COMMUNICATION

A rotational study of the AlaAla dipeptide

I. León,*[a] E. R. Alonso,[a,b,c] S. Mata[a] and J. L Alonso[a]

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Herein we present the first rotational study of the AlaAla dipeptide, brought into the gas phase by laser ablation. Two different structures have been unveiled in the isolated environment of a supersonic expansion by Fourier transform microwave spectroscopy. These structures have been identified through their rotational and ^{14}N quadrupole coupling constants. The flexibility of the -NH_2 and -COOH ends allow the formation of strong intramolecular interactions giving rise five- and seven-membered ring configurations.

The properties of a complex system are always intimately related to the structure of every single building block that comprises it. An example is proteins: Its structural features and folding mechanism could be understood by the investigation of the structural preferences and intramolecular interactions of small peptides that make them up.[1,2] Therefore, the knowledge of the different geometries that natural peptides, such as di- or tripeptides, adopt is necessary to gain insight on the shape of protein structures, as well as their biological functions. Despite that the solvent is relevant to the biological conditions and that other effects such as charged species play a crucial role, the basic information obtained from smaller systems can provide initial information to understand and model complex biological assembles. Therefore, to gain some understanding about the structure and dynamics of these building blocks in their natural biological conditions, it is helpful to first understand these properties in the absence of intermolecular interactions, avoiding alterations of their intrinsic conformational preferences. Gas-phase experiments provide the isolation conditions necessary to reveal the intrinsic conformational preferences of small dipeptides, in which non-bonded intramolecular interactions play an essential role.

To date, natural peptides derived from aromatic amino acids have been investigated in the gas phase using double-resonance laser spectroscopy.[3–11] Because the majority of these techniques use multiphoton resonantly enhanced ionization, they have been applied to molecules containing UV chromophores. However, experimental studies of non-aromatic peptides containing, for example, amino acids as relevant as glycine, or alanine have not been possible using the above techniques due to the absence of a chromophore. Rotational spectroscopy, considered as the most robust technique to discern among subtle changes in structure and to distinguish between different conformers in the spectra, is not constrained by the need of a chromophore. Laser ablation coupled with Fourier transform microwave spectroscopy carried out in a supersonic expansion allows the rotational study of biomolecules with high meltig point.[12–15] To date, detailed conformational information has been reported on almost all of the proteinogenic amino acids[16] and even some dipeptides[17–20] using this experimental approach.

In the present work, L-Alanyl-L-alanine (AlaAla, see the molecular sketch in Figure 1a) has been submitted to a high-resolution rotational study owing to its key importance as a relevant dipeptide. Ala-Ala is a thermally fragile solid (m.p. 285 ºC) that cannot be brought into the vapor phase by heating methods. We have assembled a chirped-pulse Fourier transform microwave spectrometer (CP-FTMW) with a laser ablation (LA) source to overcome the vaporization difficulties. It has allowed us to transfer the Ala-Ala molecules into the gas-phase, using the third harmonic (355 nm) of a picosecond laser, entrained in a supersonic expansion of neon. Then, a high-power excitation pulse of 200 W polarized the molecules from 3.5 to 8 GHz. Up to 120000 individual free induction decays at a 2 Hz repetition rate were averaged in the time domain, and Fourier transformed to the frequency domain.

A small section of the broadband rotational spectrum is shown in Figure 1a. To ease the identification of Ala-Ala

[a] Grupo de Espectrocopía Molecular (GEM), Edificio Quifima, Laboratorios de Espectroscopia y Bioespectroscopia, Unidad Asociada CSIC, Parque Científico UVa, Universidad de Valladolid, 47011, Valladolid, Spain.
E-mail: iker.leon@uva.es
[b] Instituto Biofisika (UPV/EHU, CSIC), University of the Basque Country, Leioa, E-48940, Spain.
[c] Fundación Biofísica Bizkaia/Biofisika Bizkaia Fundazioa (FBB), Barrio Sarriena s/n, Leioa, E-48940, Spain.
Electronic Supplementary Information (ESI) available. See DOI: 10.1039/x0xx00000x

