Identification and Quantification of Any Isoforms of Carbohydrates by 2D UV-MS Fingerprinting of Cold Ions

Erik Saparbaev,[1] Vladimid Kopysov,[1] Viktoriia Aladinskaia,[1] Vincent Ferrieres,[2] Laurent Legentil,[2] and Oleg V. Boyarkin[1]*

[1]Laboratoire de Chimie Physique Moléculaire, École Polytechnique Fédérale de Lausanne, Station-6, 1015 Lausanne, Switzerland

[2] Ecole Nationale Supérieure de Chimie de Rennes, 11 allée de Beaulieu - CS 5083, 35708 Rennes, France.

ABSTRACT

Biological functionality of isomeric carbohydrates may differ drastically, making their identifications indispensible in many applications of life science. Due to the large number of isoforms, structural assignment of saccharides is challenging and often requires a use of different orthogonal analytical techniques. We demonstrate that isomeric carbohydrates of any isoforms can be distinguished and quantified using solely the library-based method of 2D UV-MS photofragmentation of cold ions. The two-dimensional "fingerprint" identities of UV transparent saccharides were revealed by photofragmentation of their non-covalent complexes with aromatic molecules. We assess the accuracy of the method by comparing the known relative concentrations of isomeric carbohydrates mixed in solution with the concentrations that were mathematically determined from the measured in the gas phase fingerprints of the complexes. For the tested sets with up to five isomers of di- to heptasaccharides, the root-mean-square deviation of 3-5% was typically achieved. This indicates the expected level of accuracy in analysis of unknown mixtures for isomeric carbohydrates of similar complexity.

Carbohydrates are ubiquitous in nature. Their various biological functions range from being a fuel for our brain and signalling the state of living cells to antimicrobial defence of infants[1] and shielding viruses from recognition by human immune system.[2] The tremendous isomeric diversity of carbohydrates allows nature to assign glycans of specific structures to each function of this variety, but it also makes identification of these isoforms challenging.[3] The composition of glycans of the same mass may differ by the monosaccharide units this polymer is composed of, by their order, linkage (1→3, 1→4, etc. and α/β) and branching. The structural difference between the isomeric monosaccharide units includes orientation of hydroxyl groups (epimers and α/β anomers), the ring size (e.g., pyranose and furanose) and the absolute configuration (*D* or *L*). Natural modifications, such as N-acetylation, phosphorylation, etc., further multiply the isomeric diversity of carbohydrates *in vivo*.

There is no single analytical approach capable of distinguishing carbohydrates in all their numerous isoforms.[4] The time-proven methods of x-ray diffraction, NMR[5] and, since recently, cryo-electron microscopy[6,7] are capable to provide direct structural information for carbohydrates, although each of these techniques has its own limitations and practical drawbacks. The most common method of structural analysis of isomeric carbohydrates involves combination of enzymatic digestion for a controlled cleavage of large oligosaccharides, liquid chromatography (LC) to separate the produced isomers and single/tandem mass spectrometry (MS) to identify them.[8] Despite its wide time-proven applicability, the method has certain limitations, such as the use of compound-specific conditions for separation, the need for chemical derivatization (e.g., by reductive amination and/or permethylation) and, often, an insufficient resolution in separating multicomponent mixtures of oligosaccharides.[9,10] Ion mobility spectrometry (IMS) is another technique that is capable of separating isomeric ions that exhibit different collision cross sections.[11] Structural difference between isomeric glycans is often insufficient however for their separation in complex mixtures,[12] although the technique is demonstrating a rapidly increasing selectivity.[13] More generally, the values measured by chromatography and ion mobility (retention and arrival times, respectively) are not fundamental to molecules and dependent on specific experimental conditions and their stability (e.g., type of chromatographic column, pressure, temperature, etc.),[14-16] This may limit reproducibility in high-resolution measurements, making a use of internal calibrants indispensible.

2D UV-MS fingerprinting is a recently developed approach that is capable of highly accurate identification of isomeric biomolecules.[17-22] This analytical technique integrates UV

photofragmentation (UVPD) spectroscopy of cryogenically cooled ions with broadband high-resolution (e.g. Orbitrap-based) mass spectrometry. It measures abundances of all UV-induced photo fragments at once in function of UV wavelength, thus providing a two-dimensional data array, named 2D UV-MS fingerprint, as a characteristic of an ion. Cryogenic cooling of mid-size biomolecular ions often enables vibrational resolution in their electronic spectra,[23] which makes them highly distinguishable. The UVPD mass spectra may also exhibit certain isomeric specificity. The synergy of the two techniques, spectroscopy and mass spectrometry, makes 2D UV-MS fingerprints highly individual to structure of ions, including isomers and conformers. Different from LC and IMS, optical spectroscopy reflects molecular transitions between quantum states, which are fundamental to ions. This makes 2D UV-MS fingerprints the highly reproducible standards that can be shared across different laboratories. The high performance of 2D UV-MS cryogenic fingerprinting was earlier demonstrated in library-based identification and accurate quantification of conformers and isoforms of different peptides in their multicomponent mixtures,[17,18,20,24] as well as of small isomeric drugs on a concentration scale of a few ng/mL.[19] Here we demonstrate the use of 2D UV-MS cryogenic fingerprinting for identification and quantification of natural carbohydrates for all types of their numerous isoforms. The tested sets of isomers range from mono- to hepta-saccharides and from two to five compounds in a set. The accuracy of relative library-based quantifications for some of the sets was determined by fingerprinting several multi-compound solution mixtures of isomeric molecules. Finally, we demonstrate that in certain cases the 2D UV-MS method can be reduced to the technically less demanding one-dimensional UV fragment spectroscopy, yet without a significant loss of accuracy in the quantification.

A straightforward application of 2D UV-MS fingerprinting to glycans is challenging, because they do not absorb in UV. Instead, we make use of non-covalent complexes of glycans with protonated aromatic molecules, such as phenylalanine, tyrosine and its decarboxylated analog tyramine.[22,25-27] These complexes can be readily prepared by dissolving a glycan and an aromatic molecule in a standard for MS methanol/water solution. The complexes appear to be strong enough to survive electrospray ionization that brings them to the gas phase. Our estimate suggests that, under mild conditions of our electrospray, up to half of the dissolved N-acetylated glycans bind to the aromatics and appear in the gas phase as protonated complexes. The isomeric structure of a guest carbohydrate is communicated to the host aromatic molecule by non-covalent interactions, which induce changes in the UV absorption of the latter. In particular, an interplay of different types of H-bonds in isomeric

complexes makes the position of their UV absorption onset very sensitive to the isoform of a guest carbohydrate.[22] We also found that the number of low-energy conformers is much lower for non-covalent complexes of glycans with aromatic molecules than for glycan–alkaline ion adducts,[27,28] which are often used for identifications of isomeric carbohydrates by IR spectroscopy.[29] This fortunate circumstance slows spectral congestion in UV spectra of cold carbohydrates upon increase of their size. The conserved vibronic structure enables high spectral isomeric specificity of 2D UV-MS fingerprints. The MS isomeric specificity arises from the phenomena of UV-induced proton transfer from host ion to guest molecule in the complexes.[30] The transfer results in fragmentation of glycans, which often yields isomer-specific pattern of fragment-MS.

