AIP | The Journal of Chemical Physics

Photoelectron spectra and structures of three cyclic dipeptides: PhePhe, TyrPro, and HisGly

Anoja P. Wickrama Arachchilage, Feng Wang, Vitaliy Feyer, Oksana Plekan, and Kevin C. Prince

Citation: J. Chem. Phys. **136**, 124301 (2012); doi: 10.1063/1.3693763
View online: http://dx.doi.org/10.1063/1.3693763
View Table of Contents: http://jcp.aip.org/resource/1/JCPSA6/v136/i12
Published by the American Institute of Physics.

Photoelectron spectra and structures of three cyclic dipeptides: PhePhe, TyrPro, and HisGly

Anoja P. Wickrama Arachchilage,[1] Feng Wang,[1] Vitaliy Feyer,[2,3] Oksana Plekan,[2,a)] and Kevin C. Prince[2,4]

[1]*eChemistry Laboratory, Faculty of Life and Social Sciences, Swinburne University of Technology, Hawthorn, Melbourne, Victoria 3122, Australia*
[2]*Sincrotrone Trieste, in Area Science Park, I-34149 Basovizza, Trieste, Italy*
[3]*Forschungszentrum Jülich GmbH - IFF - IEE, 52425 Jülich, Germany*
[4]*Istituto Officina dei Materiali, Consiglio Nazionale delle Ricerche, Area Science Park, I-34149 Trieste, Italy*

(Received 4 January 2012; accepted 25 February 2012; published online 22 March 2012)

We have investigated the electronic structure of three cyclic dipeptides: cyclo(Histidyl-Glycyl) (cHisGly), cyclo(Tyrosyl-Prolyl) (cTyrPro), and cyclo(Phenylalanyl-Phenylalanyl) (cPhePhe) in the vapor phase, by means of photoemission spectroscopy and theoretical modeling. The last compound was evaporated from the solid linear dipeptide, but cyclised, losing water to form cPhePhe in the gas phase. The results are compared with our previous studies of three other cyclopeptides. Experimental valence and core level spectra have been interpreted in the light of calculations to identify the basic chemical properties associated with the central diketopiperazine ring, and with the additional functional groups. The valence spectra are generally characterized by a restricted set of outer valence orbitals separated by a gap from most other valence orbitals. The theoretically simulated core and valence spectra of all three cyclic dipeptides agree reasonably well with the experimental spectra. The central ring and the side chains act as independent chromophores whose spectra do not influence one another, except for prolyl dipeptides, where the pyrrole ring is fused with the central ring. In this case, significant changes in the valence and core level spectra were observed, and explained by stronger hybridization of the valence orbitals. © *2012 American Institute of Physics*. [http://dx.doi.org/10.1063/1.3693763]

I. INTRODUCTION

We have previously reported a theoretical and experimental photoemission investigation of three cyclic dipeptides,[1] namely the smallest dipeptide, cyclo(Glycyl-Glycyl) (diketopiperazine, or DKP hereafter), cyclo(Leucyl-Prolyl), and cyclo(Phenylalanyl-Prolyl). Cyclic dipeptides are a subset of the much larger class of peptides, and many are biologically active. These dipeptides consist of a ring containing two peptide bonds and this ring structure confers high stability[2] and resistance to human digestion. This last property allows the dipeptides to be used as scaffolds for drugs, and they also possess a number of interesting biological properties including antiviral, antibiotic, and anti-tumour activity. Some marine organisms produce cyclopeptides,[3] and the compounds have shown promise for several pharmaceutical and other applications.[4–7] The rigid central ring limits conformational freedom,[7] making them tractable models of larger peptides.[3] We note that the adsorption of peptides on surfaces is a subject of wide current interest, and this includes dipeptides:[8–10] to investigate them in the condensed state, it is of great help to have reference data for the isolated molecules.

Structure dictates function, but intrinsic properties of biologically active molecules may be modified or masked by their aqueous environment or by their interactions with it. Investigation of free molecules in the gas phase permits the identification of intrinsic properties without perturbations due to interaction with solvents. The aim of this study is to determine the intrinsic structure-property relationship of a number of cyclic dipeptides, using the experimental[11–14] and theoretical methods[15–17] previously applied to modelling both valence and inner-shell electronic states of biomolecules.

The compounds selected for study, shown in Figure 1, are cyclo(Histidyl-Glycyl) (cHisGly), cyclo(Tyrosyl-Prolyl) (cTyrPro), and cyclo(Phenylalanyl-Phenylalanyl) (cPhePhe). These all consist of a core of the six-member DKP ring, and differ by the side chains that the constituent amino acids contain: hydrogen and imidazole in cHisGly, phenol and pyrrolidine in cTyrPro, and two phenyl groups in cPhePhe. Lucietto *et al.*[18] reported the results of an NMR and structure modelling study of cHisGly, as well as data on its pharmacological properties. They concluded that the energetically favourable conformers adopt a boat configuration in solution, with respect to the DKP ring. Zhu *et al.*[19] calculated the structures of all of the cyclic peptides consisting of the same two amino acids, including cPhePhe, and found that they adopted the boat conformation. We recently summarised other previous studies of cyclic dipeptides.[1]

The amino acids constituting the dipeptides under study here are glycine, proline, phenylalanine, tyrosine, and histidine, and the core level spectra and accompanying theoretical calculations of the first four of these have been reported recently,[13,14] as well as the spectra of the linear

[a)]Present address: Aarhus University, Department of Physics and Astronomy, Ny Munkegade 120, 8000 Aarhus C, Denmark.

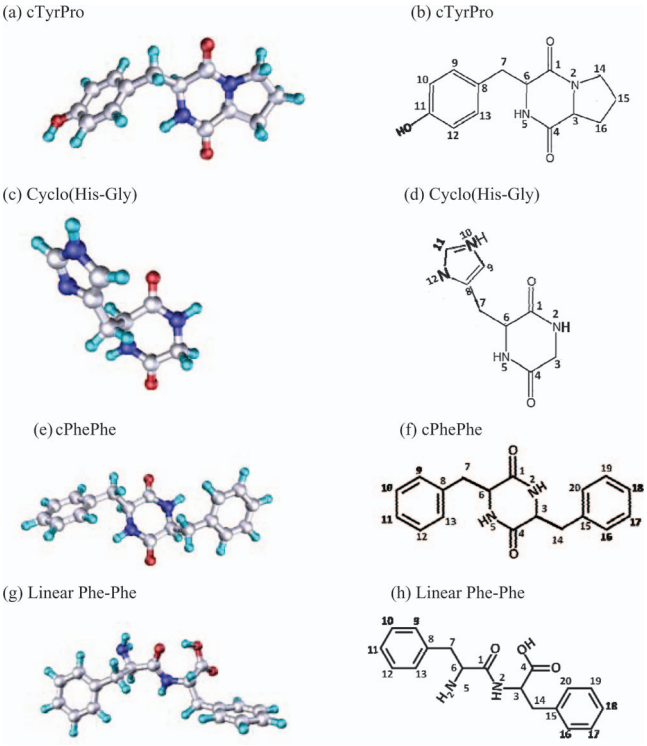

FIG. 1. (a)–(h) Three-dimensional views and schematic structures of cTyrPro, cHisGly, cPhePhe, and lPhePhe. Double click the three-dimensional figures on the left to view the interactive 3D structures.

