



Original software publication

SULISO: The Bath suite of vibrational characterization and isotope effect calculation software



Ian H. Williams*, Philippe B. Wilson

Department of Chemistry, University of Bath, Bath, BA2 7AY, United Kingdom

ARTICLE INFO

Article history:

Received 31 August 2016

Received in revised form

22 October 2016

Accepted 8 November 2016

Keywords:

Isotope effects

Vibrational frequencies

Computational chemistry

Fortran

ABSTRACT

Isotope effects are subtle but powerful probes of chemical reaction mechanisms and environmental conditions, with applications across chemical, biological and earth sciences. Their meaningful interpretation often relies on calculations based upon fundamental theories for their origin. The SULISO suite consists of four programs for the calculation of vibrational frequencies and isotope effects. CAMVIB is a broad vibrational characterization code developed for analysis of calculated harmonic frequencies and of normal modes in terms of internal coordinates. LIPFR calculates isotopic partition function ratios for pairs of isotopically substituted whole molecules, corresponding to conventional methodology, whereas UJISO is designed to perform similar calculations on subsets of atoms from very large systems. CUTOFF is a utility which truncates a force-constant matrix for a large system to obtain a smaller matrix appropriate for a specified subset of atoms.

© 2016 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY license
(<http://creativecommons.org/licenses/by/4.0/>).

Code metadata

Current Code version	v 2.0
Permanent link to code/repository used of this code version	https://github.com/ElsevierSoftwareX/SOFTX-D-16-00074
Legal Code License	MIT
Code Versioning system used	none
Software Code Language used	fortran
Compilation requirements, Operating environments & dependencies	UNIX
If available Link to developer documentation / manual	https://github.com/pbw20/SULISO_manuals
Support email for questions	Philippe.B.Wilson@bath.edu

Software metadata

Current software version	2.0
Permanent link to executables of this version	https://github.com/ElsevierSoftwareX/SOFTX-D-16-00074
Legal Software License	MIT
Computing platform / Operating System	UNIX
Installation requirements & dependencies	None.
If available Link to user manual – if formally published include a reference to the publication in the reference list	https://github.com/ElsevierSoftwareX/SOFTX-D-16-00074
Support email for questions	Philippe.B.Wilson@bath.edu

1. Introduction

Isotope effects (IEs) on rate or equilibrium constants are exquisitely subtle but yet extraordinarily powerful experimental probes of reaction mechanisms, environmental conditions and

structural changes throughout chemistry geochemistry and biochemistry [1]. IEs are employed in studies of chemical reactions and equilibria in a wide range of contexts in both gaseous and condensed-phases including enzyme-catalyzed processes. Substitution of one isotope of an element with another isotope (of the same element but with different mass) in general causes small but significant variations in the properties of a molecule or a material which are manifest in ratios of rate constants or equilibrium constants that differ from unity for reactions or equilibria involving the isotopic variants.

* Corresponding author.

E-mail address: I.H.Williams@bath.ac.uk (I.H. Williams).

It is vital to have robust computational protocols in place for their reliable calculation and interpretation. Existing programs for IE calculations were designed for isolated molecules and are inappropriate for subsets of atoms embedded in large supramolecular systems, such as commonly arise in quantum-mechanical/molecular-mechanical (QM/MM) simulations of condensed matter, including enzyme-catalyzed reactions and chemical reactions in solution [2]. The SULISO suite¹ comprises four programs to facilitate IE calculations based upon Cartesian force-constant (Hessian) matrices computed externally by means of electronic-structure (e.g. Gaussian09 [3]) or molecular-simulation (e.g. fDynamo [4]) codes. CAMVIB is an extensive vibrational calculation program, with numerous features for characterization and correlation of frequencies and normal modes. LIPFR calculates isotopic partition function ratios for pairs of isotopically substituted whole molecules, corresponding to conventional methodology, whereas UJISO is designed to perform similar calculations on subsets of atoms from very large systems. CUTOFF is a utility which truncates a force-constant matrix for a large system to obtain a smaller matrix appropriate for a specified subset of atoms; this is helpful in the determination of the minimal but sufficient size of Hessian for subsets of supramolecular systems. UJISO and LIPFR both act as end-of-line codes, taking input from CAMVIB and/or CUTOFF to calculate isotope effects from the vibrational frequency data.

