ARTICLE

Tracking Local and Global Structural Changes in a Protein by Cold Ion Spectroscopy

Andrei Zviagin, [a] Vladimir Kopysov, [a] Natalia S. Nagornova, [a] and Oleg V. Boyarkin *[a]

Characterization of native structures of proteins in the gas phase remains challenging due to the unpredictable conformational changes the molecules undergo during desolvation and ionization. We spectroscopically studied cryogenically cooled protonated protein ubiquitin and its microhydrated complexes prepared in the gas-phase in a range of charge states under different ionization conditions. The UV spectra appear vibrationally resolved for the unfolded protein, but becomes redshifted and smooth for the native-like structures of ubiquitin. This spectroscopic change results from the H-bonding of hydroxyl of Tyr to amide group of Glu-51 in the compact structures; the minimum length of this bond was estimated to ~1.7 Å. IR spectroscopy reflects the global structural change by observing redshifts of free NH/OH-stretch vibrational transitions. Evaporative cooling of microhydrated complexes of ubiquitin keeps the protein chilly during the ionization, enabling the native-like conformers with up to eight protons to survive in the gas phase.

Introduction

Characterization of 3D native structure of proteins remains in the focus of life science. Nuclear magnetic resonance, X-ray crystallography and cryogenic electron microscopy within their known limitations are capable of solving 3D structures of proteins in condense phase. Mass spectrometry (MS) -based methods can characterize molecules in the gas phase with unprecedent sensitivity, selectivity and on short timescales, which allows for high throughput screening of biological samples in a variety of -omics applications. Although MS, when combined with different methods of molecular fragmentation (e.g., collisional, electron transfer or photo), is the key approach for sequencing of proteins, the ability of MS itself to reveal 3D structures of this large biomolecules remains fairly limited. Moreover, the stress that proteins experience in MS measurements during the transfer from solution to the gas phase may crucially alter their native geometry. Along with other MS-based approaches, ion mobility (IM) technique allows for monitoring these changes and selecting specific conformers prior the MS investigations. IM determines collisional cross-section of ions, which characterizes their geometry globally only, without revealing any local structural changes however. Cold ion photofragmentation spectroscopy (CIS) is a recent addition to the toolbox of the gas-phase structural studies.[1,2] Cooling suppresses thermal broadening in UV/IR spectra,[3,4] making them vibrationally resolved, while MS detection of charged fragments adds many benefits of this technique. The resolved spectra allow for indirect but stringent structural characterization of biomolecules. Accurate gas-phase geometries of biomolecules as large as decapeptides and of their non-covalent complexes have been determined by validating quantum chemistry computations with CIS data.[5,6] The extension of spectroscopy to proteins is questionable however due to their large number of vibrations. UV and IR spectra of a small protein cytochrome c, for instance, exhibit no vibrationally resolved structure even under cryogenic cooling.[7] Here we report the gas-phase UV and IR spectra of a small protein ubiquitin (Ubi), as well as of its microhydrated complexes, produced in a range of protonated charge states, from different solutions and under different conditions of electrospray ionization. Referring to the known global geometries of Ubi, evaluated in IM studies, the spectra have been related to some local and global structural changes in the protein.

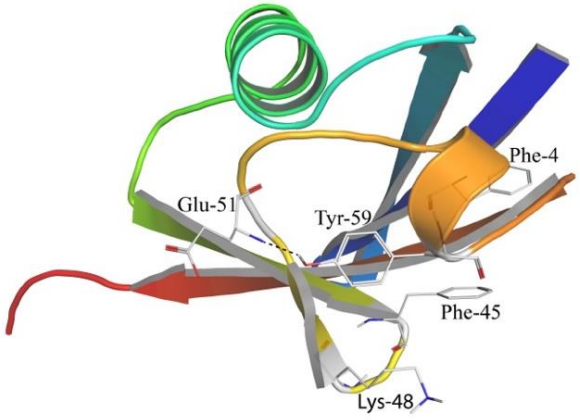

Fig. 1 3D back-bone structure of ubiquitin measured by X-ray diffraction (1UBQ).[8] Lys-48, Glu-51 and aromatic residues are shown explicitly. The dashed line indicates the hydrogen bond between the hydroxyl of Tyr-59 and the amide group of Glu-51.

[a] *Laboratoire de Chimie Physique Moléculaire, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland.*
** Email: oleg.boiarkin@epfl.ch*

† Electronic Supplementary Information (ESI) available: MS spectra of the produced Ubi ions; detailed UV spectra of Phe absorption; redshifts of UV band origins in non-covalent complexes of phenol and in ubiquitin; UV spectrum of Ubi^{5+}. See DOI: 10.1039/x0xx00000x

