Probing structural information of gas-phase peptides by near-edge X-ray absorption mass spectrometry

Simon Dörner,[*,†] Lucas Schwob,[*,†] Kaan Atak,[†] Kaja Schubert,[†] Rebecca Boll,[‡] Thomas Schlathölter,[¶] Martin Timm,[§] Christine Bülow,[§] Vicente Zamudio-Bayer,[§] Bernd von Issendorff,[‖] J. Tobias Lau,[§,‖] Simone Techert,[†,⊥] and Sadia Bari[*,†]

†*Deutsches Elektronen-Synchrotron DESY, Notkestrasse 85, 22607 Hamburg, Germany*

‡*European XFEL GmbH, Holzkoppel 4, 22869 Schenefeld, Germany*

¶*Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 4, 9747AG Groningen, The Netherlands*

§*Abteilung Hochempfindliche Röntgenspektroskopie, Helmholtz-Zentrum Berlin für Materialien und Energie, Albert-Einstein-Strasse 15, 12489 Berlin, Germany*

‖*Physikalisches Institut, Albert-Ludwigs-Universität Freiburg, Hermann-Herder-Strasse 3, 79104 Freiburg, Germany*

⊥*Institut für Röntgenphysik, Universität Göttingen, Friedrich-Hund-Platz 1, 37077 Göttingen, Germany*

E-mail: simon.doerner@desy.de; lucas.schwob@desy.de; sadia.bari@desy.de

S.Dörner, L.Schwob and S.Bari are corresponding authors

S.Dörner and L.Schwob are equal contributors to this manuscript

Abstract

Near-edge X-ray absorption mass spectrometry (NEXAMS) is an action-spectroscopy technique of growing interest for investigations into the spatial and electronic structure of biomolecules. It has been used successfully to give insights into different aspects of the photodissociation of peptides and to probe the conformation of proteins. It is a current question whether the fragmentation pathways are sensitive towards effects of conformational isomerism, tautomerism and intramolecular interactions in gas-phase peptides. To address this issue we studied the cationic fragments of cryogenically cooled gas-phase leucine enkephalin ($[LeuEnk+H]^+$) and methionine enkephalin ($[MetEnk+H]^+$) produced upon soft X-ray photon absorption at the carbon, nitrogen and oxygen K-edges. The interpretation of the experimental ion yield spectra was supported by density-functional theory and restricted-open-shell configuration interaction with singles (DFT/ROCIS) calculations. The analysis revealed several effects that could not be rationalized based on the peptide's amino-acid sequences alone. Clear differences between the partial ion yields measured for both peptides upon C 1s $\rightarrow$ π^*(C=C) excitations in the aromatic amino-acid side chains give evidence for a sulfur-aromatic interaction between the methionine and phenylalanine side chain of $[MetEnk+H]^+$. Furthermore, a peak associated with N 1s $\rightarrow$ π^*(C=N) transitions, linked to a tautomeric keto-to-enol conversion of peptide bonds, was only present in the photon energy resolved ion yield spectra of $[MetEnk+H]^+$.

Introduction

Proteins and peptides are chains of amino-acid residues linked by peptide bonds. They have a variety of biological functions in living organisms, including the catalysis of biochemical reactions, signal transduction and providing structure to cells. Their functions depend on both the amino-acid sequence and the three-dimensional conformation resulting from intramolecular interactions like hydrogen bonds. Investigations of protein and peptide spatial structures are of great importance for research in cell biology, biocatalysis and drug develop-

ment. Since the spatial and electronic structures of a molecule are inextricably linked, it is crucial to obtain detailed information about the electronic structure of proteins and peptides as well. This, however, remains a major experimental challenge.

The past two decades saw major advances in the use of near-edge X-ray absorption fine structure (NEXAFS) spectroscopy for studies on both the spatial and electronic structure of biomolecular systems. In NEXAFS spectroscopy, the photoabsorption cross sections for photoionization and resonant photoexcitation of electrons from the atomic core level to unoccupied molecular orbitals (MOs) are determined by measuring, for example, the Auger-electron yield or the X-ray fluorescence while scanning the photon energy across an atomic absorption edge. The positions and intensities of spectral features in a NEXAFS spectrum allow for the identification of chemical bonds and groups and the characterization of their chemical environments.

NEXAFS spectroscopy in the condensed phase has been used to investigate the electronic and spatial structure of amino acids, peptides and proteins[1–11] as well as nucleobases, nucleotides and strands of deoxyribonucleic acid (DNA).[12–20] NEXAFS spectra for thin films of all standard amino acids are available for the carbon, nitrogen and oxygen K-edges.[3,21] The spectral changes caused by the formation of a peptide bond have been identified and characterized.[1,22] It has been proposed that, in a building block approach, NEXAFS spectra of polypeptides can be approximated from the spectra of their constituents.[2,5,7] Furthermore, calculations for neutral diglycine suggest that NEXAFS is sensitive towards the conformation of the peptide bond.[1]

There is a growing interest in NEXAFS studies on biomolecular systems in the gas phase, in order to investigate their intrinsic properties. Electrospray ionization (ESI) has proven to be a reliable technique to produce charged gas-phase biomolecules. Working with charged sample molecules enables the use of advanced ion manipulation tools, including quadrupole mass filters and ion traps as well as separation techniques based on ion mobility, thus allowing the study of fragile molecules in well-defined states and controlled molecular environment. By

combining gas-phase NEXAFS spectroscopy with ESI tandem mass spectrometry, the mass-to-charge ratios (m/z) and the yields of the photoionization and photodissociation products of biomolecular ions can be measured as a function of the photon energy. To highlight the position of this technique at the overlap between NEXAFS spectroscopy and (biomolecular) mass spectrometry (MS), it is also referred to as near-edge X-ray absorption mass spectrometry (NEXAMS[8,23–25]). The photon energy dependent ion yields can reveal correlations between electronic transitions and structurally informative fragmentation pathways. In this direction, similar to the 2D UV-MS fingerprinting technique,[26,27] NEXAMS could help identifying unknown compound and could, furthermore, be complementary to IR spectroscopy by determining certain aspects of a molecular structure[28,29] and, therefore, constraining the search of the calculated conformational space. Additionally, when the dissociation becomes site-selective,[25] NEXAMS may then be employed to follow ultrafast charge migration and energy transfer in pump-probe schemes at XUV/X-ray Free-Electron Laser facilities.

In the first NEXAMS study on a protein, Milosavljević et al.[9] demonstrated that absorption of a single soft X-ray photon by gas-phase protonated cytochrome *c* predominantly leads to non-dissociative photoionization. In contrast, absorption of soft X-ray photons by gas-phase protonated leucine enkephalin (LeuEnk; Tyr-Gly-Gly-Phe-Leu) at the carbon K-edge induces fragmentation into immonium ions, short amino-acid sequence ions and side-chain fragments.[8] By measuring the photofragmentation mass spectra of different peptides and proteins in the C 1s excitation and C 1s ionization regime, Egorov et al.[30] showed that the correlation between the size of a polypeptide and the degree of fragmentation is based on intramolecular vibrational redistribution (IVR) over the molecule's degrees-of-freedom. The authors also proposed that, in peptides, dissociation driven by IVR is in competition with fast local fragmentation processes involving Auger decay. Furthermore, a study on gas-phase protonated collagen peptides highlighted the involvement of radical-induced dissociation with a low energetic barrier (1-2 eV) in the relaxation processes following photoionization.[31] Recently, we presented a NEXAMS study on gas-phase protonated methio-

nine enkephalin (MetEnk; Tyr-Gly-Gly-Phe-Met), demonstrating site-selective dissociation of a peptide upon site-specific photoexcitation of the methionine sulfur atom at the sulfur L-edge.[25]

By measuring the ionization products of multiply protonated ubiquitin (76 residues) at the carbon K-edge, Milosavljević et al.[32] demonstrated an interplay between the electronic and spatial structure of a protein. While, overall, the ionization threshold increased with the protonation state, over the region of charge states +4 to +8 the unfolding of the protein compensated for the introduced charge, leading to the effective Coulomb field being virtually constant. In contrast, for the smaller, largely α-helical polypeptide melittin (26 residues), even though the conformation changed drastically over the charge states of +2 to +4, the ionization energy still increased.[11] Additionally, the authors observed an increased rate of secondary ionization by inelastic scattering of Auger electrons in the compact conformer.

The current challenge is to find evidence for the sensitivity of NEXAMS towards the conformation of and intramolecular interactions within small peptides, *i.e.* investigate whereas the related absorption bands are accompanied by specific cleavage in peptides. In order to address this question we carried out a comprehensive investigation into the yields of cationic dissociation products of gas-phase protonated LeuEnk and MetEnk ($[\text{LeuEnk+H}]^+$ and $[\text{MetEnk+H}]^+$) measured upon soft X-ray photon absorption at the carbon, nitrogen and oxygen K-edges. The analysis of the experimental data is supported by X-ray absorption spectra calculated using density-functional theory (DFT) and restricted-open-shell configuration interaction with singles (ROCIS).

The only difference between the amino-acid sequences of LeuEnk and MetEnk is the C-terminal residue, which is leucine or methionine, respectively (Figure 1a and 1b). LeuEnk is a well-established standard in ESI-based biomolecular mass spectrometry and has been investigated extensively using non-ionizing activation methods, including collision- (CID[33–35]), surface- (SID[36]) and ultraviolet photon-induced dissociation,[37–39] IR-UV double resonance photofragment spectroscopy[40] and infrared multiphoton dissociation (IRMPD[41]) as well as

ionizing activation methods, involving vacuum-ultraviolet (VUV) photon-,[42,43] soft X-ray photon-,[8] keV ion-[44] and high-intensity femtosecond IR pulse-induced dissociation.[45] The ion yields measured upon soft X-ray absorption by gas-phase $[LeuEnk+H]^+$ at the carbon K-edge[8] are qualitatively in agreement with data reported for gas-phase multiply protonated proteins.[9,10,30] The characteristic spectral features of the peptide backbone and the aromatic residues tyrosine and phenylalanine were observed in the total ion yield spectrum as well as in the partial ion yield spectra of all reported fragments. From the peak intensity ratios, the conclusion was drawn that C 1s $\rightarrow \pi^*$(CONH) excitations in the backbone promote the formation of small immonium ions and side-chain fragments and C 1s $\rightarrow \pi^*$(C=C) excitations in the aromatic side chains contribute more strongly to the formation of larger backbone fragments.

Methods

Near-edge X-ray absorption mass spectrometry

The experiments were performed at the NanoclusterTrap endstation[46,47] at the UE52_PGM beamline at the BESSY II synchrotron (HZB Berlin, Germany). The sample molecules were introduced into the gas phase by electrospray ionization in positive ion mode. The ions were focused into a beam using a radio-frequency ion funnel and guided through a radio-frequency hexapole ion guide to a quadrupole mass filter. For the measurements on $[MetEnk+H]^+$ we combined the NanoclusterTrap with our self-made electrospray ionization source and high-fluence ion funnel.[48] The data sets of $[LeuEnk+H]^+$ were gathered during a different beamtime, where the electrospray ionization source and ion funnel of the NanoclusterTrap were used.

After being filtered according to their mass-to-charge ratio (m/z), the ions passed a 90° bender to prevent neutral molecules from entering the interaction region. Subsequently, the

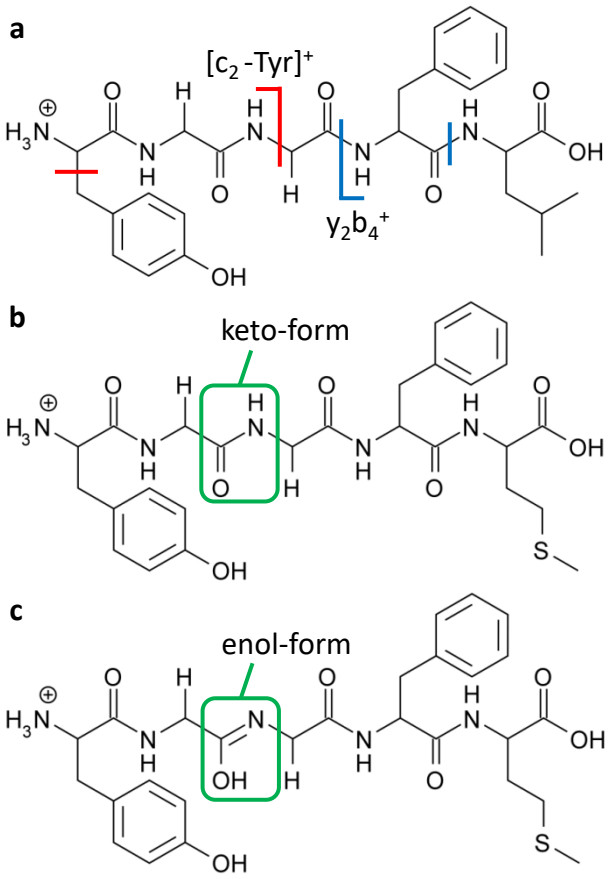

Figure 1: Structural formulas of **(a)** $[\text{LeuEnk}+\text{H}]^+$, **(b)** $[\text{MetEnk}+\text{H}]^+$ and **(c)** $[\text{MetEnk}+\text{H}]^+$ with the Gly-Gly peptide bond in enol-form. All peptides are N-terminal-protonated. The red lines and blue lines indicate the cleavage sites of the $[c_2\text{ -Tyr}]^+$ and $y_2b_4^+$ fragment respectively. The green boxes indicate the enol and keto-form of the Gly-Gly peptide bond of $[\text{MetEnk}+\text{H}]^+$.

ions were accumulated and stored in a cryogenic linear ion trap while being cooled to ∼20 K with helium buffer gas. After accumulation, the ions were irradiated with the soft X-ray photons for 100 ms. Afterwards, trapped cationic products were analyzed using a reflectron time-of-flight mass spectrometer. The photon flux was measured by a photodiode behind the interaction region and was in the order of 10^{12} s^{-1}. The design of the NanoclusterTrap instrument constraints the detection of the photoproducts only on a small m/z window. In order to detect, for both peptides, most of the immonium ions and side-chain fragments, which have reported to be the main photoproducts observed for $[LeuEnk+H]^+$ at the carbon K-edge,[8] the trapping parameters were adjusted for a mass range of about m/z 50 - 180. Consequently, expected backbone fragments with m/z-ratios outside of the detection window (b_3, b_2, GF) were not detected.