peptide Glycyl-Glycine.[20] Here we investigate cyclic dimers of these amino acids. cHisGly can also be considered as DKP modified by the addition of an imidazole ring. In cTyrPro, the pentagonal pyrrolidine ring is fused to the DKP ring to produce a rather rigid double ring structure, a motif also found in two of our previously studied compounds, cyclo(Prolyl-Phenylalanyl) and cyclo(Leucyl-Prolyl). In proline, the four lowest energy conformers are constructed from two structural elements: the rotation of the carboxylic acid group and the puckering of the pyrrolidine ring.[21] The carboxylic group is not present in the dipeptide, so we expect that the main conformational element will be the puckering of the proline ring to an "up" or "down" geometry with respect to the twist in the DKP ring, giving rise to two possible conformers. In addition, the phenol moiety of cTyrPro contributes to the conformational landscape. The third compound, cPhePhe, contains two aromatic side groups; like DKP, it is formed from two identical amino acids.

II. EXPERIMENTAL AND COMPUTATIONAL METHODS

The measurements were performed at the Gas Phase Photoemission beamline, Elettra, Trieste,[22] using apparatus and calibration methods described previously.[11–13] The samples were supplied by Bachem (www.bachem.com) and used without further purification. The compounds used were those nominated except for cPhePhe, for which the sample inserted in the furnace and heated was Phenylalanyl-Phenylalanine a linear, non-cyclic dipeptide. They were evaporated at temperatures of 473 (cHisGly), 448 (cTyrPro), and 443 (lPhePhe) K, respectively, and checked for signs of thermal decomposition (spectral changes as a function of time, discoloration after heating, etc.). No evidence was found for decomposition of cHisGly and cTyrPro, but for the third compound, excessive degassing of water was observed. This was interpreted as a sign of structural alteration, and it will be shown below that this was indeed the case, as cyclisation with water elimination was occurring. No other sign of thermal change was observed for this compound, such as discoloration or spectral changes as a function of time, once the water was eliminated. Details of the energy resolution and calibration procedure have been given in Refs. 23 and 24, in brief, the resolution was estimated to be 0.32, 0.46, and 0.78 eV at $h\nu$ 382 (C 1s), 495 (N 1s), and 628 eV (O 1s), respectively.

The computational methods have been discussed previously,[1] and will not be repeated here but are summarized briefly. All the geometries were optimized using the B3LYP/cc-pVTZ model, followed by harmonic vibrational frequency calculations, which is incorporated in GAUSSIAN 03 computational package.[25] The interactive 3D-pdf structures for the molecules, shown in Figure 1 are produced as described by Selvam *et al.*[26] Adobe Acrobat 8.1 or higher is suitable for viewing. Single point calculation was performed, based on the LB94/et-pVQZ model,[27,28] which is incorporated in the Amsterdam Density Functional (ADF) computational chemistry package[29] to produce core vertical ionization energies. The outer valence vertical ionization energies of the dipeptides were calculated using the outer valence Green's function OVGF/6-31G* model,[30–34] while the complete valence vertical ionization energies were calculated using the HF/6-311G** model.

III. RESULTS

Figure 1 displays the optimized structures of the three dipeptides in three-dimensional space using a recently developed interactive 3D-pdf technique.[26] Double clicking Figures 1(a), 1(c), 1(e), or 1(g) online or the pdf file will allow viewing of the structures embedded in the pdf file. As can be seen from their 3D structures, all the cyclic dipeptides share a common six-member DKP ring, which prefers a boat conformation. The dipeptides, cTyrPro, cHisGly, and linear PhePhe, exhibit C_1 point group symmetry, while cPhePhe adopts a boat conformation, with C_2 point group symmetry.

Table I of the supplementary material[35] presents a summary of some characteristic geometric parameters of the dipeptides. Although all the cyclic dipeptides contain a six-member DKP ring, their attached functional groups are not the same. For example, cHisGly is composed of a DKP ring, connected to a five-member imidazole ring via the C(8)-C(7)-C(6) carbon bridge. The DKP ring of cTyrPro is fused with a pyrrole ring sharing the N(2)-C(3) bond, while the other end of the C(8)-C(7)-C(6) carbon bridge is a *para*-phenol moiety. As stated, cPhePhe exhibits C_2 point group symmetry: the symmetric DKP ring links two phenyl moieties through the C-C-C bridge. Regardless of the type of functional moiety linked to the DKP ring, the bridge dihedral angle, C(1)-C(6)-C(7)-C(8), does not change significantly for cTyrPro and

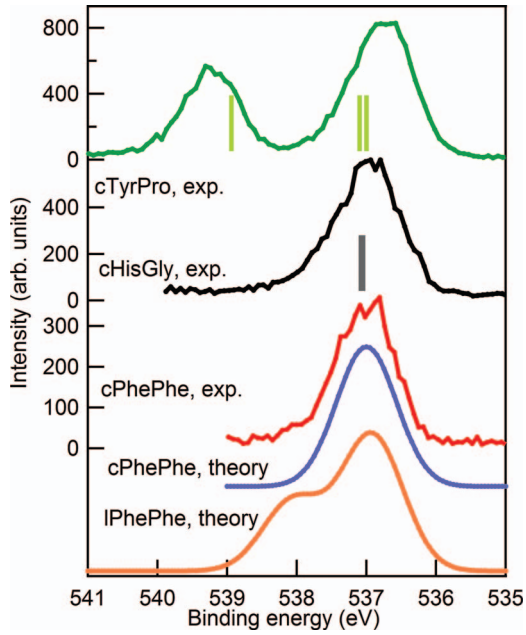

FIG. 2. O 1s spectra of cPhePhe, cHisGly, and cTyrPro. Theoretical curves (cPhePhe, linear PhePhe) and histograms (cHisGly, cTyrPro) are also shown. The FWHM of the simulated spectra of cPhePhe and linear PhePhe is 0.87 eV. The shifts of the theoretical spectra are: cPhePhe, +2.95 eV; linear PhePhe, +2.2 eV; cHisGly, +2.97 eV; cTyrPro, +3.04 eV.

cPhePhe but is different in cHisGly, where the imidazole ring is bonded to the DKP ring.

