ISOLDE

Interactive Structure Optimisation by Local Direct Exploration

Introduction

ISOLDE, in its essence, is a toolkit for model-building and real-space refinement of moderate-to-low resolution molecular structures into experimental maps. The guiding philosophy behind its development is that the experience of building molecular models should come as close as possible to interacting with a real-world physical object. While the rules governing atomic interactions may be alien to our macroscopic day-to-day experience, codifying what we know of those rules into real-time interactive simulations brings them to life and allows us explore and adjust structures in a highly intuitive way. ISOLDE makes use of current- and next-generation graphics and computational resources to make this possible. Since molecules are inherently three-dimensional structures, interaction using truly three-dimensional input is a key priority. ISOLDE already has basic support for 3D haptic devices such as the Novint Falcon, Geomagic Touch, and other interfaces supported by the CHAI3D library. Support for more common 3D input devices such as the 3dConnexion SpaceNavigator is planned, as well as for current and upcoming immersive virtual and augmented reality systems. That being said, rich support for good old mouse-driven interaction will always be a high priority.

While sufficient functionality is now in place to make it useful for real-world tasks, please keep in mind that ISOLDE remains in very early pre-alpha status and will be constantly growing and changing over the coming months and years. Suggestions and constructive criticism are welcome, as are reports of any and all ways you may manage to break it.

System requirements

ChimeraX and ISOLDE should in theory run on most reasonably recent desktops and high-end laptops. In order to achieve decent performance, however, a recent high-end graphics card is **strongly** recommended. This is not necessary simply for graphics rendering: the OpenMM libraries responsible for handling the simulations themselves rely heavily on the massively-parallel computing capabilities of modern graphical processing units (GPUs) for optimum performance. If you have a compatible GPU, it should be detected automatically. Note that professional-level (e.g. Quadro or Tesla) cards are generally not required – gaming-level cards are more than adequate and far more cost-effective for simulation purposes.

ISOLDE is implemented as a plugin to ChimeraX and for the most part conforms to its operating system requirements. Haptic device support has so far only been tested in Linux.

Linux: ChimeraX nightly builds are compiled in Debian 8. It runs without problems in Fedora versions 21 onwards. The primary challenge with other Linux distributions will likely be GCC (or, more specifically, the standard C++ libraries contained therein). ChimeraX is compiled using GCC

4.9, and will fail with something similar to the following error on systems with earlier versions:

ImportError: /lib64/libstdc++.so.6: version `CXXABI_1.3.8' not found

I can confirm that in CentOS 7 (and, most likely, other operating systems), this can be resolved by downloading and compiling GCC 4.9.2 (**warning:** this is not provided in the repositories for a good reason. Installing it as root to overwrite your old version will likely break your operating system beyond repair. Install it to your home directory or /opt only) and placing the following bash script on your execution path:

```
#!/bin/bash

CHIMERA_HOME=/path/to/chimerax

export LD_LIBRARY_PATH=/path/to/gcc-4.9.2/lib64

$CHIMERA_HOME/bin/ChimeraX $*
```

MacOS: While I have not yet had the opportunity to test it personally, I am assured by the ChimeraX developers that both ChimeraX and ISOLDE run well in recent MacOS versions.

Windows: ChimeraX itself fully supports Windows. It is planned that ISOLDE will ultimately run as easily here as in any other OS.

Installation

ISOLDE is implemented as a plugin to ChimeraX, so your main challenge is downloading and running it. If you have an OpenCL-compatible Nvidia or Radeon GPU make sure you have the latest manufacturers' drivers installed. Note that ChimeraX itself is in early pre-release and changing rapidly, and that any given nightly build may contain bugs. It can be obtained from http://preview.cgl.ucsf.edu/chimerax/docs/download/chimerax.html. In Linux, simply untar to the directory of your choice, and run /path/to/installation/dir/chimerax/bin/ChimeraX. See the notes under System requirements if it fails to start. For other operating systems, simply run the startup scripts.

Once ChimeraX is running, ISOLDE can be obtained via the Toolshed (Tools/General/Toolshed) (Figure 1). After downloading, ISOLDE should be automatically installed to the menu under Tools/General/ISOLDE.