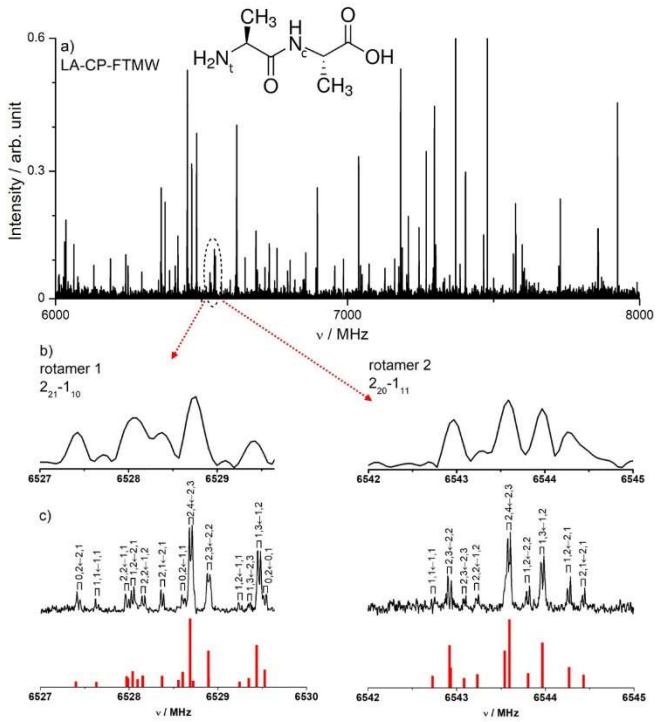

Fig. 1 (a) A small section of the LA-CP-FTMW rotational spectrum of AlaAla dipeptide, with its chemical structural formula shown on top. (b) Some examples of individual transitions of the two detected conformers showing the characteristic hyperfine structure of molecules with nitrogen atoms. (c) The 2_{21}-1_{10} and 2_{20}-1_{11} rotational transitions of the two detected conformers of AlaAla, highlighting their hyperfine structure resolved entirely using the LA-MB-FTMW spectrometer. Each hyperfine component labeled with the corresponding quantum numbers $I',F' \leftarrow I'',F''$ is split by the Doppler effect. The predicted components are also shown at the bottom, showing the excellent agreement.

conformers in the spectrum, we first identified and removed the lines of the common species as well as the known photofragmented products found in previous studies of amino acids and peptides.[17,21,22] Many of the remaining lines were seen to present a not well resolved hyperfine structure arising from nuclear quadrupole coupling interactions such that they might be attributed to AlaAla (see Figure 1b). AlaAla possesses two ^{14}N nuclei (N_c and N_t of the molecular sketch in Figure 1a) with a nuclear quadrupole moment I = 1, which interacts with the electric field gradient created by the rest of the molecule at the nuclei. The ^{14}N nuclear quadrupole coupling splits the rotational energy levels decreasing the overall intensity of each rotational transition and giving rise to a very complex hyperfine structure.[23] All these lines disappeared when the laser ablation was turned off, confirming that they belong to the AlaAla dipeptide. Immediately, this approach allowed us to assign a total of 46 rotational transitions for a first rotameric AlaAla species, labeled as rotamer 1, consisting of a *b*-type R-branch progression and some weaker *c*-type R-branch transitions. No attempt was made to assign the complex hyperfine structure, and only the center of the rotational frequencies was measured. The measured rotational transitions (see Table S01 of the ESI) were fitted to a rigid rotor Hamiltonian[24] to give a preliminary

