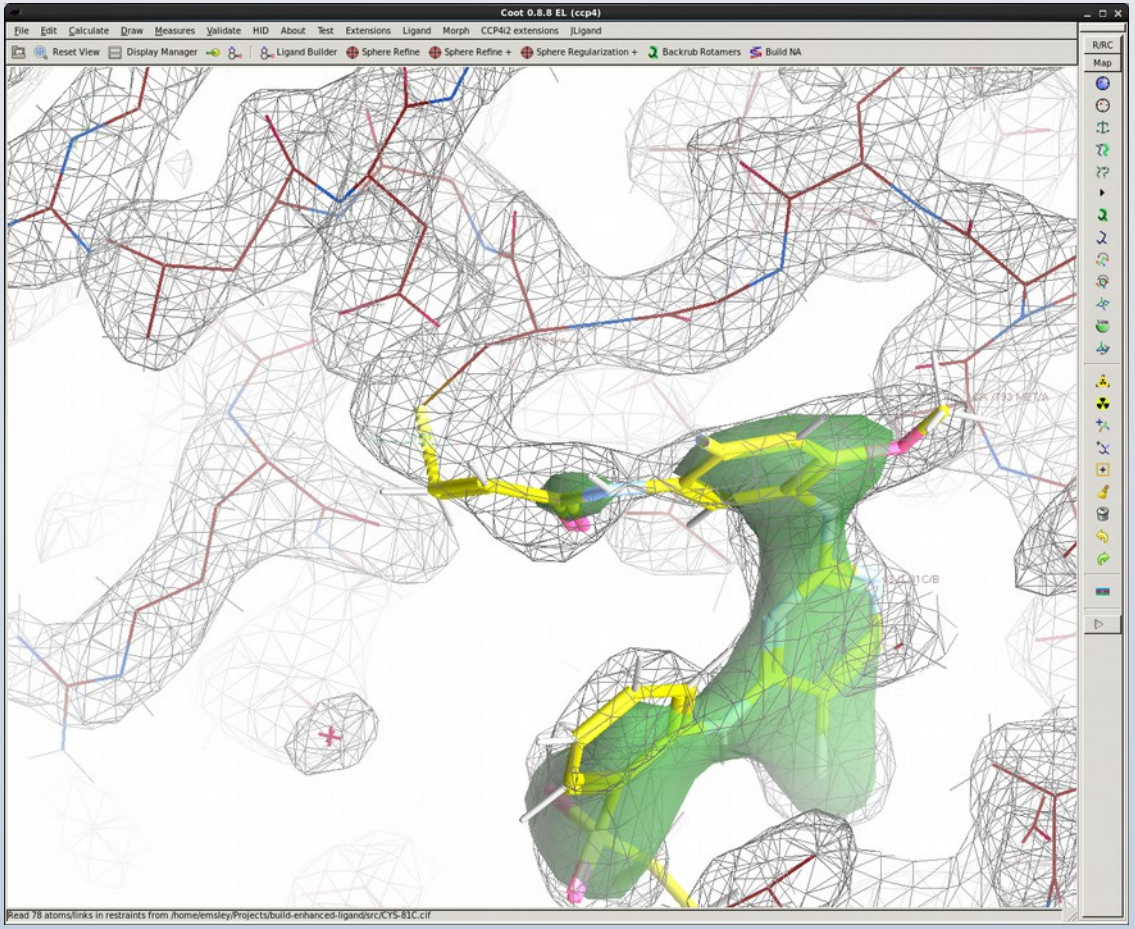


Acedrg and JLigand are incompatible

JLigand – After Hand-edit: Fixed bonds



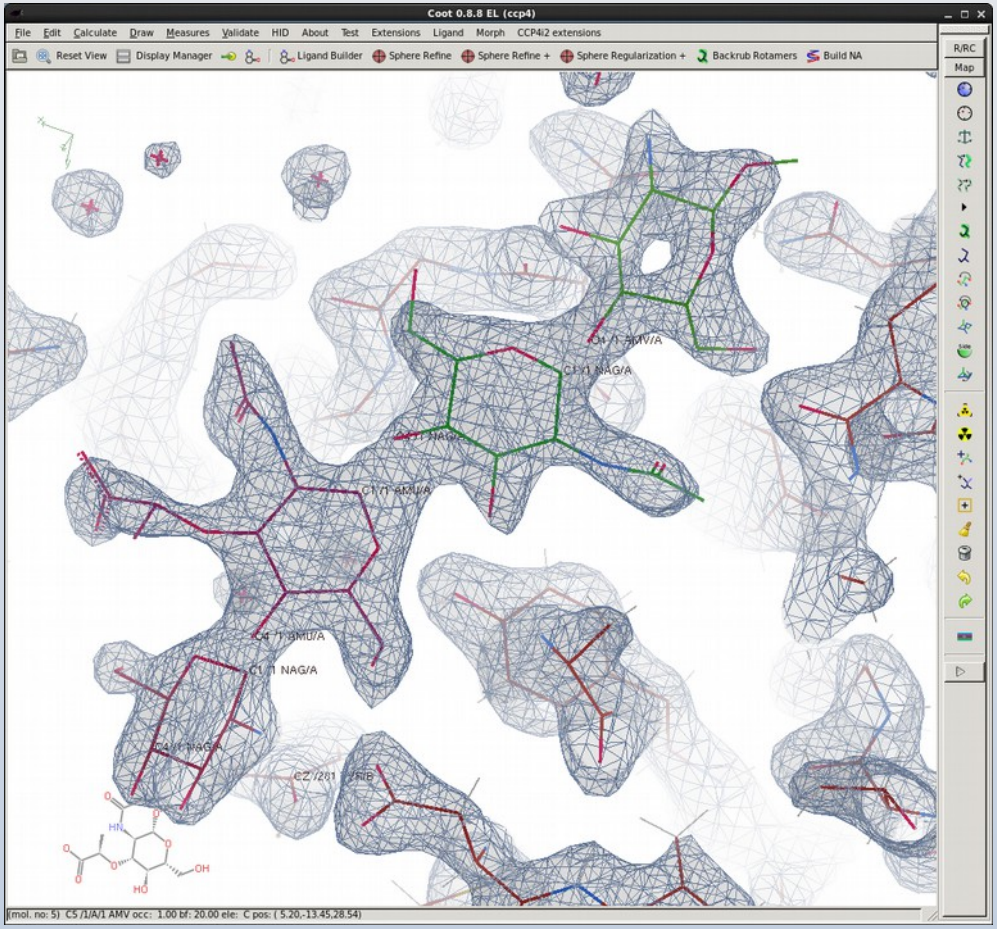
Puzzles 5-1: Bad Link



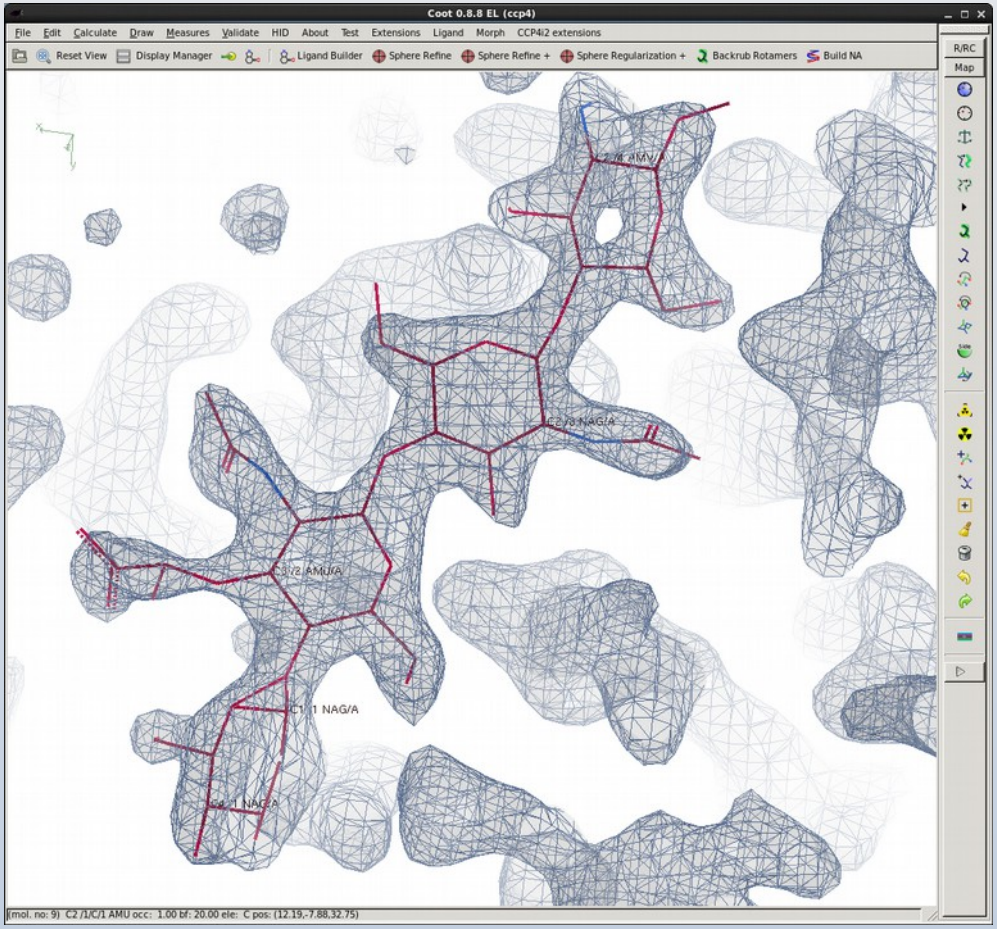
Puzzle 5-2

- Non-covalently bonded glycan
- Very tricky
- Coot's ligand finding/fitting not suitable
  - (it works on single monomers)
- After finding the blob, fit them one by one with β1-4 links
  - Different molecules, wrong residue numbers and chain-ids
  - Tedious to fix