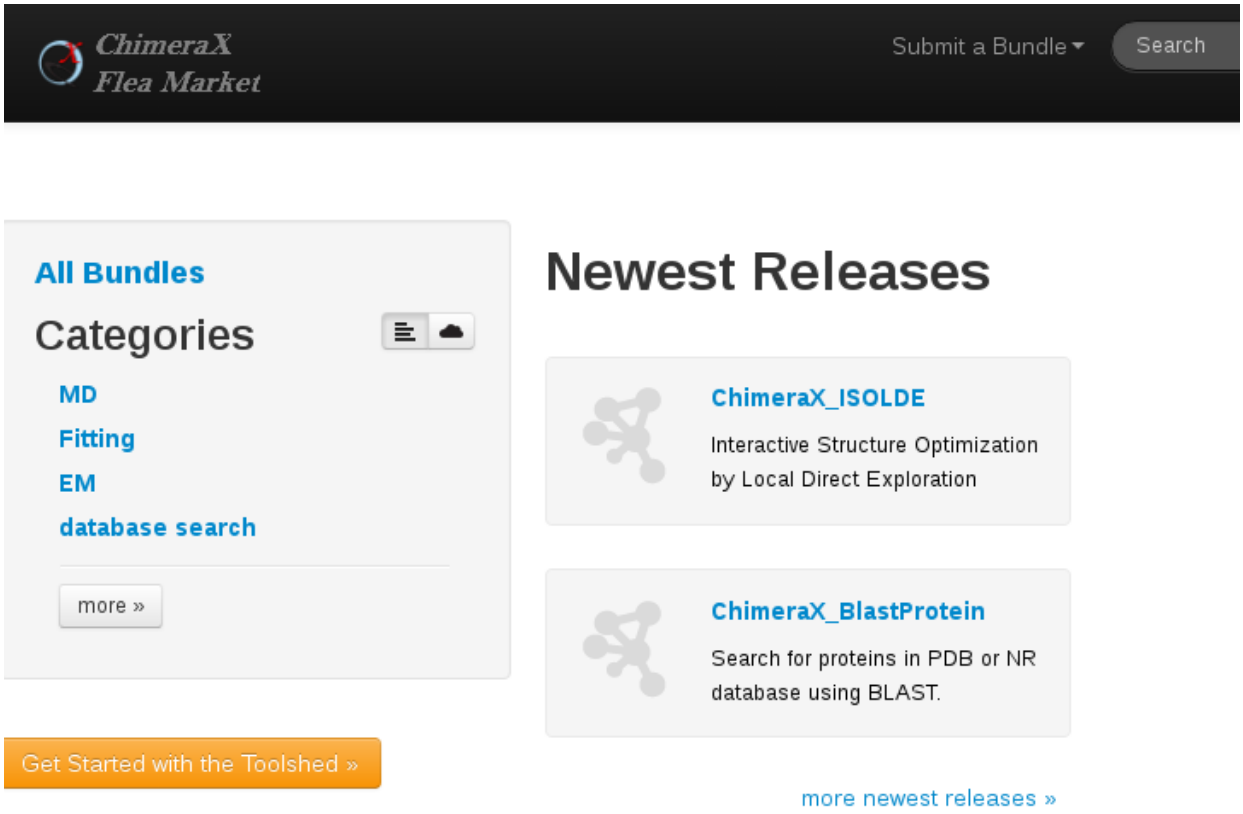

Figure 1: The ChimeraX Toolshed (aka Flea Market) for downloading plugins

Using ISOLDE

Structure requirements

In order to work with ISOLDE, you will need a PDB or mmCIF with every residue complete with all atoms including hydrogens. All non-standard bonds including disulfides must be defined. At chain breaks, it is allowable to have incomplete N- and C-termini, as long as every other atom in the terminal residue is present (Figure 2). Ultimately I plan to provide for automated addition of hydrogens within the ISOLDE environment, but for now you will need to do this elsewhere. Note that if using phenix.reduce to add hydrogens you will need to add the N-terminal H atom(s) manually, either using your favourite GUI editor or by directly editing the coordinate file. Don't worry too much about the geometry at this stage – as long as the distance to the nitrogen isn't too extreme it will be corrected near-instantly on starting a simulation.

The heavy computational work of running the simulations themselves is handled by OpenMM, which by default currently supports the AMBER force field. This means at present simulations are limited to those residues supported by AMBER: protein, water, and some monoatomic ions only. The OpenMM developers are actively working to add support for the far more extensive CHARMM36 force field, which will add support for nucleic acids (including all naturally-occurring modified bases), various post-translational modifications, most small ions, etc. Ultimately I envision that ISOLDE will also include a streamlined interface for parameterisation of ligands and unusual residues and, eventually, hybrid QM/MM simulations for handling of particularly unusual/interesting sites.

Please note that altlocs are currently not supported (although it is in principle possible to support them in this environment). This will be explored once all the more critical functionality is in place.

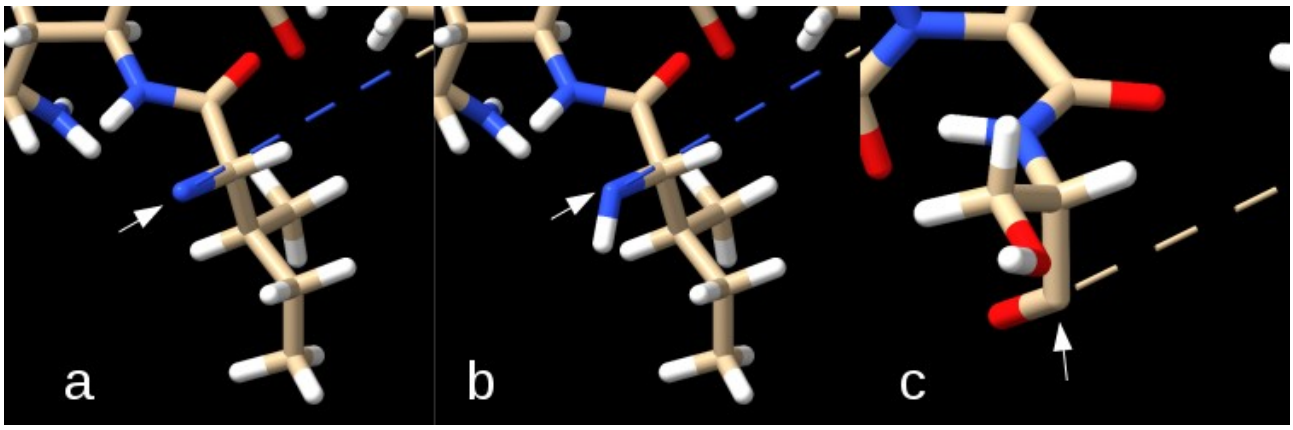

Figure 2: Handling of chain breaks in ISOLDE. The N-terminal residue shown in (a) would be disallowed, since it is missing the backbone H. The residues shown in (b) and (c) are allowed. In each frame the white arrow points to the terminal heavy atom, and the dashed line leads to the next residue present in the chain.

Quick overview of ChimeraX

An annotated overview of the ChimeraX main GUI is shown in Figure 3. The left toolbar sets the behaviour of the right mouse button. While these are worthwhile exploring, they are not used by ISOLDE and should in fact be avoided entirely while a simulation is running since they will override ISOLDE's use of the right mouse button for dragging atoms. The structure display controls are (in pairs from left): show/hide atoms; show/hide ribbons; show/hide surface, followed by the atom representation types (stick/VDW/ball-and-stick) and colour options (by element/by chain/random). The general graphics controls include buttons for setting background colour and lighting quality, resetting the display extent and orientation, and taking snapshots or movies. The density map toolbar contains tools for adjusting the display of volumetric data. At present, functions available through these toolbars and the menus are relatively limited, and functions entered via the command line are far more powerful. There is also a fully-featured Python shell for more complex tasks. Further documentation for ChimeraX may be found at http://preview.cgl.ucsf.edu/chimerax/docs/.