Ubiquitin is a highly conserved eukaryotic protein, which is involved to the processes of intracellular protein degradation. It contains 76 residues, including one tyrosine and two phenylalanine aromatic residues (Fig. 1). It is known to be highly stable in a wide range of pH and temperatures due to a rich network of hydrogen bonds and a large hydrophobic core. The secondary structure of Ubi includes one large and one short α-helices, a mixed β-sheet and a disordered C-terminus domain.[9] The tertiary gas-phase structure of this benchmark protein in different charge states and under different conditions of desolvation-ionization was extensively studied by H/D exchange,[10-13] proton transfer reactions[14-16] and, particularly, by IM-MS methods.[17-23] Overall, most of these studies led to a few main consistent conclusions. The global gas-phase tertiary structure of ubiquitin differs for different charge states and it is strongly influenced by the composition of the solution and by the conditions of the electrospray ionization. The typically "harsh" conditions of ESI that maximize concentration of ions in the gas phase imply their significant collisional heating. Under such conditions only the proteins in low charge states conserve their folded, solution-like, structures. Upon increasing the number of protons, these compact geometries, progressively, partially and then fully unfold to the extended conformers. This trend is explained by interplay of the intramolecular non-covalent bonds and the repulsive electrostatic force, which increases with the number of protons. Ubiquitin in the charge states with $n \leq 6$ protons is compact in the gas phase and closely resembles the folded native structure of the protein in solution.[20] The ions with $n \geq 12$ can only be observed as the extended conformers, while Ubi with n=7-11 exhibits broad transient distributions with compact, partially and fully unfolded conformers. "Gentle" ESI conditions minimize collisional heating of the ions and hinder the unfolding. Under such conditions, the ions with n=7 retain their native compact structures, although already for n=8 the proteins become partially unfolded and only a fraction of them remains compact.[20] In opposite, the "harsh" ESI and denaturing solutions (e.g., methanol/water) assist in the partial unfolding: the charge distribution shifts toward high states due to a facile access to internal protonation sites, such that for the same charge state, the conformational distribution shifts toward (partially) unfolded structures.[16,24] This rich suite of global structural information will be used herein for calibrating spectroscopic data obtained below for Ubi under similar conditions.

EXPERIMENTAL SECTION

Our cold ion spectrometer has been described in detail elsewhere.[2,25,26] Briefly, multiply protonated ubiquitin ions are brought to the gas phase directly from solution and transferred into the low-pressure region of the spectrometer by an orthogonal, double ion-funnel (DIF) nano electrospray ionization (ESI) source or by an inline triple skimmer "super gentle" nano-ESI for the experiments with the microhydrated protein. The latter source allows for adjustment of ESI conditions (e.g., by increasing the inter skimmer potentials) from the "super gentle" to gentle and harsh, but produces lower ion current than the former source. The settings of the triple skimmer source for the gentle and harsh conditions were determined by obtaining UV spectra that are nearly identical to the ones measured with DIF ESI under the same category of the conditions (gentle or harsh).

The electrosprayed ions of interest in a certain charge state are pre-selected by a quadrupole mass filter and guided into a T=6 K octupole ion trap,[27] where they get trapped and cooled in collisions with He buffer gas. Once cooled, the ions are fragmented by a UV laser pulse of ns duration for electronic spectroscopy or by IR and UV laser pulses for vibrational spectroscopy. The ions are then released from the trap and analyzed by a quadrupole mass spectrometer. The loading of the cold ion trap is performed at 20 Hz repetition rate, while the laser sources work at 10 Hz. The appearing unresolved mass peaks of the UV photofragments correspond to the mass-loss of 45-100 Da by the bare protein. An electronic photofragmentation spectrum is recorded by measuring the dissociation yield of all these fragment ions as a function of UV wavelength in the loading cycles with the laser pulses, while the number of parent ions is measured in the alternative, "laser-free" cycles. The vibrational "gain" spectra of the bare protein are measured by fixing UV laser wavelength at the red from the electronic band origin, while monitoring the appearance of photofragments as a function of the wavelength of the preceding IR laser pulse. The photodissociation vibrational spectra of the microhydrated proteins were measured by detecting the complexes that loose 3-5 water molecules upon absorption of IR photons. Each data point in a spectrum is averaged over 10 measurements and normalized on the average signal of parent ions and the average UV energy of the respective 10 laser shots. We typically measure each spectrum three times and, finally, average them.

Ubiquitin from bovine erythrocytes (≥98 % purity) was purchased and used without further purification. A stock solution of ubiquitin was prepared by dissolving it in pure water. For electrospray ionization, the aqueous stock solution was further diluted to yield a 50 μM solution either in water or in a water/methanol (50/50) mixture with addition of 1% of acetic acid. Water, methanol, and acetic acid are of LC-MS grade.

The use of the nano electrospray capillaries with small diameter of opening (the typical flow rate as small as 80-120 nL/min) is crucial for preventing formation of Ubi dimers at the used concentration of < 50 μM.[28] The lack of dimers was confirmed by lack in mass spectra of any noticeable peaks that correspond to the ions with non-integer m/z; these dimers could arise from clustering of highly abundant ions in n and (n+1) charge states (Fig S1).

RESULTS AND DISCUSSION

UV spectroscopy of bare Ubi

Figure 2 shows a panoramic UV photofragmentation spectra of the +6 to +12 charge states of cold ubiquitin in the spectral range that overlaps the electronic band origins of both Tyr and Phe aromatic residues.[4] Except for the +6 and +7 states, the ions were electrosprayed from a denaturing solution (methanol/water/acetic acid, 50/50/1). Ubi in all the charge states ex-

perienced "harsh" conditions of ESI, which were tuned to maximize the number of ions. The spectra are dominated by Tyr absorption, which is 5-6 time stronger than that of Phe.[29] For the number of protons *n*=6, the absorption extends to the red from the band origin of Tyr by as much as ~1200 cm^{-1} and, apart from a few "bumps" and one broad peak, appears vibrationally unresolved. The lack of prominent sharp peaks is in line with the general expectation that no/little vibrationally-resolved structure can be observed for a large molecule like Ubi. IM studies firmly assign Ubi^{6+} in the gas phase to compact, native-like, structures.[20,30]