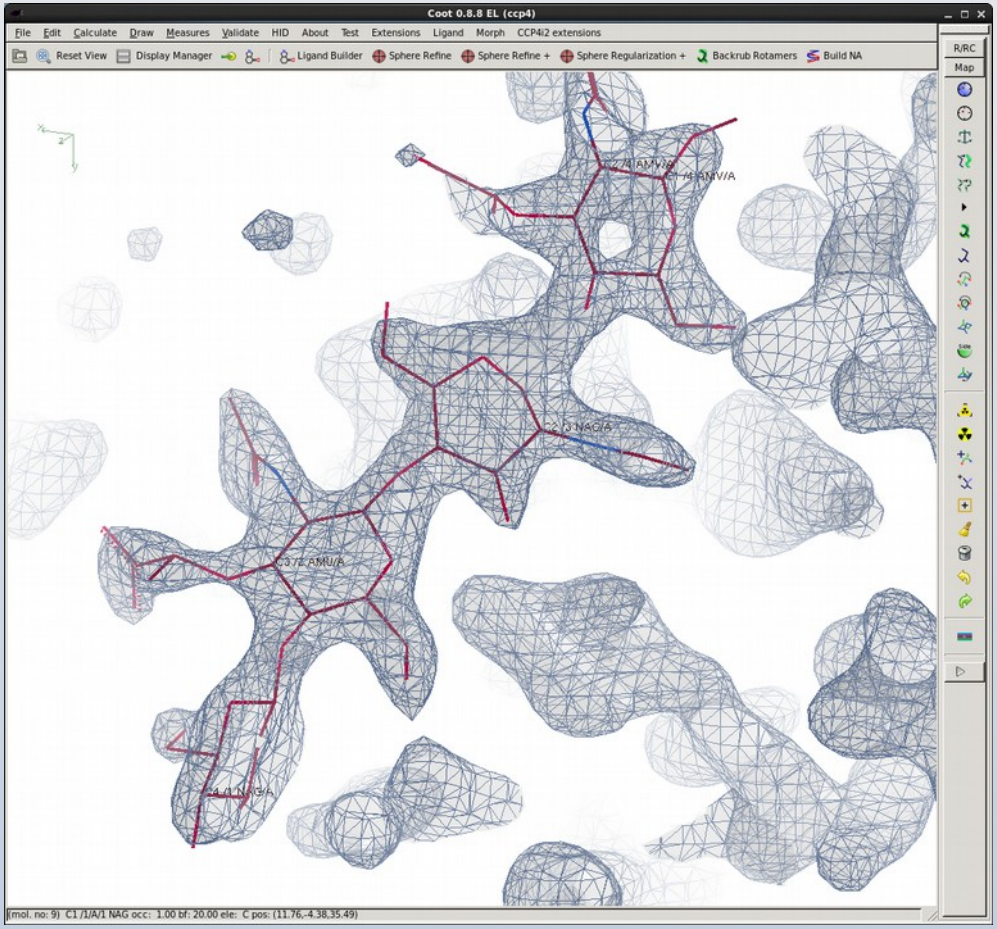
Puzzles 5-2


Coot 0.8.8 EL (ccp4)
File Edit Calculate Draw Measures Validate HID About Test Extensions Ligand Morph CCP4i2 extensions
Reset View
Display Manager
Ligand Builder
Sphere Refine
Sphere Refine +
Sphere Regularization +
Backrub Rotamers
Build NA
R/RC
Map
C4 /1 AMV/A
C1 /1 NAG/A
C1 /1 AMV/A
C4 /1 AMV/A
C1 /1 NAG/A
C4 /1 NAG/A
CZ /201 /NIB
(mol. no: 5) C5 /1/A/1 AMV occ: 1.00 bf: 20.00 ele: C pos: (5.20,-13.45,28.54)

