International Journal of Mass Spectrometry 297 (2010) 152–161

Contents lists available at ScienceDirect

International Journal of Mass Spectrometry

journal homepage: www.elsevier.com/locate/ijms

Assessment of density functionals for predicting the infrared spectrum of sodiated octa-glycine

David Semrouni [a], Carine Clavaguéra [a,*], Jean-Pierre Dognon [b], Gilles Ohanessian [a,*]

[a] Laboratoire des Mécanismes Réactionnels, Département de Chimie, Ecole Polytechnique, CNRS, 91128 Palaiseau Cedex, France
[b] CEA/SACLAY, UMR 3299 CEA/CNRS SIS2M, Laboratoire de chimie de coordination des éléments f, F-91191 Gif-sur-Yvette, France

ARTICLE INFO

Article history:
Received 19 June 2010
Received in revised form 28 July 2010
Accepted 4 August 2010
Available online 24 August 2010

Density functional theory
Infrared spectra
Peptides

ABSTRACT

The sodiated peptide GGGGGGGG–Na^+ or G_8–Na^+ has a remarkable structure with a highly coordinated sodium ion and an acidic OH that is strongly hydrogen-bound with the N-terminus. The presence of the sodium ion makes this hydrogen bond unusually strong and makes proton transfer easy, leading to an equally stable, salt bridge isomer. The performances of a variety of density functionals in describing the geometries, energetics and infrared spectra of these two isomers were investigated. Usual density functionals were tested and moreover, more recent functionals such as dispersion-corrected ones and Truhlar's M06 series were also considered. The computed infrared spectra are compared with ab initio results and InfraRed Multiple Photon Dissociation (IRMPD) experiments. Two functionals in the M06 series have been proved to be quite efficient. A large number of functionals seems to be inadequate to compute infrared spectra for peptide in the amide N–H stretching region. In addition, a detailed analysis of the sodium-peptide interaction and of the hydrogen bond between the two peptide terminations points out a distinct electronic structure for the two isomers.

1. Introduction

In the last years, InfraRed Multiple Photon Dissociation (IRMPD) spectroscopy has emerged as an efficient way to obtain infrared signatures of gaseous molecular and versatile ions. It has been applied successfully to organic, inorganic and biological species, enabling the distinction between isomers and, in some instances between conformers [1–4]. Since the species under scrutiny are gaseous ions rather than neutral molecules in the condensed phase, a given functional group bears vibrational frequencies that are different. Thus the large databases of infrared frequencies that have been gathered over the years are useful indicators but cannot be used as they stand. Assigning IRMPD bands to specific vibrational modes therefore relies heavily on molecular modeling. In most cases, this is done using quantum chemical methods based on the double harmonic approximation (mechanics and electrical). The purpose of the work described below is to provide a panorama of such methods for one specific test cases, sodiated octa-glycine G_8–Na^+, that has been characterized in detail recently [5].

Oligo-glycines G_n owe their value as model peptides to their relative simplicity due to the absence of side chains. Since Na^+ makes essentially electrostatic and polarization interactions with little or no charge transfer, small G_n–Na^+ complexes involve peptide folding around the sodium ion to maximize interactions of amide oxygens with Na^+ [6–8]. In medium-size peptides, a delicate balance occurs between metal–peptide binding and internal stabilization of the peptide through multiple hydrogen bonds. As a result, the most stable structures are much more difficult to predict when the number of residues increases. In a previous study, we investigated the structure of the sodiated octa-glycine G_8–Na^+ [5]. IRMPD spectra have been obtained at room temperature in the 1000–1900 cm^{-1} and 2900–3700 cm^{-1} regions. The coupling of the polarizable force field AMOEBA and the replica-exchange molecular dynamics (REMD) method allowed for an efficient exploration of the potential energy surface. Geometries, energetics and IR spectra were obtained at quantum chemical levels including DFT with dispersion effects and MP2. G_8–Na^+ is a medium-size peptide and has the advantage of combining various non-covalent interactions such as a very strong hydrogen bond between the N- and the C-termini, seven-membered C_7 and ten-membered C_{10} motifs, and several sodium-amide interactions. The unusually strong hydrogen bonding interaction between the N and C termini leads to a very large red-shift of the acidic O–H stretch of more than 1500 cm^{-1}. Such enormous red-shifts have been previously observed for some proton-bound dimers [9]. Based on our previous results showing that only MP2/SVP calculations can match the IRMPD spectra, the main idea behind this work is to evaluate the ability of recent DFT methods to reproduce the experimental spectra.

* Corresponding authors.
E-mail addresses: carine.clavaguera@dcmr.polytechnique.fr (C. Clavaguéra), gilles.ohanessian@polytechnique.fr (G. Ohanessian).

1387-3806/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijms.2010.08.006

The description of long-range interactions in term of geometry, energy and infrared spectrum remains a challenge of theoretical chemistry. On the one hand, the standard DFT functionals fails to describe dispersion interactions [10]. On the other hand, the MP2 theory provides an approximate treatment of dispersion, and remains tractable for systems involving up to ca. 100 atoms, on the condition that the resolution of identity (RI) approximation is used. Several studies have been carried out to evaluate the ability of MP2 to describe energetic and structure [11,12] but only few of them deal with vibrational frequency estimates [13,14]. It has been shown that in some cases, MP2 overestimates the dispersion effects with extended basis sets as indicated by overly compact structures [11].

In the last years, the DFT-D approach has been developed by adding an empirical dispersion term for which the parameters are specific to each functional. It has been proved to be an efficient way to describe dispersion effects [15]. Recently, Truhlar's group have proposed new functionals, called the M06 series, that have been parameterized to take into account various physical effects such as non-covalent interactions. They have been shown to provide a real improvement on the description of chemical properties in DFT [16]. M06-L is the meta-GGA local functional in the series, derived from the M05 one and including the spin kinetic energy densities and a self-interaction correction [17]. M06 is a meta hybrid functional with 27% of Hartree–Fock exchange for general purpose [18]. M06-2X and M06-HF are specific meta hybrid functionals with 54% and 100% of Hartree–Fock exchange, respectively. M06-2X has improved performance for thermochemistry [19] and M06-HF is recommended to provide the correct asymptotic behavior of the exchange-correlation potential for charge-transfer problems, for example [20]. The B2PLYP double-hybrid density functional has been also presented in which the semi-local GGA correlation is replaced by a standard MP2-type perturbation theory correction using Kohn–Sham orbitals and eigenvalues [21]. Moreover, long-range dispersion effects are added by a R^{-6} classical potential in the B2PLYP-D functional [22]. However, these functionals are CPU time consuming and most of time limited to single-point energy calculations. Korth and Grimme have built "artificial" and randomly generated molecules containing 8 atoms to test a variety of density functionals [23]. The CCSD(T)/complete basis set level was taken as reference for energetics. They showed that the double-hybrid functionals such as B2PLYP-D perform best, followed by the M06-2X meta-hybrid functional, to predict relative energies for thermochemical benchmarking purpose. The assignment of infrared spectra of peptides in the gas phase was investigated by MP2 theory and DFT with and without dispersion corrections by Bouteiller et al. [13]. They pointed out that MP2 (with SVP basis set) outperforms a set of the density functionals with an average error of 5 cm^{-1}. B3LYP-D and B97-D functionals were also recommended for DFT vibrational frequency computations and specific scaling factors were proposed for each method and each modes (NH_2, NH and CO stretches, NH_2 and NH in-plane bends). The theoretical treatment of dispersive forces was investigated by evaluating geometries, energies, and frequencies in the case of aromatic–aromatic interactions in isolated capped dipeptides by Gloaguen et al. [14]. The authors recommended the use of the B97-D functional to obtain reliable infrared spectra.