EXPERIMENTAL

Chemicals. All the carbohydrates, except for furanoside Gal*f*NAc, were purchased from Carbosynth (>95% purity), ROTH (>98% purity) and Acros Organics (>98% purity), and used without any further purification. The Gal*f*NAc glycan is not commercially available and was synthesized according to the procedure described elsewhere[31] and purified to >95%. Aromatic host molecules of >98% purity were purchased from Sigma-Aldrich and TCI. All the solvents were of HPLC grade and acetic acid was of >99% purity. Protonated gas-phase carbohydrate-aromatic complexes were produced from $5 \cdot 10^{-5}$ M equimolar solutions of a carbohydrate and an aromatic molecule in the 1:1 water-methanol mixture with 0.5 % of acetic acid.

Apparatus. Apart from a few minor changes, the details of our experimental setup has been describes elsewhere.[32] Briefly, the ions, produced in the gas phase from solution by a nano-electrospray ionization (n-ESI) source, enter an electrodynamic ion funnel (IF) orthogonally to its axis through a 100 mm long stainless steel capillary of 0.7 mm ID. The ions are then turned by 90° to travel along the IF axis at pressure of 6 mbar and exit the funnel through the conductance limit of 2.5 mm diameter. They subsequently pass through two skimmers (1.5 and 2 mm in diameter), which allow for a gradual reduction of pressure from 0.5 mbar to 10^{-4} mbar. The ions are then accumulated and thermalized in an octupole ion trap for 97 ms, before they are released and pass through a quadrupole mass filter (Q1), which is set to select the ions of interest. Mass-selected ions are then turned by 90° using an electrostatic bender, focused by a stack of three electrostatic lenses and moved through an RF octupole guide into a cold octupole trap,[33] which is kept at 6 K. The trap is driven by two 1

MHz sinus waveforms with peak-to-peak amplitudes of 50-100 V. The stored in the trap non-covalent complexes are cooled down to ~10 K in collisions with He buffer gas,[33] which is pulsed into the trap shortly before arrival of the ions. Approximately 85 ms later the ions are irradiated by a pulse of UV light (1.8±0.7 mJ/pulse, 5 ns duration, ~6 cm^{-1} spectral linewidth), produced by a UV optical parametric oscillator (OPO; NT 342C, EKSPLA) and undergo photo fragmentation. The fragment and parent ions were 90° turned by the second electrostatic quadrupole bender either toward a highly sensitive quadrupole mass spectrometer (QMS) or toward high-resolution broadband Orbitrap-based MS (Exactive, Thermo Fisher). UVPD spectra and 2D UV-MS fingerprints are measured by continuously recording the yield of the photofragments using the QMS and the Orbitrap-based MS, respectively, while scanning UV wavelength. The repetition rate of the cooling/fragmentation cycle was determined by the 10 Hz repetition rate of the OPO, such that the parent ions experienced only one OPO shot in each cycle. At each UV wavelength the yield for a single fragment (with QMS) or the entire fragment mass spectrum was measured in 10 cycles and averaged to give a data point in UVPD spectrum or a fragment MS in 2D UV-MS fingerprint, respectively.

Special measures have been taken to narrow the variation of pulse energy of the OPO over the wide tuning range (210-350 nm). The angular positions of non-linear crystals in the OPO doubling/mixing stages were detuned from their optimal phasematching angles, such that the output energy at each wavelength was close to the minimum energy attainable within the desired spectral range. This minimizes the potential contributions to UVPD from non-linear absorption and therefore increases reproducibility of the fingerprints. In addition, the measured ion signals were normalized to OPO pulse energy, which was measured by a broadband pyroelectric detector.

In measurements with Orbitrap-MS, the whole fragment mass spectrum was normalized to the total ion signal detected in the cycle. In QMS measurements we employ 20 Hz cooling cycle and detect fragment or parent ions in 10 Hz alternative cycles with and without UV pulse, respectively.

Data processing. The measured 2D UV-MS data arrays (intensity vs UV wavelength and m/z) were, first, pre-processed for detection of mass-peaks using the PeakbyPeak software package (Spectraswiss). The reduced 2D data arrays of the known isomeric compounds were stored as a library tagged by the exact mass. 2D UV-MS fingerprints of solution mixtures of (unknown) isomeric molecules were pre-processed the same way and then mathematically decomposed in the basis set of the 2D matrices from the library tagged by the mass of the

isomers. The details of the matrix decomposition procedure are described elsewhere.[17,19] The same software is used for decomposition of UVPD spectra measured with QMS.

RESULTS AND DISCUSSION

Monosaccharide isomers. The first but essential step toward structural identification of isomeric oligosaccharides is to distinguish their structural units – cyclic monosaccharides. These building blocks are classified by five types of isomerism as epimers, α/β-anomers, *D/L*-enantiomers, aldose/ketose and ring-size isomers.

Aldose-ketose. Glucose and galactose example the most common mammalian aldose; ketoses (e.g. fructose) occurs widely in plants. Aldoses have a terminal formyl COH group, while ketoses contain a carbonyl group bound to two carbons. Different chemical properties of these isomers allow their reliable identification by a variety of techniques.[34] Nevertheless, for the sake of completeness we tested the UVPD spectroscopy approach for distinguishing isomers of this type. Figures 1a and 1b show UVPD optical spectra of protonated aromatic molecule tyramine (Trm) in non-covalent complexes with *D*-glucose (*D*-Glc) and *D*-fructose (*D*-Fru), respectively. Despite the very close positions of the onsets of UV absorption, the very different shapes, widths and positions of the measured vibronic transitions make the spectra readily distinguishable. Regarding the relative simplicity and similarity of the UV fragment MS for both isomeric complexes, a use of a more complex method of 2D UV-MS fingerprinting seems to be unnecessary in this case.

Epimers. Different from aldose-ketose isomers, epimers are stereoisomers that differ in absolute configuration about single of their several asymmetric carbons. For example, N-acetyl-D-glucopyranose (GlcNAc) and N-acetyl-D-mannopyranose (ManNAc) differ only in the configuration at the C2 position, whereas N-acetyl-D-galactopyranose (GalNAc) is an epimer of N-acetyl-D-glucopyranose at the C4 position. Figure 1 (c-e) compares UVPD optical spectra of these three isoforms. The revealed differences in the spectral shapes and/or the onsets of UV absorption make the spectra unambiguously specific to these isomers. In opposite, the UVPD MS of the isomers do not exhibit any clear specificity.

When complexes of all isomers of a library exhibit sharp and well separated (compared with the width of the rising edge of the absorptions) onsets of their UV absorptions, the positions of the onsets alone can be used as reliable spectroscopic tags of these isomers. The onsets of unknown library compounds that are mixed together in solution can be located by measuring the UVPD yields at a few but “critical” pre-selected wavelengths only, instead of

measuring the whole continuous spectrum at, typically, hundreds of wavelengths.[17,19,20] Such approach may shorten the measurements to a few seconds, which would allow for performing isomeric identifications online with HPLC.

Anomers. Anomerism is a particular case of epimerism, where structural variations are about the hemiacetal/hemiketal carbon in a carbohydrate ring (e.g. about C1 for aldoses). Cyclisation of an open-chain carbohydrate results in one of the two possible anomeric forms, notated as α- and β-. The anomers coexist at equilibrium in aqueous solutions and spontaneously interconvert upon reopening-closing the ring. Methylation of the anomeric 1-OH group inhibits this conversion. Noticeable differences between IR spectra of α/β anomers were earlier revealed for neutral and protonated complexes of aromatics with MeGalNAc and with GalNAc.[26,27] Figures 1f-1g illustrate how much UVPD spectra of complexes of $TrmH^+$ with methylated glycans α-MeGalNAc and β-MeGalNAc differ in the onsets of UV absorption, but also in many details of highly structured absorption bands. These differences ensure an unambiguous spectral identification of the two anomers. An earlier detailed study revealed that in complexes with GalNAc the anomeric 1-OH group, which position is the only structural difference between the α/β isomers, is not involved to any non-covalent bonds with the host aromatic and therefore cannot directly influence the UV absorption of Trm. Instead, the isomeric information is communicated by the guest to the host through an inductive polarization of the H-bond, formed between the hydrogen of 6-OH in the glycan and the oxygen atom of the tyramine side chain.[27] A stronger polarization of the bond results in a larger blue shift of UV absorption by the aromatic, whose ring is conjugated with the hydroxyl.

