

# LIPFR

**Philippe B. Wilson and Ian H. Williams**

*University of Bath*

LIPFR was created by IHW (University of Bath). Its function is to calculate isotopic partition function ratios (hereafter referred to as IPFRs) from a Hessian obtained by means of an electronic-structure calculation for a full molecular system comprising  $N$  atoms and corresponding to a stationary point on a potential energy surface.

LIPFR is written in Fortran90, with some elements of earlier versions, but compiles fluidly under standard GNU compilers (gfortran...). In the event of compiler errors, g77 generally overcomes these appropriately. The present version is dimensioned by using the external CAM\_SIZE file which can be edited to change the number of atoms and valence coordinates (MAXNAT=1000).

```
! Header file for CAMVIB, CAMISO, LIPFR and UJISO to set array
! sizes as functions of the maximum number of atoms.
!
! 31 April 1997 - AJT
! 16 October 2016 - IHW

      parameter(MAXNAT =1000)

! These should not be touched:
      parameter(MAXNC  = MAXNAT * 3)
      parameter(MAXNI  = MAXNAT * 3 + MAXNAT / 2)
      parameter(MAXNINT= MAXNC  - 6)
      parameter(MAXDIAG= MAXNC  * (MAXNC + 1) / 2)
      parameter(MAXAPO = MAXNAT + 1)

! For CAMISO:
      parameter(MAXFCMS = MAXDIAG * 3)
      parameter(MAXXYZ =  MAXNAT * 10)
```

**Figure 1.** The CAM\_SIZE file, used to dimension the arrays in CAMVIB, LIPFR, and UJISO. Parameter MAXNAT (currently set at 1000) determines the maximum number of atoms. For most purposes a much smaller number may be chosen.

The aim of this user guide is to gently introduce the individual to the functionalities of LIPFR and provide explicit descriptions of the input and output one can expect from standard LIPFR runs. We will begin by discussing the format of the input, suggesting possible edits to the source code, which is available from the SULISO\_source repository on Github: [https://github.com/pbw20/SULISO\\_source](https://github.com/pbw20/SULISO_source)

## Overview of LIPFR input

- P. B. Wilson & I. H. Williams

NT is the number of temperatures (I2) and the values of  
T are the absolute temperatures in Kelvin (5F10.4) to be used in the IPFR evaluations.

3. Header section, line 3: (FORMAT 3I3,74A)

IS, IG, IFFREQ, TITLE

IS is a symmetry number (as appropriate, or set equal to 1 if symmetry is not required),

IG is the electronic degeneracy (usually = 1),

IFFREQ is the number of frequencies to be input (= 0 if a Hessian is supplied, but otherwise as appropriate).

TITLE is text string used as a description for the file or calculation.

4. Coordinate input, one line per atom, from line 4 to line 4 + N, where N is the number of atoms: (FORMAT A4, 1X, 3F12.7).

IA, (X(J, N), J=1, 3)

IA is the element symbol (right-justified in the A4 field) for an atom, and is followed by a space; X are the x, y, and z Cartesian coordinates in Ångström (3F12.7).

The following element symbols are currently accepted, although the UJISO source code may be edited to make additions or changes:

1H, 2H, 3H, 12C, 13C, 14C, 14N, 15N, 16O, 17O, 18O, F, 32S, 34S, 35CL, 37CL, 6LI, 7LI, 9BE, 10B, 11B, SI, P, 79BR, 81BR, NA, SP, MG, AL, CL, K, CA, ZN, BR, RB, I, CS, B, H, C, N, O, S, HE, NE, AR, KR, 33S, 36S, 11C

An element symbol preceded by an atomic mass specify a particular isotope (e.g. <sup>1</sup>H) whereas an element symbol without a mass specifies the chemical atomic weight (e.g. H = 1.00797). The special symbol 'SP' specifies a mass of 999.

5. Data separation line, line 5 + N : (FORMAT A4)

Consists of an asterisk '\*' (left-justified in the A4 field) to separate the coordinate input from the Hessian. (This character is a special value for 'IA' in item 4.)

