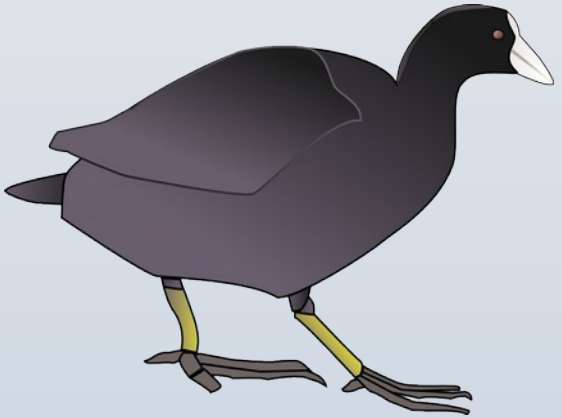
Coot Updates



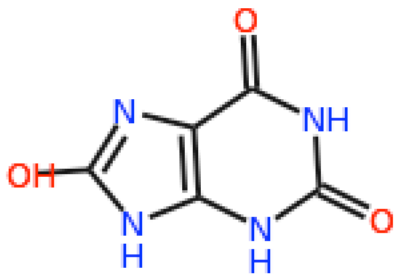
Paul Emsley
Sept 2016

Coot 0.8.4 Released

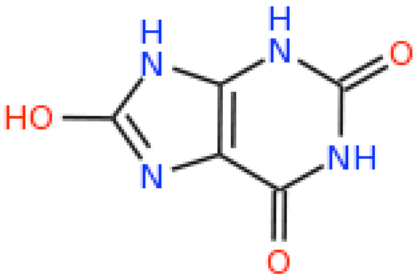
- Patterson from intensities
- Fill-partial-residue uses Backrub-rotamers
- Sequence dialog is now dynamically updated
- “Outliers Only” in Ramachanran Plot
- Improved lighting
  - solid objects look better
- Trans-torsion restraints
  - now in `mini-rsr` too
- PDBx 4-char H names
  - tracking mmdb2
- COD Atom types synced with Acedrg

Coot Updates 0.8.4

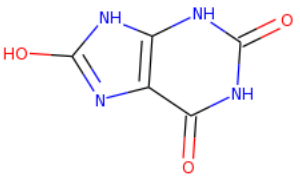
- Better PDBx dict ⟶ RDKit-mol when FLEVing
  - one code path had missing ring info
- Lidia CPK atom colouring
- Lidia "Flip" and "Rotate" molecule functions added
- Lidia Improved letter placement



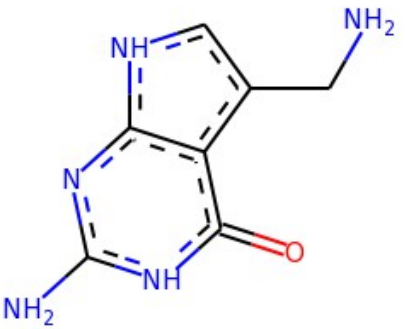
Old Coot



New Coot

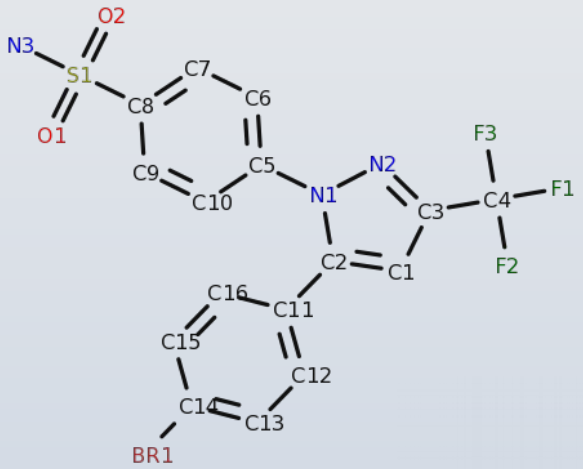


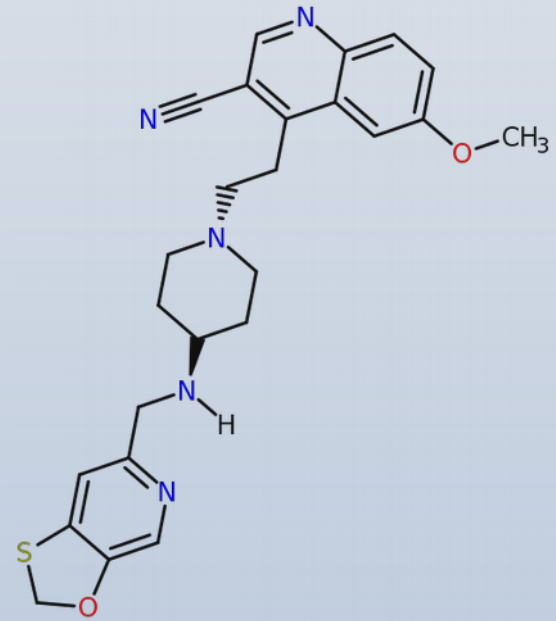
RDKit



Coot Updates 0.8.7: Lidia

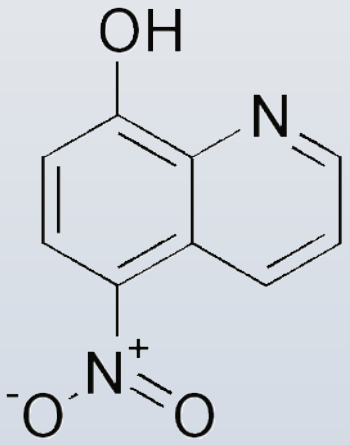
- Atom Name Mode
- Canvas scaling
- Editing improvements
  - bond dragging
- CH_3 superatoms
- Embedded python
  - network functions (“Fetch Molecule”)
  - QED
- Now configure works with `–with-lidia-only`
  - compiles `utils`, `geometry`, `lidia-core`, `lbg`
  - 38 other directories not compiled



