Electronic Supplementary Information for: Setting up the HyDRA blind challenge for the microhydration of organic molecules

Taija L. Fischer, Margarethe Bödecker, Anne Zehnacker-Rentien, Ricardo A. Mata, Martin A. Suhm

Profiles of the ten monohydrates and their acceptor molecules (Table S1, Figure S1) chosen as the training set for the HyDRA challenge. The profiles include some key references and routinely computed and experimental properties which can be helpful in assessing the results of a chosen strategy to predict the hydrogen-bonded water OH stretching fundamental. For the computations, the B3LYP method with the D3 dispersion correction including three body term (abc) is combined with the def2-TZVP basis set in the double-harmonic approximation, using the ORCA quantum chemistry package. The training set was selected from proposals listed under qmbench.net, The proponents are listed in the profiles and the estimated wavenumber accuracy is based on their information.

We do not claim that the monohydrate structures shown in the profiles are the global minimum arrangements, but they correspond to the lowest structure found for the employed computational approach and are consistent with the indicated literature data.

Table S1: Alphabetically ordered acceptor molecules with their three-letter acronym and CAS registry number.

acceptor molecule	abbreviation	CAS RN
Acetone	ACE	67-64-1
Acetophenone	APH	98-86-2
Aniline	ANL	62-53-3
Cyclobutanone	CBU	1191-95-3
Dibenzofuran	DBF	132-64-9
Di-*tert*-butyl nitroxide	DBN	2406-25-9
Imidazole	IMZ	288-32-4
o-Cyanophenol	OCP	611-20-1
1-Phenylethanol	POH	98-85-1
1,2,4,5-Tetrafluorobenzene	TFB	327-54-8

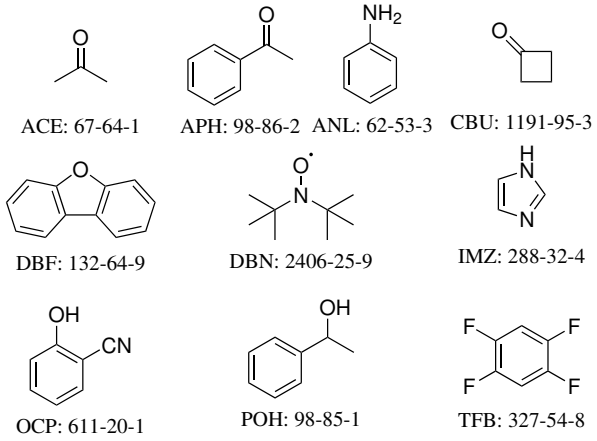
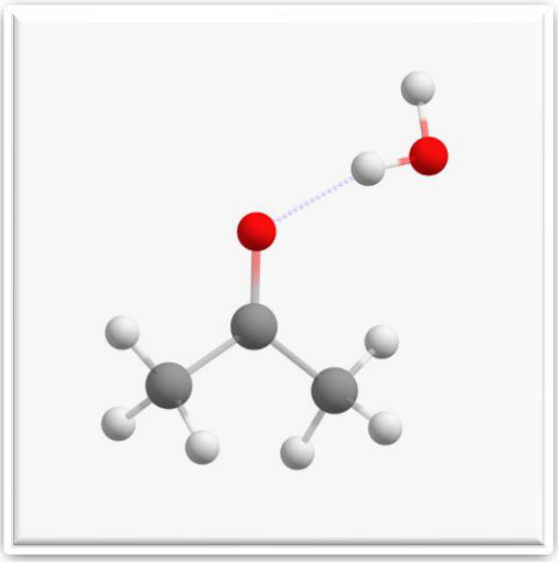


Figure S1: Alphabetically ordered structural formulas of the ten acceptor molecules from Table S1.

