We have developed a data repository to store information associated with the crystallographic structure solution of each target. The repository uses a MySQL database back end for data storage and dynamically generated front-end HTML pages for user input. The system is implemented with the Python scripting language, facilitating integration with the *cctbx* (Grosse-Kunstleve et al., 2000) and PHENIX (Adams et al., 2002) components. We have designed interfaces for experimental data deposition in accordance with NIH guidelines, automated data analysis, modifying existing data, and viewing data and results. Users can also enter key files from each stage of the structure solution, such as the positions of anomalous scatterers, experimental phases, density-modified phases, and the final atomic model. The data-analysis interface allows users to automatically run the HySS program (Grosse-Kunsteleve and Adams, 2003) for substructure determination, followed by phasing with SOLVE (Terwilliger and Brendzen, 1999), and density modification and model building with RESOLVE (Terwilliger, 2000; Terwilliger, 2002; Terwilliger, 2003). In favorable cases, a fairly complete model is generated, which can then be extended and refined manually (e.g., BSGC target 1323B; a 75% model was automatically built using this procedure). When only lower-resolution data are available (2.8 Å or worse), a partial model is often generated. However, results usually indicate that data are sufficient to solve the structure, helping to minimize the time spent collecting additional data sets. To facilitate the efficient use of computing resources, we have implemented a mechanism that runs automated data analysis programs on a remote Linux cluster, using the SOAP protocol available in Python. All necessary data files are automatically transferred to the cluster, where a batch job is then submitted. After job completion, the results are transferred back to the Web server where they are made available in the repository HTML pages.

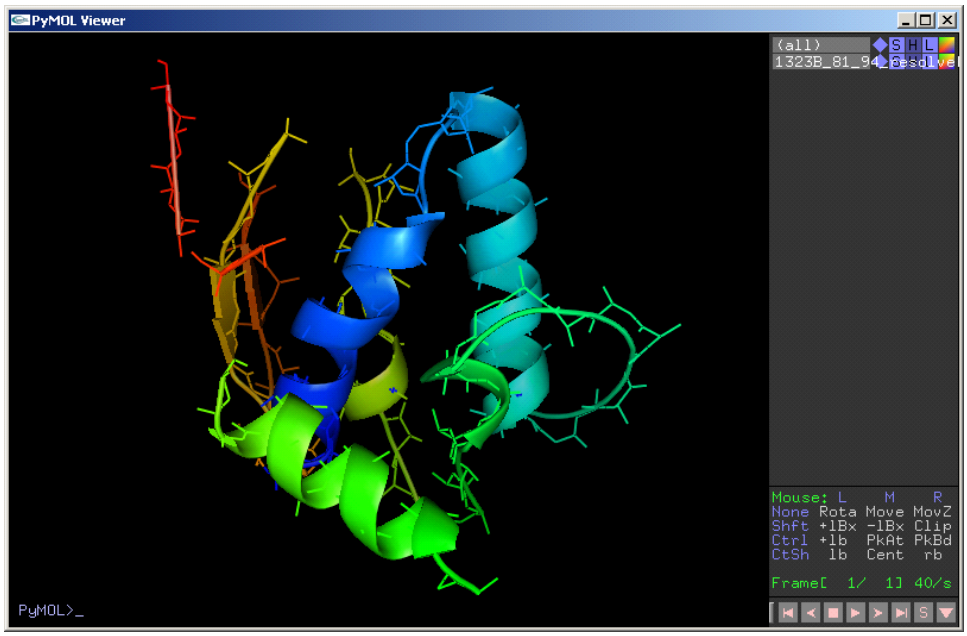

Figure 1: 75% complete model automatically generated for BSGC target 1323B using HySS, SOLVE and RESOLVE via the Data Repository interface. Image generated using PyMOL.

References

- Adams PD, Grosse-Kunstleve RW, Hung L-W, Ioerger TR, McCoy AJ, Moriarty NW, Read RJ, Sacchettini JC, Sauter NK, Terwilliger TC: PHENIX: building new software for automated crystallographic structure determination *Acta Cryst.* 2002, D58:1948-1954.
- Grosse-Kunstleve RW, Adams PD: Substructure search procedures for macromolecular structures. *Acta Cryst.* 2003, D59:1966-1973.
- Grosse-Kunstleve RW, Sauter NK, Moriarty NW, Adams PD: The Computational Crystallography Toolbox: crystallographic algorithms in a reusable software framework *J. Appl. Cryst.* 2002, 35:126-136.
- Terwilliger TC: Maximum likelihood density modification, *Acta Cryst.* 2000, D56, 965-972.
- Terwilliger TC: Automated structure solution, density modification and model building. *Acta Cryst.* 2002, D58, 1937-1940.
- Terwilliger TC: Automated side-chain model building and sequence assignment by template matching. *Acta. Cryst.* 2003, D59, 45-49.
- Terwilliger TC, Berendzen J: Automated MAD and MIR structure solution. *Acta Cryst*. 1999, D55, 849-861.