COLD ION SPECTROSCOPY FOR STRUCTURAL IDENTIFICATIONS OF BIOMOLECULES

Boyarkin O.V., Laboratoire de Chimie Physique Moléculaire, EPFL, oleg.boiarkin@epfl.ch

Over the last decade, the spectroscopy of cryogenically cold ions isolated in the gas phase has been developed as a new tool for structural elucidations of biological molecules. Cooling to as low as ~10 K allows for vibrational resolution in UV spectra of small to midsize pharmaceutical drugs, peptides, carbohydrates, lipids, as well as of clusters of these biomolecules. The rich in details UV spectra of these cold gas-phase species become very specific not only to chemical composition, but also to their geometry (e.g., isomers or conformers). Absorption of a UV photon often results in photofragmentation of a protonated parent biomolecules, thus producing charged fragments with the mass-to-charge ratios that are different from that of the parent. This allows for a convenient way of measuring of the UV "action" spectra by mass-spectrometric detection of these fragments. When all the fragments are detected at once at a fixed UV wavelength, the resulting fragment mass spectrum (fragment intensity vs m/z ratio) also exhibits high specificity to the chemical composition and to the geometry of the parent.

Here we describe the pioneering analytical method that fuses UV action spectroscopy and UV photofragment mass spectrometry to a powerful analytical method, called 2D UV-MS fingerprinting. The method is illustrated by its applications to identifications and relative quantifications of isomers for different types of biomolecules, as well as of conformers of peptides.

INTRODUCTION

Conformational/isomeric heterogeneity of biomolecules in vivo is a key factor that enables a high level of 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 the 3D structures of conformers and revealing the mechanisms of their interconversion in vivo is therefore one of the key problems that have to be solved to enable the intelligent modeling of functionality of biomolecules on a molecular level. Currently, none of the known methods alone, owing to 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 from the geometry they adopt in vivo [2, 3]. This powerful method remains the main "work horse" for solving structures of large biomolecules, although crystal structures of only rigid molecules with low conformational heterogeneity may have any direct relevance to their native conformation. Recently developed to the level of a laboratory technique, the method of cryogenic electron microscopy (cryo-EM) solves some of these problems [4]. This technique can be applied for structural determinations of

large biomolecules that are 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 still 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 that can be as large as proteins [5]. One of the weak points of NMR is its limited 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 using time-resolved 2D IR spectroscopy of biomolecules in solutions [7]. This approach 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 its poor spectral resolution in solutions.

One common problem for the condensed phase methods is a need for a highly purified sample in relatively large quantities. The development of electrospray ionization, which allows for transfer of large, non-volatile biomolecules directly from the 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 reduces the required amount and purity of samples. While MS alone provides little information regarding the 3D structure of biomolecular ions, there are, at least, two associated approaches that may deliver certain structural information. The electron transfer/capture dissociation methods demonstrate certain selectivity for the tertiary structure of proteins and allow an evaluation of their overall folding state, although by far, these methods cannot serve to unambiguously validate the 3D structure [8]. The technique of ion mobility, which measures drift times of ionic conformers dragged through buffer gas by a weak electrical field, is capable of separating 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 methods described above can alone provide an unambiguous assignment of each calculated single conformational structure, the current developments in structural elucidations are two-fold. Continuous development in computer technology and computational theory are making calculations faster, which allows for a more accurate treatment of larger molecular systems. Experimentally, one would like to combine different proven techniques and implement new approaches for measuring more and more structural constraints that are suitable for pre-selecting calculated molecular geometries and for a 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 that is capable of providing 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] research direction in cold ion spectroscopy, which targets the analytical identification of isomeric biomolecules.

Figure 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]

VIBRATIONALLY RESOLVED SPECTROSCOPY OF COLD BIOMOLECULES

Spectroscopy measures the energy and intensity of optical transitions between the quantum states of molecules, and therefore reflects their molecular structure. UV spectra of large molecules containing aromatic groups reflect 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 through calculations. Instead, high-resolution UV spectra, molecular “fingerprints”, which are rich in details, can be used to identify differently structured molecules, including isomers and conformers. In addition, UV excitation can be used for the photofragmentation of ions, serving as a convenient and highly sensitive detection tool.