In the present work, we report our investigations in the assessment of a variety of density functionals for predicting the structure and the infrared spectrum of sodiated octa-glycine. For this molecular size, and with the prospect of extending the conclusions to even larger peptides, triple-zeta basis sets cannot be used to compute IR spectra, therefore double-zeta plus polarization basis sets have been used throughout. Due to their close relative energies, both charge solvation (CS) with a hydrogen-bound acidic OH and salt bridge (SB) isomers are considered for relative energetics. Understanding of their electronic structure is obtained thanks to energy decomposition analysis, Hirshfeld and Voronoi Deformation Density (VDD) charges and electron localization function (ELF) analysis to determine explicitly the nature of the hydrogen bond.

2. Theoretical calculations

RI-MP2 calculations have been previously reported on this system [5]. Density functional theory (DFT) calculations were performed using various functionals with the TURBOMOLE program package [24,25], except for the functionals of the M06 series. Geometry optimizations were performed at DFT levels using the resolution-of-the-identity [26,27] (RI) approximation with the split valence plus polarization double-ζ (SVP) basis set. One GGA functional, i.e., PBE [28], one meta-GGA functional, i.e., TPSS [29] and three hybrid functionals, i.e., TPSSH [30], PBE0 [31] and B3LYP [32] were used. Furthermore, B3LYP has been shown to give inaccurate energetics for G_8–Na^+ [33] and A_n–H^+ conformers [34]. Schwabe and Grimme have also shown the role of the lack of dispersion effects by observing an increase of the errors while the system sizes get larger [22]. Accordingly, four density functionals corrected by the empiric dispersion term proposed by Grimme [35] were used: B3LYP-D, PBE-D, TPSS-D and B97-D [36,37], see (Eqs. (1) and (2)).

$$E_{\text{DFT-D}} = E_{\text{DFT}} + E_{disp} \tag{1}$$

$$E_{disp} = -s_6 \sum_{j} \sum_{i \neq j} f_{damp}(R_{ij}) \frac{C_{6(ij)}}{R_{ij}^6} \tag{2}$$

$$\text{with} \quad f_{damp}(R_{ij}) = \frac{1}{1 + e^{-d(R_{ij}/R_r - 1)}} \quad \text{and} \quad C_{6(ij)} = \sqrt{C_{6(i)} C_{6(j)}}$$

with $C_{6(ij)}$, the dispersion coefficient for atom pair ij and the damping function depending on the damping factor d, the interatomic distance R_{ij} and R_r, the sum of these atoms van der Waals radii.

RI-DFT and RI-MP2 energy calculations on the optimized geometries, and RI-CC2 energy calculations on the RI-MP2 geometries were performed with the valence plus polarization triple-ζ (TZVPP) basis set. Vibrational frequencies were calculated at the same level as the geometry optimization without the RI approximation. The calculated band intensities were convoluted assuming a Lorentzian profile with a full width at half maximum of 15 cm^{-1}. The vibrational frequencies were scaled by a factor depending on the method. The scaling factors are provided in Table 1. For MP2 and some functionals, one scaling factor can be applied in the entire spectral range and has been determined on the adjustment of the free stretch N–H band. For the other functionals, different scaling factors have to be used below 2000 cm^{-1} (named the fingerprint region) and above 3000 cm^{-1}. In these cases, the scaling factor used in the fingerprint region was determined such as to fit the maximum of the amide I band position. These scaling

Table 1
Scaling factors used for the two spectral ranges (i.e., 1100–1800 cm^{-1} and 2900–3600 cm^{-1}) at MP2 and various DFT levels using SVP basis set.

Method	1100–1800 cm^{-1}	2900–3600 cm^{-1}
MP2	0.943	
B3LYP/B3LYP-D	0.960	
B97-D	0.985	0.975
PBE/PBE-D	0.985	0.980
PBE0	0.940	0.945
TPSS/TPSS-D	0.985	0.980
TPSSH	0.960	0.965
SSB-D	0.975	0.968
M06-L	0.958	
M06	0.955	0.960
M06-2X	0.950	
M06-HF	0.945	0.950

factors are similar to the ones available in previous studies at MP2/SVP, B3LYP/SVP and B3LYP-D/SVP levels [13,14]. Finally, the relative enthalpies and Gibbs free energies at 298 K were determined at the DFT/TZVPP//DFT/SVP, MP2/TZVPP//MP2/SVP and CC2/TZVPP//MP2/SVP levels of theory.

Geometry optimizations were also performed on both isomers using the M06-series functionals [17–20], i.e., M06, M06-L, M06-2X and M06-HF, associated to the cc-pVDZ basis set, with the GAMESS program package [38,39]. Single-point energy calculations were performed at the optimized geometries with the cc-pVTZ basis set.

Finally, geometries and vibrational frequencies were computed using the SSB-D functional with the DZP basis set in the ADF 2009.01 program package [40,41]. The SSB-D functional is based on a small correction of the exchange part of the PBE functional combined with the gradient expansion exchange term from Keal and Tozer and Grimme's dispersion correction. The SSB-D functional was designed for biological applications in order to improve the representation of hydrogen bonding, π–π stacking, spin-state splitting, accuracy of geometries, reaction barriers [42].

Energy decomposition analysis (EDA) based on the works by Morokuma [43,44], and Ziegler and Rauk [45] was carried out using the ADF 2009.01 program package [40,41] at the TURBOMOLE geometries with the M06-L functional. The interaction energy ΔE_{int} between the peptide and the sodium ion is decomposed into electrostatic, Pauli repulsion and orbital mixing components as in Eq. (3).

$$\Delta E_{int} = \Delta E_{elstat} + \Delta E_{\mathrm{Pauli}} + \Delta E_{orb} \quad (3)$$

Slater-type orbitals (STOs) were employed as basis functions in SCF calculations. The TZ2P basis sets used have triple-ζ quality augmented by two sets of polarization functions. Detailed information on chemical bonding was provided with an electron localization function (ELF) analysis of the ADF results using the DGrid 4.5 program package [46].

3. Results and discussion

3.1. Structures and energetics of G_8–Na^+

The lowest energy CS and SB isomers were found to be the two most stable structures for the G_8–Na^+ complex. They are stabilized by a hydrogen bond between the N-terminus and the C-terminus [5]. The two structures obtained at the MP2/SVP level are shown in Fig. 1. They are very similar to each other and the major difference is located on the nature of the terminations, neutral or charged. At the MP2 level of theory, we were able to locate the two isomers as minima on the potential energy surface. However, in most of the DFT calculations, it was impossible to locate one of the isomers. Indeed, only B3LYP-D, M06, M06-L and M06-HF functionals supplied both isomers. All the other functionals (i.e., PBE, TPSS, TPSSH, PBE0, B3LYP, PBE-D, TPSS-D, SSB-D and B97-D) provided only the SB structure with a systematic proton transfer occurring during the geometry optimization of the CS one. The M06-2X results are noteworthy with a proton transfer from SB to CS, locating only the CS isomer. To exclude a possible basis set effect due to the lack of diffuse functions, the aug-cc-pVDZ basis set was used in cases of B97-D and PBE-D functionals, however this did not prevent the proton transfer from CS to SB. It is not possible to exclude strictly the existence of a second isomer in any of these cases, however our results indicate that it is at best a very shallow minimum. The fact that the most reliable density functionals agree with MP2 in predicting the existence of two minima lends confidence into this results, and points to deficiencies in other density functionals.

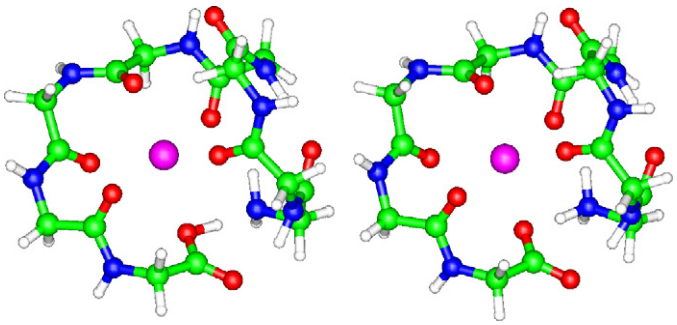

Fig. 1. CS (left side) and SB (right side) MP2 geometries.