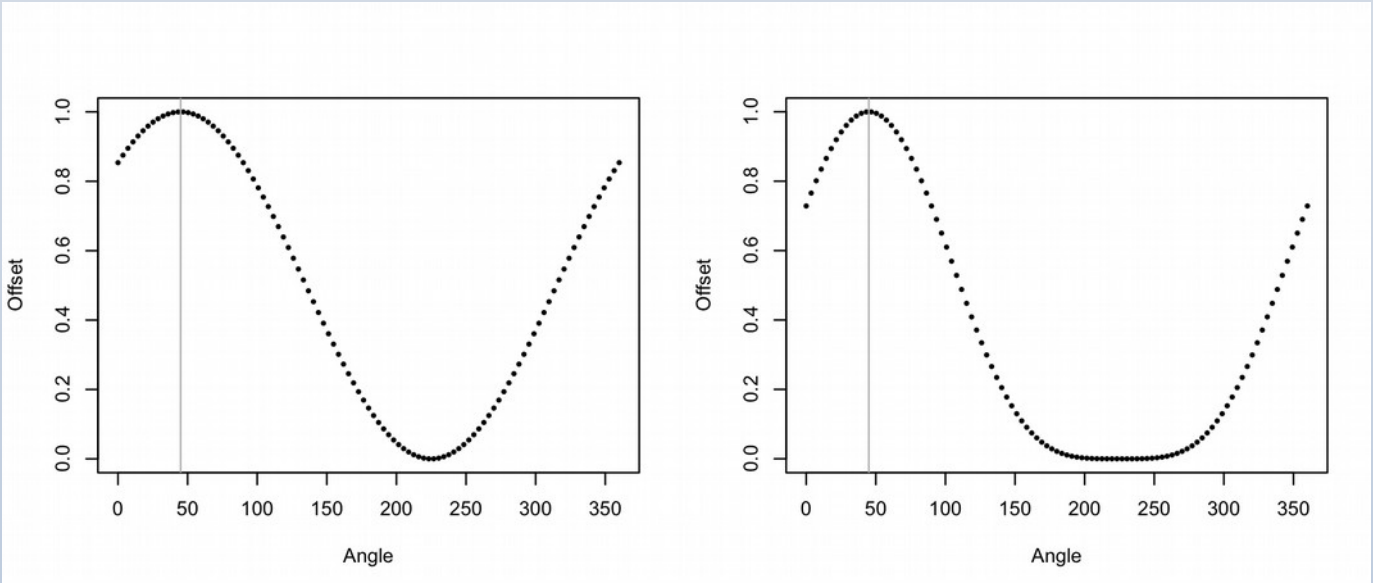


More Lidia Updates

- Typesetting nitros:



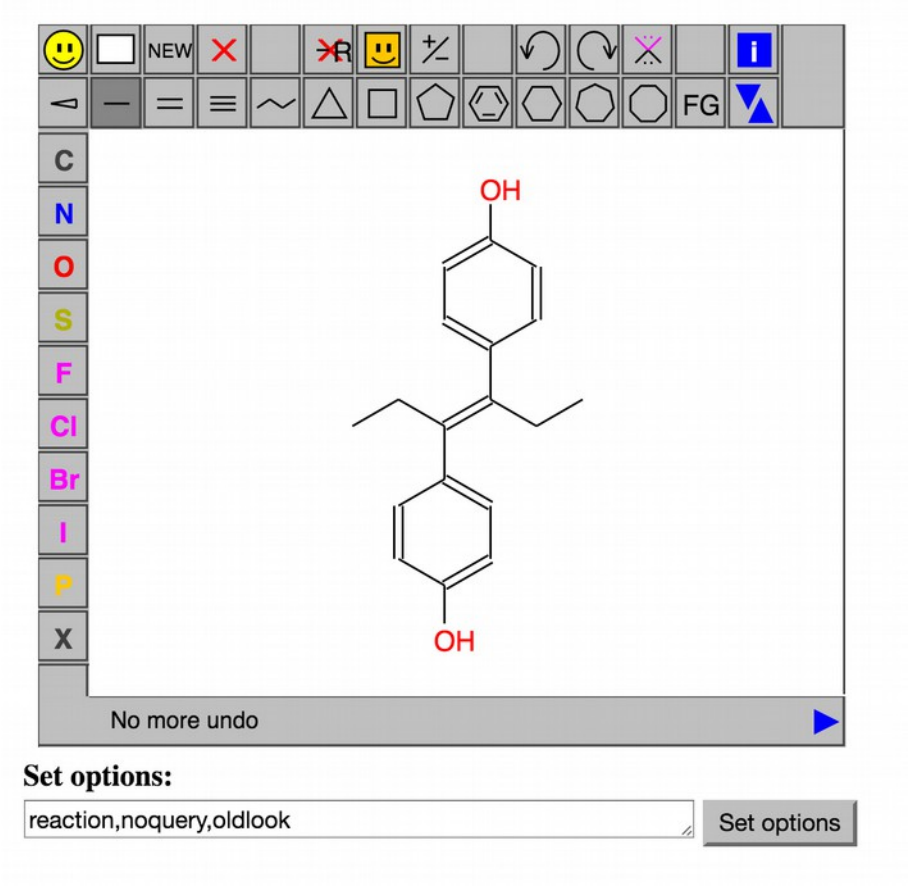
This image from Wikipedia



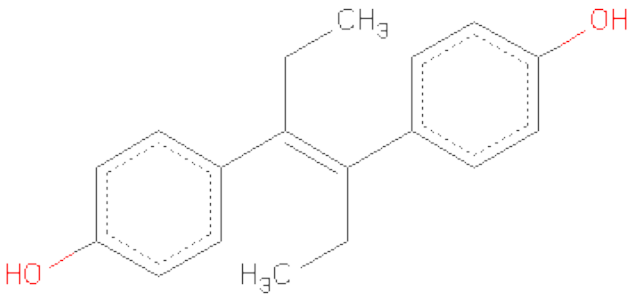
Note: canvas Y axis is “inverted”

Chemical Diagram Examples

diethylstilbestrol: DES, DB00255



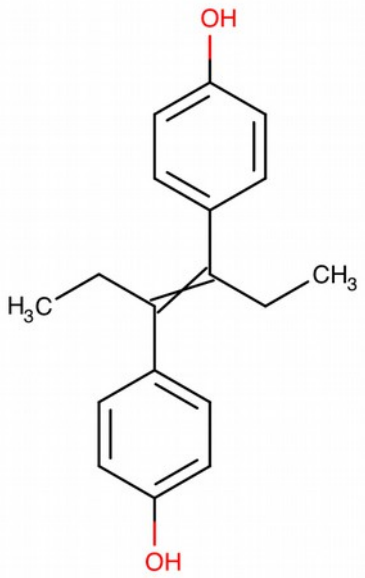
JSME



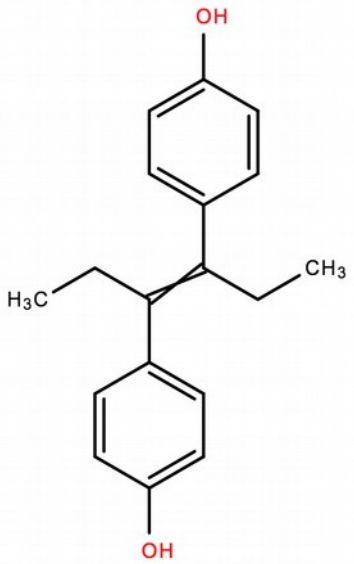
PDBe

Chemical Diagram Examples

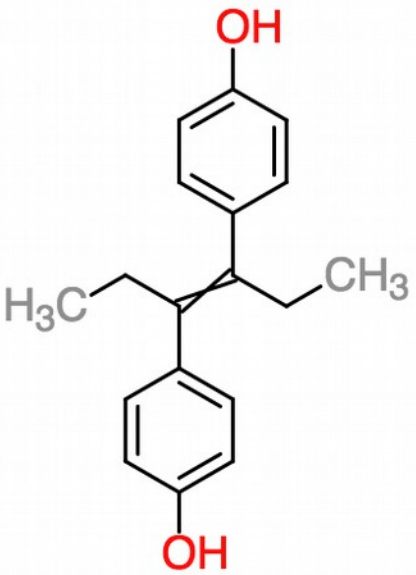
diethylstilbestrol: DES, DB00255
Mol File from DrugBank (MarvinSketch output)
Alkene double bond: cis or trans (either)



