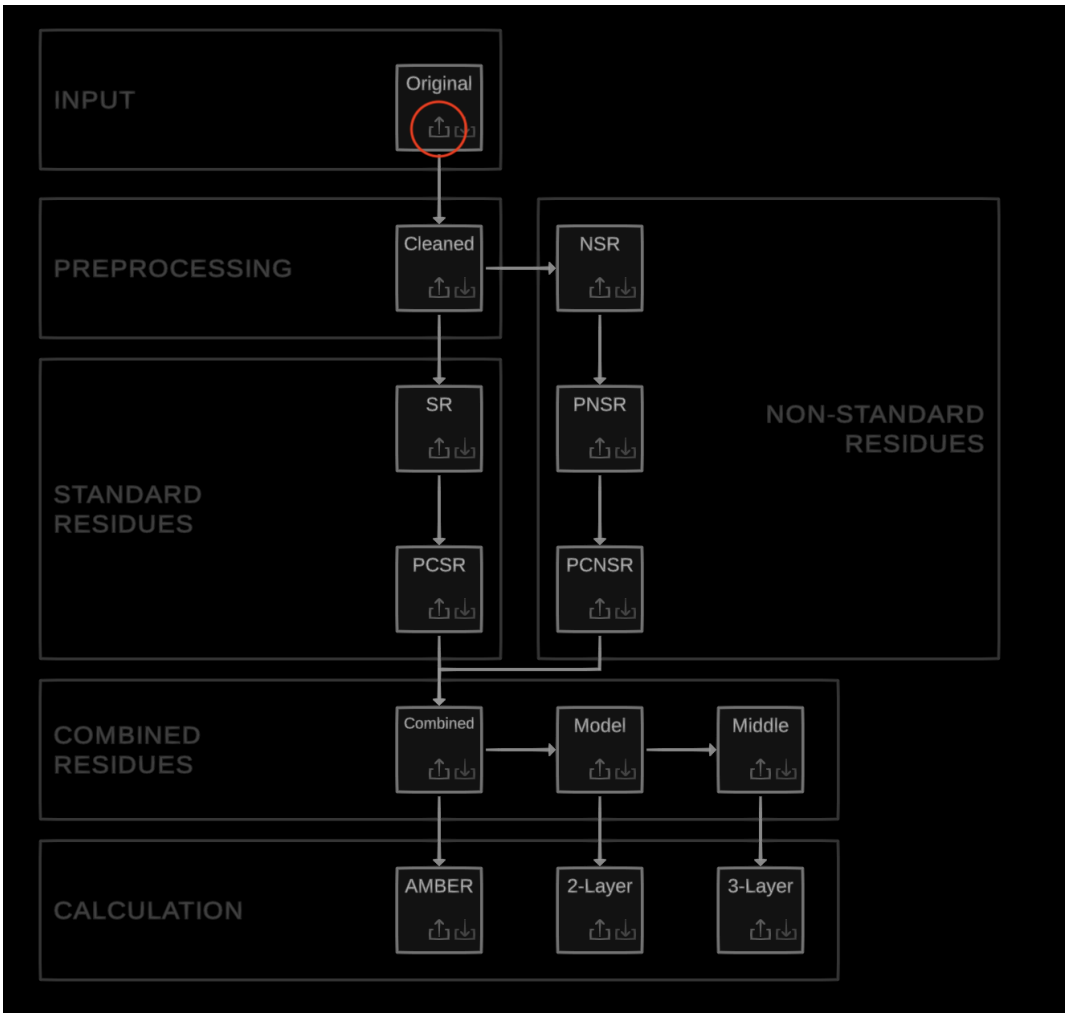
Creating a Gaussian Input File with ONIOM Flow

Once a project is created and opened, the following window will be presented. The initial PDB File may be opened by selecting the Open Geometry button circled in red.