Briefly, the same functional groups as those in the previously studied dipeptides have very nearly the same structural parameters. For example, the perimeters of the central DKP ring and the benzene ring do not vary by more than 0.03 and 0.01 Å, respectively.

In Figures 2 and 4, we show the experimental and theoretical O, N, and C 1s photoemission spectra of the three compounds, and key data are summarized in Tables I–III; Table 2 of the supplementary material[35] gives a complete listing. Theoretical spectra have been offset by the energies indicated in the captions and tables, to give the best overall agreement with experiment. In the calculated O 1s spectra of linear and cyclic PhePhe, the linear form contains two peaks with an intensity ratio of 2:1, with the larger peak due to the carbonyl and peptide oxygen atoms, and the weaker peak due to the hydroxy oxygen; the spectrum is similar to that of the linear peptide GlyGly.[20] Our theoretical spectrum of cPhePhe agrees with the experimental data (within an offset of 3.00 ± 0.05 eV for all compounds), which shows a single peak at binding energy 537.0 eV, whereas the lPhePhe spectrum is quite different from the experimental data. We therefore conclude that on heating the linear form cyclised with the loss of water, so that the cyclic form is produced in the vapor, and we refer to this sample as cPhePhe from here on.

The O 1s spectrum of cHisGly also shows a peak at 537.0 eV, and again this is a single peak as expected from the similarity of the chemical environments of the two oxygen atoms. The theory predicts that the splitting is only 0.27 eV, clearly too small to be resolved experimentally.

The spectrum of cTyrPro contains two O 1s peaks, of which the stronger at 536.8 eV is assigned to the two peptide oxygen atoms, while the weaker feature at 539.2 eV is due to the oxygen atom in the phenol side chain. The theoretical splitting of the two peptide oxygen peaks is 0.1 eV, again too small to be resolved. The experimental intensity ratio is 1.8, slightly below the stoichiometric value of 2. We attribute this small difference to the transfer of oscillator strength from the main peak to satellites, which are stronger for peptide oxygen compared to the hydroxyl oxygen.

All three compounds show peaks due to the peptide oxygen atoms, with binding energies within 0.2 eV of one another, and these values are very similar to those of the cyclopeptides c(Leucyl-Prolyl) and c(Phenylalanyl-Prolyl) studied previously.[1] We conclude that these O 1s binding energies do not vary much within this class of compounds, and are substantially independent of the amino acid side chain, provided it is larger than a hydrogen atom. DKP, where the side chains are hydrogen atoms, showed higher O 1s binding energy (537.45 eV), attributed to reduced screening in the final state.

The experimental and calculated N 1s core level spectra are shown in Figure 3 and energies are listed in Table II. The vertical bars are the calculated N 1s core levels for cHisGly and cTyrPro whereas simulated core N 1s spectra are given for cPhePhe and lPhePhe. For cPhePhe, theory predicts a single spectral peak, and this is observed. The alternative linear structure contains a peptide and an amino nitrogen atom, and the calculated spectrum has two peaks, which were not observed experimentally. This is further evidence that the sample is cyclic in the gas phase.

cHisGly contains four nitrogen atoms, two of which are in the central DKP ring, and are chemically very similar. The other two are amino and imino nitrogen atoms in the imidazole side chain. Thus we expect three peaks with a stoichiometric ratio of 1:2:1, and experimentally three peaks are observed, with fitted intensities of 0.85:2.24:0.90. This is reasonable agreement, considering that the peaks are not fully

TABLE I. Experimental and theoretical O 1s vertical ionization energies of cPhePhe, cHisGly, and cTyrPro and theoretical values for linear PhePhe (eV). Theoretical values have been calculated with the LB94/et-pVQZ model. O 1s values are shifted by +2.95 eV (cPhePhe); +2.2 eV (linear PhePhe); +2.97 eV (cHisGly); +3.04 eV (cTyrPro).

cPhePhe		lPhePhe	cHisGly		cTyrPro	
Expt.	Theory	Theory	Expt.	Theory	Expt.	Theory
537.0	O(1) 537.00	O(1) 537.02	537.0	O(1) 536.98	536.8	O(1) 537.00
	O(4) 537.0	O(4) 536.82		O(4) 537.01		O(4) 537.10
		O(4)-H 538.06			539.2	O(11) 538.94

TABLE II. Experimental and theoretical N 1s vertical ionization energies of cPhePhe, cHisGly, and cTyrPro and theoretical values for linear PhePhe (eV).). Theoretical values have been calculated with the LB94/et-pVQZ model. N 1s values are shifted by +1.31 eV (cPhePhe); +1.35 eV (lPhePhe); +1.85 (cHisGly); +1.07 (cTyrPro).

cPhePhe (Expt.)	cPhePhe (Theory)	lPhePhe Theory	cHisGly (Expt.)	cHisGly Theory	cTyrPro (Expt.)	cTyrPro (Theory)
405.65	N(2) 405.65	N(2) 406.38	404.45	N(12) 403.74	405.45	N(2) 406.44
	N(5) 405.65	N(5) 403.08				N(5) 405.29
			405.80	N(2) 405.44		
				N(5) 405.52		
			406.75	N(10) 405.85		

resolved. The calculation underestimates the splitting between the peptide and amino nitrogen core levels, N(2) and N(5) versus N(10).

cTyrPro also displays a single peak due to the two peptide nitrogen atoms at 405.45 eV, and the calculated splitting is 0.15 eV. The center of the peak is 0.2 eV lower in binding energy than the peak of cPhePhe, and about 0.3 eV lower than the corresponding peak in cHisGly. Thus the order of the peptide N 1s binding energy is cTyrPro < cPhePhe < cHisGly.

