FASTModelfree

Roger Cole and J. Patrick Loria
Yale University

License

This program is free software; you can redistribute and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation; either version 2 of the License or and later version.

This program is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details.

You should have received a copy of the GNU General Public License along with this program; if not, write to the:

Free Software Foundation
59 Temple Place - Suite 330
Boston, MA 02111-1307, USA

Address any questions or comments to:

Prof. Patrick Loria
Department of Chemistry
Yale University
PO Box 208107
New Haven, CT 06520

email: patrick.loria@yale.edu

Required Software

Executable versions of FASTModelfree are available from the author's website. (http://xbeams.chem.yale.edu/~loria)

FASTModelfree is written in the Perl programming language and requires the optional Perl/Tk module. The base Perl package is now a default component of most UN*X operating systems. Unfortunately, the base Perl package lacks the ability to create graphical user interfaces. As a result, the optional Perl/Tk module was used to create the graphical user interface for FASTModelfree. If the graphical interface is undesired, the FASTModelfree configuration files can also be edited by hand and the program run without and use of the graphical user interface.

The official site for obtaining source code for both Perl and all Perl modules is CPAN (http://www.cpan.org). Since compiling Perl and the Perl/Tk module from source can be somewhat tedious, precompiled binary distributions of each are available for most common operating systems. In addition to CPAN, good resources for binary distributions include the O'Reilly Perl Resource (http://www.perl.com) and the Perl/Tk Archive (http://www.perltk.org). To facilitate the installation of this package, links are available from the Loria Group web site to both source code and binaries for Intel processors running RedHat Linux.

FASTModelfree interfaces with the program Modelfree 4.1, which is available for download from the laboratory of Arthur G. Palmer (http://cpmcnet.columbia.edu/dept/gsas/biochem/labs/palmer/software.html)

FASTModelfree was developed on an Intel processor based machine running RedHat Linux 7.3.

Configuration

The two components of FASTModelfree are the programs fastMF and setupFMF. The first line of each of these programs must point to the Perl binary, which may vary slightly from system to system. Depending on local configuration, users may also have to edit the line near the beginning of the program that specifies the location of the Modelfree 4.1 executable. For example, the following

```
#!/usr/bin/perl
#
# FASTModelfree 1.01
#

$modelfree="/usr/local/modelfree4.1/modelfree4";
```

would specify that the Perl binary is /usr/bin/perl and the modelfree4 executable is /usr/local/modelfree4.1/modelfree4.

Preparation of Input Files

Prior to beginning an analysis, several files must be created. Separate files for R1, R2 and NOE data files for each static field should be made. The form of the file should be one line per residue, with residue number, relaxation rate and error specified. The fields may be separated by any combination of spaces and tabs. An example is illustrated below (the first line should not be included in the data file):

Res #   Rate   Err
2   1.526   0.134
5   1.586   0.042
6   1.583   0.02
9   1.678   0.013

If an axially symmetric diffusion tensor is desired, a coordinate PDB file must also be properly prepared. Preparation of the PDB file requires several steps. First, all H atoms must be added to the structure file. This can be done by a number of programs, including the freely available program MolMol (http://www.mol.biol.ethz.ch/wutrich/software/molmol). If MolMol is used, the necessary H atoms may be added via the following MolMol commands

ReadPdb input.pdb
CalcAtom "H*"
SelectAtom 'atom.name "*Q*'
RemoveAtom 'selected'
WritePdb output.pdb

Next, the coordinate system of the PDB file should be moved to the center of mass of the molecule. Finally, an estimate of the diffusion tensor should be made and the axis of the PDB coordinate system should be rotated so that they are aligned with the principal axis of the diffusion tensor.

These final two transformations may be accomplished with the programs "pdbinertia" and "r2r1_diffusion" which are available from the website of Arthur G. Palmer. (http://www.mol.biol.ethz.ch/wutrich/software/molmol) Details for the operation of these programs can be found on the Palmer laboratory website.

The user should be aware that the program ModelFree will not tolerate anything other than ATOM statements. All other statements, headers, etc. should be removed before running ModelFree.