6. *Either* (if IFFREQ = 0):

Hessian input lines, from line 6 + N to line 6 + N + m: (FORMAT: 3E20.10)

(F(I), I=1, NNC)

This is the lower triangle of a square symmetric Hessian matrix, stored as a linear array. The number of unique elements is  $NNC = N \times (N + 1)/2$ . The number of three-column lines of data in this section is therefore given by:

$$\text{Number of lines of Hessian, } m = \frac{3N}{2} \times \frac{3N + 1}{3}$$

The number of lines of Hessian output from a calculation can be used to verify the reasonableness of the input, by comparing this with the number of atoms using the formula above.

*Or* (if IFFREQ = NFREQ):

Frequency input lines, line 6 + N to line 6 + N + m: (FORMAT(6F12.4))

(FREQ(I), I=1, NFREQ)

FREQ are a selection vibrational frequencies (as wavenumbers /  $\text{cm}^{-1}$ ) from a prior calculation on this molecular species. NFREQ does not need to be equal to either  $3N$  or  $3N - 6$  but may have any integer value.

Number of lines of frequencies,  $m = \text{ROUNDUP}(\text{NFREQ}/6, 0)$

[The total number of lines of data up to this point is  $q$ .]

7. Frequency control line for parent species, line  $q + 1$ : (FORMAT 20I4)

ITS, NX, (IX(L), L=1, NX)

ITS = 0 if the species is **not** to be treated as a transition structure. Normally this is because the species is an equilibrium structure corresponding to a minimum on the potential energy surface.

ITS =  $n$  (a positive integer) specifies the ordinal number of the reaction coordinate frequency (RCF) in the list of all  $3N$  vibrational frequencies for a species which **is** to be treated as a transition structure. Usually a transition structure corresponds to a first-order saddle point and possesses a single imaginary frequency which appears first in the ordered list of vibrational frequencies; in this case ITS = 1. Sometimes a species may correspond to a higher-order saddle point, but it may be desired to treat other imaginary frequencies as if they were real; for example, if the RCF appears as third in the list, then ITS = 3, and this frequency alone will be excluded from the partition function but used instead to determine the tunnelling correction.

NX quantifies the number of frequencies to be excluded from the IPFR evaluation.

If the species corresponds to a genuine zero-gradient stationary point, then 6 eigenvalues of the full  $3N$ -dimensional Hessian should be zero. In practice, these values may not be exactly equal to zero for a number of reasons, including numerical rounding errors, and it is helpful to identify them explicitly. Usually, NX = 6, and the corresponding degrees of freedom are considered as translations and rotations which are excluded from evaluations of vibrational IPFRs and zero-point energies.

The third and subsequent integers IX (format I4) on this line refer to the ordinal numbers of the frequencies to be excluded from the calculations. If the molecular species is an equilibrium structure with  $3N - 6$  real vibrational frequencies, this line would read as:

\_\_\_0\_\_\_6\_\_\_1\_\_\_2\_\_\_3\_\_\_4\_\_\_5\_\_\_6

8. Isotopologue 1 title specification, line  $q + 2$ : (FORMAT 3I3,74A)

IS, IG, IFFREQ, TITLE

This is the same as for item 3, but with appropriate data to describe the first isotopically substituted species.

9. Isotopologue 1 atom specification lines, from line  $q + 2$  to line  $(N/16)$  where  $N/16$  refers to the 16 slots of atoms on one line of the specification, and is therefore equal to the nearest integer, rounded up. For example, an 18-atom species would be specified by one line of 16 atoms and another of 2 atoms, i.e. 2 lines (FORMAT (16(A4,1X))).

(IAT(I), I=1, NAT)

Each of the  $n$  atoms specified in the coordinate input lines (from line 5 to line  $5 + n$ ) is included in the format of 16 slots of format A4, separated by one space between each atom (1X). This is where the isotopic substitution can take place. Therefore, for the following atoms in the coordinate input:

12C, 1H, 1H, 1H

The isotopologue 1 atom specification line for a triply deuterated isotopologue would be:

`_12C__2H__2H__2H`

where again the underscores are used to represent the formatting.

