**Cold Ion Spectroscopy for Structural Identifications of Biomolecules**

Oleg V. Boyarkin[†]

*Laboratoire de Chimie Physique Moléculaire, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland.*

[†] Electronic address : oleg.boiarkin@epfl.ch

**Abstract**

Over last decade the spectroscopy of cryogenically cold ions isolated in the gas phase has been developed to a new tool for structural elucidations of biological molecules. Cooling allows for vibrational resolution in UV and IR spectra of small to midsize peptides, enabling different multi-laser techniques of conformer-specific spectroscopy. We first briefly overview some fundamental and technical aspects of the cold-ion spectroscopy (CIS) approach and illustrate its application to protonated peptides, carbohydrates and to non-covalent complexes of these biomolecules. The challenges and limitations of CIS in view of its relevance to life-science studies are critically assessed. Finally, we discuss and illustrate some approaches of CIS to analytical identifications of biomolecules, in particular the recently developed method of 2D UV-MS fingerprinting, which combines CIS with high-resolution mass spectrometry.

**Contents**

## INTRODUCTION

Conformational/isomeric heterogeneity of biomolecules in vivo is a key factor that enables high diversity and dynamical flexibility of their biological functions. A change in three-dimensional structure of a protein in response to environmental conditions or biological signaling may alter its activity from live to death.[1] Determination of 3D structures of the conformers and revealing the mechanisms of their interconversion *in vivo* is therefore one of the key problems that have to be solved to enable an intelligent modeling of functionality of biomolecules on a fundamental, molecular level. Currently, none of the known methods alone, owing their limitations, is capable of unambiguous structural determinations for the variety of biomolecules. X-ray diffraction may determine 3D structures of crystallized biomolecules with high resolution, but the crystallization may drastically alter the 3D geometry of molecules they adopt in vivo.[2-3] This powerful method remains the main working "horse" for solving structures of large biomolecules, although crystal structures of only rigid molecules with low conformational heterogeneity may have direct relevance to their native conformations. The recently developed to the level of a laboratory technique method of cryogenic electron microscopy (cryo-EM) solves some of these problems.[4] It can be applied for structural determinations of large biomolecules, rapidly frozen to cryogenic temperatures in a thin layer of solvent. In addition to the problems of radiative damage and limited resolution, which are continuously improving, this method still suffers from a limited ability to resolve multiple conformers of biomolecules. Alone with cryo-EM, NMR spectroscopy in vitro is, eventually, the only proven method capable of providing direct structural constraints in the form of internuclear distances for biomolecules in their quasi-natural environments – in aqueous solutions or in hydrophobic solvents (e.g. for membrane proteins and peptides). In conjunction with "inexpensive" low-level methods of computational quantum chemistry these constraints allow for solving "native" structures of biomolecules as large as proteins.[5] One of the severe limitations of NMR is its low ability to resolve multiple conformational structures that may interconvert on the time scale of the measurements.[6] To a certain extent this limitation can be overcome by time-resolved 2D IR spectroscopy of biomolecules in solutions.[7] This approach, basically, detects anharmonic couplings of molecular vibrations; the detected correlation can be used to constrain structural computations. Although demonstrated for several different systems, the 2D IR approach is by far

lower in its performance (e.g. size of molecules and accuracy) than NMR, primarily due to a poor spectral resolution in solutions.

One common problem for the condense phase methods is a need for highly purified sample in relatively large quantities. The development of electrospray ionization, which allows for transfer of large, non-volatile biomolecules directly from solution to the gas phase, enabled the use of mass spectrometry (MS) for structural determinations. High sensitivity and m/z resolution of MS instrumentation drastically reduce the required amount and purity of samples. While MS alone is low informative for 3D structure of biomolecular ions, there are, at least, two associated approaches, which may deliver certain structural information. For instance, electron transfer/capture dissociation methods demonstrate certain selectivity to tertiary structure of proteins, allowing an evaluation of their overall folding state, although by far these methods cannot serve for unambiguous validations of 3D structures.[8] The technique of ion mobility, which measures drift times of ionic conformers dragged through buffer gas by a weak electrical field, is capable to separate conformers with sufficiently different collisional cross sections.[9] Each drift time can be converted to an ionic collisional cross section, thus providing a direct but single structural constraint for validating calculated 3D geometry of a conformer (conformational family). As none of the mentioned above methods alone can provide unambiguous assignment of each calculated single conformational structure, the current developments in structural elucidations are two-fold. The continuous developments in computer technology and computational theory are making calculations faster, which allows for treatment of larger molecular systems and with a better accuracy. Experimentally, one would like to combine different proven techniques and implement new approaches for measuring more and more structural constraints, suitable for pre-selecting calculated molecular geometries and for highly confident validation of the final optimized structures. Laser spectroscopy, when combined with cryogenic cooling of ions and their mass spectrometric detection is one of the recent tools capable to provide large numbers of structural constraints for gas-phase biomolecules. The ions of interest can be stored in a cryogenic ion trap, cooled in collisions with a buffer gas (e.g., He) and then spectroscopically interrogated by pulses of one or several lasers. The early stage developments in this field have been reviewed elsewhere.[10] Herein, after giving some tutorial background information and a short historical reference, we review the most recent developments in IR/UV cold

ion spectroscopy for structural determinations and identifications of biomolecules. We then shortly overview the recently launched[11] new research direction in cold ion spectroscopy – analytical identifications of isomeric biomolecules.

## I. Vibrationally-resolved Spectroscopy of Biomolecules

Spectroscopy measures the energy and intensity of optical transitions between quantum states of molecules, therefore reflecting their molecular structure. IR spectroscopy in particular can measure vibrational frequencies, which directly reflect nuclear motions in molecules. Large biomolecules are typically classified as asymmetric tops, such that all fundamental transitions are formally IR active. Frequencies of these transitions can serve for a stringent multipoint validation of calculated structures, although this requires calculation of the frequencies for each candidate structure. UV spectra of large molecules containing aromatic groups reflect more complex vibronic transitions (rotational resolution is unrealistic already for midsize biomolecules), which are governed by the Frank-Condon principle and which are challenging to reproduce by calculations. This makes a straightforward use of UV spectra for structural elucidations very challenging. Instead, high-resolution UV spectra, molecular "fingerprints", which are reach in details, can be used for identification of structurally different molecules, including isomers and conformers. In addition, well-resolved UV transitions can be used for spectroscopic tagging of different conformers in IR-UV double resonance vibrational spectroscopy. Finally, UV excitation can be used for photofragmentation of ions, serving as a convenient and highly sensitive detection tool.

### *1.1 Suppression of Inhomogeneous Broadening by Cooling*

One common problem of optical spectroscopy is the spectral congestion that may arises (i) from thermal population of many molecular (including conformational) quantum states, (ii) from broadening/ splitting of transitions to excited vibrational/vibronic states due to their short lifetime. Overlap of these numerous transitions prohibits vibrational resolution in room temperature vibrational/electronic spectra of mid- to large size molecules, drastically reducing the obtained informational content. The first experimental challenge in spectroscopy of biomolecules is therefore to minimize this congestion. While the lifetime broadening (splitting) is intrinsic to excited states and cannot be suppressed, the inhomogeneous

thermal congestion can be fully removed by condensing all the ions to a single rovibrational state. Cryogenic cooling of rovibrational degrees of freedom in large ions isolated in the gas phase may indeed greatly suppress this broadening (Figure 1.1).

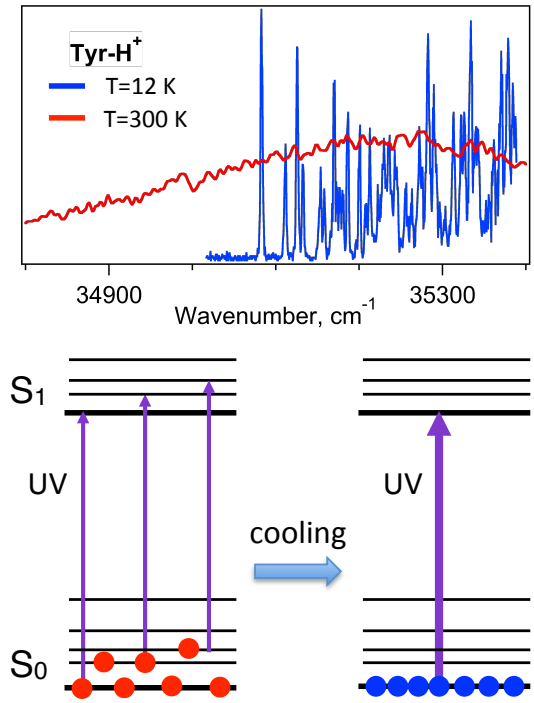


**Figure 1.1** Schematic energy diagram illustrating suppression of spectroscopic thermal congestion by cryogenic cooling. The top panel compares UV spectra of gas-phase protonated tyrosine measured at room temperature (300 K) and cooled to vibrational temperature of 12 K in a cryogenic 22-pole ion trap.[12]

But, how low one should cool biomolecular ions to attain vibrational resolution in their spectra? Cryogenic rotational cooling is essential for small and midsize biomolecules only. Indeed, in the absence of lifetime broadening the width of a vibrational transition, $\Delta\nu_{rot}$, is determined by rotational thermal distribution and, roughly, by the largest rotational constant of a molecule as:

$$\Delta\nu_{rot} = \sqrt{8T_{rot}B}, \qquad (1)$$

where $T_{rot}$ and B are rotational temperature and the constant, respectively. For instance, for a decapeptide gramicidin S, B≈ $1.2\cdot10^{-3}$ $cm^{-1}$,[13] such that already at T=100 K the rotational broadening reduces to ~1 $cm^{-1}$. The broadening scales down upon increasing the number of residues $N$ in a peptide approximately as $\Delta\nu_{rot} \propto \sqrt{T_{rot}}/N$. Extrapolating the above numbers to a small protein cytochrome C (104 residues), one gets ~ 0.1 $cm^{-1}$ for the rotational broadening, which is on the scale of the linewidth of typical high-resolution tunable ns pulsed lasers. Vibrational cooling

is more essential for suppressing thermal congestion. Typically, the energy of the lowest excited vibrational states in biomolecules doesn't exceed 15-20 $cm^{-1}$ (large amplitude motions of molecular groups). To limit the thermal population of such levels, let say, to 10% of the ground state population, one would need a vibrational temperature of 9-12 K. Such cooling will reduce the intensities of vibrational "hot" bands (the transitions, originated from thermally populated non-ground vibrational levels) relative to the intensity of the fundamental transitions (the transitions, originated from the vibrational ground state) to below 10%. A more stringent condition of 1% would require only a slightly lower vibrational temperature of 6-8 K. Overall, the above consideration suggests that, the internal cooling to below ~10 K will largely remove thermal congestion in IR and UV spectra of midsize and larger biomolecules. Further cooling, for instance, by embedding biomolecules to He nanodroplets (0.37 K), unlikely will simplify the spectra, although can be useful for some spectroscopic applications.

The last source of thermal congestion in IR and UV spectra is conformational heterogeneity of biomolecules. This source of broadening can be essential even at a cryogenic temperature, provided the cooling is of non-adiabatic nature. Relative population of the conformers that are separated by the energy barriers much higher than this temperature, is not a function of the temperature itself, but rather of the cooling rate.[14] An ultimately slow (adiabatic) cooling should condense all the molecules in the global vibrational ground state, although the, typically, fast (compared with the rates of conformational interconversions) non-adiabatic collisional cooling in RF ion traps results in trapping of cold ions in higher local potential minima too. A prompt collisional cooling also allows for bracketing the energies of high-energy conformers, which may be spectroscopically observed in cryogenic ion traps. Indeed, biomolecules are often transferred to the gas phase using the electro spray ionization (ESI) technique. The internal temperature of the ions in ESI, typically, does not exceed T=500K.[15] Taking this value and assuming that for a reliable detection the spectroscopic transitions due to minor conformers should be, at least, of 10% intensity ($n_i$) relative to the similar peaks of the lowest energy conformer ($n_0$), one can evaluate the maximum relative potential energy of the conformers that can be detected as:[16]

$$\Delta U \le kT \cdot \ln(n_0 / n_i) \cong 2.3\,\text{kcal/mol}. \qquad (2)$$

The meaning of the 2.3 kcal/mol cutoff is that, no gas-phase molecular conformers with the potential energy (relative to the most stable conformer) above this value can be detected in a typical cold ion spectroscopy experiment. This energy cutoff can be quite useful in limiting the conformational search of suitable gas-phase molecular structures generated by quantum chemistry computations. A lack of a calculated structure, whose energy is below 2.3 kcal/mol and for which the calculated IR spectrum matches to the experimentally measured one, may indicate that this conformer originates from a ”kinetically trapped” solution geometry.[17] Apparently, the suggested energy cutoff rule can be applied only, if the computational accuracy is proven to be much better than 2.3 kcal/mol.

Collisional cooling of small to midsize biomolecules in cryogenic traps is, typically, sufficiently fast and does not suppress the thermal congestion, caused by conformational heterogeneity. A positive side of this effect is that a researcher has access to high-energy conformers, some of which may, roughly, reflect the geometry of molecules in solution. This opens additional opportunities for revealing the landscape and interconversion of the conformers, although requires more complex spectroscopic tools to attain conformational selectivity.

Overall, internal cooling is more essential for UV, than for IR spectroscopy. The latter is, typically, performed in the 1000-3700 $cm^{-1}$ spectral regions of characteristic high-frequency vibrations (CO, NH, OH groups), where transitions from thermally populated lowest-frequency vibrational modes appear as “hot bands”. Already at a moderate cooling, only a few modes, for which $E_{vib} \geq kT_{vib}$, remain sufficiently populated, making thermal broadening less significant than, for instance, the homogeneous broadening due to anharmonic mode couplings (including hydrogen bonds). Lifetime broadening due to intramolecular vibrational energy redistribution (IVR) in biomolecules thus remains the intrinsic limit for spectral resolution in IR spectra, and it cannot be suppressed by cooing. In UV spectra, each low frequency mode, in addition to a hot band, manifests itself as closely spaced vibrational progression that originate from the thermally populated level with the number of peaks in each progression governed by the Frank-Condon factor. An overlap of such progressions quickly congests the UV spectra upon an increase of the energy of UV photons above the electronic band origin.[18] Lifetime broadening of the electronic band origin and a few lowest energy vibronic transitions is insignificant, provided the excited electronic states do not undergo fast internal conversion or

intersystem crossings. The excitation of these states, indeed, doesn't induce any IVR on the excited electronic surface. The contribution of IVR to spectral broadening of vibronic transitions quickly increases however upon increasing vibrational energy of the excited electronic state.

### *1.2 Types of Cold Ion Traps*

A convenient way to cool a large number of biomolecular ions produced by electrospray ionization sources is to store them in a cold ion trap and then remove internal energy in inelastic collisions with a cold buffer gas. There are many types of ion storage devices, which have been demonstrated as suitable for collisional cooling of molecular ions to as low as a temperature of a few K. For a deeper understanding of the principles, the advantages and limitations of different cold ion storage devices we can recommend, for instance, a nice book chapter by Gerlich,[19] as well as some later review and research articles on this topic.[10, 20-26] Here we only briefly summarize the most essential fundamental and practical aspects of cooling large biomolecular ions in the most popular radio frequency-driven linear and 3-D Paul ion traps.

Radio frequency-driven (RF) quadrupole Paul traps,[27] made of stainless steel, are readily commercially available and convenient for interfacing with time of flight mass spectrometers. There are several drawbacks of these traps too. First, the commercial traps still need to be mechanically modified to get access to the stored ions by laser beams. The number of stored ions is also limited due to a small volume of these traps. Most important, however, is the fact that they capable of only a limited collisional cooling of the stored ions. Most of the publications report vibrational ion temperatures above 40-50 K with the traps cooled to below 10 K.[28] A recent report by Fuji et al. suggests that it might be an effect of the extremely low thermal conductivity of stainless steel at low temperatures. They solved the problem by making a Paul trap of gold-plated copper and demonstrated the vibrational temperature of T=12 K for the cooled ions (Figure 1.2).[29] We believe, it is the RF self-heating of the trap internal surfaces rather than the low thermal conductivity of stainless steel that limits the lowest attainable temperature of standard Paul traps. A gold plating of such stainless steel traps should solve the problem in a most economical way, yet to be verified.

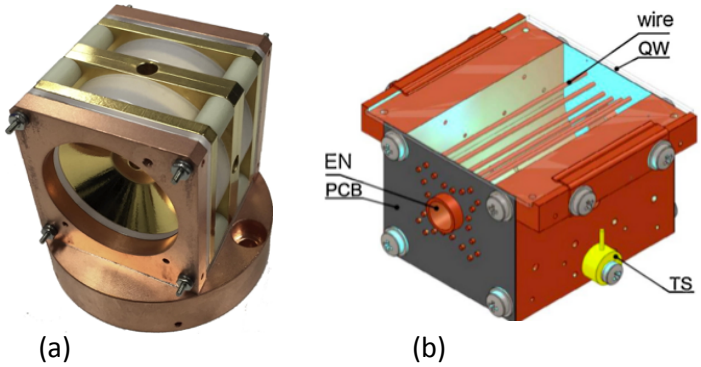


**Figure 1.2** (**a**) Photo of the modified 3D quadrupole cryogenic ion trap from Tokyo (reproduced with permission from ref[29]) and (**b**) the schematic of the wired linear quadrupole cryogenic ion trap designed in Prague (reproduced with permission from ref.[30]).

Linear cold ion traps used for spectroscopy of ions, including heavy biological ones (peptides and proteins), are not available commercially. Several different designs, which differ, first, of all, in the number of poles, were developed over last two decades, after pioneering works by Gerlich.[19] Regardless of the number of poles, which ranges from 4 to 22, they offer an easy laser beam access to the stored ions and vibrational cooling to below 10 K. Historically, the first successful experiments on spectroscopy of electrosprayed biological ions have been performed in 22-pole cold traps,[12] which were initially developed for spectroscopy of small ions, relevant to astrophysical research. Typically, 22-pole traps offer a large "effective" ion storage volume and low internal, but substantially higher translational temperatures.[19, 21] From point of view of laser spectroscopy, a severe practical drawback of 22-pole traps is insufficiently stable radial confinement of cold ions. Because of a very flat effective radial potential in these and other multipole traps (Figure 1.3), the radial position of ion cloud in such traps is very sensitive to mechanical imperfections and contaminations of the trap poles.

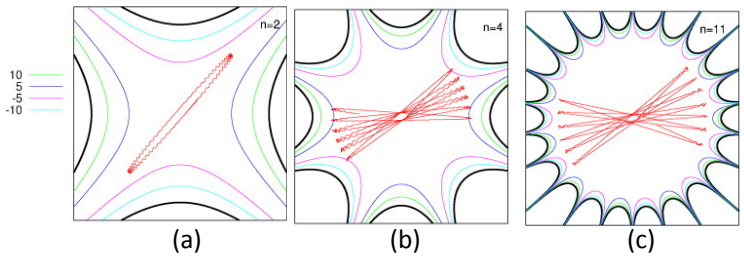

**Figure 1.3** Equipotential lines of "time-effective" electric potentials in (a) quadrupole, (b) octupole and (c) 22-pole ion traps. Red wiggling lines illustrate typical ion trajectories (reproduced with permission after an adaptation from ref.[21]).

These practically unavoidable circumstances can displace the ion cloud out of the trap axis, making a reliable day-to-day overlap of laser beams with the cloud difficult. Development of lower order linear cold ion traps, such as octupoles[22-23, 31] and even quadrupoles,[30] which all have a well-defined radial minimum (on the trap axis) of the effective RF potential, have solved this practical problem (Figures 1.2 and 1.4). Vibrational temperature of, for instance, a protonated dipeptide, estimated from intensity of a UV hot band (0←1 vibronic transition) relative to the intensity of the respective 1←0 vibronic transition, appears to be below 10 K with a 6 K octupole ion trap.[23] This cooling capability is even slightly better than that in a 22-pole under similar conditions. Observation of weakly bound He-ion clusters formed in the octupole ion trap is another, qualitative, manifestation of a low achievable internal ion temperature. Such complexes were routinely observed in a linear, state-of the art, "wired" quadrupole ion trap, in which four sets of thin wires instead of four solid rods were used to reproduce quadrupole RF filed.[30] The use of wires allows for a radial optical access to the stored ions, which can be convenient, for instance, for collecting laser-induced ionic fluorescence.