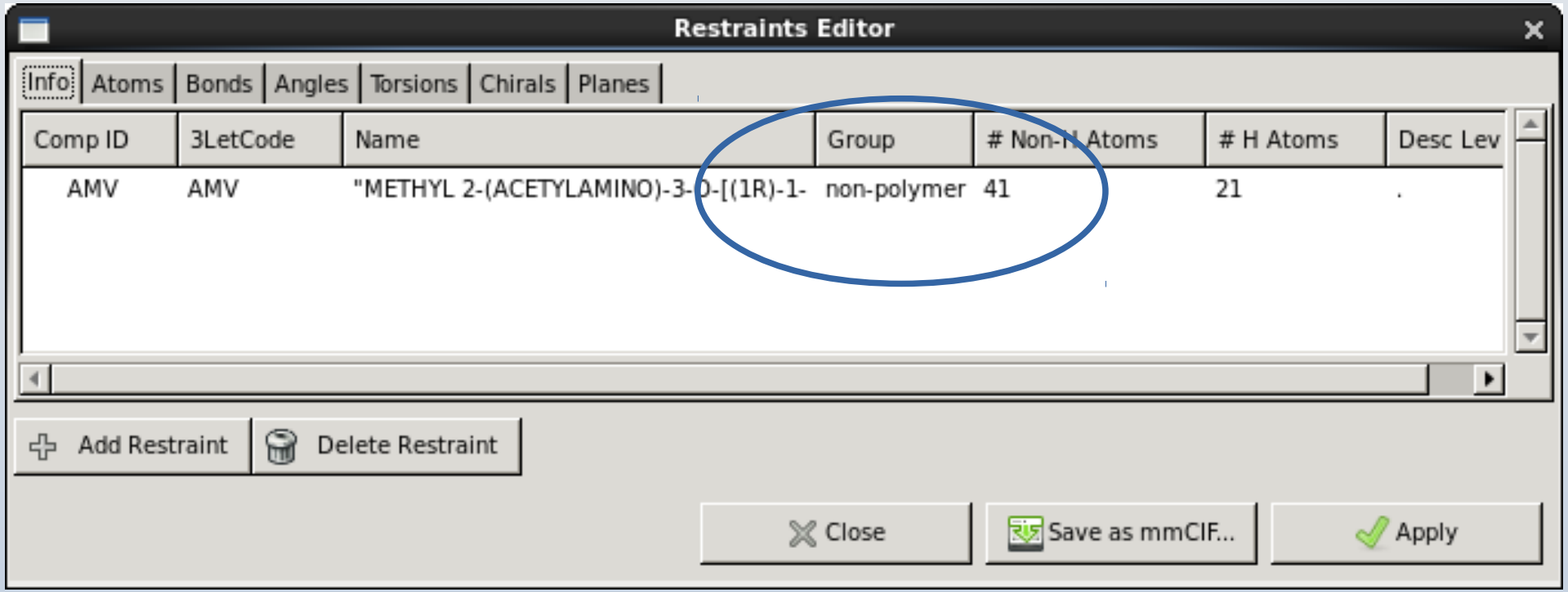
Puzzles 5-2


Coot 0.8.8 EL (ccp4)
File Edit Calculate Draw Measures Validate HID About Test Extensions Ligand Morph CCP4i2 extensions
Reset View
Display Manager
Ligand Builder
Sphere Refine
Sphere Refine +
Sphere Regularization +
Backrub Rotamers
Build NA
R/RC
Map
2 /3 AMU/A
C2 /3 NAG/A
3 /2 AMU/A
C1 /1 NAG/A
C4 /1 NAG/A