Selecting the green arrow from the geometry box opens this dialogue:

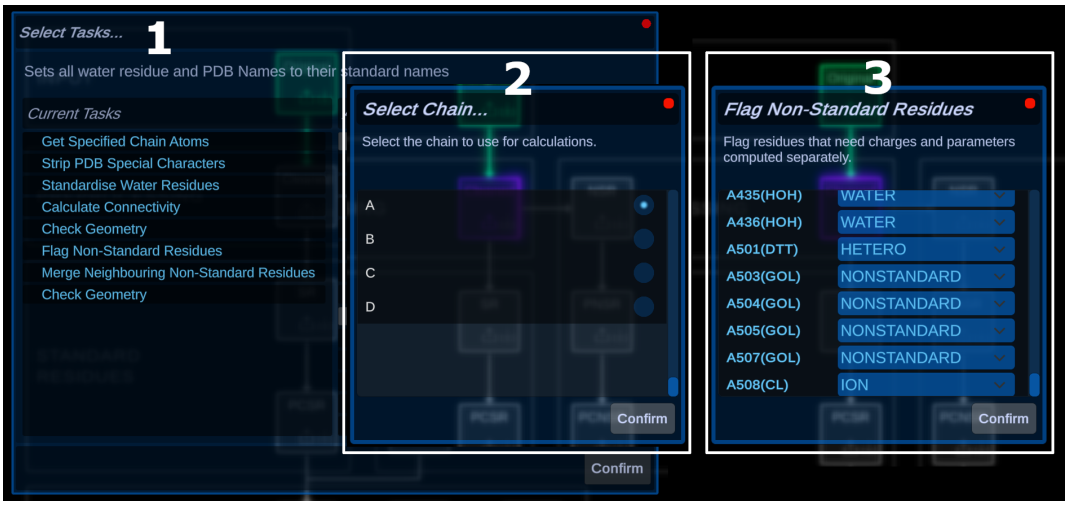
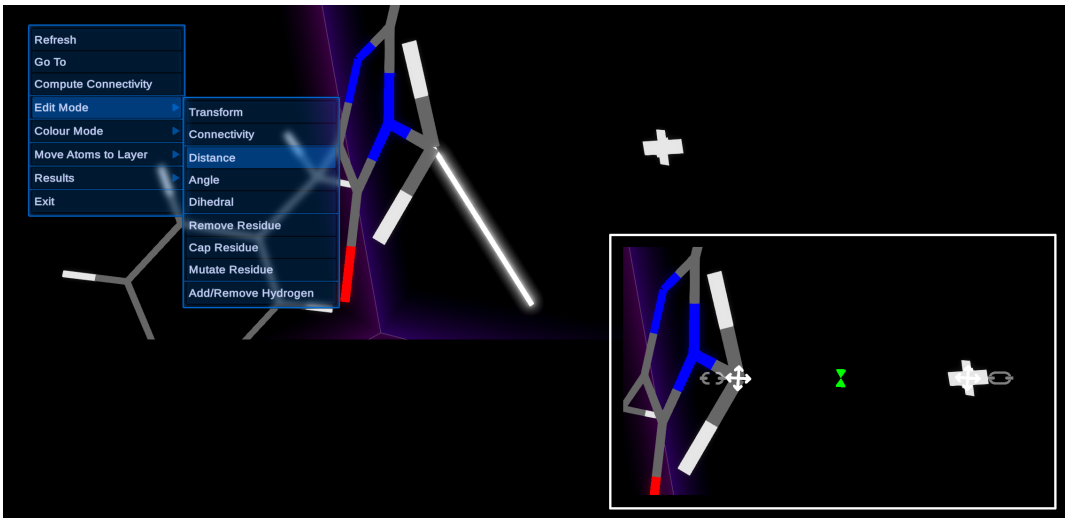


Figure 1: 1: The Select Task Dialogue showing the default tasks. 2: The Select Chains Dialogue. 3: The Flag Non-Standard Residues Dialogue.