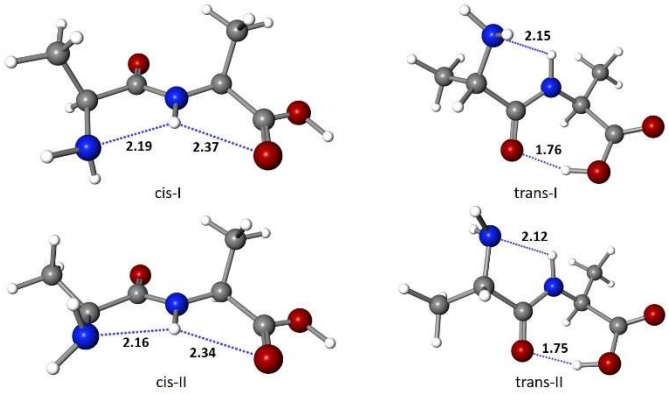

Fig. 2 The four most stable structures of AlaAla calculated at MP2/6-311++G(d,p) level. The bond distances are given in Angstrom.

set of values A=2001.8, B=610.9, and C=523.4 MHz for the rotational constants. After removing all the lines due to this rotamer, it was easy to observe the μ_a-type progressions of a second rotamer. Additionally, μ_b-type transitions were also observed. A total of 42 transitions were measured (see Table S02 of the ESI) and fitted, resulting in the rotational constants A=1965.5, B=644.0, and C=569.8 MHz. Almost no lines remained in the rotational spectrum when subtracting those of both rotamers, confirming that no other species are present.

A more detailed spectroscopic portrait can be obtained analyzing the ^{14}N hyperfine structure of the above rotamers, not well resolved in the broadband spectrum. We measured some selected rotational transitions with the higher resolution of our narrowband LA-MB-FTMW spectrometer to fully resolve the nuclear quadrupole hyperfine structure (see Figure 1c) arising from the two $^{14}N_c$ and $^{14}N_t$ nuclei. The nuclear quadrupole coupling constants (χ_{aa}, χ_{bb}, χ_{cc}) extracted from the analysis provide information on the electronic environment of the nitrogen nuclei, as well as the orientations of the corresponding NH_2 groups. A total of 45 and 49 hyperfine components were measured for rotamers 1 and 2, respectively (see Tables S03 to S04 of the ESI) and analyzed using a Watson´s S-reduced semirigid rotor Hamiltonian in the I_r-representation supplemented with a term to account for the nuclear quadrupole coupling contribution.[23] The quadrupole coupling Hamiltonian was set up in the coupled basis set (I_1I_2IJF), I_1+I_2=I, I+J=F. The energy levels involved in each transition are thus labeled with the quantum numbers J, K_{-1}, K_{+1}, I, and F. Figure 1c shows how the hyperfine structure of the transitions in Figure 1b are now fully resolved thanks to the higher resolution reached using our narrowband LA-MB-FTMW spectrometer. The rotational constants A, B, and C, and the diagonal elements of the nuclear quadrupole coupling tensor (χ_{aa}, χ_{bb}, and χ_{cc}) for the two ^{14}N nuclei are accurately determined and summarized in Table 1.

The procedure to ascribe the two observed rotamers to the corresponding conformers of Ala-Ala begins with the generation of a set of possible conformers. More than 70 structures were obtained within an energetic window of 2500 cm^{-1} in an exhaustive conformational search using molecular

Table 1. Experimental and calculated spectroscopic parameters for the detected conformers of AlaAla together with the calculated energies at the MP2/6-311++G(d,p) and B3LYP-D3(BJ)/def2tvtz using the Grimme's dispersion correction (D3) with Becke-Johnson (BJ) damping for the four most stable conformers.

