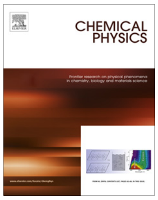
Chemical Physics 435 (2014) 21–28

Contents lists available at ScienceDirect

Chemical Physics

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



Computational Vibrational Spectroscopy of glycine in aqueous solution – Fundamental considerations towards feasible methodologies

Oliver M.D. Lutz [a], Christoph B. Messner [a], Thomas S. Hofer [b,*], Lorenz R. Canaval [b], Guenther K. Bonn [a], Christian W. Huck [a]

[a] Institute of Analytical Chemistry and Radiochemistry, Center for Chemistry and Biomedicine, Leopold-Franzens University, Innsbruck, Innrain 80-82, 6020 Innsbruck, Austria
[b] Theoretical Chemistry Division, Institute of General, Inorganic and Theoretical Chemistry, Center for Chemistry and Biomedicine, Leopold-Franzens University, Innsbruck, Innrain 80-82, 6020 Innsbruck, Austria

ARTICLE INFO

Article history:
Received 22 January 2014
In final form 5 March 2014
Available online 14 March 2014

Keywords:
Vibrational Spectroscopy
Molecular dynamics
VSCF
VPT2
QM/MM
QMCF/MD
FT-IR

ABSTRACT

In this work, the mid-infrared spectrum of aqueous glycine is predicted by a number of computational approaches. Velocity autocorrelation functions are applied to *ab initio* QMCF-MD and QM/MM-MD simulations in order to obtain IR power spectra. Furthermore, continuum solvation model augmented geometry optimizations are studied by anharmonic calculations relying on the PT2-VSCF and the VPT2 formalism. In this context, the potential based EFP hydration technique is discussed and the importance of a Monte Carlo search in conjunction with PT2-VSCF calculations is critically assessed. All results are directly compared to newly recorded experimental FT-IR spectroscopic data, elucidating the qualities of the respective methodology. Moreover, the computational approaches are discussed regarding their usefulness for the interpretation of experimental spectra.

1. Introduction

Computational infrared (IR) spectroscopic data can be obtained from either frequency calculations at geometries corresponding to a stationary point (i.e., typically an energy minimum) or molecular dynamics (MD) based approaches, with corresponding methodologies being readily implemented in a large number of computational chemistry software packages. One main reason for the computation of vibrational frequencies and their corresponding bond force constants is the fact, that the results are directly comparable to experimental infrared and Raman spectra, thus enabling an estimation of the theoretical approaches' quality [1–3]. Conversely, as long as the quality of the theoretical method is on par with the experiment, the computationally derived spectroscopic data may be utilized for the interpretation of the experimental results. This inverse approach is especially helpful in connection with complex samples where a thorough band assignment, solely relying on an empirical basis, is prohibitive.

Glycine (Gly), being the most simple amino acid, serves as a benchmark molecule in this work. A characteristic feature of amino acids is their preferential formation of a zwitterion when subjected to a polar solvation environment with high permittivity such as water [4]. A reliable description of this characteristic feature requires high quality computational approaches, namely *ab initio* molecular dynamics (MD) solvation treatments or sophisticated stationary point calculations incorporating implicit solvation models or even explicitly considered solvent molecules [4,5]. Glycine is a widely studied molecule [4–14], however, most of the works either focus on gas phase studies of the uncharged molecule or on solvation studies by means of either molecular-mechanics (MM) or quantum–mechanical/molecular-mechanical (QM/MM) hybrid MD simulations. Alper et al. [10] and Sun et al. [7] presented spectrosopic data from *ab initio* potential fitted MM-MD simulations and density functional theory (DFT) MD simulations of the solvated molecule, respectively. Investigating spectroscopic differences between the solid and the liquid state, Chowdhry et al. [15] compared frequencies derived from PW91 and PBE DFT based stationary point calculations of glycine with a polarizable continuum model (PCM) augmented Becke three parameter (B3LYP) based optimization.

The majority of articles investigating spectroscopic properties of glycine confine themselves on the harmonic approximation with further reaching frequency corrections [16,17] being the exception. Very recently, Barone et al. [14] provided theoretically derived infrared spectroscopic data of various glycine conformers by

---

* Corresponding author.
E-mail address: t.hofer@uibk.ac.at (T.S. Hofer).

http://dx.doi.org/10.1016/j.chemphys.2014.03.003
0301-0104/© 2014 Elsevier B.V. All rights reserved.

means of the extended vibrational second order perturbative treatment (VPT2). A different, equally profound discussion involving gaseous glycine, was presented by Shemesh and coworkers, elucidating the adsorption characteristics of glycine on the Si (100)-2×1 surface via anharmonic calculations based on the vibrational self-consistent field (VSCF) approach [18].

The aim of this work is the investigation of various computational methods for their qualities in terms of how they resemble the experimentally observed IR spectrum of aqueous glycine, thus critically assessing the usefulness of theoretical calculations for the interpretation of experimentally derived data. Since many conventional *ab initio* computations treat the system of interest *in vacuo*, such approaches are hardly useful for the interpretation of solution based experiments. In this work, the aqueous environment is modeled on the one hand by explicit water molecules described with the effective fragment potential (EFP) approach [19,20], stabilizing an energetically favoured structure and on the other hand by the conductor-like variant of the polarizable continuum solvation model (cPCM) [21,22]. Moreover, the quantum–mechanical charge-field (QMCF) MD package [23–27] is utilized to execute simulations of aqueous glycine at different levels of theory, enabling a direct comparison of the respective approaches.