Photon energy scans on $[LeuEnk+H]^+$ and $[MetEnk+H]^+$ were performed, scanning the carbon (280 - 320 eV), nitrogen (395 - 420 eV) and oxygen K-edges (530 - 550 eV) in steps of 125 meV. The monochromator bandwidth was set to 150 meV at the carbon K-edge and to 250 meV at the nitrogen and oxygen K-edges. To remove contributions by fragments arising from collisions with the buffer gas, for each scan a background spectrum without photons was recorded and subtracted from each individual mass spectrum. The mass resolution $m/\Delta m$ was 590 at m/z 77, 214 at m/z 107, 317 at m/z 120 and 800 at m/z 147. Each mass spectrum was normalized to the photon flux. The peptide samples consisted of 30 μM of peptide in 1:1 water/methanol with 1 vol% of formic acid. Leucine-enkephalin acetate salt hydrate (purity $\geq$ 95%), [Met^5]-enkephalin acetate salt hydrate (purity $\geq$ 95%) as well as methanol and formic acid were purchased from Merck/Sigma-Aldrich and were used without additional purification.

Data treatment and fitting

Photon-energy-dependent ion-yield spectra, in the following referred to as partial NEXAMS spectra, were calculated from the photon-flux-normalized mass spectra measured at the different photon energies by summarizing the data points within a window of Δm/z = $\pm$0.4 around the center of a mass peak. The total NEXAMS spectrum of a scan, calculated by summing up the partial NEXAMS spectra of all detected fragments (m/z 57 - 147), was used as an approximation for the total ion yield.

The experimental broadening of the spectroscopic peak can be attributed to several factors. Contributions to the broadening from the 1s-core-hole lifetime have values for carbon, nitrogen and oxygen which are expected to be in the order of 0.1 eV,[49] 0.1 eV[50] and 0.15 eV,[51] respectively. The photon energy bandwidth, defined by the opening of the exit slit of the beamline, accounts for additional broadening of 0.15 eV at the carbon K-edge and 0.25 eV at the nitrogen and oxygen K-edges. Other main sources of broadening are intrinsic to the size of the molecule and to X-ray spectroscopy. The former leads to the overlap of electronic excitations due to multiple virtually equivalent photon absorbing sites and the latter to vibrational excitations in the core-hole excited state (Franck-Condon principle). Finally, the contribution of the vibrational excitations in the initial state of the peptides ions is limited by the cooling of the ions in the cryogenic trap.

In order to identify and separate overlapping spectral features, the total NEXAMS spectra were fitted using Gaussian functions. The number of functions as well as the starting points and expected boundaries for the fit parameters were chosen based on reference data from literature and on the DFT/ROCIS-calculated X-ray absorption spectra (see the 'Total NEXAMS spectra' section). In addition, a sigmoid function was applied to fit the peptide's respective core-ionization edge. The core-ionization edge of an element within a molecule is the superposition of the ionization edges of all corresponding atoms, influenced by their chemical environments. Therefore, instead of the sigmoid function's point of inflection E_{infl} the x-intercept of the tangent of E_{infl} was considered the peptide's ionization potential E_{IP}.

The area A_{peak} under a Gaussian function was calculated directly from the fitting parameters. Thus, A_{peak} is the ion yield attributed to an ensemble of electronic transitions integrated over the whole spectrum. In contrast, the area of an ionization edge A_{edge} was calculated by integrating the respective sigmoid fit function between zero and $E_{IP} + 2\Delta E$ with $\Delta E = E_{infl} - E_{IP}$. The sum of the areas of all Gaussian functions plus the area of the ionization edge was considered as the total area under a spectrum A_{spec}. The partial NEXAMS spectra of individual fragments were fitted in the same way, with the exception that the peak positions and peak widths were taken from the fits of the total NEXAMS spectra.

The trapping efficiency of the NanoclusterTrap is not constant across the trapping window. Therefore, all NEXAMS spectra were normalized. The total NEXAMS spectra were normalized to the maximum intensity of the ionization edge and the partial NEXAMS spectra of individual fragments were normalized to A_{spec}.

DFT/ROCIS calculations

All calculations were carried out with the ORCA program package.[52] The molecular geometries were optimized by DFT using the Ahlrichs TZVP basis set[53] in combination with the B3LYP exchange-correlation functional.[54,55] The optimized geometries are shown in the supporting information (SI:Figures 1/7). DFT/ROCIS calculations using parameters $c_1 = 0.18$, $c_2 = 0.20$ and $c_3 = 0.40$,[56] the TZVP basis set and the B3LYP functional were performed to obtain the transition energies and dipole moments for the carbon, nitrogen and oxygen K-edges. The resolution-of-identity approximation[57–61] in combination with an auxiliary basis set generated by the AutoAux generation scheme[62] was used during all calculations. Furthermore, all DFT calculations were performed on a dense integration grid (ORCA Grid4). In all calculations relativistic effects were taken into account using zeroth-order regular approximation (ZORA).[63] Vibronic effects were not taken into account.

In order to determine the energy shifts of the calculated transition energies, inherent to the DFT/ROCIS calculation, theoretical X-ray absorption spectra were calculated and matched with the experimental total NEXAMS spectra. The theoretical absorption spectra were produced by broadening the calculated transition lines using a pseudo-Voigt profile, approximated by linear combination of a Gaussian (FWHM = photon energy bandwidth + 0.4 eV) and Lorentzian term (FWHM = 1s-core-hole-lifetime broadening).[64–66] The coefficient for the linear combination and the FWHM of the Voigt profile f_V were calculated according to the work of Thompson et al.[67] The resulting f_V values are ∼0.6 eV, ∼0.7 eV and ∼0.75 eV for the carbon, nitrogen and oxygen K-edge measurement.

Results and discussion

Fragment Identification

At all absorption edges, qualitatively, the same fragments were detected. Thus, for the identification of the detected fragments the mass spectra measured at 288 eV (Figure 2) are representative for all scans. Compared to $[MetEnk+H]^+$ the trapping window for $[LeuEnk+H]^+$ was shifted about m/z 10 to lower masses. The mass spectra are in line with the carbon K-edge data reported by González-Magaña et al.[8] and largely in agreement with spectra produced by keV ion-induced dissociation[44] and VUV photon-induced dissociation at photon energies higher than the ionization threshold.[43]

The most intense peaks at m/z 91, 107, 120 and 136 are common for molecules containing phenylalanine (Phe) and tyrosine (Tyr) residues. They are attributed to the Phe and Tyr immonium ions (F_{im}: 120 u, Y_{im}: 136 u) and the respective side-chain fragments (F_{sc}: 91 u, Y_{sc}: 107 u).[68,69] The less intense peaks at m/z 57 and 86 measured for $[LeuEnk+H]^+$ stem from the leucine (Leu) side chain (L_{sc}: 57 u) and the leucine immonium ion (L_{im}: 86 u). For $[MetEnk+H]^+$, instead, peaks matching the methionine (Met) side chain (M_{sc}: 75 u) and

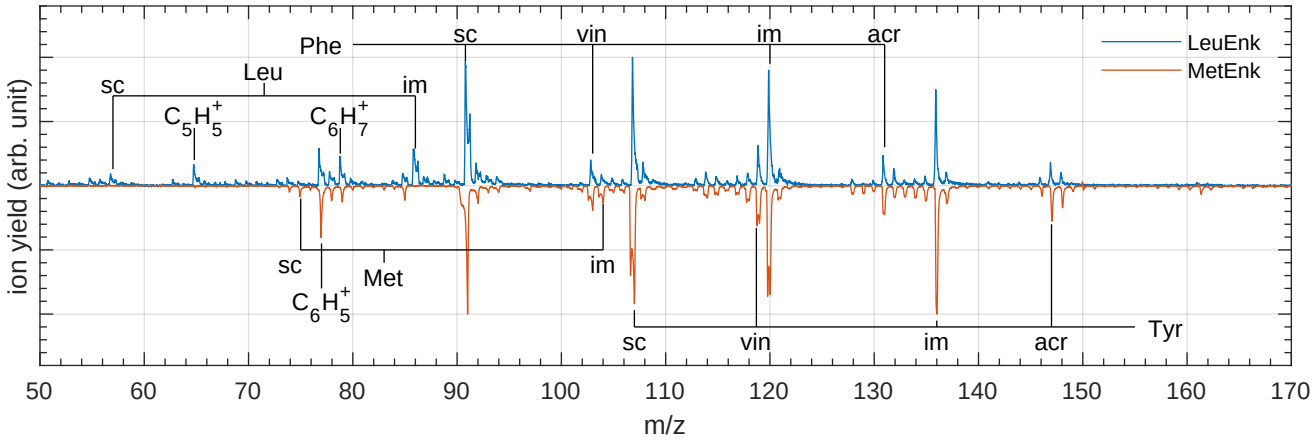

Figure 2: Photodissociation mass spectra of gas-phase $[LeuEnk+H]^+$ (blue line) and $[MetEnk+H]^+$ (red line; inverted) measured at 288 eV. sc: side-chain fragment, vin: vinyl cation, im: immonium ion, acr: acryloyl cation.

methionine immonium ion (M_{im}: 104 u) are present.

The peaks at m/z 131 and 147 have previously been assigned to the $[c_2-Y_{sc}]^+$[8,45] and $y_2b_4^+$ fragment[38,39] of $[LeuEnk+H]^+$ (see Figure 1a). The fragments are labeled according to the nomenclature established by Roepstorff and Fohlmann.[70] Alternatively, these peaks might be attributed to acryloyl cations of Phe and Tyr (F_{acr}: 131 u, Y_{acr}: 147 u). The peaks at m/z 103 and 119 indicate vinyl cations of Phe and Tyr (F_{vin}: 103 u, Y_{vin}: 119 u). The vinyl and acryloyl cations of Phe have been reported for low energy CID of Ag^+-phenylalanine complexes.[71] Peaks at m/z 119 and 147 have been observed upon UV-induced dissociation of cold protonated Tyr.[72] The low energy CID spectrum of protonated 3-iodotyrosine[73] shows corresponding peaks at m/z 245 (119 + 126) and m/z 273 (147 + 126).

The peaks at m/z 65, 77 and 79 can be assigned to different cyclic fragments of the aromatic side chains. From organic mass spectrometry it is known that a toluene cation (F_{sc}) can rearrange into a tropylium cation[74,75] which can then undergo a loss of C_2H_2, leading to the formation of the cyclopentadienyl cation ($C_5H_5^+$: 65 u).[76,77] This cation also appears in the electron ionization dissociation (EID) spectrum of Phe and toluene.[78] The Y_{sc} fragment can rearrange into a hydroxy-tropylium cation and a subsequent loss of CO can lead to the arenium ion ($C_6H_7^+$: 79 u).[79,80] Fragments at m/z 79 have been observed upon UV/VUV-

induced dissociation of protonated Tyr[81] and EID of p-cresol.[78] A peak at m/z 77 is a strong indication for the phenyl cation ($C_6H_5^+$: 77 u), which has been measured as a fragment of both Tyr and Phe upon EID[78] and photon-induced dissociation in the UV/VUV range.[81] By EID of 1-phenylethanol it has been shown, that a loss of H_2 from the arenium ion can lead to the phenyl cation.[82] However, direct bond cleavage and decomposition of larger Phe and Tyr related fragments are plausible origins for this cation as well. The remaining weak peaks across the mass spectra are due to small internal and terminal fragments. A list of the discussed fragments as well as their chemical structures can be found in the supporting information (SI:Table 9 and SI:Figure 16).

Total NEXAMS spectra

Carbon K-edge

The total NEXAMS spectra of $[LeuEnk+H]^+$ (Figure 3a) and $[MetEnk+H]^+$ (Figure 3c) across the carbon K-edge appear very similar at first. The spectra show two intense narrow main features at ∼285 eV and ∼288 eV, a small feature at ∼287 eV, creating a shoulder on the left flank of the second main peak, and two broad features centered at ∼291.5 eV and ∼294 eV. The spectrum of $[LeuEnk+H]^+$ was fitted with three intense narrow peaks CK1a, CK1b and CK3 located at 285.0 eV, 285.4 eV and 288.1 eV, two weak narrow peaks CK2 and CK4 at 287.2 eV and 288.6 eV and two broad features CK5 and CK6 centered at 290.1 eV and 291.8 eV. The respective peak positions determined for $[MetEnk+H]^+$ are CK1a: 285.1 eV, CK1b: 285.4 eV, CK2: 287.3 eV, CK3: 288.2 eV, CK4: 288.7 eV, CK5: 290.1 eV and CK6: 291.8 eV. The assignments of these peaks, listed in Table 1, are discussed in the following.