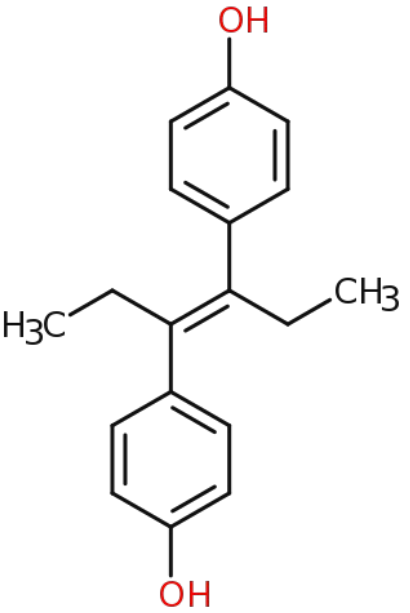
Marvin JS



Ketcher



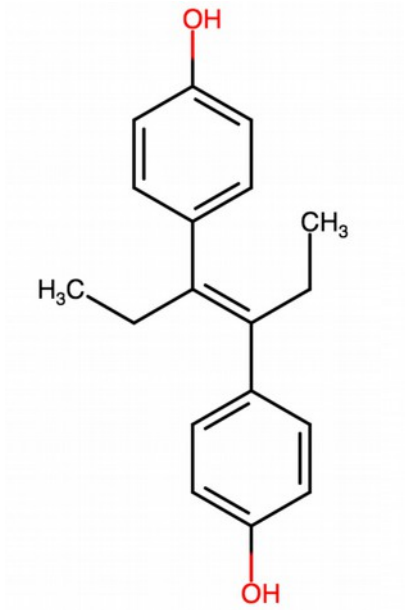
ChemDoodle



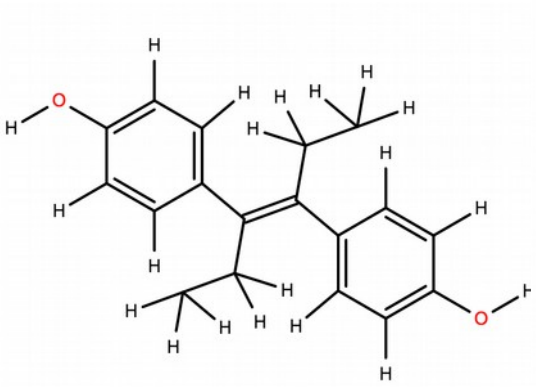
Lidia

Chemical Diagram Examples

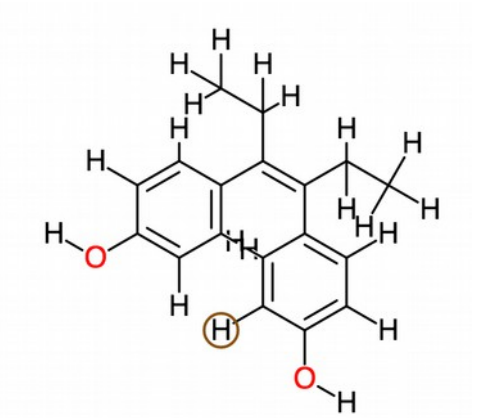
diethylstilbestrol: DES, DB00255
Mol File from PDBeChem
(after 3D → 2D)



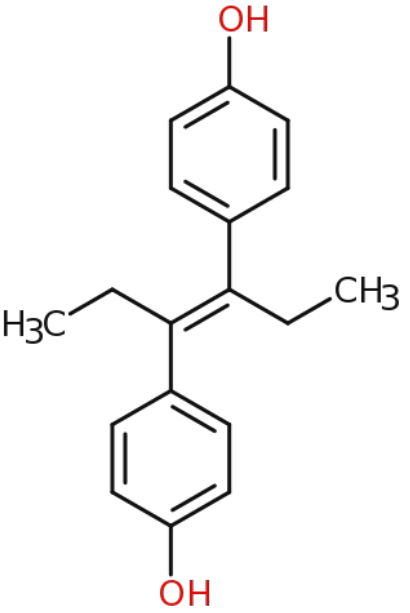
Marvin JS



Ketcher



ChemDoodle



Lidia

What Can Lidia Do?

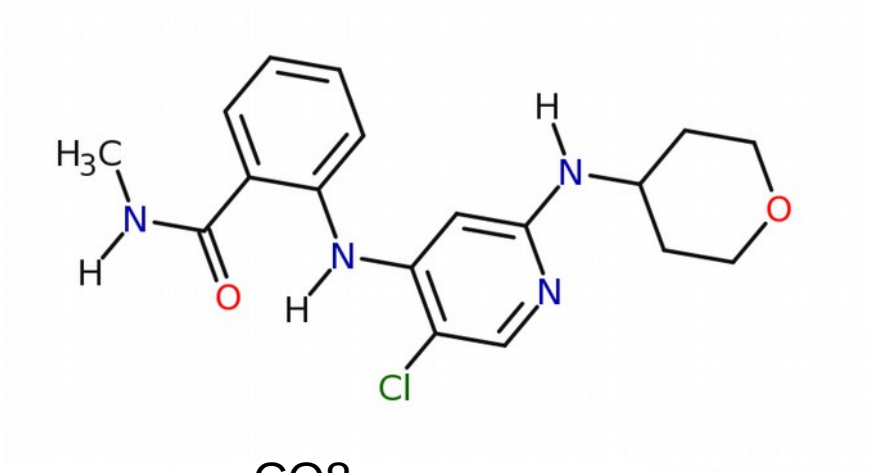
- Run Acedrg, Pyrogen, prodrg
- Import from molecule name
- Drag and drop
- Remove Hydrogen atoms
- QED/”Lipinski” descriptors
- Clean up
- Import from CCD, Refmac monomer library comp_id
- Output SMILES, mol, png, pdf, svg

What Can't Lidia Do?

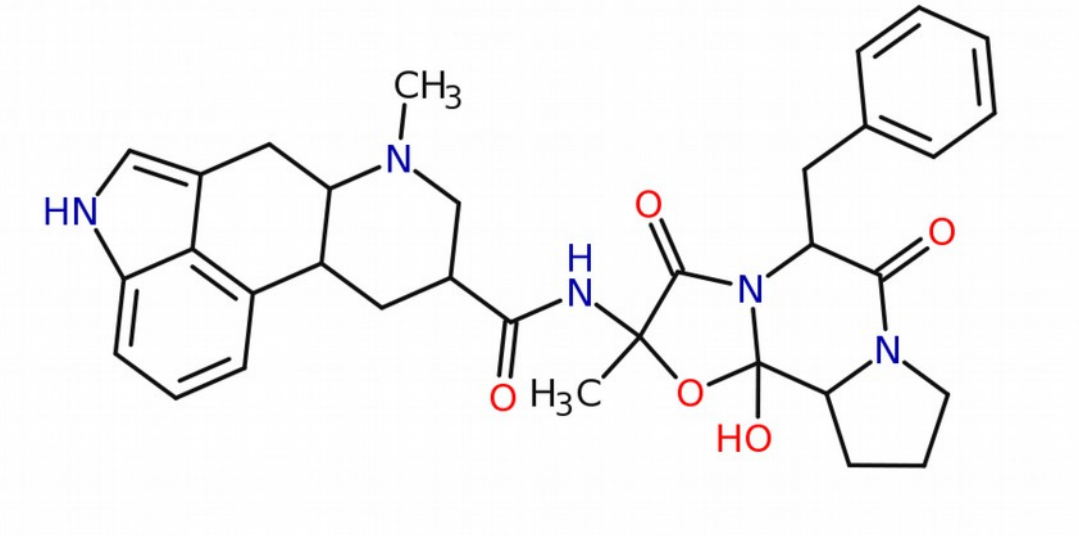
- Multiple molecules
- Reactions
- Add Hydrogen atoms
- Represent unresolved chirality

