

UCSF

UC San Francisco Previously Published Works

Title

Fourier Transform-Ion Cyclotron Resonance Mass Spectrometry as a Platform for Characterizing Multimeric Membrane Protein Complexes

Permalink

<https://escholarship.org/uc/item/1mf09795>

Journal

Journal of The American Society for Mass Spectrometry, 29(1)

ISSN

1044-0305

Authors

Lippens, Jennifer L
Nshanian, Michael
Spahr, Chris
[et al.](#)

Publication Date

2018

DOI

10.1007/s13361-017-1799-4

Peer reviewed

RESEARCH ARTICLE

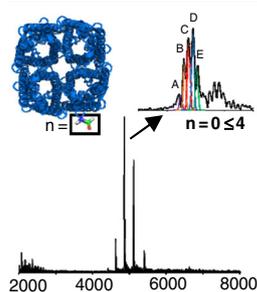
Fourier Transform-Ion Cyclotron Resonance Mass Spectrometry as a Platform for Characterizing Multimeric Membrane Protein Complexes

Jennifer L. Lippens,¹ Michael Nshanian,² Chris Spahr,¹ Pascal F. Egea,³ Joseph A. Loo,^{2,3} Iain D. G. Campuzano¹

¹Discovery Attribute Sciences, Amgen, Thousand Oaks, CA 91320, USA

²Department of Chemistry and Biochemistry, University of California-Los Angeles, Los Angeles, CA 90095, USA

³Department of Biological Chemistry and Molecular Biology Institute, University of California-Los Angeles, Los Angeles, CA 90095, USA



Abstract. Membrane protein characterization is consistently hampered by challenges with expression, purification, and solubilization. Among several biophysical techniques employed for their characterization, native-mass spectrometry (MS) has emerged as a powerful tool for the analysis of membrane proteins and complexes. Here, two MS platforms, the FT-ICR and Q-ToF, have been explored to analyze the homotetrameric water channel protein, AquaporinZ (AqpZ), under non-denaturing conditions. This 97 kDa membrane protein complex can be readily liberated from the octylglucoside (OG) detergent micelle under a range of instrument conditions on both MS platforms. Increasing the applied collision energy of the FT-ICR collision cell yielded varying degrees of tetramer (97 kDa) liberation from the OG micelles, as well as dissociation into the trimeric (72 kDa) and monomeric (24 kDa) substituents. Tandem-MS on the Q-ToF yielded higher intensity tetramer signal and, depending on the m/z region selected, the observed monomer signal varied in intensity. Precursor ion selection of an m/z range above the expected protein signal distribution, followed by mild collisional activation, is able to efficiently liberate AqpZ with a high S/N ratio. The tetrameric charge state distribution obtained on both instruments demonstrated superpositioning of multiple proteoforms due to varying degrees of N-terminal formylation.

Keywords: Native-mass spectrometry, Membrane protein, Fourier transform ion cyclotron resonance

Received: 20 June 2017/Revised: 25 August 2017/Accepted: 26 August 2017

Introduction

Membrane proteins are responsible for many cellular processes, such as ion transport and molecular recognition [1, 2]. Currently, membrane proteins constitute at least 50% of prospective therapeutic targets [3, 4], making their structural characterization a high priority. Their structure and binding properties are vital pieces of required information for comprehending their capacity as potential therapeutic targets. Yet, in spite of high interest, membrane proteins represent less

than 10% of the total solved structures in the Protein Data Bank [4–7], and only 707 unique membrane protein structures are named in the Structural Biology Knowledgebase (<http://blanco.biomol.uci.edu/mpstruc>) [8]. The dearth of solved structures and complete characterization of these proteins comes as a direct consequence of their demanding nature when it comes to expression, purification, and solubilization [9], which make them particularly challenging to investigate by traditional structural elucidation techniques such as nuclear magnetic resonance (NMR) and X-ray crystallography [10–14]. Over the past decade, native-mass spectrometry (native-MS) [9, 15] has proven to be a highly enabling and complementary biophysical technique for membrane protein characterization.

Native-MS provides a unique platform for rapid characterization with the benefit of low sample consumption, a particularly important attribute considering the poor expression of

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13361-017-1799-4>) contains supplementary material, which is available to authorized users.

Correspondence to: Joseph Loo; e-mail: jloo@chem.ucla.edu, Iain Campuzano; e-mail: iainc@amgen.com

Published online: 02 October 2017

these proteins [9]. While monomeric membrane proteins, such as bacteriorhodopsin and bovine rhodopsin, have been investigated by native-MS [16], a vital turning point in the utilization of native-MS for membrane protein characterization came in 2008 with the observation of the intact heterotetrameric membrane and soluble protein complex, BtuC₂D₂ [17]. This analysis has been subsequently followed by numerous successful native-MS investigations of various membrane protein complexes, such as *Thermus thermophilus* V-type ATPase [18] and the trimeric OmpF porin [15, 19]. However, this native-MS strategy still requires the membrane protein to be solubilized, typically in a detergent micelle, which can in fact continue to encapsulate the protein even after introduction of the protein to the gas phase. Therefore, mild collision induced dissociation (CID) is frequently employed to promote protein liberation from the detergent micelle and subsequent molecular weight (MW) determination [20]. Other activation/dissociation methods, such as infrared multiphoton dissociation (IRMPD), have also been demonstrated to efficiently liberate a membrane protein from a detergent micelle [21, 22]. Activation of the protein-detergent assembly through the application of collision energy (50 to 70 V) typically results in improved signal-to-noise (S/N) and high protein-to-detergent (P/D; Eq. 1) signal ratios.

$$P/D = \frac{\text{Intensity of Protein Signal}}{\text{Intensity of Detergent Signal}} \quad (1)$$

Yet, with the application of collisional activation, it is critical to strike a balance between release of the membrane protein from the detergent micelle and dissociation of the intact membrane protein [20, 22, 23]. This is of particular importance for the analysis of membrane protein complexes, i.e., protein-protein and protein-ligand complexes, as preservation of these fragile interactions in the gas phase while simultaneously disrupting the detergent micelle can be challenging. Therefore, in many cases, complete removal of detergent micelle signal from the spectrum may not be achieved. Besides detergents, there are other solubilization methods currently under investigation for their ability to provide high quality MS data [24–26]. Some of these methods include nanodiscs (a phospholipid bilayer contained within two membrane scaffold proteins) [27, 28], styrene-maleic acid copolymers (SMA; polymers that self-insert into the membrane and extract regions of intact membrane containing membrane proteins and associated lipids) [29], and amphipols (amphipathic copolymers that bind to the membrane protein) [25, 26]. Nevertheless, detergents remain the most common membrane mimetic by which membrane proteins are solubilized and analyzed by native-MS.

To date, reported native-MS membrane protein research has been heavily dominated by the use of quadrupole time-of-flight (Q-ToF) MS analyzers. Recently, however, Orbital trapping instruments (Orbitrap) have become increasingly employed and are an attractive alternative to Q-ToF devices due to their apparent high resolving power and improvement in spectral

quality without the need for extensive post-data acquisition processing [22, 23, 30]. However, it should be noted that real time enhanced Fourier transformation (eFT) data processing occurs on these platforms [31, 32]. Given the variety of available MS instrumentation, it is important that the distinct advantages of individual MS platforms be highlighted and leveraged to provide comprehensive native-MS characterization. Here we present the characterization of a multimeric membrane protein complex, *Escherichia coli* AquaporinZ (AqpZ), a homotetramer (97 kDa) and highly efficient water channel protein, using a Fourier transform-ion cyclotron resonance (FT-ICR) mass spectrometer. We aim to provide an unbiased comparison of the data quality between this platform and a typical Q-ToF platform regularly employed for membrane protein analysis, as well as highlight how each platform was leveraged for AqpZ characterization.

Experimental

Materials

The detergent *n*-octyl β-D-glucopyranoside (OG) used for solubilization was purchased from Anatrace (O311S; Maumee, OH, USA). Gold-coated nano-electrospray ionization (ESI) needles were purchased from Waters MS-Technologies (M956232AD1, long thin wall; Manchester, UK). Sulfur hexafluoride (SF₆) was purchased from Airgas (Palmdale, CA, USA) and used as the collision gas in both the collision cell (FT-ICR) and Trap (Synapt G2). BioRad P6 spin columns used for buffer exchange were purchased from BioRad (732-6221; Hercules, CA, USA). Trypsin was a product of Sigma-Aldrich (T-6567; St. Louis, MO, USA). Chymotrypsin was purchased from Thermo Scientific (90056; Rockford, IL, USA). Lys-C was a product from Wako (129-02541; Osaka, Japan). The preparation and purification of the AqpZ protein complex are detailed in the [Supporting Information](#).