One common problem of optical spectroscopy is the spectral congestion that may arise from (i) the thermal population of many molecular (including conformational) quantum states, or (ii) the broadening/splitting of transitions to excited vibrational/vibronic states due to their

Figure 2. 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 [22]

short lifetime. Overlap of these numerous transitions prohibits vibrational resolution in the 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 therefore greatly suppress this broadening (Figure 1).

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 that are known to be suitable for collisional cooling of molecular ions to temperatures of a few K. For a deeper understanding of the principles of operation, the advantages and limitations of different cold ion storage devices, we recommend a helpful book chapter by Gerlich [13], as well as some later reviews and research articles on this topic [10, 14–20]. Here we briefly summarize the most essential and practical aspects of cooling large biomolecular ions in the most popular radio frequency-driven linear ion traps.

Linear cold ion traps used for the spectroscopy of ions, including heavy biological ions (peptides and proteins), are

not available commercially. Several different designs, which differ in the number of poles, were developed over the last two decades, after the pioneering work by Gerlich [13]. Regardless of the number of poles, which range 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 were performed in 22-pole cold traps [12], which were initially developed for the spectroscopy of small ions of astrophysical interest. From the point of view of laser spectroscopy, a severe practical drawback of the 22-pole traps is their insufficiently stable radial confinement of cold ions. Because of a very flat effective radial potential in these and other multipole traps, the radial position of the ion cloud in such traps is very sensitive to mechanical imperfections and contamination of the trap poles.

These practically unavoidable circumstances can displace the ion cloud out of the trap axis, which hinders the reliable day-to-day overlap of laser beams with the cloud. Development of lower order linear cold ion traps, such as octupoles [16, 17, 21], which have well-defined radial minima (on the trap axis) of the effective RF potential, have solved this practical problem. For instance, the vibrational temperature of a protonated dipeptide, estimated from the intensity of a UV hot band ($0 \leftarrow 1$ vibronic transition) relative to the intensity of the respective $1 \leftarrow 0$ vibronic transition, appears to be below 10 K, with a 6 K octupole ion trap [17].

As biomolecules are usually large polyatomics with a variety of chemical bonds, the absorption of a single UV photon usually breaks some of these bonds. 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 using mass spectrometry for reporting the electronic absorption of the ions. The stronger is the absorption the higher is the UV-induced number of fragments and depletion of the parent ions. An ion, however, 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 2 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 [22].

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

Выставка перенесена на осень

МОСКВА
Экспоцентр '20

16-18
сентября

+7 (495) 78-601-78 | expocontrol.online

ЭКСПО КОНТРОЛЬ

12-я тематическая выставка-форум
метрологического и испытательного
оборудования и приборов

в рамках выставки

**METROLOGY
DAYS**
MOSCOW

Законодательная, научная и
практическая метрология

открытые лекции и семинары

При поддержке:

Figure 3. 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 [23]

with the IR-UV technique. Still, one may rely on the "universal" VUV/UV absorption by peptide bonds to perform photofragmentation spectroscopy (Figure 3). This recent demonstration makes cold ion spectroscopy a versatile research tool for structural studies of (eventually) any small to midsize peptide, as well as of many other types of aromatic molecules [23–24].

High-resolution Cold Ion Spectrometer in Lausanne

Implementation of all the described above spectroscopic techniques for studying cryogenically cold biological ions requires three basic types of hardware: lasers, mass spectrometers and cryogenic ion traps. As an example, Figure 4 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 a 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, which works at a pressure of ~10 mbar. The ions are guided further through another ion funnel, held at a pressure of ~1.5 mbar, and accumulate (for ~200–500 ms) in a hexapole ion trap [17]. 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 octupole 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 via 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

Figure 4. Schematic view of high-resolution Cold Ion Spectrometer in Lausanne [11–12]

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 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 using 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 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 wave-numbers, which are measured by a wavelength meter.