Ring-size isomers. Upon cyclization of open-chain glycan molecules, monosaccharides can form rings with different number of carbon centers. For example, N-acetyl-galactosamine may form five- or six-membered rings, named furanose (Gal*f*NAc) and pyranose (Gal*p*NAc), respectively. Although, forming pyranoses is energetically more favorable, the furanose isomers are also observed in nature as units of polysaccharides.[35] Figures 1g and 1h compare UVPD spectra of protonated tyramine in complexes with methylated N-acetyl-galactosamine in pyranose (β-MeGal*p*NAc) and furanose (β-MeGal*f*NAc) forms. The methylation was required to lock the carbohydrates in their β-anomeric form. Overall, the significant, ~5 nm, spectral shift of the electronic band origins and the very different vibronic structures of the two spectra make them readily distinguishable. The spectra thus can be used for a reliable visual identification of the isomers.

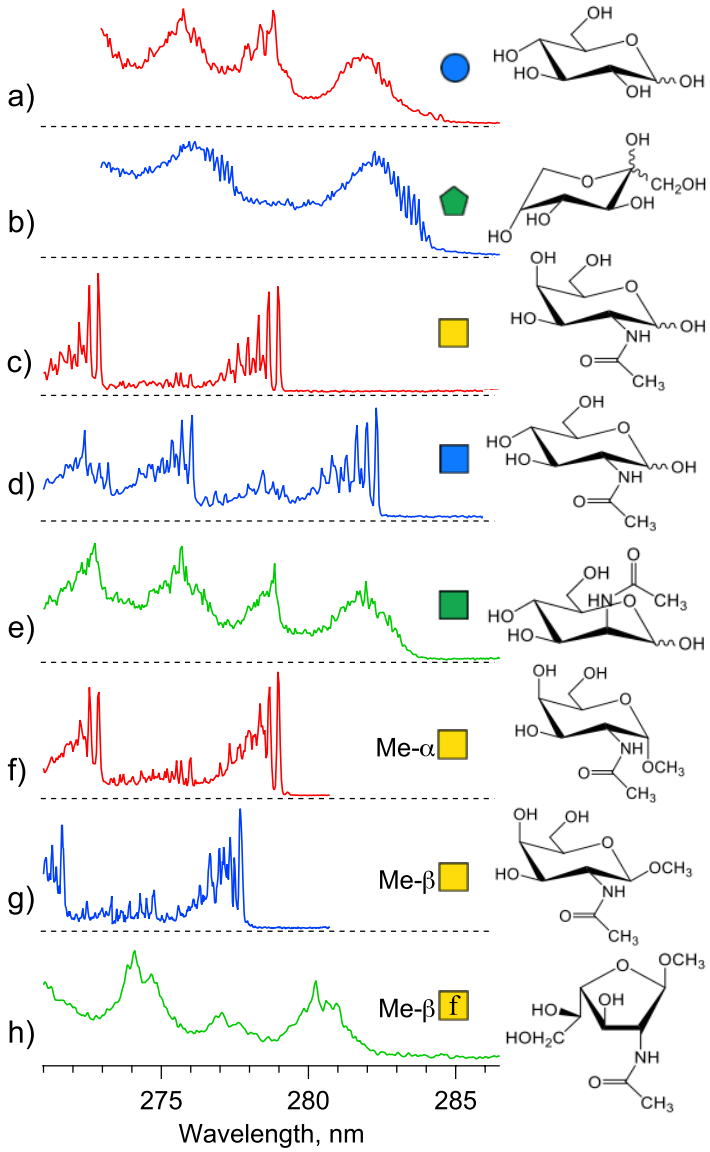

Figure 1. UVPD spectra of cold non-covalent complexes of protonated tyramine ($TrmH^+$) with monosaccharides: a) *D*-Glc, b) *D*-Fru, c) *D*-GalNAc, d) *D*-GlcNAc, e) *D*-ManNAc, f) α-MeGalNAc, g) β-MeGalNAc, h) β-MeGal*f*NAc. The spectra (a) and (b) were measured by detecting $TrmH^+$ photofragment with QMS; the spectra (c) to (h) were generated from the 2D UV-MS fingerprints measured with the Orbitrap-based MS by integrating them over m/z dimension. The standard color/shape-coded symbols of the saccharides and their structures are shown on the right.

Enantiomers. When a carbohydrate and its mirror image cannot be superimposed, such two isomers are called *D* and *L* enantiomers of this chiral molecule. Except for circular dichroism,[36,37] enantiomers of glycans have exactly the same physical and optical properties and therefore identical spectroscopy. Complexes of *D*/*L* glycans with tyramine, which is an achiral molecule, must adopt exactly the same geometry, such that their UV spectra is to be indistinguishable too. Figures 2a and 2b illustrate this fundamental property of enantiomers,

comparing UVPD spectra of Trm in complexes with *D*/*L*-glucose. Apart from some small deviations that characterize the reproducibility of the experiment, the two spectra are, indeed, identical. The *D*/*L* isoforms may drastically differ however in interactions with chiral molecules. This structural property allows, for instance, performing separation of enantiomers by chiral HPLC or IMS.[38] Similarly, non-covalent complexes of *D*/*L* glycans with a chiral aromatic should have different geometries and therefore distinguishable spectroscopy.

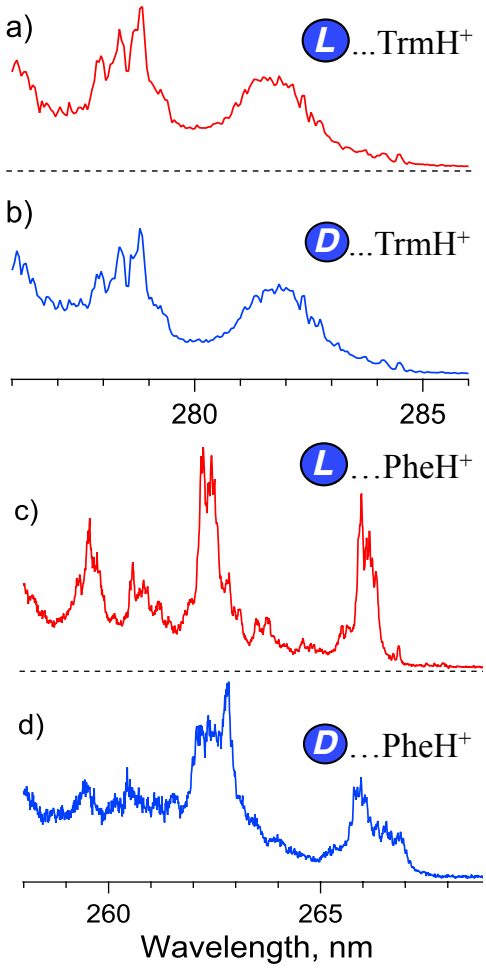

Figure 2. Photofragmentation UV spectra of the complexes of enantiomeric *L*/*D* glucose with non-chiral (Trm) and chiral (*L*-Phe) chromophores.