Native Mass Spectrometry Instrumentation and Sample Preparation

A 15-Tesla solariX FT-ICR mass spectrometer (Bruker Daltonics; Bremen, Germany) and a Synapt G2 quadrupole mobility MS instrument with a travelling wave ion guide (TWIG) (Waters MS-Technologies; Manchester, UK) were used to generate native-MS spectra. Specific acquisition collision voltages, source Skimmer 1 and collision cell voltages (FT-ICR) and sampling cone and Trap voltages (Q-ToF), are explicitly discussed with the results obtained from those settings on each instrument. For other FT-ICR and Q-ToF instrument settings please refer to the [Supporting Information](#). For all AqpZ experiments, SF₆ was used as the collision gas in the collision cell (FT-ICR) and trap (Q-ToF). In general, 60 μM AqpZ stock was buffer exchanged after SEC purification ([Supporting Information](#)) using the BioRad P6 spin columns into 200 mM ammonium acetate containing 1.1% (w/v) OG (two times the critical micelle concentration; 2× CMC). This stock solution

was diluted to 15–30 μM for native-MS analysis, depending on the instrumentation used. Following buffer exchange, the AqpZ solution was loaded into a gold-coated borosilicate needle and analyzed in positive ion mode by nanoflow-ESI.

Denatured Protein Mass Spectrometry

LC/MS was performed with an Agilent 1200 series HPLC, using a ZORBAX 300SB-C3 (2.1×50 mm, $1.8 \mu\text{M}$) column (857750-909; Agilent Technologies, Santa Clara, CA, USA) in line with a Synapt G2 mass spectrometer (Waters MS-Technologies; Manchester, UK) during the gradient. This gradient employed 0.1% formic acid in H_2O as mobile phase A and 0.1% formic acid in acetonitrile as mobile phase B. The gradient was maintained at 5% B from 0–5 min, then increased to 95% B and held for 5–9 min. After this period, the gradient was reduced to 5% B for the remaining 6 min. The flow rate for the gradient was 0.250 mL/min and the column temperature was 40 °C. Zero-charge MW values for the AqpZ species were determined using MaxEnt 1 [33] in the MassLynx software (Waters MS-Technologies, UK). For deconvolution, the output mass range was set to 10,000:40,000 Da (chosen based on the expected mass of the monomeric species); 0.20 Da/channel resolution; minimum intensity ratio left and right at 50%; width at half height for uniform Gaussian model at 0.750 Da with iterations to convergence.

LC-MS/MS of Proteolytically Digested AqpZ

Trypsin and chymotrypsin digestion conditions were as follows: 15 μg of AqpZ was combined with digestion buffer and urea to final concentrations of 37.5 mM Tris (pH 7.5), 2 M urea and 10 mM hydroxylamine. Lys-C digest conditions were as follows: 15 μg of AqpZ in final concentrations of 75 mM Tris (pH 7.5), 4 M urea, and 20 mM hydroxylamine. Each digestion was brought to a total volume of 30 μL with deionized water after the addition of 1 μL of the respective enzyme to maintain an enzyme-to-substrate ratio (w/w) of 1:15. All digestions were incubated overnight at 37 °C, after which 30 μL of 0.1% formic acid was added, followed by 25 mM tris (2-carboxyethyl) phosphine (TCEP) for an incubation period of 15 min at 37 °C prior to LC-MS/MS analysis.

The proteolytic digests were measured by LC-MS/MS using a Waters NanoAcquity LC and LC-MS/MS interfaced to a Thermo Orbitrap Velos MS (Thermo Scientific; Waltham, MA, USA). The HPLC utilized a Symmetry C18 $180 \mu\text{m} \times 20$ mm trapping column (part# 186006527; Waters Corporation) in-line and preceding a ZORBAX 300SB-C18 $5 \mu\text{m}, 250 \times 0.5$ mm analytical column (part# 5064-8266; Agilent Technologies). For peptide separations, mobile phases A and B consisted of 0.1% formic acid in H_2O and 0.1% formic acid in acetonitrile, respectively. Digests were loaded for 5 min onto the trap column using 3% B at a flow rate of 8 $\mu\text{L}/\text{min}$. Peptides were separated at 12 $\mu\text{L}/\text{min}$ on the analytical column using the following gradient: 0 min at 3% B, up to 45% B over 85 min, up to 97% B over 1 min, isocratic at 97% B for 6 min, down to

3% B over 3 min, and then isocratic at 3% B for 20 min. AqpZ digested with trypsin, chymotrypsin, or Lys-C was analyzed by LC-MS/MS using a full MS scan of m/z 300–2000 at 30K resolution (FWHM), followed by low resolution CID scans of the top six most abundant precursors. A spray voltage of 4.5 kV, a S-Lens rf level of 50, an isolation width of 2.0 Da, a collision energy (CE) of 35 V, and a 15 s dynamic exclusion were used.

Targeted CID and higher-energy collisional dissociation (HCD) experiments were performed on the singly charged ions (unmodified: 453.2224 Da and modified: 480.2182 Da) of the peptide MFR (amino acids 1–3 of the full length monomer) obtained from the tryptic digest. CID and HCD mass spectra were collected at 7.5K resolution using CE values of 35 V and 29 V, respectively. Other instrumental parameters were kept consistent, as above (spray voltage, isolation width, etc.).

Results and Discussion

Regions of Collisional Activation for Detergent Micelle Disruption on the FT-ICR and Q-ToF Platforms

Historically, the process of membrane protein liberation from detergent micelles by native-MS is achieved by increasing the overall activation energy of the MS system (typically sample cone and collision cell) until the micelle is disrupted, and the protein is liberated [20]. This technique has been vastly employed for the analysis of many membrane proteins [20, 34, 35]. Each unique MS platform has various regions where voltage settings can be manipulated to provide the necessary energy to achieve protein liberation. The source region of the FT-ICR houses a 300 mm heated glass capillary that provides a long desolvation region, which introduces the protein ions to the high vacuum regions of the instrument from the ambient atmosphere. Following this capillary, there are two funnel stacks, each followed by a skimmer, the voltage of which can be manipulated to aid in ion transmission as well as to generate various levels of in-source dissociation. One of the two FT-ICR voltages investigated in this study, Skimmer 1 (Skim1), is located at the end of the funnel 1 stack (Figure 1a) and creates a voltage gradient such that if the voltage is high the ions will be accelerated into the funnel 2 stack, potentially producing in-source dissociation. The second voltage manipulated on this platform to achieve protein ejection was applied in the collision cell, which is located after the quadrupole and prior to the magnet. Application of collision energy and resultant ion-neutral collisions disrupt the detergent micelle to allow for protein ejection so that the membrane protein complex is ultimately detected. Increased application of this voltage can also generate dissociation of the protein into its constituent subunits (monomer and trimer).

Alternatively, the nanoflow-ESI source of the Synapt G2 quadrupole-ion mobility MS instrument does not utilize a heated capillary for ion transfer from ambient conditions to the vacuum regions of the instrument, but instead utilizes dual

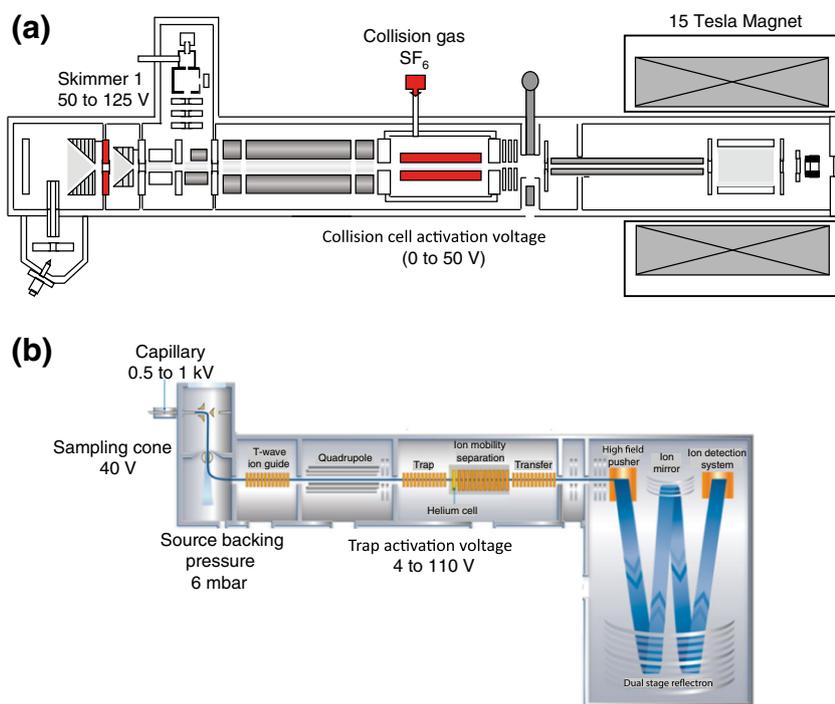


Figure 1. Instrument schematics of both **(a)** the FT-ICR, highlighting the Skimmer 1 and collision cell regions, and **(b)** the Synapt G2 quadrupole ion mobility ToF, highlighting the sampling cone and TrapCE voltage regions. Voltage values shown reflect those used in the experiments detailed in this manuscript