Lidia: Chemical Diagrams: CCD → 2D

- OK, not flawless



CQ8



2GM

Chirality Conversion

- Encoded in SMILES strings with @, @@
- Encoded in mmCIF wwPDB as R,S stereocentres
- Encoded in mmCIF as signed volume of local neighbours
- Encoded by wedge bonds in MOL file
- Encoded by inference from 3D coordinates
- Stored as mmdb Residues object
- Stored as parsed dictionary object
- Stored as molfile parsed object
- Stored as CIP-based CW, CCW Chirality attribute in RDKit

Pyrogen Updates

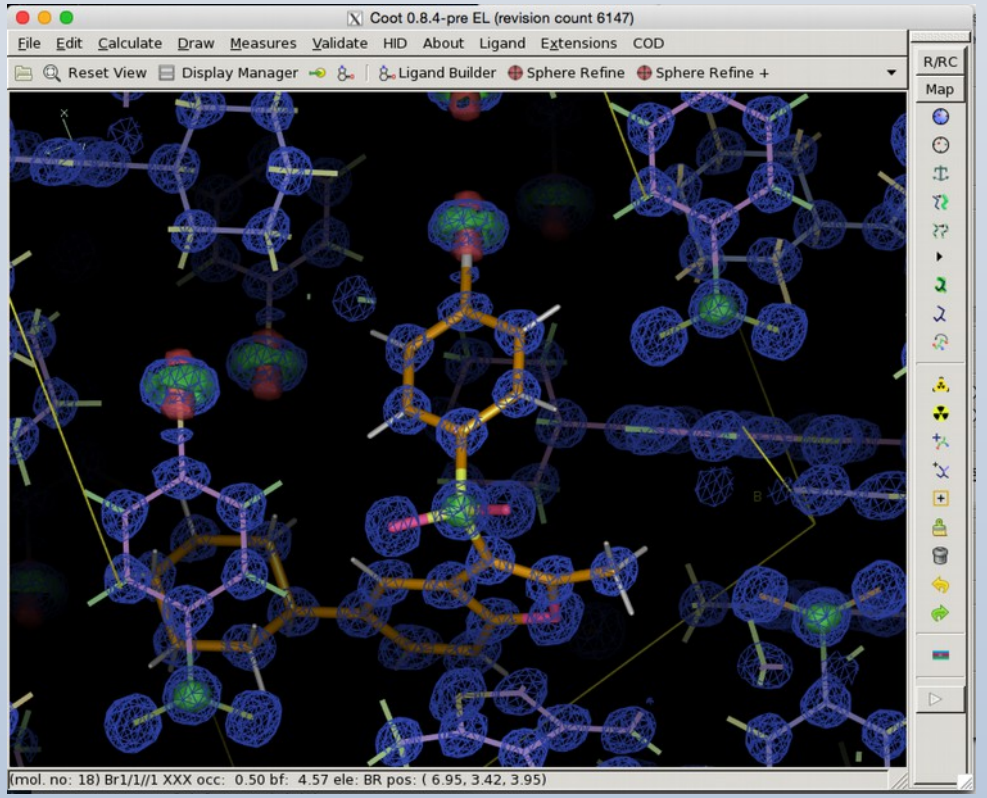
- tautomer mode returned a list of SMILES strings
- tautomer results were different when the input was pdbx CCD cif and when the input was the SMILES from the CCD cif
  - the SMILES string generated more kekulizable results
  - 1 *vs.* 24 (for 8HX)
  - now generate the same results
  - call the tautomer enumeration with a molecule that does not contain hydrogen atoms

Pyrogen Updates

- Pyrogen no longer inherits the monomer group from a matching dictionary
  - *e.g.* BrC1=CC=CN2C1=NC=C2N[C@@H](S(=O)(C3=CN=CS3)=O)F
  - matched a TRP and was called `L-peptide` despite no N, CA, C, O, H equivalents

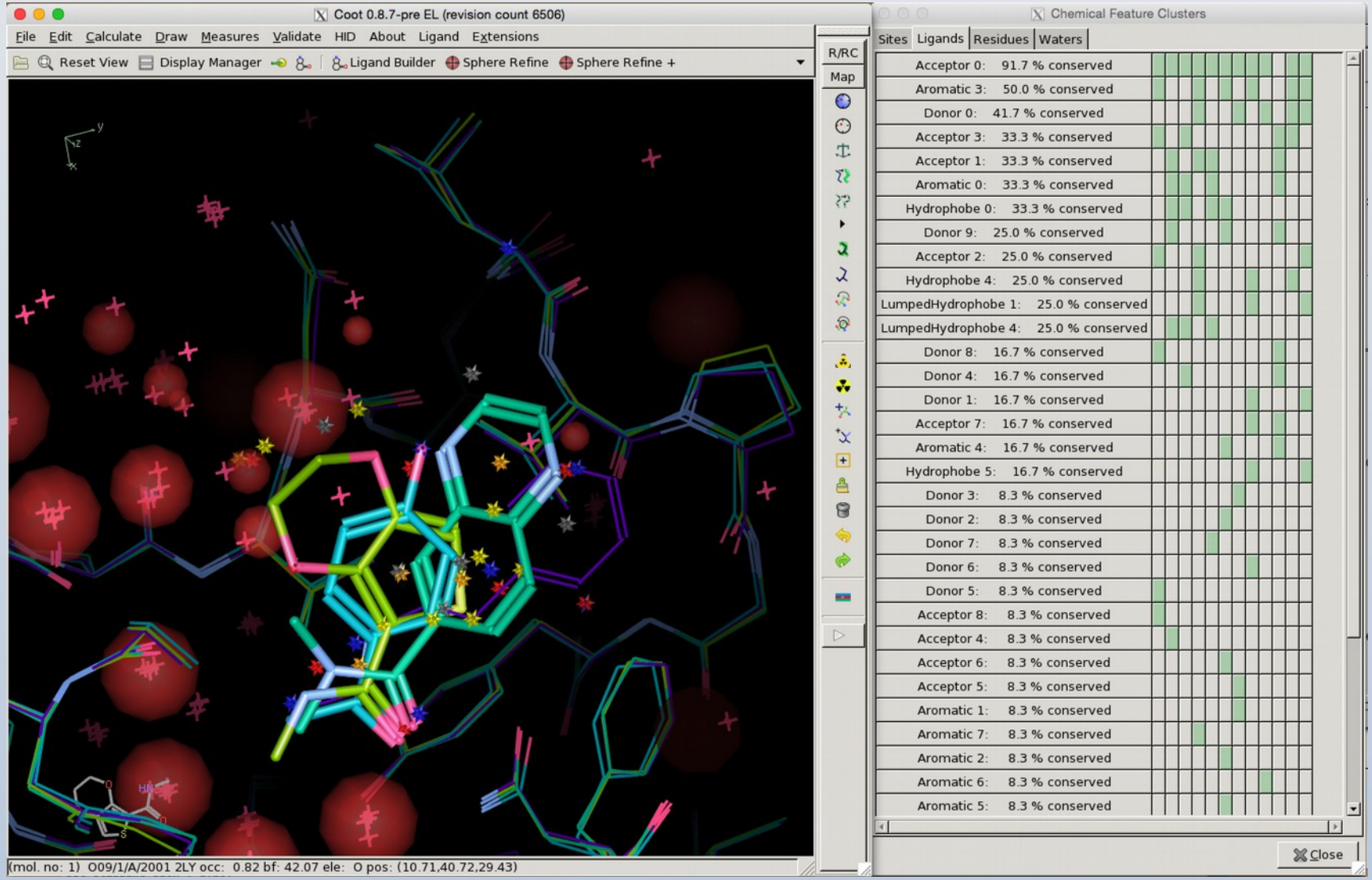
Coot Updates [continued]