Acetone



CAS-No.	67-64-1
molecular formula	C_3H_6O
molecular weight /g mol^{-1}	58.08
calc. lowest monomer wavenumber/cm^{-1}	14*

Dimer:

exp. OH_b wavenumber / cm^{-1}	3538[1]
estimated accuracy / cm^{-1}	<1
calc. OH_b wavenumber / cm^{-1}	3613*
calc. OH_b band strength / km mol^{-1}	452*

Proposing and supporting researchers:

Taija L. Fischer Hannes C. Gottschalk Till Wagner
Martin A. Suhm Arman Nejad

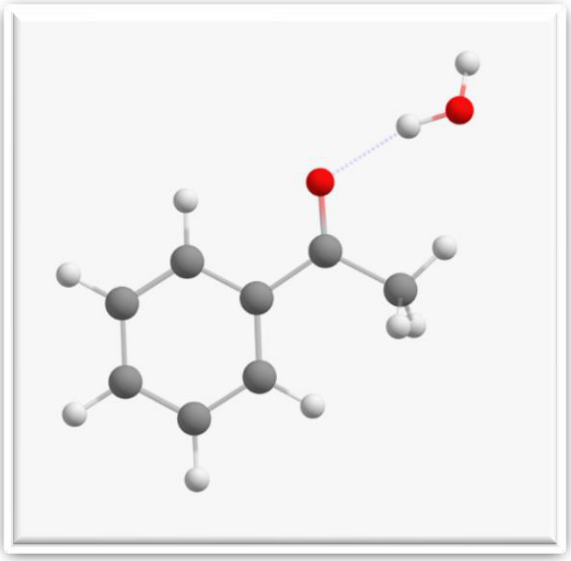
References:

[1-IR] Fischer, T. L.; Wagner, T.; Gottschalk, H. C.; Nejad, A.; Suhm, M. A. A Rather Universal Vibrational Resonance in 1:1 Hydrates of Carbonyl Compounds, *J. Phys. Chem. Lett.* **2021**, *12(1),* 138−144. (DOI: 10.1021/acs.jpclett.0c03197)

[2-IR@Matrix] Zhang, X. K.; Lewars, E. G.; March, R. E.; Parnis, J. M. Vibrational spectrum of the acetone-water complex: a matrix isolation and FTIR and the theoretical study, *J. Phys. Chem.* **1993**, *17*, 4320−4325. (DOI: 10.1021/j100119a012)

* Calculation Method: B3LYP abc D3BJ def2-TZVP double harmonic

Acetophenone



CAS-No.	98-86-2
molecular formula	C_8H_8O
molecular weight /g mol^{-1}	120.15
calc. lowest monomer wavenumber/cm^{-1}	27*

Dimer:

exp. OH_b wavenumber / cm^{-1}	3536[1]
estimated accuracy / cm^{-1}	<1
calc. OH_b wavenumber / cm^{-1}	3609*
calc. OH_b band strength / km mol^{-1}	633*

Proposing and supporting researchers:

Taija L. Fischer Hannes C. Gottschalk Till Wagner
Martin A. Suhm Arman Nejad

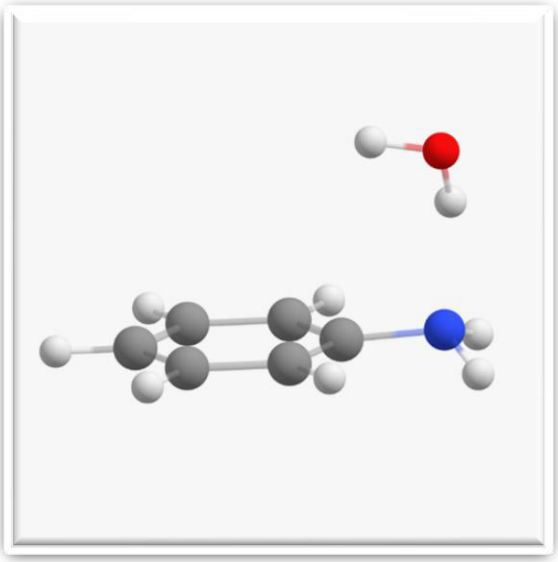
References:

[1-IR] Fischer, T. L.; Wagner, T.; Gottschalk, H. C.; Nejad, A.; Suhm, M. A. A Rather Universal Vibrational Resonance in 1:1 Hydrates of Carbonyl Compounds, *J. Phys. Chem. Lett.* **2021**, *12(1),* 138−144. (DOI: 10.1021/acs.jpclett.0c03197)

[2-Microwave] Lei, J.; Zhang, J.; Feng, G.; Grabow, J.-U.; Gou, Q. Conformational preference determinded by inequivalent n-pairs: rotational studies on acetophenone and its monohydrate, *Phys. Chem. Chem. Phys.* **2019**, *21,* 22888−22894. (DOI: 10.1039/C9CP03904)

* Calculation Method: B3LYP abc D3BJ def2-TZVP double harmonic

Aniline



CAS-No.	62-53-3
molecular formula	C_6H_7N
molecular weight /g mol^{-1}	93.13
calc. lowest monomer wavenumber/cm^{-1}	220*

Dimer:

exp. OH_b wavenumber / cm^{-1}	3524[1]
estimated accuracy / cm^{-1}	1
calc. OH_b wavenumber / cm^{-1}	3637*
calc. OH_b band strength / km mol^{-1}	314*

Proposing and supporting researchers:

José A. Fernández
Iker León
Imanol Usabiaga

References:

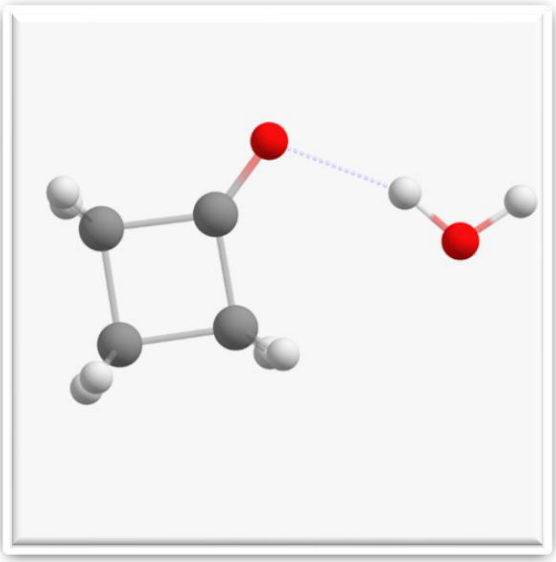
[1-IR-UV] Leon, I.; Arnaiz, P.F.; Usabiaga, I.; Fernandez, J. A. Mass resolved IR spectroscopy of anline-water aggregates, *Phys. Chem. Chem. Phys.* **2016**, *18,* 27336. (DOI: 10.1039/c6cp04373a)

[2-Microwave] Spoerel, U.; Stahl, W. The Aniline-Water Complex, *J. Mol. Spectrosc.* **1998**, *190,* 278−289. (DOI: 10.1006/jmsp.1998.7600)

[3-REMPI] Piani, G.; Pasquini, M.; López-Tocón, I.; Pietraperzia, G.; Becucci, M.; Castellucci, E. The aniline-water and aniline-methanol complexes in the S_1 excited state, *Chem. Phys.* **2006**, *330,* 138−145. (DOI: 10.1016/j.chemphys.2006.08.004)

* Calculation Method: B3LYP abc D3BJ def2-TZVP double harmonic

Cyclobutanone



CAS-No.	1191-95-3
molecular formula	C_4H_6O
molecular weight /g mol^{-1}	70.09
calc. lowest monomer wavenumber/cm^{-1}	402*

Dimer:

exp. OH_b wavenumber / cm^{-1}	3548[1]
estimated accuracy / cm^{-1}	<1
calc. OH_b wavenumber / cm^{-1}	3636*
calc. OH_b band strength / km mol^{-1}	362*

Proposing and supporting researchers:

Taija L. Fischer Hannes C. Gottschalk Till Wagner
Martin A. Suhm Arman Nejad

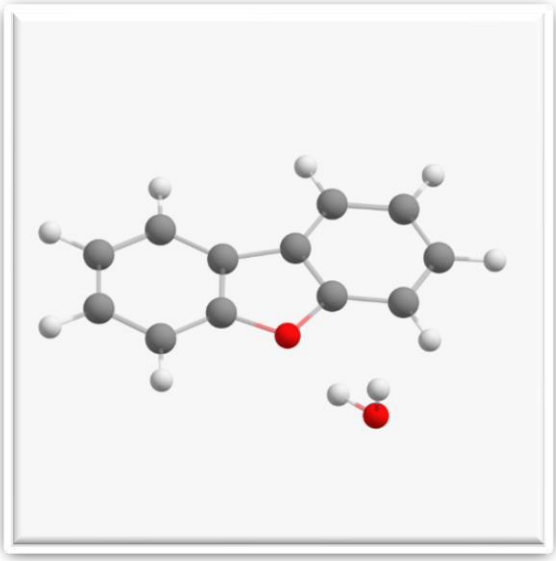
References:

[1-IR] Fischer, T. L.; Wagner, T.; Gottschalk, H. C.; Nejad, A.; Suhm, M. A. A Rather Universal Vibrational Resonance in 1:1 Hydrates of Carbonyl Compounds, *J. Phys. Chem. Lett.* **2021**, *12(1),* 138−144. (DOI: 10.1021/acs.jpclett.0c03197)

[2-Microwave] Melandri, S.; Maris, A.; Giuliano, B. M. ; Caminat, W. Water-ketones hydrogen bonding: The rotational spectrum of cyclobutanone-water, *J. Chem. Phys.* **2005,** *123*, 164304. (DOI: 10.1063/1.2078767)

* Calculation Method: B3LYP abc D3BJ def2-TZVP double harmonic

Dibenzofuran



CAS-No.	132-64-9
molecular formula	$C_{12}H_8O$
molecular weight /g mol^{-1}	168.19
calc. lowest monomer wavenumber/cm^{-1}	105*

Dimer:

exp. OH_b wavenumber / cm^{-1}	3623[1]
estimated accuracy / cm^{-1}	<1
calc. OH_b wavenumber / cm^{-1}	3738*
calc. OH_b band strength / km mol^{-1}	181*

Proposing and supporting researchers:

Dominic Bernhard
Markus Gerhards

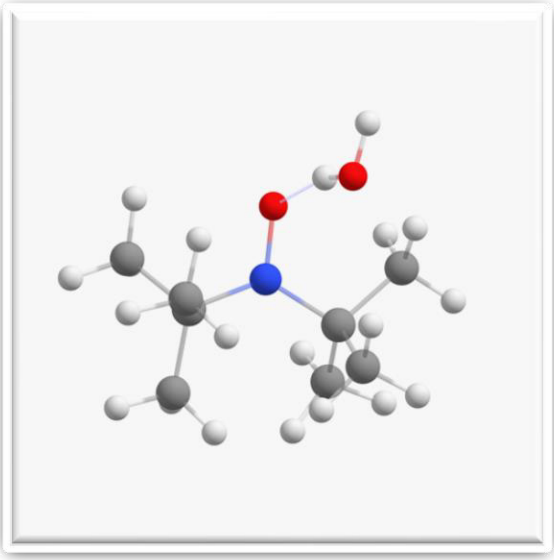
References:

[1-FTIR/IR-UV/Microwave] Bernhard, D.; Fatima, M.; Poblotzki, A.; Steber, A. L.; Pérez, C.; Suhm, M. A.; Schnell, M.; Gerhards, M. Dispersion-controlled docking preference: multi-spectroscopic study on complexes of dibenzofuran with alcohols and water, *Phys. Chem. Chem. Phys.* **2019**, *21,* 16032—16046.

[2-Fluorescence] Auty, A. R.; Jones, A. C.; Philips, D. Fluorescence excitation spectra and decay times of jet-cooled dibenzofuran and the dibenzofuran-water complex, *Chem. Phys. Lett.* **1984**, *112 (6),* 529—533.