orthogonal ion sampling (Z-spray) that can be held from 2-6 mbar. To this point, the backing pressure of the FT-ICR platform cannot be altered from the manufacturer setting. For membrane protein analysis, the backing pressure on the Q-ToF is typically set to 6.0 mbar [20], which was the backing pressure maintained for the experiments described herein. Although lower pressures can be used, the resulting spectra will consist largely of membrane protein-detergent complex signals, and thus very heterogenous [20]. Several groups have developed modifications to the source region of the Synapt G1 and G2 instruments to allow for source heating more similar to the FT-ICR [36, 37]. However, given the isolation of these modifications to specific laboratories, we opted to evaluate the data quality and conditions necessary for optimal protein liberation using the manufacturer provided instrument sources.

Given the two fundamentally different designs of these source regions, a direct comparison of these two voltage settings is challenging. The sampling cone voltage, analogous to Skim1 (FT-ICR), can be manipulated to aid in ion transmission as well as to generate various levels of in-source dissociation. However, this voltage is in much closer proximity to the nano-ESI emitter than that of Skim1 (FT-ICR). The other voltage setting manipulated on this instrument is the collision energy in the Trap (TrapCE), which is the offset between the exit of the quadrupole and the entrance of the collision cell (Q-ToF TWIG cell; FT-ICR hexapole). The ability to manipulate the energy imparted to these ions at various locations in each instrument allows for thorough investigation of which voltage parameters

generate optimal liberation of the membrane proteins from the detergent micelles. Additionally, it allows for determination of which conditions generate intense dissociation of the liberated membrane proteins – leading to subunit stoichiometric information. Exploration of optimal voltage conditions is dependent upon the given pressure settings (backing and Trap pressure) of the instrument and could potentially vary upon changes in these values.

FT-ICR Native-MS Analysis of the AqpZ Tetramer

Native-MS spectra of detergent solubilized membrane proteins in the absence of collisional activation are typically populated by detergent micelle distribution signal, highly polydisperse in nature (Figures S1–S4, Supporting Information) [38]. As previously mentioned, ion activation is used to increase the protein-to-detergent (P/D; Equation 1) ratio. Multiple Skim1 voltages were explored (50, 75, 100, and 125 V) which, in the absence of an applied voltage to the collision cell, yielded a polydisperse detergent micelle distribution between m/z 1000 and 3000 at all voltages (Figures S1–S4, Supporting Information) and produced minimal tetramer protein complex ejection solely at 125 V (Figure S1, Supporting Information). At this voltage, tetramer signal was observed at 5.4% relative intensity to the base peak (Figure S1, Supporting Information). It was instead the combination of the Skim1 and collision cell voltages that generated consistent AqpZ ejection. Increasing the applied CE in the collision cell above 20 V produced varying degrees of AqpZ tetramer ejection from the OG detergent

micelle. Higher Skim1 voltages or increase of the CE voltage, (Skim1 75 V with CE 30 V, Skim1 100 V with CE 30–50 V, and Skim1 125 V with CE 20–50 V; Figures S1–S4, Supporting Information) generated dissociation of the tetramer into lower MW multimeric species (trimer and monomer; Figures S1 and S2, Supporting Information). The trimeric species, when observed, displayed a maximum of five charge states ($z = 8+$ to $12+$), which at higher CE voltages were more abundant than the tetramer signal (m/z 4200–6000). Further increase of the collision cell activation, i.e., up to 50 V, caused further loss of tetramer signal by way of increased dissociation and ultimately led to an inability to observe any signal corresponding to intact AqpZ. With the exception of several Skim1 and CE voltage combinations, the detergent micelle signal remained relatively intense compared with the protein signal (low P/D) throughout the analysis.

In general, a relatively wide range of Skim1 and collision cell voltages produced some level of intact tetrameric AqpZ signal (Skim1 50 V with CE 30 and 50 V and Skim1 75, 100, and 125 V with CE ranging from 10 to 40 V; Figures S1–S4, Supporting Information). However, many of these combinations also generated dissociation of the tetramer, while very few settings produced solely intact tetramer (Skim1 50 V with CE 30–40 V, Skim1 75 V with CE 20 and 40 V, Skim1 100 V with CE 10–20 V, and Skim1 125 V with CE 10 V; Figures S1–S3, Supporting Information). Given the observation of subunit ejection under a variety of source Skim1 and collision cell voltage settings, it is clearly a balance of the applied voltages between these two regions that result in efficient AqpZ liberation, highly consistent with previously described Q-ToF data [34].

Optimal liberation of AqpZ was obtained with the combination of Skim1 50 V and CE activation at 30 V, which were considered optimal conditions due to both the AqpZ tetramer signal intensity and the absence of any detectable monomer/subunit ejection. Under these conditions, four major charge states ($z = 18+$ to $21+$) were detected, which correlated with the unmodified tetramer (4 M1, *vide infra*) average theoretical MW of 97,074.24 Da (calculated from elemental composition) and a mass accuracy of +89 ppm (Figure 2a and Figure S4). These mild activation conditions provided a spectrum with a S/N of 111, a P/D of 1.6, and an average charge state of $z = 19.6+$. Previous native-MS investigation in alternative detergents, such as C₈E₄ and *n*-Nonyl- β -D-glucopyranoside (NG), yielded lower average charge states, as expected for these types of detergents [34, 39]. However, the AqpZ preparation and purification described herein were performed in OG (Supporting Information); in order to decrease sample handling prior to MS, OG was also employed during the native-MS analysis.

Examination of the individual tetrameric charge states under optimal protein liberation conditions (Skim1 50 V and CE 30 V) revealed clear superpositioning of five distinct possible proteoforms [40] (Figure 2b). Mass deconvolution of these different proteoforms provided the MWs listed in Figure 2b. Using the m/z values of the different proteoforms, tetrameric masses that maintained an average of $\Delta 32.23 \pm 0.04$ Da

between each proteoform were obtained. Mass error compared with the theoretical mass of the intact tetramer (with incorporation of the modified monomer) could not be accurately determined for the last four proteoforms (proteoforms B–E in Figure 2) because of the contribution of adjacent proteoforms to the m/z ratios of the individual proteoforms. However, this error is expected to contribute to the slight shift in the predicted distribution compared with the experimental observations (Figure 2b, inset). At higher CE voltages where lesser multimeric species were observed, their inspection also revealed complementary superpositioning (i.e., trimer consisting of four potential combinations of monomeric species), with the monomeric species showing two distinct signals for each discrete charge state. Mass deconvolution (by MaxEnt 1; [33]) of these two species provided average MWs of 24,270.65 and 24,297.04 Da, designated as M1 and M2, respectively, which provided an overall mass difference of $\Delta 26.39$ Da. These MWs correlate well to those obtained from the denaturing LC/MS experiments (*vide infra*). The less intense signal to the right of the main AqpZ signal, between m/z 4860 and 4870, differs from the expected AqpZ tetramer MW by an average of $\Delta 31.40$ Da. Two oxidation events could potentially account for this signal.

Q-ToF Native-MS Analysis of the AqpZ Tetramer

Q-ToF platforms are the most widely utilized analyzers for native-MS of membrane proteins solubilized under a variety of conditions [20, 26, 34], hence AqpZ was also analyzed on an ion mobility Q-ToF system. For these experiments, a single sampling cone voltage of 40 V was employed and the TWIG Trap activation voltage was varied, ranging from 10 to 110 V. This produced various levels of protein liberation and subunit ejection from the OG-detergent micelles from 50 to 110 V (Figure S5, Supporting Information). While some previous studies [20, 34] have utilized higher sampling cone voltages (typically 200 V), here variation of this voltage, in both the presence or absence of Trap CE activation, did not provide the same variation of AqpZ tetramer dissociation as observed with variation of the source Skim1 voltage on the FT-ICR. Instead, its variation solely affected the observation of detergent micelle signal and the stability of the signal observed. Therefore, a low sampling cone voltage was maintained through the experiments to aid in maintaining mild activation conditions optimal for observation of intact tetramer. Consequently, TrapCE variation was the sole voltage parameter investigated for its effect on the observation of AqpZ tetramer signal.