	Experimental		Theory (MP2/B3LYP-D3(BJ))			
	Rotamer 1	Rotamer 2	*cis-I*	*trans-I*	*cis-II*	*trans-II*
A[a]	2001.80883(47)[g]	1965.49749(43)	1976/1991	1947/1968	2129/2137	1993/2012
B	610.94120(28)	644.00226(41)	613/611	647/645	587/586	644/642
C	523.42105(16)	569.81025(13)	529/524	576/569	504/502	551/546
$\lvert\mu_a\rvert$	Not observed	observed	0.1/0.2	5.8/6.0	0.1/0.1	5.9/6.0
$\lvert\mu_b\rvert$	observed	observed	1.9/1.9	3.2/3.3	1.8/1.8	3.7/3.7
$\lvert\mu_c\rvert$	observed	Not observed	0.8/0.7	0.5/0.3	1.0/1.1	1.0/0.9
N_{NH}/χ_{aa}	2.3061(48)	1.5850(46)	2.36/2.37	1.59/1.63	2.40/2.42	1.71/1.7
N_{NH}/χ_{bb}	0.5706(55)	0.6004(46)	0.61/0.60	0.64/0.71	1.06/1.10	0.41/0.4
N_{NH}/χ_{cc}	-2.8768(55)	-2.1854(46)	-2.97/-2.97	-2.23/-2.35	-3.46/-3.53	-2.12/-2.2
N_{NH2}/χ_{aa}	-2.1839(53)	-0.9778(75)	-2.24/-2.27	-0.83/-0.93	-4.19/-4.39	-3.68/-3.8
N_{NH2}/χ_{bb}	2.6833(58)	-0.2737(78)	2.75/2.88	-0.63/-0.51	2.17/2.23	1.96/2.0
N_{NH2}/χ_{cc}	-0.4993(58)	1.2515(78)	-0.51/-0.61	1.46/1.45	2.01/2.15	1.72/1.8
σ[b]	2.8	3.0				
N[c]	45	49				
ΔE[d]			125/602	0/0	393/790	221/153
ΔE_{ZPE}[e]			0/435	91/0	246/589	269/121
ΔG[f]			0/130	404/0	207/260	503/66

[a]*A*, *B*, and *C* represent the rotational constants (in MHz); μ_a, μ_b, and μ_c are the electric dipole moment components (in D); χ_{aa}, χ_{bb}, and χ_{cc} are the nuclear quadrupole coupling constants of each nitrogen atom; [b]RMS deviation of the fit (in kHz). [c]Number of measured transitions. [d]Calculated relative energies (in cm^{-1}) at the MP2/B3LYP-D3(BJ) level of theory respect to the global minimum. [e] Calculated relative energies (in cm^{-1}) at MP2/B3LYP-D3(BJ) respect to the global minimum, taking into account the zero-point energy (ZPE). [f] Calculated Gibbs energies (in cm^{-1}) at MP2/B3LYP-D3(BJ) and 298 K. [g]Standard error in parentheses in units of the last digit.

mechanics methods.[25] Then *ab initio* calculations at MP2[26]/6-311++G(d,p)[27] and B3LYP-D3(BJ)/def2tvtz[28–30] using the Grimme's dispersion correction (D3) with Becke-Johnson damping(BJ)[31] were employed to reoptimize all the structures, resulting in ten conformers below 1000 cm^{-1} in good agreement with Ref. [32], with the four structures bellow 500 cm^{-1} depicted in Figure 2. Their predicted spectroscopic parameters relevant for conformational identification, namely, the rotational and quadrupole coupling constants and the electric dipole moment components along the principal inertial axes are also collected in Table 1 for comparison. The comparison between experimental and calculated values found in Table 1, unambiguously allow us to assign the rotamers 1 and 2 as conformers *cis-I* and *trans-I*, respectively. Note the significant differences in the quadrupole coupling constants between *cis-I* and *cis-II* (and *trans-I* vs. *trans-II*), which allows our technique to discern, not only the conformers but also the orientation of the NH_2 group.

The absence of the low-energy conformers *cis-II* and *trans-II* in the supersonic expansion could be due to conformational relaxation processes. Rotation about the C-N_t terminal bond connects both structures into *cis-I* and *trans-I,* respectively. Conformational relaxation takes place through collisions with the noble carrier gas in the adiabatic expansion when the interconversion barrier is low enough. By examining the behavior of a variety of systems, Ruoff et al.[33,34] noted that for conformational species separated by an energy barrier lower than 400 cm^{-1}, efficient relaxation occurs. Similar behavior has been observed for proteinogenic amino acids.[31] To test whether conformer interconversion is taking place, we have examined the interconversion barriers. As can be seen in Figure 3, the calculated interconversion barriers between structures *cis-II* and *cis-I* and structures *trans-II* and *trans-I* are about 100 cm^{-1} and could explain the absence of *cis-II* and *trans-II* in the supersonic expansion.