Figures 2c and 2d show UVPD spectra of *L*-Phe, which is a chiral molecule, in complexes with *D*/*L*-Glc. Different from Trm, the side chain of Phe does not have an OH group, whose H-binding to glycans induces the most significant differences in UV spectra (e.g. shifts of the electronic band origins) of the complexes.[22] Nevertheless, the change of the host molecule makes the spectra of *D*/*L* -Glc…$PheH^+$ complexes clearly different, which is likely due to interplay of the relatively weak OH-π and CH-π interactions.

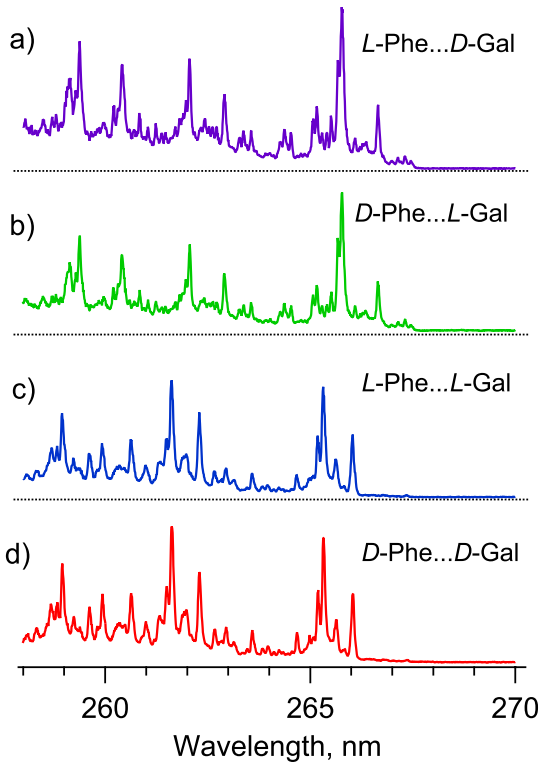

Figure 3. UVPD spectra of protonated *L*/*D*-Phe with *L*- and *D*-enantiomers of Gal.

Fundamentally, the spectra of complexes of a chiral aromatic with a carbohydrate must remain the same upon a simultaneous change of chirality of both partners: $D \leftrightarrow L$. Figure 3 compares UVPD spectra of Gal–PheH$^+$ complexes for all four combinations of chirality. Within the accuracy of the experiment, the spectra of *L*-*D* and *D*-*L* complexes look, indeed, identical (Figures 3a and 3b) Similarly, the complexes with *D*-*D* and *L*-*L* combinations of enantiomers are also indistinguishable by spectroscopy (Figures 3c and 3d), while their spectroscopic signature is clearly different from that of the complexes with the molecules of the mixed chirality. Apart from curiosity, the analytical implication of this comparison is in the possibility to detect appearance of a previously unobserved enantiomer (e.g. in asymmetric catalysis) using UV spectroscopy of its existing mirror image in complex with a chiral host molecule. The complementary enantiomer then can be identified by spectroscopy of the host of the opposite chirality in complex with the suspected isomer.

Overall, the presented above data demonstrate that monosaccharides of any type of isoformerism can be distinguished by cold-ion UVPD spectroscopy of the easily prepared in solution glycan–aromatic non-covalent complexes.

Disaccharide isomers. In di–, oligo– and polysaccharides the glycosidic linkage may occur (i) between the same or different monosaccharaides and in different order (*compositional* isomers), (ii) at different points of attachments (*connectivity* isomers) and (iii) with different orientations of glycosidic bonds (*configurational* isomers). In addition, large glycans may form not only linear, but also branched polymeric structures, which further add to the diversity of the connectivity isomers. Furthermore, glycans can naturally undergo different modifications, such as, for instance, N-acetylation, phosphorylation and sulfation. The modifications can be multiple and occur on different units of the polymers (*regioisomers*), thus increasing the number of possible isomers. All these additional features multiply the already high isomeric diversity of monosaccharaide building blocks.

Compositional. A disaccharide with monoisotopic mass of 342.12 Da can be composed, for instance, of either glucose or galactose or mannose, linked in different orders. UVPD fragment mass spectra of the complexes of these disaccharides with protonated Trm or Tyr are still low informative and contain, mainly, the protonated aromatics and their fragments (Figure 4). This is similar to the UVPD MS of the monosaccharaides that have no NAc group. The presence of this group, which has high proton affinity, enables proton transfer from the host aromatic and, subsequently, fragmentation of the guest glycan. The isomer-specific content of the 2D UV-MS fingerprints for the non-acetylated disaccharides remains, largely, in the spectroscopic domain. Figures 4a-4c illustrate the difference between the UV spectra of compositional isomers of disaccharides. All three glycans exhibit clearly distinct spectroscopic signatures, which can be used for identification of the isomers.

Configurational. Maltose and cellobiose (α/β-Glc–1,4-Glc) differ only by the orientation of glycosidic linkage between a pair of monosaccharaide units (figures 4c and 4d, respectively). Nevertheless, their spectroscopic signatures in complexes with $TyrH^+$ differ substantially. The pair of isomers of mannobiose (a/β-Man–1,4-Man) in figures 4b and 4f also exhibits well distinguishable UV spectra that can be used for identification of these disaccharides.

Connectivity. In these isomers the two monosaccharaides are linked by a glycosidic bond between carbon atoms at different positions (e.g. 1→2, 1→3, 1→4 etc.). In the case of the disaccharides containing, for instance, only mannoses as building blocks and just for the α-configuration of the glycosidic bond, four connectivity isomers are naturally abundant: α-Man-1,2-Man, α-Man-1,3-Man, α-Man-1,4-Man, α-Man-1,6-Man. 2D UV-MS fingerprinting method is capable of distinguishing all these isoforms. Figures 4e-4h show the

UVPD spectra of these isomers in complexes with *L*-tyrosine. The spectra differ significantly from each other in the positions of the onset of UV adsorption and/or in the shapes of the vibronic bands. Overall, each of the spectra in Figure 4 is unique and can be used as a spectroscopic tag for identification of the isomers.