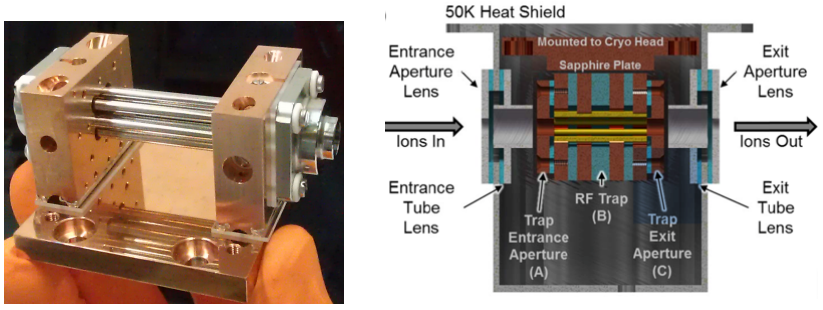


**Figure 1.4** (**a**) Photo of the octupole cryogenic ion trap developed in Lausanne[23], and (**b**) schematic of the octupole cryogenic trap from Yale (right; reproduced with permission from ref.[31]).

Recently, the group of Polfer reported on design of a cryogenic linear quadrupole ion trap,[32] which combines mass-selectivity of Paul traps with large trapping volume and easy laser beam access of linear traps. The cooling substantially

narrowed vibrational transitions in the measured IR spectra, although the achieved ion temperature has not been accessed.

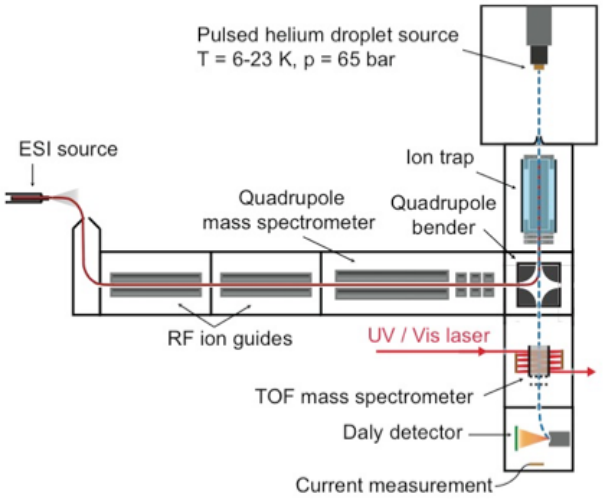


**Figure 1.5** Experimental setup in FHIMP (Berlin) for spectroscopy of ions embedded to He nanodroplet (T=0.4 K). **(**Reproduced from Ref. [33] - Published by the PCCP Owner Societies).

An ultimate cooling of biological ions to sub-Kelvin temperatures was demonstrated by von Helden.[33] In this approach, a beam of He nanodroplets picks up the ions stored in a room temperature ion cell (Figure 1.5). Once in a droplet "nanofridge," an ion is instantaneously cooled by evaporating a fraction of the droplet, presumably, close to the 0.38 K of the droplets. The benefits of such low temperature are not clear, however. The reported IR spectra of protonated peptides and proteins exhibit the level of spectral congestion and width of vibrational transitions, which are even higher than in the spectra of similar species cooled in linear 4-6 K ion traps. The spectral broadening can be attributed to the linewidth of the used free-electron laser and/or to lifetime broadening due to vibrational energy relaxation from ions to droplets. The latter phenomenon intrinsically broadens IR transitions and cannot be suppressed by lowering ion temperature.

## *1.3 Spectroscopic Techniques*

### 1.3. *IRMPD Vibrational Spectroscopy*

Low concentration of ions in cold traps (e.g., typically $10^5$-$10^6$ $cm^{-3}$ for an octupole trap) forbids a use of direct absorption for spectroscopy of charged molecules. Fragmentation of ions induced by absorption of laser light can be

employed instead to monitor spectroscopic transitions in biomolecules. Several types of such "action" spectroscopy techniques were developed and applied to biological ions. Perhaps, the most versatile and simple method of action spectroscopy is infrared multiple photon dissociation (IRMPD), which employs the phenomena of absorption of several photons by polyatomic molecules from a highly intense pulse of an IR laser (Figure 1.6).

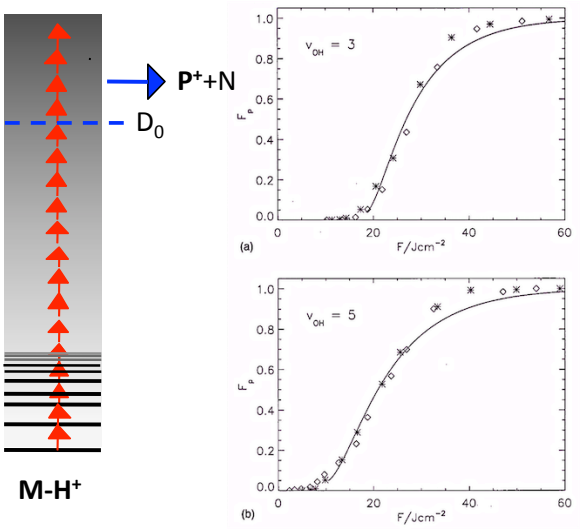


**Figure 1.6** Left: Schematic energy diagram for IRMPD of protonated molecules $[M-H]^+$, yielding charge fragments $P^+$ and neutrals. Right: Experimentally determined IRMPD product yields for methanol molecules preexcited to $\nu_{OH}$=3 (a) and for $\nu_{OH}$=5 (b) as function of $CO_2$ laser fluence $F$ (symbols) and their biexponential fits (solid line). Reproduced with permission from ref[34].

Once the total imported energy exceeds the lowest dissociation threshold of the molecule, it may fragment on a timescale determined by the statistical (e.g., RRKM) theories (unless the laser pulse is very short (ps or less)). The statistical nature of IRMPD process leads to a few practically important principles. First, regardless of the nature of the excited molecular transition, the weakest bond will always break first (or the most efficiently). This golden rule of IRMPD reflects the phenomena of the intramolecular vibrational energy redistribution (IVR) in polyatomics, which suggests that, statistically, the imported to a molecule vibrational energy will be promptly and equally redistributed among all the molecular modes.[35] Second, for an efficient IRMPD on a short (typically μs-ms) timescale of observations, large molecules (e.g., peptides, proteins) have to be significantly overexcited above the lowest dissociation threshold.[36] The apparent advantage of IRMPD approach for

monitoring vibrational transitions in polyatomics is its universality in applying to biomolecules of any composition, including the peptides that have no UV chromophores. Apart from a need for a high-energy tunable IR laser source, the associated drawbacks of IRMPD include a lack of conformational selectivity, spectral broadening of IR transitions and a substantial distortion of relative intensities of the transitions, which is caused by the non-linear nature of IRMPD. The latter may cause a complete suppression of weak IR transitions, misleading comparisons with computed spectra. This effect, however, can be reduced by carefully controlling the fluence of the dissociating IR laser. In general, the IRMPD yield is a complex function of $\Phi\cdot\sigma$ ($\Phi$- laser fluence, $\sigma$-absorption cross section of the pumped fundamental transition), which can be schematically characterized by an initial exponential growth below a threshold value $(\Phi\sigma)_0$, a nearly linear increase and, finally, above a saturation value, $(\Phi\sigma)_\infty$, by a near exponential asymptotic approach of the dissociation yield to unit (Figure 1.6). For large polyatomics, both critical values of $\Phi\cdot\sigma$ are governed by the density of vibrational states in the molecule and are difficult to quantify. A practical recipe for recording a nearly truly linear (as a function of $\Phi$) IR spectrum is (i) to measure a preliminary IRMPD spectrum; (ii) for any found (reference) IR transition to record dissociation yield as a function of laser fluence and to evaluate $(\Phi\sigma)_0$ and $(\Phi\sigma)_\infty$ parameters; (iii) fix the fluence at a low, $\Phi_L$, to record truly spectra of the strong transitions, for which $\Phi_L\sigma < (\Phi\sigma)_\infty$, and then at a high, $\Phi_H$, to reveal the weak transitions; (iv) combine the two spectra, using the reference transition for their scaling. Interestingly, IRMPD at low fluence ($\Phi\cdot\sigma << (\Phi\sigma)_0$), when the non-linearity of the process is significant, may cause narrowing of a transition: the yield at the maximum of the transition band will pop-up, while its low-intense wings will be further reduced in the IRMPD spectrum. This artifact can be used to improve spectral resolution in IRMPD spectra.

Several groups reported the use of widely tunable high-energy free-electron IR lasers for recording vibrational MPD spectra of charged biological molecules at room temperature.[37-39] Typically, an IRPMD spectrum, which reflects vibrational transitions in all abundant ionic conformers, is fitted by a sum of the frequency-scaled spectra, calculated in harmonic approximation for a few computed lowest energy structures (Figure 1.7). Apparently, such fits may leave a substantial ambiguity in validating both, the number and the identity of the involved conformers.

Nevertheless, the simplicity of the technique makes is particularly attractive for practical analytical applications.

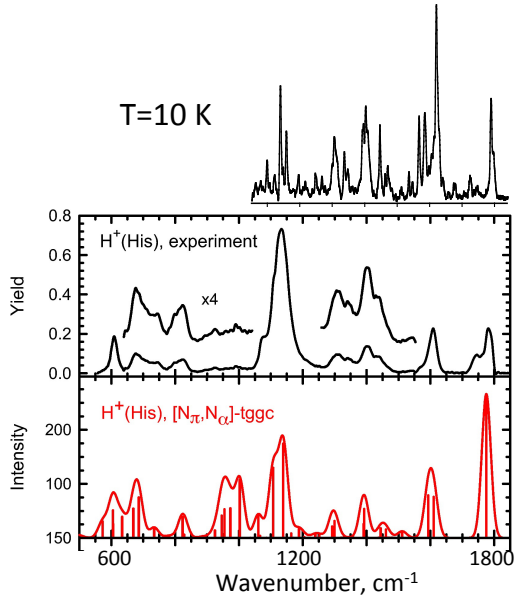


**Figure 1.7** Room temperature IRMPD spectrum of His-$H^+$ and its fits by IR transitions calculated for one of the candidate structure (lower panels, reproduced with permission from ref.[38]), compared with IR spectrum of this amino acid, measured at T=10 K[16] (on top).

### *1.3.2 Tagging Spectroscopy*

Tagging action spectroscopy employs non-covalent binding of "messenger" molecules or atoms (neutral or ionized) to the ionic (or neutral) analyte molecule.[37, 40-41] The absorption of UV or IR light is then manifested as a dissociation of such complexes, which reduces their concentration (Figure 1.8). An action spectrum of the complex is recorded by monitoring this reduction (depletion) as a function the dissociating laser wavelength. Cryogenic cooling of ions enables a use of many different neutral atoms and molecules for gas-phase tagging in cold ion traps. Light species of low polarizability may cause only little structural distortions of an ion they are complexed with.[31] Photofragmentation IR spectra of such clusters are often almost identical to the spectra of bare ions and, therefore, suitable for structural determinations of the latter. The ionic structure and spectra are particularly little distorted in the complexes with helium atoms.[23, 41] Due to low binding energies of He, an efficient formation of such clusters requires very low, typically below 10 K,

vibrational temperature of ions; a small size and high charge state of an ion facilitate formation of the clusters. An observation of substantial concentration of He-tagged ions in a cryogenic ion trap, in its turn, manifests a low ion temperature achieved in the trap. Several groups reported on photofragmentation spectroscopy of He-tagged ions. The group of Roithova and Gerlich, for instance, cools ions in their elegant "wired" quadrupole cryogenic trap (Figure 1.2b).[30] They demonstrated the distortion-free IR spectroscopy of He-tagged dications generated from benzene, but revealed certain structural changes induced by complexation of $HCCl_2^+$ dications with He.[41] The group of Johnson, who were among the pioneers of the tagging spectroscopy,[40, 42] uses tagging by $H_2$, Ar and Ne in a cryogenic Paul trap at T>10 K and by He in a 4.5 K octupole ion trap (Figure 1.4b) for IR spectroscopy of protonated water clusters; they quantified the distortions of vibrational spectra induced by the tags.[31] The group of Asmis recorded conformer-specific IR photodissociation spectra of protonated water clusters employing $H_2$ tagging in a ring-electrode cold trap.[43-44]

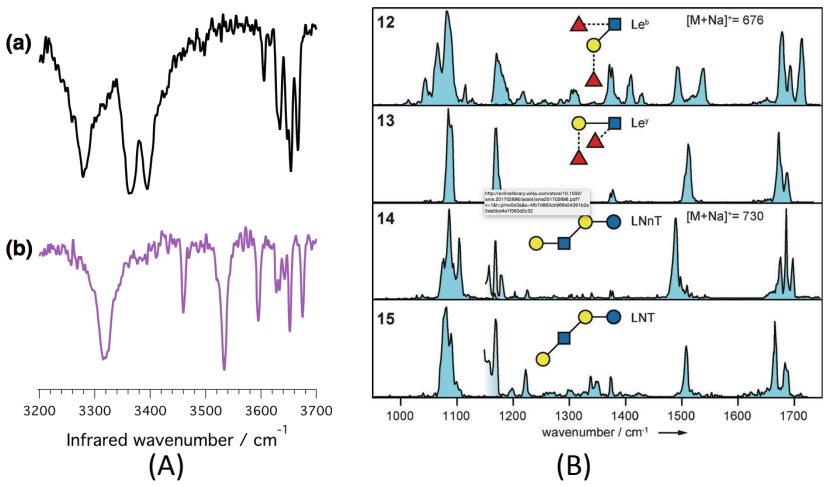


**Figure 1.8** Infrared signatures for (**A**) the α (a) and β (b) isomers of sodiated GalNAc(1–3)Gal recorded at a temperature of 69 K with $N_2$ as a messenger tag (reproduced with permission from ref.[45]); (**B**) tetrasaccharides $Le^b$ (12), $Le^y$ (13), lacto-N-neotetraose (14) and lacto-N-tetraose (15) cooled in He nanodroplets (reproduced with permission from ref.[46]).

Complexes of neutral biomolecules with charged tags, for instance with $Na^+$, which is often present in solvents in sufficient concentrations, can be formed directly in solution. In such cases the messenger ion makes the complex charged. Alternatively, other alkali metals (e.g., Li Rb, Cs) can be added into solution as salts to form similar complexes. The binding energy of alkali metals to neutral molecules

is, typically, too high to break the clusters by a singe IR photon.[47] Such complexes are therefore suitable only for IRMPD or UV photofragmentation spectroscopy. A severe drawback of this approach is unavoidable distortions of the analyte 3D geometry by the interaction with such a harmful tag. Although these distortions limit the use of tagging by alkali metal ions for structural determinations, the technique can be conveniently employed for identifications of analyte molecules by recording and comparing their spectroscopic signatures. As a bonus, these complexes can be studied in room temperature ion traps. The group of Polfer used alkali ion tagging and IRMPD spectroscopy for qualitative identifications of isomeric glycans in room temperature experiments.[37] They demonstrated that the measured spectra are visually different for different isomers, although did not quantify the differences. Once measured for a set of known isomers of a molecule (labeled by its m/z), such "library" IRMPD spectra, potentially, can be compared with a spectrum of an unknown analyte for its identification.[48] Recently, the group of Rizzo further developed this approach by combining $Na^+$ tagging of saccharides in solution with $H_2$ or $N_2$ gas-phase tagging of such sodiated charged analytes in a cold ion trap for IR depletion spectroscopy.[45, 49] In addition, the isomers were characterized by ion mobility drift times. Because the binding energy of $H_2$ and $N_2$ tags to a charged sodiated analyte is much lower than that of alkali ions, a much lower fluence of IR laser was required for a multiple-photon detachment of single neutral tags. The recorded in 3 μm region IR depletion spectra of such doubly tagged glycans appeared visibly different for up to tetra-saccharides (Figure 1.8A). Particular examples of the quantitative identifications (e.g, for solution mixtures of the isomers) yet have to be demonstrated in these studies to assess the analytical performance of this approach. The group of von Helden demonstrated cooling and tagging of up to tetra-glycans by ultracold (0.4 K) helium nanodroplets.[46] The droplets, doped by the analyte ions, were evaporated upon absorption of, likely, several IR photons delivered by a free-electron IR laser, tunable in 12-6 μm spectral range. Similar to the spectra reported by Rizzo's group for the 3 μm range, the recorded spectra appeared visibly distinguishable for different isomers of the tested saccharides (Figure 1.8B), yet without quantitative comparisons of the spectra. An exceptional technical complexity of the method (generation of He droplets and the use of a free-electron laser) questions its practical analytical value, however.

### *1.3.3 Photofragmentation UV Spectroscopy*

As biomolecules are, typically, large polyatomics with a variety of chemical bonds, the absorption of a single UV photon usually breaks some of them. The outcome of UV excitation can be a reduced number of parent ions and an appearance of the fragments, for which m/z ratio differs from that of the parents. These charged fragments or the depleted parent ions can be selectively detected by means of mass spectrometry: the stronger is the absorption the high is the UV-induced number of fragments or the depletion of the parent ions. Vibrational "pre-heating" of cold ions by an IR photon can significantly alter their UV absorption.[18] The detection of this change allows for conformer-selective or non-selective IR spectroscopy of cold ions, using the described below IR-UV double resonance schemes. An ion therefore must contain a UV chromophore. In the case of peptides and proteins this can be one of the aromatic residues (His, Phe, Tyr or Trp), which all absorb in UV, although differently. Figure 1.9 compares the gas-phase UV absorptions of the four protonated aromatic amino acids, produced by electrospray ionization from their equimolar solution mixture and measured by photodissociation spectroscopy.[16]

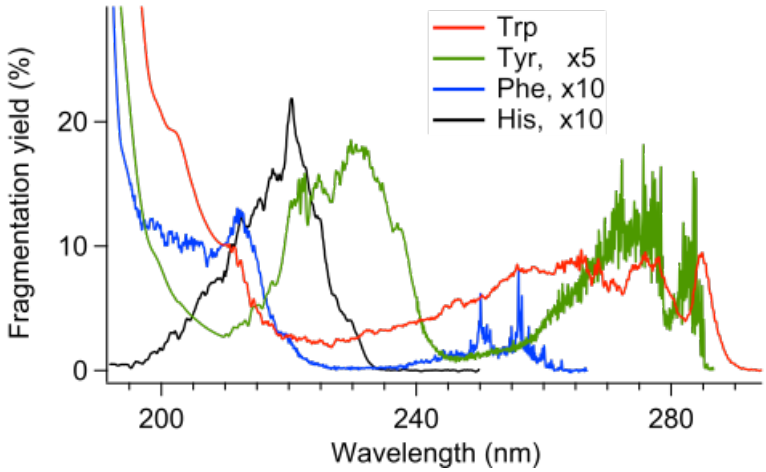


**Figure 1.9**. Photofragmentation spectra of four protonated aromatic amino acids recorded by integrating all major charged fragments, which were detected by an Orbitrap mass analyzer under similar settings of the cold ion spectrometer and DUV/UV OPO. The spectra are normalized to the total number of ions and to the OPO pulse energy of 1 mJ. The spectral resolution is 0.04 nm. Reproduced with permission from ref.[16]

Among these aromatics, tryptophan is by far the strongest UV absorber in the first absorption band, followed by Tyr, His and Phe. Despite this arrangement of absorption intensities, most of the published works on IR cold-ion spectroscopy of

biomolecules was done with Phe and, less often, with Tyr as a chromophore in peptides. Apart from the higher natural abundance of Phe (4%) and Tyr (3.3%) compared with His (2.9%) and Trp (1.3%), the main reason for the "popularity" of the two former aromatic residues is the frequent observations of vibrationally-resolved structure in UV spectra of the peptides that contain these two chromophores, Phe in particular. Sharp UV transitions, assigned to Phe absorption have been observed even in a small protein ubiquitin.[50] Tyr is less prone to conserve structured UV spectra even in relatively small peptides. This is, likely, due to the larger number of Franc-Condon active vibrations of this residue and the ability of its OH group to be involved to non-covalent interactions. An often lack of any structure in the absorption bands of tryptophane is, likely, due to the known high susceptibility of the Trp electronic terms to non-covalent interactions of its indole ring (e.g., proton-$\pi$ □□□□□□□□s). These interactions may drastically shorten the lifetime of the $S_1$ excited state, which leads to a large homogeneous broadening in UV spectra of Trp.[12] Recently, the UV photodissociation of the gas-phase His amino acid was also characterized.[16] In contrast to $TyrH^+$ and $PheH^+$, but similar to $TrpH^+$, the UV spectrum of $HisH^+$ exhibits no vibrationally resolved structure and appears red-shifted by 12 nm to the red, relative to the absorption of this amino acid in aqueous solutions. Opposite to His-$H^+$, desolvation of the three other protonated aromatic amino acids shifts their gas-phase UV spectra slightly to blue. IR cold-ion spectroscopy revealed three low energy conformers of His-$H^+$, which are protonated on the imidazole ring, but not on N-terminus, as the other three aromatic amino acids. It is exactly this distinct protonation site that makes UV absorption of His-$H^+$ highly sensitive to the local environment of the ring and explains the unusual spectral properties of this aromatic amino acid.