Peaks CK1a and CK1b are both attributed to C 1s → π*(C=C) transitions in the aromatic side chains of the Tyr and Phe residues.[21,22,86] In the DFT/ROCIS-calculated data of both peptides the energies for transitions in the Tyr side chain were slightly higher

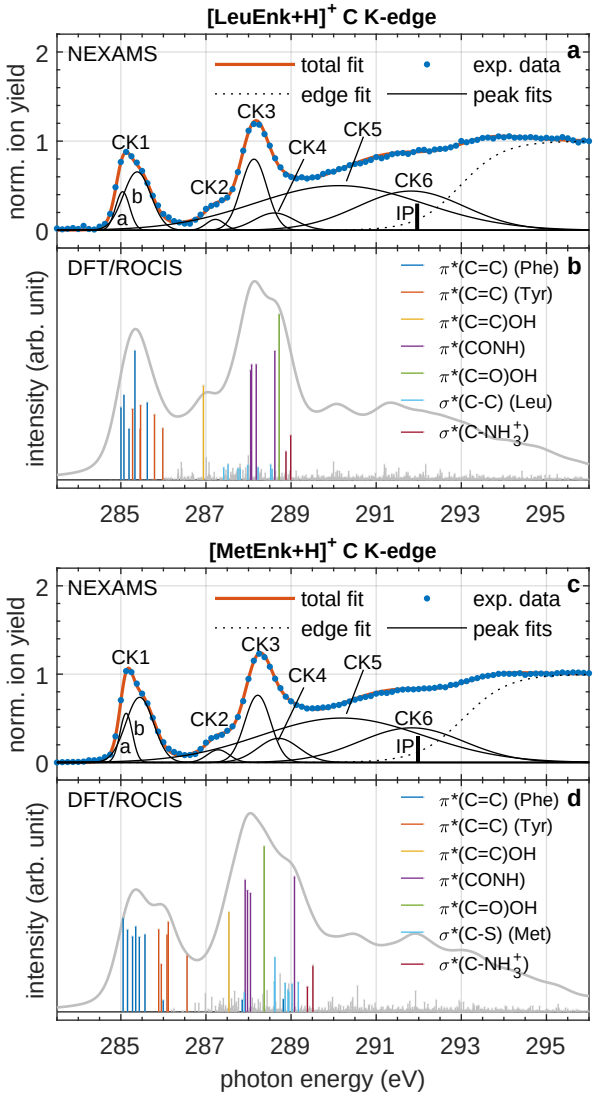

Figure 3: Experimental carbon K-edge ion yield spectra of **(a)** $[\text{LeuEnk+H}]^+$ and **(c)** $[\text{MetEnk+H}]^+$ as well as DFT/ROCIS-calculated carbon K-edge X-ray absorption lines of **(b)** $[\text{LeuEnk+H}]^+$ (shifted +11.3 eV) and **(d)** $[\text{MetEnk+H}]^+$ (shifted +11.7 eV). The experimental spectra are normalized to the maximum intensity of the C 1s ionization edge. The calculated discrete absorption lines were broadened using a pseudo-Voigt profile with 0.6 eV FWHM. IP: ionization potential

Table 1: Experimental energy positions (exp) and proposed assignments of the spectral features observed in the total carbon K-edge NEXAMS spectra of $[LeuEnk+H]^+$ and $[MetEnk+H]^+$ as well as energies of the corresponding prominent DFT/ROCIS-calculated absorption lines (calc).

peak	position [eV]				assignment	references
	$[LeuEnk+H]^+$		$[MetEnk+H]^+$			
	exp	calc	exp	calc		
CK1a	285.0	285.0 – 285.6	285.1	285.1 – 285.6	$\rightarrow \pi^*$(C=C)[a]	[83,84]
CK1b	285.4	285.3 – 286.0	285.4	285.9 – 286.1, 286.6	$\rightarrow \pi^*$(C=C)[b]	[84]
CK2	287.2	286.9	287.3	287.5	$\rightarrow \pi^*$(C=C)OH	[21,22,85,86]
CK3	288.1	288.0 – 288.6	288.2	287.9 – 288.0, 289.1	$\rightarrow \pi^*$(CONH)	[1,4,22]
CK4	288.6	288.7	288.7	288.4	$\rightarrow \pi^*$(C=O)OH	[4,21]
		288.9, 289.0		289.4, 289.5	$\rightarrow \sigma^*$(C−NH_3^+)	[21]
		287.4 – 288.6			$\rightarrow \sigma^*$(C−C) Leu	
				287.9, 288.6, 288.9 – 289.2	$\rightarrow \sigma^*$(C−S) Met	
CK5	290.1		290.1		($\rightarrow \sigma^*$)	
CK6	291.8		291.8		($\rightarrow \sigma^*$)	
IP	292.0		292.1		C 1s IP	

[a] mainly Phe side chain; [b] mainly Tyr side chain
IP: ionization potential

than for the respective transitions in the Phe side chain. This suggests an assignment of CK1a to transitions in the Phe side chain and of CK1b to transitions in the Tyr side chain. However, across all fragments no significant differences between the ratios of these peaks to each other were observed. Therefore, in the following both peaks are combined into one feature CK1 with its maximum at 285.1 eV. Features at similar positions as CK1 have also been reported for gas-phase benzene[83] as well as thin films of benzene and phenol.[84] Feature CK2 is attributed to C 1s $\rightarrow \pi^*$(C=C)OH transitions in the Tyr side chain only.[21,22,85,86] The second main feature CK3 originates from C 1s $\rightarrow \pi^*$(CONH) transitions in the peptide backbone.[1,4,22]

The energy position of peak CK4 is in agreement with C 1s $\rightarrow \sigma^*$(C−NH_3^+) transitions reported for thin film data of Tyr (288.8 eV)[21] as well as C 1s $\rightarrow \pi^*$(C=O)OH transitions in the carboxylic acid groups of Leu,[21] Met[21] and isoleucine[4] (all 288.6 eV). The DFT/ROCIS-calculated data show transitions into MOs with high π^*(C=O)OH contribution at ∼288.7 eV for $[LeuEnk+H]^+$ (LUMO+11 SI:Figure 4) and at ∼288.4 eV for $[MetEnk+H]^+$ (LUMO+8 SI:Figure 10, LUMO+9; SI:Figure 11). First significant transitions into σ^* orbitals of the C−NH_3^+ group appear at 288.9 eV for $[LeuEnk+H]^+$ (LUMO SI:Figure 2, LUMO+1

SI:Figure 3) and at 289.4 eV for $[MetEnk+H]^+$ (LUMO SI:Figure 8, LUMO+1 SI:Figure 9) in the calculated data. Since the N-terminal regions of the DFT-optimized geometries of both peptides are fairly similar, the shift between these energies is a theoretical evidence for a high sensitivity of the energy of the $\sigma^*(C-NH_3^+)$ orbital towards the peptide conformation. Noteworthy electronic transitions into MOs of the Met side chain (mainly LUMO+20, LUMO+23, LUMO+27 and LUMO+28 SI:Figures 12-15) and MOs of the Leu side chain (mainly LUMO+23 and LUMO+26 SI:Figures 5/6) also appear in the energy region of peak CK4 at 288.0 - 289.2 eV and 287.4 - 288.6 eV, respectively. However, for both peptides the intensities calculated for C 1s $\rightarrow \pi^*$(C=O)OH transitions are one order of magnitude higher than for the transitions into MOs of the Met or Leu side chains. In conclusion, because of a better agreement with our calculations and peak positions reported in literature we consider the assignment of peak CK4 to C 1s $\rightarrow \pi^*$(C=O)OH transitions as more plausible. Features CK5 and CK6 cannot be assigned to single transitions and represent the accumulation of various transitions into σ^* orbitals and potentially Rydberg states and shape resonances.

The area of peak CK1 normalized to the maximum intensity of the C 1s ionization edge is $\sim$0.71 for $[LeuEnk+H]^+$ and $\sim$0.80 for $[MetEnk+H]^+$. This means, overall, the ion yield upon excitations in the aromatic side chains is $\sim$14.3% higher for $[MetEnk+H]^+$ than for $[LeuEnk+H]^+$. Experimental data reported for Leu and Met do not show transitions in the energy range of CK1.[3,21] The DFT/ROCIS-calculated carbon K-edge absorption spectra of both peptides show only transitions related to the aromatic side chains in this energy regime. Thus, we propose that there is an intramolecular interaction in play, promoting fragmentation upon photoexcitation in the aromatic side chains. This idea is discussed in more detail in the 'Evidence for Sulfur-Aromatic Interactions' section based on the partial ion yields of different fragments.

Nitrogen K-edge

The total NEXAMS spectra of $[LeuEnk+H]^+$ and $[MetEnk+H]^+$ measured at the nitrogen K-edge are shown in Figure 4a and Figure 4c. For both peptides there are two narrow peaks at ∼401.4 eV and ∼402.8 eV and a broad but distinct peak at ∼406 eV. $[MetEnk+H]^+$ additionally exhibits one weak, isolated feature at 399.2 eV. From about 408 to 411 eV the intensities in both spectra increase only slightly because the falling slope of the 406 eV feature is compensated by the rising slope of the N 1s ionization edge. The total NEXAMS spectrum of $[LeuEnk+H]^+$ was fitted with three peaks NK1, NK2 and NK3 positioned at 401.4 eV, 402.6 eV and 406.0 eV. The positions of peaks NK1 and NK2 were identical for $[MetEnk+H]^+$ and $[LeuEnk+H]^+$. Peak NK3 of $[MetEnk+H]^+$ was shifted by 0.4 eV to 406.4 eV. The isolated feature at 399.2 eV in the spectrum of $[MetEnk+H]^+$ was fitted with one Gaussian peak NK4 centered at the same energy position. During the fitting procedure it became apparent that in the case of $[MetEnk+H]^+$ there is a shoulder on the low-energy slope of the ∼406 eV feature. Therefore, another peak NK5 at 404.4 eV was added. In the following the assignments of all peaks are discussed. The assignments are listed in Table 2.

Table 2: Experimental energy positions (exp) and proposed assignments of the spectral features observed in the total nitrogen K-edge NEXAMS spectra of $[LeuEnk+H]^+$ and $[MetEnk+H]^+$ as well as energies of the corresponding prominent DFT/ROCIS-calculated absorption lines (calc).

peak	position [eV]				assignment	references
	$[LeuEnk+H]^+$		$[MetEnk+H]^+$			
	exp	calc	exp	calc		
NK1	401.4	400.6 − 401.7	401.4	401.2 − 401.6, 402.7	$\rightarrow \pi^*$(CONH)	[1,3,4]
NK2	402.6		402.6		$\rightarrow \pi^*$(N−C_α)	[6,87,88]
NK3	406.0		406.4		$(\rightarrow \sigma^*)$	[1,4]
NK4	-		399.2		$\rightarrow \pi^*$(C=N)	[89]
					$NH_2 \rightarrow \sigma^*$(N−H)	[90]
NK5	-	403.1	404.4	403.7, 403.9	$NH_3^+ \rightarrow \sigma^*$(N−H)	[1,4]
IP	407.4		407.6		N 1s IP	

IP: ionization potential

Peak NK1 is attributed to N 1s $\rightarrow \pi^*$(CONH) transitions in peptide bonds.[1,3,4] It is important to note that because of the delocalization of the π-electrons in a peptide bond (O=C−NH $\rightleftharpoons$ O^-−C=N^+H), both the C=O and C−N bond have partial π-bond character.

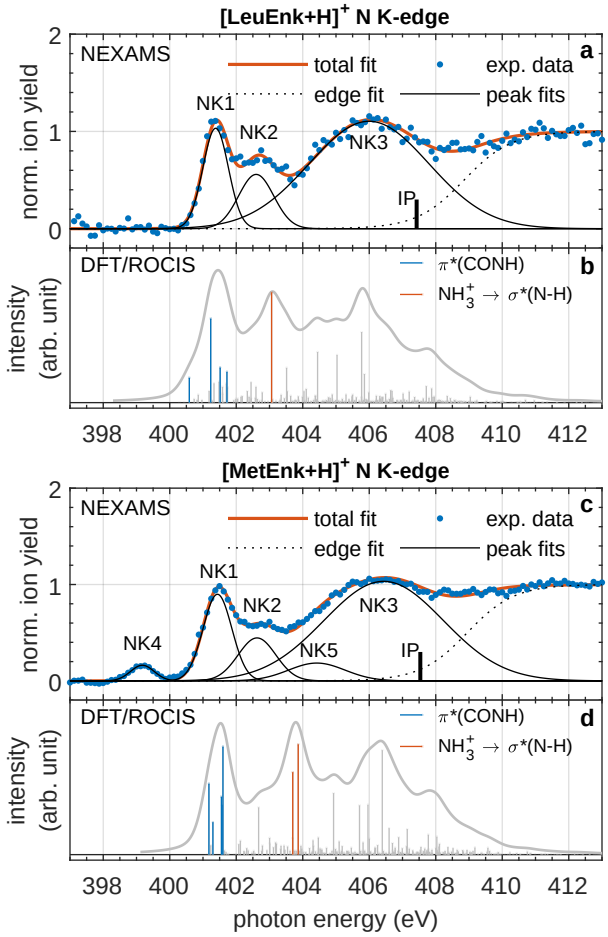

Figure 4: Experimental nitrogen K-edge ion yield spectra of **(a)** $[\text{LeuEnk+H}]^+$ and **(c)** $[\text{MetEnk+H}]^+$ as well as DFT/ROCIS-calculated carbon K-edge X-ray absorption lines of **(b)** $[\text{LeuEnk+H}]^+$ (shifted +12.2 eV) and **(d)** $[\text{MetEnk+H}]^+$ (shifted +12.6 eV). The experimental spectra are normalized to the maximum intensity of the N 1s ionization edge. The calculated discrete absorption lines were broadened using a pseudo-Voigt profile with 0.7 eV FWHM. IP: ionization potential

The assignment of peak NK2 is controversial. Features at this position have been assigned to N 1s $\rightarrow \pi^*(\mathrm{N{-}C_\alpha})$ transitions in gas-phase alanine (402.4 eV[87]) and diglycine (402.44 eV[6]) as well as N 1s $\rightarrow \pi^*(-\mathrm{NH{=}})$ transitions in thin films of leucoemeraldine ($[\mathrm{C_6H_4NH}]_\mathrm{n}$; 402.1 eV).[88] For different homopolypeptides Zubavichus et al.[4] suggested N 1s $\rightarrow \pi^*$(CONH) transitions (402.7 eV) that are energetically shifted because of an α-helical peptide conformation. Another common assignment for peaks around 402.5 eV are transitions into 3p-Rydberg orbitals.[1] The DFT/ROCIS calculations for $[\mathrm{MetEnk{+}H}]^+$ (Figure 4d) show a small but distinct feature at 402.7 eV. The probed MOs are, however, highly delocalized. Ultimately, we rated the assignment of NK2 to N 1s $\rightarrow \pi^*(\mathrm{N{-}C_\alpha})$ transitions as most plausible.