2. Problems and background

The traditional theory of IEs based on molecular partition functions and the Bigeleisen equation [1] assume that each molecular system corresponds to a zero-gradient stationary point on a potential energy surface. Reliable optimized geometries and Hessians are readily available from modern quantum-chemical packages employing second-derivative methods, provided that convergence thresholds are set suitably tightly. However, since the turn of the millennium, it has become common to perform simulations for very large molecular systems (either molecules in solution or within enzyme active sites) but to compute explicit Hessians for only a subset of the total number of atoms in the system. For example, relaxation of a specified subset of atoms to a local minimum (or saddle point) may be performed within a frozen environment of the remaining atoms. In the context of QM/MM methods, the mobile subset and the frozen environment may be the same as the QM and MM regions, although different selections may also be made. In these circumstances the N_s subset atoms do not in themselves constitute a stationary structure in which vibrational degrees of freedom are separable from translations and rotation. The constraining influence of the environment means that diagonalization of the mass-weighted $3N_s \times 3N_s$ Hessian in Cartesian coordinates generally yields $3N_s$ non-zero eigenvalues which include six corresponding to libration of the whole subset with respect to its environment: “translation” and “rotation” of the subset as a whole are not free or separable motions but are coupled with the internal vibrational degrees of freedom.

It appears that other programs for IE calculations (e.g. QUIVER [5] and ISOEFF98 [6]) employ the Bigeleisen equation for $3N_s - 6$ internal degrees of freedom even for atomic subsets embedded within a supramolecular environment, and this has been achieved in published work either by projecting out the six external degrees of freedom or by simply ignoring the six lowest frequencies. In principle these procedures are both invalid and unnecessary, and should be replaced by methodology that uses all $3N_s$ frequencies of

the embedded subset, as described elsewhere [7]. Moreover, other programs tend to be organized in order to provide an equilibrium IE or a kinetic IE for a particular reaction involving single structures for each of the initial and final states of the system. This approach is valid for gas-phase processes but arguably is not very suitable for reactions in condensed phases. We have argued elsewhere that it is essential to determine the isotopic sensitivity of a large flexible supramolecular system by averaging over thermally accessible conformations [8]. This may be done by consideration of the (“reduced”) isotopic partition function ratio [1] (IPFR), f , for each conformation of each system. The average IE is obtained as the quotient of average IPFRs for the initial state (IS) and final state (FS).

$$\langle \text{IE} \rangle = \langle f_{\text{IS}} \rangle / \langle f_{\text{FS}} \rangle \quad (1)$$

where each individual value of f involves a product of partition functions for $3N_s$ degrees of freedom (or $3N_s - 1$ for a first-order saddle point together with a one-dimensional tunneling factor along the direction of the transition vector, within the range of validity of the harmonic approximation with respect to the de Broglie wavelength). Our programs LIPFR and UJISO are designed to compute IPFR values for a simple structure but for as many different isotopic substitutions as are desired; the results for many different conformations are collated and averaged by means of a spreadsheet.

3. Software framework

3.1. Software architecture

The four programs are written in FORTRAN and output from CAMVIB and CUTOFF is used as input for LIPFR and UJISO, respectively. Each code is split into individual subroutines which carry out individual part of the calculation explicitly, the different functionalities being called by a series of keywords. (See Fig. 1.)

3.2. Software functionalities

CAMVIB functions with three main aims. Firstly, normal coordinates and vibrational frequencies can be calculated for any number of atoms up to a limit set by a parameter MAXNAT in the included file CAM_SIZE. Secondly, the code allows for transformation of Cartesian force constants to internal or symmetry force constants, which can all individually be scaled and manipulated. CAMVIB can also compute compliance constants and relaxed force constants. All of these options are requested by a series of keywords within the program.

CUTOFF performs the function of truncating a $3N \times 3N$ Hessian to a smaller matrix representing the atomic positions specified in the input file. This outputs a new coordinate system and Hessian for input to UJISO.