It may happen that a peptide of interest contains no aromatic residues. Indeed, the natural abundance of aromatic amino acids suggests that, for instance, 56% of pentapeptides have none of these chromophores. Vibrational spectra of a good fraction of oligopeptides therefore cannot be probed by IR-UV technique. Still, one may rely on the "universal" VUV/UV absorption by peptide bonds to perform IR-UV vibrational spectroscopy (Figure 1.10). This recent demonstration makes cold ion spectroscopy a versatile research tool for structural studies of, eventually, any small to midsize peptides, as well as of many other types of aromatic molecules.

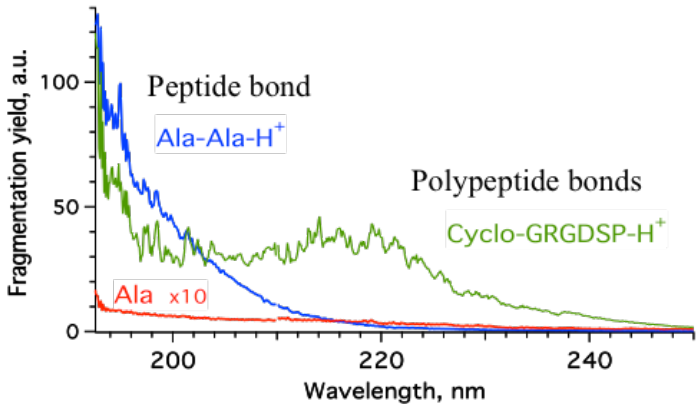


**Figure 1.10**. Photofragmentation electronic spectra of the biomolecules that do not contain any chromophore groups: singly protonated amino acid (magnified by a factor of 10), dipeptide and cyclic hexapeptide.[51]

### *1.3.4 IR-UV Depletion and "Gain" Vibrational Spectroscopy*

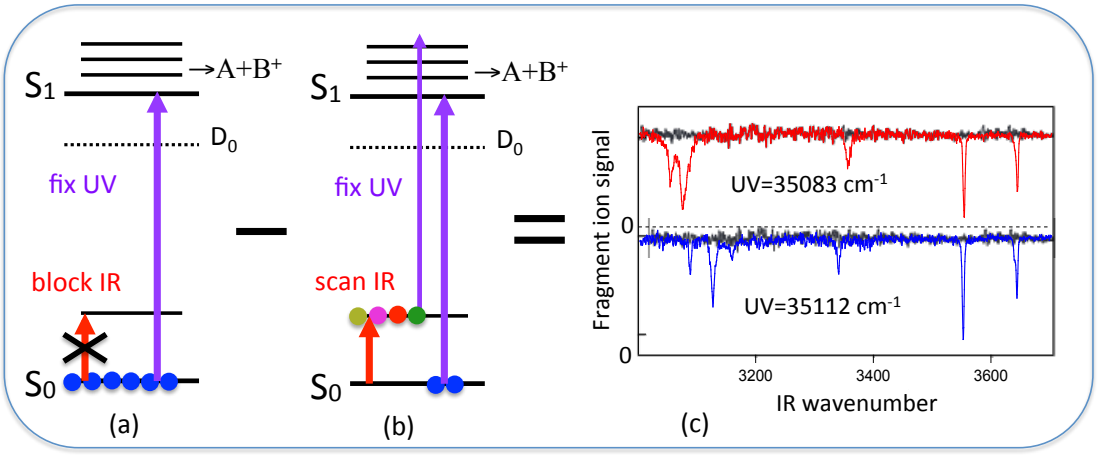


**Figure 1.11** Schematic energy diagrams (**a**, **b**) and IR spectra of two conformers of Tyr-H$^+$ (**c**), illustrating IR-UV vibrational depletion spectroscopy. A depletion IR spectrum results from the differences of ion signals measured with UV laser wavenumber fixed on a vibronic peak, while the scanned IR laser is blocked (scheme (**a**); black traces in (**c**)) and unblocked (scheme (**b**); red and blue traces in (**c**)). The vibrationally excited ions reside in different vibrational states of the same total vibrational energy, as it is color coded in (**b**). Compared with the ground state, only the ion in a few of these excited states can absorb UV light at each fixed wavenumber.

Vibrationally resolved UV spectra allow for measurements of conformation specific IR spectra of biomolecules using an IR-UV depletion technique. Initially developed for neutrals,[52-56] this approach quickly became the main driving horse for conformer-selective IR cold-ion spectroscopy.[57-58] Figure 1.11 shows a schematic energy-level diagram of IR-UV excitation. In this approach the wavelength of a fragmentation UV laser is fixed at a well-resolved vibronic transition originated from the vibrational

ground state of a cold ionic conformer. A preceding IR pulse, when absorbed by this conformer, heats up the ions internally, transferring a fraction of the ground state population to many different vibrationally excited states of near the same energy (determined by the linewidth and energy of IR photon). The subsequent UV pulse therefore interacts with a lower number of ions, which remain in the ground state. This results in a drop of the UV dissociation yield from this state. Although the ions in the vibrationally excited states may also absorb UV photons and contribute to photofragmentation, this contribution is small to negligible for large molecules. Indeed, at the energy level of an IR excitation, the density of vibrational eigenstates in biomolecules is astronomically high (e.g., ~$10^{37}$ states per $cm^{-1}$ for excitation of NH stretch vibration in a decapeptide[18]). With a typical ~1 $cm^{-1}$ linewidth of an IR laser, the excited molecules will reside in a similarly large number of different vibrational quantum eigenstates. Because of anharmonicity of molecular vibrations, the frequencies of UV transitions originated from these different states are not the same. This leads to intramolecular statistical broadening of a UV transition, such that only a small fraction of the ions absorbs at a fixed UV wavelength.[18] Overall, the vibrational preheating may result in a significant decrease of UV photofragmentation at the wavelength of a UV peak: the depletions as high as 80% have been observed.[59] To some extent, the spectral broadening of UV absorption upon IR preheating is similar to the difference in the absorptions by warm and cold ions. Based on this effect, an IR spectrum is generated by monitoring the reduction of UV photofragmentation yield (at m/z of one, a few or all fragment ions) as a function of IR wavenumber in the alternative IR OPO on/off measuring cycles. IR absorption is manifested as dips in the fragmentation at the IR wavenumbers of vibrational transitions of the conformer labeled by the UV laser. It is important that IR absorptions of only the ions that share the same vibronic ground state for both IR and UV transitions will appear in the spectrum. By fixing UV wavelength at different resolved vibronic transitions, one may subsequently measure one by one IR spectra of all highly abundant conformers.

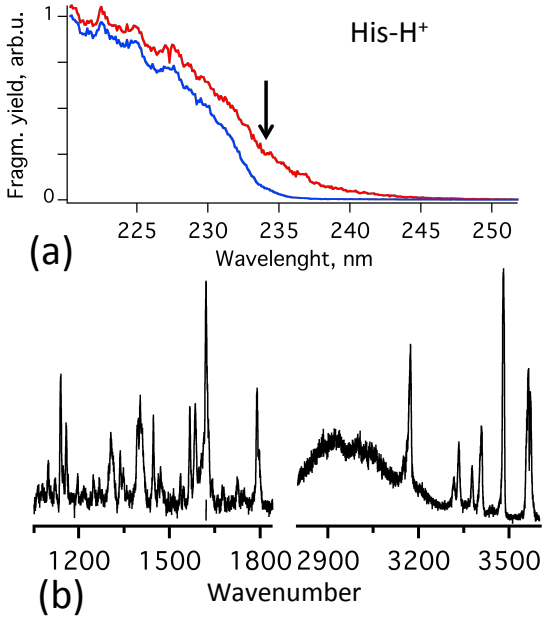


**Figure 1.12**. IR gain spectroscopy of protonated His amino acid. (**a**) IR pre-excitation redshifts the UV photofragmentation spectra (compare: blue trace – IR OPO is blocked; red trace – IR OPO is at 3482 $cm^{-1}$). (**b**) The difference of the red and blue trace as a function of IR OPO wavenumber.[16]

Alternatively, UV laser wavelength can be fixed slightly to the red from a strong UV transition or the UV band origin, where dissociation yield is insignificant. IR pre-excitation of any conformer broadens and/or redshifts its UV absorption bands, leading to an overlap of the transition and the UV wavelength and therefore to an increase in UV fragmentation. Such an IR wavelength dependence, called "gain" spectrum and measured in a single IR wavenumber scan, reflects the IR transitions of all abundant conformers of an ion.[16, 60-62] The particular wavelength of the UV laser is not important and should be selected as a compromise between the level of the signal in the cycle of measurements with IR pre-heated ions and the baseline fragmentation originated from cold ions in the alternative cycle. In contrast to IR-UV depletion spectroscopy, gain spectra may have, essentially, zero baseline, which drastically improves signal to noise ratio of the measurements (Figure 1.12).

### *1.3.5 IR-IR-UV Hole-burning Vibrational Spectroscopy*

A pre-condition of using IR-UV depletion technique for conformer-selective vibrational spectroscopy is the resolution of individual vibronic transitions in a UV photodissociation spectrum of an ion. This however may become challenging due to overlap of numerous transitions in large biomolecules and/or the molecules with high conformational heterogeneity, or intrinsically impossible due to lifetime broadening

of vibronic transitions even in small ions. One then can use vibrational transitions instead to label different conformers, provide such transitions appear resolved in IR spectrum of an ion. This approach, called IR-IR-UV hole-burning, employs two IR and one UV laser pulses, separated by 10-100 ns delays. The wavelength of the UV fragmentation laser is set to the red from the onset of UV absorption, similar to that for the discussed above IR gain spectroscopy. A pulse from the first, pump, IR laser selectively saturates a vibrational transition, specific for one conformer, such that UV fragmentation yield of this conformer becomes independent of the wavelength of the subsequent probe IR laser. This constant baseline signal can only be increased by excitation of any other conformers, except the one selected by the pump laser. Scanning the probe laser while keeping IR pump and UV fragmentation lasers fixed, generates the IR gain spectrum of all conformers, except the selected one, on top of this baseline. The vibrational spectrum of the selected conformer is generated as the difference of the UV fragmentation yields, measured in two alternative cycles (IR probe laser On/Off), as a function of IR probe laser wavelength (Figure 1.13). The use of IR-IR hole-burning technique for conformer-selective vibrational spectroscopy of cold protonated peptides has been demonstrated on several occasions. [16, 61-62]

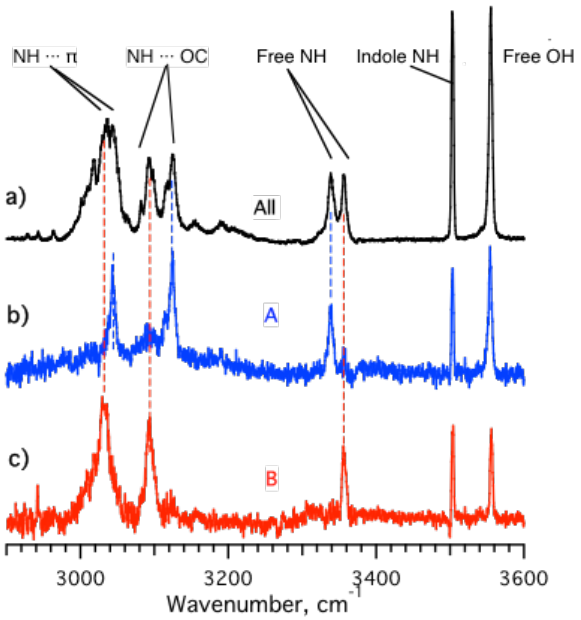


**Figure 1.13**. IR spectra of $TrpH^+$ in 3-μm region: (**a**) the IR−UV gain spectrum of all available conformers; (**b** and **c**) conformer-specific IR spectra of conformers A and B generated with the pump OPO wavenumber fixed at 3339 and 3356.7 cm−1, respectively, and UV fragmentation laser fixed at 290.698 nm.[61]

### *1.3.6 High-resolution Cold Ion Spectrometer in Lausanne*

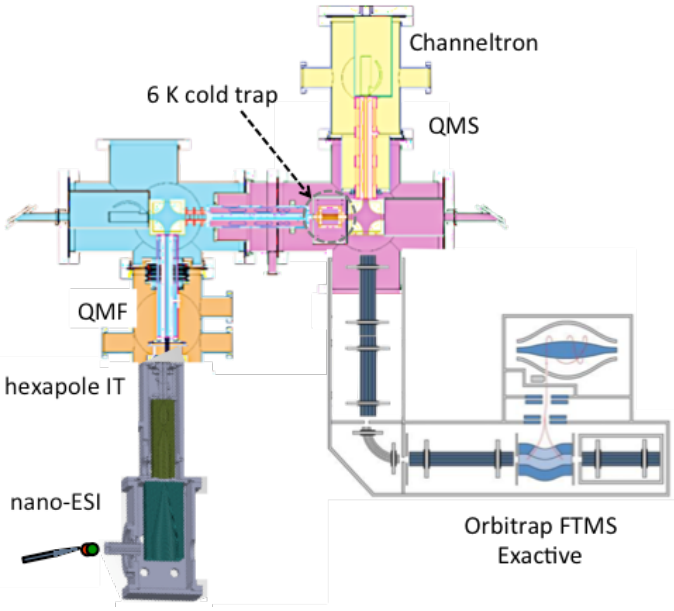


**Figure 1.14.** Schematic view of high-resolution Cold Ion Spectrometer in Lausanne.[11-12]

Implementation of all the described above spectroscopic techniques for studying of cryogenically cold biological ions requires three basic types of hardware: lasers, mass spectrometers and cryogenic ion traps. As an example, Figure 1.14 illustrates the layout of the cold-ion spectroscopy instrument, built in Lausanne for high-resolution spectroscopy and mass spectrometry of cold ions. Protonated peptides are generated from solution using a nano electrospray ion source. The ions are, first, guided into the vacuum chamber through a 0.7 mm ID 90 mm long stainless steel capillary and orthogonally injected into an electrodynamic ion funnel that works at a pressure of ~10 mbar. The ions are further r guided further through another ion funnel, held at a pressure of ~1.5 mbar, and accumulated (for ~200-500 ms) in a hexapole ion trap.[23] The trapped ions are then released from the hexapole and transmitted through a low-resolution (~100) quadrupole mass filter, which selects parent ions of a particular mass to-charge ratio (m/z). An rf-only octopole guides them further into an octopole ion trap,[2] which is cooled to 6 K by a closed-cycle refrigerator. Ions in the trap are cooled in collisions with a pulse of helium atoms, which are introduced into the trap shortly before the arrival of the ion packet. Approximately 40 ms later, when ions

are cooled and most of the helium has been pumped out, the ions are interrogated in the trap with IR and UV laser pulses. A few ms later, the content of the cold trap is released by lowering the potential of the trap exit cap and the fragment and parent ions are guided either into the quadrupole mass spectrometer (QMS) of low resolution or into the high-resolution broadband Orbitrap-based mass spectrometer Exactive for mass analysis. In the latter case, the ions can also be ejected from the cold trap without a laser excitation and transferred into the original HCD cell of the Exactive instrument for collisional activation and subsequent MS analysis of the fragments. The cold ion spectrometer is coupled to the Exactive MS by replacing the original ion source section of the MS with a two-section octopole ion guide. In the case of the ion analysis by QMS, the duty cycles of the lasers, cold ion spectrometer (typically 10 or 20 Hz) and data acquisition electronics are synchronized by trigger pulses from a master pulse generator. When the Orbitrap analyzer is used, this generator is triggered by a small fraction of the electrical pulse, applied to the entrance split lens of the analyzer. In addition to the QMS ion signal or to the entire fragment mass spectrum of Exactive, the computer-controlled data acquisition system records UV/IR laser power, measured by a pyroelectric detector, and laser wavenumbers, measured by a wavelength meter.

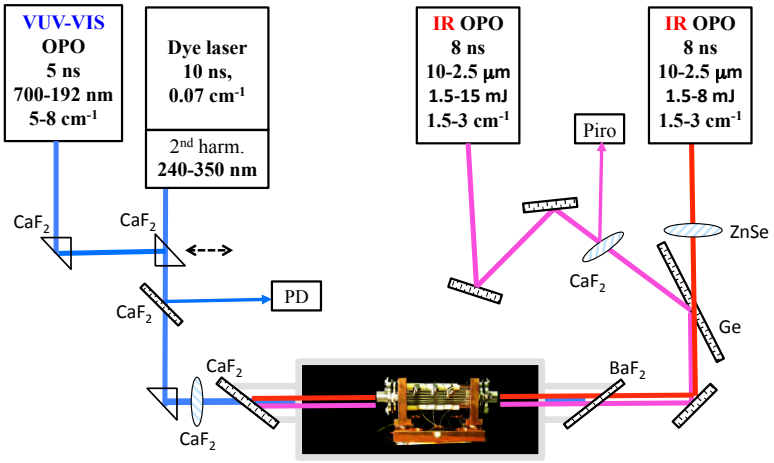


**Figure 1.14** Optical layout of high-resolution Cold Ion Spectrometer in Lausanne.

Figure 1.15 shows optical layout in the experiments performed with the Cold Ion Spectrometer in Lausanne. UV photofragmentation is performed using 2-3 mJ output of a frequency-doubled dye laser (0.15 $cm^{-1}$ linewidth), pumped by 6-7 ns frequency-tripled pulses of a Nd:YAG laser. Alternatively, the photofragmentation can be performed by a mid-resolution (5-8 $cm^{-1}$), but widely tunable (350-192 nm)

fast (~100 ms per a change of wavelength independent of the step size) UV/VUV optical parametric oscillator (OPO), which delivers 0.3 to 6 mJ (depending on wavelength) pulses at 10 Hz repetition rate. For IR spectroscopy, one (IR-UV depletion) or two (IR-IR-UV hole burning) IR OPO/OPA can be used. These sources deliver up to 15 mJ pulses at 10 Hz repetition rate in 10-2.5 μm ranges (3 $cm^{-1}$ to 1.5 $cm^{-1}$ linewidth). The IR OPO and the UV laser beams counter propagate, entering and exiting the vacuum chamber through $BaF_2$ and $CaF_2$ windows. The two IR laser beams of the orthogonal linear polarizations are combined onto a Ge plate, which is put at Brewster angle relative to one of the beams. Each beam is focused onto the center of the cold ion trap by a $BaF_2$ lens of F=50 cm. The UV beam is loosely focused by a $CaF_2$ lens of F= 75 cm onto the trap axis ~3 cm outside of its entrance to ensure an overlap of the entire ion cloud in the trap by the UV beam. Broadband pyroelectric detectors are used to record relative pulse energies of one UV laser source and one IR OPO. A small fraction of the OPO signal wave (red light) is directed to a wavemeter (0.1 $cm^{-1}$ resolution), and the measured wavelength is recorded along with the ion signal during IR OPO wavelength scan. Together with the once measured wavelength of the pumping Nd:YAG laser, this allows for an accurate wavelength calibration of IR spectra. Alternatively, when the high-resolution dye laser is scanned, its wavelength is directly recorded by the wavemeter and used for labeling of UV spectra.

### *1.4 Conformational Assignment of Spectral Transitions*

Once vibrationally resolved UV and IR gain spectra of a biomolecule are available, a reasonable question that may occur is, how many highly abundant conformers do these spectra reflect? Some transitions in the spectra may originate from the same single conformer, but some of the spectral peaks may reflect closely overlapping transitions in more than one conformer. A straightforward but time-consuming workflow for the assignment of IR spectra is to compare IR-UV depletion spectra measured at each UV peak reasonably close to the UV band origin. A more economical way is to build, iteratively, a cross-correlation between the IR and UV transitions. One may, for instance, first, fix IR laser on a peak in the gain spectrum and scan wavelength of the subsequent UV laser pulse. Those UV transitions, which

originate only from the conformers excited by IR pulse, all belong to the same conformer(s). A single IR-UV depletion spectrum measured at any of these UV peaks will therefore represent IR spectra of this conformational family. The respective IR and UV transitions can be excluded from the further iteration steps. This procedure can be repeated until all the peaks in both UV and IR depletion spectra will be assigned to one or another conformational family.