The spectral feature NK3 involves mainly N 1s $\rightarrow \sigma^*(\mathrm{N{-}C_\alpha})$ and N 1s $\rightarrow \sigma^*$(N−H) transitions.[1,4] In the energy regime of NK3 the DFT/ROCIS calculations for both peptides (Figure 4b and Figure 4d) show prominent absorption lines attributed to transitions from the N-terminal nitrogen atom into MOs delocalized over the $\mathrm{NH_3}^+$ group and the tyrosine residue. Most noteworthy are intense absorption lines at 405.8 eV for $[\mathrm{LeuEnk{+}H}]^+$ and 406.4 eV of $[\mathrm{MetEnk{+}H}]^+$, leading to distinct peaks at about the same photon energies. In the experimental spectra of both peptides such peaks were not observed. Furthermore, the calculated absorption spectrum of $[\mathrm{LeuEnk{+}H}]^+$ features a peak at 403.1 eV stemming mainly from a single absorption line attributed to N 1s $\rightarrow \sigma^*$(N−H) transitions in the $\mathrm{NH_3}^+$. For $[\mathrm{MetEnk{+}H}]^+$ this peak is shifted to 403.8 eV. Obvious corresponding peaks were not observed in the experimental spectra. However, peak NK5 in the ion yield spectrum of $[\mathrm{MetEnk{+}H}]^+$ might be connected to these transitions.

For peak NK4 we consider N 1s $\rightarrow \pi^*$(C=N) transitions in peptide bonds in enol-form and N 1s $\rightarrow \sigma^*$(N−H) transitions in unprotonated amino groups as plausible reasons. The assignment is discussed in detail in the 'Discussion of the 399.2 eV feature of $[\mathrm{MetEnk{+}H}]^+$' section.

Oxygen K-edge

At the oxygen K-edge, qualitatively, the total NEXAMS spectra of both peptides (Figure 5a/c) exhibit the same features: a dominant feature at ∼532 eV and three weaker features centered at ∼535eV, ∼538 eV and ∼543 eV. For $[LeuEnk+H]^+$ the fitting procedure led to one intense peak OK1 at 532.1 eV and two less intense peaks OK2 and OK3 located at 534.4 eV and 539.0 eV. The respective peak positions determined for $[MetEnk+H]^+$ are OK1: 532.2 eV, OK2: 535.2 eV and OK3: 538.9 eV. The assigments of these peaks are discussed in the following and are listed in Table 3.

Table 3: Experimental energy positions (exp) and proposed assignments of the spectral features observed in the total oxygen K-edge NEXAMS spectra of $[LeuEnk+H]^+$ and $[MetEnk+H]^+$ as well as energies of the corresponding prominent DFT/ROCIS-calculated absorption lines (calc).

peak	position [eV]				assignment	references
	$[LeuEnk+H]^+$		$[MetEnk+H]^+$			
	exp	calc	exp	calc		
OK1	532.1	531.9 – 532.1, 532.8	532.2	531.9 – 532.1, 533.3	CO → π*(CONH)	[1,6,91]
		532.5		532.4	CO → π*(C=O)OH	[1,6]
OK2	534.4	534.7 – 534.9	535.2	534.8, 535.0	OH → π*(C=O)OH	[1,6,92]
		534.3		535.0 – 535.2	OH → π*(C=C)OH	[93,94]
OK3	539.0		538.9		(→ σ*)	
IP	539.9		539.9		O 1s IP	

IP: ionization potential

Peak OK1 can be assigned to transitions of carbonyl oxygen 1s electrons into π*(CONH) and π*(C=O)OH orbitals. Transitions of these types have been reported for gas-phase diglycine (π*(CONH): 532.2 eV;[1,6] π*(C=O)OH: 531.7 eV,[1] 531.85 eV[6]). Furthermore, O 1s → π*(CONH) transitions have been measured for N-methylacetamide (∼532 eV[95,96]) and O 1s → π*(C=O)OH transitions for formic acid (532.17 eV[92]). A peak in the energy region of OK1 has also been observed for gas-phase substance P (532.15 eV[91]), a peptide composed of eleven amino-acid residues. It should be noted that the calculated absorption line attributed to O 1s → π*(CONH) transitions in the Tyr-Gly peptide bond of $[MetEnk+H]^+$ is shifted by 0.5 eV to higher energies compared to the energy of the respective transition in $[LeuEnk+H]^+$, which leads to a peak at 533.3 eV in the calculated absorption spectrum of

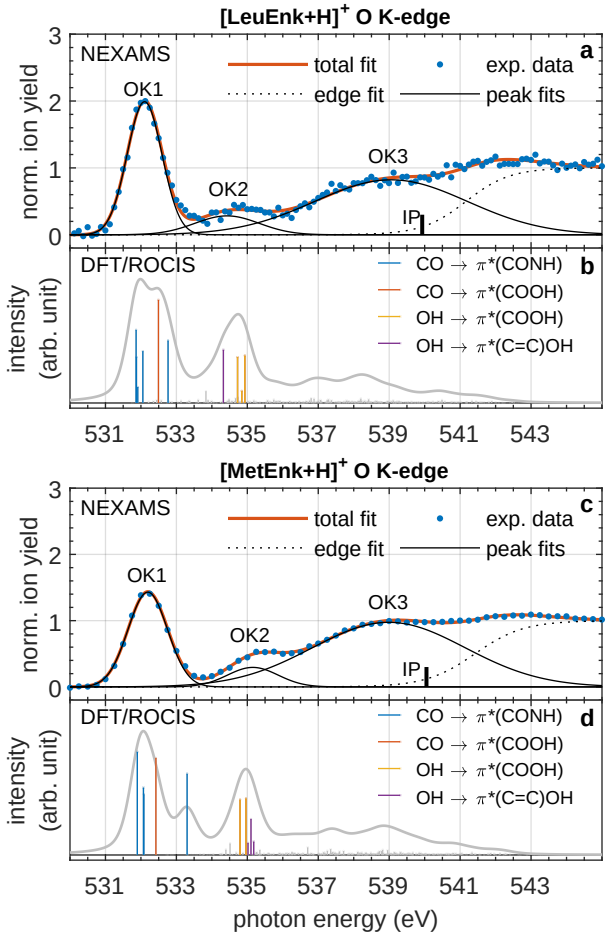

Figure 5: Experimental oxygen K-edge ion yield spectra of **(a)** $[\text{LeuEnk+H}]^+$ and **(c)** $[\text{MetEnk+H}]^+$ as well as DFT/ROCIS-calculated carbon K-edge X-ray absorption lines of **(b)** $[\text{LeuEnk+H}]^+$ (shifted +14.1 eV) and **(d)** $[\text{MetEnk+H}]^+$ (shifted +14.6 eV). The experimental spectra are normalized to the maximum intensity of the O 1s ionization edge. The calculated discrete absorption lines were broadened using a pseudo-Voigt profile with 0.75 eV FWHM. IP: ionization potential

$[MetEnk+H]^+$. This effect was, however, not observed in the experimental spectra.

The position of peak OK2 in the total NEXAMS spectrum of $[LeuEnk+H]^+$ (534.4 eV) is shifted 0.8 eV towards lower photon energies in comparison to $[MetEnk+H]^+$ (535.2 eV). The DFT/ROCIS-calculated oxygen K-edge absorption spectrum of $[LeuEnk+H]^+$ shows a feature at ∼534.7 eV which is caused mainly by excitations of hydroxy oxygen 1s electrons into adjacent π*-orbitals of the Tyr side chain and the carboxylic acid group. For $[MetEnk+H]^+$ this feature appears at ∼535 eV. An assignment to O 1s → π*(C=O)OH transitions is in line with experimental data of gas-phase formic acid (535.37 eV[92]), gas-phase diglycine (535.3 eV,[1] 535.4 eV[6]) and gas-phase Met (535.4 eV).[97] O 1s → π*(C=C)OH transitions have been reported for gas-phase phenol (534.9 eV[93]) and hydroquinone (534.6 eV[93]) as well as tyrosine-terminated propanethiol assembled on a gold surface (535.5 eV[94]).

The strong shift of peak OK2 in the experimental data might be due to hydrogen bonding. For water[98–100] and different alcohols[101] it has been demonstrated that there can be significant differences, especially in the 535 eV region, between the oxygen K-edge absorption spectra of an H-donating and H-accepting molecule. The 535 eV region of the ion yield spectrum of $[MetEnk+H]^+$ could furthermore be influenced by O 1s → σ*(C−OH) transitions in peptide bonds in enol-form. The broad feature OK3 represents the entirety of excitations into σ* orbitals and, potentially, excitations leading to Rydberg states and shape resonances.

The areas of peaks OK1 and OK2 normalized to the maximum intensity of the O 1s ionization edge are 2.45 and 0.64 for $[LeuEnk+H]^+$ and 2.00 and 0.53 for $[MetEnk+H]^+$. This means the total measured ion yield upon resonant O 1s → π*(CONH) and O 1s → π*(C=O)OH excitations is ∼18% lower for $[MetEnk+H]^+$ than for $[LeuEnk+H]^+$. With ∼17% this relative difference is very similar for peak OK2. For the immonium ions and side-chain fragments of Tyr and Phe, which govern the biggest part of the total measured ion yield, the area of peak OK1 is 12 to 38% smaller for $[MetEnk+H]^+$ than for $[LeuEnk+H]^+$ (see SI:Tables 4/8). This shows that the difference between the total NEXAMS spectra is not caused by the decreased yield of a single fragment. These findings are surprising because

similar effects were not observed for the transitions into π*(CONH) orbitals at the carbon and nitrogen absorption edges. Also the DTF/ROCIS-calculated oxygen K-edge absorption spectra (Figure 5b/d) do not reflect this difference. The reduced ion yield for $[MetEnk+H]^+$ upon O 1s → π*(C=O) excitation either means that overall the rate of these excitations is reduced or that the rates in both peptides are on a similar level but in $[MetEnk+H]^+$ there are competing dissociation channels leading to fragments with m/z ratios that are not within the trapping window. The reason for this effect could be connected to C=O···H−N hydrogen bonds between different peptide bonds. More information about the gas-phase structure of protonated MetEnk is needed for a conclusive interpretation of these data.

Partial NEXAMS spectra measured for $[MetEnk+H]^+$

At each absorption edge, qualitatively, all partial NEXAMS spectra exhibit the same spectral features as the respective total NEXAMS spectrum (see the 'Total NEXAMS spectra' section). The intensities of the features, however, show clear differences. The normalized yields of all measured immonium ions and complete side-chain fragments as well as the arenium ion ($C_6H_7^+$) and phenyl cation ($C_6H_5^+$) of $[MetEnk+H]^+$ are shown in Figure 6. The non-normalized peak areas and intensities determined for these fragments by fitting are listed in the supporting information (SI:Tables 2-4,6-8).

At all absorption edges the ratio between the maximum intensities of the π*(CONH) feature and the ionization edge is smaller for the complete side-chain fragments than for the immonium ions. For the side-chain daughter-fragments this ratio is even smaller. This effect is in general agreement with observations made at the carbon K-edge of $[LeuEnk+H]^+$.[8] Overall, the normalized yields of F_{im} and M_{im} only show small differences (Figures 6a,d,g). In comparison, at all absorption edges, the normalized yields of Y_{im} measured upon excitations into π*(CONH) orbitals are increased. Immonium ions are protonated fragments. Thus, migration of a proton to the fragmentation site is necessary in order to produce these ions.