Figure 5 shows the 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), which is

pumped by 6–7 ns frequency-tripled pulses of a Nd:YAG laser. Alternatively, the photofragmentation can be performed by a mid-resolution ($5\text{--}8\text{ cm}^{-1}$), but widely tunable ($350\text{--}192\text{ nm}$) fast ($\sim 100\text{ ms}$ 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 a 10 Hz repetition rate. For IR spectroscopy, one (IR-UV depletion) or two (IR-IR-UV hole burning) IR OPO/OA can be used. These sources deliver up to 15 mJ pulses at a 10 Hz repetition rate over $10\text{--}2.5\text{ }\mu\text{m}$ 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 the 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 $F=50\text{ cm}$. The UV beam is loosely focused by a CaF_2 lens of $F=75\text{ cm}$ onto the trap axis $\sim 3\text{ cm}$ outside of its entrance to ensure an overlap of the entire ion cloud in the trap by the UV beam. Broadband pyroelectric detectors were 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 the scan of IR OPO wavelength. 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 is used for labeling the UV spectra.

ANALYTICAL APPLICATIONS OF COLD ION SPECTROSCOPY

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 structural identifications of biomolecules, without compromising the relevance of these results for biology. Multistage high-resolution mass spectrometry (MS^n), in combination with different dissociation techniques demonstrates an impressively accurate determination of, for instance, peptides and protein sequences. Because MS is based on measuring the mass-to-charge ratio of molecular ions and their fragments, the identification of isobaric molecules, isomers in particular, remains among the challenges for this technique [25]. Coupling of MS to high resolution LC or ion mobility (IM) often enables the separation of isomeric ions with sufficiently different retention or arrival times, respectively [26–28]. This separation does not yet ensure a reliable identification of the isobars, because both of these values are not funda-

Figure 5. Optical layout of high-resolution Cold Ion Spectrometer in Lausanne

mental to molecules and are sensitive to experimental conditions, which may not be reproduced with sufficient accuracy [29–30]. 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 photofragments that are specific to their 3D structure but absent (or not abundant) in the products of thermal-like dissociation [31–32]. In particular, this phenomenon may lead to a detectable difference between the photofragmentation mass spectra of isomeric ions and even between the spectra of conformers of an ion [31]. 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 the identification of biomolecules named 2D UV-MS fingerprinting [11]. Compared with the used in mass spectrometry UV fragmentation (UVPD) [33], performed at single VUV/UV wavelengths, 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. The spectroscopy in this region therefore adds a great number of details into the ionic 2D UV-MS fingerprints. The differences between these fingerprints may allow for recognition of 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 5. The fingerprints (Figure 6) are measured by recording fragment mass spectra at each UV laser pulse of a

Figure 6. Photofragmentation UV-MS fingerprint of a protonated dipeptide [34]

10 Hz dye laser (Figure 5), while scanning its wavelength. After normalization of the fingerprints to the UV laser fluence and to the total ion signal, correcting the baseline and removing the occasional “fake” peaks in the recorded MS, the rectified 2D data arrays (matrices) of isobaric ions are stored as a library for use in the identification of unknown isomeric ions that belong to the library. The rectification procedure is currently automated and uses a custom-modified commercial software package (Peak-by-Peak; Spectroswiss Sarl).

Library approach

Peptides and drugs. The identification and relative quantification of single isobaric ions, or their mixtures, can be accomplished using a numerical decomposition of the data arrays of the unknown ionic sample in the basic set of the library's matrices [11, 35–36]. 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.

Figure 7 shows rectified 2D UV-MS fingerprints of five model isomeric phospho-peptides TSAAATSY-H^+ , which differ only by their 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 contrast, the measured 2D UV-MS fingerprints visually differ for all five positional isomers. In order to assess the accuracy of the isomer identification by this new approach, 12 different solution mixtures that contain between one and all five isomers were prepared and their 2D UV-MS were measured.

Figure 7. Color-coded 2D UV-MS photo fragmentation fingerprints of (1–5) five singly protonated isomeric octopeptides (TpSAAATSY, pTSAAATSY, TSAAATSpY, TSAAApTSY and TSAAATpSY, respectively) and (M) 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]

The decomposition of the matrices of these test mixtures in the basis set of the five components correctly identified the peptides in all the mixtures and reproduced the relative concentrations of the peptides in the solutions with RMSD of 6.5% (Figure 8a). This number implies that, on average, any of the five library isomers that are present in a solution mixture with other isomers at relative concentrations of $> 13\%$ can be identified at the confidence level of 95% [11].