(if IFFREQ = NF):

This line is followed by *m* lines of frequencies for this isotopologue (also applies to subsequent isotopologues). Sample input files for a frequency-based IPFR calculation, as well as traditional Hessian-based calculations are included in the following GitHub repository:

[https://github.com/pbw20/SULISO\\_sample\\_files](https://github.com/pbw20/SULISO_sample_files)

#### 10. Frequency control line for isotopologue 1: (FORMAT 2014)

As for item 7 for the parent structure, this line specifies whether or not the species is to be treated as a transition structure, and whether any frequencies are to be excluded from the IPFR calculation. The guidelines for item 7 can be followed here.

#### 11. Subsequent isotopologue lines.

A repeat of items 8, 9 and 10 for subsequent isotopologues, therefore a minimum of 3 new lines per additional isotopologue. Remember that the number of isotopologues for IPFRs to be calculated here, including the parent structure, must match that specified in the Header section line 1.

#### 12. Final line: (FORMAT A4)

The final line of the program consists of the word STOP, followed by a line break, in order to inform the program that no more input remains to be read.

## Overview of LIPFR output

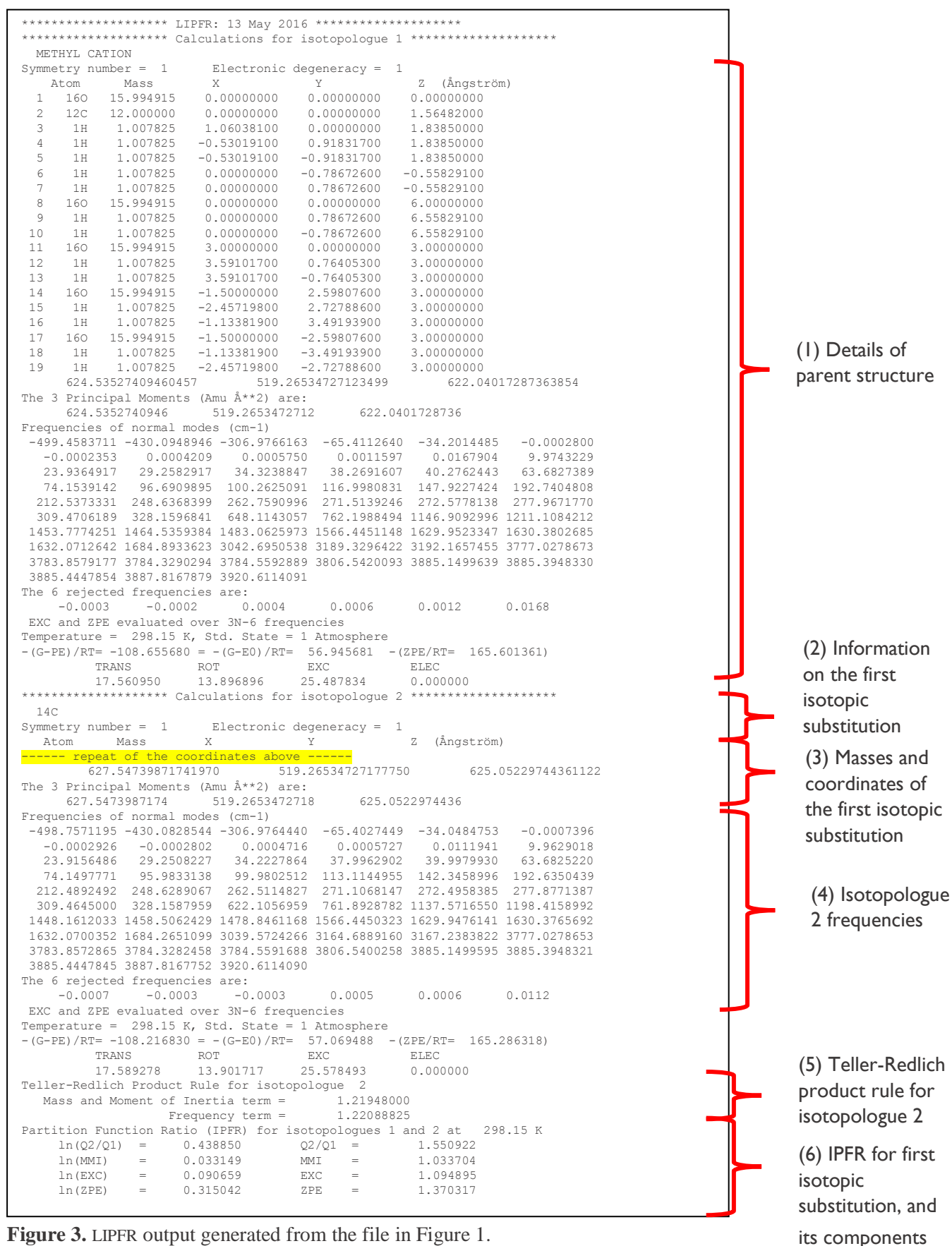


Figure 3. LIPFR output generated from the file in Figure 1.