- Add Alt-Confs to Ligand via torsion-angle selection
- [demo]
- Now has maps from COD (calc sfs)



Coot Updates [continued]

- Chemical Feature Clustering
  - needs SciPy



Coot Updates

- Release in the next week

RCSB SDF files still broken

- Atom names in caps are wrong
  - probably
- PDBe files also
- comp_id SDF files do not contain atom names
  - Accession code-based model ligand SDF files *do* contain atoms

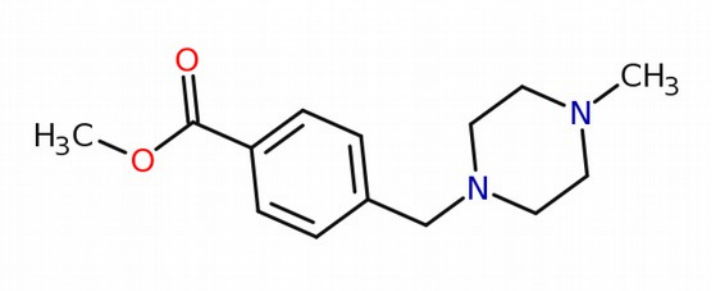
(blank)

The Ligand Fitting Challenge

Problem Set 1: Autumn Equinox 2016

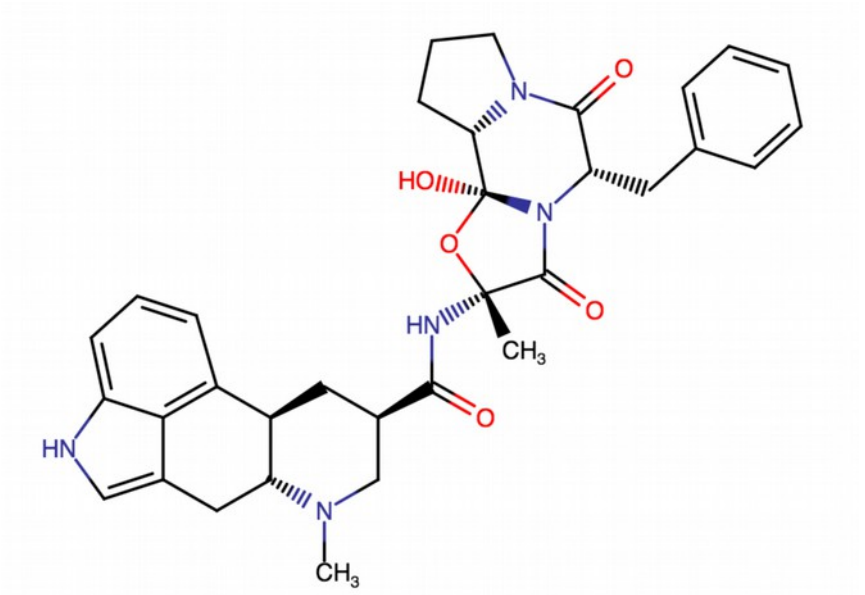
Ligand Fitting Challenge: Inputs

- mmCIF model, a SMILES string and Fobs mtz



- Problem 1:
  - Bromodomain
  - Resolution: 2.1 Å, poor density for piperazine
  - confusing missing waters
  - WONKA examples dataset

- Problem 2:
  - GPCR
  - Ligand: ergotamine derivative
  - DB00320
  - Resolution 2.8 Å
  - Chiral protonated N6
  - Is ergotamine in the PDB?
    - how do I transfer the atom names?



Ligand Fitting Challenges

- Inside the CCP4 world these problems need intervention and thought
  - Can the correct fitting be automated using other tools?
- Given a ligand SMILES, how do I find out to which ligand (about which the PDB already knows) it is most similar?
  - Change the atom names of my new ligand to match that
- If I edit a ligand, how can I impose the same/similar atom names to the original ligand? (A ligand about which PDBeChem may not already know)

Judit

Paul E

Paul R.

Ben

All: Reflections, Action Points

How I did it...

- Acedrg
  - for SMILES ⟶ extract string ⟶ dictionary
- Coot
  - Refmac to make the map
  - Find Unmodelled blobs
  - Read Acedrg dictionary
  - Find Ligands (which fails)
  - Add waters
  - Find Ligands (which ~ works) – flip and symmetry shuffle
  - Merge Molecules
  - Refine with Refmac
  - Hydrogenate
  - FLEV

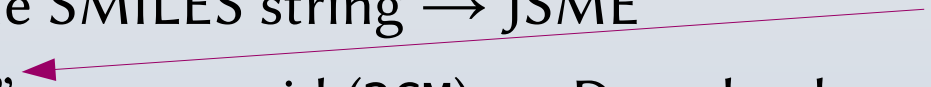
Didn't work for problem-2, Fcalcs already

Extra Work for 2GM

- An analysis of the 2GM ligand with the pocket shows a missing NH^+ “hydrogen bond” to COO^- from an ASP
- How do we add it?
  - I edited the SMILES string
  - I didn't know which N was which so I charged each of them
  - (with a bit more thought I could have worked it out)
  - Reran Acedrg, fit the ligand again
  - Wrong Chirality!
    - Edit Restraints → Chirals → N1 → negative
  - Rerefine