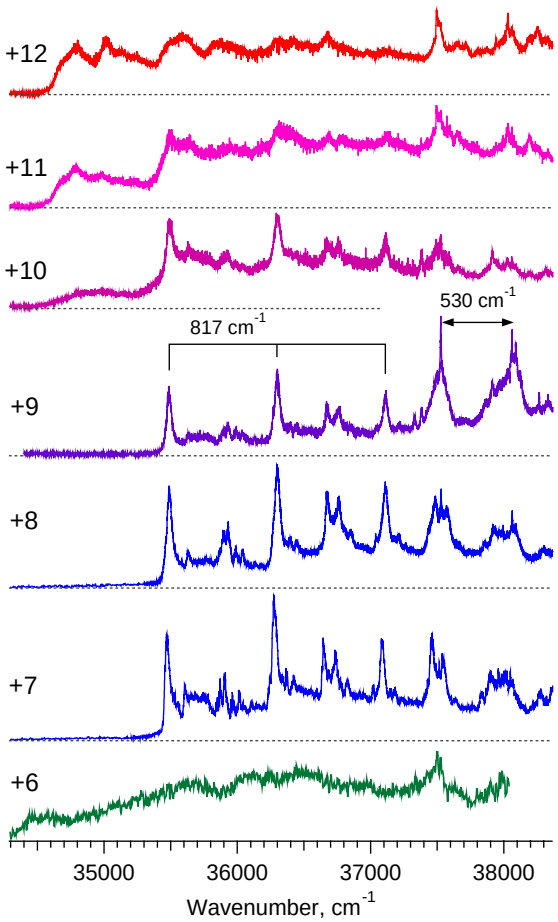

Fig. 2 UV photofragmentation spectra of protonated $[Ubi+n\cdot H]^{n+}$, *n*=6-12, recorded for the ions produced under harsh settings of the ESI source from methanol/water solution for *n*=8-12 and from water for *n*=6-7. The numbers on the left indicate the charge states of Ubi; the numbers above the spectrum of the +9 state indicate the spacing in the vibrational progressions that are specific to Tyr and Phe.

We, thus, can tentatively correlate a broadband spectrum with folded conformers of Ubi^{6+}. In contrast to the spectrum of the protein with *n*=6, a striking feature in all other spectra in figure 2 is the clearly resolved vibrational bands, which are the most prominent for the *n*=7-9 states. The largest well-resolved peaks in these spectra are of ~40 cm^{-1} wide. They exhibit a Franck-Condon progression with the characteristic for the Tyr aromatic ring frequency of the "breathing" mode (~810 cm^{-1}).[4] The maximum of the first absorption band (35469.4 cm^{-1}) in the spectra of the +7 to +9 states of Ubi appears to be very close to the band origin of the neutral Tyr.[31,32] The position of the electronic band origin in Tyr is known to be very sensitive to local environment of its side chain. Non-covalent interactions with hydroxyl of the residue, as well as the proton–π and the weaker H–π couplings with the aromatic ring can induce large spectral shifts of the UV absorption onset.[33-35] The lack of a spectral shift (*n*=7-9) implies the lack of substantial non-covalent interactions of the Tyr side chain with its environment. This local isolation of the side chain is consistent with the IM observations of, mainly, the partially unfolded structures for +7 to +9 charge states of Ubi in the gas phase.[20,30]

The spectral resolution progressively degrades and the onset of the absorption redshifts for the +10 to +12 charge states. Different from the smooth spectrum of the +6 state, the arising redshifted bands remain, however, structured. For *n*=10-12, Ubi in the gas phase was found to adopt, mostly, the unfolded structures. We tentatively attribute the redshifted bands to those conformers, in which the additional charges occupy the protonation sites that are close to the side chain of Tyr. This may lead to a long-range proton-π interaction, which is known to shift the UV band origin of Tyr to the red.[36,37] The conformational families that differ by occupation of the available in the vicinity of Tyr side chain protonation sites should exhibit distinct onsets of UV absorption. The overlap of these similar but shifted absorptions of different intensities may generate the stepwise spectra observed for *n*=10-12.

In addition to the large peaks of Tyr, a few well-reproducible spikes appear on top of the Tyr absorption bands above ~37500 cm^{-1} (Figs. 2, S2, S3). Regarding the size of the studied molecule, the width of some of these spikes is amazingly small, below 2 cm^{-1}. The most intense of them appears in the spectrum for *n*=9 at 37527 cm^{-1}. This is very close to the electronic band origin of Phe residue in, for instance, helical peptides, where aromatic side chain of Phe is free of any non-covalent interactions.[38] The sharp peaks are spaced by ~530 cm^{-1}, which is a highly conservative frequency of the in-plane bending of the Phe ring.[39] Based on these observations, we firmly assign these peaks to the absorption by Phe residues of the protein. The Phe peaks of different intensities remain sharp and can be tracked in the spectra of all the charged states. We cannot distinguish however whether these electronic transitions originate from the two different Phe residues of ubiquitin or from the same Phe but residing in different conformers of the protein. The lack of a shift for *n*=6 and 7 (e.g., Fig. S2 red traces) implies no interactions of the Phe aromatic rings with the environment in these charge states. A small shift with no broadening of the peaks for *n*=8 and 9 (Fig. S3) suggests some weak interactions of the rings and therefore a (slightly) different environment of the chromophores in these states compared with the *n*=6 and 7 states. The detected small spectral redshifts, splitting and broadening of the sharp peaks for *n*=10-12 (Fig. S3) indicate certain changes in the local environment of the Phe residues in Ubi in these charge states too.