Figure 4 shows the theoretical and experimental C 1s spectra, and the energies are listed in Table III. In the spectrum of cPhePhe there are three peaks: the molecule contains 18 carbon atoms, with nine chemically distinct types of atoms for the conformation with C_2 symmetry. However, the chemical environments of the carbon atoms in the two phenyl rings are very similar, and the core levels of C(7) and C(14) are unlikely to be much shifted from the energies of the phenyl carbon atoms, so we expect a peak due to 14 carbon atoms. C(1) and C(4) are chemically equivalent, and so are C(3) and C(6). These considerations predict three peaks with a stoichiometric ratio of 2:2:14, as predicted by theory. The experimental ratio, obtained by fitting the peaks and integrating the areas, is 1.9:2.2:13.9, in good agreement with the calculation and expectation.

The theoretical spectra for both linear and cyclic PhePhe are shown, and although the spectrum for the linear form agrees better with the experimental data, we have already established above that the cyclic form is present in the gas phase. The calculations generally tend to predict a more compressed energy scale than that observed in the experiment, and this is the case for cPhePhe.

In cHisGly, there are eight carbon atoms, giving rise to three peaks in the experimental spectrum. The two carbonyl carbon atoms of the central ring, C(1) and C(4), appear as a single peak at 293.75 eV; the theoretical splitting is 70 meV. The peak at 292.00 eV is assigned to the carbon atoms in the central ring, C(3) and C(6), with a contribution from C(11), located between the two nitrogen atoms of the imidazole ring. The peak at lowest binding energy is composed of contributions from C(7), C(8), and C(9). The latter two are bonded to nitrogen in the imidazole ring, while C(7) is the only carbon atom bonded solely to carbon and hydrogen, with no bonds to nitrogen or oxygen.

The cTyrPro molecule contains 14 carbon atoms and three peaks are observed experimentally. The peak at 293.34 eV is assigned to the carbonyl atoms of the central ring and it has an intensity ratio of 2.1:11.9 with respect to the other peaks, close to the expected value of 2:12. The binding energy is lower than that of cHisGly and cPhePhe and closer to the values measured for cyclo(Leucyl-Prolyl), 293.52 eV, and cyclo(Phenylalanyl-Prolyl), 293.40 eV.[1] cTyrPro has in common with these compounds a fused pyrrolidine ring so we conclude that this structural feature tends to decrease the C 1s binding energy by a small amount. The second peak consists of C 1s emission from C(3), C(6), and C(11), which are located in the central ring, or bonded to oxygen in the phenol ring, C(11). A peak due to C(14), the carbon atom in the pyrrolidine ring which is bonded to nitrogen, is not resolved, but theory predicts it to lie between the two main peaks.

TABLE III. Experimental and theoretical C 1s vertical ionization energies of cPhePhe, cHisGly, and cTyrPro and theoretical values for linear PhePhe (eV). Theoretical values have been calculated with the LB94/et-pVQZ model. C 1s values are shifted by +0.50 eV (cPhePhe); +0.66 eV (lPhePhe); +0.75 eV (cHisGly,); +0.24 eV (cTyrPro).

	cTyrPro			cHisGly				cPhePhe			
Atom	Theory	Expt.	Relative intensity	Atom	Theory	Expt.	Relative intensity	Atom	Theory	Expt.	Relative intensity
Other 8 atoms	289.70-290.18	290.26	7.37	C(7), C(8), C(9)	290.24-290.39	290.96	2.84	Other 14 atoms	289.70-290.11	290.34	13.94
C(14)	290.92	-		C(6)	291.46			C(3), C(6)	291.60	291.96	2.17
C(3)	291.46	} 291.65	4.51	C(3)	291.40	} 292.00	} 3.2	C(1), C(4)	292.64	293.73	2.11
C(6)	291.51			C(11)	291.03						
C(11)	291.44			C(1)	292.63	} 293.75	} 1.96				
C(1)	292.54	} 293.34	2.11	C(4)	292.56						
C(4)	292.39										

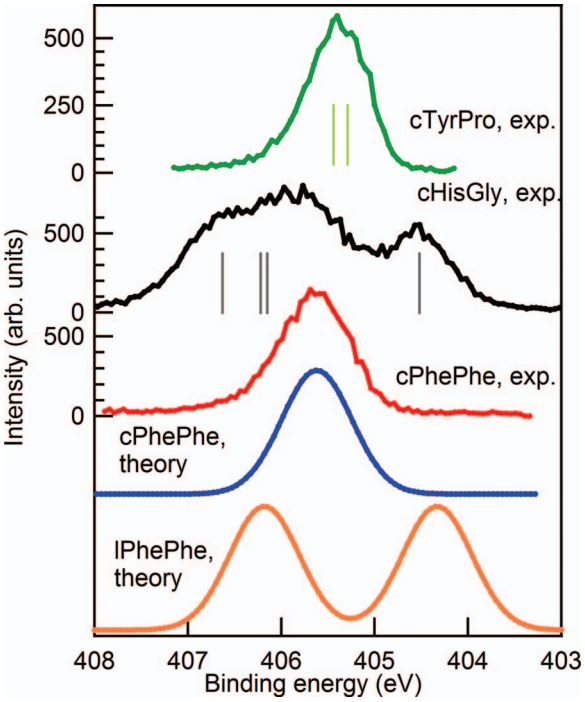

FIG. 3. N 1s spectra of cPhePhe, cHisGly, and cTyrPro. Theoretical curves (cPhePhe, linear PhePhe) and histograms (cHisGly, cTyrPro) are also shown. The shifts of the theoretical spectra are: cPhePhe, +1.31 eV; linear PhePhe, +1.35 eV; cHisGly, +1.85 eV; cTyrPro, +1.07 eV.

Finally the main peak is due to all carbon atoms which are not bonded to an electronegative O or N neighbor.

The binding energies are reasonably well predicted, with a contraction of the scale for the carboxylic carbon atoms, but the intensity ratios of the peaks are less well reproduced by the calculation. This could have two origins: firstly the method of plotting is very simple, with a single Gaussian width being used for all peaks, whereas in reality the peak shape is a Franck-Condon envelope. Secondly, the main peak is composed mainly of emission from the aromatic carbon atoms of the phenol ring. The satellite intensity from phenol and related molecules is stronger than for unconjugated atoms due to π-π excitations[36] so some loss of intensity from the main peak is expected, although this was not a major effect for cPhPhe.