Table 2
Energies, enthalpies and free energies at 298 K (in kcal/mol) of the SB relative to the CS isomer at various quantum chemistry levels.

Method	ΔE	ΔH	ΔG
CC2 [a]	−2.2	−1.9	−1.3
MP2	−1.4	−1.1	−0.5
B3LYP-D	−0.8	−0.5	−1.5
M06	+0.6	+1.6	+0.5
M06-L	−0.8	−0.5	−0.6
M06-HF	−2.1	−2.1	−2.7

Δ G calculated using MP2 frequencies.
[a] At MP2 geometries.

The description of the structures is given using the sequence numbering from the N- to the C-terminus with superscripts, i.e., NH^2 and CO^2 stand for the peptidic N–H and C=O bonds of the second residue. For residue 1, the N–H bonds are in the NH_2 group rather than an amide group and for residue 8, CO^8 is an acidic rather than an amide C= O. The structure of G_8–Na^+ is characterized by a high coordination for the sodium ion by five peptidic carbonyl groups and the C–OH or C–O^- at the C-terminus. The acidic C= O is hydrogen-bound to the peptidic NH^2. A strong direct interaction exists between the N- and C-termini, either $^+$N–H···O^- (SB isomer) or N···H–O (CS isomer). In both isomers, the second N–H at the N-terminus is bound to CO^4 and in the SB case, the third one is free. NH^4 interacts through a hydrogen bond in a ten-membered C_{10} motif with CO^1 and NH^5 is involved in a hydrogen bond in a seven-membered C_7 motif with CO^3. The other peptidic N–H (i.e., NH^3, NH^6, NH^7 and NH^8) are free. Relative energies, enthalpies and free energies at 298 K of the CS and SB isomers are given in Table 2 taking the CS isomer as the reference. In most cases, the SB isomer has the lowest energy. It is the case at the CC2/TZVPP level with a ΔG of −1.3 kcal/mol which we consider as the reference. MP2 values differ from the CC2 ones by less that 1 kcal/mol. The B3LYP-D and M06 series energies are consistent with the post-Hartree–Fock ones, giving confidence in these DFT calculations. B3LYP-D and M06-L energies are very close to the CC2 ones, while M06 gives CS as more stable than SB. All free energy differences are in a ±2 kcal/mol window with respect to the CC2 values. We consider that both systems are about equally stable in the limit of the calculation accuracy. In addition, the transition state for interconversion of the CS and SB isomers was located at the B3LYP-D/SVP level [5]. B3LYP-D/TZVPP and MP2/TZVPP calculations performed at the B3LYP-D/SVP geometries yield a very shallow barrier of 0.3 and -0.06 kcal/mol respectively, relative to the CS minimum. Despite this small or negligible barrier, experiments provide evidence of the CS isomer only.

3.2. Analysis of the electronic structure of the CS and SB isomers

Geometric, energetic and population analysis parameters are provided in Table 3 at the M06-L level. The total bonding energy is given with respect to the (G_8 + Na^+) ionic dissociation limit (the peptide structure being the same as that in the sodium complex

Table 3
Characteristic distances for the N-terminus/C-terminus interaction in Å, bonding Energy (BE) analysis without relaxation of the G_8 structure at dissociation (all values in kcal/mol). Pauli + Electrostatic + Orbital = BE, and population analysis.

	CS M06-L	SB M06-L
d(O–H)	1.056	1.551
d(N–H)	1.584	1.094
BE	−106.5	−110.1
Pauli repulsion	11.0	14.4
Electrostatic	−78.6	−85.1
Orbital	−38.9	−39.4
VDD Nter	0.13	0.32
VDD Cter	−0.06	−0.47
Hirshfeld Nter	0.12	0.38
Hirshfeld Cter	−0.07	−0.37
ELF attractor	0.95	0.90

at equilibrium) with an absolute value larger for SB than for CS. Electrostatic and orbital interactions are of the same magnitude in both systems corresponding to about two thirds and one third of the sum of these attractive contributions for electrostatic and orbital terms, respectively. No strong difference can be pointed out to distinguish the isomers because the interaction energy is a global property while their differences are highly localized: the chemical nature of the terminations does not have a strong influence on the total sodium–peptide binding energy. Moreover, the distance between the N-terminus nitrogen and the sodium ion is very similar in the two isomers, 3.78 Å for CS and 3.74 Å for SB. Other theoretical frameworks must used to focus on the terminations such as population analyses. Hirshfeld and Voronoi Deformation Density (VDD) atomic charges are both based on the molecular electron density. They were proved to be less dependent upon basis sets than Mulliken charges and often give similar results [47]. The VDD calculates the flow of electron density between fragments due to bond formation. The population analyses on the N-terminus and C-terminus groups reflect their chemical differences. The global charges correspond to the sum of the atomic charges at the termination groups defined as CO_2H/NH_2 in the CS isomer and CO_2^-/NH_3^+ in the SB one. For SB, a clear charge separation is found with a positive charge for the N-terminus (+0.32 or +0.38 electron) and a negative charge for C-terminus (−0.47 or −0.37 electron). The terminations are close to neutral in CS with charges around +0.1 and −0.1. These DFT population frameworks are able to discriminate between the two isomers in spite of the very similar geometries and interactions with the sodium ion.

The electron localization function (ELF) provides a "picture" of the electronic structure showing the regions of the molecular space where the electron pairs localize. It is possible to split the molecular space into basins corresponding to the core electrons and to the shared and unshared (lone pair) electron density of the valence shell. The core basins are centered on the nuclei and the valence basins are in the remaining space. The notation typically used is C(A) for the core basin of atom A, V(A) for a lone pair of atom A, V(A,B) for a two center bond between A and B [48]. The ELF function ranges from 0 (same spin pairs) to 1 (opposite spin pairs) and the local maxima of ELF define localization domains. In previous studies, a topological description of the hydrogen bond was proposed confirming that it is possible to evaluate the strength of the hydrogen bonds on the basis of the ELF criteria [49,50]. The formation of the hydrogen bond changes the properties of the ELF basins. It enables to discuss the bonding in both the CS and SB isomers. The CS isomer is characterized by the formation of an OH···N bond between the hydrogen atom and the nitrogen lone pair. The hydro-

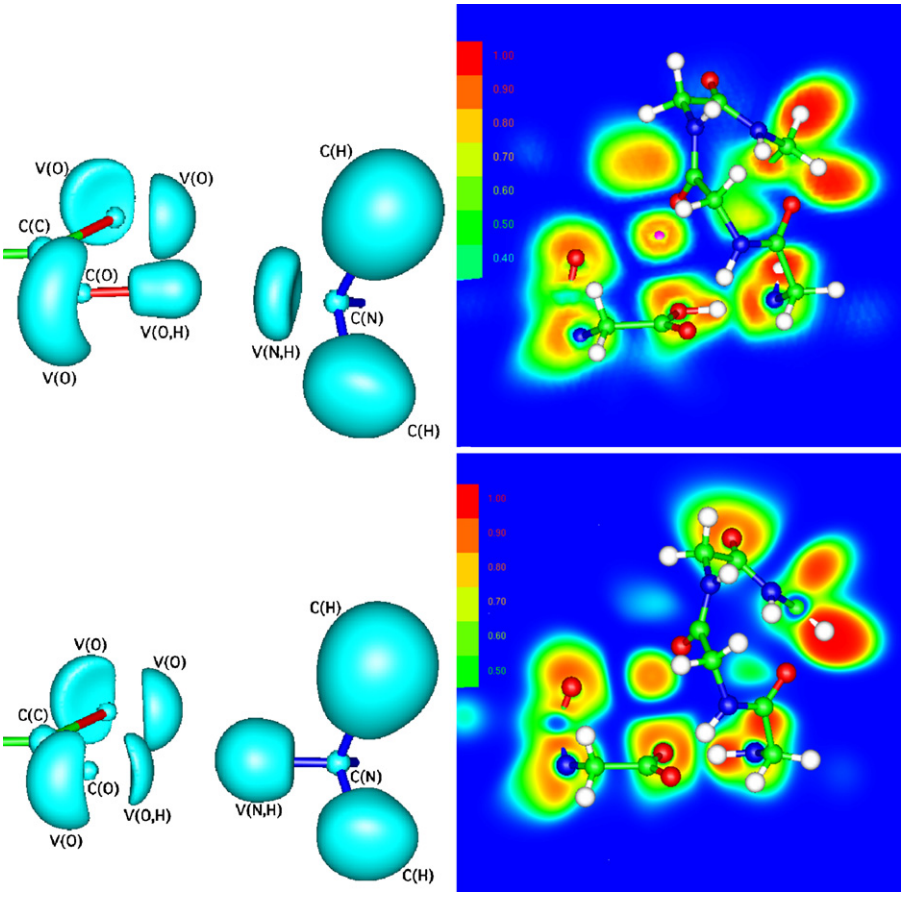

Fig. 2. ELF localization domains and cut-plane ELF representations for CS (top) and SB (bottom) isomers.