An important issue for this 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. However, a truncated fingerprint, measured for a mixture only at a few critical wavelengths may still remain very specific. Figure 8b 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 converge close to the RMSD of the full decomposition for 100 wavelengths. Analysis employing a few critical wavelengths drastically reduces the RMSD relative to results acquired with randomly selected wavelengths that approach the RMSD of the full analysis at 10–20 wavelengths. These critical wavelengths relate to the UV band origins and to the prominent peaks in the UV spectra of all library compounds. This result demonstrates that the UV-MS finger-

Figure 8. Calculated relative concentrations of five isomeric (phosphorylated and singly protonated) peptides (dots) as a function of their relative solution concentrations (0–100%) for 12 different solution mixtures (a). The RMSD of 6.5% was determined taking deviations of the calculated values from the solution concentrations (solid line). The two dashed lines show the interval of the 99.7% confidence level ($3 \times \text{RMSD}$) in calculating concentrations. 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) (b). 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 that are related to the band origins of the five library peptides [11]

printing can potentially be coupled to LC for online identifications in a broad range of applications.

In addition to these 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 fin-

gerprinting. 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 the hormone protein amylin, isoforms of mono- to hepta-saccharides, etc. Table 1 summarizes some of the results in identifica-

Table 1. 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)	[35]
Heptapeptide IAPP [31–37] (6 peptides with 4 isoforms: L/D-Asp/L/D-isoAsp in 2 positions)	Fragment of hormone peptide amylin	2.2 ^b (<4) ^c	[36]
Conformers of TyrAla-H+	Model dipeptide	0.9 ^d (blind decomposition)	[34]
Conformers of decapeptide gramicidin S	Antibiotic	– (blind decomposition) ^e	[34]
Five epimeric, anomeric and regioisomeric disaccharides ^f	Carbohydrates	5.3	[37]

^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.

^f For 12 multicomponent solution mixtures.

Figure 9. 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

tion of isomeric molecules by 2D UV-MS fingerprinting method.

Oligosaccharides. Saccharides are the biomolecules that are particularly rich 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 the development of complementary/alternative methods for the identification of isoforms of carbohydrates. Similar to HPLC in its spirit and limitations, ion mobility-mass spectrometry in different implementations was employed for the fast separation of some isoforms of carbohydrates [38–39].

There are two main problems in using 2D UV-MS fingerprinting for identifying isoforms of saccharides. First, both protonated and deprotonated glycans do not absorb in the near UV spectral region. UV absorption of deprotonated saccharides is intrinsically lifetime broadened 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 VUV spectra of ionic sugars, therefore, exhibit low isomeric specificity, which makes their identification by UV spectroscopy challenging. Second, carbohydrates exhibit low protonation efficiency, even in the 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 [40]. The complexes appear to be sufficiently stable to survive

under conditions of a mild to moderately harsh electrospray ionization source. Figure 9 shows the photodissociation and UV spectra of the isomeric trisaccharides Lewis X and Lewis A bound in aqueous solutions to the $TyrH^+$ 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 [37]. Regarding our previous experience with the decomposition of mixtures of isomeric peptides, such spectral differences alone should allow an identification of the isomers in their mixtures with an accuracy of a few percent. Interestingly, the presence of an amino group in the saccharides results in proton transfer from AAA to a non-covalently bound carbohydrates. Because the pKa of carbohydrates is lower than that of AAAs [41–43], this proton transfer may only happen either as a spontaneous or a UV-induced process in the gas phase [44, 45]. A valuable outcome of this transfer is the cleavage of internal glycan bonds. This results in the appearance of charge fragments that are characteristic to the isoforms of the saccharides. The produced photofragmentation mass spectra additionally assist in the structural characterization of glycans under study. It is important that the same instrumentation, workflow and software of the 2D UV-MS method that were developed for the identification of isomeric peptides and drugs can be used for similar identifications of saccharides.

Blind identifications of isomers and conformers

The identifications of isoforms using 2D UV-MS fingerprinting approach requires creating a library of all the isoforms/exact isobars that are expected in a process. It may readily happen that the library needed for the identification of a specific type of isomers 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). It may happen, that even the number of different isomers that are present in a mixture is unknown. It turns out that 2D UV-MS fingerprinting may detect the presence and the number of these unknown compounds and recover their UV absorption and fragmentation mass spectra using mathematical algorithms of a matrix analysis [11, 34]. The principle condition for such “blind” decomposition of the mixture fingerprint is that both the UV absorption and fragment mass spectra of all the mixed molecules are substantially different 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).