Similarly, IR transitions can be assigned with IR-IR-UV hole burning technique, which employs correlations between IR transitions in a gain and in the hole burning IR spectra. Assignment of UV peaks can also be performed using UV-UV hole burning in a similar way,[54, 63-65] although this requires either ejecting a specific fragment ion[63] or detection of a neutral fragment.[64]

## II. Cold Ion Spectroscopy for Validating Structures of Biomolecules

### *2.1 Achievements and Limitations*

A significant volume of data have been generated over last decade by cold ion spectroscopy for validation of structures, analysis of photophysics, non-covalent interactions, etc., of biomolecular ions, including protonated amino acids, peptides and their non-covalent complexes, drug molecules, and carbohydrates. Many of these results have been reviewed in a few older and recent publications.[10, 24, 41, 66] Overall, accurate geometries of low-energy conformers have been solved for several protonated amino acids, midsize peptides and micro-hydrated complexes of some of these biomolecules. To our knowledge, the largest biomolecule, for which the accurate structures of the low-energy conformers have been unambiguously validated, remains decapeptide gramicidin s and its doubly hydrated non-covalent complex (Figure 2.1).[13, 59, 67-71]

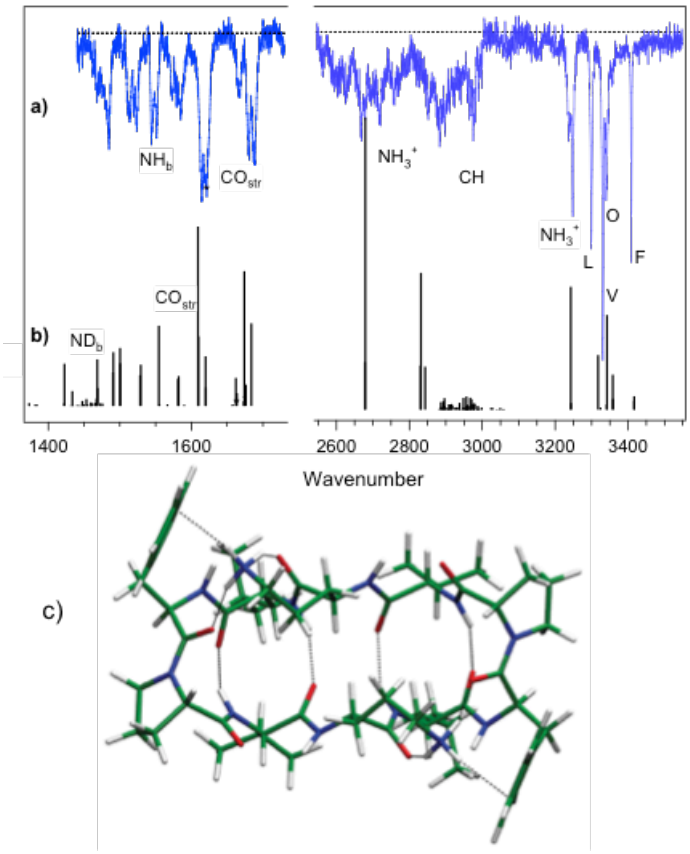


**Figure 2.1** (**a**) IR spectrum of the most abundant conformer of doubly protonated decapeptide gramicidin S, cooled to T=12 K, and (**b**) its theoretical wavenumber-scaled spectrum, calculated for the most stable structure of the peptide (**c**). [13, 59]

These studies revealed, at least, three main limitations of cold ion spectroscopy in solving structures of large biomolecules. On the experimental side, the desired vibrational resolution in IR spectra can be intrinsically limited even in a relatively small species due to anharmonic couplings of vibrational states. Such couplings lead to lifetime broadening of vibrational transitions to these states and it cannot be suppressed. It is worth mentioning that there is no IVR (or perhaps only a very limited IVR) following IR or UV excitation of large molecules by laser pulses of, typically, ns duration. Instead, superpositions of the coupled modes (quantum eigenstates) are excited, such that the excitation samples, eventually, the entire molecule. The currently common use of the IVR term in the frequency-resolved spectroscopy implies this static phenomenon rather than the truly dynamic effect. Anharmonic couplings are particularly strong for the vibrations of the nuclear that are involved to hydrogen bonds. In addition to this, the number of IR transitions near linearly increases with weight of biomolecules. Most of strong characteristic

transitions remain concentrated within limited spectral ranges (e.g., around 3 and 6 μm). Because of this, the increased spectral density of the transitions unavoidably suppresses vibrational resolution in the spectra, making them low informative and, eventually, useless for validating calculated molecular structures.

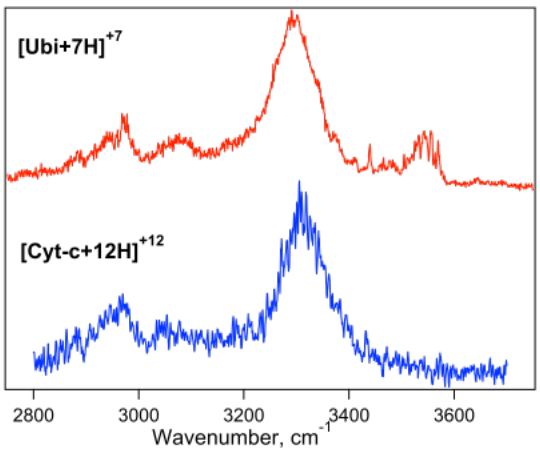


**Figure 2.2**. IR spectra of protonated ubiquitin (+7 state) and cytochrome C (+12 state) proteins, cooled in a 6K cryogenic ion trap.[18]

As an example, Figure 2.2 illustrates spectral congestion in IR spectra for two small protonated proteins: ubiquitin and cytochrome C. Although cooled in a 6 K ion trap, the spectra look almost structurless and, to some extend, as in solution phase. Apparently, the low informational content of such spectra prohibits their use for strict validation of the complex peptide structures. A way to move up in the the size of biomolecules for which IR spectra still can be vibrationally resolved is seen in extending spectroscopy to low-frequency vibrations in the terahertz spectral region.[72] Both IVR and vibrational inhomogeneous broadening effects should become strongly suppressed, when the total vibrational energy of a molecule is comparable with frequencies of its few vibrations only.

Even if many individual vibrational transitions in such large biomolecules were resolved, the practically affordable calculations nowadays would not be able to reproduce the transitions with the required "spectroscopic" accuracy. Although for midsize molecules (e.g., oligopeptides) such calculations might be feasible, the conformational search of a limited pool of candidate structures, to which spectral calculations will be applied, still remains a challenge. Additional, different from IR transitions, spectroscopic structural constraints may narrow the search however. Positions of electronic band origins, measured by UV spectroscopy in aromatic

peptides, may reveal an involvement of a chromophore into non-covalent couplings (e.g., proton-π interactions or hydrogen bonds), constraining conformational search to those calculated geometries, where such interactions are present.[13, 70] FRET measurements can provide an estimate of interchromophore spacing in cold ions,[73] and isotopic labeling allows for assignment of some IR transitions.[13, 70] Additional non-spectroscopic structural filters, such as, for instance, collisional cross-sections, measured by ion mobility spectrometry or photodissociation mass spectra may also help to increase the size of "treatable" biomolecules, but unlikely substantially.

Over years, spectroscopy of both neutral and charged biomolecules accumulated a great number of observations. Tables 1 and 2 summarize and reference many of the validated observations for UV and IR spectroscopy of amino acids and peptides. UV band origins of Trp, Tyr, Phe and His aromatic amino acids (Fig. 1.9) experience different spectral shifts (Tyr, Phe) or spectral broadening (Trp) upon interactions of their aromatic rings with a charge or upon involvement of side-chains of these residues in formation of hydrogen bonds. The characteristic frequencies of, for instance, OH- and NH-stretch vibrations of residues, in certain cases arise within specific spectral intervals, governed by the residue identity, hydrogen bond pattern, etc. The data in Tables 1 and 2 therefore can be use for "intuitive" qualitative structural determinations to guide and narrow a conformational search within calculated candidate structures of peptides.

### 2.2 *Outlook*

Despite the accumulated spectroscopic data and the additional non-spectroscopic constraints, in our opinion, the size of protonated biomolecules, which structure can be solved by cold ion spectroscopy, realistically, is limited to molecules with 10-15 residues only. Several other types of biomolecules, such as, for instance, nucleotides, glycans and drugs that are below this size-limit also can be studied by cold ion spectroscopy for structural elucidations. Nevertheless, for most of the biomolecules, which are much larger in size, this approach cannot provide vibrational spectral resolution, becoming limited in its ability to assist in structural elucidations. Moreover, the relevance of gas-phase structural studies to biology is questionable. The intrinsic geometries of biomolecules, indeed, may not resemble the native geometries of the molecules they adopt in aqueous solution. A simple example to illustrate this problem is the fact that non-cyclic peptides are, typically, protonated on

C-terminus in the gas phase, but adopt zwitterion geometry (with deprotonated C-terminus) in solution phase. Certainly, a charge may crucially influence the geometry of a peptide. This example alone makes elusive the concept of "kinetically trapped" molecular states, which suggests that some conformers of electrosprayed ions may retain their native structures, being trapped in local minima of the gas-phase potential energy surface. For large biomolecules (e.g., proteins), which are hydrated, mainly, on the surface, the overall structural motive can be, indeed, conserved upon desolvation in the gas phase, but the size of proteins becomes prohibitive for using IR spectroscopy in structural determinations. Although this contradiction between the limited size of structurally treatable molecules and the likely irrelevance of intrinsic structures of small molecules to biology limits the power of cold ion spectroscopy, there still several directions, where the method may remain successful. For instance, almost all medical drugs, a good fraction of metabolites, carbohydrates, etc., are well below the foreseen size limit, and, likely, can be studied by cold ion spectroscopy. A reasonable compromise between the accuracy in 3D structural determinations and size of biomolecules seemed to be in studying microhydrated ions in a way, similar to what was demonstrated, for instance, with water clusters of decapeptide gramicidine S (Figure 2.3).[70] This would require an extensive measurements to track molecular geometry by increasing the number of water molecules until the geometry of a biomolecule will begin to converges, presumably, to a solution like structure. Another field where vibrational spectral resolution offered by cryogenic cooling of ions can be quite helpful is analytical identifications of biomolecules. Below, we review these applications of cold ion spectroscopy, which are directly relevant to life-science studies.

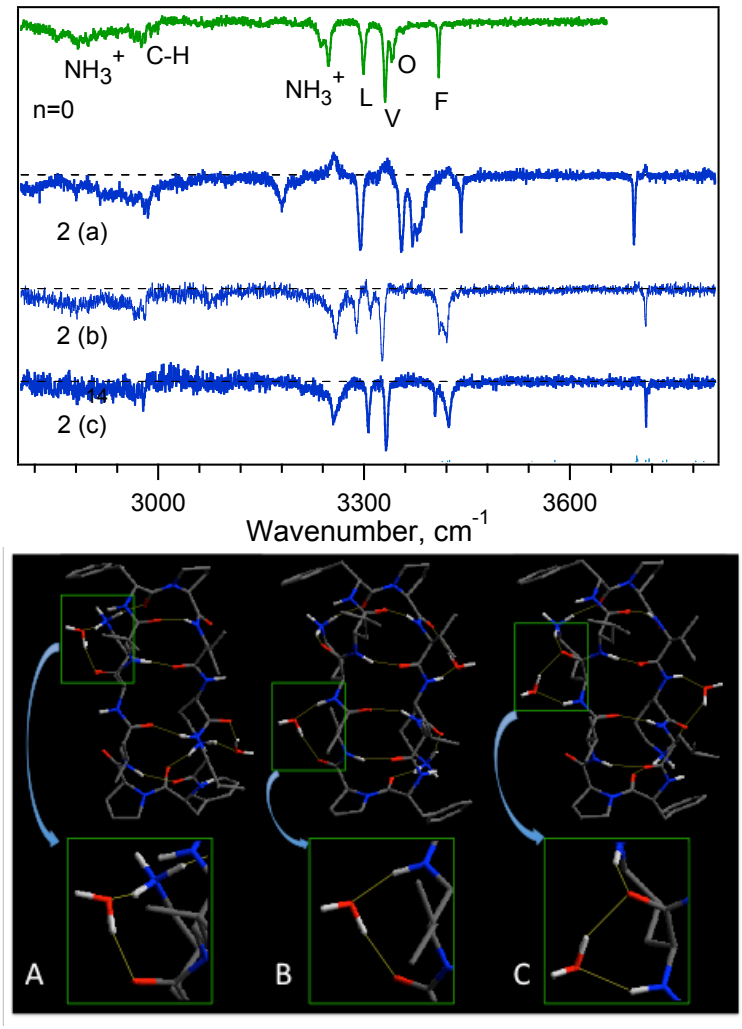


**Figure 2.3** Conformer-specific vibrational spectra of doubly protonated decapeptide gramicidin S retaining two water molecules and cooled to T=12 K (upper panel),[70] and the calculated structures of the three low-energy conformers of the complex (lower panel), for which the computed IR spectra exhibit the best match to the measured spectra.[71]

## III. Analytical Applications of Cold Ion Spectroscopy

### *3.1 2D UV-MS Fingerprinting of Cold Ions.*

Identification of biomolecules is one of the most demanding objectives in life-science research. Because the identity of a molecule in the gas phase and in vivo remains the same, all the powerful methods of the gas phase can be applied for the identifications without compromising the relevance of results to biology. Multistage high-resolution mass spectrometry $(MS)^n$ in combination with different dissociation techniques demonstrates impressively accurate determination of, for instance, peptide and

protein sequences. Because MS is based on measuring mass-to-charge ratio of molecular ions and their fragments, the identification of isobaric molecules, isomers in particular, remains among the challenges of this technique.[74] Coupling of MS to high resolution LC or ion mobility (IM) often enables separation of isomeric ions with sufficiently different retention or arrival time, respectively.[75-77] The separation doesn't yet ensure a reliable identification of the isobars, because Both of these values are not fundamental to molecules and are sensitive to experimental conditions, which may not be always reproduced with sufficient accuracy.[78-79] In contrast, vibronic transitions in isolated molecules and ions reflect differences in energies between their initial and terminal quantum states, making vibrationally resolved electronic spectra unique fundamental fingerprints of such species. Electronic excitation of ions may produce specific to their 3D structure photofragments, which are absent (or low abundant) in the products of thermal-like dissociation.[57, 80] In particular, this phenomenon may lead to a detectable difference between photofragmentation mass spectra of isomeric ions and even between the spectra of conformers of an ion.[57] Prompted by these two observations, we have combined high-resolution UV cold ion spectroscopy with broadband high-resolution mass spectrometry as a new highly specific method for identification of biomolecules, named 2D UV-MS fingerprinting.[11] Compared with the widely used in mass spectrometry UV fragmentation (UVPD),[81] performed at a single VUV/UV wavelength, the fragmentation mass spectra (MS) in 2D UV-MS fingerprinting are recorded at each wavelength in a large spectral range. UV spectra are particularly structured in the regions of electronic band origins of studied cold ions, which adds a great number of details into ionic 2D UV-MS fingerprints. This may allow distinguishing very similar species, including isomers and even conformers of cold ions. To implement 2D UV-MS fingerprinting approach, we have interfaced our cold ion spectroscopy instrument to a commercial Fourier-transform Orbitrap mass spectrometer Exactive, as described above and as illustrated in Figure 1.1.13. The fingerprints (Figure 3.1) are measured by recording fragment mass spectra at each UV laser pulse of a 10 Hz dye laser (Fig. 1.15) while scanning its wavelength. After normalization of the fingerprints to UV laser fluence and to total ion signal, correcting baseline and removing occasional "fake" peaks in the recorded MS, the rectified 2D data arrays (matrices) of isobaric ions are stored as a library for using in identifications of unknown isomeric ions that belong to the library. The rectification

procedure is currently automated using a custom-modified commercial software package (Peak-by-Peak; Spectroswiss Sarl).

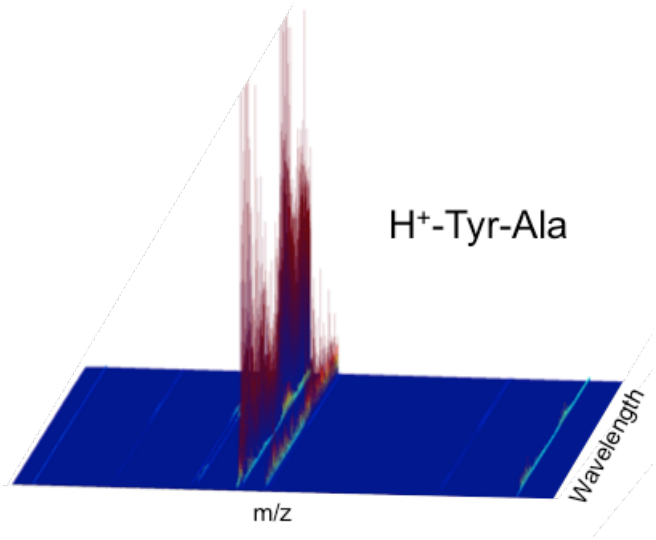


**Figure 3.1**. Photofragmentation UV-MS fingerprint of a protonated dipeptide.[82]

### *3.1.1 Library approach*

***(a) Peptides and drugs.*** The identification and relative quantification of single isobaric ions or their mixtures can be accomplished by a numerical decomposition of the data arrays of the unknown ionic sample in the basis set of the library's matrices.[11, 83-84] Once the difference between the linear combination of the basis set and the tested matrix is minimized, the normalized coefficients of this decomposition yield the relative contributions of the library's species to the analyzed sample. The decomposition procedure is performed using software written in Matlab package.

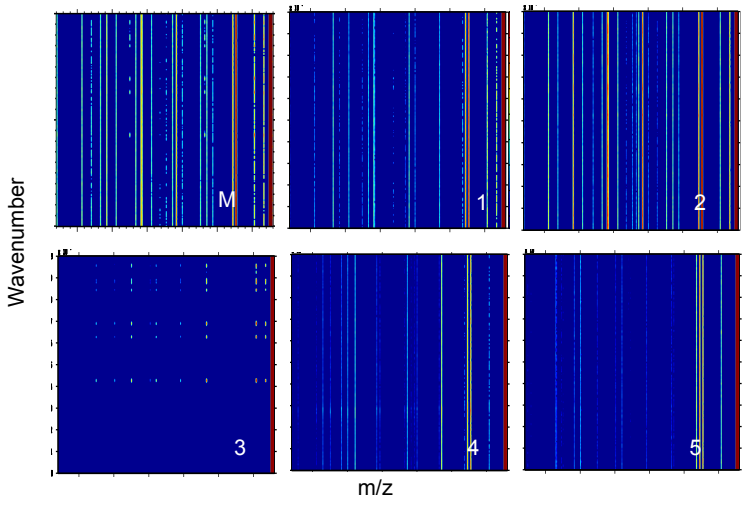

**Figure 3.2**. Color-coded 2D UV-MS photo fragmentation fingerprints of (**1-5**) five singly protonated isomeric octopeptides (T**p**SAAATSY, **p**TSAAATSY, TSAAATS**p**Y, TSAAA**p**TSY and TSAAAT**p**SY, respectively) and (**M**) of a solution mixture of the peptides 3, 4 and 5 with the relative concentrations of 50%, 25% and 25%. The library-based decomposition determines the relative concentrations of the peptides 1-5 as 2.6%, 1.5%, 49%, 23% and 24%, respectively.[11]

Figure 3.2 shows rectified 2D UV-MS fingerprints of five model isomeric phospho-peptides TSAAATSY-$H^+$, which differ only by the phosphorylation site ($Thr_1$, $Ser_2$, $Thr_6$, $Ser_7$ or $Y_8$). A standard UPLC analysis failed to determine the number and the identity of the isomers in their equimolar mixture. In opposite, the measured 2D UV-MS fingerprints visually differ for all five positional isomers. In order to assess the accuracy of identification of the isomers by the new approach, 12 different solution mixtures that contain from one to all five isomers were prepared and their 2D UV-MS were measured. The decomposition of the matrices of these test mixtures in the basis set of the five components correctly identifies the peptides in all the mixtures and reproduces the relative concentrations of the peptides in the solutions with RMSD of 6.5% (Figure 3.3a). This number implies that, in average, any of the five library isomers with relative concentration of > 13% in solution mixtures with other isomers can be identified at the confidence level of 95%.[11]

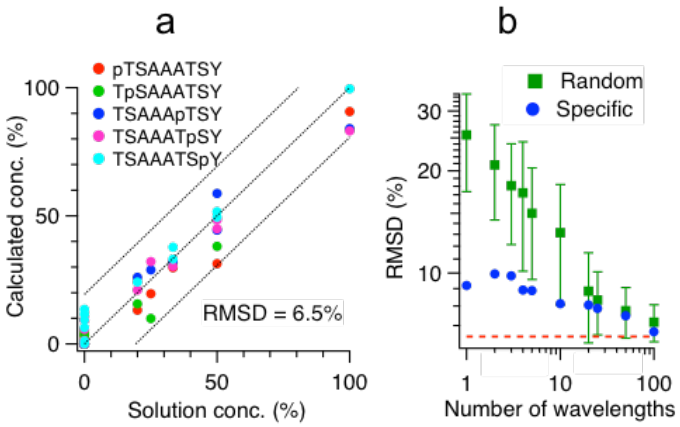


**Figure 3.3.** (**a**) Calculated relative concentrations of five library isomeric phosphorylated singly protonated peptides (dots) as a function of their relative solution concentrations (0−100%) for 12 different solution mixtures. The RMSD of 6.5% was determined taking deviations of the calculated from the solution concentrations (solid line). The two dashed lines show the interval of the 99.7% confidence level (3×RMSD) in calculating concentrations. (**b**) RMSD values calculated for all the mixtures as in part **a** and plotted as a function of the number of wavelengths, retained in matrix decomposition (log scales). Each green dot represents

an average over 100 randomly generated wavelength sets, and the error bar indicates the dispersion of the average for each set. Blue dots are the RMSDs, calculated for the specific wavelengths, related to the band origins of the five library peptides.[11]

An important issue for the method is whether it can be used for online identification, when the time for a fingerprint measurement is limited by the LC peak width to a few to tens of seconds. High-resolution fingerprints of the library components, which we measure at thousands of UV wavelengths on a time scale of an hour, can only be obtained offline. A truncated fingerprint, measured for a mixture only at a few but critical wavelengths may still remain very specific, however. Figure 3.3b illustrates how the RMSD for the tested mixtures depends on the number of wavelengths, retained in the truncated matrices of the mixtures and the library. A random choice of a few wavelengths results in unacceptably high RMSDs, which, however, converge close to the RMSD of the full decomposition for 100 wavelengths. Analysis employing a few but critical wavelengths drastically reduces the RMSD relative to results acquired with randomly selected wavelengths, approaching the RMSD of the full analysis at 10−20 wavelengths. This result demonstrates that the UV-MS fingerprinting, potentially, can be coupled to LC for online identifications in a broad range of applications.