Optimal AqpZ liberation was observed using a TWIG Trap activation voltage of 70 V, as this condition provided distinguishable tetramer signal with minimal dissociation into trimeric or monomeric substituents. However, even under optimal liberation conditions, the spectrum also maintained an intense detergent micelle distribution in the lower m/z range that overpowered the minimal protein tetramer signal (P/D of 0.04) (Figure 3a). Under these conditions, five major charge

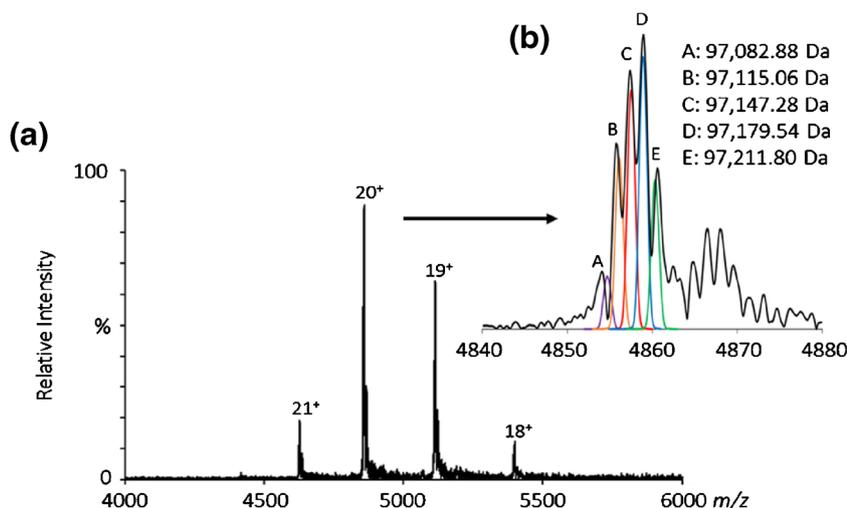


Figure 2. AqpZ native mass spectra acquired on the FT-ICR; **(a)** mass spectrum collected at Skim 1 50 V and collision cell at 30 V, m/z 4000–6000 showing the tetrameric complex and **(b)** Zoom in of the $z = 20+$ charge state displaying the superpositioning of proteoforms corresponding tetrameric MWs. These proteoforms represent N-terminal formylation (vide infra) and were observed for each tetrameric charge state. The colored distribution corresponds to the predicted observance of the various proteoforms normalized to the intensity of the tetrameric distribution observed. The proteoforms are colored according to the following monomeric species combinations, for which theoretical MWs, calculated from elemental composition, are provided for each: 4 M1: 97,074.24 Da (purple); 3 M1/1 M2: 97,102.25 Da (orange); 2 M1/2 M2: 97,130.26 Da (red); 1 M1/3 M2: 97,158.27 Da (blue) and 4 M2: 97,186.28 Da (green). MW values provided for species A-E in the inset were calculated from the individual apex m/z values. The predicted proteoform distribution m/z values as compared to the apices of the experimental peaks differ by 0.58, 0.38, 0.16, -0.06 , and -0.27 , respectively. The peak distribution to the right of the AqpZ signal differs from the expected AqpZ tetrameric MW by an average of 31.40 Da. This mass difference could potentially be explained by two oxidation events

states ($z = 16+$ to $20+$) were detected, which correlated with the calculated tetramer MW, 97,074.24 Da (Figure 3b, inset). Increased tetramer signal from tandem-MS (Figure 3b) facilitated examination of the discrete tetrameric charge states revealing similar superpositioning to that observed in the FT-ICR data (Figure 2). However, this inspection revealed poorly resolved yet distinguishable superpositioning of five proteoforms (Figure 3b, inset), which was in agreement with the results observed on the FT-ICR. Here the ppm error of the observed proteoforms m/z could not be accurately determined as the proteoforms remained poorly resolved, thus contribution from the adjacent proteoforms is highly probable. However, in this analysis the different proteoforms correlated to tetrameric masses that maintained an average of $\Delta 43.39 \pm 21.74$ Da between the apex of each proteoform (Figure 3b, inset). Given the minimal resolution between proteoforms, it is important to note that the data presented throughout this work is not post-processed in any way (smoothed, background subtracted, etc). Attempts to smooth the data resulted in loss of resolution between these proteoforms depending on the smoothing parameters utilized.

Similar to the observations on the FT-ICR, increasing ion activation (TWIG voltage of 90–110 V) caused a shift in abundance of the tetrameric species to favor the trimeric and monomeric species, which eventually overpowered the signal for the tetrameric complex. Under these conditions, mass deconvolution by MaxEnt1 of the two monomeric species provided MWs of

24,269.48 Da and 24,297.29 Da for M1 and M2, respectively, which provided an overall mass difference of $\Delta 27.81$ Da. Yet while this deconvolution allowed for MW determination of the two monomeric species, it did not explain the physical difference between the two monomeric species, which was ultimately responsible for the superpositioning observed in the tetramer. This superpositioning was due to various combinations of the two monomeric species (vide infra).

Utilization of Tandem-MS to Increase S/N and P/D Ratios through Targeted m/z Experiments

Tandem-MS can be used for interrogation of individual charge states as well as to simplify the analysis of highly polydisperse spectra [22, 41, 42]. It has been previously reported during the investigation of bacteriorhodopsin in OG that employment of tandem-MS on both FT-ICR and Q-ToF instruments achieved higher P/D and S/N ratios for the protein [22]. This was accomplished through selection of several m/z ratios, both slightly higher and in the middle of the expected m/z range of the charge state distribution for the protein. Here, tandem-MS was employed in this fashion to determine if similar increases in S/N and P/D ratios would be observed for a larger multimeric membrane protein complex.

In full-MS mode, FT-ICR analysis generally provided low P/D ratios at the various Skim1 and CE voltages investigated than the Q-ToF. Under optimal ejection conditions, tandem-

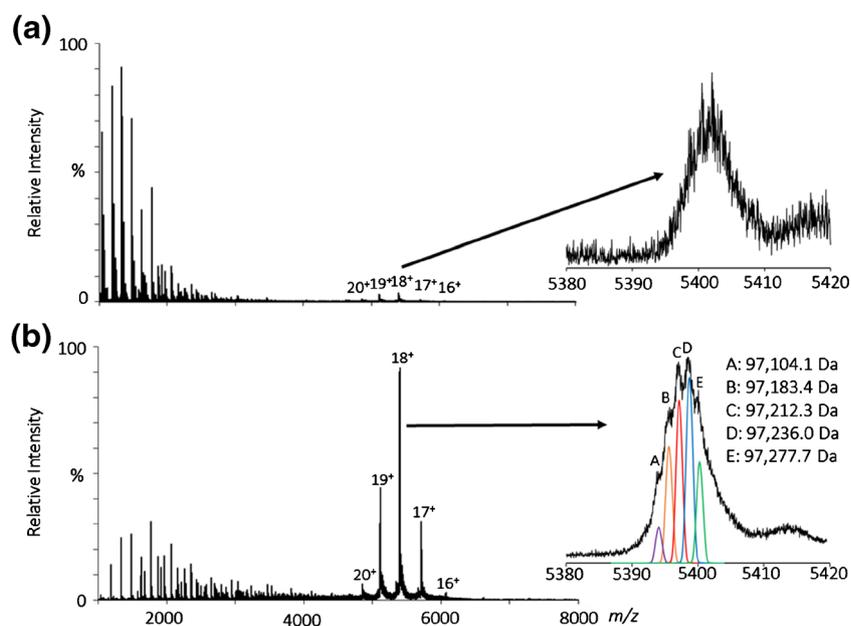


Figure 3. Native-MS Q-ToF data of AqpZ. **(a)** Full-MS data showing low P/D with intense detergent signal from m/z 1000–2500. Sampling cone and TWIG Trap voltage at 40 V and 70 V, respectively. Inset shows the lack of proteoform resolution observed for the individual charge states and highlights that in full-MS mode the tetramer is still slightly adducted; and **(b)** tandem-MS data with the quadrupole selection at m/z 6500 showing improved P/D. Acquisition conditions were otherwise kept consistent for both spectra. Inset shows the superpositioning for the $z = 18+$ charge state of different proteoforms for this tetrameric protein complex with the corresponding tetrameric MWs. The colored distribution corresponds to the predicted observance of the various proteoforms normalized to the intensity of the tetrameric distribution observed on the FT-ICR. The proteoforms are colored according to the following monomeric species combinations: 4 M1 (purple); 3 M1/1 M2 (orange); 2 M1/2 M2 (red); 1 M1/3 M2 (blue), and 4 M2 (green). The Q-ToF data presented in this figure, and throughout the manuscript, has not been smoothed or post-processed. This lack of post-processing clearly demonstrates that tandem-MS followed by collisional activation improves the proteoform resolution compared with full-MS followed by activation for this large multimeric complex