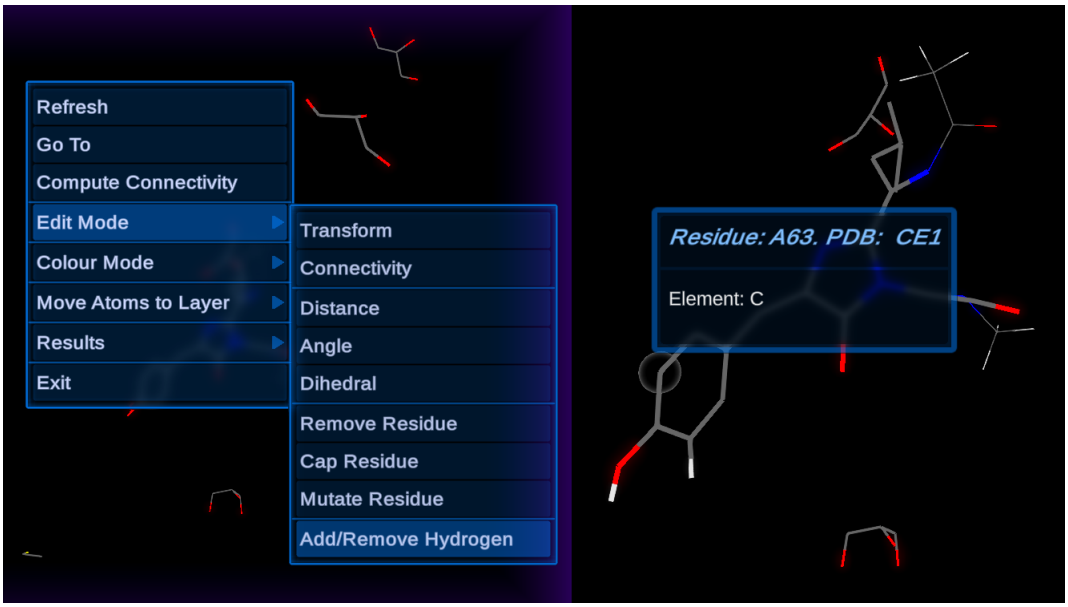
The default tasks are sufficient for this step. Since proteins are usually not crystalised as monomers and it is unnecessary to run computations on multimers, a chain needs to be isolated. Here, chain A is selected. The Strip PDB Special Characters task removes non-ASCII characters from PDB names that would otherwise crash Gaussian and other software. If duplicate PDB Names are created as a result, they are renumbered. Water residues and their oxygen atoms are renamed to be consistent with the AMBER forcefield at this point. The connectivity is then computed. Residue types are automatically detected, but it is worth going through the list to make sure nothing is left as UNKNOWN. Finally, all non-standard residues that are neighbouring and connected are joined together and renamed. The Check Geometry task at this step ensures that PDB Names are ok, but checks may be edited in the Flow.xml and Tasks.xml settings files.

Moving to the right side, the Get Non-Standard Residues task isolates and caps residues including the chromophore residues. This step is necessary since standard residues are trivially parameterised with lookups. It is important to check that caps are correctly attached. Occasionally, atom distances between the chromophore residues and neighbouring atoms are too long and are missed by the connectivity calculator. In that case, the distances can be modified with the Atoms Visualiser.

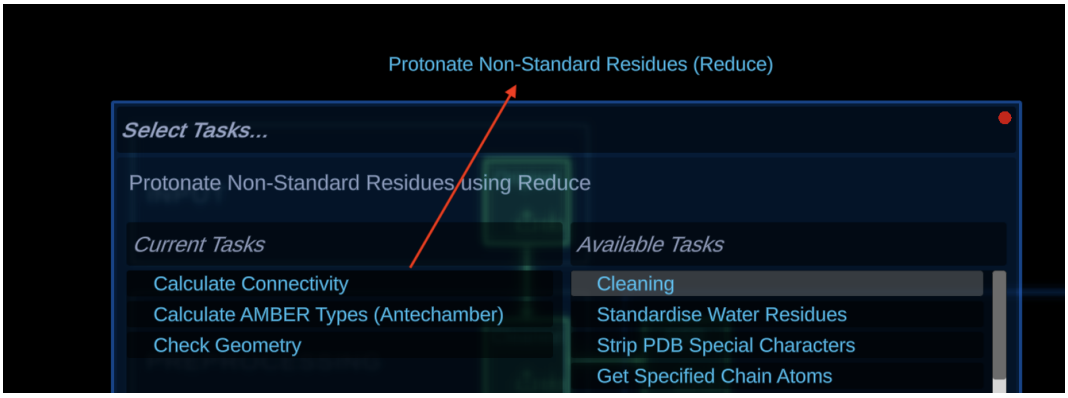


Normally, changing the distance will move both atoms by an equal amount towards or away from each other. If the pivot point (hourglass shape in between the atoms) is moved onto the heavy atom, only the hydrogen will move. The icons beside the atoms indicate how neighbouring atoms should be moved with three modes: Unlocked - only that atom moves; Locked - all atoms attached will move by the same amount; Spring - nearby atoms will move a fraction of the amount moved.