Puzzles 5-2

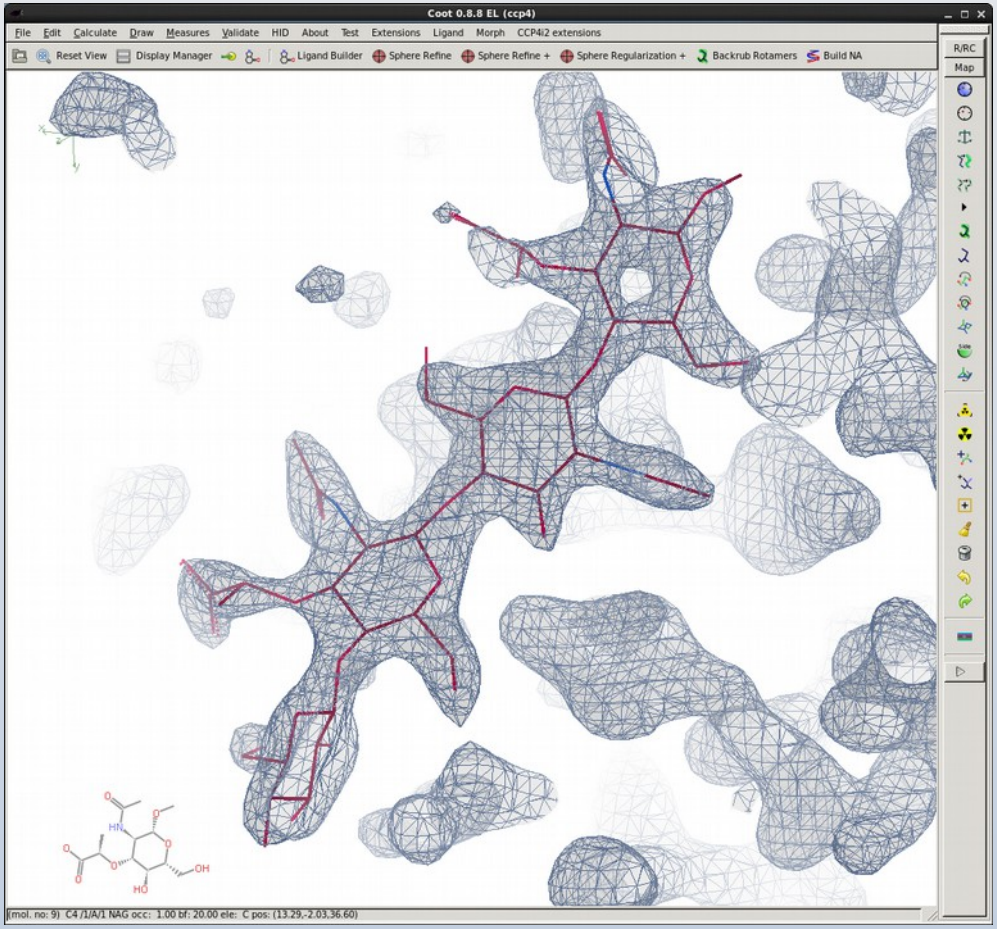

Coot 0.8.8 EL (ccp4)
File Edit Calculate Draw Measures Validate HID About Test Extensions Ligand Morph CCP4i2 extensions
Reset View
Display Manager
Ligand Builder
Sphere Refine
Sphere Refine +
Sphere Regularization +
Backrub Rotamers
Build NA
R/RC
Map