Mouse interactions

Select atom(s): Ctrl-click (and drag)

Add to selection: Ctrl-shift-click (and drag)

Rotate: left click-and-drag

Translate: middle click-and-drag

Zoom: scroll wheel

Set centre of rotation: Shift-right-click on an atom

The right mouse button performs various tasks depending on which of the icons on the mouse mode

toolbar has been chosen, and is co-opted as “tug this atom” during ISOLDE simulations.

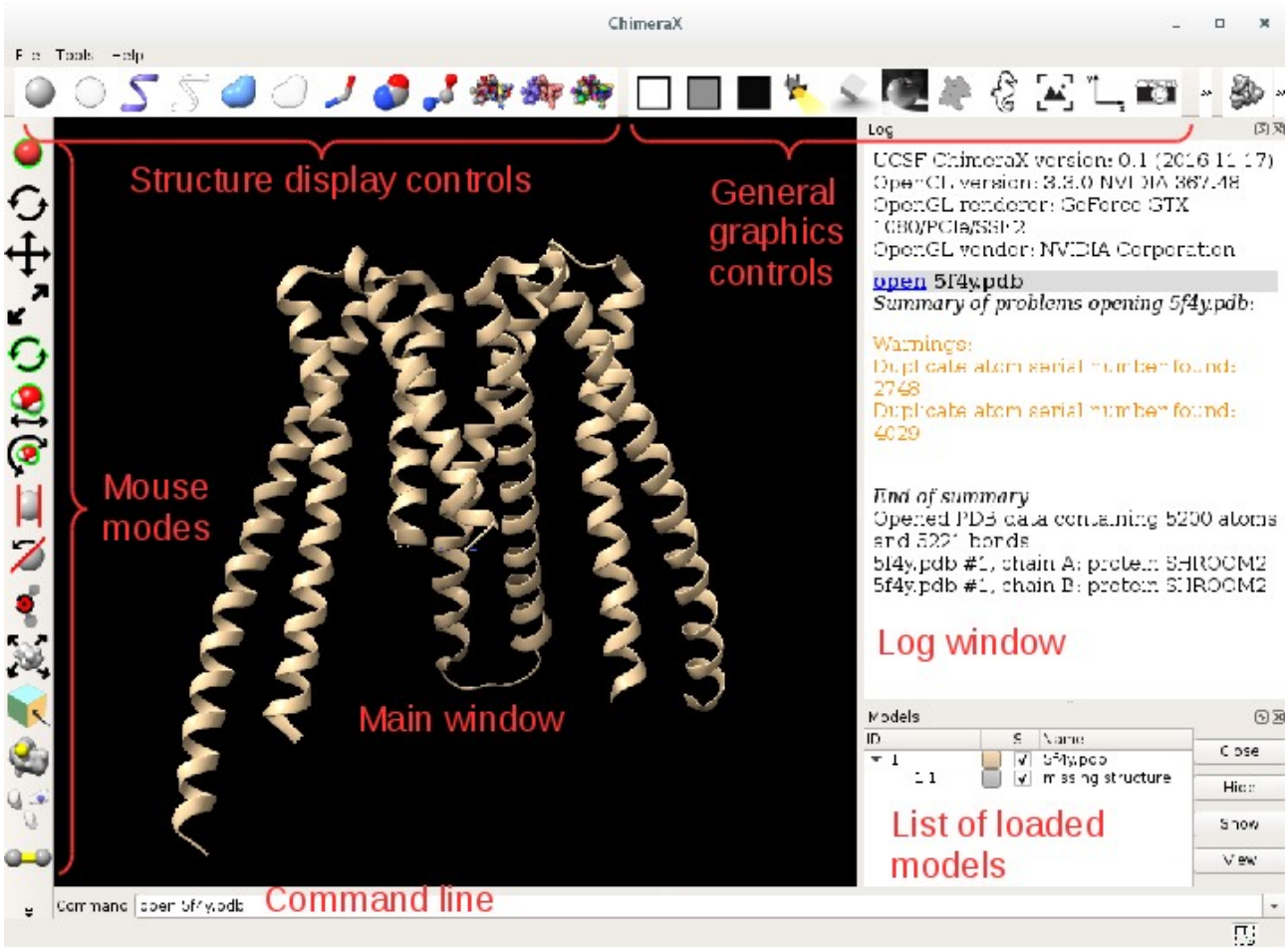

Figure 3: Annotated overview of the main ChimeraX GUI.

Before you start

Start ChimeraX and load your structure and any associated maps. At present only real-space maps are supported, and they must cover the whole structure (preferably with 10-15Å of padding if they are crystallographic maps). Support for periodic maps and calculation directly from structure factors is planned but not yet implemented. **Note: for crystallographic work it is extremely important that the maps are generated with the free reflections omitted**. Don't worry about adjusting their visualisation and contours (although you may do so if you wish using the density and mouse mode toolbars) – basic handling of this is available in the ISOLDE GUI.

Setting up the display

While in my experience inclusion of a cartoon representation can significantly aid interpretation in a live interactive simulation, at present the cartoon implementation in ChimeraX is not optimised and significantly slows simulations and is best turned off. Switching to an all-atom stick representation can be done by clicking the first and fourth icons in the structure display controls toolbar. To reduce visual clutter I recommend hiding all non-polar hydrogens, which can be achieved by entering:

```
hide HC
```

in the command line. Alternatively, the same effect can be achieved entirely through the command line by entering:

```
hide cartoon
```

```
show ~HC
```

I also recommend switching from perspective to orthographic view by entering:

```
camera ortho
```

at the command line. The command:

```
camera mono
```

will switch it back if needed.

A simple ISOLDE simulation

After starting ISOLDE (Tools/General/ISOLDE) you should see a new window similar to that shown in Figure 4. At its simplest, starting an interactive simulation can be a case of selecting some atoms in the main ChimeraX window, switching ISOLDE from “Crystallography mode” to “Free mode (no maps)”, and clicking “Go” (Figure 4). After a second or two the simulation will start. You will note that the selection of mobile atoms has grown somewhat compared to your initial selection. In general this is preferable – for a start, the simulation requires whole residues. Additionally, one must keep in mind that small simulations (particularly those starting from coordinates with significant residual errors) tend to be unstable – there must be sufficient room for large stresses to relax. Finally, direct contact between a fixed atom and one experiencing large interactive tugging forces often causes undesirable effects. Therefore, in almost all cases it is best to add an extra shell of mobile residues to act as a soft “buffer zone” around your region of interest. For similar reasons to the above, if you are loading for the first time a structure in which you expect large errors (in particular atomic clashes), it is advisable to first run a brief simulation of the entire structure to allow these to settle prior to commencing interactive work.