UJISO and LIPFR calculate isotope effects with input originating from either program CUTOFF or CAMVIB, respectively. UJISO is particularly tailored to post-CUTOFF use to calculated IEs from subset Hessians, while LIPFR acts upon full systems. Multiple isotopes can be substituted and IPFRs calculated for each structure and Hessian. The calculations can also be carried out at a number of temperatures. In order to confirm the integrity of the IPFR calculations, the Teller–Redlich product rule [1] is employed: the independently calculated “masses and moments of inertia” and “vibrational product” terms should be exactly equal.

¹ In the Roman era the city of Bath was known as Aquae Sulis.

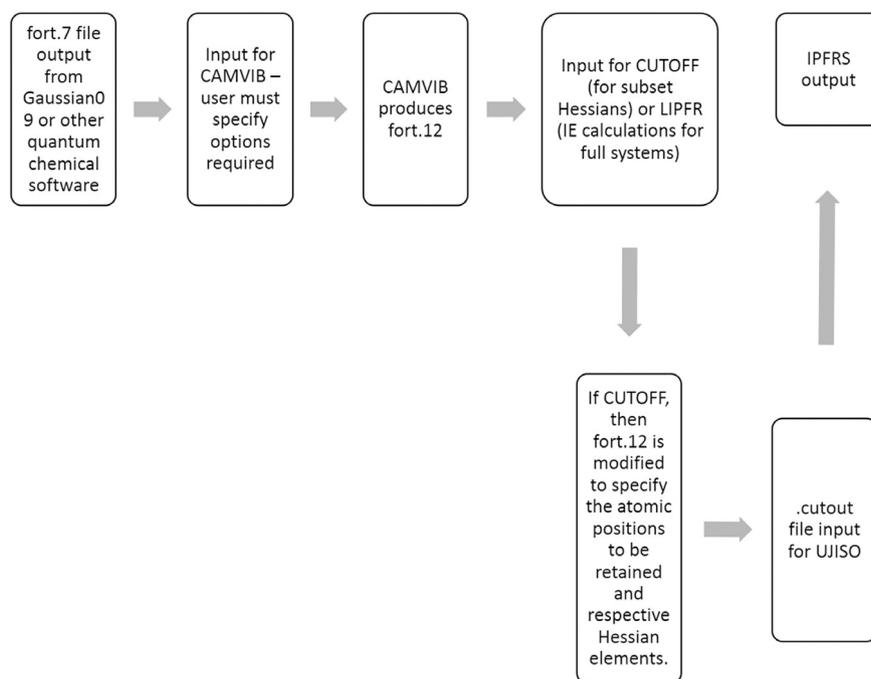


Fig. 1. Processing in the SULISO suite, allowing user manipulation between each stage of the process.

```

VIB
GAUS REDU IFCS VECT FULL DISO
Methyl cation in a solvent cage of five water molecules
1 1 0
12C 0.000000 0.000000 0.000000
1H 1.084985 0.000000 0.000000
1H -0.542493 -0.939625 0.000000
1H -0.542493 0.939625 0.000000
16O 0.000000 0.000000 2.040000
1H 0.000000 -0.786727 2.598282
1H 0.000000 0.786727 2.598282
&&&
*
S 1 2 0 0 C1H2
&&&
B 5 1 2 0 TBPB
&&&
&&&
*
0.6689155869D+00 -0.1603723016D-08 0.6906256398D+00
&&&
STOP

```

! to start a CAMVIB job
! keywords to specify options
! title
! symmetry number, electronic degeneracy...
! Cartesian coordinates for atom 1
! Cartesian coordinates for atom 7
! plus coordinates for 12 more atoms
! end of Cartesian coordinates
! valence coordinate for stretching + label
! more valence coordinates
! valence coordinate for bending + label indicating construction
! of symmetry coordinates for a trigonal bipyramid
! plus more valence coordinates (total 58 including 7 redundancies)
! end of valence coordinates
! first 3 Hessian elements
! plus 550 more lines of lower-triangular Hessian elements
! ends the job, but VIB would allow more data to be input for another calculation

Fig. 2. Example of input file for CAMVIB.