To evaluate the intramolecular interactions in the two observed conformers we conducted an NCIplot analysis,[36,37] which shows a visual representation of the intramolecular interactions and their strength (see Figure S01). Conformer *cis-I* adopts a cis-carboxylic arrangement and is stabilized by two intramolecular hydrogen bonds, N-H···O=C and N-H···N, closing two five-membered rings. The N-H···N hydrogen bond is considerably stronger than the N-H···O=C interaction and the bond distances are 2.19 Å and 2.37 Å, respectively. Conformer *trans-I* adopts a trans-carboxylic disposition and is stabilized by an N-H···N interaction and an O-H···O=C interaction forming a five- and seven-membered rings, respectively. Both interactions have a relatively strong hydrogen bond, with the N-H···N interaction having a bond distance of 2.15 Å and that of the O-H···O=C interaction being 1.76 Å. Interestingly, these conformers nicely correlate with conformers *pI* and *npII* of GlyGly.[16] A third conformer (*pII*) that was observed in GlyGly dipeptide is not detected in AlaAla, which seems to indicate the critical role of the steric restrictions imposed by the methyl groups in AlaAla dipeptide. The *p*II conformer of GlyGly is similar to the *cis* conformer but the N-H···N intramolecular interaction is replaced by a bifurcated N-H··· O=C interaction. This structure is clearly unfavourable in Ala-Ala due to the methyl groups restrictions, forcing the molecule to adopt a distorted structure. Consecuently, there is no longer a bifurcatated N-H···O=C

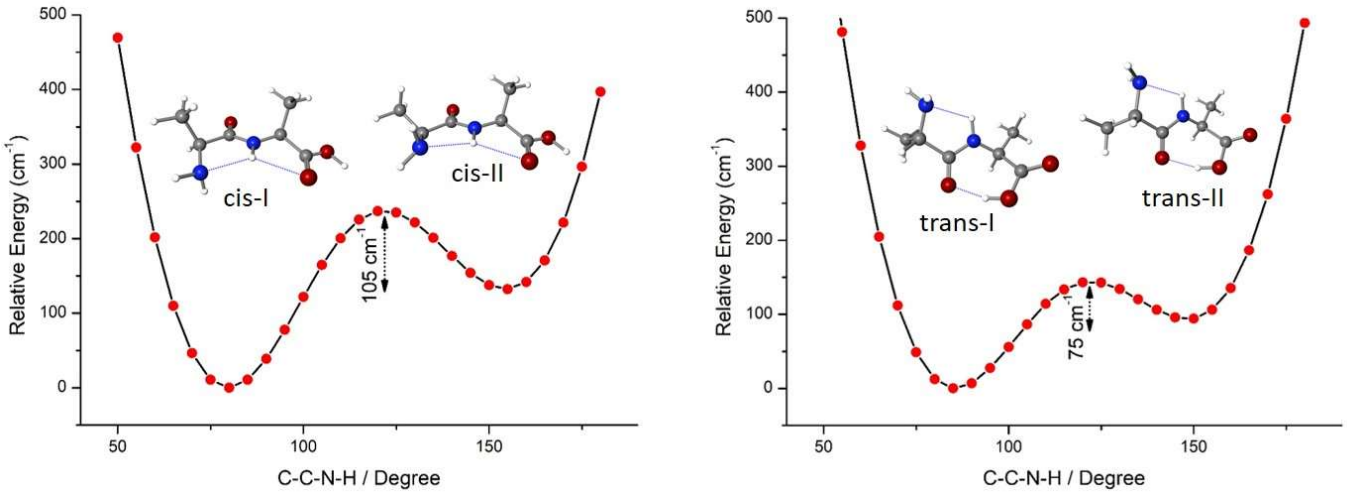

Fig. 3 The calculated relaxed potential energy surface (PES) varying the ∠ C-C-N-H dihedral angle. As can be seen, structures *cis-II* and *trans-II* experience conformer interconversion into structures *cis-I* and *trans-I*, respectively. The calculations were done using B3LYP-D3(BJ)/6-311++G(d,p).

interaction but a weaker simple interaction, and the relative stability of such structure is calculated to be at ~1000cm^{-1}.