In addition to the identifications, demonstrated with these five position-isomeric octopeptides, a few life-science relevant molecules with different types of isomerization have been studied so far by 2D UV-MS fingerprinting. This includes, for instance, stereoisomers of pain-relief peptide drug enkephaline, diastereomers of stimulant non-peptide drug ephedrine, L/D-Asp and L/D-isoAsp isoforms in a peptide derived from hormon protein amylin, isoforms of mono- to tetra-saccarides, etc. Table 3 summarizes the results in identification of isomeric molecules by 2D UV-MS fingerprinting method.

**Table 3.** Isomers/conformers of biomolecules identified by 2D UV-MS method.

| ISOMERIC MOLECULES | ORIGIN | RMSD, % | REF. |
|---|---|---|---|
| Phosphorylated hexapeptides (Five positional isomers) | Model | **6.5** | [11] |
| Enkephalines pentapeptide (L/D, two stereoisomers) | Opioid Drug | **5.4** | [11] |

| | | | |
|---|---|---|---|
| Ephedrines (165 Da) (+/-, two stereoisomers) | Sport doping | **2.9** (LOD[a]=30 ng) | [83] |
| Heptapeptide $IAPP_{31-37}$ (Six peptides with 4 isoforms: L/D-Asp/ L/D-isoAsp in 2 positions) | Fragment of hormone peptide amylin | **2.2**[b] (<**4**)[c] | [84] |
| Conformers of TyrAla-$H^+$ | Model dipeptide | **0.9**[d] (blind decomposition) | [82] |
| Conformers of decapeptide gramicidin S | Antibiotic | **-** (blind decomposition)[e] | [82] |

[a] Limit of detection. [b] For 12 binary solution mixtures. [c] Maximum error for an equimolar quarterly solution mixture. [d] Based on detection of an isomeric conformer. [e] Qualitative identification only.

**(*b) Oligosaccharides*.** Saccharides are the biomolecules, which are particularly reach in isomers. While currently the method of choice for a large scale screening of glycoproteins is HPLC, the known limitations of this technique drive interest toward development of complementary/alternative methods for identifications of isoforms of glycans. Similar to HPLC in its spirit and limitations, ion mobility-mass spectrometry in different implementations was employed for fast separation of some isoforms of carbohydrates.[85-86] Back in 2006 Polfer demonstrated a use of room temperature IRMPD spectroscopy (performed by a free-electron laser) with mass-specific detection of photofragments for qualitative identifications of disaccharides tagged by $Li^+$, although no quantification of the method performance was reported.[37] Recently, von Helden and Pagel extended the spectroscopic part of this approach to isoforms of tetrasaccharides, cooled to 0.37 K in He droplets.[46] The recorded vibrationally resolved spectra exhibit distinct isomeric specificities,although no quantitative assessment of the method selectivity was reported. Regarding the high technical complexity of both free-electron lasers and helium droplet generation, this approach may remain purely demonstrative without an analytical future. The group of Rizzo recently demonstrated the use of a technically simpler combination of IR spectroscopy, performed by a tabletop IR OPO, with cryogenic cooling in an RF ion trap for identification of $N_2$-tagged mono- to disaccharides. In addition, the tested isomeric saccharides could be pre-separated using IMS.[45, 49] A principal drawback of using weakly tagged complexes for isomeric identifications is the loss of mass spectroscopic information on molecular fragments. Another potential problem for

identifications of isoforms by the tagging approach is a non-zero baseline in the measured depletion IR spectra. This may complicate quantifications of isomeric carbohydrates. Despite all the challenges, the universality of tagging spectroscopy makes it a promising approach for analytical identifications of isomeric biological molecules.

There are two main problems in using 2D UV-MS fingerprinting for identifications of isoforms of saccharides. First, both protonated and deprotonated glycans do not absorb in near UV spectral region. UV absorption of deprotonated saccharides is intrinsically lifetime broaden due to a fast electron photodetachment. VUV absorption bands of protonated glycans, although limited in width, appear with no resolved vibrational structure. The gas phase UV spectra of ionic sugars therefore exhibit low isomeric specificity, making their identifications challenging. Second, carbohydrates exhibit low protonation efficiency even at highly acidic conditions of aqueous solutions. Instead, saccharides readily bind sodium cations, which are usually present in the solutions in sufficient concentrations. Both problems have been solved by adding an aromatic amino acid (AAA) into aqueous solutions of glycans. Already in slightly acid solutions (pH= 3-5) AAAs become protonated and can efficiently bind to the saccharides to produce protonated non-covalent glycan--AAA complexes.[87] The complexes appear to be sufficiently stable to survive under conditions of a mild to moderately harsh electrospray ionization source. Figure 3.4 shows the photodissociation and UV spectra of the isomeric trisaccharides Lewis X and Lewis A bound in aqueous solutions to $TyrH^+$ amino acid. The mass spectra contain all the same major fragments, although of different relative intensities for the different isomers. UV spectra of the isoforms are visually distinct in both the first and the second absorption bands. Regarding our previous experience with decomposition of mixtures of isomeric peptides, such spectral difference alone should allow an identification of the isomers in their mixtures with an accuracy of a few %. Interestingly, the presence of an amino group in the saccharides results in proton transfer from AAA to a non-covalently bound glycan. Because pKa of carbohydrates is lower than that of AAAs,[88-90] this proton transfer may only happen in the gas phase either as a spontaneous UV-induced process.[91-92] A valuable result of this transfer is the cleavage of inter/intra-glycan bonds, resulting in appearance of charge fragments that are characteristic to the isoforms of NAc saccharides. In contrast to the IRPD spectroscopy of tagged saccharides, the produced photofragmentation mass spectra

may additionally assist in structural characterization of glycans under study. It is important that the same instrumentation, workflow and software of the 2D UV-MS method, developed for identification of isomeric peptides and drugs, can be used for similar identifications of saccharides.

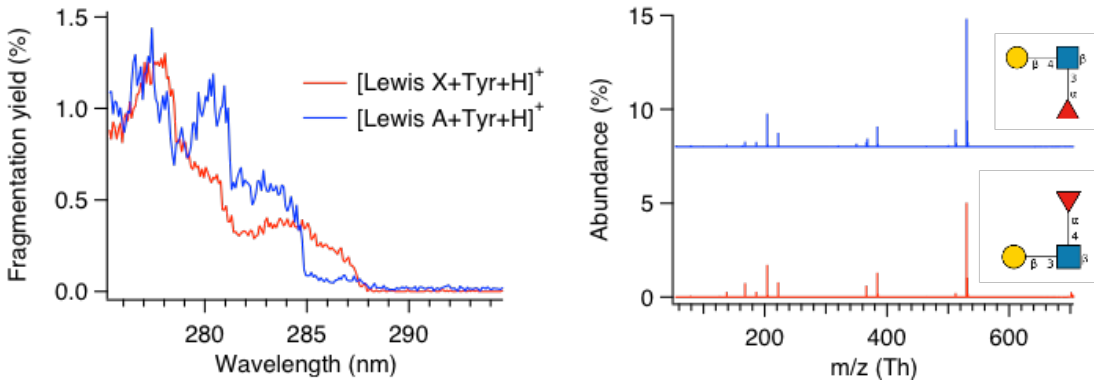


**Figure 3.4.** Photofragmentation mass and UV spectra of isomeric trisaccharides Lewis X and Lewis A complexed with $[Tyr+H]^+$. The spectra are derived from the respective 2D UV-MS fingerprints at 225 nm and 204.087 Th.

### *3.1.2 Blind identifications of isomers and conformers.*

The described above approach to identification of isoforms by 2D UV-MS fingerprinting requires creating a library of all the isoforms/exact isobars that are expected in a process. It may readily happen that such a library doesn't exist yet or it is incomplete. In the latter case a sample may contain a mixture of known library components and one or a few unknown compounds, which cannot be distinguished from the library species by mass (e.g., isomers) and/or by LC (e.g., conformers). One may even don't know, how many different species are present in the mixture. It turns out that 2D UV-MS fingerprinting may detect the presence of these unknown compounds and recover their UV absorption and fragmentation mass spectra, using mathematical algorithms of matrix analysis.[11, 82] The principle condition for such "blind" decomposition of the mixture fingerprint is that both, UV absorption and fragment mass spectra of all the mixed molecules are substantially differ from each other. Mathematically, this implies orthogonality of the 2D UV-MS data matrix of all the compounds, such that none of the matrices is a linear combination of the others. The compounds, whose UV and/or mass spectra are identical, will be determined by the blind decomposition as a single species (e.g., a conformational family).

The workflow of the blind decomposition consist of two main steps.

1. Factorization of the 2D UV-MS data matrix $\boldsymbol{D}$ into a product of a non-negative matrix $\boldsymbol{W}$ of UV spectra and a non-negative matrix $\boldsymbol{H}$ of MS spectra of different components. If $\boldsymbol{k}$ denotes the number of such components, then the $\boldsymbol{n}$-by-$\boldsymbol{m}$ matrix $\boldsymbol{D}$ can be factorized into $\boldsymbol{n}$-by-$\boldsymbol{k}$ matrix $\boldsymbol{W}$ and $\boldsymbol{k}$-by-$\boldsymbol{m}$ matrix $\boldsymbol{H}$, so that their matrix product $\boldsymbol{WH}$ is a lower-rank approximation to $\boldsymbol{D}$.[3] Since all the matrices, $\boldsymbol{D}$, $\boldsymbol{W}$ and $\boldsymbol{H}$, are non-negative, this approach was called non-negative matrix factorization (NMF).[93] One can find the best $\boldsymbol{k}$-rank approximation by minimizing the Frobenius norm of the residual:

$$\|D - W \cdot H\|_F = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{m} \left(D_{ij} - \sum_{l=1}^{k} W_{il} \cdot H_{lj}\right)^2} \qquad (2)$$

2. Use Bi-cross-validation (BCV) procedure[94] to determine the most likely number $\boldsymbol{k}$ of the mixed compounds, which differ by their UV and mass spectra. The general idea of BCV is to hold out a sub-matrix of the matrix $\boldsymbol{D}$ (*i.e.*, a set of elements of $\boldsymbol{D}$) and use the rest of $\boldsymbol{D}$ to predict the held out sub-matrix. The quality of NMF for each $\boldsymbol{k}$ is reported as a BCV error. For the $\boldsymbol{k}$ number below the true number of components, the BCV error is mainly determined by the NMF itself (low accuracy of the $\boldsymbol{k}$-rank approximation), implying that BCV error is close to the RMSD error. When the value of $\boldsymbol{k}$ becomes greater than the true number of components, the NMF fits the noise in the experimental data by the redundant components. However, the NMF of the sub-matrix $\boldsymbol{X}$ is used to estimate the sub-matrix $\boldsymbol{D_{ij}}$. The overfitting of $\boldsymbol{X}$ does not lead to a better approximation to $\boldsymbol{D_{ij}}$, because the noise functions in these two sub-matrices do not correlate. Thus, the large values of $\boldsymbol{k}$ will not result in a reduced BCV error, while the RMSD error monotonically decreases upon increasing $\boldsymbol{k}$.

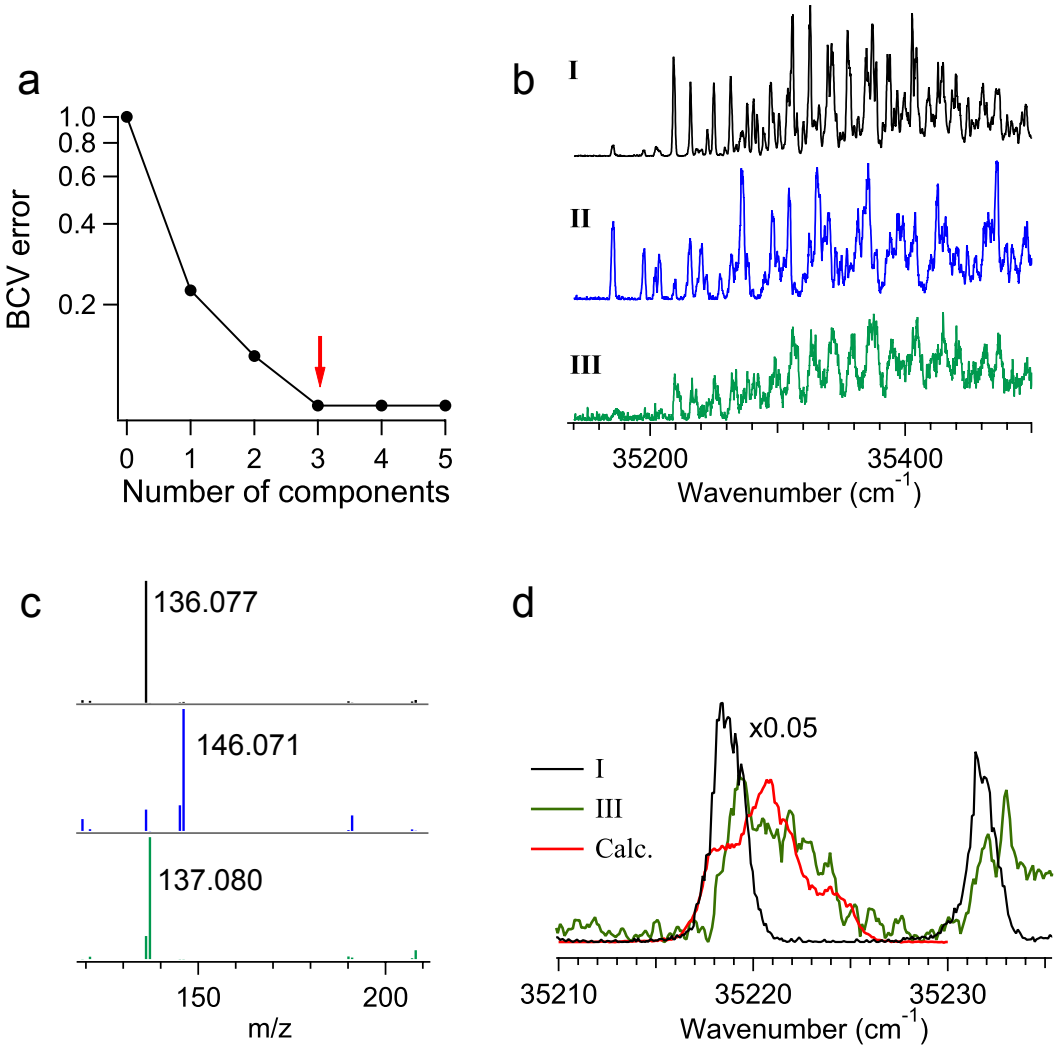


**Figure 3.5** (**a**) Plot of BCV error (log scale) for 2D UV-MS spectrum of singly protonated Tyr-Ala dipeptide (Fig. 3.1); the error reaches a minimum for three components of factorization. (**b**) The derived UV absorption spectra normalized by the relative contributions of the respective components to the UV-MS matrix (81%, 12.5% and 6.5% for the component I, II and III, respectively). (**c**) The derived photofragmentation mass spectra are normalized by the intensities of their largest peaks. (**d**) UV spectra around the band origins of the conformers I and III, and the calculated spectrum of the conformer I with single $^{13}C$ substitution. Reproduced from ref.[82]

Figure 3.3 illustrates the application of the blind decomposition approach to a dipeptide $[YA+H]^+$, which 2D UV-MS fingerprint is shown in Figure 3.1. The BCV analysis of the fingerprint (Fig. 3.5a) distinguishes 3 main components. The factorization of the fingerprint with $k$=3 yields the UV absorption and mass spectra

for each of them. The derived UV spectra of the components, denoted in Fig. 3.5b as I and II, each contains all the UV transitions, previously assigned only to one or to the other pair of conformers, respectively.[57] It turned out that the 3rd derived component is an isotopologue of the conformer I with one $^{13}C$ in the tyrosine immonium fragment. Indeed, the most abundant peaks in the otherwise similar fragment mass spectra of the components I and III (Fig. 3.5c) are shifted exactly by the mass difference of the $^{12}C$ and $^{13}C$. Second, the determined relative contribution of the component III (8%) is close to the abundance of the Y-immonium ion with one $^{13}C$ (8.9%). The spectroscopic evidence of our assignment arises from the tiny shifts of UV transitions (Fig. 3.5d) in the component III, caused by the change of the zero-point vibrational energies (ZPVE) of $S_0$ and $S_1$ electronic states upon the isotopic substitution. The identification of the isotopic component III by the blind decomposition illustrates the high power and potential of the method, which can be helpful, for instance, in identifications of unknown isomeric/isobaric metabolites.

**IV. Conclusions and Prospective**

The initial motivation for cryogenic cooling of large molecular ions was in providing their detailed spectroscopic signatures, suitable for stringent validations of structural calculations through comparisons with computed vibrational spectra. Technically, the cooling was achieved by interfacing electrospray ionization sources with 22-pole linear ion traps, which were originally developed for storage and spectroscopy of small ions, relevant to astrochemistry. Low concentration of trapped ions prohibits detection of direct absorption for spectroscopy in the traps. Instead, a photodissociation action spectroscopy, combined with high sensitive mass-spectrometric detection of the UV-induced charged fragments, was used for vibrationally resolved electronic spectroscopy of protonated amino acids and small to midsize peptides. Implementation of IR-UV depletion spectroscopy, earlier developed for neutrals, enabled conformer specific vibrational spectroscopy of biomolecules and their non-covalent clusters, although limited only to the species that exhibit vibrationally resolved UV spectra. This “classical” approach to solving intrinsic structures was successfully demonstrated on Phe and Tyr aromatic amino acids and on several peptides that contained one of these UV chromophores. Application of IR-UV depletion to protonated Trp amino acid and peptide with this

residue was more challenging, due to lack of vibrational resolution in UV spectra. Unstructured UV spectra appeared as often problem for mid- to large size peptides with Tyr and, to a lesser extend, with Phe as chromophores. With the advent of implementation of IR-IR-UV triple resonance hole burning technique, conformer specific vibrational spectroscopy was recently extended to the peptides that exhibit moderately broad structurless UV absorptions. Also recently, His was added to this toolbox as a chromophore.[16] Finally, conformer-selective vibrational spectroscopy was demonstrated on the peptides that do not have any aromatic residues, but the naturally present peptide bond UV absorption was employed instead for photofragmentation of protonated peptides.

With all these recent additions, conformer specific IR spectra can nowadays be measured, essentially, for any protonated small to midsize peptide, regardless of the UV spectral complexity. There are two main factors that limit further extension of CIS to larger peptides. First, coexistence of several abundant conformers multiplies the number of IR transitions. To certain extentthis problem can be solved by internal heating of ions followed by slow adiabatic cooling to cryogenic temperatures. Such annealing should concentrate the ions essentially, in single most stable geometry. An alternative way is a conformational pre-separation of ions, using, for instance, ion mobility as a filter. An intrinsic limitation of vibrational spectral resolution, which cannot be overcome, is, however, the increased number of IR transitions in large molecules. Regarding our experience and the body of the published data, conformer-specific IR spectroscopy seems feasible for peptides with less than 10-15 residues only. This limit, potentially, can be pushed up by moving to terahertz spectral region.

Even for smaller peptides conformational search is often a time bottleneck in structural computations. Here the experiment may help by providing additional to IR spectra structural constraints, which can confine the search. Standard techniques like isotopic labeling, substitution of residues may provide assignments of certain IR transitions. UV spectra may also give some hints, in particular for involvement of aromatic rings to non-covalent interactions. Collisional cross-sections, measured with conformational selectivity may additionally facilitate the search.