MS was investigated by selecting m/z 5900, 5500, and 5000, while maintaining the acquisition m/z range of 153–10,000 to allow for observation of potential tetramer dissociation. However, none of the precursor m/z selections investigated provided recognizable protein signal and, in fact, in all cases provided only unstable detergent signal (data not shown). A similar strategy was employed on the Q-ToF while maintaining its respective optimal ejection conditions and acquisition m/z range of 500–10,000. The quadrupole selection window on this platform is set using LM and HM resolution slider values. For these experiments, values of 4 and 10 were used for the LM and HM values, respectively, which equate to a transmission window of 24 m/z at FWHM (based on native MS/MS quad selection of BSA at m/z 4457.42). First, the m/z ratio 6500 (slightly above the tetramer charge state distribution) was selected, which increased the P/D roughly 100-fold ($P/D = 2.8$) while maintaining the same charge distribution as observed in full-MS mode (Figures 3 and 4a). Interestingly, the m/z targeted during tandem-MS had an effect not only on the observed tetramer charge state distribution, but also on the observation of AqpZ monomer signal (Figure 4). Decreasing precursor selection from m/z 6500 to m/z 5500 shifted the average tetramer charge state from $z = 18.1+$ to $z = 20.1+$, and resulted in the observation of intense monomer signal (charge states $z = 7+$ to $11+$). This observation could be

explained by the selection of higher tetrameric charge states as the m/z ratio selected is decreased. These higher charge states require less activation energy to undergo asymmetric subunit dissociation than lower charge states, thus more dissociation into its respective monomer and trimeric constituents was observed [42, 43].

Some tandem-MS studies of soluble proteins have demonstrated that the energies necessary to dissociate higher order complexes is dependent upon the charge state isolated during the tandem-MS experiment [42, 44]. For example, it was observed that if the dodecameric SP-1 charge state isolated was higher than $z = 24+$ it dissociated mainly through ejection of higher oligomers, whereas if lower than $z = 24+$, the complex dissociated mainly through ejection of monomers. Thus, depending on the charge state isolated, the complex may dissociate via several different mechanisms. Here, given the shift in average tetrameric charge state observed, we can deduce similar charge-dependent effects on dissociation. However, it is unclear why tandem-MS showed no effect in generating high S/N mass spectra for AqpZ tetramer in OG using the FT-ICR platform, whereas it is an effective strategy using the Q-ToF instrument. The roles of OG micelle size and how the FT-ICR instrument effectively removed detergent molecules from the membrane protein, as was reported for the Q-ToF system [39], will be investigated in the future.

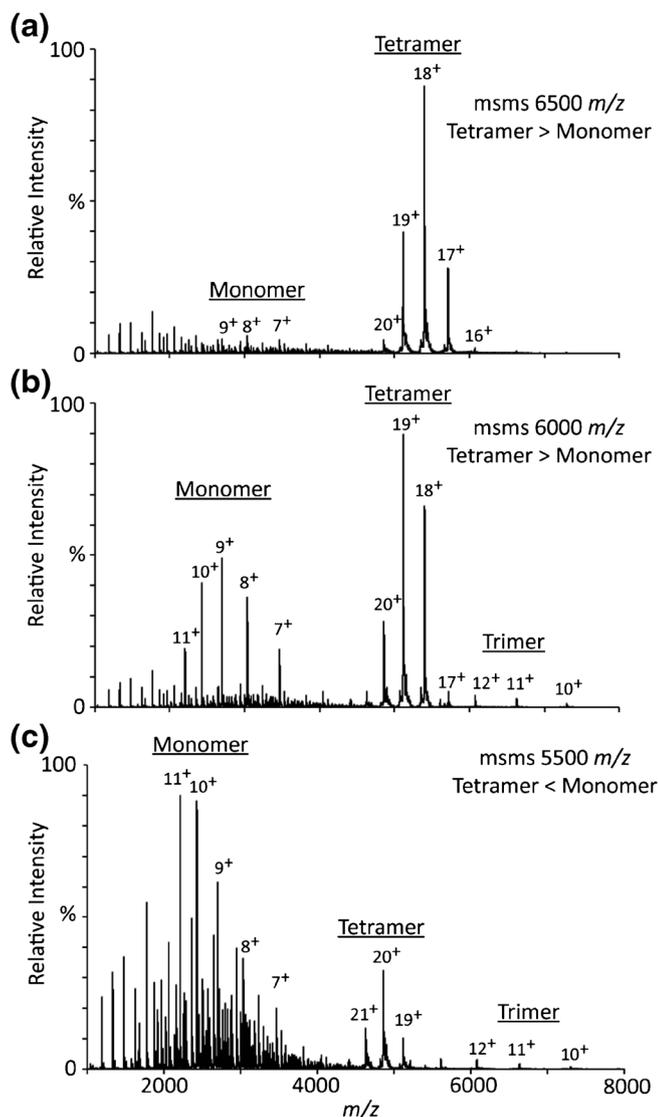


Figure 4. Use of quadrupole selection to improve the P/D ratio of AqpZ on the ion mobility Q-ToF. **(a)** Selection of m/z 6500 shows intense tetramer signal with minimal observable monomer species; **(b)** m/z 6000 maintains a closer equilibrium of tetramer and monomer species with minimal trimer species observed; and **(c)** m/z 5500 is predominantly monomer signal with low intensity tetramer signal as well as trimer signal

Denaturing LC/MS for Monomer Accurate Mass and Identification of the $\Delta 27.14$ Da Mass Addition by Proteolytic Digestion Followed by LC-MS/MS

Denaturing LC/MS was employed to obtain accurate mass measurements for the two monomeric species identified as a consequence of CID activation of the AqpZ tetramer. For this experiment, AqpZ in 200 mM ammonium acetate and 1.1% (w/v) OG was diluted to 0.5 $\mu\text{g}/\mu\text{L}$, and 0.125 μg AqpZ was injected on the C3 column and analyzed using a 15 min gradient (Experimental). The resulting spectrum (Figure S6, Supporting Information) revealed two monomeric species with MWs of 24,267.76 Da and 24,294.90 Da for M1 and M2,

respectively. The M1 species corresponded to the monomeric AqpZ sequence minus the two C-terminal amino acid residues glycine and serine. The M2 MW corresponded to the M1 sequence with a 27.14 Da mass addition. Given the various potential sources of the mass addition, such as single amino acid substitutions of serine with asparagine (+27.01 Da), lysine with arginine (+28.01 Da), or formylation (+27.99 Da), further investigation by enzymatic digestion was performed.

The solution conditions for each of the three enzymatic digestions can be found in the [Experimental](#) section. The level of sequence coverage from combining the spectral information for all three digestions equated to only 55% coverage, with the Lys-C digestion providing the least coverage (2.94%) to chymotrypsin, which predictably provided the most (54.6%) (Figure 5a). Yet while the level of sequence coverage was less than ideal, it was observed that the N-terminal peptides from both the trypsin and the chymotrypsin digests presented two very distinct peaks that differed by over 10 min in retention time (Figure S7, Supporting Information). Further investigation of the two chromatographic peaks from the tryptic digest revealed that the corresponding peptides that eluted (deconvoluted monoisotopic masses: unmodified = 452.2224 Da and modified = 480.2182 Da) differed in mass by $\Delta 27.9958$ Da, with the modified peptide maintaining the longer retention time. The fact that a 27.9958 Da mass addition was observed in the N-terminal peptides from both the tryptic and chymotryptic digests meant that the source of the mass addition could be isolated to the three N-terminal amino acid residues MFR. Given the three amino acids present in the tryptic digest N-terminal peptide and taking into account the potential modification sources of the addition, it was hypothesized that N-terminal methionine formylation (fMet) was in fact the modification present as there was only 5 ppm error between the observed mass and the theoretical mass of a formylated peptide. This hypothesis was confirmed by targeted CID of the modified and unmodified N-terminal MFR peptide from the tryptic digest (Figure 5b and c), which yielded unique b_2 , $b_3\text{-NH}_3$, and y_3 ions, contingent upon the specific peptide fragmented.