* Calculation Method: B3LYP abc D3BJ def2-TZVP double harmonic

Di-*tert*-butyl nitroxide



CAS-No.	2406-25-9
molecular formula	$C_8H_{18}NO$
molecular weight /g mol^{-1}	144.24
calc. lowest monomer wavenumber/cm^{-1}	56*

Dimer:

exp. OH_b wavenumber / cm^{-1}	3484[1]
estimated accuracy / cm^{-1}	2
calc. OH_b wavenumber / cm^{-1}	3578*
calc. OH_b band strength / km mol^{-1}	382*

Proposing and supporting researchers:

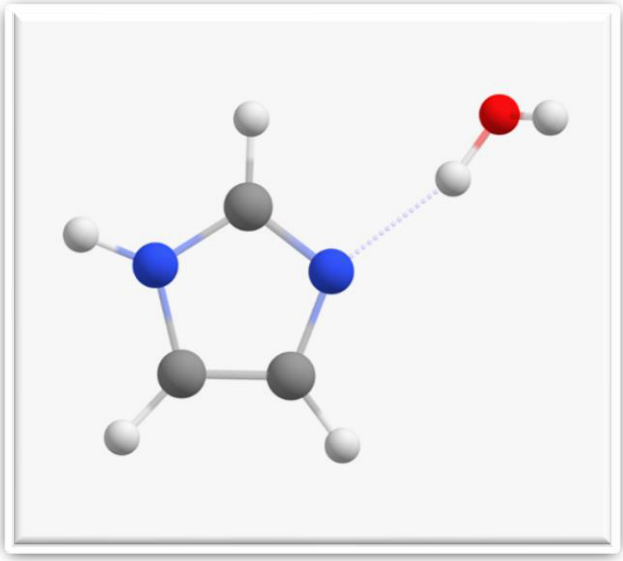
Elisa Bras
Taija L. Fischer
Martin A. Suhm

References:

[1-IR] Brás, E. M.; Fischer, T. L.; Suhm, M. A. The hydrates of TEMPO: Water vibrations reveal radical microsolvation, *Angew. Chem. Int. Ed.* **2021**, *60,* 19013–19017. (DOI: 10.1002/anie.202104496)

* Calculation Method: B3LYP abc D3BJ def2-TZVP UKS double harmonic

Imidazole



CAS-No.	288-32-4
molecular formula	$C_3H_4N_2$
molecular weight /g mol^{-1}	68.08
calc. lowest monomer wavenumber/cm^{-1}	517*

Dimer:

exp. OH_b wavenumber / cm^{-1}	3458[1]
estimated accuracy / cm^{-1}	2
calc. OH_b wavenumber / cm^{-1}	3537*
calc. OH_b band strength / km mol^{-1}	701*

Proposing and supporting researchers:

Julia Zischang
Martin A. Suhm

References:

[1-IR] Zischang, J.; Lee, J. J..; Suhm, M. A. Communication: Where does the first water molecule go in imidazole?, *J. Chem. Phys.* **2011**, *137,* 061102. (DOI: 10.1063/1.3624841)

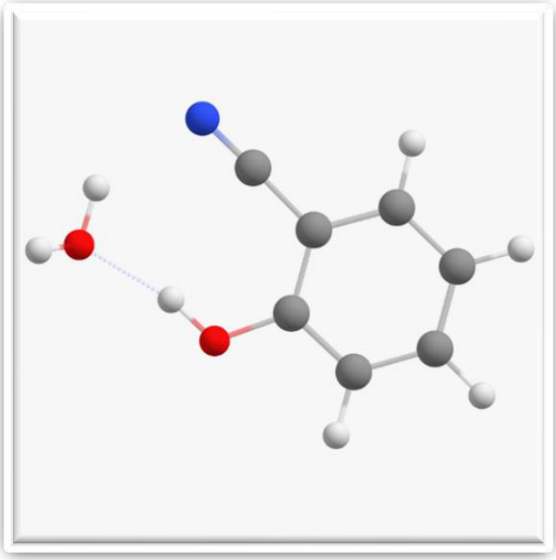
[2-Microwave] Gougoula, E.; Cole, D. J.; Walker, N. R. Bifunctional Hydrogen Bonding of Imidazole with Water Explored by Rotational Spectroscopy and DFT Calculations, *J. Phys. Chem. A* **2020,** *124,* 2649−2659. (DOI: 10.1021/acs.jpca0c00544)