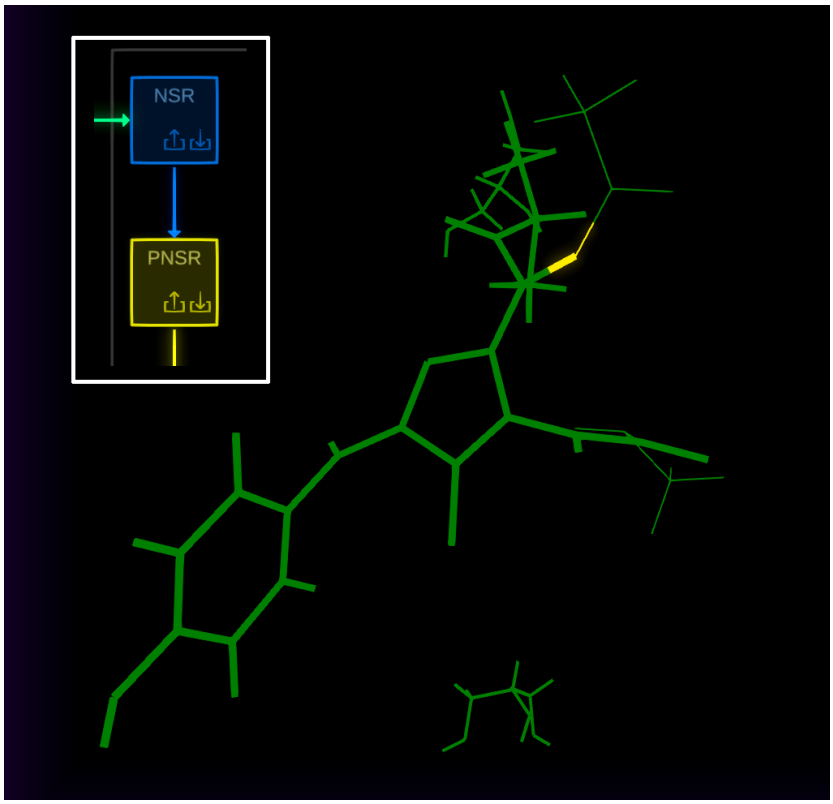
While it can be possible to automate protonation at this stage with software such as `Reduce`[1] in AMBER, we have found that it's safer to manually add protons. This is the most user time-consuming part of the process. Hydrogens are added by changing the edit mode to Add/Remove Hydrogen and clicking on heavy atoms that are missing them.



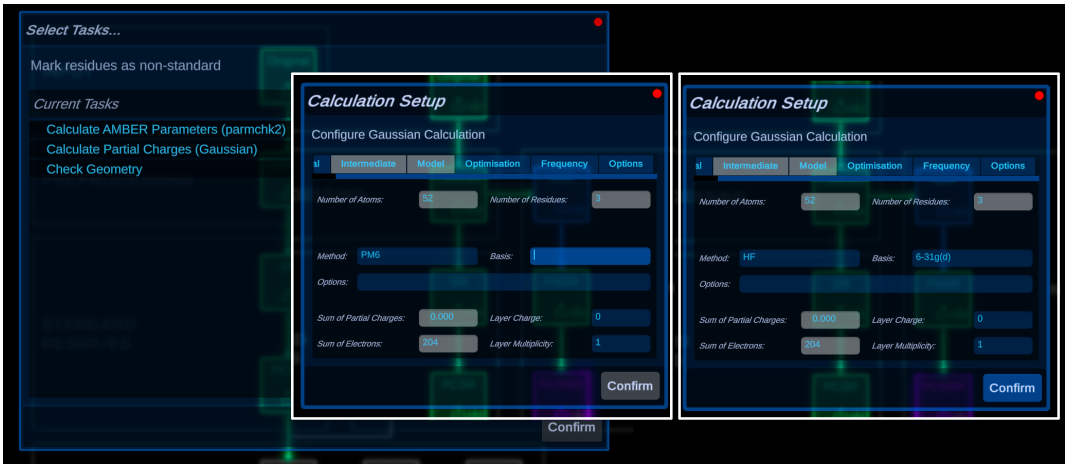
Since hydrogens were added manually, the Protonate Non-Standard Residues task should be removed by dragging it outside of the window.



