Journal of Molecular Spectroscopy xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Journal of Molecular Spectroscopy

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

The structures of proton-bound dimers of glycine with phenylalanine and pentafluorophenylalanine

Weiqiang Fu [a], Jeffrey Xiong [a], Michael J. Lecours [a], Patrick J.J. Carr [a], Rick A. Marta [a], Eric Fillion [a], Terry McMahon [a], Vincent Steinmetz [b], W. Scott Hopkins [b,*]

[a] Department of Chemistry, University of Waterloo, 200 University Ave West, Waterloo, ON N2L 3G1, Canada
[b] Laboratoire Chimie Physique – CLIO, Bâtiment 201, Porte 2, Campus Universitaire d'Orsay, Orsay 91405, France

ARTICLE INFO

Article history:
Received 12 May 2016
In revised form 5 July 2016
Accepted 14 July 2016
Available online xxxx

Keywords:
Amino acid
IRMPD
Cation-π interaction
Basin hopping
Heterodimers

ABSTRACT

Infrared multiple photon dissociation spectra of the proton-bound heterodimers of phenylalanine/glycine (Phe·H$^+$·Gly) and pentafluorophenylalanine/glycine (F_5-Phe·H$^+$·Gly) have been acquired in the 975–1975 cm^{-1} region. Exhaustive basin hopping molecular dynamics searches and subsequent DFT calculations predict four low energy intermolecular binding motifs. The spectrum of F_5-Phe·H$^+$·Gly is assigned predominantly to isomers exhibiting a N—H$^+$···O intermolecular interaction between nitrogen-protonated Gly and the carbonyl oxygen atom of F_5-Phe. In contrast, the spectrum Phe·H$^+$·Gly is found to consist of a mixture of isomers exhibiting the N—H$^+$···O between nitrogen-protonated Gly and the carbonyl oxygen of neutral Phe, and isomers exhibiting a N—H$^+$···N interaction between nitrogen-protonated Phe and the neutral Gly moiety. Cation-π effects also seem to influence the cluster geometries, especially in the case of the Phe·H$^+$·Gly global minimum structure where the phenyl ring orients with the site of protonation so as to maximize charge-quadrupole interactions.

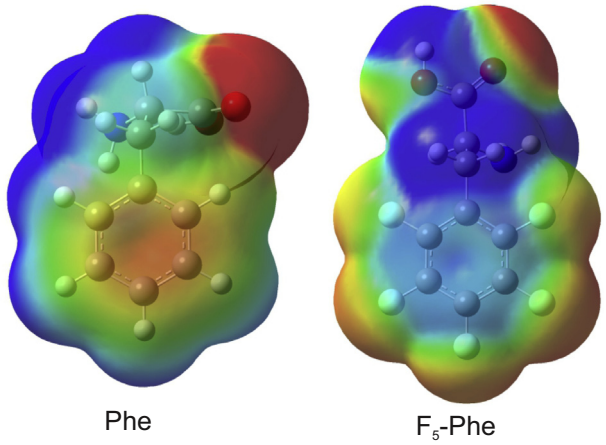
1. Introduction

Proton-bound dimers of amino acids offer a pared-down model for studying the fundamental non-covalent interactions that give rise to secondary structure in larger systems such as proteins [1]. IR spectroscopy, in particular, has found utility in many of the investigations to date owing to the fact that many structural features, *e.g.*, sites of protonation, exhibit diagnostic spectral signatures [2–12]. Of the proton-bound amino acid dimers studied, most have been homodimers (*i.e.*, composed of two of the same amino acid). Typically, these species exhibit a geometry featuring an intermolecular ionic hydrogen bond between the carbonyl of the neutral moiety and the ammonium group of the protonated moiety. This binding motif seems also to be preferred for the mixed heterodimers of alanine/glycine and serine/threonine [6,7]. There are, however, exceptions to this generalization. The spectrum of the proton-bound dimer of proline, for example, exhibits vibrational peaks that are indicative of not only the typical N—H$^+$···O interaction, but also of isomers exhibiting a N—H$^+$···N binding motif [9]. The proton-bound dimer of lysine also exhibits atypical binding geometry, seemingly adopting a salt-bridged zwitterion structure owing to interactions involving the side chain amino groups [8]. The same is true of the glycine/lysine heterodimer, in which glycine adopts a zwitterionic form owing to interaction with the side chain amino group of the lysine [6]. Further still, the serine/phenylalanine heterodimer seems to also adopt an atypical geometry, however, in this case, no definitive structural assignment has yet been made [1].