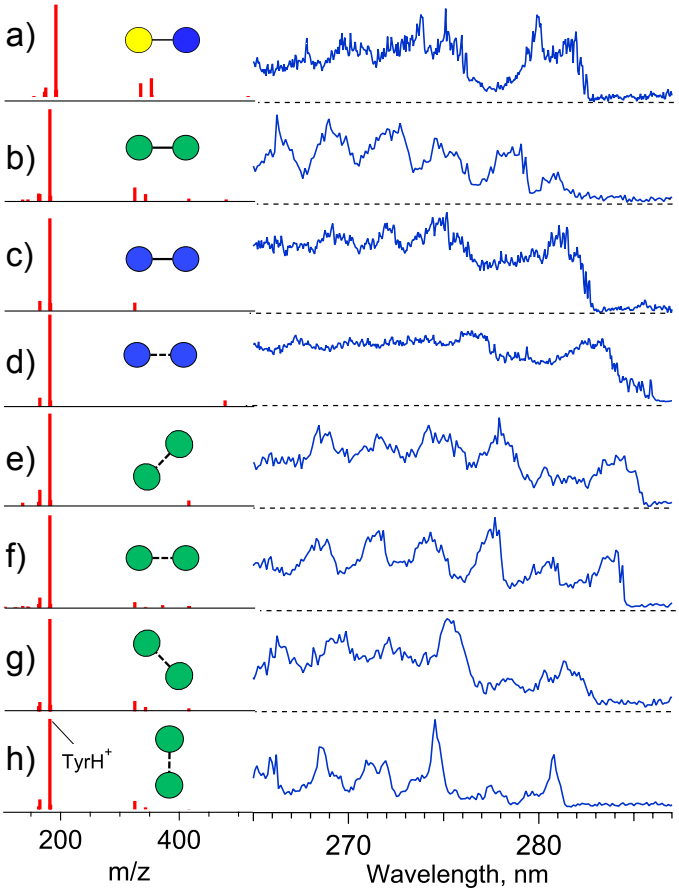

Figure 4. UVPD optical and mass spectra, generated by integrating over m/z and the wavelength, respectively, the 2D UV-MS fingerprints of complexes of tyrosine (*L*-Tyr) with isomeric disaccharides (m/z=524.31Th): a) β-Gal–1,4-Glc, b) β-Man–1,4-Man, c) β-Glc–1,4-Glc, d) α-Glc–1,4-Glc, e) α-Man–1,3-Man, f) α-Man–1,4-Man, g) α-Man–1,6-Man, h) α-Man–1,2-Man. Conventionally, dashed and solid lines used in the symbolic representation indicate α and β orientations of glycosidic bond, respectively; the orientation of the lines indicates the connectivity points.[39]

Regioisomers. Figures 5a and 5b compare UVPD optical and mass spectra of $TyrH^+$ in complexes with disaccharide regioisomers β-GalNAc-1,3-Gal and β-Gal-1,3-GalNAc, which differ only by the location of N-acetylation. Figure 5 also compares UVPD identities of the disaccharides that differ by combinations of two and even three types of isomerism: regio- and configurational (5a vs 5c), regio- and compositional (5b vs 5d) and regio-, compositional and configurational (5c vs 5e). In addition to the apparent isomeric specificity of the optical

spectra, the UVPD fragment mass spectra of different isomers also exhibit a clear difference in presence and abundance of many photo fragments. The presence of NAc group, which has high proton affinity, enables UV-driven proton transfer from the host aromatic to a guest glycan with subsequent fragmentation of the latter (Fig. 5). The transfer greatly enriches the fragment MS, which appear to be isomer-specific. The fact, that the 2D UV-MS fingerprints of the amino glycans become isomer-specific in both the wavelength and the m/z dimensions, greatly improves the reliability of the method in identifications of single isomeric carbohydrates, but also increases its accuracy in quantification of their multicomponent mixtures (see below).

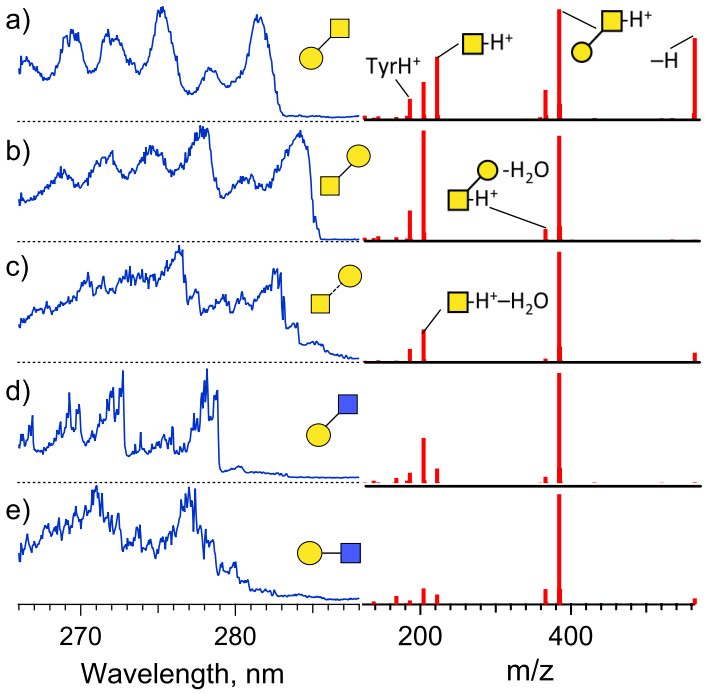

Figure 5. UVPD optical and mass spectra generated by integrating over m/z and wavelength, respectively, the 2D UV-MS fingerprints of non-covalent complexes of *L*-TyrH$^+$ with a) β-Gal-1,3-GalNAc, b) β-GalNAc-1,3-Gal, c) α-GalNAc-1,3-Gal, d) β-Gal-1,3-GlcNAc, e) β-Gal-1,4-GlcNAc isomeric N-acetylated disaccharides. Some intense fragment MS peaks are labelled using standard colour/shape presentation of carbohydrates and chemical symbols.

Oligosaccharides. The isomeric diversity of glycans grows rapidly with their size due to the increasing number of combinations for composition and connectivity of monosaccharaides. This implies that identification of oligosaccharides, as compared with mono-/disaccharides, requires more selective analytical methods. Two circumstances may, potentially, inhibit a use of 2D UV-MS fingerprinting of oligosaccharide–aromatic complexes for identification of

large glycans. Single aromatic, which size is comparable with size of a monosaccharaide, may not be capable to sense isoforms of the oligosaccharide units that are remote from the docking site of the aromatic molecule in the complexes. In addition, for larger glycans one may expect a larger number of conformers. The overlap of their slightly different UV absorptions would increase the inhomogeneous broadening of transitions in (non-conformer specific) UVPD optical spectra of the complexes. The broadening may wash out any details in the spectra, potentially, making them low isomer-specific. Figure 6 compares UVPD optical and mass spectra derived from 2D UV-MS fingerprints of trisaccharides. The optical spectra still remain sufficiently different to distinguish the three isomers easily. Compared with the disaccharides in Figure 5, UVPD of these N-acetylated trisaccharides results in rich fragment MS, which contain several highly abundant fragments. Although most of them are the same for all three isomers, the overall patterns of the MS become well distinguishable too. The glycans as large as tetra- and even hepta-saccharides also exhibit quite distinct 2D UV-MS fingerprints[22] (Figures S1 and S2). We explain the ability of a small aromatic to sense isoforms of a large glycan by non-covalent binding of the former to different sites of the latter. For each particular configuration of the complex its UVPD would reflect only the local isoform of a glycan close to the binding site. The contributions from different available configurations collectively may reflect all the isomeric features of a glycan in its 2D UV-MS fingerprint.

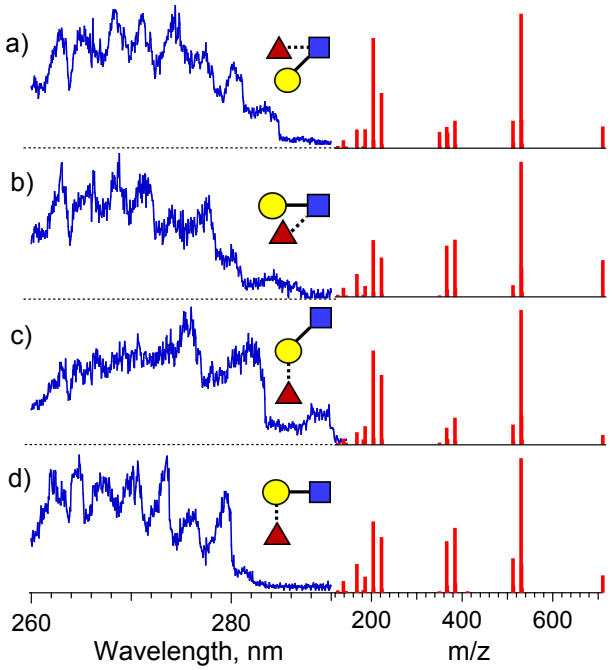

Figure 6. Photofragmentation UVPD optical and mass spectra, generated by integrating over m/z and wavelength, respectively, the 2D UV-MS fingerprints of non-covalent complexes of

L-TyrH$^+$ with trisaccharides a) Lewis A, b) Lewis X, c) Blood Group H1, d) Blood Group H2.

