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

Sébastien Morin 1,2, Troels E. Linnet 3, Mathilde Lescanne 4, Paul Schanda 4, Gary S. Thompson 5, Martin Tollinger 6, Kaare Teilm 3, Stéphane Gagné 1, Dominique Marion 4, Christian Griesinger 7, Martin Blackledge 4, and Edward J. d’Auvergne 4,7*

1 PROTEO, Université Laval, Québec G1V 0A6, Canada. 2 International AIDS Society HQ, CH-1202 Geneva, Switzerland. 3 Dept. of Biology, University of Copenhagen, DK-2200, Denmark. 4 Institut de Biologie Structurale, Grenoble F-38027, France. 5 Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds LS2 9JT, UK. 6 Institute of Organic Chemistry & CMBI, University of Innsbruck, A-6020, Austria. 7 NMR-based Structural Biology, Max Planck Institute for Biophysical Chemistry, D-37077 Göttingen, Germany.

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 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 numeric models, taking advantage of model nesting, and using averaged non-clustered results for the clustered analysis.

Availability: The software relax is written in Python with C modules and is released under the GPLv3+ licence. Source code and precompiled binaries for all major operating systems are available from http://www.nmr-relax.com.

Contact: edward@nmr-relax.com

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 R1,2 (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 (No Rex). 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 R1,2-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 time scales (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), GUARDD (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

*To whom correspondence should be addressed

© The Author(s) 2014. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.
relax, this is the most comprehensive dispersion package supporting the
greatest number of dispersion models and NMR data types.

The number of dispersion models supported by relax is
extensive (Table 1). This allows for detailed comparisons between
modern numeric and traditional analytic approaches. Different
user interfaces (UIs) can be employed to analyse dispersion
data including the prompt, scripting, and graphical user interface (GUI).
The scripting UI enables the greatest flexibility and allows for
most analysis protocols to be replicated. By implementing a novel
automated analysis and providing an easy to use GUI based on this
auto-analysis, the study of dispersion data is much simplified.

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, 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 a model clustering
- 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 library (https://gna.org/projects/minf/) 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 Kortzheev for kindly providing their
software and published dispersion data.

REFERENCES

principle. In B. N. Petrov and F. Csaki, editors, Proceedings of the 2nd
Akademia Kiado.
analysis using NESSY. BMC Bioinformatics, 12, 421.
Carver, J. and Richards, R. (1972). General 2-site solution for chemical exchange
produced dependence of T2 upon Carr-Purcell pulse separation. 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 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
by nuclear magnetic relaxation in rotating frame. Mol. Phys., 18(4),
553–559.
Probing chemical shifts of invisible states of proteins with relaxation dispersion
NMR spectroscopy: how well can we do it? J. Am. Chem. Soc., 130(8),
2667–2675.
proteins from measurements of transverse relaxation rates in solution. J. Biomol.
NMR, 14(4), 369–372.
for 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.
of the dimethylammonium ion in aqueous solution - order of reaction with respect to
Shere Khan-calculating exchange parameters in relaxation dispersion data from
28(3), 430–431.
chemical 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(1rho) relaxation outside of the fast-exchange