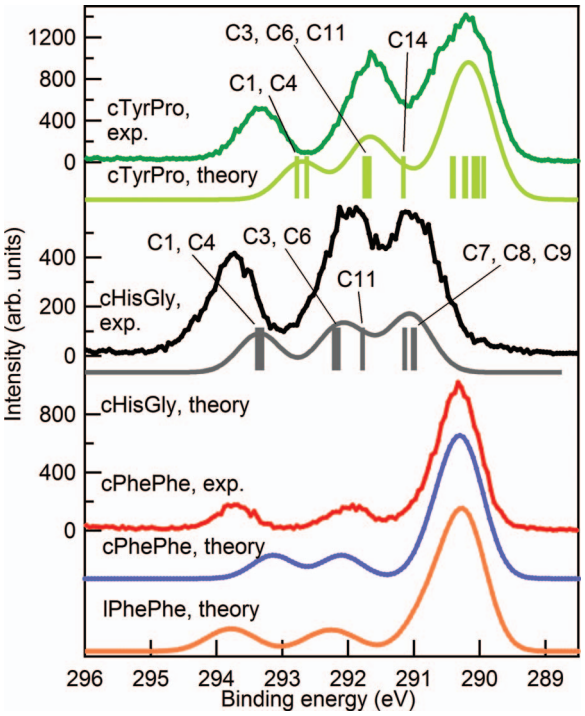

FIG. 4. C 1s spectra of cPhePhe, cHisGly, and cTyrPro. Theoretical curves and histograms are also shown. The shifts of the theoretical spectra are: cPhePhe, +0.50 eV; linear PhePhe, +0.66 eV; cHisGly, +0.75 eV; cTyrPro, +0.24 eV.

Figure 5 compares the measured and simulated valence photoemission spectra of the three compounds in the outer valence region to 16 or 18 eV of binding energy. The complete valence vertical ionization potentials (IPs) of cPhePhe, lPhePhe, cHisGly, and cTyrPro calculated using the HF/6-311G** model and the outer valence vertical ionization energies using the OVGF/6-31G* model are given in Table 3 of the supplementary material.[35] The spectroscopic pole strengths of these IPs calculated using the OVGF/6-31G* model are all larger than 0.85, indicating that the single particle approximation used in the models is valid and appropriate in this study. The outer orbital ionization energies are also shown in the figures as vertical bars and the simulated spectra have been produced by assuming the same width and intensity for all states. This is clearly a very crude approximation as widths (Franck-Condon envelopes) and cross-sections vary. However, the observed spectra are congested and not resolved in this region. The simple approximation of equal spectral line width in the simulation assists the assignment of the spectral features in the measurements.

The shape of the valence ionization spectra of the three cyclic dipeptides in Figure 5 exhibits similar patterns consisting of an intense peak in the region <11 eV, the outermost valence region, which is separated by an energy gap from the broader peaks formed by other valence ionic states. Detailed correlation of the spectra with their individual structures can be further refined by theory. Here we concentrate on the outermost molecular orbitals which are those most relevant for chemistry.

The highest energy part of the cPhePhe spectrum consists of a single peak with two weak shoulders. The calculation predicts a two peak structure, but this difference may be due to the assumption of constant width. This band has contributions from eight ionic states: the charge density of the parent molecular orbitals is shown in Figure 6. The Highest Occupied Molecular Orbital (HOMO) and indeed the first four orbitals consist of π orbitals of the phenyl rings. They are derived from the HOMO and HOMO-1 orbitals of benzene in combinations which are in- and out-of-phase, to give four orbitals. The fifth highest orbital, 37b, is mostly localized on the central ring and has mostly oxygen p (in plane) character, with some admixture of other DKP orbitals. The sixth highest orbital, 37a, also has predominantly in-plane oxygen p character, but is much more strongly mixed with the orbitals of the two phenyl side chains and the DKP ring. Finally the seventh and eighth highest molecular orbitals consist mostly of states constructed from + and − combinations of O 2p and N 2p atomic states, with some admixture of p states localized on the C–C bonds of the DKP-phenyl bonds. The first ionization potential is calculated to be 8.67 eV, in good

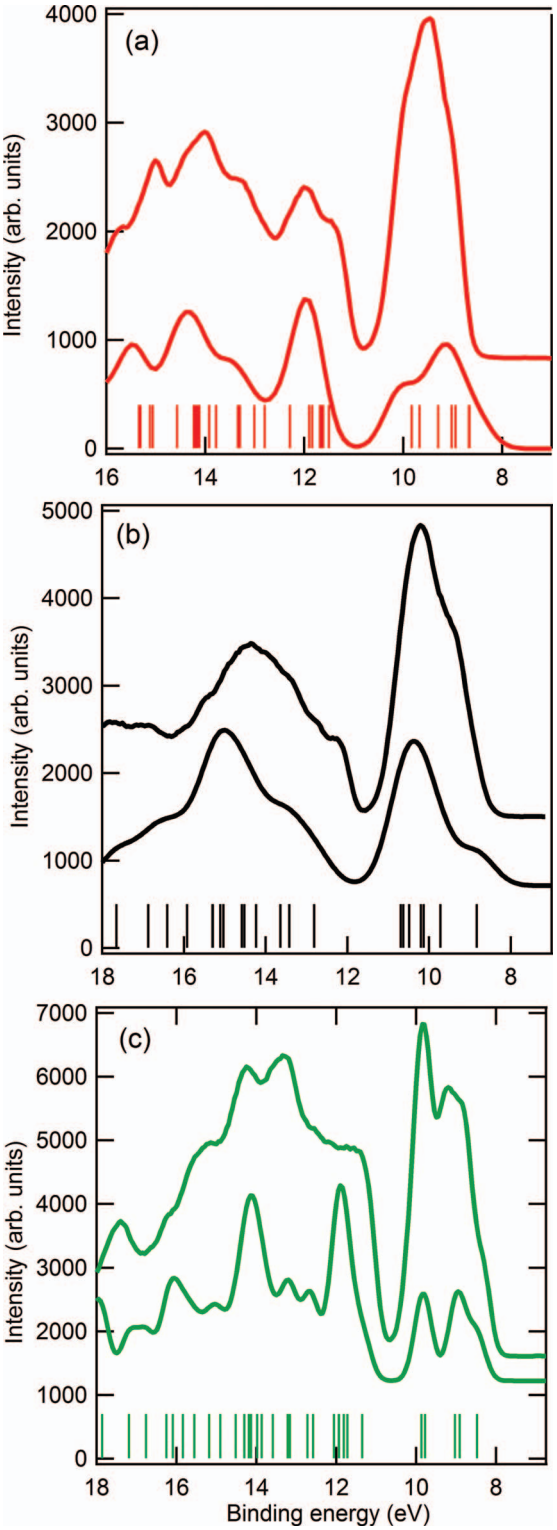

FIG. 5. Valence band spectra of cPhePhe, cHisGly, and cTyrPro. Photon energy 100 eV. (a) cPhePhe; (b) cHisGly; (c) cTyrPro. Top curves: experiment; middle curves: theory; bottom histograms: calculated ionization energies, OVGF/6-31G* model.