The rationale for selecting the phenylalanine/glycine (Phe·H$^+$·Gly) and pentafluorophenylalanine/glycine (F_5-Phe·H$^+$·Gly) protonated heterodimers for study is twofold. First, a survey of the experimentally determined binding trends observed across the homo- and heterodimers reported to date suggests that differences in gas phase basicity (GPB = $\Delta G°$) or proton affinity (PA = $\Delta H°$) between the two amino acids correlates with the observed binding motif [13]; if ΔGPB ≈ 0, the typical N—H$^+$···O interaction is likely to be observed, whereas if ΔGPB is more substantial (*e.g.*, ΔGPB = 98.8 kJ mol^{-1} for the glycine/lysine heterodimer), atypical binding is more common. In the case of the Phe·H$^+$·Gly heterodimer ΔGPB = 36.7 kJ mol^{-1}, with Phe exhibiting the larger GPB of the two amino acids (GPB_{expt} = 888.9 kJ mol^{-1}, GPB_{calc} = 896.3 kJ mol^{-1}). For the F_5-Phe·H$^+$·Gly system ΔGPB = 2.7 kJ mol^{-1} [GPB_{expt}(Gly) = 852.2 kJ mol^{-1}, GPB_{calc} (F_5-Phe) = 854.9 kJ mol^{-1}) [13]. Consequently, it is interesting to test the notion that F_5-Phe·H$^+$·Gly should exhibit typical N—H$^+$···O binding and that Phe·H$^+$·Gly should not. The second reason for studying Phe·H$^+$·Gly and F_5-Phe·H$^+$·Gly stems from the possibility that these two

* Corresponding author.
E-mail address: scott.hopkins@uwaterloo.ca (W.S. Hopkins).

http://dx.doi.org/10.1016/j.jms.2016.07.004
0022-2852/



Scheme 1. Electrostatic potential map of (left) Phe·H$^+$·Gly and (right) F_5-Phe·H$^+$·Gly as calculated at the B3LYP/6-311++G(d,p) level of theory. The heat map ranges from −0.02*e* (red) to +0.02e (blue). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

chemical systems might represent an interesting case study for cation–π interactions in non-covalently bound amino acid clusters [14,15]. This interaction, which here might better be called a charge–quadrupole interaction, has been identified as an important contributor to protein folding and molecular recognition. When fully substituted with fluorine, the negative quadrupole of benzene inverts ($\Theta_{C6H6} = -33.3 \times 10^{-4}$ C m^2; $\Theta_{C6F6} = 31.7 \times 10^{-4}$ C m^2) [16]. Complete fluorination of the phenyl ring in Phe is expected to similarly invert the local quadrupole of the π–system. This effect is visualized in Scheme 1, which shows a heat map of the electrostatic potentials for Phe·H$^+$·Gly and F_5-Phe·H$^+$·Gly mapped onto their respective total electron density distributions. Positively charged species will be attracted to the red coloured region(s) of the phenyl ring system. One can therefore hypothesize that in the Phe·H$^+$·Gly system the phenyl group will orient such that the ring face can interact attractively with the site of protonation (*i.e.*, the positive charge and the ring centre will attract), whereas in F_5-Phe·H$^+$·Gly it will be the ring edge that attractively interacts with the protonation site.

The "atypical" amino acid dimers hint that there is a more complex picture of inter- and intra-molecular interactions in amino acid clusters than that which we currently hold. There has been some progress towards separating the isomer contributions from convoluted spectra by UV/IR double resonance [1,17], mobility selection [18], and via isomer-specific and mode-selective infrared multiple photon dissociation (IRMPD) behavior [19,20]. However, these techniques are not widely utilized/available, nor are they applicable in all cases. For the most part, spectroscopists currently tend to rely on pairing traditional experimental techniques with exhaustive, high level quantum chemical calculations. This is the approach undertaken here; the outcomes of IRMPD spectroscopic experiments are supported by exhaustive basin-hopping (BH) searches of the potential energy surfaces of Phe·H$^+$·Gly and F_5-Phe·H$^+$·Gly.