Certain structural constraints can be revealed from photofragment mass spectra too. Although initially MS was considered mainly as a detection step, it became clear soon that photofragmentation MS patterns might correlate with

conformational state of molecules. There is no yet a simple straightforward workflow to reveal and rationally use these correlations, however.

Overall, over these years the toolbox of cold ion spectroscopy has become sufficiently comprehensive for its application to a variety of problems related to the intrinsic structures and photophysics of up to midsize peptides. Further developments are viewed in extending the wavelength range down to terahertz spectral region and in extensive applications of the method to other types of biomolecules, such as glycans, oligonucleotides, non-peptide drugs, etc.

The main and fundamental problem of cold ion spectroscopy remains, however in the obscure relevance of the gas-phase structures to biology, which largely deals with biomolecules in solution phase. Although desolvation may still leave intact the main structural motif for large (e.g., proteins) biomolecules, the intrinsic and the native structures of peptides are likely very different. In this respect, the only prospective of the CIS is viewed in structural elucidation of biomolecules, complexed with a few water molecules. The number of solvent molecules should be sufficient to preserve the main features of native structures, while the size of the clusters still has to be computationally treatable.

The frequent correlation of the conformational/isomeric form of biomolecules with their UV photofragmentation MS led us to the idea of using such correlations for analytical identifications of biomolecules. Although structures of a molecule in solution and in the gas phase are often quite different, the identifications still can be performed in the gas phase, where one may benefit from high spectral and mass resolutions. This approach was developed as the 2D UV-MS fingerprinting method, which was demonstrated for identifications of isoforms of peptides, drug molecules and oligosaccharides. We see further development of this and similar (e.g. IR-MS) approaches as quantitative analytical methods for identifications of different types of biomolecules. From a practical point of view, one of the current objectives is to make such methods compatible with online LC, which implies a few second timescale for interrogation of a sample. In parallel, one would need to create libraries of targeted fingerprints, which can refer to specific problems, challenging for HPLC-$MS^n$.

Finally, in the forthcoming years there can be many new unexpected developments, based on a simple fact that cooling may make spectra of biomolecules vibrationally resolved. Despite all the challenges of this still infant field, cold ion spectroscopy has a great potential in providing structural constraints for strict

validation of intrinsic molecular structures. Analytical applications of this wonderful technique are currently under development to solve the emerging problems of life science.

## V. Acknowledgments

OB thanks Dr. N. Nagornova, Dr. V. Kopysov and Dr. A. Pereverzev for their hard work in obtaining some of the presented herein data, Dr. V. Kopysov for collecting preliminary information for Tables 1 and 2; Dr. A Makarov and Yu. Tsybin for many fruitful discussions. This work was supported by Swiss National Science Foundation (grants 200020_172522 and 206021_164101), by Novartis AG (FreeNovation program) and by EPFL.

**Table 1.** Gas-phase UV absorption by aromatic amino acids/residues

| Chromophore | Molecule/ ion | Band origin, cm-1 | Characteristic vibronic freg., cm-1 | Ref. |
|---|---|---|---|---|
| His | His-H+ | 42920 | | [16] |
| Phe | cyclo-VOLFPVOLFP-2H+ (gramicidin S conf. A) | 37707.3 | | [70] |
| Phe | $[(H_2O)_2$ - cyclo-VOLFPVOLFP-2H]+2 | 37619.3 | | [70] |
| Phe | $[(H_2O)_8$ - cyclo-VOLFPVOLFP-2H]+2 | 37578.8 | | [70] |
| Phe | $[(H_2O)_{14}$- cyclo-VOLFPVOLFP-2H]+2 | 37610.3 | | [70] |
| Phe | $[(H_2O)_{30}$- cyclo-VOLFPVOLFP-2H]+2 | 37568 | | [70] |
| Phe | cyclo-VOLFPVOLFP+2H | 37528 | | [59] |
| Phe | AcPheNHMe | 37650 | | [95] |
| Phe | $AcGlyPheNH_2$ | 37635 | | [96] |
| Phe | $AcPheGlyNH_2$ | 37620 | | [96] |
| Phe | Phe | 37613 | | [97] |
| Phe | $AcPheNH_2$ | 37610 | | [96] |
| Phe | Phe | 37600 | | [97] |
| Phe | AcPheNHMe | 37593 | | [95] |
| Phe | AcPheNHMe | 37592 | | [95] |
| Phe | AcPheOMe | 37579 | 537 | [98] |
| Phe | Phe | 37570 | | [97] |
| Phe | $AcFA_{10}K$-H+ | 37568.2 | | [99] |
| Phe | $AcFA_5K$-H+ | 37560.3 | | [99] |
| Phe | Phe | 37558 | | [97] |
| Phe | $AcFA_5K$-H+ | 37541.9 | | [99] |
| Phe | Phe | 37537 | | [99] |
| Phe | $(AcPheOMe)_2$ | 37537 | 538 | [98] |
| Phe | $AcAAFNH_2$ | 37535 | | [100] |
| Phe | $AcFA_5K$-H+ | 37531.6 | | [99] |
| Phe | $AcPheNH_2$ | 37530 | | [96] |
| Phe | Phe-H+ | 37529.6 | 531 | [58] |
| Phe | $AcFA_{10}K$-H+ | 37527.7 | | [99] |
| Phe | $AcFA_5K$-H+ | 37525.6 | | [99] |
| Phe | Phe-H+ | 37520 | 531 | [58] |
| Phe | AcPheNHMe | 37518 | | [95] |
| Phe | $FA_{10}K$-H+ | 37512 | | [101] |
| Phe | $AcGlyPheNH_2$ | 37510 | | [96] |
| Phe | YGGFL-H+ | 37507 | 531 | [73] |
| Phe | $AcPheNH_2$ | 37505 | | [96] |
| Phe | YAGFL-H+ | 37504 | 532 | [73] |
| Phe | FG | 37490 | | [102] |
| Phe | $AcAFANH_2$ | 37490 | | [100] |
| Phe | $AcPheGlyNH_2$ | 37485 | | [96] |
| Phe | $FA_{10}K$+2H | 37480 | | [101] |
| Phe | $AcFAANH_2$ | 37425 | | [100] |
| Phe | AcPheNHMe | 37414 | | [95] |
| Phe | RPPGFSPFR+2H (bradykinin) | 37359 | | [103] |
| Phe | Y-dAGF-dL-H+ | 37261 | 534 | [73] |
| Phe | RPPGFSPFR+2H (bradykinin) | 37224 | | [103] |
| phospho-Tyr | $Ac\text{-}pYA_3SK$-H+ | 37170 | | [80] |
| phospho-Tyr | $Ac\text{-}SA_3pYK$-H+ | 37167 | | [80] |

| | | | | |
|---|---|---|---|---|
| phospho-Tyr | Ac-pYV$_3$SK-H+ | 36950 | | [80] |
| phospho-Tyr | pYA-H+ | 36945 | | [80] |
| phospho-Tyr | ApY-H+ | 36930 | | [80] |
| phospho-Tyr | ADEpYLIPQQ+2H | 36295 | | [80] |
| Tyr | Tyr | 35687 | 803 | [102] |
| Tyr | YAGFL-H+ | 35684 | 810 | [73] |
| Tyr | YGGFL-H+ | 35683 | 810 | [73] |
| Tyr | Tyr | 35492 | | [97] |
| Tyr | Tyr | 35485 | 809 | [102] |
| Tyr | NVGS(*D*-Asp)TY-NH$_3$+ | 35410 | | [84] |
| Tyr | AY | 35404 | | [102] |
| Tyr | NVGS(*D*-isoAsp)TY-NH$_3$+ | 35014 | | [84] |
| Tyr | Y-dAGF-dL-H+ | 35399 | 810 | [73] |
| Tyr | Ac-YA$_3$Sk-H+ | 35398 | | [80] |
| Tyr | GY | 35387 | | [102] |
| Tyr | YA | 35345 | | [102] |
| Tyr | YA-H+ | 35317.1 | | [57] |
| Tyr | Ac-YA$_3$pTK-H+ | 35307 | | [80] |
| Tyr | NVGS(*L*-Asp)TY-NH$_3$+ | 35298 | | [84] |
| Tyr | YA-H+ | 35277.2 | | [57] |
| Tyr | Tyr-H+ | 35234.9 | | [58] |
| Tyr | YA-H+ | 35223.2 | | [57] |
| Tyr | Tyr-H+ | 35186.4 | | [58] |
| Tyr | YA-H+ | 35174.8 | | [57] |
| Tyr | Ac-YA$_3$pSK-H+ | 35140 | | [80] |
| Tyr | Tyr-H+ | 35111.4 | | [58] |
| Tyr | Tyr-H+ | 35083 | | [12] |
| Tyr | Tyr-H+ | 35081.3 | | [58] |
| Trp | TrpH+ | 34700 | ~750 | [12] |
| Tyr | AY-H+ | 34525.2 | | [57] |