Figure 10. Application of the blind decomposition approach to a dipeptide $[YA+H]^+$: a – plot of BCV error (log scale) for 2D UV-MS spectrum of singly protonated Tyr-Ala dipeptide (Figure 6); 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 [34]

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

The first step is factorization of the 2D UV-MS data matrix D into a product of a non-negative matrix W of UV spectra and a non-negative matrix H of MS spectra of different components. If k denotes the number of such components, then the n -by- m matrix D can be factorized into a n -by- k matrix W and k -by- m matrix H , so that their matrix product WH is a lower-rank approximation to D [3]. Since all the matrices, D , W and H , are non-negative, this approach was called non-negative matrix factorization (NMF) [46]. One can find the best k -rank approximation by minimizing the Frobenius norm of the residual:

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

The second step. Use Bi-cross-validation (BCV) procedure [47] to determine the most likely number 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 D (i.e., a set of elements of D) and use the rest of D to predict the held out sub-matrix. The quality of NMF for each k is reported as a BCV error. If the k number is below the true number of com-

ponents, the BCV error is mainly determined by the NMF itself (low accuracy of the k -rank approximation), implying that the BCV error is close to the RMSD error. When the value of 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 X is used to estimate the sub-matrix D_{ij} . The overfitting of X does not lead to a better approximation to D_{ij} , because the noise functions in these two sub-matrices do not correlate. Thus, the large values of k will not result in a reduced BCV error, while the RMSD error monotonically decreases upon increasing k .

Figure 10 illustrates the application of the blind decomposition approach to a dipeptide $[YA+H]^+$, whose 2D UV-MS fingerprint is shown in Figure 6. The BCV analysis of the fingerprint (Figure 11a) 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 Figure 10b as I and II, each contain all of the UV transitions, previously assigned only to one or to the other pair of conformers, respectively [31]. 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 (Figure 10c) 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 the UV transitions (Figure 11d) in the component III that are 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 this method, which could be helpful, for instance, in the identification of unknown isomeric/isobaric metabolites.

CONCLUSIONS AND PROSPECTIVE

The frequent correlation of the conformational/isomeric form of biomolecules with their UV photofragmentation MS can be employed for the analytical identification of biomolecules. Although structures of a molecule in solution and in the gas phase are often quite different, these identifications can be

performed in the gas phase, in which one may benefit from the high spectral and mass resolutions. This approach was developed as the 2D UV-MS fingerprinting method, which has been demonstrated for the identification of isomeric peptides, drug molecules and oligosaccharides. We see further development of this and similar (e.g. IR-MS) approaches as quantitative analytical methods for the identification 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 the interrogation of a sample. In parallel, one would need to create libraries of fingerprints for the families of isomers that are challenging for separating by HPLC-MSⁿ.

Author thanks Dr. V.Kopysov for his crucial contributions to the developments of the 2D UV-MS method, Dr. N.Nagornova, Dr. A.Per-everzev and E.Saparbayev for their hard work in obtaining some of the presented herein data; A.Makarov, M.Belov and M.Gorshkov for 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.