2. Methods

2.1. IR spectra based on frequency computations at energy minima

The evaluation of spectroscopic data based on second derivatives of the energy with respect to the nuclear coordinates (i.e., the Hessian matrix) obtained at geometries corresponding to an energy minimum is readily implemented in many molecular mechanics and electronic structure programs. The resulting force constants, by approximation, rely on the potential energy V expressed as a second-order Taylor series:

$$V(Q) \approx V_0(Q) + g(Q) \cdot \Delta Q + \frac{1}{2} \Delta Q^T \mathbb{H} \Delta Q \tag{1}$$

with the gradient $g(Q)$ being zero at the energy minimum and

$$H_{r,s} = \frac{\partial^2 V(Q)}{\partial q_r \partial q_s} \tag{2}$$

Diagonalization of $\mathbb{H}$ yields the diagonal eigenvalue matrix $\mathbb{h}$ and the unitary matrix $\mathbb{U}$ containing the respective eigenvectors corresponding to the normal modes:

$$\mathbb{H} = \mathbb{U}^T \mathbb{h} \mathbb{U} \tag{3}$$

Typically, empirically derived scaling factors are applied to correct errors resulting from the harmonic approximation, the inadequate description of correlated electron motion in the large part of methods as well as from the quasi *in vacuo* treatment of the systems [28]. Equally, revised scaling factors have been established for a number of popular and widely employed basis sets [29]. It has to be mentioned, however, that the aforementioned scaling factors need to be treated with caution before being applied to complex systems since they have been fitted against a set of small molecules [30]. In order to account for anharmonicities on a more system-independent basis, popular electronic structure programs such as GAMESS [16], Gaussian [17] and NWChem [31] provide corrective techniques such as the vibrational self-consistent field (VSCF) approach and the vibrational second order perturbative treatment (VPT2) [1,32–36].

2.1.1. Energy minimizations

A crucial point when obtaining vibrational data via frequency computations is the chosen level of theory. Since the aforementioned anharmonic calculations may only be applied to previously optimized structures with subsequently conducted harmonic vibrational analyses, the necessary electronic energies discussed herein have been evaluated at three different levels of theory, namely Hartree–Fock (HF), Møller-Plesset second order perturbation theory (MP2) and Becke three parameter Lee–Yang–Parr density functional theory (B3LYP). Higher correlated methods such as configuration interaction (CI) or coupled-cluster (CC) methods [37] have not been considered due to their, yet, infeasible computational requirements, especially in conjunction with the calculation of second, third and fourth derivatives of the energy. Semi-empirical methods and fully empirical force-field approaches have been left aside due to their strongly parameterized nature and the resulting system dependency. Since the quality of a subsequent frequency evaluation is strongly dependent on a thoroughly minimized system geometry [38], the following options have been chosen in accordance with the recommendations found for the utilized electronic structure programs: GAMESS May 1st 2012 [16] "OPTTOL = 0.000001" and Gaussian 09 rev. C.01 [17] "Opt = VeryTight". For DFT calculations, the recommended keyword "Integral (Grid = UltraFineGrid)" was added for the Gaussian based optimizations.

2.1.2. Monte Carlo search

As it will be discussed later, especially the PT2-VSCF corrective technique strongly relies on a thoroughly minimized geometry. Employing conventional minimization algorithms, the successful discovery of a local minimum reflecting a substantially deep potential well or, at best, the global energy minimum is strongly dependent on the complexity of the studied system. Monte Carlo methods [39,40] are considered excellent approaches for an unbiased scanning of the complete potential energy surface (PES) since they do not rely on gradients of the energy. In this context, the GAMESS program package offers a valuable tool available via the keyword "GLOBOP", enabling a straightforward and highly customizable Monte Carlo search. The solvent molecules whose positions are moved in this approach are defined via the EFP technique [19,20]. GAMESS, in its current revision, offers two internally stored EFPs for water based on RHF and DFT calculations. In this work, the HF based global optimization was conducted with RHF EFP water molecules while the B3LYP and the MP2 optimizations made use of the DFT derived potentials. The initial global optimization has been extended by means of a simulated annealing approach (temperature gradient from 20000 K to 300 K in 50 steps). Each of the 50 steps contained 200 different solvation structures and each of these structures has been minimized locally [40], thus reflecting the basin-hopping technique [41]. An acceptance ratio of 0.535 was observed during the global optimization. The resulting minimum energy geometry has been employed exclusively for the subsequent PT2-VSCF calculations since Gaussian 09 rev. C.01 does not support the EFP solvation technique.

2.1.3. Anharmonic calculation formalisms

Starting from a thoroughly optimized system geometry, the VPT2 method enables a fully automated generation of third and semi-diagonal fourth derivatives and a subsequent computation of anharmonic frequencies as long as analytical second derivatives are available for the respective calculation formalism [42,43]. For Gaussian 09 C.01, this is the case for HF, MP2 and B3LYP levels of theory [17].

The VSCF approach [44,45] is available through GAMESS [16] as well as within the NWChem program suite [31], enabling the consideration of anharmonic effects occuring on an interatomic as well as on an intermolecular basis. The approach utilizes the assumption of the full vibrational wavefunction being fully factorizable into single-mode wavefunctions [46] and thus