Once the simulation has started, ctrl-click somewhere in empty space if you wish to remove the green “selected atoms” highlighting. You can tug any mobile atom with the right mouse button. Peptide bonds are restrained and unable to flip by default (any residue within 30 degrees of *cis* is restrained to *cis*, with all others restrained to *trans*). It is desirable to leave these restraints in place for almost all interactive simulations, since without them it is remarkably easy to accidentally introduce a flip. You may turn any or all peptide bond restraints on or off via the “Custom restraints” tab. Turning off peptide bond restraints may be desirable when settling the whole structure non-interactively prior to saving the coordinates for refinement. You can flip a peptide bond from *cis* to *trans* or *vice versa* on the “Rebuild” tab (while a simulation is running) by first selecting at least one atom from the associated residue then clicking “flip”. Note that the remainder of the buttons on this tab (associated with rotamers) are currently non-functional.

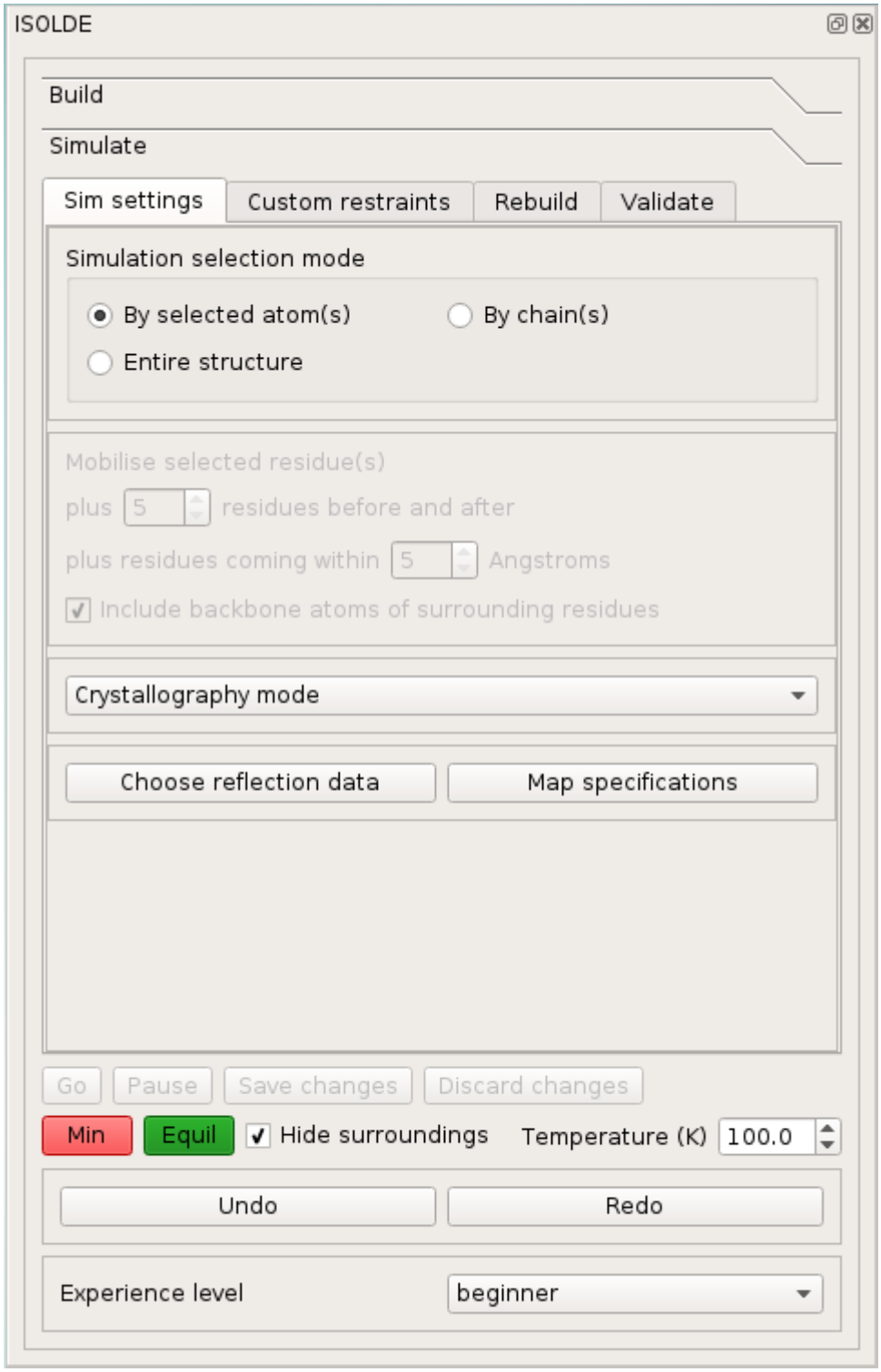

Figure 4: The ISOLDE window on startup. The "Build" page is currently an empty stub for future use. The "Sim settings" tab contains all the settings necessary to define how a simulation will run. Three main modes are envisioned: Crystallography, EM and Free (no maps). At present only the latter two are implemented, since crystallography requires extensive extra functionality for handling of periodic symmetry, reciprocal space maps, etc. The other three tabs will handle: restraints (peptide bond, secondary structure, NCS etc.); script-driven rearrangements (pep-flips, rotamers, register shifts, loop modelling etc.); and validation (Ramachandran, rotamers, cross-correlation, etc.).