[3-IR@He] Choi, M. Y.; Miller, R. E. Infrared Laser Spectroscopy of Imidazole Complexes in Helium Nanodroplets: Monomer, Dimer, and Binary Water Complexes, *J. Phys. Chem. A* **2006,** 9344−9351. (DOI: 10.1021/jp0624146)

[4-IR@Matrix] Van Bael, M. K.; Smets, J.; Schoone, K.; Houben, L.; McCarthy, W.; Adamowicz, L.; Nowak, M. J.; Maes, G. Matrix-Isolation FTIR Studies and Theoretical Calcilations of Hydrogen-Bonded Complexes of Imidazol. A Comparison between Experimental Results and Different Calculation Methods, *J. Phys. Chem. A* **1997**, *101,* 2397−2413. (DOI: 10.1021/jp963111)

* Calculation Method: B3LYP abc D3BJ def2-TZVP double harmonic

o-Cyanophenol



CAS-No.	611-20-1
molecular formula	C_7H_5ON
molecular weight /g mol^{-1}	119.12
calc. lowest monomer wavenumber/cm^{-1}	126*

Dimer:

exp. OH_b wavenumber / cm^{-1}	3595[1]
estimated accuracy / cm^{-1}	<1
calc. OH_b wavenumber / cm^{-1}	3684*
calc. OH_b band strength / km mol^{-1}	165*

Proposing and supporting researchers:

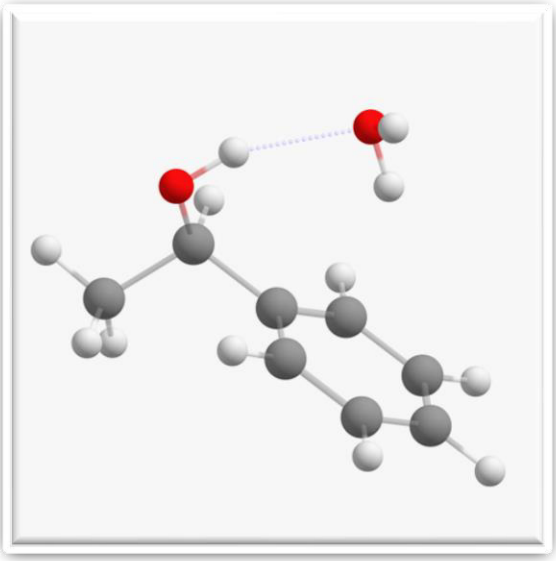
Anne Zehnacker
Michel Broquier
Valérie Brenner

References:

[1-IR] Broquier, M.; Lahmani, F.; Zehnacker-Rentien, A.; Brenner, V.; Millié, Ph.; Peremans, A. Hydrogen-Bonded Bridges in Complexes of *o*-Cyanophenol: Laser-Induced Fluorescence and IR/UV Double-Resonance Studies, *J. Phys. Chem. A* **2001**, *105(28),* 6841−6850**.**

* Calculation Method: B3LYP abc D3BJ def2-TZVP double harmonic

1-Phenylethanol



CAS-No.	98-85-1
molecular formula	$C_8H_{10}O$
molecular weight /g mol^{-1}	122.16
calc. lowest monomer wavenumber/cm^{-1}	32*

Dimer:

exp. OH_b wavenumber / cm^{-1}	3620[1]
estimated accuracy / cm^{-1}	1-2
calc. OH_b wavenumber / cm^{-1}	3743*
calc. OH_b band strength / km mol^{-1}	81*

Proposing and supporting researchers:

Anne Zehnacker
Katia Le Barbu-Debus

References:

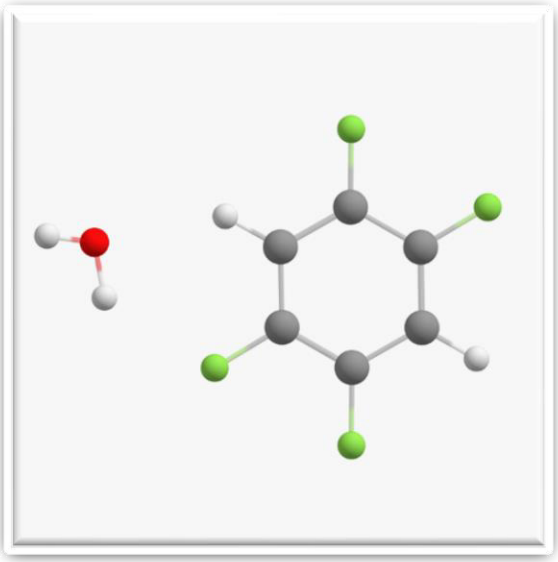
[1-IR-UV] Le Barbu, K.; Lahmani, F., Mons, M.; Broquier, M., Zehnacker, A. IR-UV investigation of the structure of the 1-phenylethanol chromophore and its hydrated complexes, *Phys. Chem. Chem. Phys.* **2001**, *3,* 4684−4688. (DOI: 10.1039/b105036m)

[2-IR] Gottschalk, H. C.; Fischer, T. L.; Meyer, V.; Hildebrandt, R.; Schmitt, U.; Suhm, M. A. A Sustainable Slit Jet FTIR Spectrometer for Hydrate Complexes and Beyond, *Instruments* **2021**, *5,* 12. (DOI: 10.3390/instruments5010012)

[3-R2PI] Giadrini Guioni, A.; Piccirillo, S.; Scuderi, D.; Satta, M.; Di Palma, T. M.; Speranza, M. Chirality and intermolecular forces: studies using R2PI experiments in supersonic beams, *Phys. Chem. Chem. Phys.* **2000**, *2,* 4139−4142. (DOI: 10.1039/b004138f)

* Calculation Method: B3LYP abc D3BJ def2-TZVP double harmonic

1,2,4,5-Tetrafluorobenzene



CAS-No.	327-54-8
molecular formula	$C_6H_2F_4$
molecular weight /g mol^{-1}	150.08
calc. lowest monomer wavenumber/cm^{-1}	129*

Dimer:

exp. OH_b wavenumber / cm^{-1}	3647[1]
estimated accuracy / cm^{-1}	1
calc. OH_b wavenumber / cm^{-1}	3778*
calc. OH_b band strength / km mol^{-1}	11*

Proposing and supporting researchers:

A. Fujii
T. Ebata
V. Venkatesan
N. Mikami

References:

[1-IR-UV] Venkatesan, V.; Fujii, A.; Ebata, T.; Mikami, N. A direct experimental evidence for an aromatic C−H···O hydrogen bond by fluorescence-detected infrared spectroscopy, *Chem. Phys. Lett.* **2004**, *394,* 45−48. (DOI: 10.1016/j.cplett.2004.06.101)

* Calculation Method: B3LYP abc D3BJ def2-TZVP double harmonic

Online submission of theoretical data

One of the main challenges in designing data submission forms for quantum chemical blind-challenges is to find a format capable of accommodating a wide variety of computational approaches. Not only does the primary data need to be harvested, one also needs to provide a framework suited for the different types in dependence of the computational protocol chosen. This is needed to guarantee reproducibility, but also to allow for additional data which might be later of interest to the challenge. We opted for a minimal online form, whereby the data to be provided consisted of:

- **Supplementary Information pdf (mandatory)** – templates were provided both in LaTeX and MS Word format. The participants were asked to provide:
    - **Computational details (mandatory)**, which should include all necessary details for reproducibility.
    - **Additional computed data (optional),** which consisted of a table to be filled in with all OH-stretching fundamentals for the test set, IR intensities and Raman scattering activities, isotopolog information, relative energies for local minima and spectral properties. A final section gave the opportunity to list other computed quantities for both sets (training and test) that the participants wished to share.
- **General information about the submission (mandatory)** – in a free text field the participants were asked to supply the reference value (symmetric OH stretch of water) to which the shift of the monohydrates should be calculated. Some overview of the methods could also be provided in connection to the reference.