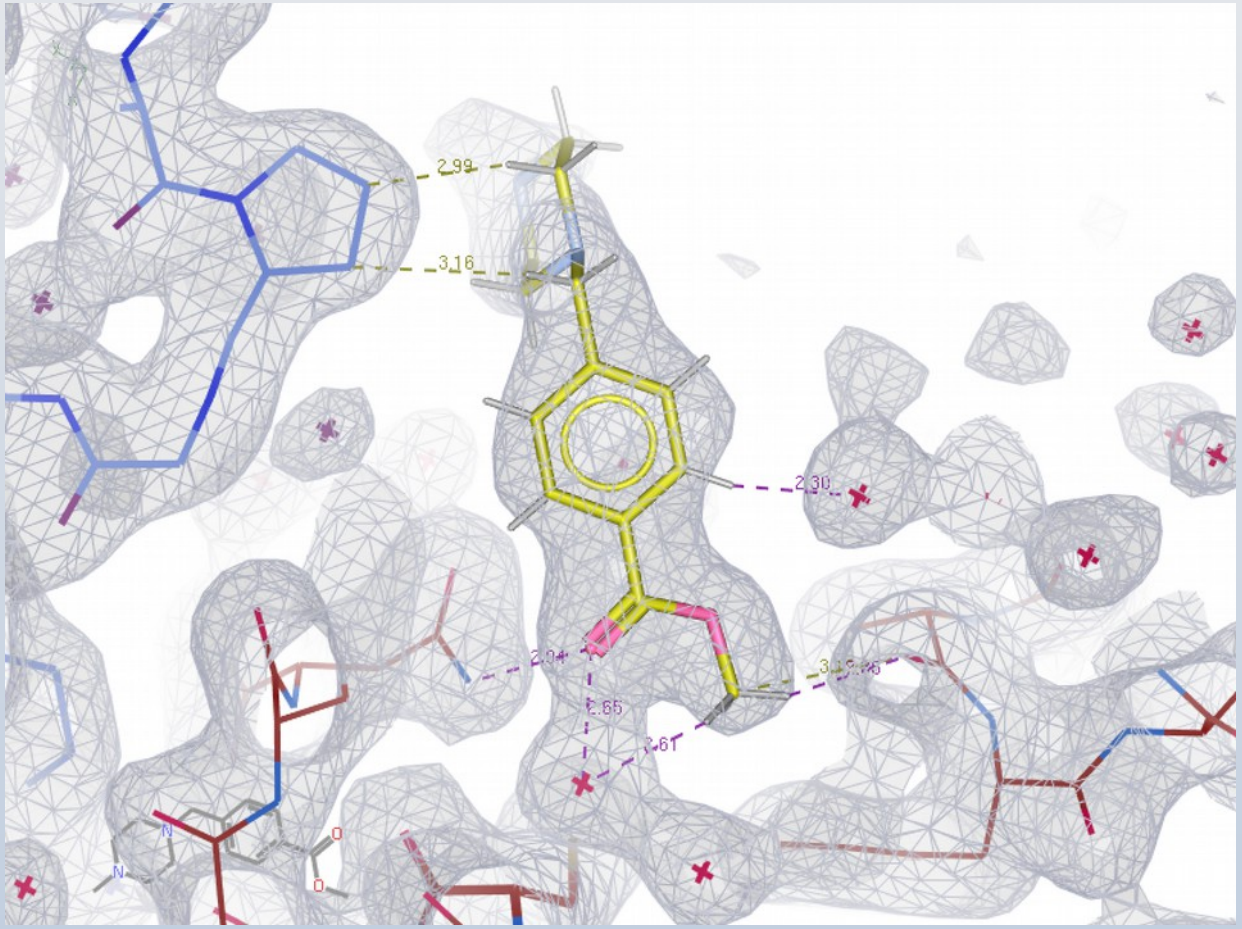
Extra Work for Atom Names

- PDBeChem
  - http://www.ebi.ac.uk/pdbe-srv/pdbechem/
  - Paste the SMILES string ⟶ JSME
  - “Search” ⟶ comp_id (2GM) ⟶ Download

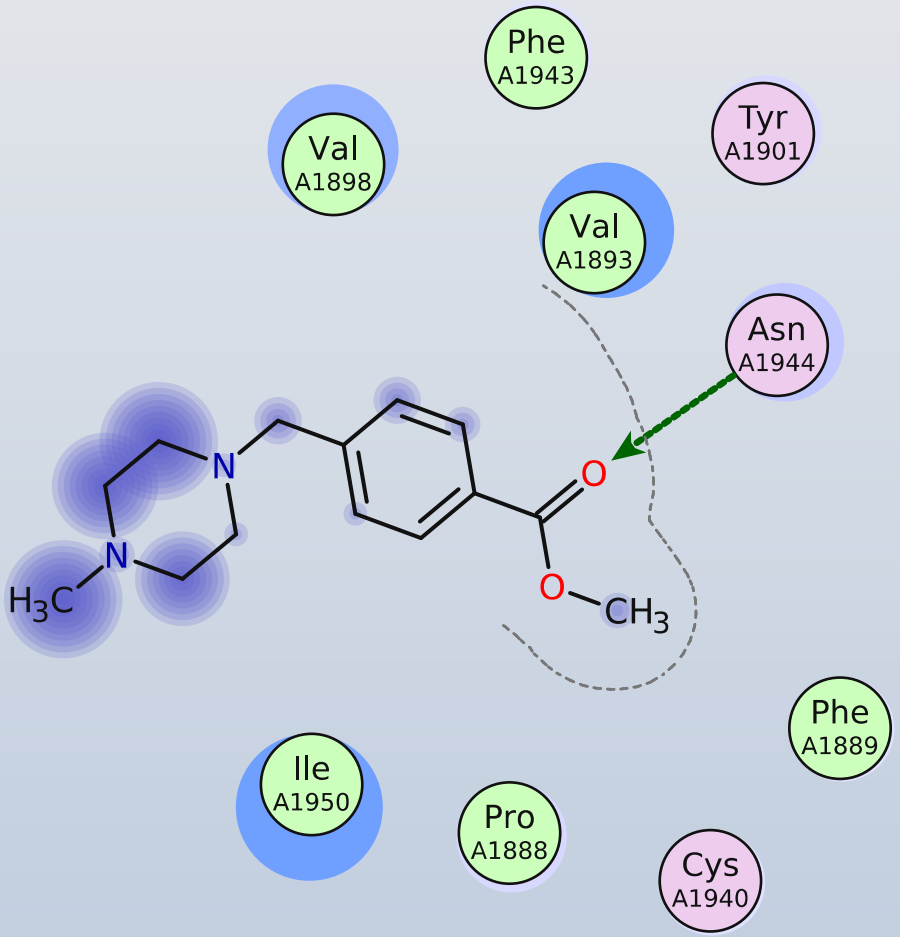


- `$ pyrogen -Mn --comp_id AAA -R 2GM.cif problem-2.smi` Graph match atom names
- `$ acedrg -m AAA-pyrogen.cif`
- This is atypical though
  - non-trivial to generate a SMILES string/MOL for a matching substructure
  - pyrogen fails to parse PDBe's `2GM.cif` :-(