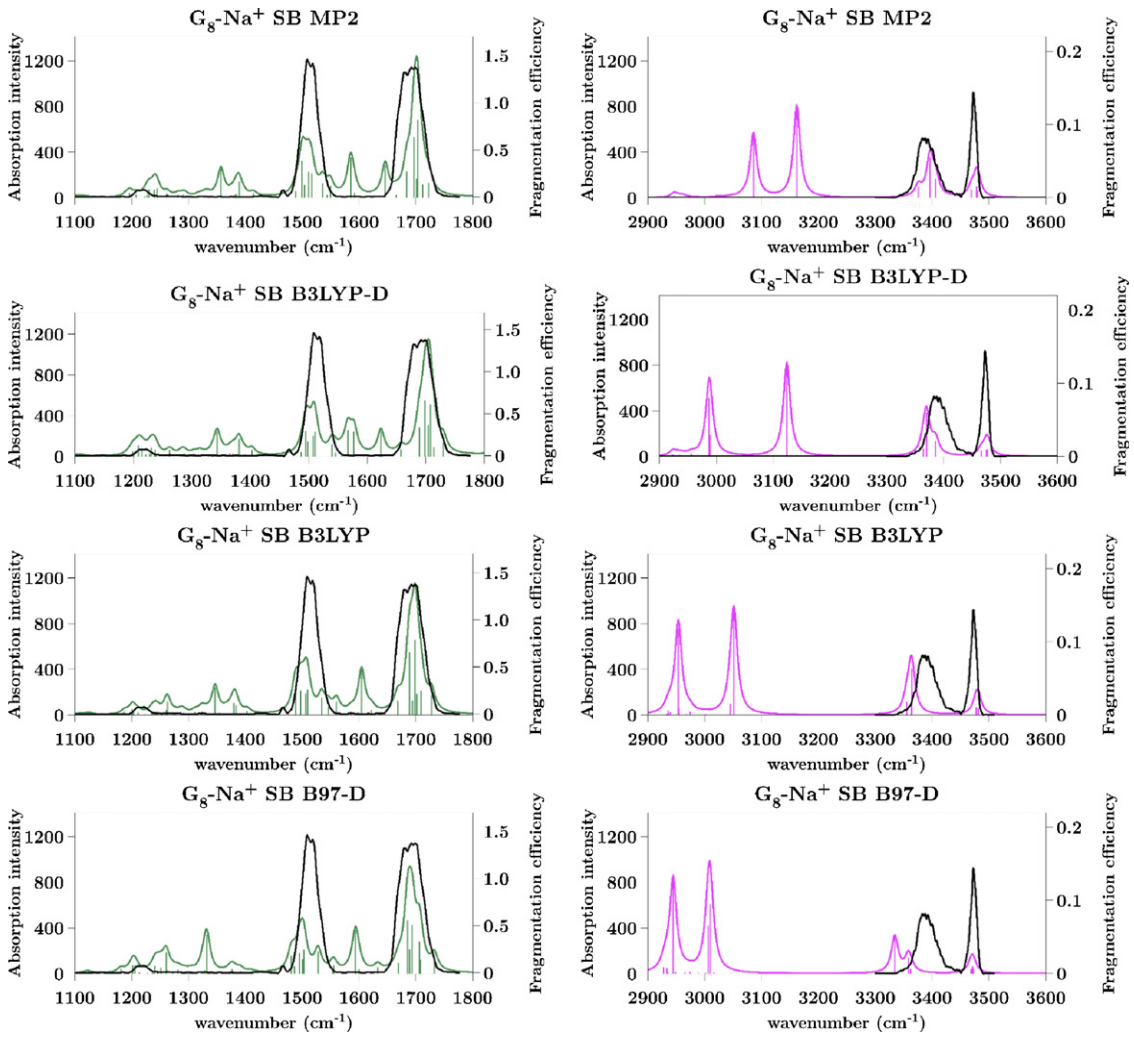

Fig. 3. G_8–Na^+ SB spectra at MP2, B3LYP-D, B3LYP, and B97-D levels. The IRMPD spectrum is given in black.

gen atom is bound to O and located in the V(O,H) basin which has a boundary with the nitrogen lone pair basin V(N). As a consequence, the formation of a V(N,H) basin is observed (top of Fig. 2). For the SB isomer where a proton transfer occurs, the V(O,H) basin turns into a lone pair basin V(O) and the hydrogen atom is relocated in a V(N,H) basin with a boundary with V(O). A V(O,H) basin is created as an indication of the presence of a hydrogen bond with the oxygen atom (bottom of Fig. 2).

The smallest value of the ELF maximum in SB (0.90) with respect to the CS one (0.95) corresponds to the delocalization of the spin-paired electron density. The ELF distribution along the O···H···N bond path shows lower values for the SB isomer (minima: 0.32) than for the CS isomer (minima: 0.46). This can be observed in the 2D picture of the cut plane through the O···H···N atoms of interest for both isomers in Fig. 2. Although relatively small, these differences remain significant as already discussed in the literature (see Ref. [50] for example). So, we can expect a weaker hydrogen bond in the SB isomer. These results indicate that the two isomers have clearly distinct electronic properties and are not just two numerical minima on a very flat portion of the potential energy surface.

3.3. Spectrum analysis

3.3.1. Comparison of computed spectra for the SB isomer

The SB isomer is available for almost all the functionals and the resulting spectra for this structure are provided in Figs. 3–5. It was already demonstrated that the structure of G_8–Na^+ cannot be a salt bridge on the basis of the IRMPD spectrum [5] and MP2 results. Several significant bands were used to discriminate between the two isomers: the presence of intense bands due to HNH bending modes of the ammonium in between the experimental amide I and II bands and two intense bands for N–H stretches of NH_3^+ that are hydrogen-bonded to carbonyl oxygens in the 3000–3200 cm^{-1} area. Although this structure is not consistent with IRMPD results, it is the one which provides the largest data for the assessment of density functionals to reproduce infrared spectra. The MP2 spectrum has been taken as reference to establish the performance of a large diversity of functionals; the IRMPD bands are also shown in order to make visual comparison easier but this comparison must be limited to modes not strongly affected by the position of the proton at the C- and N-termini. Because of the breadth of the amide I and II bands for an octa-peptide, most functionals yield spectra that are in relatively good agreement with the MP2 reference in the fingerprint region. The higher sensitivity of frequencies to hydrogen bonding in the N–H stretch region makes this region more revealing.