Parameter Input

The graphical user interface may be invoked with the command "setupFMF". The user may then specify the various parameters as desired. These parameters include:

FASTModelfree Parameters

Name for Output Files: This is the name that will be given to some of the output files created by FASTModelfree

SSECutoff: This is the α-critical value for the synthetic distribution of Γ_i. The Γ_i for a given spin is compared to the Γ_i of the α-critical value as part of the protocol to determine if a particular model is appropriate for the given spin. The most commonly used value is 0.95.

FTest Cutoff: This parameter is similar to the SSECutoff, but refers to the critical value used during the statistical F-test needed to assign two parameter models. The value most commonly used is 0.80.

Model 1 Only: In some cases the authors have noticed a significant improvement in speed when initially only those spins which fit model 1 are used to estimate the diffusion tensor. In cases where there are a large number of spins or multiple static magnetic fields it can be beneficial to perform analysis using only model 1, then after reaching a tensor which is self consistent with the set of spins assigned to model 1 begin the calculation again with this new tensor and allow all models to be fitted.

S^2 Cutoff: Spins whose calculated S^2 value is below this cutoff will be excluded during diffusion tensor optimizations.

Random Number Seed: Seed number given to the random number generator.

Number of MC Simulations: Number of Monte-Carlo simulations used for statistical error analysis.

MFPAR Parameters

^{15}N Magnetogyric Ratio: Magnetogyric ratio of ^{15}N. Typically -2.71.

N-H Bond Distance: N-H Bond distance, in Angstroms.

^{15}N CSA: ^{15}N chemical shift anisotropy, in ppm

Data Files

The necessary input data files and magnetic field strength may be selected here. Clicking on the box next to the text entry will open a file selection window. A PDB file is not required for use with an isotropic diffusion tensor.

Tensor Optimization Parameters

Diffusion Tensor: Axially symmetric or isotropic diffusion tensor may be selected.

Optimize Tensor?: If this option is set to "No", then only a single round of model selection will be performed and subsequently the diffusion tensor will not be optimized. In general, this should be set to "yes".

The remaining dialogue boxes allow specification of parameters related to the diffusion tensor. For each parameter, initial estimates, gridsearch ranges and number of gridsearch steps may be specified. The convergence limit indicates the maximum amount of change allowed in the tensor compared to the previous iteration to determine whether the system has converged to a diffusion tensor which is self consistent with the set of assigned models.

Running FASTModelfree

Once all parameters have been entered as desired, the user should select the SAVE CONFIG or SAVE CONFIG and EXIT button. This will write a file in the current directory named FMF.config which contains all of the desired parameters. This file is intended to be human readable, although all parameters may be specified within the graphical user interface.

The user may now invoke the second part of FASTModelfree by running the module "fastMF". This module reads the FMF.config file and begins the actual data anlaysis. A great deal of diagnostic output is created by this module, which details the decisions made in assign models and optimizing the tensor. Given that a typical run may take several hours, it is often helpful to run this module in the background and save the output into a log file. This may be accomplished with the UN*X command:

fastMF > mf.log &

which will run the fastMF module in the background and store the output in the file "mf.log".

FASTModelfree Output

The output from FASTModelfree is grouped into several files. First, the module "fastMF" will

directly output a great deal of information on the particulars of the model assignment. In addition, several other files will be created to futher detail both the process and results. These files are typically named according to the name for output files specified during setup. In this example, the name for output files specified in the graphical user interface is assumed to be RNaseA.

RNaseA.log: This will detail the model assignments and tensor data for each individual iteration, as well as whether the system has converged within the user defined limits.

RNaseA.x.par: The model assignments and spin parameters for each iteration are stored in separate tab delimited files. These files contain the final assigned models and relevant motional parameters.

RNaseA.x.pdb: These files are created only if an axially symmetric diffusion tensor is used. They contain the original PDB file with the coordinate system rotated to the princpal axis of the diffusion tensor used during the current iteration.

FASTModelfree Examples

Example spin relaxation data for the protein RNaseA is included in the distribution. Reasonable input parameters are already set, and the example dataset should run without any user modifications.

FASTModelfree without the GUI

It is possible to use FASTModelfree without using the accompanying GUI. The program setupFMF creates an output file named "FMF.config" which contains all necessary parameters to run FASTModelfree. FMF.config is human readable. A highly commented version is included with the distribution which should allow direct modification of the FMF.config file and subsequent use of the fastMF program.