relax: the analysis of biomolecular kinetics and thermodynamics using NMR relaxation dispersion data

Morin, Sebastien; Linnet, Troels Emtekær; Lescanne, Mathilde; Schanda, Paul; Thompson, Gary S.; Tollinger, Martin; Teilum, Kaare; Gagne, Stephane; Marion, Dominique; Griesinger, Christian; Blackledge, Martin; d’Auvergne, Edward J.

Published in:
Bioinformatics

DOI:
10.1093/bioinformatics/btu166

Publication date:
2014

Document version
Early version, also known as pre-print

Citation for published version (APA):

Download date: 10. sep., 2021
relax: the analysis of biomolecular kinetics and thermodynamics using NMR relaxation dispersion data

Sébastien Morin, Troels E. Linnet, Mathilde Lescanne, Paul Schanda, Gary S. Thompson, Martin Tollinger, Kaare Teilum, Stéphane Gagné, Dominique Marion, Christian Griesinger, Martin Blackledge, and Edward J. d’Auvergne

ABSTRACT

Nuclear Magnetic Resonance (NMR) is a powerful tool for observing the motion of biomolecules at the atomic level. One technique, the analysis of relaxation dispersion phenomenon, is highly suited for studying the kinetics and thermodynamics of biological processes. Built on top of the relax computational environment for NMR dynamics is a new dispersion analysis designed to be comprehensive, accurate and easy to use. The software supports more models, both numeric and analytic, than current solutions. An automated protocol, available for scripting and driving the GUI, is designed to simplify the analysis of dispersion data for NMR spectroscopists. Decreases in optimisation time are granted by parallelisation for running on computer clusters and by skipping an initial grid search by using parameters from one solution as the starting point for another – using analytic model results for the starting point for another – using analytic model results for the starting point for another – using analytic model results for the starting point for another – using analytic model results for the starting point for another – using analytic model results for the starting point for another – using analytic model results for the starting point for another – using analytic model results for the starting point for another – using analytic model results for the starting point for another – using analytic model results for the starting point for another – using analytic model results for the starting point for another – using analytic model results for numerically different solutions. The simplest is that of no motion (NoRex). For SQ CPMG-type experiments, analytic models include the original Luz and Meiboom (1963) multiple-site fast exchange models (LM63), the Carver and Richards (1972) and population-skewed Ishima and Torchia (1999) 2-site models for most time scales (CR72, IT99), and the Tollinger et al. (2001) 2-site very slow exchange model (TSMFK01). The CR72 model has been extended by Korzhnev et al. (2004) for MMQ data. For R1p-type data analytic equations include the Meiboom (1961) 2-site fast exchange model for on-resonance data (M61), extended by Davis et al. (1994) to off-resonance data (DPL94), and the Trott and Palmer (2002) and Miloushev and Palmer (2005) 2-site models for non-fast and all timescales (TP02, MP05). Different numeric solutions (NS) can be designed for SQ or MMQ data.

Biological macromolecules are intricate machines and their functions are closely related to their motions. These motions can be studied experimentally at the atomic level by NMR spectroscopy. Many important biological processes occur on the µs to ms timescale and, for atoms exchanging between different states, NMR relaxation dispersion can be observed. By studying this exchange process, kinetic and thermodynamic information can be obtained.

Using experimental data the solution to these equations reveals both populations of the molecular states (thermodynamics) and rates of exchange between them (kinetics). Though the general solution valid for all motions remains intractable, analytic solutions with restricted motions are available and are frequently used. The equations can also be solved numerically.