The SI pdf had to be uploaded to server side (interface provided), the free text field information was saved in a JSON file (data.json) with an entry number and code which was automatically attributed. There is more information about this file later. For each of the systems, the participants could fill in the following information:

- **Global minimum OHb wavenumber (mandatory)** – the only value that was deemed mandatory. Any system was considered as successfully submitted if this field was filled.
- **Uncertainty (optional)** – the uncertainty associated with the fundamental value. This could be an educated guess, result from some intrinsic method analysis or from the statistics available through the training set.
- **XYZ file (optional)** – structure file from which the estimate was computed. This was an optional field since not all methods are based necessarily on a single structure (e.g., ab initio molecular dynamics)
- **Comment (optional)** – a free text field whereby participants could provide some detail about the specific submission.

The interface for acetone is shown in Figure S1 as an example (before and after submitting data). Once the wavenumber was submitted and changes to the form were saved, the panel would be highlighted clearly showing the participant that the data would be used for the challenge.
The xyz files were named server side following a strict convention, and placed in an individual directory for each entry. The data for each system was included also in the aforementioned data.json file. All of the forms and data operations were coded in PHP. On a less technical note, our design decisions did leave much of the reporting responsibility on the participant side. The guidelines could have gone further, e.g., imposing the uncertainty as mandatory. We hope that after review of this challenge and other initiatives one will be able to improve the model for data acquisition.
One of the main faults raised by the participants is that this “one entry-one data set” model did not allow for the submission of several sets under a single login. We required the participants to register a different email address for each set and fill the forms once again. Some users straightforwardly registered different accounts providing the same SI. In other cases, the groups needed to be contacted to flesh out consistent data sets, as some opted to mix different levels of theory in their “optimal” submission.

Figure S1 - User interface for submitting data to the HyDRA challenge. Top panel: unfilled form for acetone. Bottom panel: after submitting the values, the panel is highlighted. A structure file has been submitted. To avoid issues with the naming nomenclature, the xyz file name is automatically changed server side to the name of the system including the three letter code and the CAS number (Acetone-ACE-67-64-1.xyz).

An example for a data.json file associated with a fictitious entry featuring 3 training and 3 test systems is shown below. It consists of three primary keys “Comments”, “Test” and “Training”. The first corresponds to the comment provided at the beginning of the form (**General information about the submission**), while the other two keys hold the keys-values for the forms of the individual systems (Figure S1).

data.json:


```
{
    "Comments": "The reference OH stretch fundamental for water was computed at the same level of theory, combining
the harmonic computed value and a low-dimensional anharmonic correction as detailed in the SI. The resulting value i
s 3652 cm-1.",
    "Test": {
        "1,3-Dimethyl-2-imidazolidinone (DMI 80-73-9)": {
            "Comment": "",
            "Global minimum OHb wavenumber": 3518,
            "Uncertainty": 10
        },
        "1-Phenylcyclohexane-cis-1,2-diol (PCD 125132-75-4)": {
            "Comment": "There were two isomers close in energy. Data for the second isomer is provided in the SI.",
            "Global minimum OHb wavenumber": 3631,
            "Uncertainty": 10
        },
        "Tetrahydrothiophene (THT 110-01-0)": {
            "Comment": "",
            "Global minimum OHb wavenumber": 3516,
            "Uncertainty": 10
        }
    },
    "Training": {
        "1,2,4,5-Tetrafluorobenzene (TFB 327-54-8)": {
            "Comment": "",
            "Global minimum OHb wavenumber": 3653,
            "Uncertainty": 10
        },
        "1-Phenylethanol (POH 98-85-1)": {
            "Comment": "”,
            "Global minimum OHb wavenumber": 3652,
            "Uncertainty": 10
        },
        "o-Cyanophenol (OCP 611-20-1)": {
            "Comment": "",
            "Global minimum OHb wavenumber": 3581,
            "Uncertainty": 10
        }
    }
}
```