Along with 2D UV-MS fingerprinting, over last few years different types of gas-phase IR spectroscopy also have been tested for recognition of single isomeric carbohydrates[29,40-42] Visually distinct IRPD and IRMPD spectra were demonstrated for some isoforms of glycans, making a promising ground for further developments of these approaches as analytical methods. The latter requires however not only some visually distinct identities of isomers, but a way to quantify them in solution mixtures and an assessment of accuracy for such quantifications. While this is likely still in agenda of IR spectroscopy-based approaches, the library-based 2D UV-MS fingerprinting technique has already demonstrated accurate quantifications of relative concentration of isomeric peptides and drug molecules in their multicomponent solution mixtures.[17-20,32] Herein we use the same instrumentation and workflow for quantification of isomeric carbohydrates mixed in solution, but perform fingerprinting of their non-covalent complexes with aromatic molecules.

Quantification of isomers. First, 2D UV-MS fingerprints of the complexes of an aromatic molecule with each isomer of interest were measured and processed to create a library of 2D data arrays labeled by the mass of these isomers. Next, a fingerprint of the isomeric complex produced from a solution mixture of unknown number of the isomeric glycans of unknown relative concentrations was measured too. For each solution mixture its 2D data array was decomposed in the basis set of the isomers of the respective library using mathematical methods of non-negative matrix decomposition. The fit coefficients then give the relative concentrations of the isomers in the solution mixture.

We applied this workflow to 12 solution mixtures of one to three (out of five) isomeric disaccharides and an aromatic molecule tyramine (Trm), which was added to the solution for forming complexes with the carbohydrates. The use of Trm instead of Tyr as a host aromatic was not critical, although the lack of the –COOH group in tyramine makes UVPD optical spectra of the complexes more structured and therefore more isomer-specific. Figures 7a and 7b illustrate the quality of matrix decompositions for a 50:50 % solution mixture of β-Gal-1,3-GlcNAc and β-Gal-1,4-GlcNAc disaccharides. The figures compare the optical and mass spectra visualized by integrating over m/z and wavelength dimensions, respectively, the measured and the calculated 2D UV-MS matrices. The five fit coefficients of the decomposition imply 0%, 1.2%, 0%, 47.5% and 51.3% relative concentrations for the five isomeric disaccharides listed in Fig 7c (from top to bottom, respectively). The identity of the

two mixed isomers was determined correctly; their calculated concentrations differ from the prepared solution ones by 2.5% at most. The analysis shows that, in average over the prepared pool of 12 mixtures (Table S1), relative concentrations of these five isomers have been determined with 5.3% accuracy (Fig. 7c). The error is substantially higher (8% and 15%) for the two mixtures that include β-Gal-1,3-GalNAc glycan, for which an exceptionally low yield of UVPD (Figure S3) has been measured. For most of the mixtures the error is below 3%, however. It is exactly the synergy of spectroscopy and mass spectrometry that allows for such accurate quantification of isomeric carbohydrates in their mixtures.

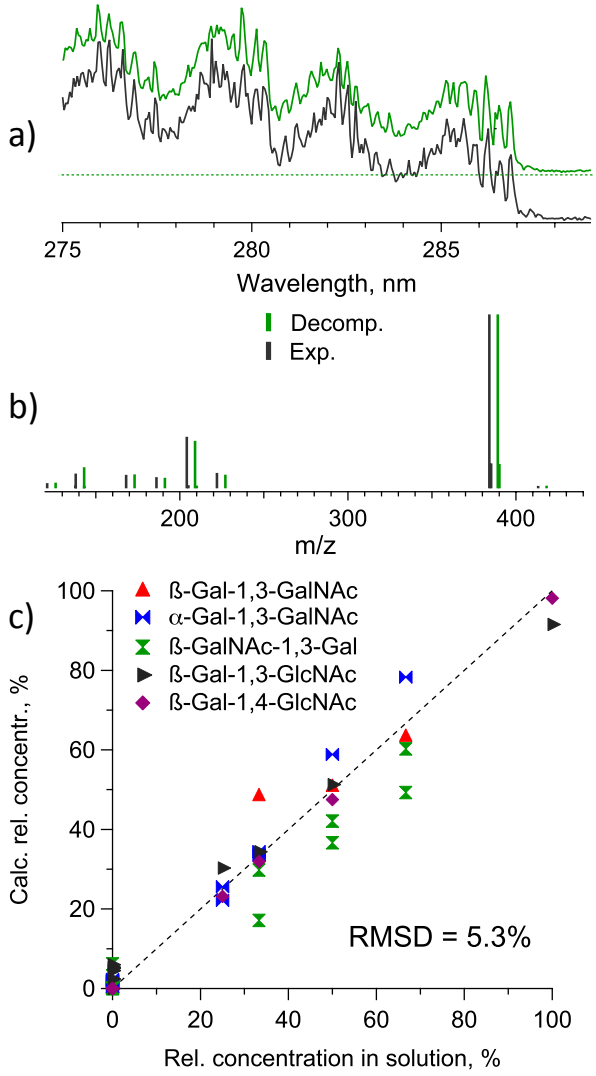

Figure 7. Comparison of the experimental (black) and decomposition (green) UVPD (**a**) optical and (**b**) mass spectra for the equimolar solution mixture of β-Gal-1,3-GlcNAc and β-Gal-1,4-GlcNAc disaccharides and of protonated Trm host aromatic. For graphical clarity the decomposition spectra are offset up and by +4 Th in a) and b), respectively. (**c**) Calculated relative concentrations for 12 mixtures of one to three isomers from the 5-component library versus their real solution concentrations (Table S1).

As we noted above (e.g. Fig. 5), UVPD of complexes of protonated aromatics with carbohydrates that have no NAc groups often yields trivial mass spectra, dominated by the aromatic and its fragments. Such spectra exhibit low isomeric specificity, which largely comes from optical spectra. This observation suggests that, perhaps, measuring a UVPD spectrum alone might be a sufficiently good but simpler alternative to the technically more demanding 2D UV-MS fingerprinting. In order to simulate this approach we compared the accuracy of 2D UV-MS decomposition with the accuracy of one-dimensional fits of optical spectra. The latter have been generated from the 2D UV-MS fingerprints of the library isomers and of their solution mixture by integrating the 2D data arrays over m/z dimension. We did such a simulation for two heptasaccharides, $Man_6GlcNAc$-I and $Man_6GlcNAc$-II, and their 2:1 solution mixture (Fig. S2). Both the 2D UV-MS and 1D UVPD-only deconvolutions earlier determined the relative concentrations as 71:29 and 68:32, respectively.[22] This example demonstrates that even in the case of large oligosaccharides, which exhibit structureless and broad UVPD optical spectra, the small differences between them, collectively, enable high accuracy in quantification of the mixed isomers.