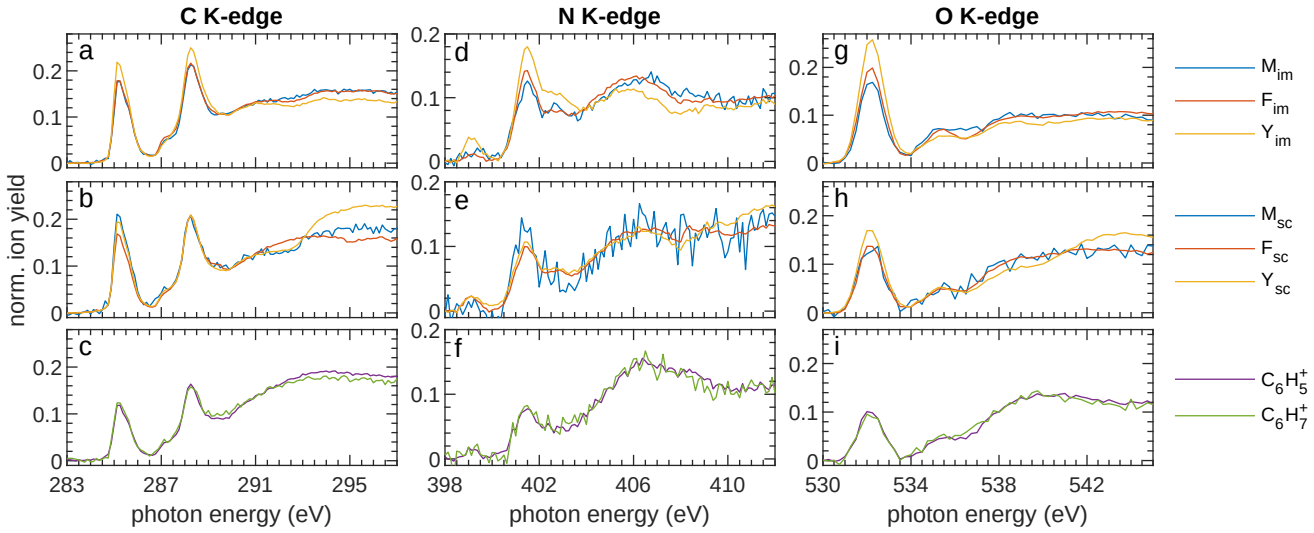

Figure 6: Normalized partial ion yields of the phenylalanine (F), tyrosine (Y) and methionine (M) immonium ions (im) and complete side-chain fragments (sc) as well as the phenyl cation ($C_6H_5^+$) and arenium ion ($C_6H_7^+$) measured upon photofragmentation of protonated methionine enkephalin at the (a-c) carbon, (d-f) nitrogen and (g-i) oxygen K-edges. Each spectrum is normalized to its total area.

The Tyr residue is the N-terminal residue of $[MetEnk+H]^+$. The N-terminal amino group is, commonly, the most favorable protonation site in a peptide without basic amino-acid side chains. Therefore, the increased yields of Y_{im} observed after excitations into the peptide bond is very likely connected to the close proximity to the N-terminal proton.

The normalized yields of the arenium ion and phenyl cation are almost identical at all absorption edges (Figures 6c,f,i). This suggests that both fragments stem from the same fragmentation pathway. The fragmentation of Y_{sc} into the arenium ion and the subsequent formation of the phenyl cation has been demonstrated in a CID study by Bauerle and Brodbelt [80] on cationic methylene-substituted phenol (equivalent to Y_{sc}). In contrast, there is no evidence for the arenium ion as a fragment of the Phe residue or toluene. From this we conclude that upon soft X-ray photon absorption in $[MetEnk+H]^+$ the phenyl cation is produced mainly as a daughter-fragment of the Tyr side chain with the arenium ion as an intermediate.

All amino acid side chains in $[MetEnk+H]^+$ contain at least three carbon atoms. Thus,

at the carbon K-edge more prominent effects caused by excitations in the amino-acid side chains can be expected than at the other absorption edges. The normalized yields of the complete side-chain fragments (F_{sc}, Y_{sc} and M_{sc}) indeed show the biggest differences at the carbon K-edge (Figure 6b). Upon direct ionization the normalized yield of Y_{sc} is higher and the normalized yield of F_{sc} is lower than for M_{sc}. In the σ^* and ionization edge regime the yield of F_{sc} appears to rise in the form of one continuous spectral feature, while the yields of Y_{sc} and M_{sc} are clearly rising in two steps. Visually, this creates the effects that the ionization potentials in the partial NEXAMS spectra of these fragments are at different energy positions. However, the fitting procedure gave more consistent results when keeping the energy positions of all spectral features and the ionization edge at similar positions.

The numbers of carbon atoms in the Phe and Tyr side chain are identical. The rate of direct ionizations should therefore not be the reason for the different normalized yields of these fragments at the ionization edge. One explanation might be that the additional degree of freedom in Y_{sc} compared to F_{sc} makes this fragment more robust towards dissociation into daughter fragments. Also the localized positive Coulomb field of the N-terminal proton or an involvement of the proton itself might have an influence on the yield of Y_{sc} upon direct ionization. A positive charge created by a direct ionization of the C_α atom of the Tyr residue is in immediate proximity to the N-terminus. At this distance Coulomb repulsion might come into play.

At the oxygen K-edge the normalized yields of F_{sc} and M_{sc} are very similar (Figure 6h). In comparison, the normalized yield of Y_{sc} is higher upon O 1s $\rightarrow \pi^*$(CONH) excitation and direct ionization but lower in the σ^* regime. When normalizing the spectra to the maximum of the π^*(CONH) feature the only strong difference between the ion yields is a significantly lower yield for Y_{sc} in the σ^* region. Since in this energy region the yields of the side-chain daughter fragments ($C_6H_7^+$ and $C_6H_5^+$) are the highest, it could be possible that the yield of Y_{sc} is decreased because excitations from the tyrosine hydroxy oxygen atom to σ^* orbitals lead to fragmentation within the Tyr side chain.

The normalized ion yield spectra of F_{sc}, Y_{sc} and M_{sc} measured at the nitrogen K-edge are again very similar (Figure 6e). Only the normalized yield of Y_{sc} is slightly higher at the ionization edge. None of the amino acid side chains in MetEnk contain nitrogen. Thus, fragmentation at the nitrogen K-edge takes place exclusively after excitations in the peptide backbone and the production of side-chain fragments should be driven mainly by IVR.

Evidence for Sulfur-Aromatic Interactions

In order to identify the cause for the increased ion yield upon C 1s $\rightarrow \pi^*$(C=C) excitation of $[MetEnk+H]^+$ compared to $[LeuEnk+H]^+$ the ratios between the normalized areas of peak CK1 determined for both peptides were calculated for different fragments and compared to each other (Figure 7a). The peak areas of fragments M_{sc} and M_{im} were compared to those of L_{sc} and L_{im}.

The only area ratio smaller than one can be observed between M_{im} and L_{im}. Most fragments show area ratios between 1.02 and 1.15. With values of (1.26 ± 0.03) and (1.58 ± 0.11) the area ratios of F_{sc} and between M_{sc} and L_{sc} are significantly higher. The area ratios for peak CK3 (C 1s $\rightarrow \pi^*$(CONH)) of these fragments are (1.03 ± 0.03) and (1.03 ± 0.07). This shows that the high area ratios for CK1 are not caused by a biased normalization.

The increased relative yield of M_{sc} (compared to L_{sc}) upon C 1s $\rightarrow \pi^*$(C=C) excitation is surprising because the Met residue does not contain an aromatic carbon ring. Since the Met and Phe residue are direct neighbors, the reason for the increased yields of M_{sc} and F_{sc} might be a direct intramolecular charge or energy transfer between the sulfur atom and the aromatic π-system of the Phe side chain. Another indication for a direct interaction between these side chains instead of a charge migration or energy transfer through the peptide backbone is the fact that no increased yields were observed for F_{im} and M_{im}. Sulfur-aromatic interactions play a unique role in structure stabilizing binding motifs in proteins and were focus of numerous computational and crystallographic studies.[102–107] To our best knowledge,

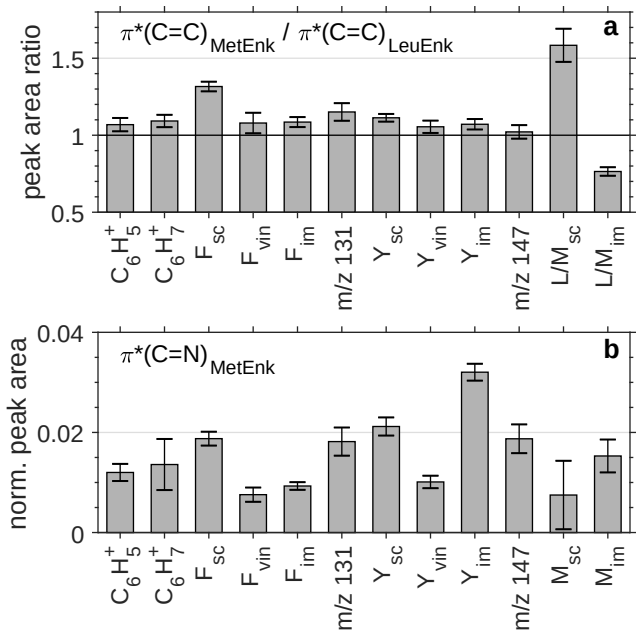

Figure 7: **(a)** Ratios between normalized partial ion yields of different fragments of $[MetEnk+H]^+$ and $[LeuEnk+H]^+$ measured upon C 1s $\rightarrow$ π^*(C=C) excitation. **(b)** Normalized partial ion yields of different fragments of $[MetEnk+H]^+$ measured upon N 1s $\rightarrow$ π^*(C=N) excitation.

effects of sulfur-aromatic interactions in peptides have so far never been observed in an X-ray absorption spectroscopy experiment.

Discussion of the 399.2 eV feature of $[MetEnk+H]^+$

Features around the energy position of NK4 (see Figure 4c) are characteristic for N 1s $\rightarrow$ π^*(C=N) transitions, as reported for example for plasma-polymerized propylamin (399.1 eV [108]) and allylamin (399.1 eV,[108] 398.7 eV [109]) as well as plasma-nitrogenated polyolefin surfaces (398.9 eV).[110] The common chemical structure of protonated MetEnk (Figure 1b), where all peptide bonds are in keto-form ($-$C(=O)NH$-$; also known as amide-form), does not contain any C=N groups. However, another tautomer of the peptide bond is the enol-form (also known as imidic-acid-form), where the hydrogen atom is bound to the oxygen atom ($-$C(OH)=N$-$). The energetically more favorable tautomer is the keto-form.[111,112] An X-ray

absorption spectroscopic signature of the keto-to-enol conversion of peptide bonds has been previously observed in the N K-edge XPS spectrum of the tripeptide glypromate isolated in an argon matrix (399.2 eV[89]) and in the N K-edge XAS spectra of S-Layer protein substrates interacting with metal atoms (398.4 eV[113]). In both cases, the enol-form was not the natural tautomer and was found to be produced either by UV irradiation or by protein-metal interactions, respectively.

To evaluate the effects of a tautomeric keto-to-enol conversion of a peptide bond on the theoretical nitrogen K-edge absorption spectrum of $[MetEnk+H]^+$ we performed DFT/ROCIS calculations for N-terminal protonated MetEnk with its Gly-Gly bond in enol-tautomer form (Figure 1c). The Gly-Gly bond was chosen for this structural change because a variety of computational studies have been performed on the keto-enol-tautomerism of peptide bonds in polyglycine peptides.[114–116] The calculated spectrum shows a very intense peak at 398.8 eV (Figure 8c) originating exclusively from N 1s $\rightarrow \pi^*$(C=N) transitions in the altered peptide bond. The ratio between the maximum intensities of this peak and the π^*(CONH) feature is about 0.8:1. This is in harsh contrast to the experimental data where this ratio is about 0.17:1. Consequently, assuming that peak NK4 is in fact due to transitions into MOs of a peptide bond in enol-form, the equilibrium between the two tautomers still lies heavily to the keto-form.

It could be possible that enol-tautomers have already been present in the MetEnk sample solution. The solvent used was a 1:1 mixture of water and methanol and the pH value was 2.7. Calculations have shown that both water and methanol drastically reduce the potential barrier between the peptide bond's enol- and keto-form.[114] Furthermore, the enol-form is a known intermediate in the acid-catalyzed proton exchange in peptide bonds.[117–119] However, based on this explanation the enol-tautomer of $[LeuEnk+H]^+$ should be expected as well, because the compositions of all sample solutions were identical. It remains unclear why the structural change leading to peak NK4 did not take place in $[LeuEnk+H]^+$ but since the peptides were measured using different ion funnels it might have been induced by collisions

of precursor ions with residual gas molecules in the ion-funnel region where the pressure is still in the order of millibars.

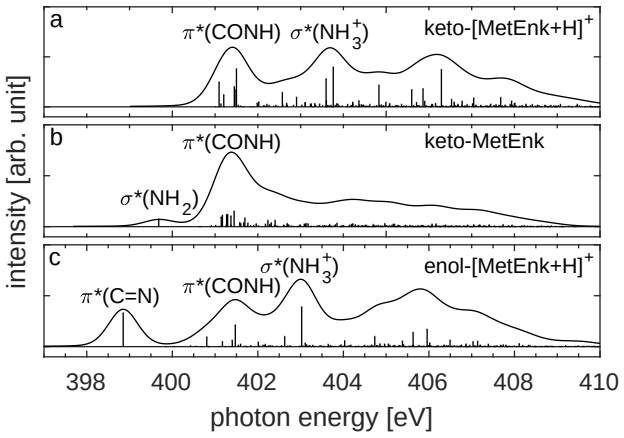

Figure 8: DFT/ROCIS-calculated X-ray absorption spectra for **(a)** N-terminal protonated MetEnk with all peptide bonds in keto-form (shifted +12.6 eV), **(b)** neutral MetEnk with all peptide bonds in keto-form (shifted +12.3 eV) and **(c)** N-terminal protonated MetEnk with the Gly-Gly peptide bond in enol-form (shifted +12.2 eV). The calculated discrete absorption lines were broadened using a pseudo-Voigt profile with 0.7 eV FWHM.