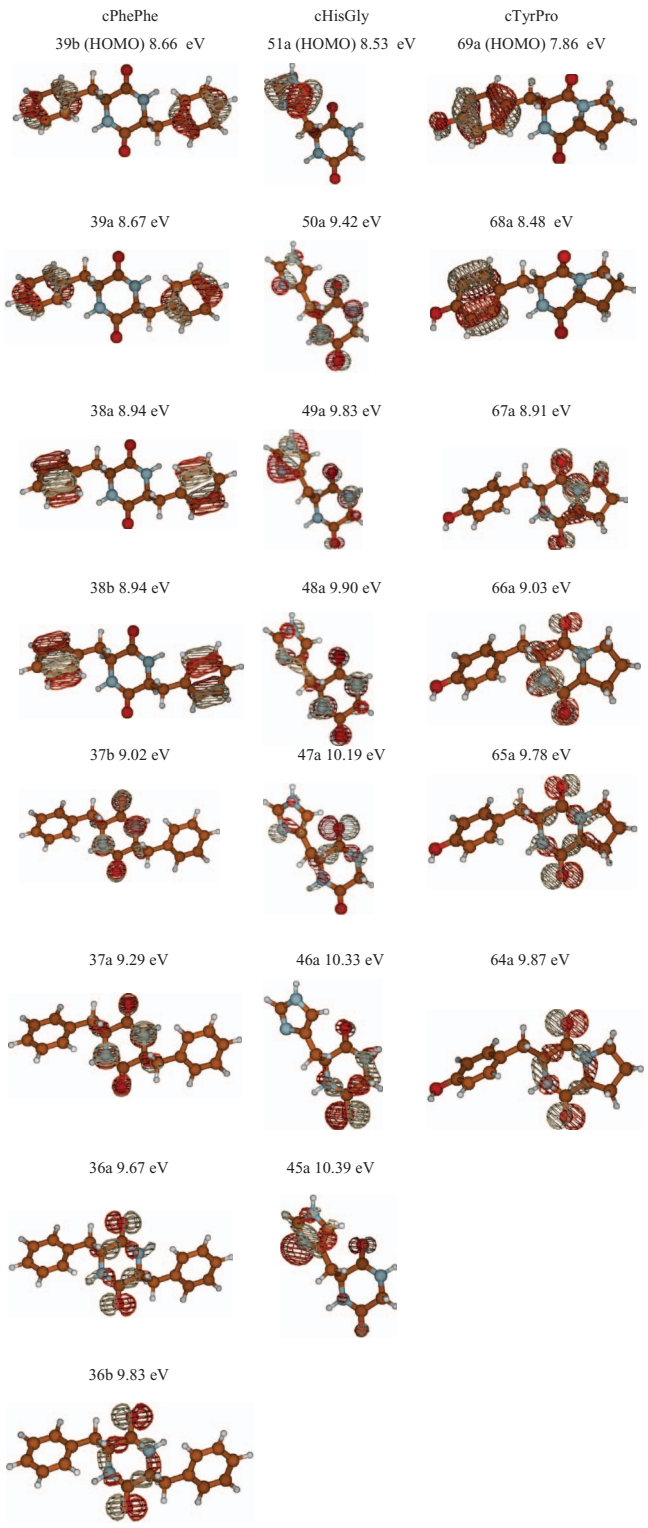

FIG. 6. The charge distribution of the frontier valence orbitals of the dipeptides: (a) cPhePhe; (b) cHisGly; (c) cTyrPro.

agreement with the experimental value of 8.93 eV. The calculation slightly overestimates the energy gap between the eight highest state and the rest of the valence band, for example, the energy difference between MOs 36b and 35a is 1.67 eV from the calculation but the energy is 1.23 eV according to the experiment. This discrepancy was observed in our previous study[1] for three other cyclopeptides, and is again observed in the present samples.

For cHisGly, the upper band due to ionization of frontier orbitals consists of seven states. The HOMO (MO 51a) and HOMO-6 (orbital 45a) are localized mostly on the unsaturated bonds in the imidazole ring. The HOMO-1 (orbital 50a) is in contrast an almost pure DKP orbital, with substantial

oxygen and nitrogen *p* character. Similarly orbitals 49a and 47a are mainly localized on the DKP ring, but contain a greater admixture of side chain orbitals. Orbital 48a has its greatest weight on the imidazole ring, but includes substantial contributions from the DKP ring.

The experimental first IP of cHisGly, determined by fitting this part of the spectrum with three peaks, is 9.41 eV, in poor agreement with the theoretical value of 8.53 eV, Table 3 of supplementary material. For cTyrPro and cPhePhe, the theoretical values are 0.3–0.4 eV lower than the experimental values, but for cHisGly, the difference is 0.88 eV. To investigate the origin of this discrepancy, we compare our results with those for related systems. As noted, the HOMO is localized on the imidazole group, so histidine and imidazole are good models. Recently Wilson *et al.*[36] determined a value of 8.2 ± 0.1 eV for the first IP of histidine, due to ionization of a MO localized on the imidazole moiety.[37,38] A major difference between the method of Wilson *et al.* and the present experiments and calculations is that they determined the adiabatic IP, whereas in the present work, we consider the vertical IP, which necessarily lies at equal or higher energy. We can estimate the adiabatic IP from our measurements by extrapolating the linear part of the spectrum near threshold to zero: we obtain a value of 8.4 ± 0.1 eV. Although this is similar to the value of Wilson *et al.*, and in very good agreement with the calculated value, the agreement is partly fortuitous, since the calculated value is the vertical ionization energy, and the experimental value is the adiabatic energy. Returning to the discrepancy between the experimental value determined by fitting peaks and the calculated value, we believe that the disagreement is due to a combination of factors: the high density of states does not allow precise fitting of individual peaks, and we do not know the shape of the Franck-Condon envelopes of the ionic states. This seems to affect the modeling of the HOMO more than the other higher energy states.

We note that the reported theoretical values of the vertical first IP of histidine using the OVGF model are 8.43 and 8.55 eV for different conformers,[37,38] similar to the present calculated value for the first IP of cHisGly, 8.53 eV.