The B3LYP-D and MP2 spectra are very similar (Fig. 3), the major difference exists in the N–H stretch range with a red-shift of the intense NH_3^+ stretch bands due to stronger hydrogen bonds with carbonyl groups. This shift involves an artificial coupling of one band with C–H stretch modes, resulting in an intense doublet just below 3000 cm^{-1}. The B3LYP spectrum is comparable to the B3LYP-D one with a larger red-shift of the NH_3^+ bands giving a coupling

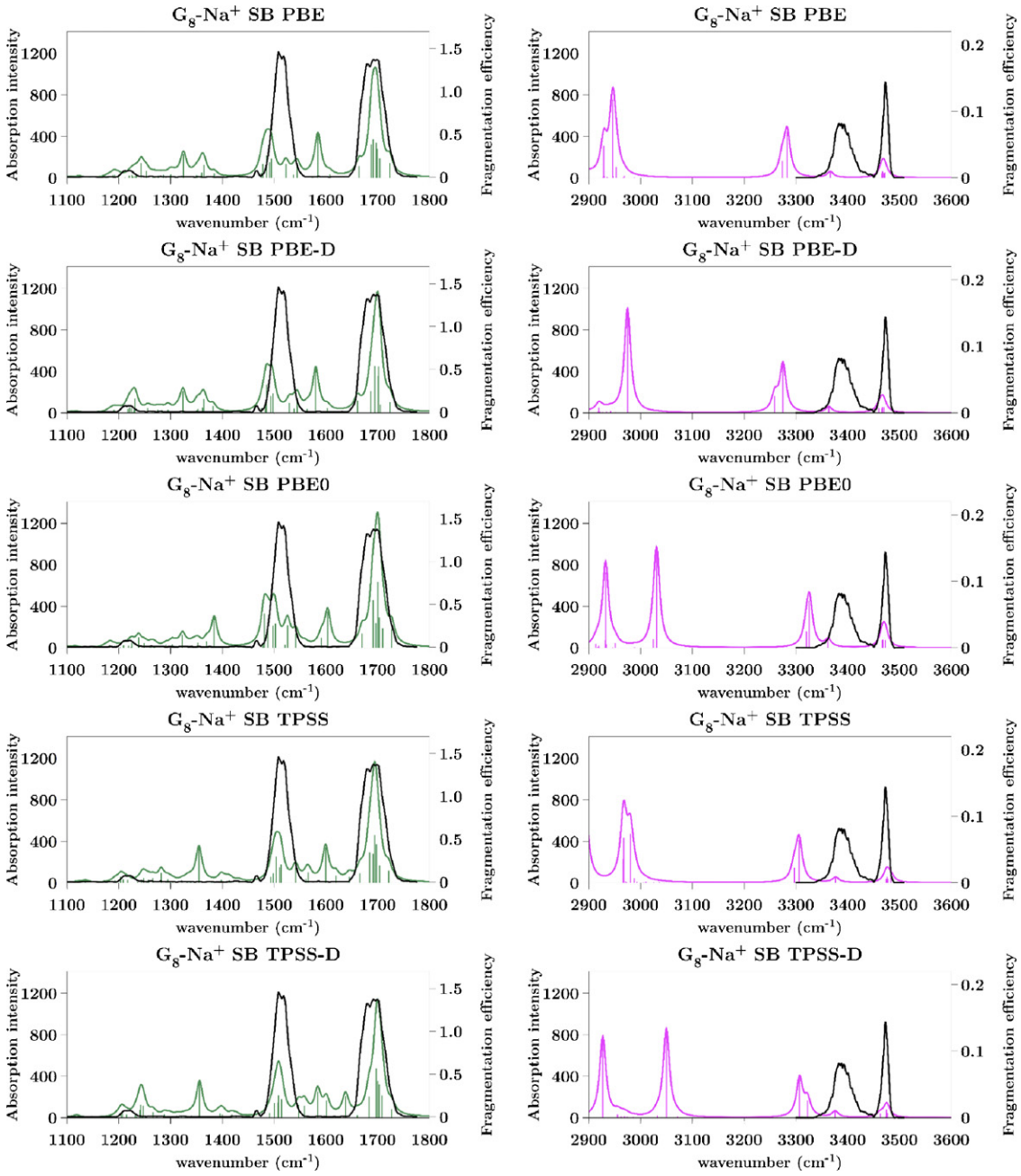

Fig. 4. G_8–Na^+ SB spectra at PBE, PBE-D, PBE0, TPSS and TPSS-D levels. The IRMPD spectrum is given in black.

of both bands with C–H stretch modes (Fig. 4) and a small red-shift of the H-bound peptidic N–H bands. These results point out the role of dispersion effects, treated here by including an empirical correction in the popular B3LYP functional. The B97-D spectrum (Fig. 4) matches the B3LYP one with a red-shift of the H-bound N–H stretch modes, particularly for one band situated outside of the IRMPD band. This result emphasizes some limitations of this dispersion-corrected GGA functional possibly due to the absence of Hartree–Fock exchange.

PBE, PBE-D and PBE0 spectra (Fig. 4) are very similar to each other, especially for the amide I and amide II range with a small red-shift of the amide II band. For the three functionals, the N–H stretch modes of the ammonium are red-shifted between 2900 and 3050 cm^{-1} with a minor improvement in PBE-D. Moreover, the H-bound N–H band is not in agreement with the MP2 one, even if the position of the band is slightly better in PBE0 thanks to the presence of Hartree–Fock exchange. We can conclude that this series of functionals is ineffective to reproduce the vibrational frequencies in the N–H stretch region with an unique scaling factor.

TPSS, TPSS-D, TPSSH and SSB-D spectra (Fig. 5) are very similar to each other in the fingerprint region and in agreement with the MP2 spectrum. In the N–H/O–H stretch range, the TPSS spectrum presents the same features than the PBE one. The TPSS-D and TPSSH spectra give a slight better agreement thanks to the inclusion of dispersion effects or Hartree–Fock exchange. As the SSB-D functional was adjusted on hydrogen-bound systems, the spectrum presents an improvement in the reproduction of the ammonium N–H stretch modes in comparison with MP2 results. For all these functionals, as for the PBE series, the position of the free N–H stretch band is well reproduced but not for the H-bound N–H modes. They cannot be recommended for frequency calculations in this range for charged peptides.