**Table 2.** Experimentally observed vibrations in peptides/amino acids.

| Group | Type | Residue | Exper. freq., $cm^{-1}$ | Calc. freq., $cm^{-1}$ | Inter-action | Molecule | Comment | Ref. |
|---|---|---|---|---|---|---|---|---|
| OH | stretch | Phospho | 3670 | | free | Ac-pYA3SK+H | | [80] |
| OH | stretch | Phospho | 3669 | | free | pYA+H | | [80] |
| OH | stretch | C-term | 3669 | | free | pYA+H | | [80] |
| OH | stretch | Phospho | 3654 | | free | pYA+H | | [80] |
| OH | stretch | Tyr | 3649 | | free | Ac-YA3pSK+H | | [80] |
| OH | stretch | Tyr | 3649 | | free | Ac-YA3Sk+H | | [80] |
| OH | stretch | Tyr | 3649 | | free | Ac-YA3pTK+H | | [80] |
| OH | stretch | Tyr | 3645.2 | 3647.6 | free | Tyr+H | Conf C | [58] |
| OH | stretch | Tyr | 3644.2 | 3647.4 | free | Tyr+H | Conf D | [58] |
| OH | stretch | Tyr | 3643.5 | 3647.3 | free | Tyr+H | Conf B | [58] |
| OH | stretch | Tyr | 3643.2 | 3647.8 | free | Tyr+H | Conf A | [58] |
| OH | stretch | Tyr | 3642.3 | 3640.8 | free | YA+H | Conf D | [57] |
| OH | stretch | Tyr | 3641.8 | 3641.4 | free | YA+H | Conf C | [57] |
| OH | stretch | Tyr | 3641.7 | 3641.7 | free | YA+H | Conf A | [57] |
| OH | stretch | Tyr | 3641.7 | 3642.2 | free | YA+H | Conf B | [57] |
| OH | stretch | Tyr | 3637.9 | 3638.85 | free | AY+H | | [57] |
| OH | stretch | C-term | 3588.5 | | free | WG | Conf A | [104] |
| OH | stretch | C-term | 3585 | | | WGG | Conf B | [104] |
| OH | stretch | C-term | 3583 | | free | GW | Conf B | [104] |
| OH | stretch | C-term | 3582 | | free | PW | Conf B | [105] |
| OH | stretch | C-term | 3581 | | free | Phe | Band A | [106] |
| OH | stretch | C-term | 3579 | | free | Phe | Band D | [106] |
| OH | stretch | C-term | 3575.3 | | free | AcFA10K+H | Conf A | [99] |
| OH | stretch | C-term | 3574.1 | | free | FA10K+H | | [101] |
| OH | stretch | C-term | 3574 | | free | SarSar+H | H2-tagged | [107] |
| OH | stretch | C-term | 3573 | | free | Ac-YA3pSK+H | | [80] |
| OH | stretch | C-term | 3573 | | free | Ac-YA3pTK+H | | [80] |
| OH | stretch | C-term | 3571.2 | | free | AcFA5K+H | Conf A | [99] |
| OH | stretch | C-term | 3570 | | free | SarGly+H | H2-tagged | [107] |
| OH | stretch | C-term | 3569 | | free | GG+H | H2-tagged | [107] |
| OH | stretch | C-term | 3569 | | free | GlySar+H | H2-tagged | [107] |
| OH | stretch | C-term | 3569 | | free | His+H | Conf A/A* | [16] |
| OH | stretch | C-term | 3568.3 | 3568.9 | free | YA+H | Conf A | [57] |
| OH | stretch | C-term | 3567.9 | 3568 | free | YA+H | Conf C | [57] |
| OH | stretch | C-term | 3567 | | free | Phe | Band C | [57] |
| OH | stretch | C-term | 3566.8 | 3566.8 | free | YA+H | Conf B | [57] |
| OH | stretch | C-term | 3566.8 | 3566.9 | free | YA+H | Conf D | [57] |
| OH | stretch | C-term | 3566 | 3566 | free | GG+H | H2-tagged | [40] |
| OH | stretch | C-term | 3564.5 | 3563.9 | free | AY+H | | [57] |
| OH | stretch | C-term | 3564 | | free | His+H | Conf B | [16] |
| OH | stretch | C-term | 3563 | | free | SarSar+H | H2-tagged | [107] |
| OH | stretch | C-term | 3561 | | free | GlySar+H | H2-tagged | [107] |
| OH | stretch | C-term | 3554.9 | 3560.15 | free | Tyr+H | Conf C | [58] |
| OH | stretch | C-term | 3554 | 3552.8 | free | Phe+H | Conf B | [58] |
| OH | stretch | C-term | 3553.65 | 3560.5 | free | Tyr+H | Conf A | [58] |
| OH | stretch | C-term | 3553.2 | 3558.6 | free | Tyr+H | Conf D | [58] |
| OH | stretch | C-term | 3552.7 | 3558.6 | free | Tyr+H | Conf B | [58] |
| OH | stretch | C-term | 3552.6 | 3550.6 | free | Phe+H | Conf A | [58] |
| NH | stretch | NH2 | 3534 | 3535 | free | AcAFANH2 | | [108] |
| NH | stretch | NH2 | 3524 | 3526 | free | AcAAFNH2 | | [108] |
| NH | stretch | NH2 | 3523 | 3527 | free | AcFGNH2 | Conf A | [96] |
| NH | stretch | Trp | 3522 | | | GW | Conf B | [104] |
| NH | stretch | NH2 | 3522 | 3519 | free | AcFAANH2 | | [108] |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| NH | stretch | Trp | 3520 | | free | GW | Conf A | [104] |
| NH | stretch | Trp | 3520 | | free | WGG | Conf B | [104] |
| NH | stretch | Trp | 3520 | | free | PW | Conf B | [105] |
| NH | stretch | NH2 | 3520 | 3516 | free | AcGFNH2 | Conf A | [96] |
| NH | stretch | Trp | 3519.5 | | free | WG | Conf A | [104] |
| NH | stretch | Trp | 3519 | | free | PW | Conf A | [105] |
| NH | stretch | NH2 | 3519 | 3519 | free | AcFGNH2 | Conf B | [96] |
| OH | stretch | C-term | 3506 | | bound | Ac-pYA3SK+H | | [80] |
| NH | stretch | imidazole | 3482 | | free | His+H | Conf A/A*/B | [16] |
| NH | stretch | Amide | 3467 | | free | AcPheNHMe | Conf I | [95] |
| NH | stretch | Amide | 3465 | | free | AcPheNHMe | Conf II | [95] |
| NH | stretch | Amide | 3458 | | free | AcPheOMe | Monomer | [98] |
| NH | stretch | Amide | 3454.6 | | free | AcFA10K+H | Conf A | [99] |
| NH | stretch | Amide | 3451.7 | | free | FA10K+H | | [101] |
| NH | stretch | Amide | 3449.65 | | free | AcFA5K+H | Conf A | [99] |
| NH | stretch | Amide | 3449 | 3457 | bound | AcFGNH2 | Conf A | [96] |
| NH | stretch | Amide | 3447 | 3441 | bound | AcFAANH2 | | [108] |
| NH | stretch | Amide | 3444 | 3432 | C5 | AcFGNH2 | Conf A | [96] |
| NH | stretch | Amide | 3442 | 3455 | bound | AcFGNH2 | Conf B | [96] |
| NH | stretch | Amide | 3440 | 3452 | bound | AcAAFNH2 | | [108] |
| NH | stretch | Amide | 3438 | | free | AcPheNHMe | Conf I | [95] |
| NH | stretch | Amide | 3438 | | free | AcPheNHMe | Conf III | [95] |
| NH | stretch | Amide | 3434.3 | | C5 | AcFA5K+H | Conf A | [99] |
| NH | stretch | Amide | 3429 | 3442 | bound | AcAFANH2 | | [108] |
| NH | stretch | N-term | 3428 | | | Phe | Band X | [106] |
| NH | stretch | Amide | 3426 | | free | AcPheNHMe | Dimer | [95] |
| NH | stretch | Amide | 3423 | 3439 | C5 | AcFAANH2 | | [108] |
| NH | stretch | Amide | 3420 | | | WG | Conf A | [104] |
| NH | stretch | N-term | 3420 | | | Phe | Band B | [106] |
| NH | stretch | Amide | 3417.5 | 3416.9 | bound | AY+H | | [57] |
| NH | stretch | Phe | 3416.3 | | | cyclo-VOLFPVOLFP +2H | Conf. B | [109] |
| NH | stretch | N-term | 3411 | | | Phe | Band B | [106] |
| NH | stretch | N-term | 3410 | | free | His+H | Con A/A* | [16] |
| NH | stretch | Phe | 3409 | 3417 | free | cyclo-VOLFPVOLFP+2H | Conf A | [13] |
| NH | stretch | NH2 | 3408 | 3410 | C10 | AcAFANH2 | | [108] |
| NH | stretch | N-term | 3405 | | | Phe | Band C | [106] |
| NH | stretch | Amide | 3402 | | bound | AcFA10K+H | Conf A | [99] |
| NH | stretch | N-term | 3402 | | | Phe | Band D | [106] |
| NH | stretch | Phe | 3399 | | | cyclo-VOLFPVOLFP+2H | Conf B | [109] |
| NH | stretch | Amide | 3396 | | | GW | Conf B | [104] |
| NH | stretch | Amide | 3395 | | | WGG | Conf B | [104] |
| NH | stretch | Amide | 3390.5 | | bound | FA10K+H | | [101] |
| NH | stretch | Amide | 3389 | | β-sheet | AcPheOMe | Dimer | [98] |
| NH | stretch | NH2 | 3389 | 3391 | C10 | AcAAFNH2 | | [108] |
| NH | stretch | Amide | 3388 | | | WGG | Conf B | [104] |
| NH | stretch | NH2 | 3388 | 3391 | C10 | AcGFNH2 | Conf B | [96] |
| NH | stretch | Amide | 3385.8 | 3381.3 | bound | YA+H | Conf B | [57] |
| NH | stretch | Amide | 3383.1 | 3379.6 | bound | YA+H | Conf D | [57] |
| NH | stretch | NH2 | 3381 | 3378 | C7 | AcFGNH2 | Conf A | [96] |
| NH | stretch | N-term | 3377 | | free | His+H | Conf B | [16] |
| NH | stretch | N-term | 3375 | | free | GlySar+H | H2-tagged | [107] |
| NH | stretch | Amide | 3374 | 3389 | C10 | AcAFANH2 | | [108] |
| NH | stretch | N-term | 3373.45 | 3369.5 | free | YA+H | Conf D | [57] |
| NH | stretch | N-term | 3373 | | free | GG+H | $H_2$-tagged | [107] |
| NH | stretch | Amide | 3372.1 | 3371.8 | bound | YA+H | Conf A | [57] |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| NH | stretch | Amide | 3372.1 | 3371 | bound | YA+H | Conf C | [57] |
| NH | stretch | N-term | 3371.7 | 3368.7 | free | YA+H | Conf B | [57] |
| NH | stretch | Amide | 3370 | | free | SarGly+H | $H_2$-tagged | [107] |
| NH | stretch | N-term | 3370 | | | Phe | Band C | [106] |
| NH | stretch | Amide | 3369 | | bound | PW | Conf B | [105] |
| NH | stretch | Amide | 3367.2 | | C10/C13 | AcFA10K+H | Conf A | [99] |
| NH | stretch | Amide | 3363 | | α–helix | FA10K+H | | [101] |
| NH | stretch | Amide | 3363 | | β-sheet | AcPheOMe | dimer | [98] |
| NH | stretch | Amide | 3362.3 | | C10/C13 | AcFA5K+H | Conf A | [99] |
| NH | stretch | Amide | 3362 | | bound | AcFNHMe | Conf II | [95] |
| NH | stretch | NH2 | 3362 | 3364 | C7 | AcFAANH2 | | [108] |
| NH | stretch | Amide | 3360 | | bound | GG+H | $H_2$-tagged | [107] |
| NH | stretch | Amide | 3359.9 | | C10/C13 | AcFA10K+H | Conf A | [99] |
| NH | stretch | N-term | 3359 | 3372 | free | GG+H | $H_2$-tagged | [40] |
| NH | stretch | Amide | 3359 | 3357 | bound | GG+H | $H_2$-tagged | [40] |
| NH | stretch | Amide | 3359 | 3342 | C7 | AcGFNH2 | Con A | [96] |
| NH | stretch | N-term | 3357.85 | 3352.1 | free | Phe+H | Con B | [58] |
| NH | stretch | N-term | 3356.2 | 3360 | free | Tyr+H | Conf C | [58] |
| NH | stretch | Amide | 3354.8 | | α-helix | FA10K+H | | [101] |
| NH | stretch | N-term | 3352.9 | 3359.2 | free | Tyr+H | Conf A | [58] |
| NH | stretch | N-term | 3351.8 | 3350.2 | free | YA+H | Conf C | [57] |
| NH | stretch | Amide | 3351.3 | | C10/C13 | AcFA5K+H | Conf A | [99] |
| NH | stretch | N-term | 3351 | | | Phe | Band B | [106] |
| NH | stretch | Amide | 3350.6 | | a-helix | FA10K+H | | [101] |
| NH | stretch | Amide | 3350 | | bound | AcPheNHMe | Conf III | [95] |
| NH | stretch | N-term | 3349.7 | 3350 | free | YA+H | Conf A | [57] |
| NH | stretch | N-term | 3349.1 | 3347.6 | free | AY+H | | [57] |
| NH | stretch | Amide | 3346.4 | | α-helix | AcFA10K+H | Conf A | [99] |
| NH | stretch | Amide | 3346 | | α-helix | FA10K+H | | [101] |
| NH | stretch | N-term | 3342 | | | Phe | Band A | [106] |
| NH | stretch | N-term | 3342 | | | Phe | Band D | [106] |
| NH | stretch | Amide | 3341.4 | | α-helix | FA10K+H | | [101] |
| NH | stretch | N-term | 3341.13 | 3343.3 | free | Tyr+H | Conf D | [58] |
| NH | stretch | N-term | 3341 | | | Phe | BandX | [106] |
| NH | stretch | Orn | 3340 | 3358 | free | cyclo-VOLFPVOLFP+2H | Conf A | [13] |
| NH | stretch | N-term | 3339.95 | 3334.6 | free | Phe+H | Conf A | [58] |
| NH | stretch | Amide | 3338.6 | | α-helix | AcFA10K+H | Conf A | [99] |
| NH | stretch | N-term | 3338.1 | 3343 | free | Tyr+H | Conf B | [58] |
| NH | stretch | Amide | 3338 | 3360 | C7 | AcFAANH2 | | [108] |
| NH | stretch | Amide | 3337 | | | GW | Conf A | [108] |
| NH | stretch | N-term | 3335 | | free | His+H | Conf A/A* | [16] |
| NH | stretch | Amide | 3333.4 | | α-helix | AcFA10K+H | Conf A | [99] |
| NH | stretch | N-term | 3333 | | free | SarSar+H | $H_2$-tagged | [107] |
| NH | stretch | Amide | 3330.2 | | α-helix | FA10K+H | | [101] |
| NH | stretch | Val | 3330 | 3342 | bound | cyclo-VOLFPVOLFP+2H | Conf A | [13] |
| NH | stretch | N-term | 3328 | | free | SarGly+H | $H_2$-tagged | [107] |
| NH | stretch | Amide | 3325.6 | | α-helix | AcFA10K+H | Conf A | [99] |
| NH | stretch | N-term | 3319 | | free | GG+H | $H_2$-tagged | [107] |
| NH | stretch | N-term | 3319 | | free | GlySar+H | $H_2$-tagged | [107] |
| NH | stretch | NH2 | 3319 | 3358 | C7 | AcFGNH2 | Conf A | [96] |
| NH | stretch | Amide | 3318.4 | | C10 | AcFA5K+H | Conf A | [99] |
| NH | stretch | Amide | 3318 | | β-sheet | AcPheNHMe | Dimer | [95] |
| NH | stretch | N-term | 3316 | | free | His+H | Conf B | [16] |
| NH | stretch | N-term | 3309 | 3303 | free | GG+H | $H_2$-tagged | [40] |
| NH | stretch | Amide | 3304 | | γ-turn | PW | Conf A | [105] |
| NH | stretch | Amide | 3303.1 | | C10 | FA10K+H | | [101] |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| NH | stretch | Amide | 3303 | 3324 | C7 | AcAAFNH2 | | [100] |
| NH | stretch | Amide | 3302.1 | | C10 | AcFA5K+H | Conf A | [99] |
| NH | stretch | Amide | 3299 | | β-sheet | cyclo-VOLFPVOLFP+2H | Conf A | [69] |
| NH | stretch | Leu | 3299 | 3317 | bound | VOLFPVOLFP+2H | | [13] |
| NH | stretch | Amide | 3293.3 | | C10 | AcFA10K+H | Conf A | [99] |
| OH | stretch | C-term | 3280 | | bound | Phe | Band B | [106] |
| NH | stretch | NH3+ (Orn) | 3247 | 3243 | free | cyclo-VOLFPVOLFP+2H | Conf A | [13] |
| NH | stretch | NH3+ (Orn) | 3243 | 3227 | bound | cyclo-VOLFPVOLFP+2H | Conf B | [69] |
| NH | stretch | NH3+ (Orn) | 3236 | 3243 | bound | cyclo-VOLFPVOLFP+2H | Conf A | [13] |
| OH | stretch | C-term | 3235 | | bound | Phe | Band X | [106] |
| NH | stretch | N-term | 3186.3 | 3193.7 | bound | YA+H | | [57] |
| NH | stretch | N-term | 3178.6 | 3190.4 | bound | YA+H | | [57] |
| CH | stretch | N-term | 3174 | | free | His+H | | [16] |
| NH | stretch | N-term | 3138 | 3169.3 | bound | YA+H | | [57] |
| NH | stretch | N-term | 3137.3 | 3164.8 | bound | YA+H | | [57] |
| NH | stretch | N-term | 3129.3 | 3197 | bound | Tyr+H | | [57] |
| NH | stretch | N-term | 3127.3 | 3196.5 | bound | Tyr+H | | [57] |
| NH | stretch | N-term | 3124 | 3188.4 | bound | Phe+H | | [57] |
| CH | stretch | | 3095 | | | AcPheOMe | | [98] |
| NH | stretch | N-term | 3094.8 | 3140.6 | bound | AY+H | | [57] |
| NH | stretch | N-term | 3091.5 | 3124 | bound | Phe+H | | [57] |
| NH | stretch | N-term | 3089 | 3129.1 | bound | Tyr+H | | [57] |
| NH | stretch | N-term | 3087.45 | 3127.2 | bound | Tyr+H | | [57] |
| NH | stretch | N-term | 3086.9 | 3143 | bound | Tyr+H | | [57] |
| NH | stretch | N-term | 3083 | 3130.6 | bound | Phe+H | | [57] |
| CH | stretch | | 3076 | | | AcPheOMe | | [98] |
| NH | stretch | N-term | 3075.8 | 3138.7 | bound | Tyr+H | | [57] |
| CH | stretch | | 3074 | | | AcPheOMe | | [98] |
| CH | stretch | | 3069 | | | WG | | [104] |
| NH | stretch | N-term | 3067.2 | 3102.9 | bound | AY+H | | [57] |
| NH | stretch | N-term | 3059.3 | 3098.1 | bound | Tyr+H | | [57] |
| NH | stretch | N-term | 3058.8 | 3090.8 | bound | Phe+H | | [57] |
| NH | stretch | N-term | 3054.9 | 3096.9 | bound | Tyr+H | | [57] |
| CH | stretch | | 3039 | | | AcPheOMe | | [98] |
| CH | stretch | | 3038 | | | AcPheOMe | | [98] |
| CH | stretch | | 3003 | | | AcPheOMe | | [98] |
| CH | stretch | | 3000 | | | AcPheOMe | | [98] |
| CH | stretch | | 2963 | | | AcPheOMe | | [98] |
| CH | stretch | | 2961 | | | AcPheOMe | | [98] |
| CH | stretch | | 2927 | | | AcPheOMe | | [98] |
| CH | stretch | | 2899 | | | AcPheOMe | | [98] |
| NH | stretch | NH3+ (Orn) | 2882 | 2843 | bound | cyclo-VOLFPVOLFP+2H | | Nagornova, 2011 #10} |
| CH | stretch | | 2858 | | | AcPheOMe | | [98] |
| NH | stretch | NH3+ (Orn) | 2667 | 2679 | bound | cyclo-VOLFPVOLFP+2H | | Nagornova, 2011 #10} |
| CO | stretch | C-term | 1788 | | free | SarGly+H | | [107] |
| CO | stretch | C-term | 1786 | | bound | GG+H | | [107] |
| CO | stretch | C-term | 1785 | 1785 | bound | GG+H | | [40] |
| CO | stretch | C-term | 1783 | | bound | GlySar+H | | [107] |
| CO | stretch | C-term | 1782 | | | PW | | [105] |
| CO | stretch | C-term | 1782 | | | PW | | [105] |
| CO | stretch | C-term | 1778 | | free | SarSar+H | | [107] |
| CO | stretch | C-term | 1754 | | free | SarSar+H | | [107] |
| CO | stretch | Amide | 1722 | 1751 | bound | GG+H | | [40] |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| CO | stretch | Amide | 1722 | | bound | GG+H | [107] |
| CO | stretch | Amide | 1721 | | bound | SarGly+H | [107] |
| CO | stretch | Amide | 1711 | | | AcPheNHMe | [95] |
| CO | stretch | Amide | 1707 | | | Leu-Enkephalin+H | [110] |
| CO | stretch | Amide | 1706 | | | AcPheNHMe | [95] |
| CO | stretch | Amide | 1700 | | bound | GlySar+H | [107] |
| CO | stretch | Amide | 1697 | | bound | SarSar+H | [107] |
| CO | stretch | Amide | 1695 | | | AcPheNHMe | [95] |
| CO | stretch | Amide | 1692 | | free | PW | [105] |
| C=O | stretch | | 1688 | 1691 | | cyclo-VOLFPVOLFP+2H | [13] |
| CO | stretch | Amide | 1681 | | β-sheet | GS+2H | [13] |
| C=O | stretch | | 1681 | 1680 | | cyclo-VOLFPVOLFP+2H | [13] |
| CO | stretch | Amide | 1676 | | β-sheet | AcPheNHMe | [95] |
| C=O | stretch | | 1668 | 1668 | | cyclo-VOLFPVOLFP+2H | [13] |
| NH | umbrella | N-term | 1655 | | | Leu-Enkephalin+H | [110] |
| C=O | stretch | | 1621 | 1626 | | cyclo-VOLFPVOLFP+2H | [13] |
| CO | stretch | Amide | 1619 | | bound | GS+2H | |
| C=O | stretch | | 1615 | 1615 | | cyclo-VOLFPVOLFP+2H | [13] |
| NH | bend | Amide | 1569 | | | AcPheNHMe | |
| NH | bend | Amide | 1554 | | | AcPheNHMe | |
| NH | bend | | 1551 | 1561 | | cyclo-VOLFPVOLFP+2H | [13] |
| NH | bend | Amide | 1549 | | | AcPheNHMe | |
| NH | bend | Amide | 1545 | | free | SarGly+H | |
| NH | bend | | 1544 | 1560 | broad | cyclo-VOLFPVOLFP+2H | [13] |
| NH | bend | Amide | 1536 | 1544 | bound | GG+H | [40] |
| NH | bend | Amide | 1536 | | bound | GG+H | [107] |
| NH | bend | Amide | 1536 | | | AcPheNHMe | |
| NH | bend | | 1524 | 1535 | | cyclo-VOLFPVOLFP+2H | [13] |
| NH | bend | Amide | 1516 | | | Leu-Enkephalin+H | [110] |
| NH | bend | Amide | 1514 | | | AcPheNHMe | [95] |
| NH | bend | | 1514 | 1534 | | cyclo-VOLFPVOLFP+2H | [13] |
| NH | bend | Amide | 1511 | | | PW | [105] |
| NH | bend | Amide | 1511 | | | PW | [105] |
| NH | bend | | 1485 | 1497 | | cyclo-VOLFPVOLFP+2H | [13] |

## VI. References


(1) Prusiner, S. B., Prion diseases and the BSE crisis. *Science* **1997,** *278*, 245-251.

(2) Nielsen, B. G.; Jensen, M. O.; Bohr, H. G., The probability distribution of side-chain conformations in [Leu] and [Met]enkephalin determines the potency and selectivity to mu and delta opiate receptors. *Biopolymers* **2003,** *71*, 577-592.

(3) Granier, S.; Manglik, A.; Kruse, A. C.; Kobilka, T. S.; Thian, F. S.; Weis, W. I.; Kobilka, B. K., Structure of the delta-opioid receptor bound to naltrindole. *Nature* **2012,** *485*, 400-U171.

(4) Adrian, M.; Dubochet, J.; Lepault, J.; McDowall, A. W., Cryo-electron microscopy of viruses. *Nature* **1984,** *308*, 32.

(5) Wüthrich, K., The way to NMR structures of proteins. *Nature Structural Biology* **2001,** *8*, 923.

(6) Konrat, R., NMR contributions to structural dynamics studies of intrinsically disordered proteins. *J Magn Reson* **2014,** *241*, 74-85.

(7) Hunt, N. T., 2D-IR spectroscopy: ultrafast insights into biomolecule structure and function. *Chem Soc Rev* **2009,** *38*, 1837-1848.

(8) Oh, H.; Breuker, K.; Sze, S. K.; Ge, Y.; Carpenter, B. K.; McLafferty, F. W., Secondary and tertiary structures of gaseous protein ions characterized by electron capture dissociation mass spectrometry and photofragment spectroscopy. *Proceedings of the National Academy of Sciences* **2002,** *99*, 15863.

(9) Bleiholder, C.; Dupuis, N. F.; Wyttenbach, T.; Bowers, M. T., Ion mobility-mass spectrometry reveals a conformational conversion from random assembly to beta-sheet in amyloid fibril formation. *Nat. Chem.* **2011,** *3*, 172-177.

(10) Rizzo, T. R.; Stearns, J. A.; Boyarkin, O. V., Spectroscopic studies of cold, gas-phase biomolecular ions. *International Reviews in Physical Chemistry* **2009,** *28*, 481 - 515.

(11) Kopysov, V.; Makarov, A.; Boyarkin, O. V., Colors for molecular masses: fusion of spectroscopy and mass spectrometry for identification of biomolecules. *Anal Chem* **2015,** *87*, 4607-11.

(12) Boyarkin, O. V.; Mercier, S. R.; Kamariotis, A.; Rizzo, T. R., Electronic spectroscopy of cold, protonated tryptophan and tyrosine. *J. Am. Chem. Soc.* **2006,** *128*, 2816-2817.

(13) Nagornova, N. S.; Guglielmi, M.; Doemer, M.; Tavernelli, I.; Rothlisberger, U.; Rizzo, T. R.; Boyarkin, O. V., Cold-Ion Spectroscopy Reveals the Intrinsic Structure of a Decapeptide. *Angew Chem Int Edit* **2011,** *50*, 5383-5386.

(14) Harrilal, C. P.; DeBlase, A. F.; Fischer, J. L.; Lawler, J. T.; McLuckey, S. A.; Zwier, T. S., Infrared Population Transfer Spectroscopy of Cryo-Cooled Ions: Quantitative Tests of the Effects of Collisional Cooling on the Room Temperature Conformer Populations. *J. Phys. Chem. A* **2018,** *122*, 2096-2107.

(15) Gabelica, V.; De Pauw, E., Internal energy and fragmentation of ions produced in electrospray sources. *Mass Spectrom Rev* **2005,** *24*, 566-87.

(16) Pereverzev, A. Y.; Kopysov, V.; Boyarkin, O. V., High Susceptibility of Histidine to Charge Solvation Revealed by Cold Ion Spectroscopy. *Angew Chem Int Ed Engl* **2017,** *56*, 15639-15643.

(17) Silveira, J. A.; Fort, K. L.; Kim, D.; Servage, K. A.; Pierson, N. A.; Clemmer, D. E.; Russell, D. H., From solution to the gas phase: stepwise dehydration and kinetic trapping of substance P reveals the origin of peptide conformations. *J Am Chem Soc* **2013,** *135*, 19147-53.

(18) Nagornova, N. S.; Rizzo, T. R.; Boyarkin, O. V., Exploring the Mechanism of IR-UV Double-Resonance for Quantitative Spectroscopy of Protonated Polypeptides and Proteins. *Angew Chem Int Edit* **2013,** *52*, 6002-6005.
(19) Gerlich, D., INHOMOGENEOUS RF FIELDS: A VERSATILE TOOL FOR THE STUDY OF PROCESSES WITH SLOW IONS. In *State-selected and state-to-state ion-molecule reaction dynamics, part 1. Experiment*, Baer, C.-Y. N. a. M., Ed. John Willey]Sons, Inc.: 1992; Vol. LXXXII, p 1.
(20) Gerlich, D., *The Production and Study of Ultra-Cold Molecular Ions*. 2008.
(21) Asvany, O.; Schlemmer, S., Numerical simulations of kinetic ion temperature in a cryogenic linear multipole trap. *Int. J. Mass Spectrom.* **2009,** *279*, 147-155.
(22) Hock, C.; Kim, J. B.; Weichman, M. L.; Yacovitch, T. I.; Neumark, D. M., Slow photoelectron velocity-map imaging spectroscopy of cold negative ions. *J Chem Phys* **2012,** *137*.
(23) Boyarkin, O. V.; Kopysov, V., Cryogenically cooled octupole ion trap for spectroscopy of biomolecular ions. *Rev. Sci. Instrum.* **2014,** *85*, 033105.
(24) Rizzo, T. R.; Boyarkin, O. V., Cryogenic Methods for the Spectroscopy of Large, Biomolecular Ions. *Gas-Phase Ir Spectroscopy and Structure of Biological Molecules* **2015,** *364*, 43-97.
(25) Heine, N.; Asmis, K. R., Cryogenic Ion Trap Vibrational Spectroscopy of Hydrogen-Bonded Clusters Relevant to Atmospheric Chemistry (vol 34, pg 1, 2015). *International Reviews in Physical Chemistry* **2016,** *35*, 507-507.
(26) Roithova, J.; Gray, A.; Andris, E.; Jasik, J.; Gerlich, D., Helium Tagging Infrared Photodissociation Spectroscopy of Reactive Ions. *Accounts Chem Res* **2016,** *49*, 223-230.
(27) Paul, W., Electromagnetic traps for charged and neutral particles. *Reviews of Modern Physics* **1990,** *62*, 531-540.
(28) Choi, C. M.; Choi, D. H.; Kim, N. J.; Heo, J., Effective temperature of protonated tyrosine ions in a cold quadrupole ion trap. *Int. J. Mass Spectrom.* **2012,** *314*, 18-21.
(29) Ishiuchi, S.; Wako, H.; Kato, D.; Fujii, M., High-cooling-efficiency cryogenic quadrupole ion trap and UV-UV hole burning spectroscopy of protonated tyrosine. *J Mol Spectrosc* **2017,** *332*, 45-51.
(30) Jasik, J.; Zabka, J.; Roithova, J.; Gerlich, D., Infrared spectroscopy of trapped molecular dications below 4 K. *Int. J. Mass Spectrom.* **2013,** *354*, 204-210.
(31) Johnson, C. J.; Wolk, A. B.; Fournier, J. A.; Sullivan, E. N.; Weddle, G. H.; Johnson, M. A., Communication: He-tagged vibrational spectra of the SarGlyH+ and H+(H2O)2,3 ions: Quantifying tag effects in cryogenic ion vibrational predissociation (CIVP) spectroscopy. *The Journal of Chemical Physics* **2014,** *140*, 221101.
(32) Cismesia, A. P.; Tesler, L. F.; Bell, M. R.; Bailey, L. S.; Polfer, N. C., Infrared ion spectroscopy inside a mass-selective cryogenic 2D linear ion trap. *Journal of Mass Spectrometry* **2017,** *52*, 720-727.
(33) Florez, A. I. G.; Ahn, D. S.; Gewinner, S.; Schollkopf, W.; von Helden, G., IR spectroscopy of protonated leu-enkephalin and its 18-crown-6 complex embedded in helium droplets. *Phys Chem Chem Phys* **2015,** *17*, 21902-21911.
(34) Boyarkin, O. V.; Rizzo, T. R.; Rueda, D.; Quack, M.; Seyfang, G., Nonlinear intensity dependence in the infrared multiphoton excitation and dissociation of