Another possible reason for peak NK4 are N 1s $\rightarrow \sigma^*$(N−H) transitions in an unprotonated amino group. The peak positions of the protonated and unprotonated amino groups in the XPS spectra of different zwitterionic amino acids reported by Zubavichus et al.[90] are (401.4 ± 0.2) eV and (399.7 ± 0.2) eV. When the N-terminus is unprotonated the proton must, consequently, sit at another site of the peptide. While it has been shown that the most favorable protonation site of $[LeuEnk+H]^+$ is the N-terminus,[41] the protonation site of $[MetEnk+H]^+$ might be influenced by the proton affinity of the Met residue, which is slightly higher than for the Leu residue.[120] Experimental evidence has been reported for backbone protonation in a doubly protonated synthetic peptide[121] and in singly protonated triglycine for which it was demonstrated that protonation occurs at an amide oxygen atom.[122,123]

We performed DFT/ROCIS calculations for unprotonated MetEnk in all-keto-form (Figure 1b) in order to evaluate the effect of the unprotonated amino group on the theoretical nitrogen K-edge absorption spectrum. The calculated spectrum features a peak at 399.7 eV

(Figure 8b) attributed to N 1s $\rightarrow \sigma^*$(N−H) transitions in the terminal NH_2 group. It should be noted that Zubavichus et al.[124] observed a peak at ∼400 eV appearing in the X-ray photoelectron spectra and peaks at 398.5 eV and 399.7 eV emerging in the X-ray fluorescence spectra of solid zwitterionic Tyr and Phe after partial decomposition of the samples using soft X-ray radiation. While they considered the formation of C=N bonds as well, ultimately, the authors assigned these features to transitions in unprotonated amino groups.

In order to locate the site of the electronic transition behind peak NK4 the areas of this peak in the normalized partial NEXAMS spectra of different fragments were compared (Figure 7b). With (0.032 ± 0.002) Y_{im} exhibits the largest peak area for this feature. The second highest area (0.021 ± 0.002) was observed for Y_{sc}. The peak areas of F_{sc} and the fragments at m/z 131 and 147 are at a comparable level at around 0.019. Since F_{sc} can also occur as a fragment of the Tyr side chain and the fragments at m/z 131 and 147 can be assigned to fragments $[c_2 - Y_{sc}]^+$ and Y_{acr}, all fragments with an increased area of peak NK4 can be associated with the Tyr residue. This indicates that the transitions leading to peak NK4 take place near the N-terminus. Assuming that there is one peptide bond in enol-form it would, therefore, most likely be the Tyr-Gly bond. However, since Tyr is the most N-terminal amino acid residue in MetEnk N 1s $\rightarrow \sigma^*$(N−H) transitions in an unprotonated amino group remain a plausible reason for peak NK4 as well.

Based on our DFT/ROCIS calculations alone, both presented types of transitions are possible reasons for peak NK4. Since the experimental energy position of this spectral feature is in slightly better agreement with the values for N 1s $\rightarrow \pi^*$(C=N) transitions reported in literature we consider this assignment as more likely. Nevertheless, further experiments are needed to undoubtedly assess the origin of this feature.

Conclusions

We have conducted a NEXAMS study on cryogenically cooled gas-phase $[LeuEnk+H]^+$ and $[MetEnk+H]^+$ by measuring their cationic fragments produced upon soft X-ray photon absorption at the carbon, nitrogen and oxygen K-edges. Our results show differences between the total and partial NEXAMS spectra of both peptides which could be interpreted in relation to aspects of their structure, including amino-acid side-chain interactions and tautomerism of peptide bonds.

A spectral feature at 399.2 eV was observed only in the nitrogen K-edge NEXAMS spectra of $[MetEnk+H]^+$. Increased areas of this peak in the partial NEXAMS spectra of fragments related to tyrosine suggest that the electronic transitions behind this peak takes place near the N-terminal Tyr residue. We discussed both N 1s $\rightarrow \pi^*$(C=N) transitions in the enol-form of the Tyr-Gly peptide bond and N 1s $\rightarrow \sigma^*$(N−H) transitions in unprotonated amino groups as plausible reasons for this feature. The DFT/ROCIS-calculated nitrogen K-edge absorption spectra for different MetEnk isomers gave evidence for both electronic transitions but the better agreement with peaks reported for this energy region in literature suggests that N 1s $\rightarrow \pi^*$(C=N) transitions are the more probable reason. For a final assignment measurements on individual isomers of gas-phase $[MetEnk+H]^+$ are necessary. Especially C=N bonds could serve as dedicated excitation sites in studies on local effects of site-specific photoexcitation. In this context, also an investigation on the dynamics of the UV photon-induced tautomeric keto-to-enol conversion of peptide bonds, demonstrated by Mateo-Marti and Pradier,[89] in an UV pump - soft X-ray probe experiment is of high interest.

Furthermore, the total measured ion yield for $[MetEnk+H]^+$ upon C 1s $\rightarrow \pi^*$(C=C) excitation was 13.7% higher than for $[LeuEnk+H]^+$. Among all fragments this effect was the most prominent for the complete Phe and Met side-chain fragments. Therefore, we propose an involvement of a direct sulfur-aromatic interaction between the Phe and Met side chains, promoting the formation of the respective side-chain fragments. Investigations on custom peptides with different pairs of neighboring sulfur-containing and aromatic amino

acid residues will be conducted to gain more information about the effects of sulfur-aromatic interactions on the soft X-ray photon-induced dissociation of peptides.

Conflicts of interest

There are no conflicts of interest to declare.

Acknowledgement

We thank the Helmholtz-Center Berlin for the provision of synchrotron radiation at the UE52-PGM beamline. S.Dörner, L.Schwob, K.Schubert and S.Bari acknowledge funding from the Helmholtz Initiative and Networking Fund through the Young Investigators Group Program (VH-NG-1104). Furthermore, K.Schubert, S.Techert and S.Bari were supported by the Deutsche Forschungsgemeinschaft, project B03 in the SFB 755 - Nanoscale Photonic Imaging. The NanoclusterTrap endstation at UE52_PGM is partially funded by the BMBF grant 05K16VF1.

Supporting Information Available

The following files are available free of charge.

- manuscript_EnkCNO_SI.pdf:
  - DFT-optimized geometries
  - isosurfaces of DFT-calculated LUMOs
  - intensities and areas of fitted spectral features

- assignments and structural formulas of detected fragments

References

(1) Gordon, M. L.; Cooper, G.; Morin, C.; Araki, T.; Turci, C. C.; Kaznatcheev, K.; Hitchcock, A. P. Inner-shell excitation spectroscopy of the peptide bond: Comparison of the C 1s, N 1s, and O 1s spectra of glycine, glycyl-glycine, and glycyl-glycyl-glycine. *Journal of Physical Chemistry A* **2003**, *107*, 6144–6159.

(2) Stewart-Ornstein, J.; Hitchcock, A. P.; Hernández Cruz, D.; Henklein, P.; Overhage, J.; Hilpert, K.; Hale, J. D.; Hancock, R. E. W. Using Intrinsic X-ray Absorption Spectral Differences To Identify and Map Peptides and Proteins. *The Journal of Physical Chemistry B* **2007**, *111*, 7691–7699.

(3) Zubavichus, Y.; Shaporenko, A.; Grunze, M.; Zharnikov, M. Innershell absorption spectroscopy of amino acids at all relevant absorption edges. *Journal of Physical Chemistry A* **2005**, *109*, 6998–7000.

(4) Zubavichus, Y.; Shaporenko, A.; Grunze, M.; Zharnikov, M. NEXAFS spectroscopy of homopolypeptides at all relevant absorption edges: Polyisoleucine, polytyrosine, and polyhistidine. *Journal of Physical Chemistry B* **2007**, *111*, 9803–9807.

(5) Zubavichus, Y.; Shaporenko, A.; Grunze, M.; Zharnikov, M. Is X-ray absorption spectroscopy sensitive to the amino acid composition of functional proteins? *Journal of Physical Chemistry B* **2008**, *112*, 4478–4480.

(6) Feyer, V.; Plekan, O.; Richter, R.; Coreno, M.; Prince, K. C.; Carravetta, V. Photoemission and Photoabsorption Spectroscopy of Glycyl-Glycine in the Gas Phase. *The Journal of Physical Chemistry A* **2009**, *113*, 10726–10733.

(7) Zubavichus, Y.; Shaporenko, A.; Grunze, M.; Zharnikov, M. NEXAFS spectroscopy of biological molecules: From amino acids to functional proteins. *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **2009**, *603*, 111–114.

(8) González-Magaña, O.; Reitsma, G.; Tiemens, M.; Boschman, L.; Hoekstra, R.; Schlathölter, T. Near-edge X-ray absorption mass spectrometry of a gas-phase peptide. *Journal of Physical Chemistry A* **2012**, *116*, 10745–10751.

(9) Milosavljević, A.; Canon, F.; Nicolas, C.; Miron, C.; Nahon, L.; Giuliani, A. Gas-phase protein inner-shell spectroscopy by coupling an ion trap with a soft X-ray beamline. *Journal of Physical Chemistry Letters* **2012**, *3*, 1191–1196.

(10) Egorov, D.; Bari, S.; Boll, R.; Dörner, S.; Deinert, S.; Techert, S.; Hoekstra, R.; Zamudio-Bayer, V.; Lindblad, R.; Bülow, C.; Timm, M.; von Issendorff, B.; Lau, J. T.; Schlathölter, T. Near-Edge Soft X-ray Absorption Mass Spectrometry of Protonated Melittin. *Journal of The American Society for Mass Spectrometry* **2018**, *29*, 2138–2151.

(11) Bari, S. et al. Soft X-ray Spectroscopy as a Probe for Gas-Phase Protein Structure: Electron Impact Ionization from Within. *Chemistry - A European Journal* **2018**, *24*, 7631–7636.

(12) Fujii, K.; Akamatsu, K.; Muramatsu, Y.; Yokoya, A. X-ray absorption near edge structure of DNA bases around oxygen and nitrogen K-edge. *Nuclear Instruments and Methods in Physics Research, Section B: Beam Interactions with Materials and Atoms* **2003**, *199*, 249–254.

(13) Samuel, N. T.; Lee, C.-Y.; Gamble, L. J.; Fischer, D. A.; Castner, D. G. NEXAFS characterization of DNA components and molecular-orientation of surface-bound DNA

oligomers. *Journal of Electron Spectroscopy and Related Phenomena* **2006**, *152*, 134–142.

(14) Lee, C.-Y.; Gong, P.; Harbers, G. M.; Grainger, D. W.; Castner, D. G.; Gamble, L. J. Surface Coverage and Structure of Mixed DNA/Alkylthiol Monolayers on Gold: Characterization by XPS, NEXAFS, and Fluorescence Intensity Measurements. *Analytical Chemistry* **2006**, *78*, 3316–3325.

(15) Petrovykh, D. Y.; Pérez-Dieste, V.; Opdahl, A.; Kimura-Suda, H.; Sullivan, J. M.; Tarlov, M. J.; Himpsel, F. J.; Whitman, L. J. Nucleobase Orientation and Ordering in Films of Single-Stranded DNA on Gold. *Journal of the American Chemical Society* **2006**, *128*, 2–3.

(16) Plekan, O.; Feyer, V.; Richter, R.; Coreno, M.; de Simone, M.; Prince, K.; Trofimov, A.; Gromov, E.; Zaytseva, I.; Schirmer, J. A theoretical and experimental study of the near edge X-ray absorption fine structure (NEXAFS) and X-ray photoelectron spectra (XPS) of nucleobases: Thymine and adenine. *Chemical Physics* **2008**, *347*, 360–375.

(17) Hua, W.; Gao, B.; Li, S.; Ågren, H.; Luo, Y. Refinement of DNA Structures through Near-Edge X-ray Absorption Fine Structure Analysis: Applications on Guanine and Cytosine Nucleobases, Nucleosides, and Nucleotides. *The Journal of Physical Chemistry B* **2010**, *114*, 13214–13222.

(18) Howell, C.; Hamoudi, H.; Zharnikov, M. Thymine/adenine diblock-oligonucleotide monolayers and hybrid brushes on gold: a spectroscopic study. *Biointerphases* **2013**, *8*, 6.

(19) Lin, Y.-S.; Lin, H.-R.; Liu, W.-L.; Lee, Y. T.; Tseng, C.-M.; Ni, C.-K.; Liu, C.-L.; Tsai, C.-C.; Chen, J.-L.; Hu, W.-P. Measurement and prediction of the NEXAFS

spectra of pyrimidine and purine and the dissociation following the core excitation. *Chemical Physics Letters* **2015**, *636*, 146–153.

(20) Tsud, N.; Bercha, S.; Ševčíková, K.; Acres, R. G.; Prince, K. C.; Matolín, V. Adenine adlayers on Cu(111): XPS and NEXAFS study. *The Journal of Chemical Physics* **2015**, *143*, 174704.

(21) Kaznacheyev, K.; Osanna, A.; Jacobsen, C.; Plashkevych, O.; Vahtras, O.; Ågren, H.; Carravetta, V.; Hitchcock, A. P. Innershell absorption spectroscopy of amino acids. *Journal of Physical Chemistry A* **2002**, *106*, 3153–3168.