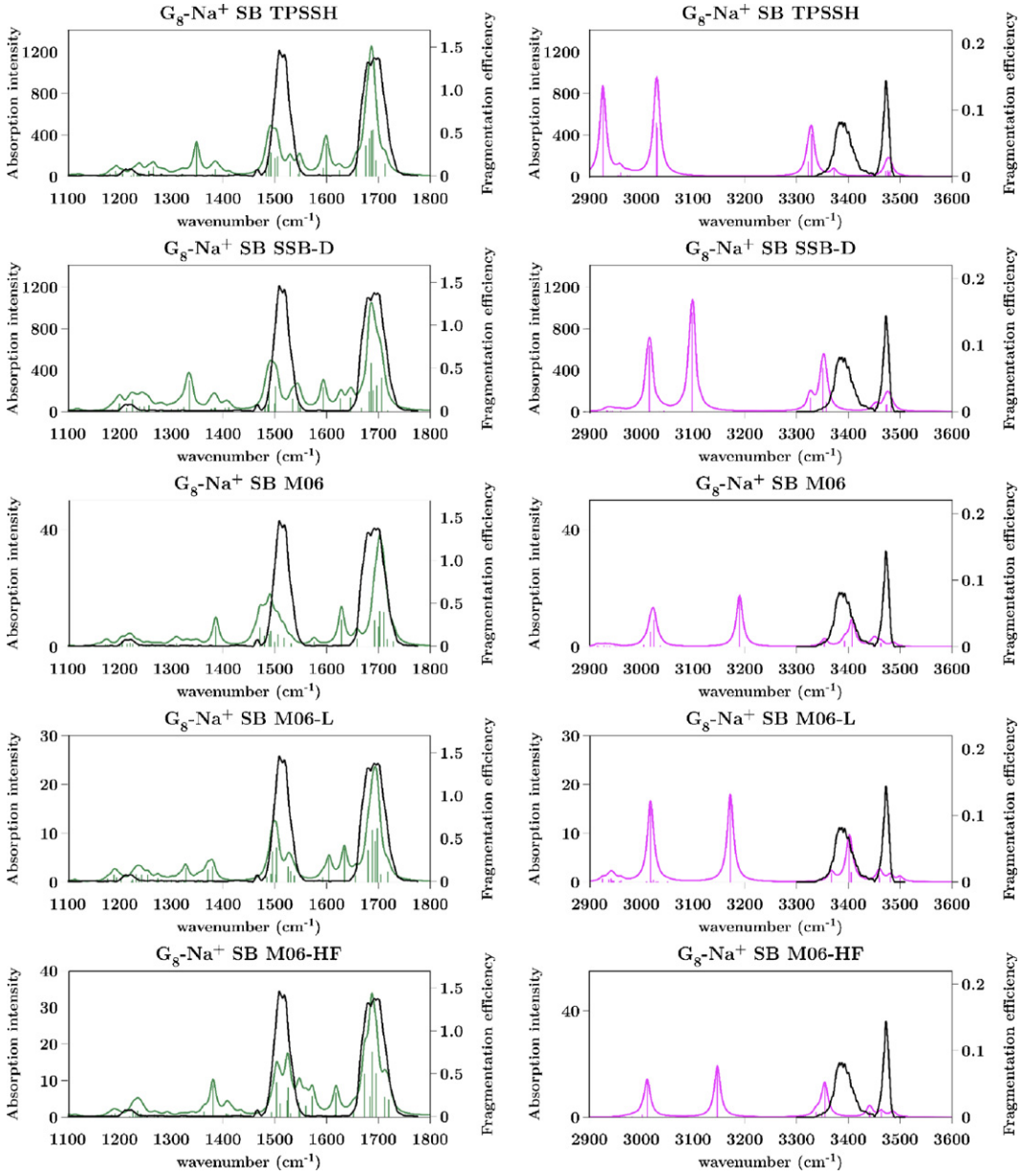

Fig. 5. G_8–Na^+ SB spectra at TPSSH, SSB-D, M06, M06-L and M06-HF levels. The IRMPD spectrum is given in black.

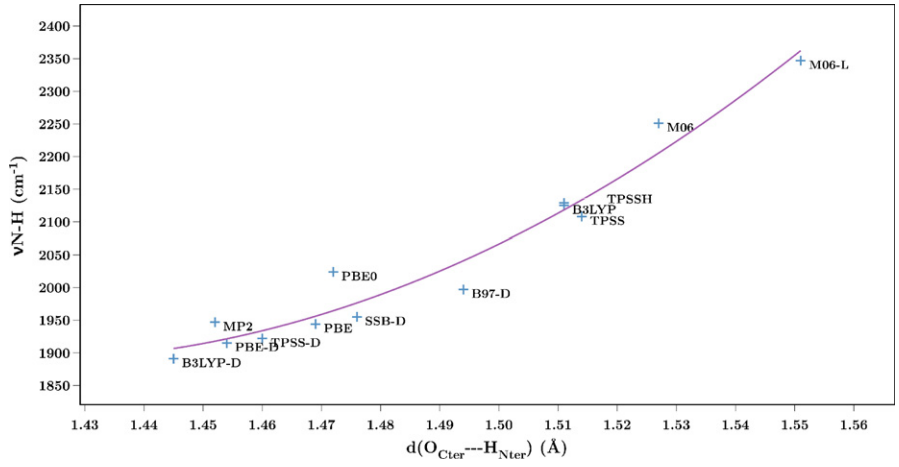

Fig. 6. Ammonium N–H stretch frequency involved in the strong hydrogen bond as a function of the O···H–N distance for the SB isomer.

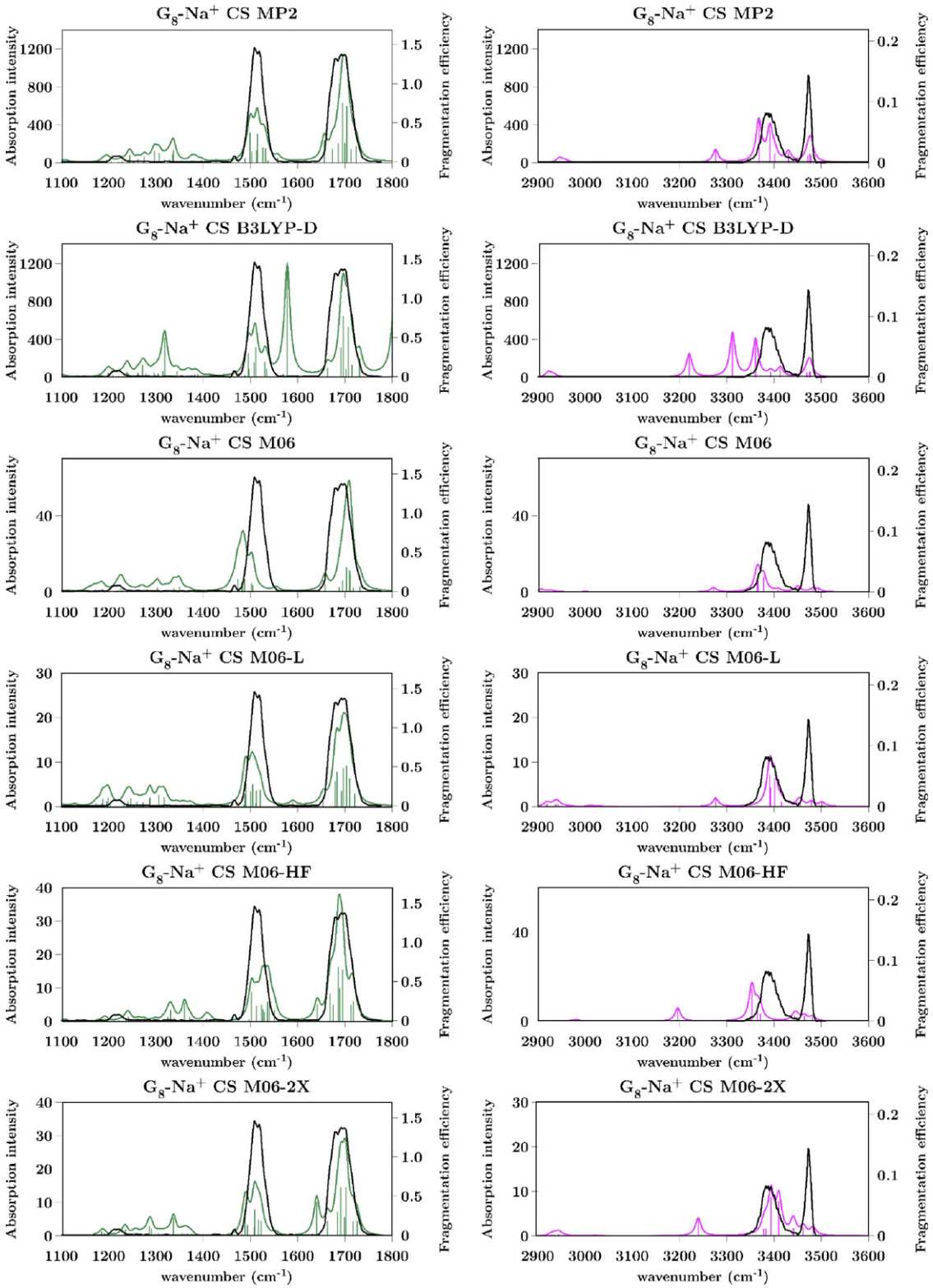

Fig. 7. G_8–Na^+ CS spectra at MP2, B3LYP-D, M06, M06-L, M06-HF and M06-2X levels. The IRMPD spectrum is given in black.

In the M06 series, the M06 spectrum shows a red-shift of the amide II band and different shifts of the N–H stretch bands of NH_3^+. The M06-L spectrum is very similar to the MP2 one in all the spectral range. The M06-HF functional provides a good agreement in the fingerprint region but a red-shift of the H-bound N–H stretch modes. In this series, the M06-L gives the best agreement moderated by the spread of the free N–H band with a red-shift about 60 cm^{-1} for the weakly-bound band.