methanol pre-excited to different energies. *The Journal of Chemical Physics* **2002,** *117*, 9793-9805.
(35) Bagratashvili, V. N.; Letokhov, V. S.; Makarov, A. A.; Ryabov, E. A., Multiple-Photon Infrared Laser Photophysics and Photochemistry. I. *Laser Chemistry* **1983,** *1*, 211-342.
(36) Boyarkin, O. V.; Ionov, S. I.; Bagratashvili, V. N., Ir Spectroscopy and Dynamics of a Strongly Vibrationally Excited Polyatomic Molecule - Cf3i. *Chem Phys Lett* **1988,** *146*, 106-112.
(37) Polfer, N. C.; Valle, J. J.; Moore, D. T.; Oomens, J.; Eyler, J. R.; Bendiak, B., Differentiation of isomers by wavelength-tunable infrared multiple-photon dissociation-mass spectrometry: Application to glucose-containing disaccharides. *Anal. Chem.* **2006,** *78*, 670-679.
(38) Citir, M.; Hinton, C. S.; Oomens, J.; Steill, J. D.; Armentrout, P. B., Infrared multiple photon dissociation spectroscopy of protonated histidine and 4-phenyl imidazole. *Int. J. Mass Spectrom.* **2012,** *330*, 6-15.
(39) Martens, J.; Koppen, V.; Berden, G.; Cuyckens, F.; Oomens, J., Combined Liquid Chromatography-Infrared Ion Spectroscopy for Identification of Regioisomeric Drug Metabolites. *Anal. Chem.* **2017,** *89*, 4359-4362.
(40) Kamrath, M. Z.; Garand, E.; Jordan, P. A.; Leavitt, C. M.; Wolk, A. B.; Van Stipdonk, M. J.; Miller, S. J.; Johnson, M. A., Vibrational Characterization of Simple Peptides Using Cryogenic Infrared Photodissociation of H-2-Tagged, Mass-Selected Ions. *J. Am. Chem. Soc.* **2011,** *133*, 6440-6448.
(41) Roithová, J.; Gray, A.; Andris, E.; Jašík, J.; Gerlich, D., Helium Tagging Infrared Photodissociation Spectroscopy of Reactive Ions. *Accounts Chem Res* **2016,** *49*, 223-230.
(42) Garand, E.; Kamrath, M. Z.; Jordan, P. A.; Wolk, A. B.; Leavitt, C. M.; McCoy, A. B.; Miller, S. J.; Johnson, M. A., Determination of Noncovalent Docking by Infrared Spectroscopy of Cold Gas-Phase Complexes. *Science* **2012,** *335*, 694-698.
(43) Heine, N.; Fagiani, M. R.; Rossi, M.; Wende, T.; Berden, G.; Blum, V.; Asmis, K. R., Isomer-Selective Detection of Hydrogen-Bond Vibrations in the Protonated Water Hexamer. *J. Am. Chem. Soc.* **2013,** *135*, 8266-8273.
(44) Fagiani, M. R.; Knorke, H.; Esser, T. K.; Heine, N.; Wolke, C. T.; Gewinner, S.; Schollkopf, W.; Gaigeot, M. P.; Spezia, R.; Johnson, M. A.; Asmis, K. R., Gas phase vibrational spectroscopy of the protonated water pentamer: the role of isomers and nuclear quantum effects. *Phys Chem Chem Phys* **2016,** *18*, 26743-26754.
(45) Masellis, C.; Khanal, N.; Kamrath, M. Z.; Clemmer, D. E.; Rizzo, T. R., Cryogenic Vibrational Spectroscopy Provides Unique Fingerprints for Glycan Identification. *J. Am. Soc. Mass. Spectr.* **2017,** *28*, 2217-2222.
(46) Mucha, E.; Florez, A. I. G.; Marianski, M.; Thomas, D. A.; Hoffmann, W.; Struwe, W. B.; Hahm, H. S.; Gewinner, S.; Schollkopf, W.; Seeberger, P. H.; von Helden, G.; Pagel, K., Glycan Fingerprinting via Cold-Ion Infrared Spectroscopy. *Angew Chem Int Edit* **2017,** *56*, 11248-11251.
(47) Yang, B.; Wu, R. R.; Polfer, N. C.; Berden, G.; Oomens, J.; Rodgers, M. T., IRMPD Action Spectroscopy of Alkali Metal Cation-Cytosine Complexes: Effects of Alkali Metal Cation Size on Gas Phase Conformation. *J. Am. Soc. Mass. Spectr.* **2013,** *24*, 1523-1533.
(48) Schindler, B.; Barnes, L.; Renois, G.; Gray, C.; Chambert, S.; Fort, S.; Flitsch, S.; Loison, C.; Allouche, A. R.; Compagnon, I., Anomeric memory of the glycosidic

bond upon fragmentation and its consequences for carbohydrate sequencing. *Nat Commun* **2017,** *8*, 973.
(49) Khanal, N.; Masellis, C.; Kamrath, M. Z.; Clemmer, D. E.; Rizzo, T. R., Glycosaminoglycan Analysis by Cryogenic Messenger-Tagging IR Spectroscopy Combined with IMS-MS. *Anal. Chem.* **2017,** *89*, 7601-7606.
(50) Nagornova, N. S.; Kopysov, V.; Boyarkin, O. V., Unfolding of a Protein Revealed by Cold Ion Spectrosopy. *J. Am. Chem. Soc.* **2018,** *in preparation*.
(51) Pereverzev, A. Y.; Kopysov, V. N.; Boyarkin, O. V., Peptide Bond Ultraviolet Absorption Enables Vibrational Cold-Ion Spectroscopy of Nonaromatic Peptides. *J Phys Chem Lett* **2018,** *9*, 5262-5266.
(52) Page, R. H.; Shen, Y. R.; Lee, Y. T., Infrared-Ultraviolet Double-Resonance Studies of Benzene Molecules in a Supersonic Beam. *J Chem Phys* **1988,** *88*, 5362-5376.
(53) Walther, T.; Bitto, H.; Minton, T. K.; Huber, J. R., Uv-Ir Double-Resonance Spectroscopy of Jet-Cooled Propynal Detected by the Fluorescence Dip Method. *Chem Phys Lett* **1994,** *231*, 64-69.
(54) Carney, J. R.; Zwier, T. S., The infrared and ultraviolet spectra of individual conformational isomers of biomolecules: Tryptamine. *J. Phys. Chem. A* **2000,** *104*, 8677-8688.
(55) Snoek, L. C.; Kroemer, R. T.; Simons, J. P., A spectroscopic and computational exploration of tryptophan-water cluster structures in the gas phase. *Phys Chem Chem Phys* **2002,** *4*, 2130-2139.
(56) Fricke, H.; Funk, A.; Schrader, T.; Gerhards, M., Investigation of Secondary Structure Elements by IR/UV Double Resonance Spectroscopy:  Analysis of an Isolated β-Sheet Model System. *J. Am. Chem. Soc.* **2008,** *130*, 4692-4698.
(57) Stearns, J. A.; Guidi, M.; Boyarkin, O. V.; Rizzo, T. R., Conformation-specific infrared and ultraviolet spectroscopy of tyrosine-based protonated dipeptides. *J Chem Phys* **2007,** *127*, 154322-7.
(58) Stearns, J. A.; Mercier, S.; Seaiby, C.; Guidi, M.; Boyarkin, O. V.; Rizzo, T. R., Conformation-specific Spectroscopy and photodissociation of cold, protonated tyrosine and phenylalanine. *J. Am. Chem. Soc.* **2007,** *129*, 11814-11820.
(59) Nagornova, N. S.; Rizzo, T. R.; Boyarkin, O. V., Highly Resolved Spectra of Gas-Phase Gramicidin S: A Benchmark for Peptide Structure Calculations. *J. Am. Chem. Soc.* **2010,** *132*, 4040-+.
(60) Inokuchi, Y.; Boyarkin, O. V.; Ebata, T.; Rizzo, T. R., UV and IR spectroscopy of cold 1,2-dimethoxybenzene complexes with alkali metal ions. *Phys Chem Chem Phys* **2012,** *14*, 4457-62.
(61) Pereverzev, A. Y.; Cheng, X.; Nagornova, N. S.; Reese, D. L.; Steele, R. P.; Boyarkin, O. V., Vibrational Signatures of Conformer-Specific Intramolecular Interactions in Protonated Tryptophan. *J. Phys. Chem A* **2016,** *120*, 5598-5608.
(62) Pereverzev, A. Y.; Boyarkin, O. V., Exploring the relevance of gas-phase structures to biology: cold ion spectroscopy of the decapeptide neurokinin A. *Phys Chem Chem Phys* **2017,** *19*, 3468-3472.
(63) Choi, C. M.; Choi, D. H.; Heo, J.; Kim, N. J.; Kim, S. K., Ultraviolet–Ultraviolet Hole Burning Spectroscopy in a Quadrupole Ion Trap: Dibenzo[18]crown-6 Complexes with Alkali Metal Cations. *Angew. Chem. Int. Ed.* **2012,** *51*, 7297-7300.
(64) Feraud, G.; Dedonder, C.; Jouvet, C.; Inokuchi, Y.; Haino, T.; Sekiya, R.; Ebata, T., Development of Ultraviolet-Ultraviolet Hole-Burning Spectroscopy for Cold Gas-Phase Ions. *J Phys Chem Lett* **2014,** *5*, 1236-40.

(65) Wako, H.; Ishiuchi, S.; Kato, D.; Feraud, G.; Dedonder-Lardeux, C.; Jouvet, C.; Fujii, M., A conformational study of protonated noradrenaline by UV-UV and IR dip double resonance laser spectroscopy combined with an electrospray and a cold ion trap method. *Phys Chem Chem Phys* **2017,** *19*, 10777-10785.
(66) Wolk, A. B.; Leavitt, C. M.; Garand, E.; Johnson, M. A., Cryogenic Ion Chemistry and Spectroscopy. *Accounts Chem Res* **2014,** *47*, 202-210.
(67) Kupser, P.; Pagel, K.; Oomens, J.; Polfer, N.; Koksch, B.; Meijer, G.; von Helden, G., Amide-I and -II Vibrations of the Cyclic β-Sheet Model Peptide Gramicidin S in the Gas Phase. *J. Am. Chem. Soc.* **2010,** *132*, 2085-2093.
(68) Rijs Anouk, M.; Kabeláč, M.; Abo-Riziq, A.; Hobza, P.; de Vries Mattanjah, S., Isolated Gramicidin Peptides Probed by IR Spectroscopy. *Chemphyschem* **2011,** *12*, 1816-1821.
(69) Joshi, K.; Semrouni, D.; Ohanessian, G.; Clavaguera, C., Structures and IR Spectra of the Gramicidin S Peptide: Pushing the Quest for Low-Energy Conformations. *J Phys Chem B* **2012,** *116*, 483-490.
(70) Nagornova, N. S.; Rizzo, T. R.; Boyarkin, O. V., Interplay of Intra- and Intermolecular H-Bonding in a Progressively Solvated Macrocyclic Peptide. *Science* **2012,** *336*, 320-323.
(71) Roy, T. K.; Nagornova, N. S.; Boyarkin, O. V.; Gerber, R. B., A Decapeptide Hydrated by Two Waters: Conformers Determined by Theory and Validated by Cold Ion Spectroscopy. *J. Phys. Chem. A* **2017,** *121*, 9401-9408.
(72) Jaeqx, S.; Oomens, J.; Cimas, A.; Gaigeot, M. P.; Rijs, A. M., Gas-Phase Peptide Structures Unraveled by Far-IR Spectroscopy: Combining IR-UV Ion-Dip Experiments with Born-Oppenheimer Molecular Dynamics Simulations. *Angew Chem Int Edit* **2014,** *53*, 3663-3666.
(73) Kopysov, V.; Boyarkin, O. V., Resonance Energy Transfer Relates the Gas-Phase Structure and Pharmacological Activity of Opioid Peptides. *Angew. Chem. Int. Ed.* **2016,** *55*, 689-692.
(74) He, F.; Emmett, M. R.; Håkansson, K.; Hendrickson, C. L.; Marshall, A. G., Theoretical and Experimental Prospects for Protein Identification Based Solely on Accurate Mass Measurement. *J Proteome Res* **2003,** *3*, 61-67.
(75) Wu, C.; Siems, W. F.; Klasmeier, J.; Hill, H. H., Separation of isomeric peptides using electrospray ionization/high-resolution ion mobility spectrometry. *Anal. Chem.* **2000,** *72*, 391-395.
(76) Winter, D.; Pipkorn, R.; Lehmann, W. D., Separation of peptide isomers and conformers by ultra performance liquid chromatography. *J Sep Sci* **2009,** *32*, 1111-1119.
(77) Shvartsburg, A. A.; Creese, A. J.; Smith, R. D.; Cooper, H. J., Separation of a Set of Peptide Sequence Isomers Using Differential Ion Mobility Spectrometry. *Anal. Chem.* **2011,** *83*, 6918-6923.
(78) Norbeck, A. D.; Monroe, M. E.; Adkins, J. N.; Anderson, K. K.; Daly, D. S.; Smith, R. D., The utility of accurate mass and LC elution time information in the analysis of complex proteomes. *J. Am. Soc. Mass. Spectr.* **2005,** *16*, 1239-1249.
(79) Riddle, L. A.; Guiochon, G., Influence of mobile phase gradients on the retention and separation of peptides from a cytochrome-c digest by reversed-phase liquid chromatography. *Chromatographia* **2006,** *64*, 121-127.
(80) Kopysov, V.; Nagomova, N. S.; Boyarkin, O. V., Identification of Tyrosine-Phosphorylated Peptides Using Cold Ion Spectroscopy. *J. Am. Chem. Soc.* **2014,** *136*, 9288-9291.

(81) Brodbelt, J. S., Photodissociation mass spectrometry: new tools for characterization of biological molecules. *Chem Soc Rev* **2014,** *43*, 2757-2783.
(82) Kopysov, V.; Makarov, A.; Boyarkin, O. V., Nonstatistical UV Fragmentation of Gas-Phase Peptides Reveals Conformers and Their Structural Features. *J Phys Chem Lett* **2016,** *7*, 1067-1071.
(83) Kopysov, V.; Makarov, A.; Boyarkin, O. V., Identification of Isomeric Ephedrines by Cold Ion UV Spectroscopy: Toward Practical Implementation. *Anal. Chem.* **2017,** *89*, 544-547.
(84) Kopysov, V.; Gorshkov, M. V.; Boyarkin, O. V., Identification of Isoforms of Aspartic Acid Residue in Peptides by 2D UV-MS Fingerprinting of Cold Ions. *Analyst* **2018**.
(85) Liu, Y.; Clemmer, D. E., Characterizing Oligosaccharides Using Injected-Ion Mobility/Mass Spectrometry. *Anal. Chem.* **1997,** *69*, 2504-2509.
(86) Hofmann, J.; Hahm, H. S.; Seeberger, P. H.; Pagel, K., Identification of carbohydrate anomers using ion mobility-mass spectrometry. *Nature* **2015,** *526*, 241-+.
(87) Vrkic, A. K.; O'Hair, R. A. J., Using non-covalent complexes to direct the fragmentation of glycosidic bonds in the gas phase† ††Gas Phase Ion Chemistry of Biomolecules, Part 39. *J. Am. Soc. Mass. Spectr.* **2004,** *15*, 715-724.
(88) Gorman, G. S.; Speir, J. P.; Turner, C. A.; Amster, I. J., Proton affinities of the 20 common .alpha.-amino acids. *J. Am. Chem. Soc.* **1992,** *114*, 3986-3988.
(89) Voleti, N.; Vairamani, M., Proton affinity differences among three N-acetylhexosamines studied by the kinetic method. *Rapid Commun Mass Spectrom* **2003,** *17*, 1089-91.
(90) Gronert, S.; Simpson, D. C.; Conner, K. M., A reevaluation of computed proton affinities for the common alpha-amino acids. *J Am Soc Mass Spectrom* **2009,** *20*, 2116-23.
(91) Agmon, N., Elementary Steps in Excited-State Proton Transfer. *J. Phys. Chem A* **2005,** *109*, 13-35.
(92) Horke, D. A.; Watts, H. M.; Smith, A. D.; Jager, E.; Springate, E.; Alexander, O.; Cacho, C.; Chapman, R. T.; Minns, R. S., Hydrogen Bonds in Excited State Proton Transfer. *Phys Rev Lett* **2016,** *117*, 163002.
(93) Lee, D. D.; Seung, H. S., Learning the parts of objects by non-negative matrix factorization. *Nature* **1999,** *401*, 788-791.
(94) Owen, A. B.; Perry, P. O., Bi-cross-validation of the SVD and the nonnegative matrix factorization. *Ann. Appl. Stat.* **2009,** *3*, 564-594.
(95) Gerhards, M.; Unterberg, C.; Gerlach, A.; Jansen, A., beta-sheet model systems in the gas phase: Structures and vibrations of Ac-Phe-NHMe and its dimer (Ac-Phe-NHMe)(2). *Phys Chem Chem Phys* **2004,** *6*, 2682-2690.
(96) Chin, W.; Dognon, J. P.; Piuzzi, F.; Tardivel, B.; Dimicoli, I.; Mons, M., Intrinsic folding of small peptide chains: spectroscopic evidence for the formation of beta-turns in the gas phase. *J Am Chem Soc* **2005,** *127*, 707-12.
(97) Martinez, S. J.; Alfano, J. C.; Levy, D. H., The Electronic Spectroscopy of the Amino-Acids Tyrosine and Phenylalanine in a Supersonic Jet. *J Mol Spectrosc* **1992,** *156*, 421-430.
(98) Gerhards, M.; Unterberg, C., Structures of the protected amino acid Ac-Phe-OMe and its dimer: A beta-sheet model system in the gas phase. *Phys Chem Chem Phys* **2002,** *4*, 1760-1765.

(99) Stearns, J. A.; Boyarkin, O. V.; Rizzo, T. R., Spectroscopic signatures of gas-phase helices: Ac-Phe-(Ala)(5)-Lys-H+ and Ac-Phe-(Ala)(10)-Lys-H+. *J. Am. Chem. Soc.* **2007,** *129*, 13820-+.
(100) Chin, W.; Piuzzi, F.; Dognon, J.-P.; Dimicoli, I.; Tardivel, B.; Mons, M., Gas Phase Formation of a 310-Helix in a Three-Residue Peptide Chain: Role of Side Chain-Backbone Interactions as Evidenced by IR−UV Double Resonance Experiments. *J. Am. Chem. Soc.* **2005,** *127*, 11900-11901.
(101) Stearns, J. A.; Boyarkin, O. V.; Rizzo, T. R., Effects of N-terminus substitution on the structure and spectroscopy of gas-phase helices. *Chimia* **2008,** *62*, 240-243.
(102) Cohen, R.; Brauer, B.; Nir, E.; Grace, L.; de Vries, M. S., Resonance-Enhanced Multiphoton Ionization Spectroscopy of Dipeptides. *J. Phys. Chem A* **2000,** *104*, 6351-6355.
(103) Papadopoulos, G.; Svendsen, A.; Boyarkin, O. V.; Rizzo, T. R., Spectroscopy of mobility-selected biomolecular ions. *Faraday Discuss* **2011,** *150*, 243-255.
(104) Hunig, I.; Kleinermanns, K., Conformers of the peptides glycine-tryptophan, tryptophan-glycine and tryptophan-glycine-glycine as revealed by double resonance laser spectroscopy. *Phys Chem Chem Phys* **2004,** *6*, 2650-2658.
(105) Haber, T.; Seefeld, K.; Kleinermanns, K., Mid- and near-infrared spectra of conformers of H-Pro-Trp-OH. *J. Phys. Chem. A* **2007,** *111*, 3038-3046.
(106) Snoek, L. C.; Robertson, E. G.; Kroemer, R. T.; Simons, J. P., Conformational landscapes in amino acids: infrared and ultraviolet ion-dip spectroscopy of phenylalanine in the gas phase. *Chem Phys Lett* **2000,** *321*, 49-56.
(107) Leavitt, C. M.; Wolk, A. B.; Kamrath, M. Z.; Garand, E.; Van Stipdonk, M. J.; Johnson, M. A., Characterizing the Intramolecular H-bond and Secondary Structure in Methylated GlyGlyH(+) with H-2 Predissociation Spectroscopy. *J. Am. Soc. Mass. Spectr.* **2011,** *22*, 1941-1952.
(108) Chin, W.; Piuzzi, F.; Dognon, J. P.; Dimicoli, L.; Tardivel, B.; Mons, M., Gas phase formation of a 3(10)-helix in a three-residue peptide chain: Role of side chain-backbone interactions as evidenced by IR-UV double resonance experiments. *J. Am. Chem. Soc.* **2005,** *127*, 11900-11901.
(109) Roy, T. K.; Kopysov, V.; Nagornova, N. S.; Rizzo, T. R.; Boyarkin, O. V.; Gerber, R. B., Conformational Structures of a Decapeptide Validated by First Principles Calculations and Cold Ion Spectroscopy. *Chemphyschem* **2015,** *16*, 1374-1378.
(110) 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. *J. Am. Chem. Soc.* **2007,** *129*, 5887-5897.