There are five frontier orbitals of cTyrPro. The experimental first IP is 8.28 eV, and the theoretical value is 7.86 eV, a similar difference to that found for cPhePhe. The HOMO (orbital 69a) and orbital 66a consist of contributions from the π system of the phenol ring, and are rather similar in orbital composition to the two highest molecular orbitals of cPhePhe and orbitals 37a plus 38b, respectively. The other three orbitals are localized mainly on the DKP ring, with some contributions from the pyrrole ring.

IV. DISCUSSION

Comparing our present core level photoemission results with those of our previous work,[1] we note that the C 1s binding energies of the two carbonyl carbons in the DKP ring, C(1) and C(4), have their highest binding energies in the parent compound (cGG). In each compound the carbonyl atoms have the highest C 1s binding energy. We have previously attributed the higher binding energy in DKP with respect to the other cyclic dipeptides to the lack of side chains, which otherwise provide electrostatic screening of the core hole in the final state. The next highest binding energies of C(1) and C(4) are found in cHisGly (293.75 eV) and cPhePhe (293.73 eV). Since the side chains of these compounds are chemically diverse, but the binding energies are rather constant, we expect that most cyclo dipeptides without a fused ring will have very similar core binding energies. The nature of the side chain does not appear to influence strongly the chemical environment at C(1) and C(4).

The lowest binding energies of C(1) and C(4), 293.34 to 293.52 eV, occurred for the three compounds in which one constituent amino acid is proline, and thus the dipeptide has a rigid, fused ring structure. This structural element appears to lower the C(1)/C(4) energy significantly. The other amino acids constituting these three peptides were tyrosine, leucine, and phenylalanine.

The other two carbon atoms of the DKP ring, C(3) and C(6) (corresponding to C(6) and C(9) in our previous work)[1] have experimental binding energies, which vary only a little with the side chains. Although chemically distinct, the spectral peaks are not resolved. As usual DKP has the highest binding energy of 292.5 eV. The proline-containing dipeptides have binding energies in the range 291.65 (cTyrPro) to 291.79 eV, while the other two have slightly higher binding energies of 292.00 (cHisGly) and 291.96 (cPhePhe) eV.

The N 1s core levels of five of the six compounds (including the three compounds in the previous study[1]) fall in the narrow range from 405.45 to 405.65 eV, including molecules with and without proline. Only DKP has a higher binding energy of 406.29 eV, again due to lack of core hole screening by side chains. Thus the chemical effects of the amino acid side chains on nitrogen appear to be even smaller for nitrogen than for carbon atoms C(1) and C(4). The carboxyl oxygen of the central ring is similar: for five out of six of the compounds, the O 1s binding energy falls in the range from 536.8 to 537.0 eV, while for cGG it is 537.45 eV.

The valence band spectra of all six dipeptides consist of 4 to 8 orbitals clustered at low binding energy, IP < 11 eV, followed by an energetic gap, and then by the remaining valence orbitals. As shown previously,[1] four of these orbitals are present in the parent compound, DKP (or cGG as denoted previously). They consist of two symmetric and two antisymmetric (under C_2 rotation) molecular orbitals, labelled 14a, 14b, 15a, and 15b. (Note that in our previous work, the caption of Figure 3(a) for cGG should read 15b, not 15a.) For dipeptides based on proline, we observed strong mixing of the orbitals of the DKP ring with the pyrrole ring of proline. This feature occurs again for cTyrPro, and its frontier orbitals resemble those of c(Phenylalanyl-Prolyl). The energy ordering is different, because the OH group in cTyrPro alters the charge distribution and energy of the three benzene-derived molecular orbitals.

For non-prolyl dipeptides, the first four molecular orbitals retain much more of their DKP character, although the lower energy molecular orbitals of the side groups change the energy ordering. For instance, orbitals 37b, 37a, 36b, and 36a of cPhePhe are very clearly derived from the four highest MOs of DKP, and retain the same energy order. The other four frontier orbitals have mostly phenyl-derived character.

For cHisGly, orbital 50a, the HOMO-1, corresponds to the HOMO of DKP, and orbitals 49a, 47a, and 46a correspond to the next three highest MOs of DKP. The remaining orbitals, 51a, 48a, and 45a are predominantly derived from the imidazole ring of histidine.

Finally we note that our observation of cyclization of linear PhePhe is relevant to some recent condensed matter studies.[9,10] These authors evaporated linear PhePhe at 400 K onto copper, whereas we used a moderately higher temperature of 443 K. It is thus possible that the species on the surface in these studies was the cyclic rather than the (assumed) linear form of the molecule, depending on whether water loss occurs at 400 K. Unfortunately we do not have data for the latter temperature.

V. CONCLUSIONS

We have carried out a theoretical and experimental study of the electronic structure of three cyclic dipeptides, and compared the results with our previous work. The theory allows us to assign all of the observed core and part of the valence level spectra. All dipeptides show core levels shifted to lower energy, compared with the parent compound, cGG or DKP. When one of the constituent amino acids is proline, the structure contains fused DKP and pyrrole rings, and compounds in this class showed significant and systematic core level shifts of the atoms in the DKP ring. In the absence of a pyrrole ring, binding energies tended to be very similar, indicating that there is no strong interaction between the side groups and the central ring. The valence band spectra show a separated group of frontier orbitals, and this appears to be characteristic of these compounds. The pyrrole containing compounds showed much stronger changes of the charge distribution of the MOs of DKP than the other dipeptides, due to hybridization, as demonstrated by the calculations, and consistent with the core level results. We conclude that most side chains (other than pyrrole) interact weakly with the central DKP moiety. Thus a building block approach can be substantially justified, in which the chemical properties can be considered to be the sum of those of the three functional groups making up the cyclic dipeptide.

ACKNOWLEDGMENT

We acknowledge the National Computational Infrastructure (NCI) at the Australian National University for an award under the Merit Allocation Scheme and Swinburne University Supercomputing Facilities, and Victorian Partnership for Advanced Computing (VPAC) for supercomputing facilities. A.P.W.A. acknowledges a Swinburne University Centenary Postgraduate Research Award. We thank our colleagues at Elettra for providing high quality synchrotron light.