Problem-1: Density Fit

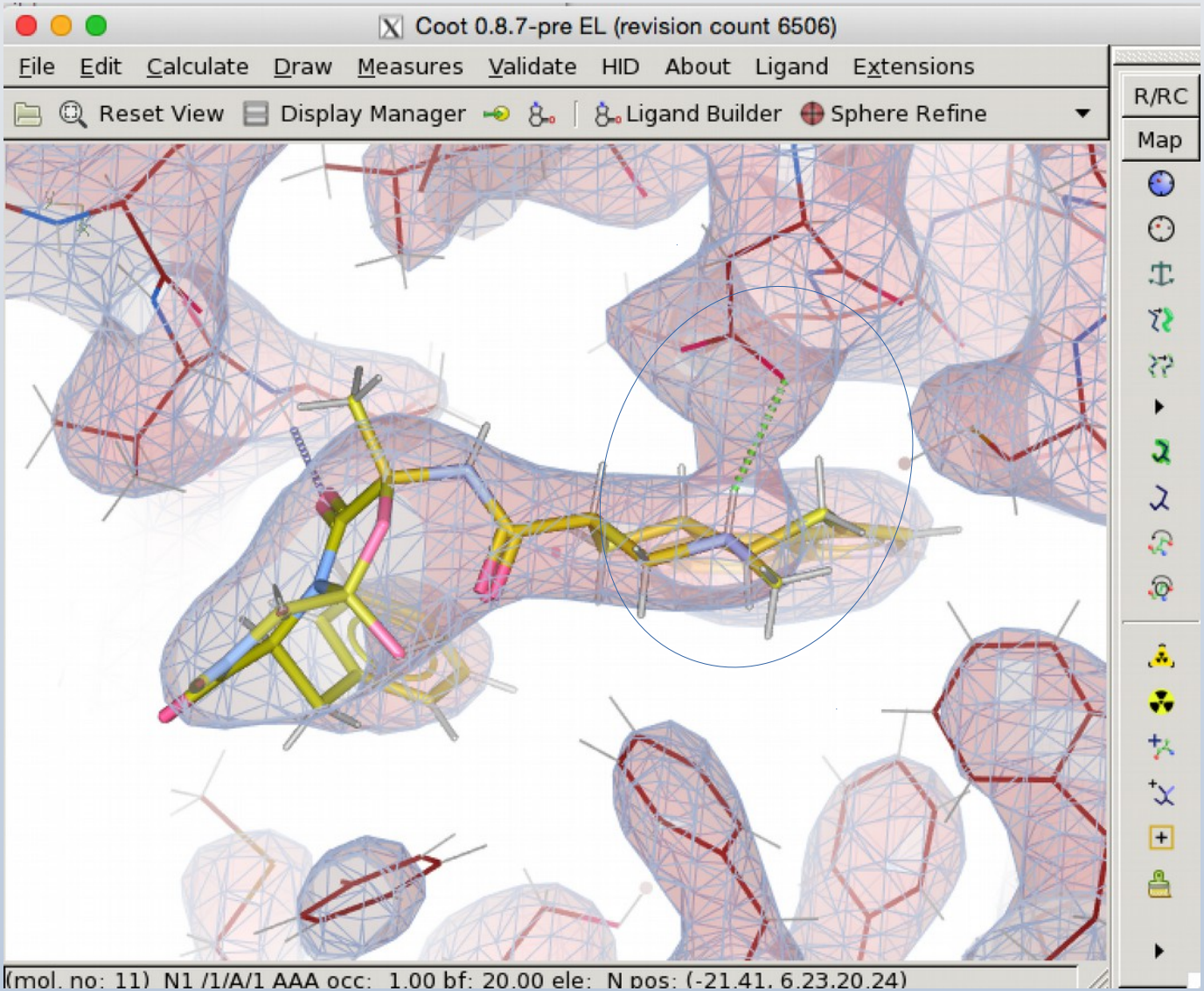


Problem-1: FLEVed

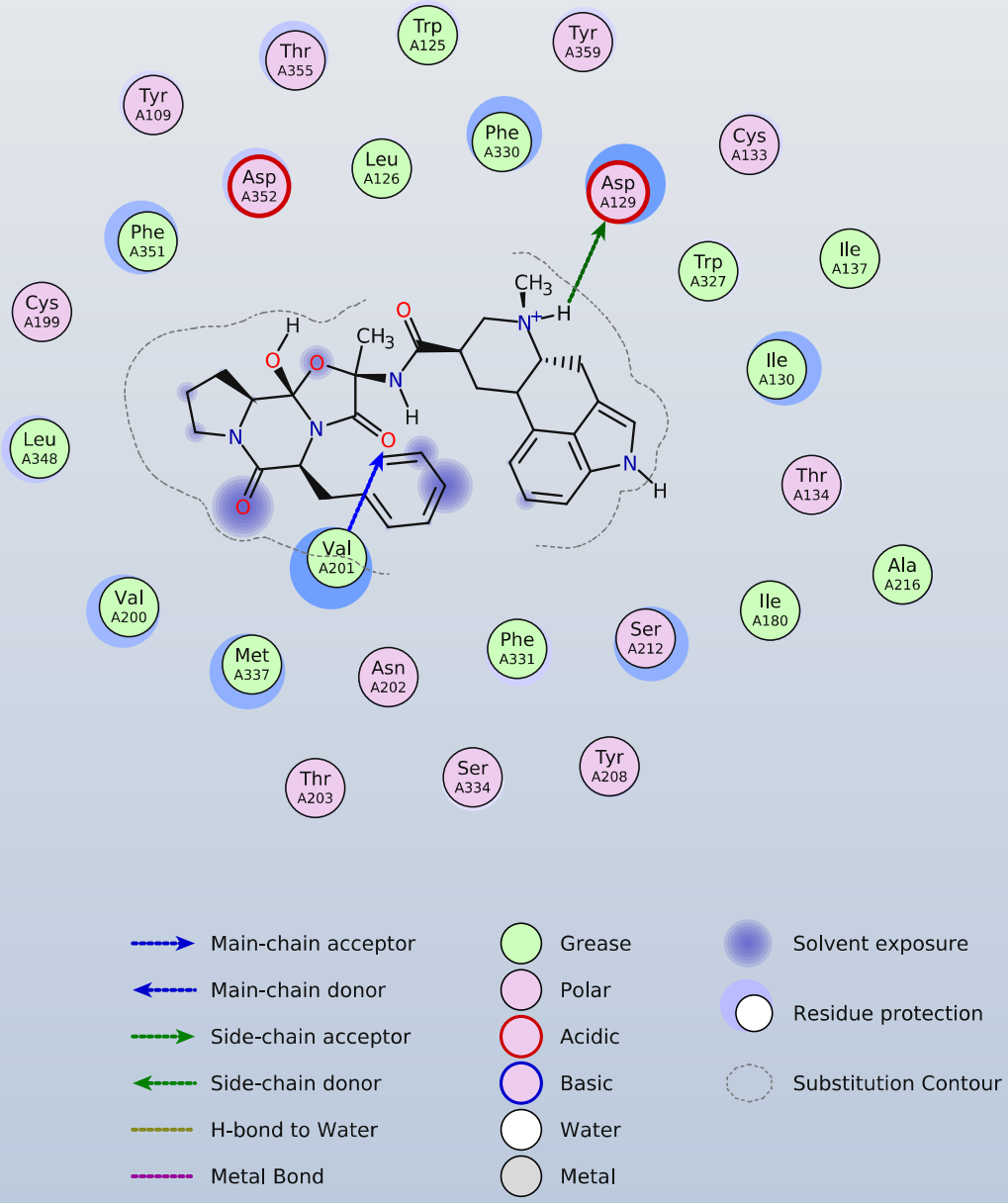


svg (png was problematic)

Problem-2: Density Fit



Problem-2 FLEVed



Ligand Fitting Trials

- Outcomes
- Action Points

(What needs to be fixed)

Acedrg-related Issues

- ...