Two NMR dispersion methods are used for analysing motions: single, zero, double, or multiple quantum (SQ, ZQ, DQ, MQ) CPMG (Carr and Purcell, 1954; Meiboom and Gill, 1958); or R1p (Deverell et al., 1970). Combined SQ, ZQ, DQ, and MQ data will be labelled as multiple-MQ or MMQ data. Various models are used to analyse different data and motions. The simplest is that of no motion (NoRex). For SQ CPMG-type experiments, analytic models include the original Luz and Meiboom (1963) multiple-site fast exchange models (LM63), the Carver and Richards (1972) and population-skewed Ishima and Torchia (1999) 2-site models for most time scales (CR72, IT99), and the Tollinger et al. (2001) 2-site very slow exchange model (TSMFK01). The CR72 model has been extended by Korzhnev et al. (2004) for MMQ data. For R1p-type data analytic equations include the Meiboom (1961) 2-site fast exchange model for on-resonance data (M61), extended by Davis et al. (1994) to off-resonance data (DPL94), and the Trott and Palmer (2002) and Miloushev and Palmer (2005) 2-site models for non-fast and all timescales (TP02, MP05). Different numeric solutions (NS) can be designed for SQ or MMQ data.

Diverse software solutions exist for analysing relaxation dispersion data including CPMGFit (http://www.palmer.hs.columbia.edu/software/CPMGfit.html), cpmg_fit (available upon request from Dmitry Korzhnev), CATIA (Hansen et al., 2008), NESSY (Bieri and Gooley, 2011), GUARDID (Kleckner and Foster, 2012), ShereKhan (Mazur et al., 2013), and GLOVE (Sugase et al., 2013). The software relax (d’Auvergne and Gooley, 2008) is a platform for studying molecular dynamics using experimental NMR data and can be used as a numerical computing environment. Herein support for relaxation dispersion within relax is presented. Distributed as part of...
Table 1. Comparison of model support for different dispersion software

<table>
<thead>
<tr>
<th>CPMG-type</th>
<th>R$_{1p}$-type</th>
</tr>
</thead>
<tbody>
<tr>
<td>No-Rex</td>
<td>No-Rex</td>
</tr>
<tr>
<td>DAX3 2-site</td>
<td>DAX3 2-site</td>
</tr>
<tr>
<td>DAX3 3-site</td>
<td>DAX3 3-site</td>
</tr>
<tr>
<td>CR72</td>
<td>CR72</td>
</tr>
<tr>
<td>IT99</td>
<td>IT99</td>
</tr>
<tr>
<td>TSMFK01</td>
<td>TSMFK01</td>
</tr>
<tr>
<td>NSSQ 2-site</td>
<td>NSSQ 2-site</td>
</tr>
<tr>
<td>MMQCR72</td>
<td>MMQCR72</td>
</tr>
<tr>
<td>NS-MMQ 2-site</td>
<td>NS-MMQ 2-site</td>
</tr>
<tr>
<td>NS-MMQ 3-site</td>
<td>NS-MMQ 3-site</td>
</tr>
<tr>
<td>M61</td>
<td>M61</td>
</tr>
<tr>
<td>DPL94</td>
<td>DPL94</td>
</tr>
<tr>
<td>TP02</td>
<td>TP02</td>
</tr>
<tr>
<td>NS R$_{1p}$ 2-site</td>
<td>NS R$_{1p}$ 2-site</td>
</tr>
<tr>
<td>NS R$_{1p}$ 3-site</td>
<td>NS R$_{1p}$ 3-site</td>
</tr>
<tr>
<td>CPMGFit</td>
<td>CPMGFit</td>
</tr>
<tr>
<td>cpmg_fit</td>
<td>cpmg_fit</td>
</tr>
<tr>
<td>CATIA</td>
<td>CATIA</td>
</tr>
<tr>
<td>NESSY</td>
<td>NESSY</td>
</tr>
<tr>
<td>GUARDD</td>
<td>GUARDD</td>
</tr>
<tr>
<td>ShereKhan</td>
<td>ShereKhan</td>
</tr>
<tr>
<td>GLOVE</td>
<td>GLOVE</td>
</tr>
<tr>
<td>relax</td>
<td>relax</td>
</tr>
</tbody>
</table>

relax, this is the most comprehensive dispersion package supporting the

The number of dispersion models supported by relax is

extensive (Table 1). This allows for detailed comparisons between modern

number of dispersion models and NMR data types.