[1] A. P. Wickrama Arachchilage, F. Wang, V. Feyer, O. Plekan, and K. C. Prince, J. Chem. Phys. **133**, 174319 (2010).
[2] C. Prasad, Peptides **16**, 151 (1995).
[3] P. Wipf, Chem. Rev. **95**, 2115 (1995).
[4] M. B. Martins and I. Carvalho, Tetrahedron **63**, 9923 (2007); K. H. Rhee, J. Gen. Appl. Microbiol. **48**, 321 (2002).
[5] M. Li, Ph.D. dissertation, Hong Kong University of Science and Technology, 2006; http://repository.ust.hk/dspace/handle/1783.1/2759.
[6] T. F. N. Kwong, L. Miao, X. Li, and P. Y. Qian, Mar. Biotechnol. **8**, 634 (2006).
[7] P. J. Milne, A. L. Hunt, K. Rostoll, J. J. van der Walt, and C. J. Graz, J. Pharm. Pharmacol. **50**, 1331 (1998).
[8] V. Feyer, O. Plekan, N. Tsud, V. Cháb, V. Matolín, and K. C. Prince, Langmuir **26**, 8606 (2010).
[9] G. Tomba, M. Lingenfelder, G. Costantini, K. Kern, F. Klappenberger, J. V. Barth, L. C. Ciacchi, and A. De Vita, J. Phys. Chem. A **111**, 12740 (2007).
[10] M. Lingenfelder, G. Tomba, G. Costantini, L. C. Ciacchi, A. De Vita, and K. Kern, Angew. Chem., Int. Ed. **46**, 4492 (2007).
[11] O. Plekan, V. Feyer, R. Richter, M. Coreno, M. de Simone, K. C. Prince, and V. Carravetta, Chem. Phys. Lett. **442**, 429 (2007).
[12] O. Plekan, V. Feyer, R. Richter, M. Coreno, M. de Simone, K. C. Prince, and V. Carravetta, J. Electron Spectrosc. Relat. Phenom. **155**, 47 (2007).
[13] O. Plekan, V. Feyer, R. Richter, M. Coreno, M. de Simone, K. C. Prince, and V. Carravetta, J. Phys. Chem. A **111**, 10998 (2007).
[14] W. Zhang, V. Carravetta, O. Plekan, V. Feyer, R. Richter, M. Coreno, and K. C. Prince, J. Chem. Phys. **131**, 035103 (2009).
[15] A. Thompson, S. Saha, F. Wang, T. Tsuchimochi, A. Nakata, Y. Imamura, and H. Nakai, Bull. Chem. Soc. Jpn. **82**, 187 (2009).
[16] S. Saha, F. Wang, J. B. MacNaughton, A. Moewes, and D. P. Chong, J. Synchrotron Radiat. **15**, 151 (2008).
[17] F. Wang, Q. Zhu, and E. Ivanova, J. Synchrotron Radiat. **15**, 624 (2008).
[18] F. R. Lucietto, P. J. Milne, G. Kilian, C. L. Frost, and M. van de Venter, Peptides **27**, 2706 (2006).
[19] Y. Zhu, M. Tang, X. Shi, and Y. Zhao, Int. J. Quantum Chem. **107**, 745 (2007).
[20] V. Feyer, O. Plekan, R. Richter, M. Coreno, K. C. Prince, and V. Carravetta, J. Phys. Chem. A **113**, 10726 (2009).
[21] W. D. Allen, E. Czinki, and A. G. Csaszar, Chem.-Eur. J. **10**, 4512 (2004).
[22] K. C. Prince, R. R. Blyth, R. Delaunay, M. Zitnik, J. Krempasky, J. Slezak, R. Camilloni, L. Avaldi, M. Coreno, G. Stefani, C. Furlani, M. de Simone, and S. Stranges, J. Synchrotron Radiat. **5**, 565 (1998).
[23] V. Feyer, O. Plekan, R. Richter, M. Coreno, K. C. Prince, and V. Carravetta, J. Phys. Chem. **112**, 7806 (2008).
[24] O. Plekan, V. Feyer, R. Richter, M. Coreno, M. de Simone, K. C. Prince, A. B. Trofimov, E. V. Gromov, I. L. Zaytseva, and J. Schirmer, Chem. Phys. **347**, 360 (2008).
[25] M. J. Frisch, G. W. Trucks, H. B. Schlegel *et al.*, GAUSSIAN 03, Revision B.04, Gaussian, Inc., Wallingford, CT, 2004.
[26] L. Selvam, V. Vasilyev, and F. Wang, J. Phys. Chem. B **113**, 11496 (2009).
[27] D. P. Chong, E. van Lenthe, S. van Gisbergen, and E. J. Baerends, J. Comput. Chem. **25**, 1030 (2004).
[28] R. van Leeuwen and E. J. Baerends, Phys. Rev. A **49**, 2421 (1994).
[29] E. J. Baerends, J. Autschbach, A. Bérces, C. Bo, P. M. Boerrigter, L. Cavallo, D. P. Chong, L. Deng, R. M. Dickson, D. E. Ellis, M. van Faassen, L. Fan, T. H. Fischer, C. F. Guerra, S. J. A. van Gisbergen, J. A. Groeneveld, O. V. Gritsenko, M. Grüning, F. E. Harris, P. van den Hoek, H. Jacobsen, L. Jensen, G. van Kessel, F. Kootstra, E. van Lenthe, D. McCormack, A. Michalak, V. P. Osinga, S. Patchkovskii, P. H. T. Philipsen, D. Post, C. C. Pye, W. Ravenek, P. Ros, P. R. T. Schipper, G. Schreckenbach, J. G. Snijders, M. Solà, M. Swart, D. Swerhone, G. te Velde, P. Vernooijs, L. Versluis, O. Visser, F. Wang, E. van Wezenbeek, G. Wiesenekker, S. K. Wolff, T. K. Woo, A. Yakovlev, and T. Ziegler, ADF 2006.01, Scientific Computing and Modelling, Vrije Universiteit, Amsterdam, The Netherlands, 2006.
[30] L. S. Cederbaum, J. Phys. B **8**, 290 (1975).
[31] L. S. Cederbaum and W. Domcke, Adv. Chem. Phys. **36**, 205 (1977).
[32] W. von Niessen, J. Schirmer, and L. S. Cederbaum, Comput. Phys. Rep. **1**, 57 (1984).
[33] V. G. Zakrzewski and W. von Niessen, J. Comp. Chem. **14**, 13 (1993).
[34] V. G. Zakrzewski and J. V. Ortiz, Int. J. Quantum Chem. **53**, 583 (1995).
[35] See supplementary material at http://dx.doi.org/10.1063/1.3693763 for a full list of the calculated data.
[36] K. R. Wilson, L. Belau, C. Nicolas, M. Jimenez-Cruz, S. R. Leone, and M. Ahmed, Int. J. Mass Spectrom. **249–250**, 155 (2006).
[37] D. Dehareng and G. Dive, Int. J. Mol. Sci. **5**, 301 (2004).
[38] D. M. Close, J. Phys. Chem. A **115**, 2900 (2011).