The observation of the vibronic structure in the spectra of ubiquitin is interesting on its own. One might expect that any structure in a UV spectrum of such a large molecule as a protein

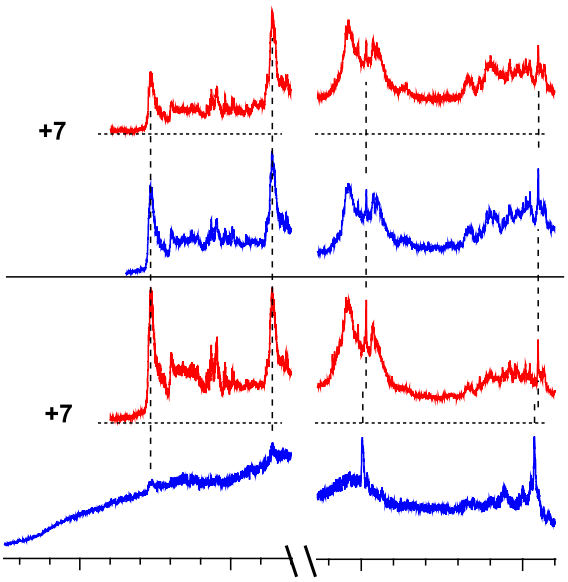

Fig. 3 Parts of UV photofragmentation spectra of $[Ub+7H]^{7+}$ recorded for the protein electrosprayed (upper panel) from water/methanol/acetic acid (50/50/1) solution and (lower panel) from water/acetic acid (100/1) solution. Blue and red spectra correspond to the gentle and harsh ESI conditions, respectively. Vertical dashed lines show the alignment/shifts of the characteristic transitions.

will be completely washed out due to an inhomogeneous spectral broadening, which arises from a large number of low-frequency vibrations. The UV-active vibrations are to be, however, associated with the changes in the protein geometry occurring upon electronic excitation. Because of the local character of electronic excitation, the major UV-induced structural changes occur around the Tyr chromophore, such that the Franck-Condon active modes are also localized in the proximity of the chromophore or are coupled to it by H-bonds. The limited number of such modes does not yet grant the observed vibrational resolution. Thermal congestion, conformational heterogeneity and lifetime broadening of the excited state may also wash out the vibrational structure. Cooling ions to cryogenic temperatures is proven to suppress the thermal spectral broadening.[40] The observation of the sharp peaks with < 2 cm^{-1} (figures 2, 3, S2 and S3) indicates that Ubi is, indeed, cold in our experiments, although we cannot quantify its vibrational temperature. Consistently with the concept of UV-active modes, the spectra of the cold protein that contains as many as 76 residues overall look for *n*=7-9 somewhat similar to the spectra of, for instance, cold gas-phase enkephalins,[37,41] which contain only five residues.

Two more observations further support the revealed above spectroscopy-structure correlation. Figure 3 shows four UV spectra of the +7-charge state, which were recorded under four different experimental conditions. The solution for ESI was either water or 50/50 water/methanol mixture with 1% of acetic acid, and the ion source conditions were either gentle (minimum RF amplitudes of the ion funnels) or harsh (maximum RF amplitudes of the ion funnels). The spectra of $[Ubi+7H]^{7+}$ appear almost identical, when ubiquitin is electrosprayed from the water/methanol solution or else from water but under harsh ion source conditions. Minimization of collisional activation in the ion source (gentle conditions) results however in a dramatic change of the UV spectrum of ubiquitin, dissolved in water. The spectrum extends to the red by more than a thousand cm^{-1} and drastically broadens. It resembles the structureless spectrum of the *n*=6 state in figure 2, although still exhibits a few partially-resolved Tyr absorption peaks. In contrast, Phe transitions broaden only twice, with the width of 5-6 cm^{-1} (figure S2), and red-shift by 13 cm^{-1} only. Such changes are, again, fully consistent with IM studies, which detect compact forms of the protein in the *n*=7 state only for "gentle" ESI from aqueous solutions.[20] The changes in absorption by Tyr clearly manifest crucial changes in its local environment; the shift and broadening of the Phe peaks suggest certain changes in their environment too.

Figure 4 compares the UV spectra measured under the harsh (red traces) and gentle (blue traces) ESI conditions for the +6 to +8 charge states of Ubi dissolved in water. The two UV spectra for *n*=6 look quite similar and resemble the broad spectrum of Ubi^{7+} produced by gentle ESI, which implies a compact structure of Ubi^{6+} under both conditions. Regardless of the conditions, the UV spectra of the +8 state exhibit the same vibrationally resolved structure, although in the "gentle" spectrum the peaks sit on a broad substrate that extends to the red by ~1100 cm^{-1}. Regarding the IM studies,[20] such superposition of the resolved and broadband spectral components in the "gentle" spectrum indicates the presence of both compact and extended structures of the protein. In contrast to Tyr, the sharp peaks of Phe do not exhibit a dramatic change (Fig. S2). The gentle production of Ubi nearly doubles the width of these peaks and redshifts them by ~12 cm^{-1} for the +7 and +6 states, but does not influence the peaks for the +8 states. This implies a change of local environment for +6 and +7 states upon collisional activation, such that Phe side chain becomes free of any non-covalent interactions, while for Ubi^{8+} the interactions (and environment) remain the same.

UV spectroscopy of hydrated Ubi

Tuning the triple skimmer ESI source to further minimize collisional heating enables retaining water molecules on ubiquitin in the gas phase. The $Ubi^{7+}(H_2O)_{10}$ complex produced under such "super gentle" ionization conditions exhibits in the region of Tyr absorption a broad spectrum (Fig. 4) that looks similar to the one measured for the +6 state of the gently electrosprayed bare protein but becomes smoother, with no features at all. This similarity prompts us to suggest that the microhydrated Ubi^{7+} appears in the gas phase in its folded, native-like, structure, at least, around Tyr chromophore. Retaining 10 or even only 5 water molecules on Ubi^{8+} also fully removes the resolved transitions, making the UV spectra of the Tyr residue smooth and eventually identical to the spectrum of the $Ubi^{7+}(H_2O)_{10}$ complex. No sharp peaks have been detected at the achieved level of S/N ratio in the Phe spectral region for all the studied complexes. A broadband structure centered on the band origin of