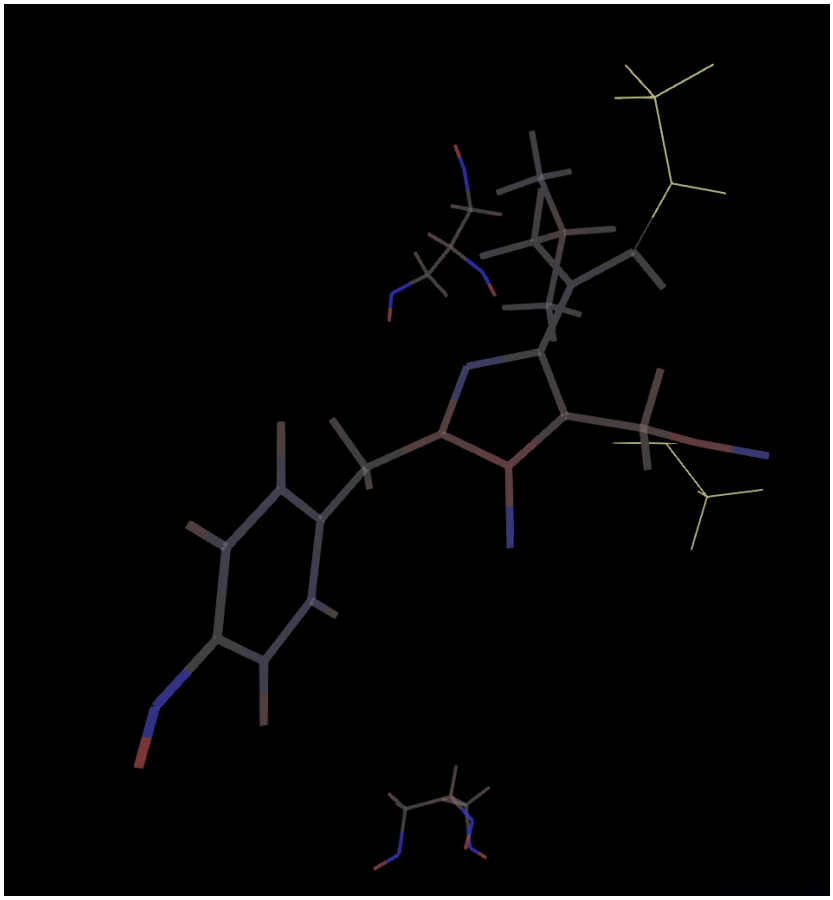
Continuing with the next task to check protons however led to warning (seen in the Flow.log file) that a Dummy atom had been read. Opening the Atoms Visualiser allows us to see what went wrong. By selecting the Has AMBER colour mode, we can see that a nitrogen atom wasn't recognised by `Antechamber`[2].



If the nitrogen is not meant to have a hydrogen, as is the case here, the problem can be resolved by adding the hydrogen and rerunning the task, then removing it before the next step. Otherwise, problems with protonation are fixed in the NSR. It is worth saving progress at this point by right clicking the background and selecting Save State. Non-standard parameters are assigned here with `parmchk2`. The next step is computing the partial charges which took 2 and a half hours running on 4 processor cores. Each non-standard residue needs to have its charge and multiplicity set, followed by a few Gaussian calculations.