Leveraging the Advantages of the Individual FT-ICR and Q-ToF Platforms

There has long been discussion in the MS community regarding the energetic differences between various MS platforms. It would be amiss not to discuss the distinct differences of the FT-ICR and Q-ToF platforms explored here that give rise to the unique experimental conditions for optimal protein liberation from the detergent micelles. The source voltages must remain at a minimum to allow for efficient transmission of the intact protein-micelle complex with minimal dissociation. Thus, it is the downstream application of collision energy (Q-ToF TWIG cell; FT-ICR hexapole) that disrupts the protein-micelle complex to liberate the protein. The intrinsically different source designs prevent an accurate comparison of the exact conditions. However, it must be considered that the ions transit

(a) Lys-C:

MFRKLAACEFGTFWLVFGGCGSAVLAAGFPELGIGFA
 GVALAFGLTVLTMAFAVGHISGGHFNPVAVTIGWAGGR
 FPAKEVVGYYVIAQVVGIVAAALLYLIASGKTGF~~GAAA~~
 SGFASNGYGEHSPGGYSMLSALVVELVLSAGFLLVIH
 GATDKFAPAGFAP~~IAIGLALTLIHLISIPVTNTSVNPARS~~
 TAV~~AI~~FQGGWALEQLWKKWVVPVGGIIGGLIYRTLLEK
RDGTLVPR

Trypsin:

MFRKLAACEFGTFWLVFGGCGSAVLAAGFPELGIGFA
 GVALAFGLTVLTMAFAVGHISGGHFNPVAVTIGWAGGR
 FPAKEVVGYYVIAQVVGIVAAALLYLIASGKTGF~~GAAA~~
 SGFASNGYGEHSPGGYSMLSALVVELVLSAGFLLVIH
 GATDKFAPAGFAP~~IAIGLALTLIHLISIPVTNTSVNPARS~~
 TAV~~AI~~FQGGWALEQLWKKWVVPVGGIIGGLIYRTLLEK
RDGTLVPR

Chymotrypsin:

MFRKLAACEFGTFWLVFGGCGSAVLAAGFPELGIGFA
GVALAFGLTVLTMAFAVGHISGGHFNPVAVTIGWAGGR
 FPAKEVVGYYVIAQVVGIVAAALLYLIASGKTGF~~GAAA~~
SGFASNGYGEHSPGGYSMLSALVVELVLSAGFLLVIH
 GATDKFAPAGFAP~~IAIGLALTLIHLISIPVTNTSVNPARS~~
 TAV~~AI~~FQGGWALEQLWKKWVVPVGGIIGGLIYRTLLEK
RDGTLVPR

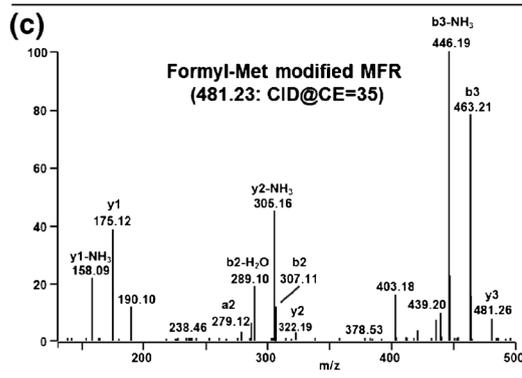
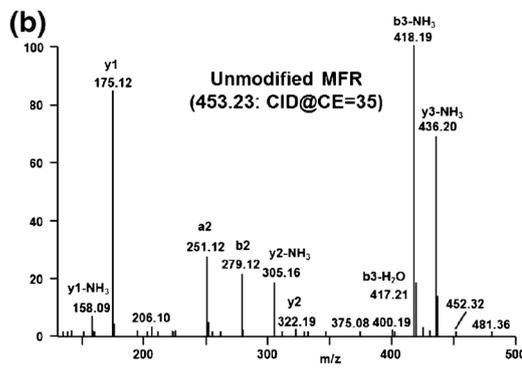


Figure 5. Characterization of the modification to the monomeric species. **(a)** Sequence coverage of AqpZ from the Lys-C, trypsin, and chymotrypsin digestions. Positive sequence identification is underlined; **(b)** targeted CID spectra of both the unmodified N-terminal peptide from the tryptic digest and; **(c)** targeted CID spectra of both the modified N-terminal peptide from the tryptic digest. Both spectra were obtained on the Orbitrap Velos MS

through a heated ion transfer capillary in the FT-ICR source. It could be hypothesized that the thermal energy imparted into the protein-micelle complex is sufficient to increase its internal energy (i.e., rotational, vibration, and translational) such that less activation energy is required for protein ejection in either the Skim1 or hexapole collision cell region. It is therefore important to note the distinct effect of increasing the Skim1 voltage on the FT-ICR, which, when the applied collision cell energy remained constant (30 V), generated spectra with varying amounts of tetramer dissociation, thus implying that some level of dissociation occurs in the source regime due to ion neutral collisions with residual gas prior to the collision cell. However, variation of the sampling cone voltage on the Q-ToF had virtually no effect on the amount of tetramer observed, regardless of the applied trapCE. These observations support the hypothesis of more efficient desolvation within the longer heated capillary of the FT-ICR, which allowed for the observation of tetramer at higher Skim1 voltages (100 and 125 V) even in the presence of minimal collision energy (10 and 20 V; Supporting Information Figures 1S and 2S).

Resolution limitations in native-MS are indeed of concern when analyzing large protein systems [45], which

could contain various levels of small modifications and/or bound ligand. We feel it is important to note that the FT-ICR enabled clear distinction of proteoform masses because there will be cases where the make-up of these multimeric membrane protein complexes could have an impact on their function and thus their overall biological significance. As well, the increase in resolution provided by the FT-ICR would enable increased confidence in the determination and assignment of small-molecule binding to membrane proteins and their complexes. As pointed out previously, this class of molecules represents a key area of interest for the pharmaceutical industry, thus making the identification of drug binding/incorporation into these proteins a key sub-area of importance. Additionally, numerous other low MW modifications may be of potential interest during the characterization of other membrane protein complexes, such as noncovalently bound lipids, bound small molecule drug candidates, and post-translational modifications [7, 23, 46, 47]. As these proteins naturally exist within a lipid membrane, the ability to observe and identify bound lipid species could provide key insights into the natural membrane environment of a protein target [2, 34].

Conclusions

This work demonstrates, to our knowledge, the first analysis of a multimeric membrane protein complex by FT-ICR. Direct comparison of the spectral quality achieved under optimal protein ejection conditions on the FT-ICR to that obtained on the Q-ToF provides strong support for the continued use of FT-ICR-MS for membrane protein analysis. While proteoform resolution was readily observed in the FT-ICR data, essentially no proteoform resolution was observed on the Q-ToF without quadrupole selection. Employment of quadrupole selection was also necessary on the Q-ToF platform to achieve comparable S/N and P/D ratios (S/N of 111, P/D of 1.6, and S/N of 85, P/D 2.8 for the FT-ICR and Q-ToF, respectively), while minimal collision energy was needed on the FT-ICR.

Resolution and m/z separation power are important characteristics to consider for any sample, but are of particular importance when analyzing complex samples with potential low MW modifications as it facilitates their observation. As well, it can provide insight as to how those modifications may affect protein subunit integration into a membrane protein complex [1]. Closer investigation of the tetramer complex signal here revealed superpositioning, which was ultimately attributed to multiple combinations of both an N-terminally formylated and a non-formylated monomer through proteolytic digests and targeted LC-MS/MS. Although not the most biologically relevant modification, the presence of fMet did indeed contribute to the complexity of the data analysis. However, the ability to better separate and more clearly assign the proteoform superpositioning here provides an optimistic outlook that with the superior m/z separation power of the FT-ICR the potential presence of low MW modifications could be determined. Given the low biological significance of the N-terminal formylation present, it is inappropriate to draw inference into the biological relevance of the various compositions based on their corresponding intensities. However, this characterization provides an important piece of information regarding the protein composition prior to any sample handling that might alter the protein structure post-cell expression and prior to native-MS investigation. Previous studies regarding N-terminal methionine excision from *E. coli* expressed proteins have shown that excision is heavily dictated by the identity of the penultimate amino acid residue [48]. In those experiments it was observed that when phenylalanine was present in the penultimate position, excision of the methionine was not observed [48], which is in accordance with the observations detailed in this analysis. It can be speculated that this is a contributing factor to why the N-terminal fMet, a common modification in *E. coli* expressed proteins [48, 49], is observed here.