1. LIPFR output begins by printing the version of the code currently being executed, in this case, the 13<sup>th</sup> May 2016 release. Then follows the data for the parent structure, in this case a methyl cation structure within a cage of water molecules.
2. Information on the first isotopic substitution (isotopologue 2) is output, including the ‘Title’ line, which was read from the input, as well as the electronic degeneracy and symmetry of the molecule.
3. The masses, atoms, and coordinates are then printed out for isotopologue 2, the coordinates should be the same as for the parent structure (isotopologue 1). In the interest of space, these have been omitted in this output in order to show more of the calculation results.
4. Frequencies calculated from the input un-mass-weighted Hessian and the specified isotopic substitutions are calculated. For heavier isotopologues the frequencies should be lower, and *vice versa*. This is another option to check the reasonableness of the calculations.
5. The product rule is printed out to test the frequency/ mass term equality which should be present in all calculations. If the two terms are equal, then there is little error in the calculated IPFR and input structures, however the holding of the Teller-Redlich product rule is a good measure of the quality of input structures and it is always a good idea to inspect it in order to check that both components are equal to all or the majority of decimal places. For more information on the Teller-Redlich product rule, see:  
 E. Teller, cf. Angus, W. R.; Bailey, C. R.; Hale, J. B.; Ingold, C. K.; Leckie, A. H.; Raisin, C. G.; Thompson, J. W.; Wilson, C. L. *J. Chem. Soc.* **1936**, 971-987.  
 Redlich, O. Z. *Physik. Chem. B* **1935**, 28, 371-382.  
 Williams, I. H., *Journal of Chemical Theory and Computation* **2012**, 8 (2), 542-553.
6. The IPFR calculated as a quotient of partition functions of the heavy isotopologue over the light, is printed out as a component of the different factors: mass (MMI), excitational (EXC), and zero-point (ZPE).

Kinetic isotope effects may be obtained from IPFRs evaluated for light (L) and heavy (H) isotopologues of RS and TS species as follows:

$$\text{KIE} = \frac{k_{\text{light}}}{k_{\text{heavy}}} = \frac{\left(\frac{Q_L^{\text{TS}}}{Q_L^{\text{RS}}}\right)}{\left(\frac{Q_H^{\text{TS}}}{Q_H^{\text{RS}}}\right)} = \left(\frac{Q_H^{\text{RS}}}{Q_L^{\text{RS}}}\right) / \left(\frac{Q_H^{\text{TS}}}{Q_L^{\text{TS}}}\right) = \frac{IPFR^{\text{RS}}}{IPFR^{\text{TS}}}$$

(where  $Q$  denotes a molecular partition function). This expression explains the convention of quoting IPFRs as “heavy/light”.

Equilibrium isotope effects may be obtained from IPFRs evaluated for light (L) and heavy (H) isotopologues of reactant and product species, by replacing TS by PS in the equation above.

To obtain a KIE averaged over multiple thermally-accessible conformations of a flexible system, within which the atomic subset considered by UJISO is embedded, it is convenient and appropriate to perform independent averaging of IPFRs over individual RS species for the reactant state and individual TS species for the transition state. This may be performed easily by collecting IPFRs from individual UJISO calculations into a spreadsheet.

$$\langle \text{KIE} \rangle = \langle IPFR^{\text{RS}} \rangle / \langle IPFR^{\text{TS}} \rangle$$

see: Ruiz-Pernía, J. J.; Williams, I. H. *Chem. Eur. J.* **2012**, 18, 9405-9414.