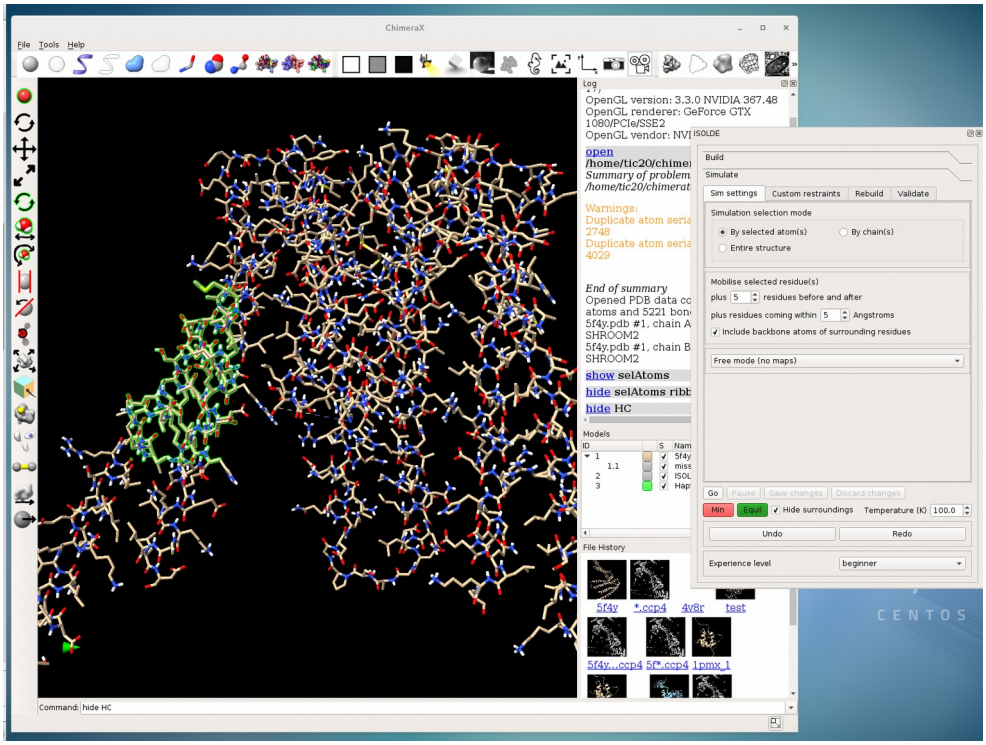

Figure 5: Starting a simple interactive simulation. The user-selected atoms (highlighted in green) will be used to define a core which will be mobilised along with a "soft shell" of surrounding residues, with a further "hard shell" of residues held fixed to maintain the context of the simulation. Note that the total number of atoms involved in the simulation dictates the speed at which it will be able to run.

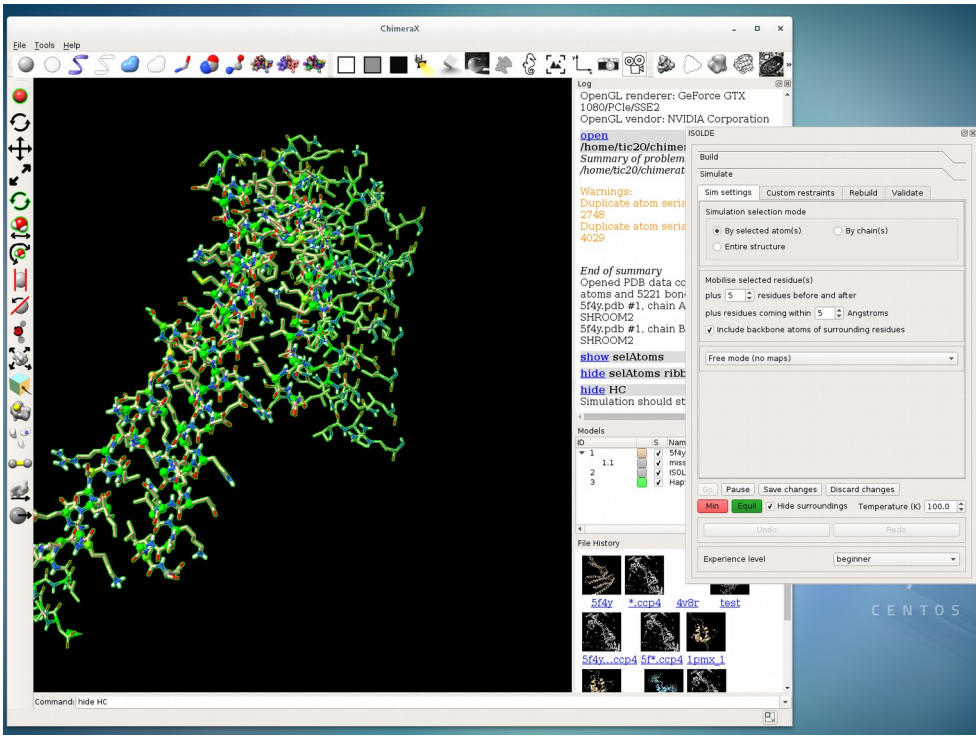

Figure 6: Beginning a simple interactive simulation. All residues not involved in the simulation are automatically hidden, and will be re-displayed on completion. Fixed atoms are shown in a slightly thinner stick representation. Any mobile (and visible) atom may be tugged by pulling on it with the right mouse button (or with a haptic device if one is connected).

Another key design goal in ISOLDE is to provide validation metrics as real-time visual feedback within the simulation itself wherever possible. As the first example of this I have implemented a real-time visualisation of Ramachandran scores (Figure 7). By default these are calculated once every ten simulation steps (approximately once per second for a moderately-sized simulation).