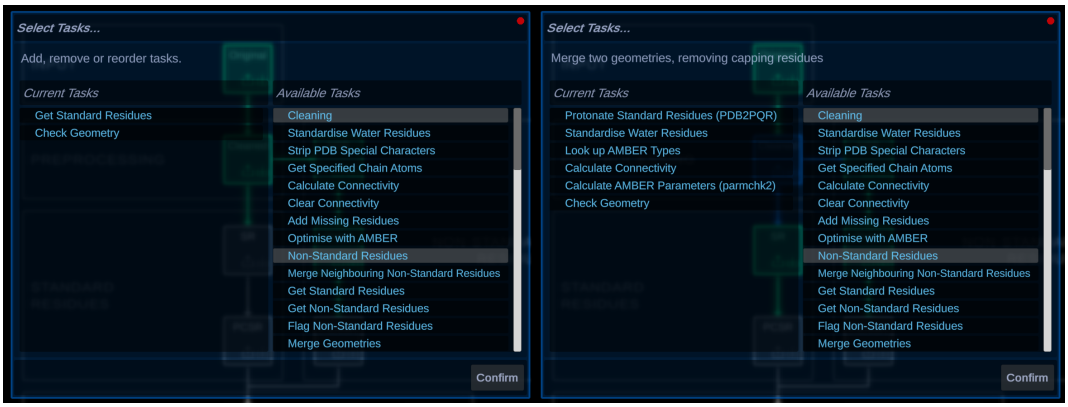


The first calculation is a loose PM6 optimisation to save time for the computationally expensive Hartree-Fock optimisation. The third is a Restricted Electrostatic Potential charge calculation from which the partial charges are assigned. Once this has finished, it is again worthwhile saving progress! The resulting geometry may be inspected to verify that the charges have been correctly assigned.



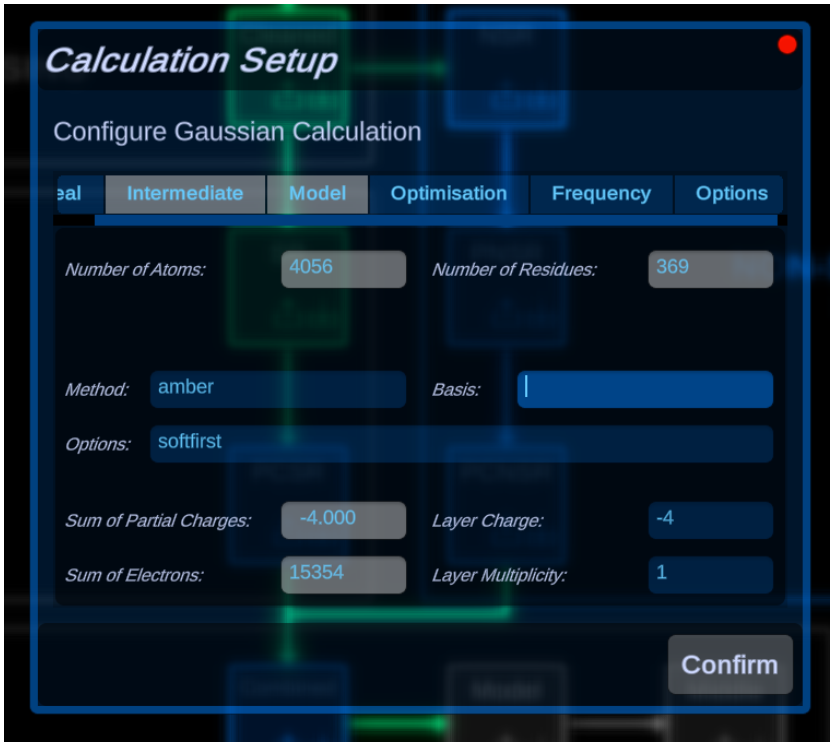
The light yellow colour of the caps indicates that no charge was assigned to them, since the software redistributes cap charges into the boundary atoms of each non-standard residue. This ensures that each residue has an integer charge.

The Standard Residue side of the flow chart is comparatively trivial.