## LIPFR Troubleshooting guide

There are a number of errors that users can come across, which will hopefully decrease in number as they familiarise themselves with the code. The most common of these are input errors. Luckily, the program will inform the user within the UNIX environment of the last read format and line number in the fortran source code, and therefore a likely area linking to the error. Throughout this manual, the formats of input have been specified to allow the user easy identification of the areas which require specific types. This should be the first reference point in the case of incomplete termination.

The advantage of the SULISO suite is that when programs terminate successfully, a line similar to that below will terminate the output:

```
***** THE END *****
```

This will ensure that users are certain of normal termination before considering results.

One error that has been noted is the format of certain coordinate/Hessian outputs from electronic structure codes. Some codes simply print out the coordinates, followed directly by the Hessian. However, sometimes the gradient is printed following the coordinates and preceding the Hessian. This consists of  $n$  lines of a similar format to the Hessian, where  $N$  is the number of atoms in the structure. This is often noticed upon processing, where erroneous frequencies are printed from the program and the input needs to be re-inspected. The simplest solution is to use the equation in Part 6 of the *Overview of LIPFR input* to deduce the expected number of lines of Hessian, and a program such as Notepad++ with a line count to identify if lines of gradient are present. When located, these can simply be deleted from the input and the programs rerun, which should fix the errors.

Additionally, in order to obtain Hessian matrices from electronic structure programs, frequency calculations must be carried out. For example, in Gaussian09, the keywords *punch=(coord,derivatives)* must be specified, along with *opt* and *freq* for a Hessian to be produced. Single point energy calculations, certainly in Gaussian09, will not yield a meaningful Hessian for isotope effect calculation.

## **Appendix – Additional examples of input and output**

The aim of this section is to provide addition information on different systems which can be treated with the programs of the suite. These can be used on any system based on a Cartesian coordinate array and Hessian matrix.

```

IPFR 2
1      418.15
1 1 0 RU 4ts
RU      -0.9722730  -0.6065750  0.1560350
16O     -1.8939790  -0.7732900  2.0939040
12C     -1.0485170  -1.6796020  2.4488440
&&&
12C      2.0129410   3.7155340   1.3270220
1H       1.6539620   4.7074760   1.0253450
1H       2.5608650   3.7809550   2.2802970
*
0.2508853214E+00    0.1175723034E-01    0.2283493745E+00
0.5990520998E-01    0.6983872957E-02    0.3947488000E+00
&&&
-0.3858175958E-02   -0.1752284733E-02   -0.6311020167E-02
0.1177052557E+00    0.2063465544E-01    0.2544330229E+00
0   6   1   2   3   4   5   6
1 1 0 2H
RU 16O 12C 12C 12C 12C 12C 12C 16O 1H 1H 1H 12C 1H 12C 12C
12C 12C 2H 12C 12C 2H 2H 2H 2H 12C 12C 1H 1H 1H 1H 12C
1H 1H 1H 12C 1H 1H 1H 12C 1H 1H 1H 16O 14N 12C 16O 12C
1H 1H 12C 1H 1H
0   6   1   2   3   4   5   6
STOP

```

Atom specification for the first 3 and last 3 of 53 atoms in the structure

First 2 and last 2 of 4240 lines of Hessian

Isotopologue specification and program termination

**Figure 1A.** LIPFR input for a ruthenium-catalysed CH activation intermediate isotope effect calculation. Note the proportionally larger Hessian when compared to the methyl cage systems in the test cases. The majority of the Hessian has been omitted for clarity in the Figure.