REFERENCES

1. **Prusiner S.B.** Prion diseases and the BSE crisis // *Science*. 1997. V. 278. P. 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. V. 71. P. 577–592.
3. **Granier S., Manglik A., Kruse A.C., Kobilka T.S., Thian F.S., Weiss W.I., Kobilka B.K.** Structure of the delta-opioid receptor bound to naltrindole // *Nature*. 2012. V. 485. P. 400–404.
4. **Adrian M., Dubochet J., Lepault J., McDowell A.W.** Cryo-electron microscopy of viruses // *Nature*. 1984. V. 308. P. 32.
5. **Wüthrich K.** The way to NMR structures of proteins // *Nature Structural Biology*. 2001. V. 8. P. 923.
6. **Konrat R.** NMR contributions to structural dynamics studies of intrinsically disordered proteins // *J. Magn. Reson.* 2014. V. 241. P. 74–85.
7. **Hunt N.T.** 2D-IR spectroscopy: ultrafast insights into biomolecule structure and function // *Chem. Soc. Rev.* 2009. V. 38. P. 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. V. 99. P. 15863.
9. **Bleilholder 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. V. 3. P. 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. V. 28. P. 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. V. 87. P. 4607–4611.
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. V. 128. P. 2816–2817.
13. **Gerlich D.** Inhomogeneous RF fields: a versatile tool for the study of processes with slow ions // **Ch.-Y. Ng, M. Baer**, editors. State selected and state to state ion molecule reaction dynamics. Part I: Experiment. V. 82. Wiley, 1992. 704 p.
14. **Gerlich D.** The production and study of ultra-cold molecular ions // *Low Temperatures and Cold Molecules*. 2008. P. 295–343.
15. **Asvany O., Schlemmer S.** Numerical simulations of kinetic ion temperature in a cryogenic linear multipole trap // *Int. J. Mass Spectrom.* 2009. V. 279. P. 147–155.
16. **Hock C., Kim J.B., Weichman M.L., Yacovitch T.I., Neuemark D.M.** Slow photoelectron velocity-map imaging spectroscopy of cold negative ions // *J. Chem. Phys.* 2012. V. 137. 17.
17. **Boyarkin O.V., Kopysov V.** Cryogenically cooled octupole ion trap for spectroscopy of biomolecular ions // *Rev. Sci. Instrum.* 2014. V. 85. P. 033105.
17. **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. V. 364. P. 43–97.

16-й МОСКОВСКИЙ МЕЖДУНАРОДНЫЙ
ИННОВАЦИОННЫЙ ФОРУМ
**ТОЧНЫЕ ИЗМЕРЕНИЯ –
ОСНОВА КАЧЕСТВА И БЕЗОПАСНОСТИ**
Москва, ВДНХ, павильон 75

16th MOSCOW INTERNATIONAL
INNOVATION FORUM
**ACCURATE MEASUREMENTS –
THE BASIS FOR QUALITY AND SAFETY**
Moscow, VDNH, Exhibition hall 75

Уважаемые участники и посетители!

В связи с эпидемиологической ситуацией по COVID-19 сроки проведения Форума переносятся на более поздний период. Точные даты будут объявлены дополнительно. Благодарим за понимание!

ОРГАНИЗАТОРЫ:

СПЕЦИАЛИЗИРОВАННЫЕ РАЗДЕЛЫ ВЫСТАВКИ

METROLEXPO
МЕТРОЛОГИЯ, ИЗМЕРЕНИЯ И ИСПЫТАНИЯ

CONTROL&DIAGNOSTIC
КОНТРОЛЬ И ДИАГНОСТИКА

LABTEST
ЛАБОРАТОРНЫЕ ИСПЫТАНИЯ

PROMAUTOMATIC
ПРОМЫШЛЕННАЯ АВТОМАТИЗАЦИЯ

RESMETERING
УЧЕТ ЭНЕРГОРЕСУРСОВ

WEIGHT SALON
ВЕСОВОЙ САЛОН

ФОРУМ 2019 г. (ЦИФРЫ И ФАКТЫ):

Участники – 256 компаний из 24 стран мира
Посетители – 4870 специалистов
Площадь экспозиции – 6500 м²
Приборы и оборудование – более 2000 ед.
Докладов и презентаций – 68 чел.
Участников съезда – 1100 чел.