In order to demonstrate this technical simplification, we used a quadrupole mass filter instead of the Orbitrap-based MS for detecting UVPD fragments. The filter was tuned to m/z=182 Th for maximum transmission of $TyrH^+$ in a low-resolution mode just sufficient to discriminate the protonated parent complexes. Figure 8ab compares UVPD spectra of anomeric disaccharides maltose and cellobiose in their complexes with the aromatic. Despite a lack of a resolved vibronic structure, the difference in many reproducible details makes the spectra suitable for identification and accurate quantification of this pair of isomers. Figure 8c illustrates the quality of the decomposition by comparing the spectra calculated and measured for the 1:1 solution mixtures of the two anomers. The fit coefficients suggest 47:53% for the relative concentration of the two glycans in this mixture. The decompositions of the spectra for 2:1 and 1:2 solution mixtures of the glycans give their relative concentrations as 64:35 and 27:73, respectively. Together, the three decompositions allow for an estimate of RMSD, which characterizes the expected accuracy in quantification of unknown sample mixtures, to 4.4%.

In comparison with 2D UV-MS fingerprinting, the evaluated above high accuracy of 1D UV deconvolution drops quickly however upon increasing the number of isomers in the library. Our simulations using 2D UV-MS fingerprints show that, for example, for a 1:1:2 mixture of three tetrasaccharides (blood group A1, A2 and A3/4; Figure S1) the 2D decomposition gives the relative concentrations of 27:24:49, while 1D UV measurements and

fit would result in a less accurate relative concentrations of 30:28:42. These data are consistent with the earlier reported simulations, that considered 2D UV-MS, 1D UV-only and 1D MS-only decompositions of 2D data arrays measured for isomeric peptides. Overall, for multicomponent libraries 2D UV-MS decompositions exhibit better stability in wide wavelength ranges, while 1D fits may give good accuracy only within specific narrow wavelength or m/z windows that have to be predetermined in advance. It is worth stressing that for the spectroscopically large systems studied herein, the high isomeric specificity of UV spectra is based on cooling ions to a cryogenic temperature, which ensures no/little population of vibrationally excited levels.

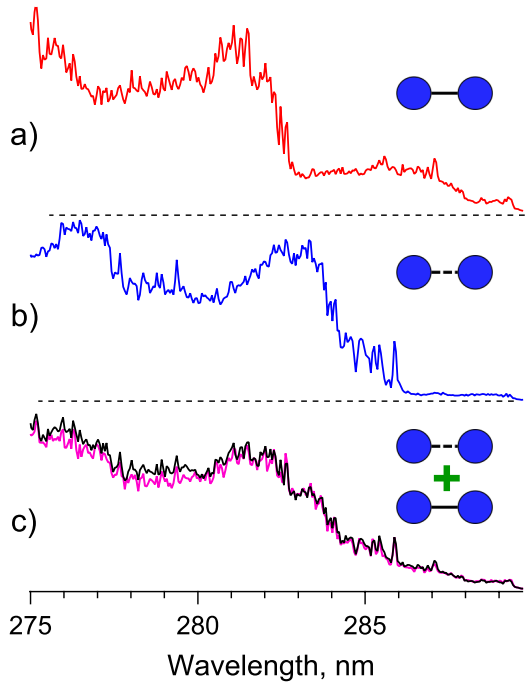

Figure 8. Measured with QMS UVPD spectra of non-covalent complexes of $TyrH^+$ with isomeric disaccharides a) α–Glc-1,4-Glc (cellobiose), b) β-Glc-1,4-Glc (maltose) and (c) with both isomers in 1:1 mixture (pink trace). The black trace in (c) is the best numerical 1D fit that adds the spectra in (a) and (b) with coefficients of 0.465 and 0.535, respectively.

▪ CONCLUSIONS

In this study, we have demonstrated how isomeric carbohydrates in any of their numerous isoforms can be distinguished and relative solution concentrations of saccharides can be quantified using our recently developed approach of 2D UV-MS fingerprinting of cold ions. Mixing carbohydrates and suitable aromatic molecules in solution efficiently produces their protonated non-covalent complexes, a good fraction of which can be brought intact to the gas phase using soft ESI. Hydrogen bonds to aromatic molecule make glycans “visible” in UV.

This enables recording the yield of photodissociation of the complexes as a 2D function of UV wavelength and of m/z of the appearing fragments. When the complexes are cooled to cryogenic temperature, such 2D fingerprints become quite specific to isomers of glycans. The sets of up to five isomers, and the carbohydrates as large as heptasaccharides, were tested. For most of the sets the expected accuracy in determination relative concentrations of isomers mixed in solution is better than 3-5%. For sets of a few isomers the 1D UVPD cold ion spectroscopy can be a sufficiently accurate while technically less demanding alternative to the 2D UV-MS sensing. Our test, in which a quadrupole mass filter replaced a broadband high-resolution Orbitrap-based MS, gives 4.4% accuracy 1D quantification of mixtures of two isomeric disaccharides. 2D UV-MS fingerprinting remains indispensible however for accurate quantifications with larger libraries and bigger glycans.

2D UV-MS fingerprinting of carbohydrates has potential for online coupling to liquid chromatography. To shorten the time of measurement to the scale of LC, the aromatic-glycan complexes can be sensed not continuously, but at a few "critical" wavelengths only, similar to what was earlier shown for peptides.[17,19,20] In such a case, yet to be demonstrated, the premixing of aromatic molecules can be performed in between an LC column and an ESI capillary.

Overall, adding carbohydrates to the sample box of 2D UV-MS fingerprinting expands its capabilities in identification of isomeric biomolecules. It is practically important that the same hardware and software can be used for very different types of biomolecules, making the method quite flexible in analytical identifications of isomers.

ASSOCIATED CONTENT

Supporting information. UVPD spectra for complexes of $TyrH^+$ with tetrasaccharides (Figure S1), UVPD spectra for complexes of $TyrH^+$ with heptasaccharides (Figure S2); UVPD spectra for complexes of $TrmH^+$ with five disaccharides (Figure S3); Tested solution mixtures of disaccharides (Table S2).

AUTHOR INFORMATION

Corresponding Author

*E-mail: oleg.boiarkin@epfl.ch

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Swiss National Science Foundation (grant 200020_172522) and Novartis AG (FreeNovation program; Dr. Hans Widmer) for financial support of this work.