Coot-related Issues

- “Ligand Expert” should be the default
- Ligand fitting Top-N-Solutions should have some GUI exposure
- Can you improve the ligand fitting algorithm so that 54V can be fit into 5cub without intervention?
  - JED-Flip is useful to recover a from incorrect solution
- Nice to Have: Better user-experience at conformer generation stage
  - (visual feedback – don't freeze)
- Enable SRS searching in Lidia
- Coot can't “Ligand ⟶ 2D” Acedrg output
- SMILES ⟶ 2D fails to preserve chirality consistent with Acedrg output
- Acedrg, pyrogen and prodrg are the 2D⟶3D engines behind lidia “Apply”
  - perhaps the default output type should be a mmCIF rather than a MOL file

CCP4i2-Related Issues

from NM

1) Raise awareness/visibility of the *Coot*/pyrogen route to geometry libraries and ligand conformers

2) Strip out home-spun RDKit code from our current 'Make Ligand" pipeline,

   - now that similar code is already exploited within ACEDRG for generating conformers

CCP4i2-Related Issues

from PE

- **Make Ligand**
  - It's often easier to navigate to a file containing the SMILES than paste a string
  - So give me that option
    - Nice to have: Drag & Drop
  - Get the three-letter-code/comp-id or molecule name from the text after the SMILES string in the .smi file
- **Refinement**
  - There's no point in showing a graph for R-factors if there are no points to plot (NCYCLES 0)

CCP4i2-Related Issues

from PE

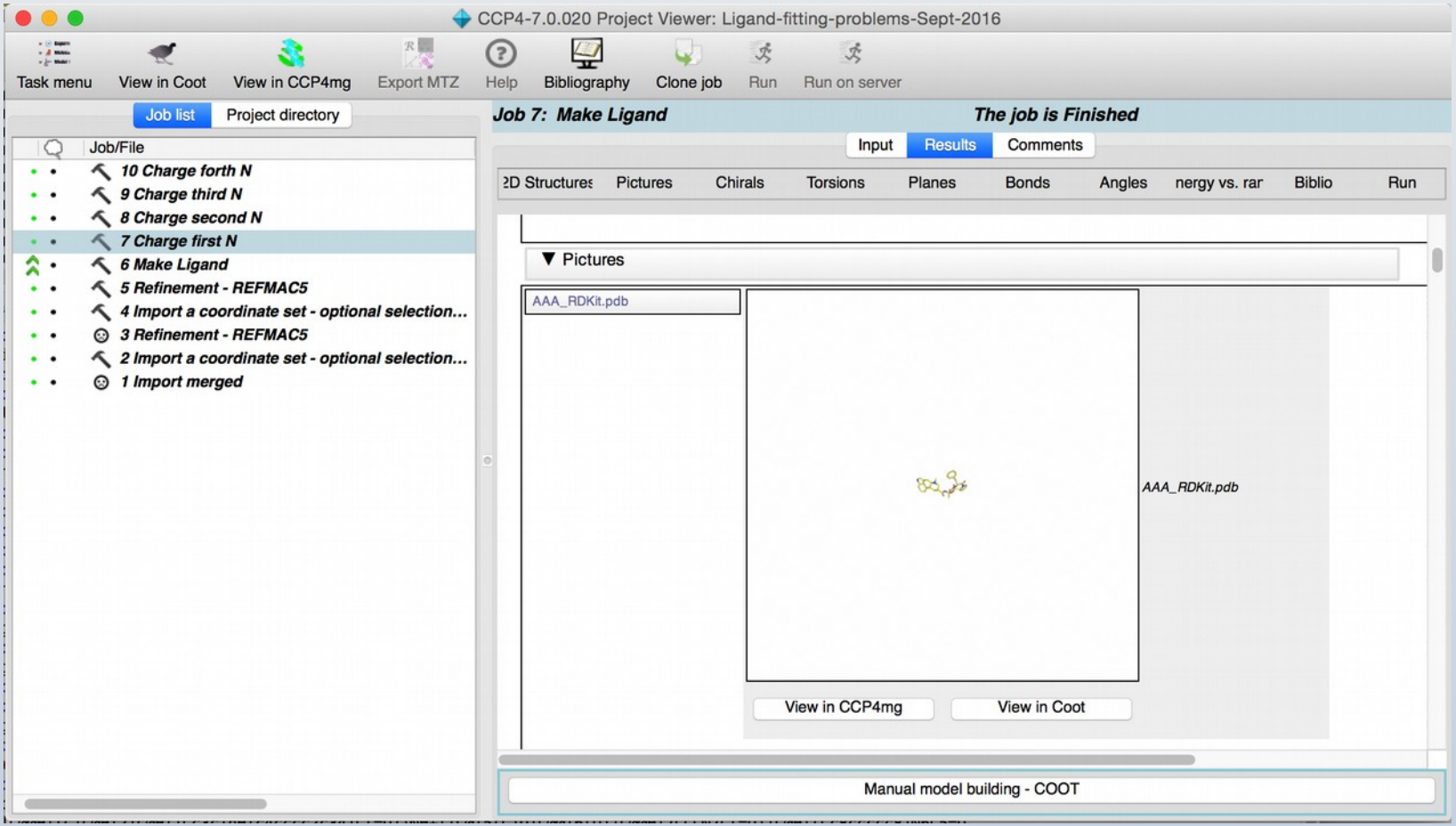
- XYZOUT.pdb contained:

```
ATOM   4958  N   THR A 355     -15.025   6.827  18.821  1.00 76.91           N
ATOM   4959  CA  THR A 355     -16.403   6.441  18.433  1.00 74.48           C
ATOM   4960  HA  THR A 355     -16.883   6.086  19.210  1.00 74.48           H
ATOM   4961  CB  THR A 355     -17.270   7.641  17.965  1.00 79.21           C
ATOM   4962  HB  THR A 355     -18.160   7.322  17.748  1.00 79.21           H
ATOM   4963  OG1 THR A 355     -16.699   8.197  16.784  1.00 86.68           O
ATOM   4964  HG1 THR A 355     -17.149   8.833  16.527  0.00 86.68           H
ATOM   4965  CG2 THR A 355     -17.451   8.726  19.019  1.00 72.64           C
ATOM   4966  HG2 THR A 355     -17.881   8.360  19.796  1.00 72.64           H
ATOM   4967  HG2 THR A 355     -17.995   9.436  18.670  1.00 72.64           H
ATOM   4968  HG2 THR A 355     -16.597   9.083  19.273  1.00 72.64           H
ATOM   4969  C   THR A 355     -16.306   5.384  17.329  1.00 76.33           C
ATOM   4970  O   THR A 355     -16.981   4.358  17.399  1.00 75.48           O
ATOM   4971  H   THR A 355     -14.644   7.757  18.719  1.00 76.91           H
```

- Coot didn't like that (hydrogen atom names)
  - exploded on refinement

CCP4i2-Related Issues

from PE



The “Pictures” image is too small to be useful