4. Implementation and empirical results

• Implementation details

Each of the programs within the suite are called through standard UNIX implementation:

```
./camvib.exe < input.vib > output.vibout
```

The convention with our codes is to use the following input extensions:

```
.vib – CAMVIB input
.cut – CUTOFF input
.uji – UJISO input
```

.lip – LIPFR input

Our programs function with any file extension chosen by the user; these are simply suggestions in order to provide meaningful reference to files in the future.

5. Illustrative examples

The following examples relate to a methyl cation surrounded by water molecules. A Gaussian09 geometry optimization and frequencies calculation generated a “punch” output file containing Cartesian coordinates and Hessian elements, which is edited by

```

1 1 0 Methyl cation in a solvent cage of five water molecules - cutoff ! title
12C* 0.0000000 0.0000000 0.0000000 ! Cartesian coordinates for atom 1: "*" denotes "keep this atom"
1H* 1.0849850 0.0000000 0.0000000
1H* -0.5424930 -0.9396250 0.0000000
1H* -0.5424930 0.9396250 0.0000000
16O 0.0000000 0.0000000 2.0400000 ! no "*", so this atom is to be "cut off"
1H 0.0000000 -0.7867270 2.5982820
1H 0.0000000 0.7867270 2.5982820
&&& ! plus coordinates for 12 more atoms
* ! end of Cartesian coordinates
0.668915586783E+00 -0.148461167860E-08 0.690625639761E+00 ! first 3 Hessian elements
&&& ! plus 550 more lines of lower-triangular Hessian elements
STOP ! ends the job, but CUTOFF would allow more data to be input for another calculation

```

Fig. 3. Example of input file for program CUTOFF.

```

UJI 1 0 1.0 ! UJISO job with 1 isotopologue in addition to the parent
1 298.15 ! one temperature and its value/K
Methyl cation: UJISO IPFR calculation from CUTOFF system, for D2 and 13C ! title
*
12C* 0.0000000 0.0000000 0.0000000 ! first of four atoms retained by CUTOFF
1H* 1.0849850 0.0000000 0.0000000
1H* -0.5424930 -0.9396250 0.0000000
1H* -0.5424930 0.9396250 0.0000000
* ! end of Cartesian coordinates
0.6689155868E+00 -0.1484611679E-08 0.6906256398E+00 ! first 3 Hessian elements
&&& ! plus 25 more lines of lower-triangular cutoff Hessian elements
1 0 ! structure is a TS; no frequencies excluded from IPFR calculation for parent
Deuterated methyl cation cut out from water cage ! title
12C 2H 2H 2H ! specification of isotopic composition for the isotopologue
1 0 ! structure is a TS; no frequencies excluded from IPFR calculation
STOP

```

Fig. 4. Example of input file for UJISO.

adding keywords, title and specification of internal coordinates, to yield a CAMVIB input file (Fig. 2).

The keywords on line 2 specify the following:

GAUS	Cartesian coordinates and Hessian are supplied in GAUSSIAN punch-file format.
REDU	A set of internal (valence) coordinates including redundancies is supplied; local symmetry coordinates are to be constructed from specific combinations of valence coordinates for bending.
IFCS	Force constants in internal (valence) coordinates are to be computed and printed.
VECT	Normal coordinates are to be printed.
FULL	Verbose output.
DISO	A fort.12 file is generated, containing information to be used subsequent by the other programs, including elements of a "cleaned-up" Hessian from spurious components of translation and rotation have been removed.

All keywords and options are described fully in the CAMVIB manual:

https://github.com/pbw20/SULISO_manuals.

The CAMVIB output for this structure, with a planar methyl moiety sandwiched midway between two equidistant water molecules positioned on either side on the local 3-fold symmetry axis, and with water molecule making a linear interaction with each of the three CH bonds, reveals it to be a transition state (TS) for S_N2 methyl transfer, possessing an imaginary frequency for displacement in the normal mode corresponding to Walden inversion.