The setup of the auto-analysis includes defining the molecular

system, loading the dispersion data directly from peak lists,

clustering atoms with the same kinetics, modifying the list of dispersion

models, and setting up Monte Carlo (MC) simulations for error propagation.

Execution involves sequential optimisation of the models, fixed model

elimination rules to remove failed models and failed MC simulations

increasing both parameter reliability and accuracy (d'Auvergne and

Gooley, 2006), and a final run whereby AIC model selection is used to judge statistical

significance (Akaike, 1973; d'Auvergne and Gooley, 2003). The optimisation is designed for absolute accuracy and robustness but, as this can take time, it has been parallelised at the spin cluster and MC simulation level to run on computer clusters using OpenMPI. Three additional methods are used to speed up calculations, all designed to skip the computationally expensive grid search. The first is model nesting – the more complex model starts with the optimised parameters of the simpler. The second is model equivalence – when two models have the same parameters. For example the CR72 model parameters are used as the starting point for the CPMG numeric models resulting in a huge computational win. The third is for spin clustering – the analysis starts with the averaged parameter values from a completed non-clustered analysis.

The dispersion analysis in relax is implemented in Python using

NumPy and the GUI using wxPython. Optimisation using the

Nelder-Mead simplex and log-barrier constraint algorithms from the

minf$x$ library (https://gna.org/projects/minf$x$) removes the need for numerical gradient approximations which add a second numeric layer to the NS models. Data visualisation is via the software Grace.

ACKNOWLEDGEMENT

Nikolai Skrynnikov is thanked for his generous feedback and code contributions for implementing many of the numeric models and Flemming Hansen and Dmitry Korzhnev for kindly providing their software and published dispersion data.

REFERENCES


principle. In B. N. Petrov and F. Csaki, editors, Proceedings of the 2nd International


using NESSY. BMC Bioinformatics, 12, 421.

carr, H. y. and Purcell, E. M. (1954). Effects of diffusion on free precession in nuclear

magnetic resonance experiments. Phys. Rev. 94(3), 630-638.

Carver, J. and Richards, R. (1972). General 2-site solution for chemical exchange

produced dependence of T2 upon Carr-Purcell pulse phase. J. Magn. Reson.,

6(1), 89–105.

d’Auvergne, E. J. and Gooley, P. R. (2003). The use of model selection in the model-free


in the model-free dynamic analysis of NMR relaxation data. J. Biomol. NMR, 35(2),

117–135.


Biomol. NMR, 40(2), 107–133.


dissociation-rate constant for inhibitor-enzyme complexes via the T1rho and T2


Probing chemical shifts of invisible states of proteins with relaxation dispersion


proteins from measurements of transverse relaxation rates in solution. J. Biomol.

NMR, 14(4), 369–372.


rigorous analysis of CPMG RD NMR data. J. Biomol. NMR, 52(1), 1–12.


Probing slow dynamics in high molecular weight proteins by methyl-TROSY NMR

spectroscopy: application to a 723-residue enzyme. J. Am. Chem. Soc., 126(12),

3964–3973.

Lau, Z. and Meiboom, S. (1963). Nuclear magnetic resonance study of protolysis of the

methylammonium ion in aqueous solution - order of reaction with respect to solvent.


ShereKhan-calculating exchange parameters in relaxation dispersion data from


28(3), 430–431.

Meiboom, S. (1961). Nuclear magnetic resonance study of proton transfer in water. J.


exchange/ general approximations and some exact solutions. J. Magn.

Reson., 177(2), 221–227.


fitting of relaxation dispersion data using the flexible software package GLOVE. J.

Biomol. NMR, 56(3), 275–283.


Trott, O. and Palmer, 3rd, A. G. (2002). R(1hho) relaxation outside of the fast-exchange


Downloaded from http://bioinformatics.oxfordjournals.org/ at Danmarks NATURØG on April 16, 2014