(22) Boese, J.; Osanna, A.; Jacobsen, C.; Kirz, J. Carbon edge XANES spectroscopy of amino acids and peptides. *Journal of Electron Spectroscopy and Related Phenomena* **1997**, *85*, 9–15.

(23) Schlathölter, T.; Hoekstra, R. In *Photophysics of Ionic Biochromophores*; Brøndsted Nielsen, S., Wyer, J. A., Eds.; Physical Chemistry in Action; Springer Berlin Heidelberg: Berlin, Heidelberg, 2013; Chapter 11, pp 209–226.

(24) Reitsma, G.; Boschman, L.; Deuzeman, M. J.; Hoekstra, S.; Hoekstra, R.; Schlathölter, T. Near edge X-ray absorption mass spectrometry on coronene. *The Journal of Chemical Physics* **2015**, *142*, 024308.

(25) Schwob, L.; Dörner, S.; Atak, K.; Schubert, K.; Timm, M.; Bülow, C.; Zamudio-Bayer, V.; von Issendorff, B.; Lau, J. T.; Techert, S.; Bari, S. Site-Selective Dissociation upon Sulfur L-Edge X-ray Absorption in a Gas-Phase Protonated Peptide. *The Journal of Physical Chemistry Letters* **2020**, *11*, 1215–1221.

(26) Kopysov, V.; Makarov, A.; Boyarkin, O. V. Colors for Molecular Masses: Fusion of Spectroscopy and Mass Spectrometry for Identification of Biomolecules. *Analytical Chemistry* **2015**, *87*, 4607–4611.

(27) Saparbaev, E.; Kopysov, V.; Yamaletdinov, R.; Pereverzev, A. Y.; Boyarkin, O. V. Interplay of H-Bonds with Aromatics in Isolated Complexes Identifies Isomeric Carbohydrates. *Angewandte Chemie International Edition* **2019**, *58*, 7346–7350.

(28) Giuliani, A.; Milosavljević, A.; Canon, F.; Nahon, L. Contribution of Synchrotron Radiation to Photoactivation Studies of Biomolecular Ions in the Gas Phase. *Mass Spectrometry Reviews2* **2014**, *33*, 424–441.

(29) Abdelmouleh, M.; Lalande, M.; El Feghaly, J.; Vizcaino, V.; Rebelo, A.; Eden, S.; Schlathölter, T.; Poully, J.-C. Mass Spectral Signatures of Complex Post-Translational Modifications in Proteins: A Proof-of-Principle Based on X-ray Irradiated Vancomycin. *Journal of the American Society for Mass Spectrometry* **2020**, *31*, 1738–1743.

(30) Egorov, D.; Schwob, L.; Lalande, M.; Hoekstra, R.; Schlathölter, T. Near edge X-ray absorption mass spectrometry of gas phase proteins: The influence of protein size. *Physical Chemistry Chemical Physics* **2016**, *18*, 26213–26223.

(31) Schwob, L.; Lalande, M.; Egorov, D.; Rangama, J.; Hoekstra, R.; Vizcaino, V.; Schlathölter, T.; Poully, J.-C. Radical-driven processes within a peptidic sequence of type I collagen upon single-photon ionisation in the gas phase. *Phys. Chem. Chem. Phys.* **2017**, *19*, 22895–22904.

(32) Milosavljević, A. R.; Nicolas, C.; Ranković, M. L.; Canon, F.; Miron, C.; Giuliani, A. K-Shell Excitation and Ionization of a Gas-Phase Protein: Interplay between Electronic Structure and Protein Folding. *Journal of Physical Chemistry Letters* **2015**, *6*, 3132–3138.

(33) Alexander, A.; Boyd, R. Experimental investigations of factors controlling the collision induced dissociation spectra of peptide ions in a tandem hybrid mass spectrometer.

I. Leucine enkephalin. *International Journal of Mass Spectrometry and Ion Processes* **1989**, *90*, 211–240.

(34) Vachet, R. W.; Bishop, B. M.; Erickson, B. W.; Glish, G. L. Novel peptide dissociation: Gas-phase intramolecular rearrangement of internal amino acid residues. *Journal of the American Chemical Society* **1997**, *119*, 5481–5488.

(35) Rakov, V. S.; Borisov, O. V.; Whitehouse, C. M. Establishing low-energy sequential decomposition pathways of leucine enkephalin and its N- and C-terminus fragments using multiple-resonance CID in quadrupolar ion guide. *Journal of the American Society for Mass Spectrometry* **2004**, *15*, 1794–1809.

(36) Laskin, J. Energetics and dynamics of fragmentation of protonated leucine enkephalin from time-and energy-resolved surface-induced dissociation studies. *Journal of Physical Chemistry A* **2006**, *110*, 8554–8562.

(37) Li, L.; Lubman, D. M. Resonant Two-Photon Ionization of Enkephalins and Related Peptides Volatilized by Using Pulsed Laser Desorption in Supersonic Beam Mass Spectrometry. *Analytical Chemistry* **1988**, *60*, 1409–1415.

(38) Tabarin, T.; Antoine, R.; Broyer, M.; Dugourd, P. Specific photodissociation of peptides with multi-stage mass spectrometry. *Rapid Communications in Mass Spectrometry* **2005**, *19*, 2883–2892.

(39) Herburger, A.; van der Linde, C.; Beyer, M. K. Photodissociation spectroscopy of protonated leucine enkephalin. *Physical Chemistry Chemical Physics* **2017**, *19*, 10786–10795.

(40) Burke, N. L.; Redwine, J. G.; Dean, J. C.; Mcluckey, S. A.; Zwier, T. S. UV and IR spectroscopy of cold protonated leucine enkephalin. *International Journal of Mass Spectrometry* **2015**, *378*, 196–205.

(41) Polfer, N. C.; Oomens, J.; Suhai, S.; Paizs, B. Infrared spectroscopy and theoretical studies on gas-phase protonated Leu-enkephalin and its fragments: Direct experimental evidence for the mobile proton. *Journal of the American Chemical Society* **2007**, *129*, 5887–5897.

(42) Ranković, M. L.; Canon, F.; Nahon, L.; Giuliani, A.; Milosavljević, A. R. VUV action spectroscopy of protonated leucine-enkephalin peptide in the 6-14 eV range. *The Journal of Chemical Physics* **2015**, *143*, 244311.

(43) Bari, S.; Gonzalez-Magaña, O.; Reitsma, G.; Werner, J.; Schippers, S.; Hoekstra, R.; Schlathölter, T. Photodissociation of protonated leucine-enkephalin in the VUV range of 8-40 eV. *Journal of Chemical Physics* **2011**, *134*, 1–9.

(44) Bari, S.; Hoekstra, R.; Schlathölter, T. Peptide fragmentation by keV ion-induced dissociation. *Physical Chemistry Chemical Physics* **2010**, *12*, 3376.

(45) Reitsma, G.; Gonzalez-Magaña, O.; Versolato, O.; Door, M.; Hoekstra, R.; Suraud, E.; Fischer, B.; Camus, N.; Kremer, M.; Moshammer, R.; Schlathölter, T. Femtosecond laser induced ionization and dissociation of gas-phase protonated leucine enkephalin. *International Journal of Mass Spectrometry* **2014**, *365-366*, 365–371.

(46) Niemeyer, M.; Hirsch, K.; Zamudio-Bayer, V.; Langenberg, A.; Vogel, M.; Kossick, M.; Ebrecht, C.; Egashira, K.; Terasaki, A.; Möller, T.; v. Issendorff, B.; Lau, J. T. Spin Coupling and Orbital Angular Momentum Quenching in Free Iron Clusters. *Physical Review Letters* **2012**, *108*, 057201.

(47) Hirsch, K.; Lau, J. T.; Klar, P.; Langenberg, A.; Probst, J.; Rittmann, J.; Vogel, M.; Zamudio-Bayer, V.; Möller, T.; von Issendorff, B. X-ray spectroscopy on size-selected clusters in an ion trap: from the molecular limit to bulk properties. *Journal of Physics B: Atomic, Molecular and Optical Physics* **2009**, *42*, 154029.

(48) Egorov, D.; Sadia, B.; Hoekstra, R.; Lawicki, A.; Hirsch, K.; Zamudio-Bayer, V.; Lau, T.; Von Issendorf, B.; Schlathölter, T. An intense electrospray ionization source for soft X-ray photoionization of gas phase protein ions. *Journal of Physics: Conference Series* **2015**, *635*, 9–10.

(49) Carroll, T. X.; Hahne, J.; Thomas, T. D.; Sæthre, L. J.; Berrah, N.; Bozek, J.; Kukk, E. Carbon 1s core-hole lifetime in CO_2. *Physical Review A* **2000**, *61*, 042503.

(50) Kempgens, B.; Kivimäki, A.; Neeb, M.; Köppe, H. M.; Bradshaw, A. M.; Feldhaus, J. A high-resolution N 1s photoionization study of the N_2 molecule in the near-threshold region. *Journal of Physics B: Atomic, Molecular and Optical Physics* **1996**, *29*, 5389–5402.

(51) Coreno, M.; de Simone, M.; Prince, K.; Richter, R.; Vondráček, M.; Avaldi, L.; Camilloni, R. Vibrationally resolved oxygen $K \rightarrow \Pi^*$ spectra of O_2 and CO. *Chemical Physics Letters* **1999**, *306*, 269–274.

(52) Neese, F. Software update: the ORCA program system, version 4.0. *WIREs Computational Molecular Science* **2018**, *8*, 4–9.

(53) Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Physical Chemistry Chemical Physics* **2005**, *7*, 3297.

(54) Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Physical Review A* **1988**, *38*, 3098–3100.

(55) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *The Journal of Chemical Physics* **1993**, *98*, 5648–5652.

(56) Roemelt, M.; Maganas, D.; DeBeer, S.; Neese, F. A combined DFT and restricted open-shell configuration interaction method including spin-orbit coupling: Application

to transition metal L-edge X-ray absorption spectroscopy. *The Journal of Chemical Physics* **2013**, *138*, 204101.

(57) Baerends, E.; Ellis, D.; Ros, P. Self-consistent molecular Hartree-Fock-Slater calculations I. The computational procedure. *Chemical Physics* **1973**, *2*, 41–51.

(58) Dunlap, B. I.; Connolly, J. W. D.; Sabin, J. R. On some approximations in applications of Xα theory. *The Journal of Chemical Physics* **1979**, *71*, 3396–3402.

(59) Vahtras, O.; Almlöf, J.; Feyereisen, M. Integral approximations for LCAO-SCF calculations. *Chemical Physics Letters* **1993**, *213*, 514–518.

(60) Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. Auxiliary basis sets to approximate Coulomb potentials. *Chemical Physics Letters* **1995**, *240*, 283–290.

(61) Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. Auxiliary basis sets for main row atoms and transition metals and their use to approximate Coulomb potentials. *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)* **1997**, *97*, 119–124.

(62) Stoychev, G. L.; Auer, A. A.; Neese, F. Automatic Generation of Auxiliary Basis Sets. *Journal of Chemical Theory and Computation* **2017**, *13*, 554–562.

(63) van Lenthe, E.; Snijders, J. G.; Baerends, E. J. The zero-order regular approximation for relativistic effects: The effect of spin-orbit coupling in closed shell molecules. *The Journal of Chemical Physics* **1996**, *105*, 6505–6516.

(64) Wertheim, G. K.; Butler, M. A.; West, K. W.; Buchanan, D. N. E. Determination of the Gaussian and Lorentzian content of experimental line shapes. *Review of Scientific Instruments* **1974**, *45*, 1369–1371.

(65) Young, R. A.; Wiles, D. B. Profile shape functions in Rietveld refinements. *Journal of Applied Crystallography* **1982**, *15*, 430–438.

(66) Ida, T.; Ando, M.; Toraya, H. Extended pseudo-Voigt function for approximating the Voigt profile. *Journal of Applied Crystallography* **2000**, *33*, 1311–1316.

(67) Thompson, P.; Cox, D. E.; Hastings, J. B. Rietveld Refinement of Debye-Scherrer Synchrotron X-ray Data from Al_2O_3. *Journal of Applied Crystallography* **1987**, *20*, 79–83.

(68) Falick, A. M.; Hines, W. M.; Medzihradszky, K. F.; Baldwin, M. A.; Gibson, B. W. Low-mass ions produced from peptides by high-energy collision-induced dissociation in tandem mass spectrometry. *Journal of the American Society for Mass Spectrometry* **1993**, *4*, 882–893.

(69) Papayannopoulos, I. A. The interpretation of collision-induced dissociation tandem mass spectra of peptides. *Mass Spectrometry Reviews* **1995**, *14*, 49–73.

(70) Roepstorff, P.; Fohlman, J. Letter to the editors. *Biological Mass Spectrometry* **1984**, *11*, 601–601.

(71) Shoeib, T.; Cunje, A.; Hopkinson, A. C.; Siu, K. W. M. Gas-phase fragmentation of the Ag^+-phenylalanine complex: Cation-π interactions and radical cation formation. *Journal of the American Society for Mass Spectrometry* **2002**, *13*, 408–416.

(72) Soorkia, S.; Broquier, M.; Grégoire, G. Conformer- and mode-specific excited state lifetimes of cold protonated tyrosine ions. *Journal of Physical Chemistry Letters* **2014**, *5*, 4349–4355.

(73) Zhao, J.; Shoeib, T.; Siu, K. M.; Hopkinson, A. C. The fragmentation of protonated tyrosine and iodotyrosines: The effect of substituents on the losses of NH_3 and of H_2O and CO. *International Journal of Mass Spectrometry* **2006**, *255-256*, 265–278.