REFERENCES

(1) Bode, L. Human milk oligosaccharides: every baby needs a sugar mama. *Glycobiology* **2012**, *22*, 1147-1162.
(2) Dwek, R. A. Glycobiology: Toward Understanding the Function of Sugars. *Chemical Reviews* **1996**, *96*, 683-720.
(3) Sharon, N.; Lis, H. Lectins as cell recognition molecules. *Science* **1989**, *246*, 227-234.
(4) Gray, C. J.; Migas, L. G.; Barran, P. E.; Pagel, K.; Seeberger, P. H.; Eyers, C. E.; Boons, G. J.; Pohl, N. L. B.; Compagnon, I.; Widmalm, G.; Flitsch, S. L. Advancing Solutions to the Carbohydrate Sequencing Challenge. *J Am. Chem. Soc.* **2019**, *141*, 14463-14479.
(5) Wormald, M. R.; Petrescu, A. J.; Pao, Y.-L.; Glithero, A.; Elliott, T.; Dwek, R. A. Conformational Studies of Oligosaccharides and Glycopeptides: Complementarity of NMR, X-ray Crystallography, and Molecular Modelling. *Chemical Reviews* **2002**, *102*, 371-386.
(6) Adrian, M.; Dubochet, J.; Lepault, J.; McDowall, A. W. Cryo-electron microscopy of viruses. *Nature* **1984**, *308*, 32.
(7) Wu, X.; Delbianco, M.; Anggara, K.; Michnowicz, T.; Pardo-Vargas, A.; Bharate, P.; Sen, S.; Pristl, M.; Rauschenbach, S.; Schlickum, U.; Abb, S.; Seeberger, P. H.; Kern, K. Imaging single glycans. *Nature* **2020**, *582*, 375-378.
(8) Han, L.; Costello, C. E. Mass spectrometry of glycans. *Biochemistry (Mosc)* **2013**, *78*, 710-720.
(9) Schenk, J.; Nagy, G.; Pohl, N. L. B.; Leghissa, A.; Smuts, J.; Schug, K. A. Identification and deconvolution of carbohydrates with gas chromatography-vacuum ultraviolet spectroscopy. *Journal of Chromatography A* **2017**, *1513*, 210-221.
(10) Tang, Y.; Wei, J.; Costello, C. E.; Lin, C. Characterization of Isomeric Glycans by Reversed Phase Liquid Chromatography-Electronic Excitation Dissociation Tandem Mass Spectrometry. *J. Am. Soc. Mass. Spectr.* **2018**, *29*, 1295-1307.
(11) May, J. C.; McLean, J. A. Ion mobility-mass spectrometry: time-dispersive instrumentation. *Anal. Chem.* **2015**, *87*, 1422-1436.
(12) Hofmann, J.; Hahm, H. S.; Seeberger, P. H.; Pagel, K. Identification of carbohydrate anomers using ion mobility-mass spectrometry. *Nature* **2015**, *526*, 241-244.
(13) Kirk, A. T.; Bohnhorst, A.; Raddatz, C. R.; Allers, M.; Zimmermann, S. Ultra-high-resolution ion mobility spectrometry-current instrumentation, limitations, and future developments. *Anal. Bioanal. Chem.* **2019**, *411*, 6229-6246.
(14) 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.
(15) 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.

(16) Tarasova, I. A.; Perlova, T. Y.; Pridatchenko, M. L.; Goloborod'ko, A. A.; Levitsky, L. I.; Evreinov, V. V.; Guryca, V.; Masselon, C. D.; Gorshkov, A. V.; Gorshkov, M. V. Inversion of chromatographic elution orders of peptides and its importance for proteomics. *J Anal Chem+* **2012**, *67*, 1014-1025.
(17) 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*, 4607-4611.
(18) 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.
(19) 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.
(20) 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**, *143*, 833-836.
(21) USA, US 10,283,336 B2, 2019.
(22) Saparbaev, 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**, *58*, 7346-7350.
(23) 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.
(24) Solovyeva, E. M.; Kopysov, V. N.; Pereverzev, A. Y.; Lobas, A. A.; Moshkovskii, S. A.; Gorshkov, M. V.; Boyarkin, O. V. Method for Identification of Threonine Isoforms in Peptides by Ultraviolet Photofragmentation of Cold Ions. *Anal Chem* **2019**, *91*, 6709-6715.
(25) 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.
(26) Cocinero, E. J.; Carcabal, P.; Vaden, T. D.; Simons, J. P.; Davis, B. G. Sensing the anomeric effect in a solvent-free environment. *Nature* **2011**, *469*, 76-79.
(27) Saparbaev, E.; Aladinskaia, V.; Yamaletdinov, R.; Pereverzev, A. Y.; Boyarkin, O. V. Revealing Single-Bond Anomeric Selectivity in Carbohydrate–Protein Interactions. *The Journal of Physical Chemistry Letters* **2020**, *11*, 3327-3331.
(28) Voss, J. M.; Kregel, S. J.; Fischer, K. C.; Garand, E. IR-IR Conformation Specific Spectroscopy of Na(+)(Glucose) Adducts. *J. Am. Soc. Mass Spectrom.* **2018**, *29*, 42-50.
(29) 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.
(30) Agmon, N. Elementary Steps in Excited-State Proton Transfer. *J. Phys. Chem A* **2005**, *109*, 13-35.
(31) Schindler, B.; Legentil, L.; Allouche, A. R.; Ferrieres, V.; Compagnon, I. Spectroscopic diagnostic for the ring-size of carbohydrates in the gas phase: furanose and pyranose forms of GalNAc. *Phys Chem Chem Phys* **2019**, *21*, 12460-12467.
(32) Boyarkin, O. V. Cold ion spectroscopy for structural identifications of biomolecules. *International Reviews in Physical Chemistry* **2018**, *37*, 559-606.
(33) Boyarkin, O. V.; Kopysov, V. Cryogenically cooled octupole ion trap for spectroscopy of biomolecular ions. *Rev. Sci. Instrum.* **2014**, *85*, 033105.

(34) Ma, C.; Sun, Z.; Chen, C.; Zhang, L.; Zhu, S. Simultaneous separation and determination of fructose, sorbitol, glucose and sucrose in fruits by HPLC-ELSD. *Food chemistry* **2014**, *145*, 784-788.
(35) Lowary, T. L. Twenty Years of Mycobacterial Glycans: Furanosides and Beyond. *Accounts Chem Res* **2016**, *49*, 1379-1388.
(36) Hong, A.; Choi, C. M.; Eun, H. J.; Jeong, C.; Heo, J.; Kim, N. J. Conformation-Specific Circular Dichroism Spectroscopy of Cold, Isolated Chiral Molecules. *Angew. Chem. Int. Ed.* **2014**, *53*, 7805-7808.
(37) Daly, S.; Rosu, F.; Gabelica, V. Mass-resolved electronic circular dichroism ion spectroscopy. *Science* **2020**, *368*, 1465-1468.
(38) Dwivedi, P.; Wu, C.; Matz, L. M.; Clowers, B. H.; Siems, W. F.; Hill, H. H. Gas-Phase Chiral Separations by Ion Mobility Spectrometry. *Anal. Chem.* **2006**, *78*, 8200-8206.
(39) Neelamegham, S.; Aoki-Kinoshita, K.; Bolton, E.; Frank, M.; Lisacek, F.; Lütteke, T.; O'Boyle, N.; Packer, N. H.; Stanley, P.; Toukach, P.; Varki, A.; Woods, R. J.; The, S. D. G. Updates to the Symbol Nomenclature for Glycans guidelines. *Glycobiology* **2019**, *29*, 620-624.
(40) Schindler, B.; Laloy-Borgna, G.; Barnes, L.; Allouche, A. R.; Bouju, E.; Dugas, V.; Demesmay, C.; Compagnon, I. Online Separation and Identification of Isomers Using Infrared Multiple Photon Dissociation Ion Spectroscopy Coupled to Liquid Chromatography: Application to the Analysis of Disaccharides Regio-Isomers and Monosaccharide Anomers. *Anal Chem* **2018**, *90*, 11741-11745.
(41) Ben Faleh, A.; Warnke, S.; Rizzo, T. R. Combining Ultrahigh-Resolution Ion-Mobility Spectrometry with Cryogenic Infrared Spectroscopy for the Analysis of Glycan Mixtures. *Anal. Chem.* **2019**, *91*, 4876-4882.
(42) Mucha, E.; Stuckmann, A.; Marianski, M.; Struwe, W. B.; Meijer, G.; Pagel, K. In-depth structural analysis of glycans in the gas phase. *Chem Sci* **2019**, *10*, 1272-1284.