Supplementary Information – HyDRA challenge

Narasimhan Viswanathan [ORCID: 0000-0003-0369-3853][1], Thomas Nevolianis [ORCID: 0000-0003-4478-5842][2], Gabriel Rath [ORCID: 0000-0003-0657-629X][3], Wassja A Kopp [ORCID: 0000-0001-8147-3464][4], and Kai Leonhard [ORCID: 0000-0001-6231-6957][5]

[1]Institute of Technical Thermodynamics, RWTH Aachen University, Schinkelstraße 8, 52072, Aachen, Germany

1 Computational Details

1.1 Computational Details of conformation space search

From the given list of molecules in the test set, a comformer space search was conducted using CREST 2.11[DOI: 10.1039/C9CP06869D, 10.1021/acs.jctc.9b00143, 10.1039/d1sc00621e] at the GFN2-xTB[DOI: 10.1021/acs.jctc.8b01176] level with default settings. From the conformers that were obtained, the twenty lowest energy conformers were chosen for further calculations. These selected conformers were then re-optimized with b3lyp/def-TZVP level of theory with Resolution of Identity (RI) approximation and d3 dispersion settings in TURBOMOLE 7.4.1 [DOI: 10.1016/0009-2614(95)00621-A]. Following this, the conformers were optimized at a higher level of b2plyp/aug-cc-pvdz [DOI: 10.1063/1.2712433] theory, with Grimme's dispersion (gd3bj key word in Gaussian) [DOI: /10.1063/1.3382344] and an extremely fine integral grid (int=grid=199974) in gaussian/16.b01. Frequencies and geometries were compared to check for symmetrical conformers. Following this, the lowest energy conformer was chosen for further calculations.

1.2 Production runs for optimization and frequency computation

The lowest energy conformer obtained was then re-optimized at the b2plyp/aug-cc-pvtz level of theory with the above mentioned keywords for dispersion, grid, and SCF convergence. In addition to these keywords, for b2plyp/aug-cc-pvtz calculations, optimization was carried out after reading the force constants from the previous b2plyp/aug-cc-pvdz calculation checkpoint file. The frequency calculation at b2plyp/aug-cc-pvtz level was carried out with selected anharmonic modes for the cluster. These modes were chosen based on visual inspection of normal modes mainly involving motion of the donor hydrogen. We optimized and calculated the frequencies for the clusters of water and one of the four molecules: FAH, THF; THT and TFE. The anharmonic modes chosen for each cluster is summarised in the following table.

Cluster	Selected modes for frequency calculation
FAH	4-6,10,14-15
THF	2,4-7,9,32,41-42
THT	1-5,7,10,32,41-42
TFE	1,4,7-8,14,25,28-30

Isotope effects on harmonic frequencies were calculated quickly without having to redo the Hessian matrix calculation. We adjusted the atomic masses in the input geometry and read in the force constants from the checkpoint file using freq=readfc. Results have been produced for all hydrogen atoms attached to an oxygen atom replaced by deuterium and for water oxygen replaced by ^{18}O.

2 Additional computed data (optional)

2.1 Symmetric OH-stretching fundamentals for test set

Code	CAS	Harmonic (cm^{-1})	Anharmonic (cm^{-1})	Anharmonic downshift (cm^{-1})
FAH	50-0-0	3701.2950	3555.54	87.18
THF	109-99-9	3606.5410	3450.70	192.02
THT	110-01-0	3623.1011	3447.20	195.52
TFE	75-89-8	3801.0626	3288.99	353.73

2.2 IR intensities and Raman scattering activities

Cluster	Harmonic intensities (km/mol)	Anharmonic intensities (km/mol)
FAH	237.00	1998.57
THF	462.95	227.04
TFE	15.01	99394.44
THT	380.20	346.28

The above table shows the harmonic and anharmonic intensities of symmetric OH bond stretch. For IR intensities of other modes, please check the gaussian .log files uploaded.

2.3 Isotopologue information

The following table shows the symmetric bond stretch frequencies of the water organic molecule clusters with hydrogen attached to oxygen atoms replaced by deuterium and the oxygen in water molecule replaced by

^{18}O. More information on the geometries can be found out in the gaussian .log files uploaded.

Name	OD (cm^{-1})	^{18}O (cm^{-1})
FAH	1204.62	1644.69
TFE	1206.12	1648.02
THT	1199.75	1622.63
THF	1199.75	1636.78

2.4 Relative energies for local minima and spectral properties

The following table shows the energy of conformers for the hydrated molecule clusters at b2plyp/aug-cc-pvdz level.

Name	Number of Conformers	Energy (Hartree/particle)
FAH	Conformer 1	-190.623567
	Conformer 2	-190.618818
THF	Conformer 1	-308.632910
	Conformer 2	-308.624238
	Conformer 3	-308.632976
THT	Conformer 1	-631.218038
		-631.218010
TFE	Conformer 1	-528.461668
	Conformer 2	-528.467446
	Conformer 3	-528.460591
	Conformer 4	-528.461157
	Conformer 5	-528.467484

In the above table, we can see that some energies are very close to each other;after analysing the geometry and frequencies, we discarded the duplicates and the calculations at the level of b2plyp/aug-cc-pvtz level was carried out.

The final geometry file of lowest energy conformer can be found in the repository of files provided.

2.5 Other computed quantities for the training and test sets

Frequency of water molecule and acetone water 1:1 cluster at b2pylp/aug-cc-pvdz and aug-cc-pvtz levels.

Method	Harmonic Frequency (cm^{-1})	Anharmonic Frequency (cm^{-1})
Water b2plyp/aug-cc-pvdz	3927.15	3735.13
Water b2plyp/aug-cc-pvtz	3812.49	2642.72
Deuterated Water	2747.84	—
^{18}O Water	3804.51	—
Acetone b2plyp/aug-cc-pvdz	3631.495	3463.464
Acetone b2plyp/aug-cc-pvtz	3634.63	3481.55

The calculation files for the above mentioned calculations are shared with Gigamove 2.0 repository. The following is the link for the same: https://gigamove.rwth-aachen.de/en/download/13c2f50fd0b2a30e828553ae0d06b6bc
This link is valid for fourteen days (till 09.03.2022).