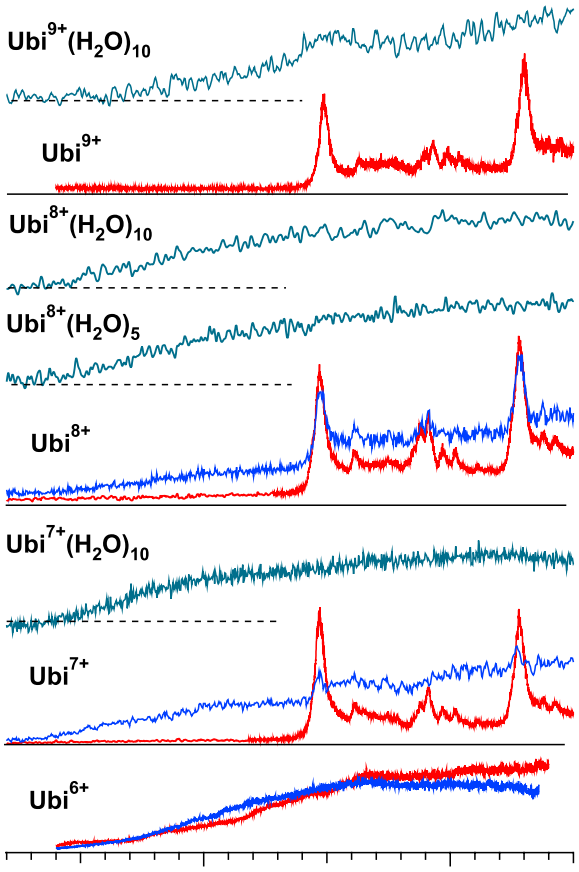

Fig. 4 UV photofragmentation spectra recorded around the UV band origin of Tyr for $[Ub+nH]^{n+}$ (*n*=6-9 from bottom to top; blue and red traces) and of the protein (*n*=7 -9), microhydrated by 5 and 10 water molecules (grey traces; as labeled). For each charge state, the blue and red spectra were measured under the gentle and harsh conditions of ESI, respectively, for ubiquitin in water/acetic acid (100/1) solution. The microhydrated ubiquitin was produced from the same solution by the "super gentle" ESI in the triple skimmer ion source.

Tyr re-appears, however, for microhydrated Ubi^{9+}. The spectrum also spans to the red less than for the *n*=6-8 states. These differences indicate certain structural changes around the chromophore, compared to the lower charge states. The microhydration is not capable anymore of retaining the intact folded structure of the proteins around the chromophore, such that a fraction of them resides in partially unfolded *n*=9 states.

We rule out a direct influence of water molecules on the UV absorption by Tyr in the complexes with only 5 and 10 water molecules, since no waters close to the Tyr residue was detected in NMR[42] and X-ray[8] studies for a partially hydrated (folded) Ubi. NMR studies explicitly observed H-binding of the Tyr side chain to the nearby Glu-51 residue even in the presence of a hydration shell of the protein.[42] These facts and our observations suggest that even for *n*= 8 and, largely, for *n*=9 the protein in the gas phase can retain its compact, solution-like, structures, provided the protein was "super gently" electrosprayed from an aqueous solution. It is unlikely that folded structures of the large protein can be preserved just by five remaining water molecules. We suggest that the retained water molecules allow adiabatic evaporative cooling of ubiquitin during ESI, thus protecting the protein from an excessive vibrational heating required for its unfolding in the gas phase.[43,44] The computations estimate that the evaporation may reduce the internal temperature of ubiquitin by as much as 60-70 K.[45] The remaining waters are to be those that have the highest binding energy to the protein, further contributing to its structural stabilization. Only the fully compact structures of Ubi in the *n*=8 and 9 states have never been observed in ion mobility studies, which revealed mixed distributions of compact and partially unfolded geometries of the protein.[30] This difference implies that the "super gentle" ESI conditions created in our experiments are the key to protect the native structure of Ubi with *n*=8 and, partially, of *n*=9 in the gas phase.

The broadening of Tyr absorption in compact structures of the cold protein can be rationally explained by (i) short lifetime of the excited electronic state of Tyr, (ii) high conformational heterogeneity of native-like structures of ubiquitin. Lifetime of the S_1 state of Tyr, potentially, could be substantially shortened due to hydrogen or charge transfer through the H-bond to Glu-51. To result in a ~1000 cm^{-1} homogeneous broadening, the lifetime is to be on the scale of ~5 fs. This would be much faster than the ~370 fs timescale of the electron transfer through the H-bond in the Phenol-$(NH_3)_5$ noncovalent complex, which is the fastest processes ever observed for S_1 state of Tyr or phenol.[46] We thus consider highly unlikely that a lifetime broadening alone can explain the observed spectral broadening in the spectra of compact Ubi. NMR studies suggest tens of conformers of Ubi in condense phase.[47,48] The position of UV band origin of Tyr is highly sensitive to the strength of non-covalent bonds.[33,34] Variation of this strength among numerous conformers would lead to different shifts of UV band origin in the ensemble of the folded structures. We propose that, it is the combination of the lifetime broadening and the conformational heterogeneity that results in the observed smooth broadening in the spectra of the compact structures for *n*=6-8. Two types of non-covalent interactions can lead to large UV spectral shifts: H-binding of the Tyr hydroxyl to amide group of Glu-51 and proton-π interaction of the aromatic ring of Tyr. The shift of the electronic band origin to the red in the spectra points to the binding of the hydrogen but not oxygen of the Tyr hydroxyl group to the amide group of Glu-51.[33] The available data for redshifts of the UV band origin in function of the length of H-bond between hydroxyl oxygen in phenylalanine and the amide group of different species (Fig. S4) suggest that, the minimum length of OH(Tyr)···NH(Glu-51) bond should be as short as ~1.7 Å. This however contradicts to the available NMR structures of the protein, in most of which the oxygen of OH is closer to the hydrogen of the Glu-51 amide (PDB structures 1D3Z (10 models),[49] 2MSG[48]). It is worth noting that these structures are not charge-specific and all differ in the position of the atoms involved to the binding. The NMR and X-ray structures do not favor the proton-π interaction too. For *n*=6-8, the nearest proton resides on the side chain of the lycine-48 residue[50] at the distance of 6-7 Å;[8,49] the occupancy of this site is unknown for *n*=9. Such ring-to-charge spacing seems to be too large to account for ~1000 cm^{-1} redshifts. Moreover, the