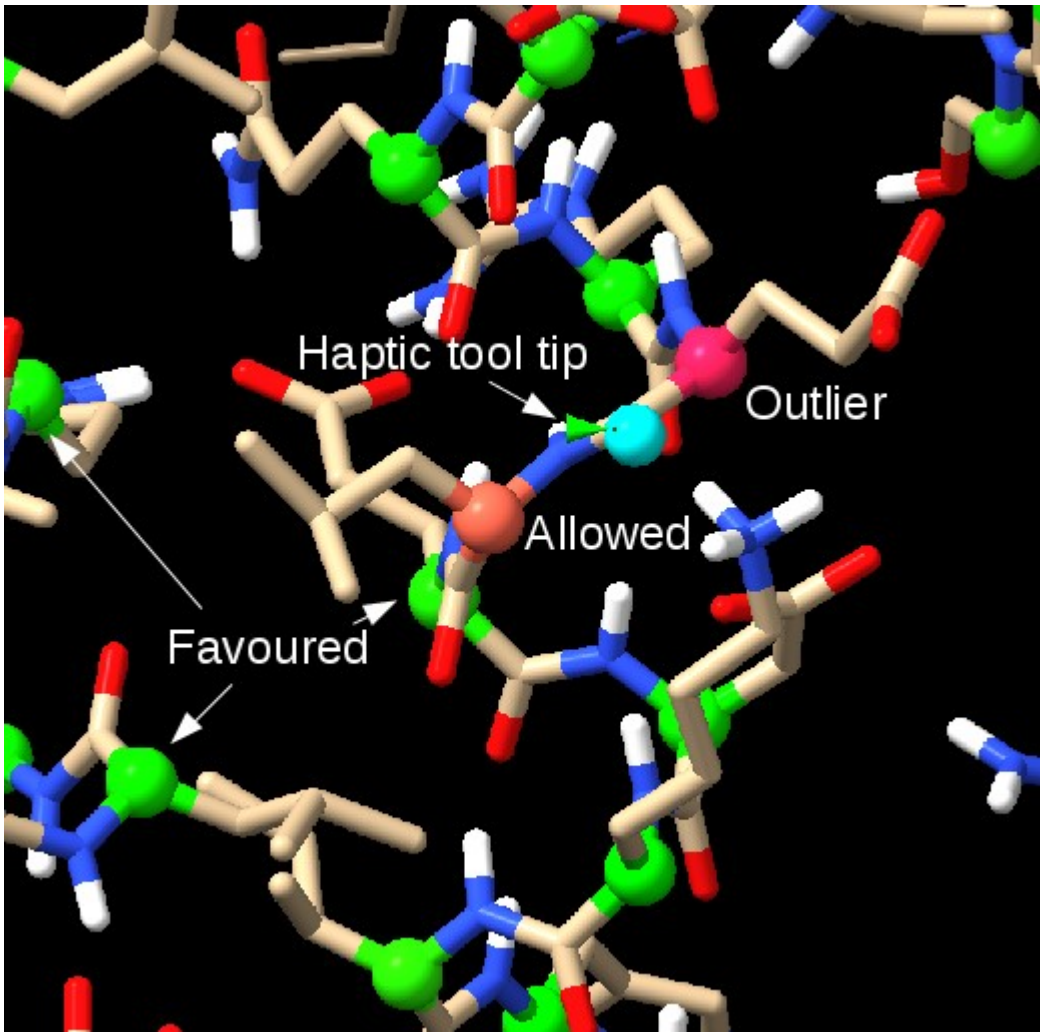

Figure 7: Real-time visualisation of Ramachandran status of protein residues. C-alpha atoms are changed to a sphere representation, and coloured according to the log of the residue's Ramachandran probability. Colours scale from green (most favoured) through yellow (2% probability) to magenta (outlier cutoff and below). Also shown is the standard representation of the haptic device tool tip (green cone) and the atom it is tugging (cyan sphere).

Of course, it is still necessary to provide more traditional Ramachandran plots (and similar tools for identifying all outliers in a structure). Such tools will be provided on the “Validate” tab, where at present you will find a simple Ramachandran plot (Figure 8). During a simulation, this plot will update live, displaying only the mobile residues. Outside of simulations, a plot can be quickly generated for a whole model or any partial selection thereof. Clicking on a plot point will select and focus the display on that residue.