ДИРЕКЦИЯ ФОРУМА

Тел./факс: +7 (495) 937-40-23

E-mail: metrol@expoprom.ru

Спешите забронировать стенд
www.metrol.expoprom.ru

18. **Heine N., Asmis K.R.** Cryogenic ion trap vibrational spectroscopy of hydrogen-bonded clusters relevant to atmospheric chemistry // *International Reviews in Physical Chemistry*. 2015. V. 34. No. 1. P. 1–34.
19. **Roithova J., Gray A., Andris E., Jasik J., Gerlich D.** Helium tagging infrared photodissociation spectroscopy of reactive ions // *Accounts Chem. Res.* 2016. V. 49. P. 223–230.
20. **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. V. 140. P. 221101.
21. **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. No. 56. P. 15639–15643.
22. **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. V. 9. P. 5262–5266.
23. **Pereverzev A.Y., Koczor-Benda Z., Sapaeva E., Kopysov V.N., Rosta E., Boyarkin O.V.** Spectroscopic evidence for peptide-bond-selective ultraviolet photodissociation // *The Journal of Physical Chemistry Letters*. 2019. P. 206–209.
24. **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. V. 3. P. 61–67.
25. **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. V. 72. P. 391–395.
26. **Winter D., Pipkorn R., Lehmann W.D.** Separation of peptide isomers and conformers by ultra performance liquid chromatography // *J. Sep. Sci.* 2009. V. 32. P. 1111–1119.
27. **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. V. 83. P. 6918–6923.
28. **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. V. 16. P. 1239–1249.
29. **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. V. 64. P. 121–127.
30. **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. V. 127. P. 154322–154327.
31. **Kopysov V., Nagomova N.S., Boyarkin O.V.** Identification of tyrosine-phosphorylated peptides using cold ion spectroscopy // *J. Am. Chem. Soc.* 2014. V. 136. P. 9288–9291.
32. **Brodbeck J.S.** Photodissociation mass spectrometry: new tools for characterization of biological molecules // *Chem. Soc. Rev.* 2014. V. 43. P. 2757–2783.
33. **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. V. 7. P. 1067–1071.
34. **Kopysov V., Makarov A., Boyarkin O.V.** Identification of isomeric ephedrine by cold ion UV spectroscopy: toward practical implementation // *Anal. Chem.* 2017. V. 89. P. 544–547.
35. **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. V. 143. P. 833–836.
36. **Sapaeva E., Kopysov V., Yamaletdinov R., Pereverzev A., Boyarkin O.V.** Interplay of H-bonds with aromatics in isolated complexes identifies isomeric carbohydrates // *Angew Chem. Int. Ed. Engl.* 2019. V. 58. P. 7346–7350.
37. **Liu Y., Clemmer D.E.** Characterizing oligosaccharides using injected-ion mobility/mass spectrometry // *Anal. Chem.* 1997. V. 69. P. 2504–2509.
38. **Hofmann J., Hahn H.S., Seeberger P.H., Pagel K.** Identification of carbohydrate anomers using ion mobility-mass spectrometry // *Nature*. 2015. V. 526. P. 241.
39. **Vrkic A.K., O'Hair R.A. J.** Using non-covalent complexes to direct the fragmentation of glycosidic bonds in the gas phase // *J. Am. Soc. Mass. Spectr.* 2004. V. 15. P. 715–724.
40. **Gorman G.S., Speir J.P., Turner C.A., Amster I.J.** Proton affinities of the 20 common α -amino acids // *J. Am. Chem. Soc.* 1992. V. 114. P. 3986–3988.
41. **Voleti N., Vairamani M.** Proton affinity differences among three N-acetylhexosamines studied by the kinetic method // *Rapid Commun. Mass. Spectrom.* 2003. V. 17. P. 1089–1091.
42. **Gronert S., Simpson D.C., Conner K.M.** A reevaluation of computed proton affinities for the common α -amino acids // *J. Am. Soc. Mass Spectrom.* 2009. V. 20. P. 2116–2123.
43. **Agmon N.** Elementary steps in excited-state proton transfer // *J. Phys. Chem. A*. 2005. V. 109. P. 13–35.
44. **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. V. 117. P. 163002.
45. **Lee D.D., Seung H.S.** Learning the parts of objects by non-negative matrix factorization // *Nature*. 1999. V. 401. P. 788–791.
46. **Owen A.B., Perry P.O.** Bi-cross-validation of the SVD and the nonnegative matrix factorization // *Ann. Appl. Stat.* 2009. V. 3. P. 564–594.

27–30.10.2020

www.chemistry-expo.ru

23-я международная
выставка химической
промышленности
и науки

ХИМИЯ

ХИМИЯ

Генеральный
информационный
партнер:

ЦЕНОВОЕ АГЕНТСТВО
ХИМ
КУРЬЕР
www.chem-courier.com
КОНФЕРЕНЦИИ

Инновации
и современные
материалы

Нефтегазохимия

Startup ChemZone

Автоматизация
и цифровизация
производства

Химмаш. Насосы

Хим-Лаб-Аналит

Зеленая химия

Индустрия пластмасс

Защита от коррозии
«КОРРУС»

12+
Реклама

Организатор:
ЭКСПОЦЕНТР

Россия, Москва, ЦВК «ЭКСПОЦЕНТР»