large redshift and broadening remains in the UV spectrum of the *n*=5 state, in which the lycine-48 side chain was found not to be protonated in the gas phase (Fig. S5). We thus rule out this interaction and suggest that the observed spectral broadening is due to OH(Tyr)···NH(Glu-51) binding. Full/partial desolvation of the protein certainly reshapes its structure and may lead to some modifications of the network of H-bonds, including the binding of Tyr hydroxyl. Vibrational spectroscopy data below provide a direct support for this suggestion.

IR spectroscopy

Vibrational spectroscopy, which reflects the global structure of a protein, provides a direct evidence of the transition from unfolded to folded structures and allows for unambiguous assignment of the Tyr-Glu hydrogen bond. Figure 5 compares the IR spectrum of Ubi^{7+} electrosprayed from water solution under the harsh conditions with the spectrum of the $Ubi^{7+}(H_2O)_5$ complex. The IR spectrum of the bare protein contains a number of sharp resolved peaks in the region of 3430-3590 cm^{-1} and a distinct peak at 3647 cm^{-1}. The sharpness of these peaks implies that they originate from non- or weakly-disturbed vibrations. Bearing this in mind and based on the known characteristic frequencies of molecular transitions in peptides,[2] the observed sharp peaks in the regions of 3430-3500 cm^{-1} and 3500-3590 cm^{-1} can be assigned to free NH and carboxylic OH stretches, respectively. The sharp peak at 3647 cm^{-1} perfectly matches to the free phenolic OH stretch of Tyr[4,36], while carboxylic OH stretches lie below 3600 cm^{-1} (Table 2 in ref[2]) and the free aliphatic OH stretches of serine and threonine are somewhat higher in frequency (3665-3690 cm^{-1}).[51,52] We therefore assigned the transition at 3647 cm^{-1} to the free hydroxyl of Tyr residue. In the native folded conformers of Ubi the hydroxyl is not free, but coupled to Glu-51.[50] Appearance of this characteristic transition is a direct evidence of the protein unfolding in the region of Tyr. Disappearance of this transition in the IR spectrum of the compact structure implies that it redshifts and becomes buried in the manifold of transitions of H-bound OH/NH groups. This allows us to estimate the anharmonic shift of the OH-stretch of the Tyr hydroxyl as, at least, ~170 cm^{-1}. Such large redshift is possible only, if the hydrogen but not oxygen of the Tyr hydroxyl is bound (to the nitrogen of the Glu-51 amide). The firm assignment of this characteristic H-bond explains the UV spectral broadening and shifts observed for the compact states of the protein.

The observation of a large number of free NH/OH stretch vibrations in the IR spectrum of the *n*=7 state of the "harshly" electrosprayed Ubi manifests a break of many intramolecular H-bonds present in native conformers of the protein.[53] Consistently with UV spectroscopy and IM studies, this implies an unfolded global structure of Ubi^{7+} ions, brought to the gas phase under the "harsh" ESI conditions. The IR spectrum changes drastically, when Ubi^{7+} is produced under the "super gentle" conditions that allow retaining of five waters on it. All the sharp transitions shift to the red and become unresolved. This change unambiguously indicates an involvement of the previously free NH/OH groups to H-bonds. Regarding the large number of these free groups, it is highly unlikely that five waters only can take up

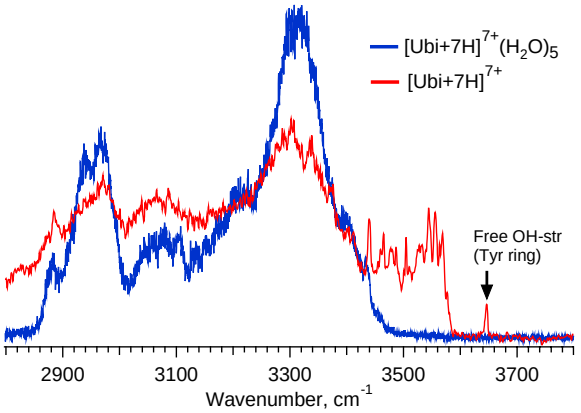

Fig. 5 IRPD spectra of $[Ubi+7H]^{7+}$ electrosprayed from water under the harsh ion source conditions (red) and of the same protein microhydrated by five water molecules (blue).

all these vacancies. We thus conclude that, the observed spectral change reflects the retaining of the network of intramolecular H-bonds in the gas-phase protein, which implies its globally compact structure.