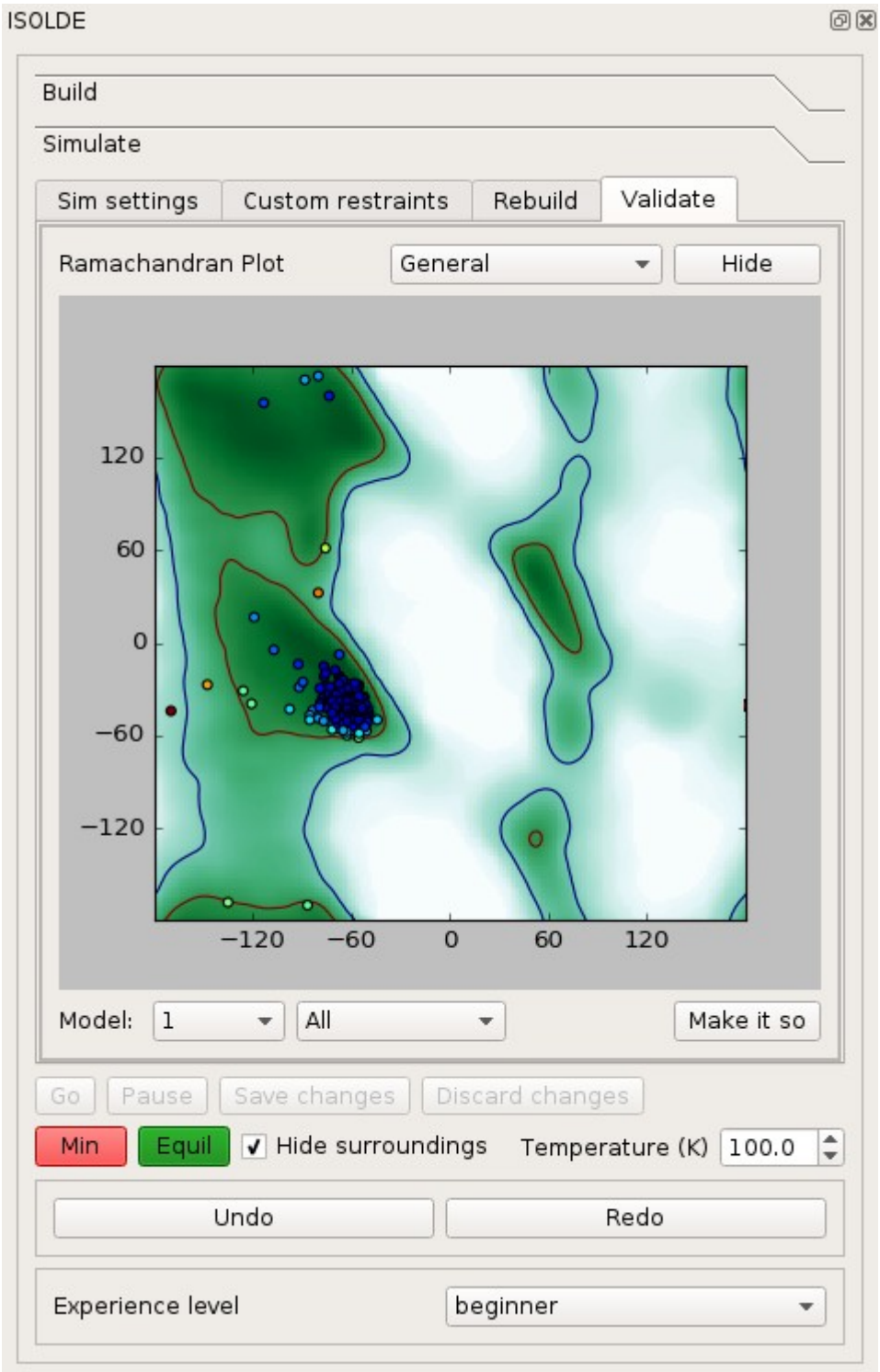

Figure 8: Ramachandran plot implementation. Contours shown as lines for the standard MolProbity outlier/allowed cutoffs, and as colour maps to give an extended picture of the very low probability (but possible) regions. Points are additionally coloured according to score. Clicking on a point will select and centre the display on the associated residue, so that if the simulation selection mode is "By selected atom(s)" clicking Go will immediately start a simulation suitable for investigation/correction.

Simulations with maps

Of course, the primary purpose of ISOLDE is to work on fitting models into maps, so let's load some up. Maps in most real-space formats can be loaded using the File/Open dialog or by entering:

```
open <filename>
```

or even:

```
open *.<extension>
```

on the command line. On the ISOLDE “Sim settings” tab, switch from “Free mode” to “Single-particle EM mode” and click “Choose map(s)”. In the frame that appears, choose the map you want from the “Model” dialog box, and give it a descriptive name (this must be unique for each map). The map will be masked to within the radius defined by the “Cutoff” parameter for simulation purposes – unless you expect to be making very large rearrangements, the default 3.0Å is usually adequate here. The coupling constant defines how strongly the map will pull on each atom (more correctly, it is a constant scaling factor on the potential energy field that the map is converted into), and for most purposes should be kept below 1.0. Very high coupling constants can lead to bad distortions of geometry, while very low values allow substantial atomic drift. As a general rule of thumb, the coupling constant should reduce as the resolution/quality of the map degrades. Style and colour of representation are quite self-explanatory and a matter of user preference. I personally find it useful to show two maps – one as wireframe at 1.0 sigma and one as a transparent surface at a higher cutoff (Figure 9). The contour value may be chosen either in terms of map sigma or absolute units (note that the contour level of the map display will not change until you change this value). Once you have set the parameters to your liking you can save them by clicking “Set”. Further maps can be added by selecting “Add map” from the bottom left combo box – note that while ISOLDE presently won't prevent you from choosing the same map twice, this will lead to unpredictable behaviour and is inadvisable. Only maps registered in this way will be used in simulations. If you wish to change the settings for a map, select it from the bottom-left combo box, adjust the settings and click “Set” again.