***** LIPFR: 13 May 2016 *****						Output header
***** Calculations for isotopologue 1 *****						
RU 4ts						Input coordinates for the first 3 and last 3 of 53 atoms in the structure
Symmetry number = 1            Electronic degeneracy = 1						
Atom	Mass	X	Y	Z (Ångström)		
1 RU	101.070000	-0.97227300	-0.60657500	0.15603500		
2 16O	15.994915	-1.89397900	-0.77329000	2.09390400		
3 12C	12.000000	-1.04851700	-1.67960200	2.44884400		
&&&						
51 12C	12.000000	2.01294100	3.71553400	1.32702200		
52 1H	1.007825	1.65396200	4.70747600	1.02534500		
53 1H	1.007825	2.56086500	3.78095500	2.28029700		
The 3 Principal Moments (Amu Å**2) are:						Moments of inertia and output frequencies for isotopologue I.
3672.4089468753            2270.1233143325            3240.9896687397						
Frequencies of normal modes (cm-1)						
-0.0089217    -0.0086011    -0.0053639    -0.0048148    -0.0046730    0.0035700						
19.4747794    26.9175719    29.5518333    44.8478110    49.4660068    58.8431581						
66.1448984    67.9118074    74.1458676    81.1583971    105.9571190    120.7041514						
137.5601225    143.3024938    161.2366827    169.6606620    171.5640406    174.1974955						
194.2479040    217.9399135    221.2668046    244.5828926    261.4612802    282.3144697						
285.7729158    293.1509487    299.0516137    313.2492215    325.0842775    341.9109691						
380.4130085    385.8891193    400.7861417    402.4618372    413.8151361    431.5838832						
440.9686477    476.4137306    491.0476906    507.0651183    549.2680654    553.3448951						
589.9470033    606.4353669    623.3742263    636.9380768    651.6264019    666.1937411						
671.1116044    680.3495082    690.4153023    706.9665936    742.8116179    757.3667453						
793.6110794    813.0830830    866.2092835    869.2923474    873.2448826    880.6479667						
886.1183761    910.2250296    920.5406839    933.6126395    936.3691437    939.5978639						
944.1638245    945.4857000    963.9235669    971.3424521    980.9044688    981.6486739						
982.0771926    987.7321121    1007.9328056    1008.9919679    1022.2858535    1032.7948841						
1041.0752795    1072.9056917    1079.9015806    1098.2957592    1101.7065328    1122.9738163						
1128.0255204    1146.8944125    1151.7363579    1174.1797150    1189.7917646    1195.3285272						
1195.9170372    1210.2253632    1222.1380844    1263.1433607    1282.5186572    1297.3842331						
1311.0097214    1314.2492449    1339.0815385    1348.5410764    1352.6744343    1360.2930306						
1364.5828402    1366.0936603    1378.7873367    1398.3243906    1401.1121690    1418.7892304						
1421.6467662    1426.0497737    1440.7554152    1445.7025288    1447.7447267    1449.0642243						
1453.2572304    1455.7734100    1462.1377402    1471.6270155    1474.5842731    1476.7875131						
1482.8318130    1488.7293875    1493.4347428    1514.6607875    1586.8769805    1596.5027834						
1646.6551901    2967.7715946    2973.8730804    2976.7774434    2979.6983454    2992.3427112						
2995.2982457    3024.3278010    3048.7667962    3053.8156323    3054.1479826    3060.6855940						
3064.2869879    3069.9174367    3072.8927017    3087.2117772    3107.0840486    3107.3856194						
3120.4875978    3127.8894605    3129.4972507    3136.9670425    3137.5225913    3141.5871565						
3145.2031936    3147.1835403    3162.3814663						
The 6 rejected frequencies are:						Zero-valued frequencies for translation and rotation
-0.0089    -0.0086    -0.0054    -0.0048    -0.0047    0.0036						
EXC and ZPE evaluated over 3N-6 frequencies						Thermodynamic properties
Temperature = 418.15 K, Std. State = 1 Atmosphere						
-(G-PE)/RT= -244.759215 = -(G-E0)/RT= 75.850597    -(ZPE/RT= 320.609812)						
TRANS            ROT            EXC            ELEC						
20.612394    16.852961    38.385242    0.000000						
The 6 rejected frequencies are:						
-0.0089    -0.0086    -0.0054    -0.0048    -0.0047    0.0036						
***** Calculations for isotopologue 2 *****						Coordinates for first 2 and last 2 atoms of isotopologue 2: these are identical to those of isotopologue I.
2H						
Symmetry number = 1            Electronic degeneracy = 1						
Atom	Mass	X	Y	Z (Ångström)		
1 RU	101.070000	-0.97227300	-0.60657500	0.15603500		
2 16O	15.994915	-1.89397900	-0.77329000	2.09390400		
&&&						
52 1H	1.007825	1.65396200	4.70747600	1.02534500		
53 1H	1.007825	2.56086500	3.78095500	2.28029700		