C3 /2 AMU/A
C4 /1 NAG/A
(mol. no: 9) C1 /1/A/1 NAG occ: 1.00 bf: 20.00 ele: C pos: (11.76,-4.38,35.49)

Puzzles 5-2

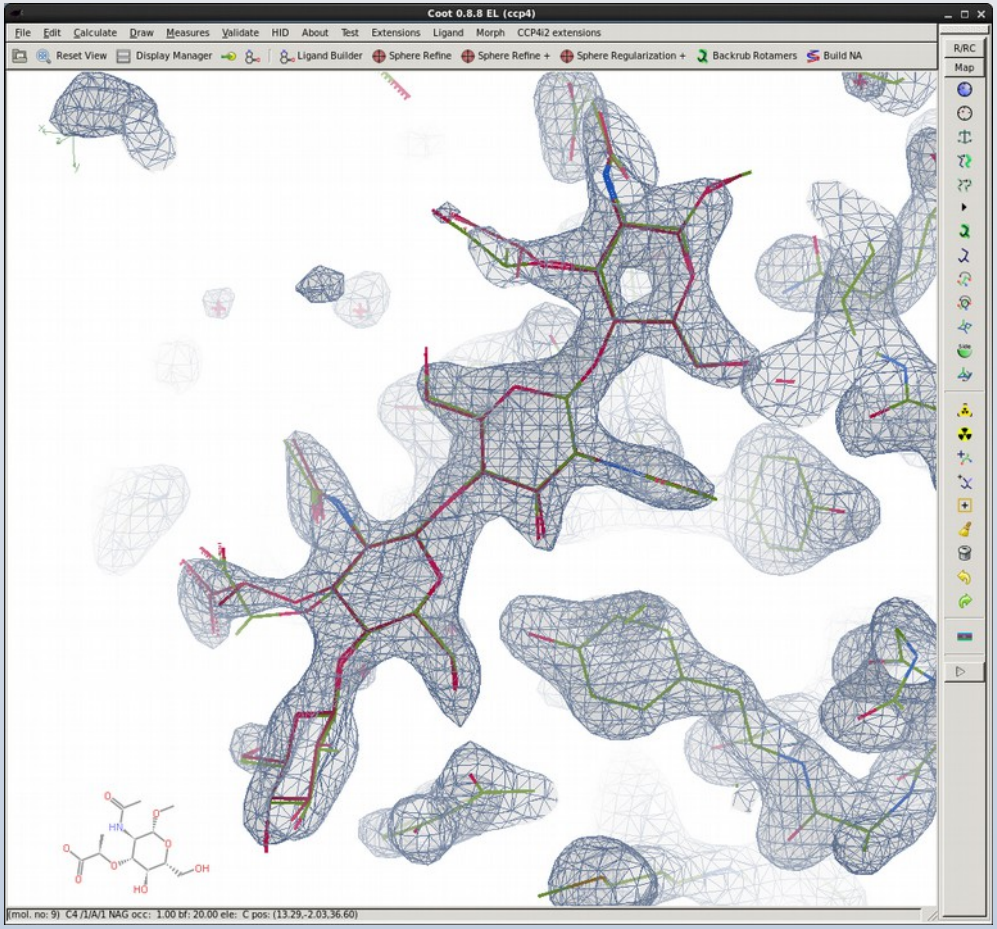


For pyranose β1-4 links, the group of AMV needs to be changed

Puzzle 5-2: Annealed



Puzzle 5-2: Comparison



Puzzles 5-2: Coot model is better?