Finally, our results also serve to benchmark theoretical calculations. Aside from the rotational and quadrupole coupling constants, our experiments can also provide information on the relative stability of the rotamers from relative intensity measurements of selected transitions (see the SI for a detailed information). The population ratio of the detected species sets conformer *trans-I* ~130 cm^{-1} higher than *cis-I* or, in other words, the population of *cis-I* is about 2 times larger than *trans-I*. The *ab initio* MP2 predicts the *cis-I* structure as the most stable candidate and the *trans-I* structure only slightly above in good agreement with our results. On the other hand, DFT methods such as B3LYP-D3(BJ) (see Table 1 and Ref. [32]), predict the *trans-I* structure as the global minimum and *cis-I* 400 cm^{-1} higher in energy, clearly far from our observations. Furthermore, DFT methods fail to reproduce the correct energetic order even taking into account the temperature prior to the expansion. For GlyGly dipeptide, the opposite occurs, and B3LYP-D3(BJ) works better than MP2.[16] The spectroscopic parameters in Table 1 as well as the structural comparison in Figure S2 using MP2 and B3LYP-D3(BJ) show a similar structure and, while MP2 reproduces better the energetics and quadrupole coupling constants, DFT gives better rotational constants. It highlights the importance of accurate experimental methods to validate current theoretical methods.

In summary, two conformers of the AlaAla dipeptide have been fully using high-resolution rotational spectroscopy coupled with laser ablation sources. These conformers show different intramolecular hydrogen bond interactions forming five- and seven-membered ring configurations. The methyl groups in AlaAla dipeptide contribute to reduce the conformational panorama. Present results consolidate the power of the LA-CP-FTMW and LA-MB-FTMW techniques to characterize the structural behavior of peptides.

Conflicts of interest

There are no conflicts to declare.

Acknowledgments

The financial fundings from Ministerio de Ciencia e Innovación (CTQ2016-76393-P), Junta de Castilla y León (Grant VA077U16) and the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC-2013-SyG, Grant Agreement n. 610256 NANOCOSMOS, are gratefully acknowledged. E.R.A. was supported by the Fundación Biofísica Bizkaia (Spain).

Notes and references

1 H. A. Lehninger, D. Nelson and M. Cox, *Lehninger Principles of Biochemistry*, W. H. Freeman, New York, 2008.
2 Y. Park and V. Helms, *Bioinformatics*, 2007, **23**, 701–708.
3 T. Häber, K. Seefeld, G. Engler, S. Grimme and K. Kleinermanns, *Phys. Chem. Chem. Phys.*, 2008, **10**, 2844–51.
4 T. Häber, K. Seefeld and K. Kleinermanns, *J. Phys. Chem. A*, 2007, **111**, 3038–3046.
5 I. Hünig, K. A. Seefeld and K. Kleinermanns, *Chem. Phys. Lett.*, 2003, **369**, 173–179.
6 J. W. E. and and D. H. Levy*, , DOI:10.1021/JP982545+.
7 W. Chin, I. Compagnon, J. P. Dognon, C. Canuel, F. Piuzzi, I. Dimicoli, G. Von Helden, G. Meijer and M. Mons, *J. Am. Chem. Soc.*, 2005, **127**, 1388–1389.
8 D. Toroz and T. van Mourik, *Mol. Phys.*, 2006, **104**, 559–570.
9 A. Baran Mandal and R. Jayakumar, .