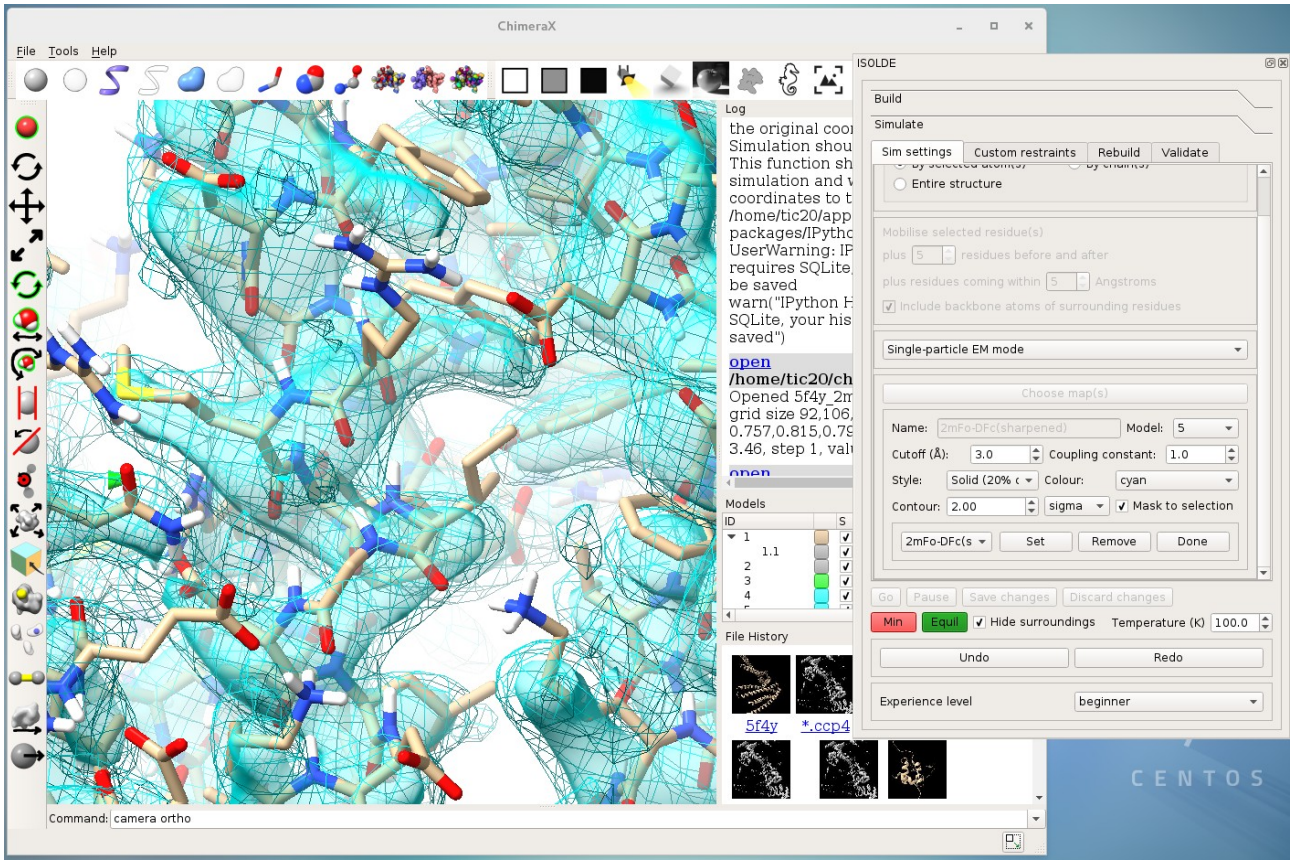

Figure 9: Selection and visualisation of maps for simulation. Here two maps have been selected, one of which is shown as wireframe at 1 sigma cutoff, and one as a transparent surface at 2 sigma. Note that at present it is unfortunately impossible to show the same map in two representations.

If you are observant you will have noted that in Figure 9 the map is somewhat out of register with the structure – a result of having run a simulation in Free mode for a while, allowing the coordinates to drift. This provides an ideal opportunity to show what ISOLDE can do. This test structure (5f4y) is 319 amino acid residues long, and I am running on a Xeon E5-2687 workstation (of which the simulation will use a single core) with a Nvidia GTX 1080 GPU. If I set the simulation mode to "Entire structure" and click "Go", there is a delay of about 10-20 seconds as the maps are masked and the simulation initialises (this can hopefully be reduced with further work). Figure 10 was captured approximately 5 seconds after the start of the simulation, by which time this gross shift in the structure is already essentially resolved. Reducing the temperature to 0K and settling for a further 20s yields a structure effectively indistinguishable from the original 5f4y, with the exception of a few flexible sidechains.

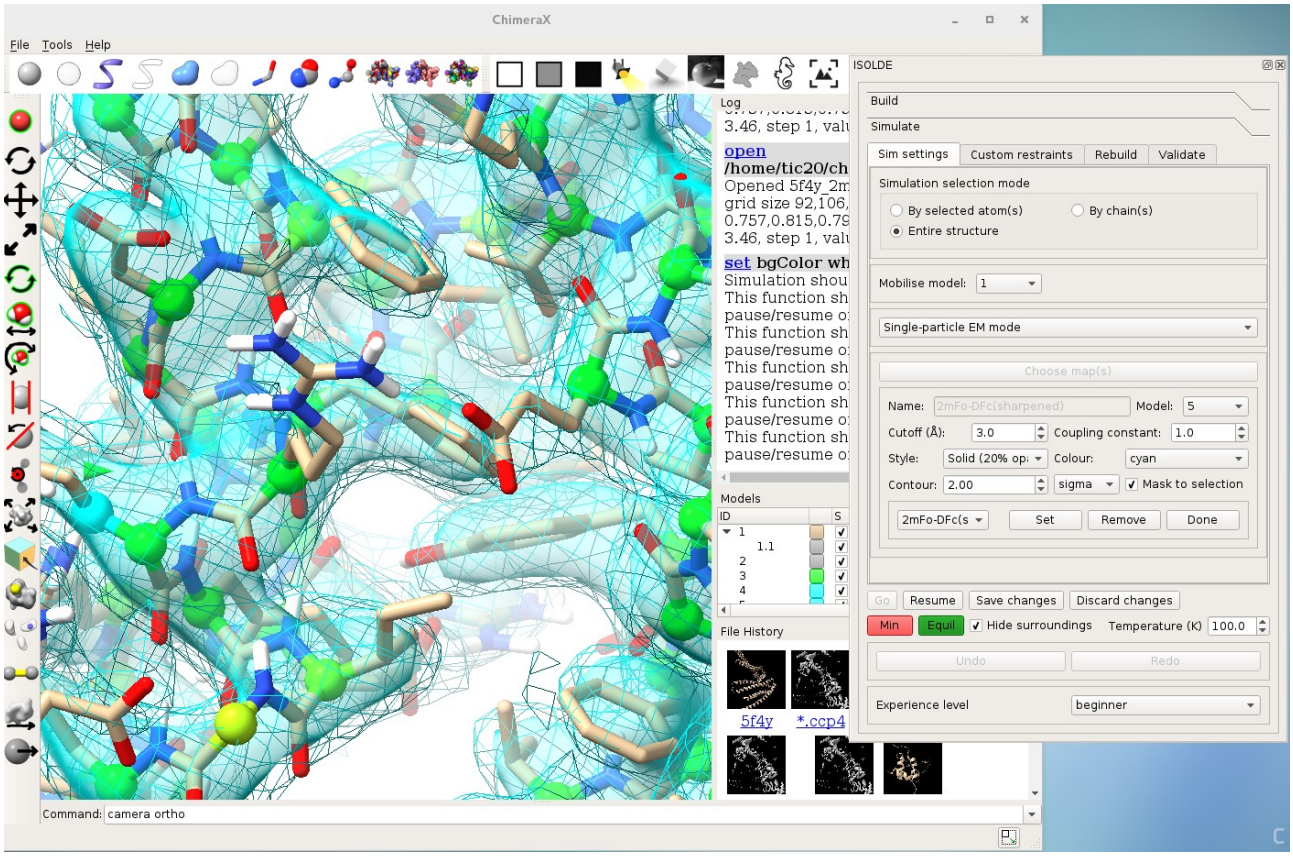

Figure 10: Snapshot approximately 5 seconds into a map-coupled simulation starting from the coordinates shown in Figure 9. The gross disparity between structure and map resolves quickly without significant distortion to the geometry.

Saving coordinates

At present, coordinate files can only be written from ChimeraX using the command line, and only mmCIF format output is supported (PDB format will be added later). To save, for example, model #1, enter:

```
save <filename>.cif #1
```

at the command line.