SEM (Symmetry Equivalent Molecules): a web-based GUI to generate and visualize the macromolecules

A. S. Z. Hussain, Ch. Kiran Kumar, C. K. Rajesh, S. S. Sheik and K. Sekar*

Bioinformatics Centre, Supercomputer Education and Research Centre, Indian Institute of Science, Bangalore, 560 012, India

Received December 30, 2002; Revised and Accepted April 9, 2003

ABSTRACT

SEM, Symmetry Equivalent Molecules, is a web-based graphical user interface to generate and visualize the symmetry equivalent molecules (proteins and nucleic acids). In addition, the program allows the users to save the three-dimensional atomic coordinates of the symmetry equivalent molecules in the local machine. The widely recognized graphics program RasMol has been deployed to visualize the reference (input atomic coordinates) and the symmetry equivalent molecules. This program is written using CGI/Perl scripts and has been interfaced with all the three-dimensional structures (solved using X-ray crystallography) available in the Protein Data Bank. The program, SEM, can be accessed over the World Wide Web interface at http://dicsoft2.physics.iisc.ernet.in/sem/ or http://144.16.71.11/sem/.

INTRODUCTION

In the present structural Bioinformatics era, advanced tools are highly essential to facilitate the display of useful information from the wealth of data available in the public repository Protein Data Bank (PDB) (1). The PDB is an archive of three-dimensional biological macromolecular structures determined experimentally using well-known techniques such as X-ray crystallography and NMR spectroscopy and is maintained by the Research Collaboratory for Structural Bioinformatics (RCSB). The three-dimensional structures are constantly being used by many researchers all over the world to unravel the underlying structure–function relationships. Towards this effort, several crystallographic modeling packages such as FRODO (2), CHAIN (3), ‘O’ (4) and SETOR (5) are available to generate and visualize the macromolecules (determined using the well known physical technique of X-ray crystallography) and its symmetry equivalent molecules. Almost all these programs have much less control to visualize the symmetry equivalent molecule(s) corresponding to a particular equivalent position present in the space group. Again, it is often difficult to save the three-dimensional atomic coordinates of the user interested symmetry equivalent molecules in the local machine for further analysis. To the best of our knowledge, there is no web-based program available to perform the above operations on all the crystal structures available in the PDB (1). To overcome this lacuna, with the least human intervention, the program SEM has been developed with the following features.

1. Symmetry related molecules around the reference molecule.
2. Molecules within the unit cell.
3. Molecules outside the unit cell.
4. Molecules within and outside the unit cell.

USAGE OF THE SOFTWARE

Towards this end, an efficient methodology has been designed (Fig. 1) to generate all the molecules within and outside the unit cell around the reference molecule. In addition, a graphics display facility has been provided to visualize the molecules and separate colour schemes have been adopted for the convenience of the users. The user needs to supply the following information associated with the given molecule in the appropriate box provided in the web page: for the three-dimensional crystal structures, available in the PDB (1), the user needs to provide the four character PDB-id code. The necessary values (unit cell constants and space group) will be automatically taken by the program from the uploaded atomic coordinate file. In addition, the user needs to choose the space group information from the pull down menu. The symmetry equivalent points for all the space groups have been incorporated in the program and the user needs to choose the correct space group using the pull down menu to get the corresponding equivalent positions from the in-built database.

*To whom correspondence should be addressed. Tel: +91 803601409; Fax: +91 803600683; Email: sekar@physics.iisc.ernet.in

Nucleic Acids Research, Vol. 31, No. 13 © Oxford University Press 2003; all rights reserved
The software is very flexible and has the option of visualizing all the symmetry equivalent molecules or any particular molecule generated within the unit cell. In addition, the user can save the three-dimensional atomic coordinates of all or any particular molecule(s) in the unit cell into their local hard disk. To perform this, the user needs to click ‘Click here to visualize the structures’ and then use the colour scheme provided in the pop-up window. Using the colour scheme, the user can choose a particular model or a set of models whose three-dimensional structure is to be viewed or to be saved in the local machine.

In addition, the program has an option to highlight a particular region of the molecule (in different representations) in the reference and its symmetry equivalent molecules. In this way, the users can see the role of a particular fragment in the crystal packing. The public domain graphics program RasMol (6) has been deployed and incorporated with the proposed package to visualize the reference and the symmetry related molecule(s). The user needs to interface the graphics program, RasMol, with the Netscape browser (only for the first time) (for instructions: http://144.16.71.11/sem/rasconf.html/). The program SEM is self-explanatory, easy to use and tested on Windows 95/98/2000, Windows NT, Linux and SGI platforms through the most popular WWW (World Wide Web) browser Netscape. The program is completely general and will work for most of the space groups available in the International Tables. The program can be used to see how the molecules are packed and arranged in the unit cell with respect to the reference molecule. In addition, this package can be used to generate the active biological tetramer from the two-homo dimers (7) like the protein quaternary structure (PQS) file server (8), which is an internet resource specially designed to generate the quaternary assembly. At the general outset, the package SEM is a very good teaching tool to the students undergoing graduate programs in structural Bioinformatics and X-ray
crystallography. The symmetry related molecules of the protein structure (9) (PDB-id code: 1UNE) generated using this program is shown in Figure 2.

The program SEM is written using CGI/Perl scripts and can be executed on our Bioinformatics Linux server [a 3.06 GHz Pentium IV processor; 1Gb (RD-RAM) of main memory]. The input data part and the pull down menu of this program are written in HTML and Java scripts. All the three-dimensional structures (using the most popular technique, X-ray diffraction) available in the PDB (1) have been incorporated in this package. The atomic coordinates are being updated every week and made available locally in our PDB-FTP server (Bioinformatics Centre, Indian Institute of Science, Bangalore, India). Hence, the users can access all the crystal structures available in the PDB at any given time. In the trial runs, the results appeared in about 30–40 s depending upon the size of the protein molecule and the network traffic. Please send your comments and suggestions for the inclusion of additional options to Dr K. Sekar (sekar@physics.iisc.ernet.in).

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the use of the Bioinformatics Centre (DIC), the Interactive Graphics Based Molecular Modeling (IGBMM) and the Supercomputer Education and Research Centre (SERC). The facilities DIC and IGBMM are supported by the Department of Biotechnology (DBT), Government of India.

REFERENCES