In conclusion, these results highlight the sensitivity of DFT vibrational frequency calculations to the inclusion of Hartree–Fock exchange or dispersion effects. Some functionals built to study hydrogen-bond complexes are not able to reproduce the MP2 data

even for some of the weakly bound N–H stretches for such an uncommon system. The recent meta-GGA M06-L functional provides encouraging results in comparison with the MP2 ones.

In addition to the spectra shown in Figs. 3–5, the very strong hydrogen bond between one of the N–H bonds at the ammonium terminus and one of the oxygens at the carboxylate terminus leads to a spectacular red-shift of more than 1000 cm^{-1} at some levels. Fig. 6 presents the computed stretch frequency for this ammonium N–H as a function of the NH$\cdots$O hydrogen bond distance. It is clear that there is a strong correlation between these two quantities. DFT-D and MP2 results provide short distances with a large red-shift for the N–H stretch band. B3LYP and other standard functionals are intermediate. On the contrary, M06 and M06-L give a longer hydrogen bond, and as a consequence, a weaker interaction with a N–H mode between 2250 and 2350 cm^{-1}. This behavior points out the difficulty to evaluate the position of this red-shifted band by standard harmonic calculations. Furthermore, adding the dispersion term involves a smaller blue-shift of the other two $^{+}$N–H bands, due to a weaker interaction of the N-terminus with CO[4]. This can be seen by comparing the spectra of TPSS and TPSS-D, B3LYP and B3LYP-D, and with a lesser effect, PBE and PBE-D. Other correlations were attempted between the N–H bond distance and the respective stretch mode frequency for free peptidic NH groups, and also between the Na$\cdots$O distance and the concerned C= O or N–H stretch frequencies. However, the variations in these distances are too small to be significant and no clear new correlation was identified.

3.3.2. Comparison of computed spectra for the CS isomer

As described in Section 3.1, the CS isomer could be found with only five functionals. The resulting IR spectra are given in Fig. 7. The MP2 spectrum of CS was proved to be in very good agreement with the IRMPD spectrum [5]. The amide I and amide II massifs are clearly identifiable in the 1650–1750 and 1480–1560 cm^{-1} ranges, respectively. No experimental band is detected in the 1250–1400 cm^{-1} range in spite of the fact that the MP2 calculations predict a large number of low-intensity frequencies. These bands correspond essentially to C–H bending modes and are not observed, probably due to small absorption cross-sections. The IRMPD band in the 3460–3490 cm^{-1} range is characteristic of several free N–H modes. The experimental broad band at 3340–3440 cm^{-1} corresponds to weakly H-bound N–H or NH_2 bands. The B3LYP-D spectrum presents an intense band in between the amide I and II bands corresponding to a strong mixing of the O–H stretch and the carboxylic C=O stretch modes. Moreover, a red-shift of the H-bound N–H stretch modes is observed in worse agreement with the IRMPD spectrum than the MP2 one. In the M06 series, the free N–H bands are spread out but mainly around the experimental position. The M06 spectrum displays the same problem as for SB with a red-shift of the amide II band. The rest of the spectrum is well reproduced. The M06-2X functional, which was unable to locate the SB isomer, is in good agreement with the IRMPD spectrum except for one C=O stretch band outside the amide I band. The M06-HF spectrum is also in rather good agreement and the unique failure is the red-shift of the H-bound N–H stretch band by 50 cm^{-1} as already seen for the SB spectrum. As for the SB case, the M06-L spectrum is in good agreement with the IRMPD spectrum, nearly as good as the MP2 one except for a spread of the free N–H bands. This functional gives also the best agreement with the MP2 data in the SB case. At both MP2 and DFT levels, the higher-frequency amide I band corresponds to a delocalized in-phase stretching mode for the amide carbonyl groups bound to Na^{+}. However, this band is not separated from the other amide I bands as it was observed for peptides coordinated to a trivalent cation [51] because of a weaker interaction of C=Os with the single-charged sodium ion than with lanthanide 3+ ions.

4. Conclusions

Some successes and failures of DFT methods to study the structure, energetics and spectroscopy of peptides were presented in the case of the sodiated octa-glycine peptide. A large number of functionals are not able to locate both the charge solvated and the salt bridge isomers of lowest energy. Only B3LYP-D and the M06 series (except for M06-2X) of functionals succeed due to a combination of Hartree–Fock exchange and dispersion effects in the first case, and a diverse set of data in the parameterization for the new meta-GGA derived functionals. DFT energy decomposition analysis at the M06-L level attests that the major part of the bonding energy between the peptide and the sodium ion comes from electrostatic interaction. Population analyses based on the electron density allow to differentiate SB from CS based on realistic values of the termination charges. The ELF analysis of M06-L densities results shows a large stability due to the strong hydrogen bond in the CS structure. The M06 series performs best among density functionals to reproduce the spectra in comparison with MP2 and IRMPD spectra. The best results are obtained with the M06-L GGA functional giving the CS and SB structures close in energy and a good agreement with the experimental spectrum. The PBE and TPSS series are not able to give the good positions of the weakly hydrogen-bound N–H stretch bands and cannot be recommended to compute IR spectra for peptides in this region.

Acknowledgments

This work was granted access to the HPC resources of [CCRT/CINES/IDRIS] under the allocation x2009085107 made by GENCI (Grand Equipement National de Calcul Intensif).

References

[1] L. MacAleese, P. Maître, Infrared spectroscopy of organometallic ions in the gas phase: from model to real world complexes, Mass Spectrom. Rev. 26 (4) (2007) 583–605.
[2] K.R. Asmis, J. Sauer, Mass-selective vibrational spectroscopy of vanadium oxide cluster ions, Mass Spectrom. Rev. 26 (4) (2007) 542–562.
[3] N.C. Polfer, J. Oomens, Vibrational spectroscopy of bare and solvated ionic complexes of biological relevance, Mass Spectrom. Rev. 28 (3) (2009) 468–494.
[4] T.D. Fridgen, Infrared consequence spectroscopy of gaseous protonated and metal ion cationized complexes, Mass Spectrom. Rev. 28 (4) (2009) 586–607.
[5] D. Semrouni, O.P. Balaj, F. Calvo, C.F. Correia, C. Clavaguéra, G. Ohanessian, Structure of sodiated octa-glycine: IRMPD spectroscopy and molecular modeling, J. Am. Soc. Mass Spectrom. 21 (2010) 728–738.
[6] M.M. Kish, C. Wesdemiotis, G. Ohanessian, The sodium ion affinity of glycylglycine, J. Phys. Chem. B 108 (9) (2004) 3086–3091.
[7] P. Wang, C. Wesdemiotis, C. Kapota, G. Ohanessian, The sodium ion affinities of simple di-, tri-, and tetrapeptides, J. Am. Soc. Mass Spectrom. 18 (3) (2007) 541–552.
[8] O.P. Balaj, C. Kapota, J. Lemaire, G. Ohanessian, Vibrational signatures of sodiated oligopeptides (GG–Na^{+}, GGG–Na^{+}, AA–Na^{+} and AAA–Na^{+}) in the gas phase, Int. J. Mass Spectrom. 269 (3) (2008) 196–209.
[9] J.R. Roscioli, L.R. McCunn, M.A. Johnson, Quantum structure of the intermolecular proton bond, Science 316 (5822) (2007) 249–254.
[10] K.E. Riley, M. Pitoňák, P. Jurečka, P. Hobza, Stabilization and structure calculations for noncovalent interactions in extended molecular systems based on wave function and density functional theories, Chem. Rev. ASAP (2010).
[11] J. Cerny, P. Hobza, Non-covalent interactions in biomacromolecules, Phys. Chem. Chem. Phys. 9 (2007) 5291–5303.
[12] K.E. Riley, P. Hobza, Assessment of the MP2 method, along with several basis sets, for the computation of interaction energies of biologically relevant hydrogen bonded and dispersion bound complexes, J. Phys. Chem. A 111 (33) (2007) 8257–8263.
[13] Y. Bouteiller, J.C. Poully, C. Desfrançois, G. Grégoire, Evaluation of MP2, DFT, and DFT-D methods for the prediction of infrared spectra of peptides, J. Phys. Chem. A 113 (22) (2009) 6301–6307.
[14] E. Gloaguen, H. Valdes, F. Pagliarulo, R. Pollet, B. Tardivel, P. Hobza, F. Piuzzi, M. Mons, Experimental and theoretical investigation of the aromatic–aromatic interaction in isolated capped dipeptides, J. Phys. Chem. A 114 (9) (2000) 2973–2982.
[15] T. Schwabe, S. Grimme, Theoretical thermodynamics for large molecules: walking the thin line between accuracy and computational cost, Acc. Chem. Res. 41 (4) (2008) 569–579.