Conclusions

To sum up, this study demonstrates that, even for biomolecules as large as proteins, cold ion spectroscopy can provide a valuable structural information. Complementary to other methods, CIS allows for tracking of global (in IR) but also of local (in UV) structural changes in the gas phase. We recorded UV and IR gas-phase spectra of cryogenically cooled protonated protein ubiquitin and its microsolvated complexes, electrosprayed from aqueous and denaturing solutions under different conditions of collisional heating in the ion source and in different charged states from +5 to +12. Some of the UV spectra exhibit vibrational resolution in the spectral regions of absorption by Tyr and Phe residues. Being calibrated by the earlier IM studies, these resolved spectroscopic fingerprints were assigned to partially/fully unfolded conformers of Ubi. These conformers dominate for Ubi in +9 to +12 charged states, but also for the +7 and +8 states, provided the ions are produced from denaturing solutions or undergo a substantial (e.g., "standard") heating (harsh conditions) during the ESI process. The unresolved spectra correspond to the protein in the native-like, folded, states, in which hydroxyl of Tyr residue forms H-bond with the amide group of Glu-51 residue. Regardless of the ESI conditions, the vast majority of the conformers exhibit the spectroscopic signature of the compact, native-like, structure for ubiquitin with *n*=6. Ubi^{7+} retains its native-like structure in the gas phase only if produced from an aqueous solution under "gentle" ESI conditions that prevent substantial heating of the protein. Consistently, IR spectroscopy reveals the break of many intramolecular H-bonds, including the one with Tyr hydroxyl, under the "harsh" conditions of ESI, but disappearance of the free NH/OH transitions in the "chilly" microhydrated protein. The increasing electrostatic repulsion partially unfolds Ubi^{8+} even under

mild heating under the "gentle" ESI conditions. Retaining a few water molecules on Ubi^{8+} and Ubi^{9+} protects these ions from heating by evaporative cooling of waters, which allows the folded structures of the protein to survive in the gas phase for long time. Our observation of the OH−NH but not HO−HN hydrogen binding between Tyr-59 and Glu-51 residues provides an example evidence that, the compact gas-phase structures are not really "native", but rather the native-like ones only.

The observed crucial role of water molecules for retaining compact structure of ubiquitin in the gas phase is practically relevant to the field of "native mass spectrometry".[54] Retaining native-like structure of large biomolecules (e.g., proteins, complexes of proteins, etc.) in the gas phase requires softening of the ESI conditions. This need for a gentle ionization contradicts however to the desire of a high mass resolution, which favors full desolvation of ions. Our study demonstrates that the appearance of microhydrated ions in the gas phase ensures their low internal energy: evaporative cooling keeps the ions "chilly" for the best protection of the solution-phase structures of biomolecules. Native MS should take this fact into account; the ionization and desolvation procedures, perhaps, can be separated.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

The authors thank the Swiss National Science Foundation (grants 200020_129649, 200021_146389 and 200020_204072) for supporting this work.