Next, an input file for CUTOFF (Fig. 3) is prepared from the CAMVIB fort.12 output by insertion of an asterisk (*) immediately after the element symbol for each atom to be retained by the cutoff procedure. In this example the first four atoms, comprising the methyl moiety, are kept while all others are removed: their Cartesian coordinates and their corresponding rows and columns of the Hessian are deleted.

The output from CUTOFF is in turn used to prepare the input for UJISO (Fig. 4) by adding some information at the head of the file and a specification of isotopologue(s) at the tail.

The UJISO output (Fig. 5) shows 12 vibrational frequencies. The first frequency (corresponding to out-of-plane bending of the planar methyl moiety) is printed as -253.7355 , but is actually imaginary; it is identified as the transition frequency for which a quantum correction to the IPFR is evaluated. The excitational (EX) and zero-point-energy (ZP) contributions to the IPFR, which appears as $Q(\text{heavy})/Q(\text{light})$, is evaluated over all of the 11 real frequencies.

The structure of the input file for LIPFR is essentially the same as for UJISO. (Full details may be found in the manual.) This program is used to compute IPFRs for a whole molecule, not a subset of a larger system: in this example we consider the methyl cation in isolation. Since its translational and rotational degrees of freedom are considered explicitly, along with the internal vibrational degrees of freedom, in general information in one line of the tail of the file differs from that for UJISO. Specifically, the line

```

Atom      Mass      X      Y      Z (Ångström)
1  12C  12.000000  0.0000000  0.0000000  0.0000000
2   2H   2.014102  1.0849850  0.0000000  0.0000000
3   2H   2.014102 -0.5424930 -0.9396250  0.0000000
4   2H   2.014102 -0.5424930  0.9396250  0.0000000

Frequencies of normal modes (cm-1)
-253.7355  207.3354  260.0611  311.3845  725.7330  745.6020
 987.7587 1053.8616 1070.9656 2196.5442 2475.5201 2479.2063

Transition frequency = 253.7355i cm-1
EX and ZP evaluated over 3N-1 frequencies excluding the transition frequency

Product Rule for isotopologue 2
  Isotopic mass term (IM) = 0.0443472301
  Frequency term (VP) = 0.0443472301

Transition frequency ratio TF1/TF2 = 1.0265648598

Partition Function Ratio for isotopologues 1 and 2 at 298.15 K
Q(heavy)/Q(light) = IM * EX * ZP
0.3040857533E+04 0.4434723010E-01 0.1379677111E+01 0.4969951090E+05

Quantum correction to transition frequency ratio = 0.1003459328E+01
Quantum corrected isotopic partition ratio = 0.3030374473E+04

```

Fig. 5. Tail of UJISO output file for cutoff methyl.

```

Atom      Mass      X      Y      Z (Ångström)
1  12C  12.000000  0.00000000  0.00000000  0.00000000
2   2H   2.014102  0.00000000  0.00000000  1.10120400
3   2H   2.014102  0.95367000  0.00000000 -0.55060200
4   2H   2.014102 -0.95367000  0.00000000 -0.55060200

The 3 Principal Moments (Amu Å**2) are:
  7.3271989715  3.6635970244  3.6636019471

Frequencies of normal modes (cm-1)
0.0000  0.0002  0.0007  0.0008  0.0012  0.0017
1017.9980 1018.0193 1107.0826 2130.8920 2403.2622 2403.2642

The 6 rejected frequencies are:
0.0000  0.0002  0.0007  0.0008  0.0012  0.0017

EXC and ZPE evaluated over 3N-6 frequencies

Temperature = 298.15 K, Std. State = 1 Atmosphere
-(G-PE)/RT= -2.757546 = -(G-E0)/RT= 21.567524 -(ZPE/RT= 24.325070)

TRANS      ROT      EXC      ELEC
14.917999  6.629925  0.019600  0.000000

Teller-Redlich Product Rule for isotopologue 2
  Mass and Moment of Inertia term = 6.06466103
  Frequency term = 6.06466103

Partition Function Ratio (IPFR) for isotopologues 1 and 2 at 298.15 K
ln(Q2/Q1) = 9.943830  Q2/Q1 = 20823.355537
ln(MMI) = 1.313226  MMI = 3.718151
ln(EXC) = 0.016093  EXC = 1.016223
ln(ZPE) = 8.614511  ZPE = 5511.053379

```

Fig. 6. Tail of LIPFR output for isolated methyl cation.

is replaced by

0 6 1 2 3 4 5 6

where the initial zero denotes that the structure of the isolated methyl cation is not a TS, the “6” refers to the number of external degrees of freedom (corresponding to translational and rotational modes with essentially zero-valued frequencies), followed by the six ordinal numbers specifying which frequencies are to be ignored in the vibrational part of the IPFR calculation.