[16] Y. Zhao, D.G. Truhlar, Density functionals with broad applicability in chemistry, Acc. Chem. Res. 41 (2) (2008) 157–167.
[17] Y. Zhao, D.G. Truhlar, A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions, J. Chem. Phys. 125 (19) (2006) 194101–194118.
[18] Y. Zhao, D. Truhlar, The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four m06-class functionals and 12 other functionals, Theor. Chem. Acc. 120 (1) (2008) 215–241.
[19] Y. Zhao, D.G. Truhlar, Exploring the limit of accuracy of the global hybrid meta density functional for main-group thermochemistry, kinetics, and noncovalent interactions, J. Chem. Theory Comput. 4 (11) (2008) 1849–1868.
[20] Y. Zhao, D.G. Truhlar, Density functional for spectroscopy: no long-range self-interaction error, good performance for Rydberg and charge-transfer states, and better performance on average than B3LYP for ground states, J. Phys. Chem. A 110 (49) (2006) 13126–13130.
[21] S. Grimme, Semiempirical hybrid density functional with perturbative second-order correlation, J. Chem. Phys. 124 (3) (2006) 034108.
[22] T. Schwabe, S. Grimme, Double-hybrid density functionals with long-range dispersion corrections: higher accuracy and extended applicability, Phys. Chem. Chem. Phys. 9 (2007) 3397–3406.
[23] M. Korth, S. Grimme, "Mindless" DFT Benchmarking, J. Chem. Theory Comput. 5 (4) (2009) 993–1003.
[24] R. Ahlrichs, M. Bar, M. Haser, H. Horn, C. Kolmel, Electronic-structure calculations on workstation computers—the program system TURBOMOLE, Chem. Phys. Lett. 162 (3) (1989) 165–169.
[25] C. Hattig, F. Weigend, CC2 excitation energy calculations on large molecules using the resolution of the identity approximation, J. Chem. Phys. 113 (13) (2000) 5154–5161.
[26] K. Eichkorn, O. Treutler, H. Öhm, M. Häser, R. Ahlrichs, Auxiliary basis sets to approximate Coulomb potentials, Chem. Phys. Lett. 242 (6) (1995) 652–660.
[27] K. Eichkorn, F. Weigend, O. Treutler, R. Ahlrichs, Auxiliary basis sets for main row atoms and transition metals and their use to approximate Coulomb potentials, Theor. Chem. Acc. 97 (1997) 119–124.
[28] J.P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple, Phys. Rev. Lett. 77 (1996) 3865–3868.
[29] J. Tao, J.P. Perdew, V.N. Staroverov, G.E. Scuseria, Climbing the density functional ladder: nonempirical meta-generalized gradient approximation designed for molecules and solids, Phys. Rev. Lett. 91 (14) (2003) 146401.
[30] V.N. Staroverov, G.E. Scuseria, J. Tao, J.P. Perdew, Comparative assessment of a new nonempirical density functional: molecules and hydrogen-bonded complexes, J. Chem. Phys. 119 (23) (2003) 12129–12137.
[31] J.P. Perdew, M. Ernzerhof, K. Burke, Rationale for mixing exact exchange with density functional approximations, J. Chem. Phys. 105 (1996) 9982–9985.
[32] A.D. Becke, Density-functional thermochemistry. III. The role of exact exchange, J. Chem. Phys. 98 (7) (1993) 5648–5652.
[33] D. Semrouni, G. Ohanessian, C. Clavaguéra, Structural, energetic and dynamical properties of sodiated oligoglycines: relevance of a polarizable force field, Phys. Chem. Chem. Phys. 12 (2010) 3450–3462.
[34] T.D. Vaden, T.S.J.A. de Boer, J.P. Simons, L.C. Snoek, S. Suhai, B. Paizs, Vibrational spectroscopy and conformational structure of protonated polyalanine peptides isolated in the gas phase, J. Phys. Chem. A 112 (20) (2008) 4608–4616.
[35] S. Grimme, Semiempirical GGA-type density functional constructed with a long-range dispersion contribution, J. Comput. Chem. 27 (15) (2006) 1787–1799.
[36] A.D. Becke, Density-functional thermochemistry. V. Systematic optimization of exchange-correlation functionals, J. Chem. Phys. 107 (20) (1997) 8554–8560.
[37] S. Grimme, Accurate description of van der waals complexes by density functional theory including empirical corrections, J. Comput. Chem. 25 (12) (2004) 1463–1473.
[38] M.W. Schmidt, K.K. Baldridge, J.A. Boatz, S.T. Elbert, M.S. Gordon, J.H. Jensen, S. Koseki, N. Matsunaga, K.A. Nguyen, S. Su, T.L. Windus, M. Dupuis, J.A. Montgomery, General atomic and molecular electronic structure system, J. Comput. Chem. 14 (1993) 1347–1363.
[39] M.S. Gordon, M.W. Schmidt, Advances in electronic structure theory: GAMESS a decade later, Elsevier, Amsterdam, 2005, pp. 1167–1189.
[40] ADF2009.01, SCM, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands, 2009, http://www.scm.com.
[41] G. te Velde, F.M. Bickelhaupt, E.J. Baerends, C.F. Guerra, S.J.A. van Gisbergen, J.G. Snijders, T. Ziegler, Chemistry with ADF, J. Comput. Chem. (2001) 931–967.
[42] M. Swart, M. Sola, F.M. Bickelhaupt, A new all-round density functional based on spin states and S_N2 barriers, J. Chem. Phys. 131 (9) (2009) 094103.
[43] K. Morokuma, Molecular orbital studies of hydrogen bonds. III, J. Chem. Phys. 55 (3) (1971) 1236–1244.
[44] K. Kitaura, K. Morokuma, A new energy decomposition scheme for molecular interactions within the Hartree–Fock approximation, Int. J. Quantum Chem. 10 (2) (1976) 325–340.
[45] T. Ziegler, A. Rauk, On the calculation of bonding energies by the Hartree Fock Slater Method, Theor. Chem. Acta 46 (1977) 1–10.
[46] M. Kohout, Program DGrid, Version 4.5, Max-Planck Institut für Chemische Physik fester Stoffe, Dresden, 2009.
[47] C.F. Guerra, J.W. Handgraaf, E.J. Baerends, F.M. Bickelhaupt, Voronoi deformation density (VDD) charges: assessment of the Mulliken, Bader, Hirshfeld, Weinhold, and VDD methods for charge analysis, J. Comput. Chem. 25 (2) (2004) 189–210.
[48] S. Noury, B. Silvi, R.J. Gillespie, Chemical bonding in hypervalent molecules: is the octet rule relevant? Inorg. Chem. 41 (8) (2002) 2164–2172.
[49] M.E. Alikhani, F. Fuster, B. Silvi, What can tell the topological analysis of elf on hydrogen bonding? Struct. Chem. 16 (3) (2005) 203–210.
[50] I. Drebushchak, S. Kozlova, Hydrogen bond in the $H_3O(Ph_3PO)^{3+}$ complex. Features of the electron density distribution, J. Struct. Chem. 51 (1) (2010) 166–169.
[51] J.S. Prell, T.G. Flick, J. Oomens, G. Berden, E.R. Williams, Coordination of trivalent metal cations to peptides: results from IRMPD spectroscopy and theory, J. Phys. Chem. A 114 (2) (2009) 854–860.