2. Experimental methods

Infrared multiple photon dissociation (IRMPD) spectra of the proton-bound dimers of Gly with Phe and F_5-Phe were recorded at the Centre Laser Infrarouge d'Orsay (CLIO) free electron laser (FEL) facility [21,22]. Details of the experimental apparatus have been reported previously [8,23]. Briefly, the Phe·H$^+$·Gly and F_5-Phe·H$^+$·Gly clusters were produced by electrospray ionization (ESI) of ca. 100 μmol L^{-1} methanol:water (50:50 vol%) solutions containing Gly and either L-Phe or L-F_5-Phe, and ca. 5 μmol L^{-1} formic acid. Nascent clusters produced via ESI were then transferred to a Bruker Esquire 3000+ ion trap mass spectrometer where clusters of interest were mass-selected and irradiated with the tunable output of the FEL over the 900–2000 cm^{-1} region. By monitoring fragmentation efficiency as a function of the FEL wavenumber, IRMPD action spectra were recorded. IRMPD efficiency is defined as:

$$Efficiency = -ln\left[\frac{I_{par}}{I_{par} + I_{frags}}\right]$$

where I_{par} is the parent ion signal intensity and I_{frags} is the signal intensity of the fragment ions.

3. Computational methods

To support experimental outcomes, a custom-written basin-hopping (BH) molecular dynamics code was used to search the potential energy surfaces (PESs) of the proton-bound dimers of Gly with Phe and F_5–Phe [24,25]. The clusters were modelled with the UFF molecular mechanics force field, using atomic partial charges that were calculated at the B3LYP/6-311++G(d,p) level of theory using the ChelpG partition scheme [26–28]. For each random structural perturbation in the BH search, the charge-carrying proton and the Gly moiety was given a random displacement of -0.7 Å $\leqslant \eta \leqslant 0.7$ Å in each of the *X*, *Y*, and *Z* directions. The Gly moiety was also given a random rotation of $-5° \leqslant \theta \leqslant 5°$ about its *x*, *y*, and *z* body-fixed axes. To further search the conformational space of each amino acid, a random dihedral angle rotation of $-5° \leqslant \varphi \leqslant 5°$ was applied to each of the C—C bonds in the saturated carbon chain. Convergence criteria for the molecular mechanics calculations were set to the default criteria specified in the Gaussian 09 computational chemistry suite of programs [29]. In total, approximately 80,000 structures were sampled by the BH algorithm for each cluster. Owing to the fact that the BH algorithm targets the low energy region(s) of the potential energy surface, we expect that low energy structures have been faithfully sampled; each unique isomer was identified several hundred times by the BH search. Nevertheless, in any finite search, there is a chance that the global minimum structure will elude identification. However, given the good agreement between our observed and calculated spectra (*vide infra*), we believe that this is not the case here. By projecting the randomly generated geometries onto stationary points of the potential energy surface, the BH algorithm produces a set of candidate structures for low energy isomers. Unique isomers (some of which were structurally very similar at the molecular mechanics level of theory) that were identified by the BH search were then carried forward for treatment at the B3LYP/6-311++G(d,p) level of theory. To ensure that each stationary point was a local minimum on the PES, normal mode analyses were conducted, which also yielded thermodynamic corrections and harmonic vibrational spectra for each cluster isomer. This treatment produced 31 low energy isomers for the Phe·H$^+$·Gly cluster (25 structures within 50 kJ mol^{-1} of the global minimum) and 33 isomers for the F_5-Phe·H$^+$·Gly cluster (27 structures within 50 kJ mol^{-1} of the global minimum) at the DFT level of theory. The thermodynamic energies and Cartesian *XYZ* coordinates for all cluster isomers found via our BH search are provided in the supplementary information that accompanies this article.

4. Results and discussion

4.1. The proton-bound Dimer of Gly and L-Phe

The IRMPD spectrum of Phe·H$^+$·Gly in the 975–1975 cm^{-1} region is shown in Fig. 1A. The only observed photofragmentation