After isolating and capping the standard residues, the software `pdb2pqr`[3] is run to assign hydrogens and their PDB names, and partial charges for all atoms. AMBER types are looked up, finishing the Standard Residue side. Charges and AMBER types may be checked with the Atoms Visualiser, but any problems should have been flagged.

Next the Standard and Non-Standard sides are combined. At this point, we have a geometry that has hydrogens, AMBER types, parameters and partial charges assigned for every residue. An optional but recommended step is an initial loose optimisation with AMBER to speed up future calculations and verify parameters. Moving down to AMBER, we can set up a calculation on the Real layer.

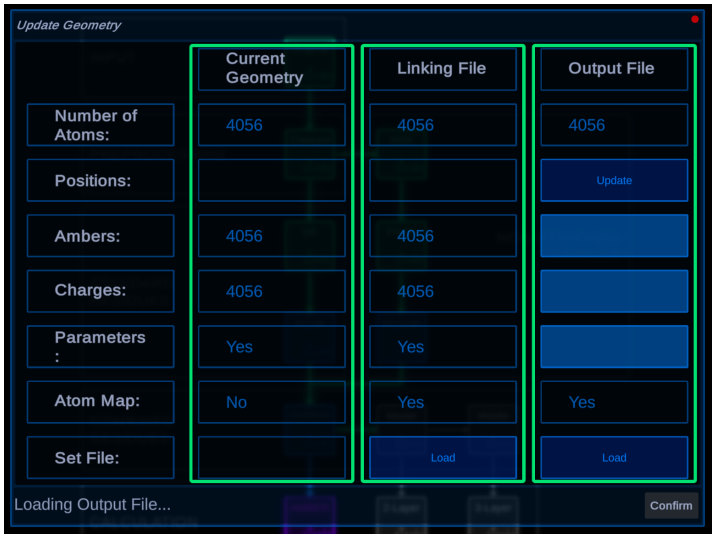


The Gaussian calculation window shows that the sum of partial charges is an integer. Confirming will not run Gaussian, but will instead prime the geometry for saving as

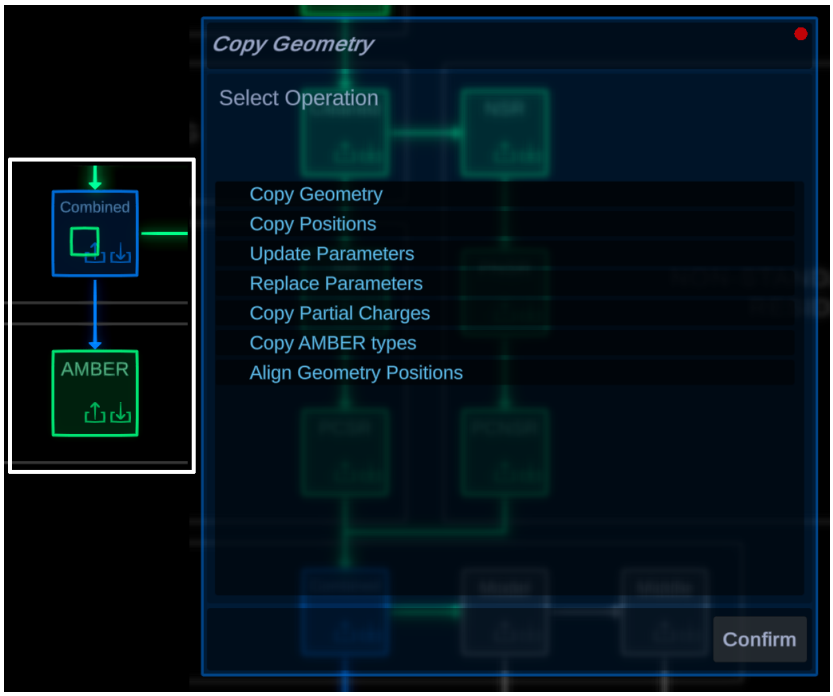
a Gaussian input. Before running, the keywords should be checked since they are sometimes removed by other steps. The AMBER=SoftFirst option is essential for these calculations.

Once the optimisation is complete, the AMBER Geometry's positions may be updated by clicking the load button. This opens a prompt to update the positions.