10 I. Hünig and K. Kleinermanns, *Phys. Chem. Chem. Phys.*, 2004, **6**, 2650–2658.
11 J. M. Bakker, C. Plutzer, I. Hünig, T. Häber, I. Compagnon, G. Von Helden, G. Meijer and K. Kleinermanns, *ChemPhysChem*, 2005, **6**, 120–128.
12 A. Lesarri, S. Mata, E. J. Cocinero, S. Blanco, J. C. López and J. L. Alonso, *Angew. Chemie Int. Ed.*, 2002, **41**, 4673–4676.
13 J. L. Alonso, M. a. Lozoya, I. Peña, J. C. López, C. Cabezas, S. Mata and S. Blanco, *Chem. Sci.*, 2014, **5**, 515.
14 E. R. Alonso, I. Peña, C. Cabezas and J. L. Alonso, *J. Phys. Chem. Lett.*, 2016, **7**, 845–850.
15 E. R. Alonso, I. León, L. Kolesniková and J. L. Alonso, *J. Phys. Chem. A*, 2019, **123**, 2756–2761.
16 I. León, E. R. Alonso, S. Mata, C. Cabezas and J. L. Alonso, *Angew. Chemie - Int. Ed.*, 2019, **58**, 16002–16007.
17 C. Cabezas, M. Varela and J. L. Alonso, *Angew. Chemie Int. Ed.*, 2017, **129**, 6520–6525.
18 C. Cabezas, M. Varela, V. Cortijo, A. I. Jiménez, I. Peña, A. M. Daly, J. C. López, C. Cativiela and J. L. Alonso, *Phys. Chem. Chem. Phys.*, 2013, **15**, 2580.
19 C. Cabezas, M. A. T. Robben, A. M. Rijs, I. Peña and J. L. Alonso, *Phys. Chem. Chem. Phys.*, 2015, **17**, 20274–20280.
20 I. León, E. R. Alonso, S. Mata, C. Cabezas, M. A. Rodríguez, J.-U. Grabow and J. L. Alonso, *Phys. Chem. Chem. Phys.*, 2017, **19**, 24985–24990.
21 I. Peña, C. Cabezas and J. L. Alonso, *Angew. Chemie - Int. Ed.*, 2015, **54**, 2991–2994.
22 C. Bermúdez, S. Mata, C. Cabezas and J. L. Alonso, *Angew. Chemie - Int. Ed.*, 2014, **53**, 11015–11018.
23 W. Gordy and R. L. Cook, *Microwave molecular spectra*, Wiley, New York, 1984.
24 H. M. Pickett, *J. Mol. Spectrosc.*, 1991, **148**, 371–377.
25 2018. Schrödinger Release 2018-3: Maestro Schrödinger, LLC, New York, NY, .
26 C. Møller and M. S. Plesset, *Phys. Rev.*, 1934, **46**, 618–622.
27 M. J. Frisch, J. A. Pople and J. S. Binkley, *J. Chem. Phys.*, 1984, **80**, 3265–3269.
28 A. D. Becke, *Phys. Rev. A*, 1988, **38**, 3098–3100.
29 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 1372–1377.
30 Y. Zhao and D. G. Truhlar, *J. Chem. Theory Comput.*, 2007, **3**, 289–300.
31 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, , DOI:10.1063/1.3382344.
32 V. Yatsyna, R. Mallat, T. Gorn, M. Schmitt, R. Feifel, A. M. Rijs and V. Zhaunerchyk, *Phys. Chem. Chem. Phys.*, 2019, **21**, 14126–14132.
33 R. S. Ruoff, T. D. Klots, T. Emilsson, H. S. Gutowsky and) T Emilsson, *J. Chem. Phys.*, 1990, **93**, 3142–3150.
34 P. D. Godfrey, R. D. Brown and F. M. Rodgers, *J. Mol. Struct.*, 1996, **376**, 65–81.
35 J. L. Alonso, I. Peña, J. C. López, E. R. Alonso and V. Vaquero, *Chem. - A Eur. J.*, 2019, **25**, 2288–2294.
36 E. R. Johnson, S. Keinan, P. Mori-Sánchez, J. Contreras-García, A. J. Cohen and W. Yang, *J. Am. Chem. Soc.*, 2010, **132**, 6498–6506.
37 J. Contreras-García, E. R. Johnson, S. Keinan, R. Chaudret, J. P. Piquemal, D. N. Beratan and W. Yang, *J. Chem. Theory Comput.*, 2011, **7**, 625–632.