Notes and references

1 T. R. Rizzo, J. A. Stearns and O. V. Boyarkin, *Int. Rev. Phys. Chem.*, 2009, **28**, 481 - 515.
2 O. V. Boyarkin, *Int. Rev. Phys. Chem.*, 2018, **37**, 559-606.
3 S. R. Mercier, O. V. Boyarkin, A. Kamariotis, M. Guglielmi, I. Tavernelli, M. Cascella, U. Rothlisberger and T. R. Rizzo, *J. Am. Chem. Soc.*, 2006, **128**, 16938-16943.
4 J. A. Stearns, S. Mercier, C. Seaiby, M. Guidi, O. V. Boyarkin and T. R. Rizzo, *J. Am. Chem. Soc.*, 2007, **129**, 11814-11820.
5 N. S. Nagornova, M. Guglielmi, M. Doemer, I. Tavernelli, U. Rothlisberger, T. R. Rizzo and O. V. Boyarkin, *Angew. Chem. Int. Ed.*, 2011, **50**, 5383-5386.
6 T. K. Roy, N. S. Nagornova, O. V. Boyarkin and R. B. Gerber, *J. Phys. Chem. A*, 2017, **121**, 9401-9408.
7 N. S. Nagornova, T. R. Rizzo and O. V. Boyarkin, *Angew. Chem. Int. Ed.*, 2013, **52**, 6002-6005.
8 S. Vijay-kumar, C. E. Bugg and W. J. Cook, *J. Mol. Biol.*, 1987, **194**, 531-544.
9 S. Vijay-Kumar, C. E. Bugg, K. D. Wilkinson and W. J. Cook, *Proc. Natl. Acad. Sci.*, 1985, **82**, 3582-3585.
10 D. Suckau, Y. Shi, S. C. Beu, M. W. Senko, J. P. Quinn, F. Wampler and F. W. McLafferty, *Proc. Natl. Acad. Sci.*, 1993, **90**, 790-793.
11 S. E. Evans, N. Lueck and E. M. Marzluff, *Int. J. Mass Spectrom.*, 2003, **222**, 175-187.
12 O. Geller and C. Lifshitz, *J. Phys. Chem. A*, 2005, **109**, 2217-2222.
13 M. A. Freitas, C. L. Hendrickson, M. R. Emmett and A. G. Marshall, *Int. J. Mass Spectrom.*, 1999, **185-187**, 565-575.
14 P. D. Schnier, D. S. Gross and E. R. Williams, *J. Am. Soc. Mass Spectrom.*, 1995, **6**, 1086-1097.
15 X. Zhang and C. J. Cassady, *J. Am. Soc. Mass Spectrom.*, 1996, **7**, 1211-1218.
16 S. J. Valentine, A. E. Counterman and D. E. Clemmer, *J. Am. Soc. Mass Spectrom.*, 1997, **8**, 954-961.
17 K. J. Laszlo, E. B. Munger and M. F. Bush, *J. Am. Chem. Soc.*, 2016, **138**, 9581-9588.
18 S. Myung, E. R. Badman, Y. J. Lee and D. E. Clemmer, *J. Phys. Chem. A*, 2002, **106**, 9976-9982.
19 T. J. El-Baba, D. W. Woodall, S. A. Raab, D. R. Fuller, A. Laganowsky, D. H. Russell and D. E. Clemmer, *J. Am. Chem. Soc.*, 2017, **139**, 6306-6309.
20 T. Wyttenbach and M. T. Bowers, *J. Phys. Chem. B*, 2011, **115**, 12266-12275.
21 B. C. Bohrer, N. Atlasevich and D. E. Clemmer, *J. Phys. Chem. B*, 2011, **115**, 4509-4515.
22 S. J. Valentine, A. E. Counterman and D. E. Clemmer, *J. Am. Soc. Mass Spectrom.*, 1997, **8**, 954-961.
23 H. Shi and D. E. Clemmer, *J. Phys. Chem. B*, 2014, **118**, 3498-3506.
24 J. Li, J. A. Taraszka, A. E. Counterman and D. E. Clemmer, *Int. J. Mass Spectrom.*, 1999, **185-187**, 37-47.
25 A. Y. Pereverzev, I. Szabó, V. N. Kopysov, E. Rosta and O. V. Boyarkin, *Phys. Chem. Chem. Phys.*, 2019, **21**, 22700-22703.
26 E. Saparbaev, V. Aladinskaia, A. Zviagin and O. V. Boyarkin, *J. Phys. Chem. Lett.*, 2021, **12**, 907-911.
27 O. V. Boyarkin and V. Kopysov, *Rev. Sci. Instrum.*, 2014, **85**.
28 Z. Xia and E. R. Williams, *Analyst*, 2019, **144**, 237-248.
29 A. Y. Pereverzev, V. Kopysov and O. V. Boyarkin, *Angew. Chem. Int. Ed.*, 2017, **56**, 15639-15643.
30 S. L. Koeniger and D. E. Clemmer, *J. Am. Soc. Mass Spectrom.*, 2007, **18**, 322-331.
31 S. J. Martinez, J. C. Alfano and D. H. Levy, *J. Mol. Spectrosc.*, 1992, **156**, 421-430.
32 R. Cohen, B. Brauer, E. Nir, L. Grace and M. S. de Vries, *J. Phys. Chem. A*, 2000, **104**, 6351-6355.
33 E. Saparbaev, V. Kopysov, R. Yamaletdinov, A. Pereverzev and O. V. Boyarkin, *Angew. Chem. Int. Ed.*, 2019, **58**, 7346-7350.
34 E. Saparbaev, V. Aladinskaia, R. Yamaletdinov, A. Y. Pereverzev and O. V. Boyarkin, *J. Phys. Chem. Lett.*, 2020, **11**, 3327-3331.
35 E. Saparbaev, R. Yamaletdinov and O. V. Boyarkin, *Anal. Chem.*, 2021, **93**, 12822-12826.
36 J. A. Stearns, M. Guidi, O. V. Boyarkin and T. R. Rizzo, *J. Chem. Phys.*, 2007, **127**, 154322-154327.
37 V. Kopysov and O. V. Boyarkin, *Angew. Chem. Int. Ed.*, 2016, **55**, 689-692.
38 J. A. Stearns, O. V. Boyarkin and T. R. Rizzo, *J. Am. Chem. Soc.*, 2007, **129**, 13820-13821.
39 G. Féraud, M. Broquier, C. Dedonder, C. Jouvet, G. Grégoire and S. Soorkia, *J. Phys. Chem. A*, 2015, **119**, 5914-5924.
40 O. V. Boyarkin, S. R. Mercier, A. Kamariotis and T. R. Rizzo, *J. Am. Chem. Soc.*, 2006, **128**, 2816-2817.
41 N. L. Burke, J. G. Redwine, J. C. Dean, S. A. McLuckey and T. S. Zwier, *Int. J. Mass Spectrom.*, 2015, **378**, 196-205.
42 N. V. Nucci, M. S. Pometun and A. J. Wand, *Nat. Struct. Mol. Biol.*, 2011, **18**, 245-249.
43 R. Signorell and A. Bertram, *Phys. Chem. Chem. Phys.*, 2009, **11**, 7759.
44 S. P. Thirumuruganandham and H. M. Urbassek, *Biochem. Res. Int.*, 2010, **2010**, 213936.
45 A. Patriksson, E. Marklund and D. van der Spoel, *Biochemistry*, 2007, **46**, 933-945.
46 M. Miyazaki, N. Washio and M. Fujii, *Chem. Phys.*, 2018, **515**, 580-585.

47 F. Lange Oliver, N.-A. Lakomek, C. Farès, F. Schröder Gunnar, F. A. Walter Korvin, S. Becker, J. Meiler, H. Grubmüller, C. Griesinger and L. de Groot Bert, *Science*, 2008, **320**, 1471-1475.
48 H. K. Fasshuber, N.-A. Lakomek, B. Habenstein, A. Loquet, C. Shi, K. Giller, S. Wolff, S. Becker and A. Lange, *Protein Sci.*, 2015, **24**, 592-598.
49 G. Cornilescu, J. L. Marquardt, M. Ottiger and A. Bax, *J. Am. Chem. Soc.*, 1998, **120**, 6836-6837.
50 L. J. Morrison and J. S. Brodbelt, *Analyst*, 2016, **141**, 166-176.
51 R. Linder, K. Seefeld, A. Vavra and K. Kleinermanns, *Chem. Phys. Lett.*, 2008, **453**, 1-6.
52 V. Scutelnic, M. A. S. Perez, M. Marianski, S. Warnke, A. Gregor, U. Rothlisberger, M. T. Bowers, C. Baldauf, G. von Helden, T. R. Rizzo and J. Seo, *J. Am. Chem. Soc.*, 2018, **140**, 7554-7560.
53 F. Cordier and S. Grzesiek, *Biochemistry*, 2004, **43**, 11295-11301.
54 S. Tamara, M. A. den Boer and A. J. R. Heck, *Chem. Rev.*, 2021, DOI: 10.1021/acs.chemrev.1c00212.