To the best of our knowledge, the observation of this low MW modification on the AqpZ tetramer has not been previously reported. Given its small mass, this formylation and other similar modifications could easily be overlooked depending upon the instrumental parameters utilized and/or the post-acquisition data processing (such as smoothing) procedures employed. Careful attention to the instrumental conditions

employed on lower resolution platforms as well as data processing settings will be absolutely essential to ensure that subtle details within the protein charge state distribution are not overlooked. It should be reiterated that the data presented from both platforms was not post-processed in any way (i.e., smoothing, background subtraction, etc). Smoothing of the FT-ICR data still allowed for the proteoform superpositioning to be observed, while the ability to retain proteoform resolution in the Q-ToF data was extremely dependent on the smoothing method and parameters applied. The ability to post-process data without loss of resolution of low level modifications provides strong support for the use of the FT-ICR platform for multimeric membrane protein complex analysis, at least in combination with a Q-ToF platform when possible. The range of applicable collision energies prior to complete complex dissociation was narrower on the FT-ICR platform than the Q-ToF, which suggests that for structural studies, the Q-ToF platform would provide superior range for probing structural stability (Figures S1–S3, *Supporting Information*). However, it is clearly demonstrated that the superior resolution of the FT-ICR makes it a strong candidate for the analysis of membrane proteins that may contain small MW modifications. In any case, we predict that due to the increased S/N and P/D levels, higher level of protein desolvation, and increased resolution provided by the FT-ICR, that it will continue to gain popularity as an important platform for the characterization of multimeric membrane protein complexes.

Acknowledgments

Support from the US National Institutes of Health (R01GM103479 and S10RR028893) and the US Department of Energy (UCLA/DOE Institute for Genomics and Proteomics; DE-FC03-02ER63421) to J.A.L. is acknowledged. The postdoctoral program at Amgen, Inc. is acknowledged for its support of this project.

References

1. Almén, M.S., Nordström, K.J.V; Fredriksson, R., Schiöth, H. B.: Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin. *BioMed Central Biol.* **7**(50) (2009). <https://doi.org/10.1186/1741-7007-7-50>
2. Barrera, N.P., Zhou, M., Robinson, C.V.: The role of lipids in defining membrane protein interactions: insights from mass spectrometry. *Trends Cell Biol.* **23**(1), 1–8 (2013)
3. Hopkins, A.L., Groom, C.R.: The druggable genome. *Nat. Rev. Drug Discov.* **1**, 727–730 (2002)
4. Arinaminpathy, Y., Khurana, E., Engelman, D.M., Gerstein, M.B.: Computational analysis of membrane proteins: the largest class of drug targets. *Drug Discov. Today.* **14**(23/24), 1130–1135 (2009)
5. Carpenter, E.P., Beis, K., Cameron, A.D., Iwata, S.: Overcoming the challenges of membrane protein crystallography. *Curr. Opin. Struct. Biol.* **18**(5), 581–586 (2008)
6. Bill, R.M., Henderson, P.J.F., Iwata, S., Kunji, E.R.S., Michel, H., Neutze, R., Newstead, S., Poolman, B., Tate, C.G., Vogel, H.: Overcoming barriers to membrane protein structure determination. *Nat. Biotechnol.* **29**(4), 335–340 (2011)
7. Landreh, M.A., Robinson, C.V.: A new window into the molecular physiology of membrane proteins. *J. Physiol.* **593**(2), 355–362 (2015)