The LIPFR output (Fig. 6) shows 6 essentially zero frequencies, corresponding to translational and rotational modes, which are excluded from the IPFR calculation. The translational and rotational contributions to the IPFR are evaluated from the atomic masses and the principal moments of inertia, and together make up

the MMI factor. The vibrational components (EXC and ZPE) of the IPFR (printed as Q2/Q1) are evaluated over the 6 real vibrational frequencies. The two terms in the Teller–Redlich product rule are exactly equal, confirming that the geometry and the Hessian are consistent with each other.

An individual IPFR value ($f_{\text{indiv}} \equiv Q_{\text{heavy}}/Q_{\text{light}}$ or Q_n/Q_1) and its factors for a single conformation of a system, from either UJISO or LIPFR, may be transferred to a spreadsheet or other program in order to evaluate other properties.

- The IPFR f_{state} for a state of the system may be obtained by averaging over an ensemble of N thermally accessible conformations. If the individual structures were generated as snapshots at regular intervals from an equilibrated molecular dynamics

trajectory, they should already represent a Boltzmann distribution and so a simple arithmetic mean provides the appropriate average IPFR.

$$f_{\text{state}} = (1/N) \sum f_{\text{indiv}} \quad (2)$$

- Either an equilibrium or a kinetic IE may be obtained as a ratio of IPFRs for initial and final structures or states.

Equilibrium IE between individual structures for reactant (R) and product (P) = f_R/f_P .

Kinetic IE between reactant state (RS) and transition state (TS) = $\langle f \rangle_{RS} / \langle f \rangle_{TS}$.

6. Conclusions

The SULISO suite of programs for vibrational characterization and isotope effect calculation has been described. In a time where IEs are increasingly used in both molecular and supramolecular research, it is vital to have robust computational protocols for their calculation, based on reliable methodologies. We have developed these programs and have demonstrated their utility in several applications [9]. We recommend their use by others.

Acknowledgment

We thank the Engineering and Physical Sciences Research Council for a DTA studentship to PBW.

References

- [1] Wolfsberg M, Van Hook WA, Paneth P, Rebelo LPN. *Isotope effects in the chemical, geological, and bio sciences*. Dordrecht: Springer; 2010.
- [2] Wilson PB, Williams IH. In: Tuñón I, Moliner V, editors. *Simulating enzyme reactivity: computational methods in enzyme catalysis*. Cambridge: Royal Society of Chemistry; 2016. ISBN 9781782624295.
- [3] Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, et al. *Gaussian 09 (Revision D.01)*. Wallingford, CT: Gaussian Inc.; 2009.
- [4] Field MJ. *A practical introduction to the simulation of molecular systems*. Cambridge: Cambridge University Press; 1999.
- [5] QUIVER, Saunders M, Laidig KE, Wolfsberg M. *J Am Chem Soc* 1989;111:8989–94.
- [6] ISOEFF98, Anisimov V, Paneth P. *J Math Chem* 1999;26:75–86.
- [7] Williams IH. *J Chem Theor Comput* 2012;8:542–53.
- [8] Ruiz-Pernía JJ, Williams IH. *Chem Eur J* 2012;18:9405–14.
- [9] (a) Wilson PB, Williams IH. *Angew Chem Int Edn* 2016;55:3192–5;
(b) Wilson PB, Weaver PJ, Greig IR, Williams IH. *J Phys Chem B* 2015;119:802–9;
(c) Wilson PB, Williams IH. *Mol Phys* 2015;113:1704–11;
(d) Kanaan N, Ruiz-Pernía JJ, Williams IH. *Chem Commun* 2008;6114–6.