(74) Von E.Doering, W.; Knox, L. H. The Cycloheptatrienylium (Tropylium) Ion. *Journal of the American Chemical Society* **1954**, *76*, 3203–3206.

(75) Rylander, P. N.; Meyerson, S.; Grubb, H. M. Organic Ions in the Gas Phase. II. The Tropylium Ion. *Journal of the American Chemical Society* **1957**, *79*, 842–846.

(76) Ausloos, P. Structure and isomerization of $C_7H_7^+$ ions formed in the charge-transfer-induced fragmentation of ethylbenzene, toluene, and norbornadiene. *Journal of the American Chemical Society* **1982**, *104*, 5259–5265.

(77) Zhao, G.; Gäumann, T. Isomeric $[C_7H_8]^{+\bullet}$ and $[C_7H_7]^+$ ions. *Organic Mass Spectrometry* **1992**, *27*, 428–434.

(78) NIST Standard Reference Database Number 69. 2018; `https://webbook.nist.gov/chemistry/`.

(79) Russell, D. H.; Freiser, B. S.; McBay, E. H.; Canada, D. C. The structure of decomposing $[C_7H_7O]^+$ ions: Benzyl versus tropylium ion structures. *Organic Mass Spectrometry* **1983**, *18*, 474–485.

(80) Bauerle, G. F.; Brodbelt, J. S. Evaluation of steric and substituent effects in phenols by competitive reactions of dimethyl ether ions in a quadrupole ion trap. *Journal of the American Society for Mass Spectrometry* **1995**, *6*, 627–633.

(81) Plekan, O.; Feyer, V.; Richter, R.; Coreno, M.; Prince, K. C. Valence photoionization and photofragmentation of aromatic amino acids. *Molecular Physics* **2008**, *106*, 1143–1153.

(82) Budzikiewicz, H.; Drabner, G.; Hammes, C. Rearrangements accompanying the fragmentation of ionized 1-phenylalkan-1-ols. *Organic Mass Spectrometry* **1993**, *28*, 1326–1328.

(83) Horsley, J. A.; Stöhr, J.; Hitchcock, A. P.; Newbury, D. C.; Johnson, A. L.; Sette, F. Resonances in the K shell excitation spectra of benzene and pyridine: Gas phase, solid, and chemisorbed states. *The Journal of Chemical Physics* **1985**, *83*, 6099–6107.

(84) Solomon, J.; Madix, R.; Stöhr, J. Orientation and absolute coverage of benzene, aniline, and phenol on Ag(110) determined by NEXAFS and XPS. *Surface Science* **1991**, *255*, 12–30.

(85) Petoral, R. M.; Uvdal, K. XPS and NEXAFS study of tyrosine-terminated propanethiol assembled on gold. *Journal of Electron Spectroscopy and Related Phenomena* **2003**, *128*, 159–164.

(86) Zhang, W.; Carravetta, V.; Plekan, O.; Feyer, V.; Richter, R.; Coreno, M.; Prince, K. C. Electronic structure of aromatic amino acids studied by soft x-ray spectroscopy. *The Journal of Chemical Physics* **2009**, *131*, 035103.

(87) Feyer, V.; Plekan, O.; Richter, R.; Coreno, M.; Prince, K. C.; Carravetta, V. Core Level Study of Alanine and Threonine. *The Journal of Physical Chemistry A* **2008**, *112*, 7806–7815.

(88) Hennig, C.; Hallmeier, K.; Szargan, R. XANES investigation of chemical states of nitrogen in polyaniline. *Synthetic Metals* **1998**, *92*, 161–166.

(89) Mateo-Marti, E.; Pradier, C. UV irradiation study of a tripeptide isolated in an argon matrix: A tautomerism process evidenced by infrared and X-ray photoemission spectroscopies. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **2013**, *109*, 247–252.

(90) Zubavichus, Y.; Fuchs, O.; Weinhardt, L.; Heske, C.; Umbach, E.; Denlinger, J. D.; Grunze, M. Soft X-Ray-Induced Decomposition of Amino Acids: An XPS, Mass Spectrometry, and NEXAFS Study. *Radiation Research* **2004**, *161*, 346–358.

(91) Milosavljević, A. R.; Jänkälä, K.; Ranković, M. L.; Canon, F.; Bozek, J.; Nicolas, C.; Giuliani, A. Oxygen K-shell spectroscopy of isolated progressively solvated peptide. *Physical Chemistry Chemical Physics* **2020**, *22*, 12909–12917.

(92) Prince, K. C.; Richter, R.; de Simone, M.; Alagia, M.; Coreno, M. Near Edge X-ray Absorption Spectra of Some Small Polyatomic Molecules. *The Journal of Physical Chemistry A* **2003**, *107*, 1955–1963.

(93) Francis, J. T.; Hitchcock, A. P. Inner-shell spectroscopy of p-benzoquinone, hydroquinone, and phenol: distinguishing quinoid and benzenoid structures. *The Journal of Physical Chemistry* **1992**, *96*, 6598–6610.

(94) Jr., R. M. P.; Uvdal, K. NEXAFS Study of Amino Acid Analogues Assembled on Gold. *Physica Scripta* **2005**, *T115*, 851.

(95) Salén, P.; Kamińska, M.; Squibb, R. J.; Richter, R.; Alagia, M.; Stranges, S.; Van Der Meulen, P.; Eland, J. H.; Feifel, R.; Zhaunerchyk, V. Selectivity in fragmentation of N-methylacetamide after resonant K-shell excitation. *Physical Chemistry Chemical Physics* **2014**, *16*, 15231–15240.

(96) Li, C.; Salén, P.; Yatsyna, V.; Schio, L.; Feifel, R.; Squibb, R.; Kamińska, M.; Larsson, M.; Richter, R.; Alagia, M.; Stranges, S.; Monti, S.; Carravetta, V.; Zhaunerchyk, V. Experimental and theoretical XPS and NEXAFS studies of N-methylacetamide and N-methyltrifluoroacetamide. *Physical Chemistry Chemical Physics* **2016**, *18*, 2210–2218.

(97) Plekan, O.; Feyer, V.; Richter, R.; Coreno, M.; de Simone, M.; Prince, K. C.; Carravetta, V. An X-ray absorption study of glycine, methionine and proline. *Journal of Electron Spectroscopy and Related Phenomena* **2007**, *155*, 47–53.

(98) Cavalleri, M.; Ogasawara, H.; Pettersson, L. G.; Nilsson, A. The interpretation of X-ray absorption spectra of water and ice. *Chemical Physics Letters* **2002**, *364*, 363–370.

(99) Myneni, S.; Luo, Y.; Näslund, L. Å.; Cavalleri, M.; Ojamäe, L.; Ogasawara, H.; Pelmenschikov, A.; Wernet, P.; Väterlein, P.; Heske, C.; Hussain, Z.; Pettersson, L. G. M.;

Nilsson, A. Spectroscopic probing of local hydrogen-bonding structures in liquid water. *Journal of Physics: Condensed Matter* **2002**, *14*, L213–L219.

(100) Hetényi, B.; De Angelis, F.; Giannozzi, P.; Car, R. Calculation of near-edge x-ray-absorption fine structure at finite temperatures: Spectral signatures of hydrogen bond breaking in liquid water. *The Journal of Chemical Physics* **2004**, *120*, 8632–8637.

(101) Pylkkänen, T.; Lehtola, J.; Hakala, M.; Sakko, A.; Monaco, G.; Huotari, S.; Hämäläinen, K. Universal Signature of Hydrogen Bonding in the Oxygen K -Edge Spectrum of Alcohols. *The Journal of Physical Chemistry B* **2010**, *114*, 13076–13083.

(102) Viguera, A. R.; Serrano, L. Side-chain interactions between sulfur-containing amino acids and phenyalanine in α-helixes. *Biochemistry* **1995**, *34*, 8771–8779.

(103) Zauhar, R. J.; Colbert, C. L.; Morgan, R. S.; Welsh, W. J. Evidence for a strong sulfur-aromatic interaction derived from crystallographic data. *Biopolymers* **2000**, *53*, 233–248.

(104) Morgan, R. S.; Tatsch, C. E.; Gushard, R. H.; Mcadon, J. M.; Warme, P. K. Chains of Alternating sulfur and π-bonded atoms in eight small proteins. *International Journal of Peptide and Protein Research* **2009**, *11*, 209–217.

(105) Morgan, R. S.; McAdon, J. M. Predictor for sulfur-aromatic interactions in globular proteins. *International Journal of Peptide and Protein Research* **2009**, *15*, 177–180.

(106) Valley, C. C.; Cembran, A.; Perlmutter, J. D.; Lewis, A. K.; Labello, N. P.; Gao, J.; Sachs, J. N. The Methionine-aromatic Motif Plays a Unique Role in Stabilizing Protein Structure. *Journal of Biological Chemistry* **2012**, *287*, 34979–34991.

(107) Forbes, C. R.; Sinha, S. K.; Ganguly, H. K.; Bai, S.; Yap, G. P.; Patel, S.; Zondlo, N. J. Insights into thiol-aromatic interactions: A stereoelectronic basis for S-H/π interactions. *Journal of the American Chemical Society* **2017**, *139*, 1842–1855.

(108) Shard, A. G.; Whittle, J. D.; Beck, A. J.; Brookes, P. N.; Bullett, N. A.; Talib, R. A.; Mistry, A.; Barton, D.; McArthur, S. L. A NEXAFS Examination of Unsaturation in Plasma Polymers of Allylamine and Propylamine. *The Journal of Physical Chemistry B* **2004**, *108*, 12472–12480.

(109) Beck, A. J.; Whittle, J. D.; Bullett, N. A.; Eves, P.; Mac Neil, S.; McArthur, S. L.; Shard, A. G. Plasma Co-Polymerisation of Two Strongly Interacting Monomers: Acrylic Acid and Allylamine. *Plasma Processes and Polymers* **2005**, *2*, 641–649.

(110) Khosravi, Z.; Kotula, S.; Lippitz, A.; Unger, W. E. S.; Klages, C.-P. IR- and NEXAFS-spectroscopic characterization of plasma-nitrogenated polyolefin surfaces. *Plasma Processes and Polymers* **2018**, *15*, 1700066.

(111) Zabicky, J. In *The Chemistry of functional groups*; Zabicky, J., Ed.; John Wiley & Sons, Ltd.: Chichester, UK, 1970; Vol. 2.

(112) Patai, S. In *The Chemistry of Amidines and Imidates*; Patai, S., Ed.; John Wiley & Sons, Ltd.: Chichester, UK, 1975.

(113) Makarova, A. A.; Grachova, E. V.; Neudachina, V. S.; Yashina, L. V.; Blüher, A.; Molodtsov, S. L.; Mertig, M.; Ehrlich, H.; Adamchuk, V. K.; Laubschat, C.; Vyalikh, D. V. Insight into Bio-metal Interface Formation in vacuo: Interplay of S-layer Protein with Copper and Iron. *Scientific Reports* **2015**, *5*, 8710.

(114) Rodriquez, C. F.; Cunje, A.; Shoeib, T.; Chu, I. K.; Hopkinson, A. C.; Siu, K. W. Solvent-assisted rearrangements between tautomers of protonated peptides. *Journal of Physical Chemistry A* **2000**, *104*, 5023–5028.

(115) Kamiya, K.; Boero, M.; Shiraishi, K.; Oshiyama, A. Enol-to-keto tautomerism of peptide groups. *Journal of Physical Chemistry B* **2006**, *110*, 4443–4450.

(116) Isaev, A. N. Keto-enol tautomerization of a peptide group with proton transfer through a water bridge. *Russian Journal of Physical Chemistry A* **2015**, *89*, 1360–1367.

(117) Hvidt, A.; Nielsen, S. O. Hydrogen Exchange in Proteins. *Advances in Protein Chemistry* **1966**, *21*, 287–386.

(118) Perrin, C. L. Proton Exchange in Amides: Surprises from Simple Systems. *Accounts of Chemical Research* **1989**, *22*, 268–275.

(119) Eriksson, M. A.; Härd, T.; Nilsson, L. On the pH dependence of amide proton exchange rates in proteins. *Biophysical Journal* **1995**, *69*, 329–339.

(120) Gorman, G. S.; Speir, J. P.; Turner, C. A.; Amster, I. J. Proton affinities of the 20 common α-amino acids. *Journal of the American Chemical Society* **1992**, *114*, 3986–3988.

(121) Cox, K. A.; Gaskell, S. J.; Morris, M.; Whiting, A. Role of the site of protonation in the low-energy decompositions of gas-phase peptide ions. *Journal of the American Society for Mass Spectrometry* **1996**, *7*, 522–531.

(122) Wu, R.; McMahon, T. B. Protonation Sites and Conformations of Peptides of Glycine ($Gly_{1-5}H^+$) by IRMPD Spectroscopy. *The Journal of Physical Chemistry B* **2009**, *113*, 8767–8775.

(123) Voss, J. M.; Fischer, K. C.; Garand, E. Revealing the structure of isolated peptides: IR-IR predissociation spectroscopy of protonated triglycine isomers. *Journal of Molecular Spectroscopy* **2018**,

(124) Zubavichus, Y.; Zharnikov, M.; Shaporenko, A.; Fuchs, O.; Weinhardt, L.; Heske, C.; Umbach, E.; Denlinger, J. D.; Grunze, M. Soft X-ray induced decomposition of phenylalanine and tyrosine: A comparative study. *Journal of Physical Chemistry A* **2004**, *108*, 4557–4565.

Graphical TOC Entry

Figure 9: For table of contents only