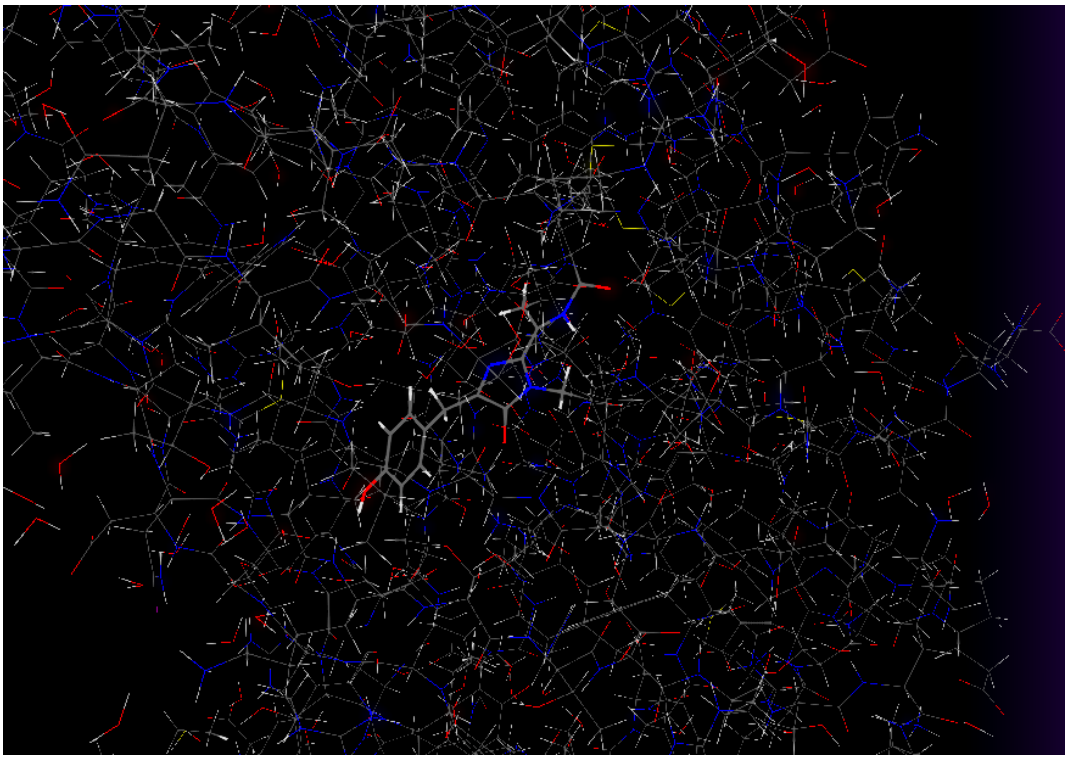
A linking file is needed to map the Gaussian atom indices to our Atom IDs. The Gaussian input file is used as the linking file. Once this is loaded, the output file is loaded and the Update Positions button becomes available. Clicking it pulls the new positions to the AMBER Geometry.



The new positions may be assigned to the Combined Geometry by dragging the AMBER Geometry box to it.



The Model region is assigned by moving uninvolved non-standard atoms to the Real Layer. This region includes the chromophore and any atom that should be computed with Quantum Mechanics (QM).



The Combined to Model arrow separates the model region and replaces dangling bonds with hydrogens at the boundary. Sometimes the new hydrogens will be placed at the wrong position.

The 2-Layer Calculation is now available, and will use the Combined Gometry with the Model parameters included. If some parameters have been lost along the way for any reason, they may be added to the calculation by dragging the relevant Geometry boxes over to the 2-Layer box and selecting Update Parameters.

References

(1) J. M. Word, S. C. Lovell, J. S. Richardson and D. C. Richardson, *Journal of Molecular Biology*, 1999, DOI: 10.1006/jmbi.1998.2401.

(2) J. Wang, W. Wang, P. a. Kollman and D. a. Case, *J. Am. Chem. Soc*, 2001.

(3) T. J. Dolinsky, J. E. Nielsen, J. A. McCammon and N. A. Baker, *Nucleic Acids Research*, 2004, DOI: 10.1093/nar/gkh381.