8. Stansfeld, P.J., Goose, J.E., Caffrey, M., Carpenter, E.P., Parker, J.L., Newstead, S., Sansom, M.S.P.: MemProtMD: automated insertion of membrane protein structures into explicit lipid membranes. *Structure*. **23**(7), 1350–1361 (2015)
9. Barrera, N.P., Robison, C.V.: Advances in the mass spectrometry of membrane proteins: from individual proteins to intact complexes. *Annu. Rev. Biochem.* **80**, 247–271 (2011)
10. le Coutre, J., Whitelegge, J.P., Gross, A., Turk, E., Wright, E.M., Kaback, H.R., Faull, K.F.: Proteomics on full-length membrane proteins using mass spectrometry. *Biochemistry*. **39**, 4237–4242 (2000)
11. Savage, D.F., Egea, P.F., Robles-Colmenares, Y., O'Connell III, J.D., Stroud, R.M.: Architecture and selectivity in aquaporins 2.5: X-ray structure of Aquaporin Z. *PLoS Biol.* **3**, 334–340 (2003)
12. Wang, S.C., Politis, A., Di Bartolo, N., Bavro, V.N., Tucker, S.J., Booth, P.J., Barrera, N.P., Robinson, C.V.: Ion mobility mass spectrometry of two tetrameric membrane protein complexes reveals compact structures and differences in stability and packing. *J. Am. Chem. Soc.* **132**, 15468–15470 (2010)
13. Chang, Y.-C., Bowie, J.U.: Measuring membrane protein stability under native conditions. *Proc. Natl. Acad. Sci.* **111**(1), 219–224 (2014)
14. Fiset, O., Paslack, C., Barnes, R., Isas, J.M., Langen, R., Heyden, M., Han, S., Schafer, L.V.: Hydration dynamics of a peripheral membrane protein. *J. Am. Chem. Soc.* **138**, 11526–11535 (2016)
15. Marcoux, J., Robison, C.V.: Twenty years of gas phase structural biology. *Structure*. **21**(9), 1541–1550 (2013)
16. Whitelegge, J.P., Gundersen, C.B., Faull, K.F.: Electrospray-ionization mass spectrometry of intact intrinsic membrane proteins. *Protein Sci.* **7**, 1423–1430 (1998)
17. Barrera, N.P., Di Bartolo, N., Booth, P.J., Robison, C.V.: Micelles protect membrane complexes from solution to vacuum. *Science*. **321**(5886), 243–246 (2008)
18. Zhou, M., Morgner, N., Barrera, N.P., Politis, A., Isaacson, S.C., Matak-Vinkovic, D., Murata, T., Bernal, R.A., Stock, D., Robison, C.V.: Mass spectrometry of intact V-type ATPases reveals bound lipids and the effects of nucleotide binding. *Science*. **334**(6054), 380–385 (2011)
19. Housden, N.G., Hopper, J.T., Lukoyanova, N., Rodrigues-Larrea, D., Wojdyła, J.A., Klein, A., Kaminska, R., Bayley, H., Saibil, H.R., Robison, C.V., Kleanthous, C.: Intrinsically disordered protein threads through the bacterial outer-membrane porin OmpF. *Science*. **340**(6140), 1570–1574 (2013)
20. Laganowsky, A., Reading, E., Hopper, J.T., Robison, C.V.: Mass spectrometry of intact membrane protein complexes. *Nat. Protoc.* **8**(4), 639–651 (2013)
21. Mikhailov, V.A., Liko, I., Mize, T.H., Bush, M.F., Benesch, J.L., Robison, C.V.: Infrared laser activation of soluble and membrane protein assemblies in the gas phase. *Anal. Chem.* **88**(14), 7060–7067 (2016)
22. Campuzano, I.D., Li, H., Bagal, D., Lippens, J.L., Svitel, J., Kurzeja, R.J., Xu, H., Schnier, P.D., Loo, J.A.: Native MS Analysis of bacteriorhodopsin and an empty nanodisc by orthogonal acceleration time-of-flight, Orbitrap, and ion cyclotron resonance. *Anal. Chem.* **88**(24), 12427–12436 (2016)
23. Gault, J., Donlan, J.A., Liko, I., Hopper, J.T., Gupta, K., Housden, N.G., Struwe, W.B., Marty, M.T., Mize, T., Bechara, C., Zhu, Y., Wu, B., Kleanthous, C., Belov, M., Damoc, E., Makarov, A., Robison, C.V.: High-resolution mass spectrometry of small molecules bound to membrane proteins. *Nat. Methods*. **13**(4), 333–336 (2016)
24. Hopper, J.T.S., Yu, Y.T.-C., Li, D., Raymond, A., Bostock, M., Liko, I., Mikhailov, V., Laganowsky, A., Benesch, J.L.P., Caffrey, M., Nietlisbach, D., Robison, C.V.: Detergent-free mass spectrometry of membrane protein complexes. *Nat. Methods*. **10**, 1206–1208 (2013)
25. Calabrese, A.N., Watkinson, T.G., Henderson, P.J.F., Radford, S.E., Ashcroft, A.E.: Amphiphils outperform dodecylmaltoide micelles in stabilizing membrane protein structure in the gas phase. *Anal. Chem.* **87**(2), 1118–1126 (2015)
26. Marty, M.T., Hoi, K.K., Robison, C.V.: Interfacing membrane mimetics with mass spectrometry. *J. Am. Chem. Soc.* **49**(11), 2459–2467 (2016)
27. Bayburt, T.H., Grinkova, Y.V., Sligar, S.G.: Self-assembly of discoidal phospholipid bilayer nanoparticles with membrane scaffold proteins. *Nano Lett.* **2**(8), 853–856 (2002)
28. Marty, M.T., Zhang, H., Cui, W., Blankenship, R.E., Gross, M.L., Sligar, S.G.: Native mass spectrometry characterizes intact nanodisc lipoprotein complexes. *Anal. Chem.* **84**(21), 8957–8960 (2012)
29. Dorr, J.M., Koorengel, M.C., Schafer, M., Prokofyev, A.V., Scheideleer, S., van der Crujisen, E.A.W., Dafforn, T.R., Baldus, M., Killian, J.A.: Detergent-free isolation, characterization, and functional reconstitution of a tetrameric K⁺ channel: the power of native nanodiscs. *Proc. Natl. Acad. Sci.* **111**(52), 18607–19612 (2014)
30. van de Waterbeemd, M., Snijder, J., Tsvetkova, I.B., Dragnea, B.G., Cornelissen, J.J., Heck, A.J.R.: Examining the heterogenous genome content of multipartite viruses BMV and CCMV by native mass spectrometry. *J. Am. Soc. Mass Spectrom.* **27**(6), 1000–1009 (2016)
31. Lange, O., Damoc, E., Wieghaus, A., Makarov, A.: Enhanced Fourier transform for Orbitrap mass spectrometry. *Int. J. Mass Spectrom.* **369**, 338–344 (2014)
32. Makarov, A., Denisov, E., Lange, O., Horning, S.: Dynamic range of mass accuracy in LTQ Orbitrap hybrid mass spectrometer. *J. Am. Soc. Mass Spectrom.* **17**(7), 977–982 (2006)
33. Ferrige, A.G., Seddon, M.J., Green, B.N., Jarvis, S.A., Skilling, J., Staunton, J.: Disentangling electrospray spectra with maximum entropy. *Rapid Commun. Mass Spectrom.* **6**(11), 707–711 (1992)
34. Laganowsky, A., Reading, E., Allison, T.M., Ulmschneider, M.B., Degiacomi, M.T., Baldwin, A.J., Robison, C.V.: Membrane proteins bind lipids selectively to modulate their structure and function. *Nature*. **510**(7503), 172–175 (2014)
35. Liu, Y., Cong, X., Liu, W., Laganowsky, A.: Characterization of membrane protein–lipid interactions by mass spectrometry ion mobility mass spectrometry. *J. Am. Soc. Mass Spectrom.* **28**(4), 579–586 (2017)
36. Cong, X., Liu, Y., Liu, W., Liang, X., Russell, D.H., Laganowsky, A.: Determining membrane protein–lipid binding thermodynamics using native mass spectrometry. *J. Am. Chem. Soc.* **138**, 4346–4349 (2016)
37. Lippens, J.L., Mangrum, J.B., McIntyre, W., Redick, B., Fabris, D.: A simple heated-capillary modification improves the analysis of non-covalent complexes by Z-spray electrospray ionization. *Rapid Commun. Mass Spectrom.* **30**, 773–783 (2016)
38. Borysik, A.J., Hewitt, D.J., Robison, C.V.: Detergent release prolongs the lifetime of native-like membrane protein conformations in the gas-phase. *J. Am. Chem. Soc.* **135**, 6078–6083 (2013)
39. Reading, E., Liko, I., Allison, T.M., Benesch, J.L.P., Laganowsky, A., Robison, C.V.: The role of the detergent micelle in preserving the structure of membrane proteins in the gas phase. *Angew. Chem.* **54**(15), 4577–4581 (2015)
40. Smith, L.M., Kelleher, N.L.: Proteoform: a single term describing protein complexity. *Nat. Methods*. **10**, 186–187 (2013)
41. Benesch, J.L., Aquilina, J.A., Ruotolo, B.T., Sobott, F., Robison, C.V.: Tandem mass spectrometry reveals the quaternary organization of macromolecular assemblies. *Chem. Biol.* **13**(6), 597–605 (2006)
42. Benesch, J.L., Ruotolo, B.T., Sobott, F., Wildgoose, J., Gilbert, A., Bateman, R., Robison, C.V.: Quadrupole-time-of-flight mass spectrometer modified for higher-energy dissociation reduces protein assemblies to peptide fragments. *Anal. Chem.* **81**(3), 1270–1274 (2009)
43. Pagel, K., Hyung, S.-J., Ruotolo, B.T., Robison, C.V.: Alternative dissociation pathways identified in charge-reduced protein complex ions. *Anal. Chem.* **82**(12), 5363–5372 (2010)
44. Erba, E.B., Ruotolo, B.T., Barsky, D., Robison, C.V.: Ion mobility-mass spectrometry reveals the influence of subunit packing and charge on the dissociation of multiprotein complexes. *Anal. Chem.* **82**(23), 9702–9710 (2010)
45. Lossl, P., Snijder, J., Heck, A.J.R.: Boundaries of mass resolution in native mass spectrometry. *J. Am. Soc. Mass Spectrom.* **25**(6), 906–917 (2014)
46. Katada, T., Ui, M.: Direct modification of the membrane adenylate cyclase system by islet-activating protein due to ADP-ribosylation of a membrane protein. *Proc. Natl. Acad. Sci.* **79**, 3129–3133 (1982)
47. Yen, H.-Y., Hopper, J.T.S., Liko, I., Allison, T.M., Zhu, Y., Wang, D., Stegmann, M., Mohammed, S., Wu, B., Robison, C.V.: Ligand binding to a G protein-coupled receptor captured in a mass spectrometer. *Sci. Adv.* **3**(6), e1701016 (2017)
48. Hirel, P.-H., Schmitter, J.-M., Dessen, P., Fayat, G., Blanquet, S.: Extent of N-terminal methionine excision from *Escherichia coli* proteins is governed by the side-chain length of the penultimate amino acid. *Proc. Natl. Acad. Sci.* **86**, 8247–8251 (1989)
49. Spector, S., Flynn, J.M., Tidor, B., Baker, T.A., Sauer, R.T.: Expression of N-formylated proteins in *Escherichia coli*. *Protein Expr. Purif.* **32**, 317–322 (2003)

SUPPORTING INFORMATION

***Escherichia coli* AquaporinZ purification.** The gene encoding *E. coli* AqpZ was cloned in a pET29b expression vector to express the protein fused to a C-terminal hexahistidine tag cleavable with thrombin. AqpZ was expressed in C43 (DE3) cells grown in 2x LB media at 37 °C. Protein expression was induced by adding IPTG at a 0.8 mM final concentration when cultures reached an OD600 of 0.5; induction was carried out for 5 hours at 37 °C before cells were harvested and washed. Cells were lysed in 500 mM NaCl/ 20 mM Tris-HCl pH=8.0/ 10 % glycerol/ 0.1 mM PMSF and 2.8 mM β -mercaptoethanol (β -ME) using a C3-Emulsiflex (Avestin) pressurized at 20,000 psi. Following lysis, the extract was clarified by low-speed centrifugation at 10,000 g for 30 minutes at 4 °C prior to separating the bacterial membranes by ultracentrifugation at 120,000 g for 2 hours at 4 °C. Bacterial membrane pellets were then solubilized in lysis buffer supplemented with 200 mM OG at 4 °C overnight. Insoluble material was pelleted by ultracentrifugation at 120,000 g for 1 hour at 4 °C. Detergent-solubilized AqpZ was then purified by immobilized-nickel affinity chromatography. After loading the solubilization extract, the resin was washed with 25 column volumes of 500 mM NaCl/ 20 mM Tris-HCl pH=8.0/ 10% glycerol/ 25 mM imidazole/ 40 mM OG/ 0.1 mM PMSF and 1.4 mM β -ME; AqpZ was then eluted in 500 mM NaCl/ 20 mM Tris-HCl pH=8.0/ 250 mM imidazole/ 40 mM OG/ 0.1 mM PMSF and 1.4 mM β -ME. AqpZ was then desalted on a PD-10 desalting column equilibrated in 150 mM NaCl/ 20 mM Tris-HCl pH=8.0/ 5% glycerol/ 1.4 mM β -ME. Following desalting, AqpZ was treated with thrombin for 16 hours at ambient temperature to remove the histidine tag then purified on a Superdex 200 HR10/30 size exclusion column (GE Healthcare) equilibrated in 150 mM NaCl/ 20 mM Tris-HCl pH=8.0/ 5% glycerol and 40 mM OG. Pure *E. coli* AqpZ eluted as a single peak corresponding to the homotetramer [1]. AqpZ fractions were pooled and passed across a small Ni-IMAC column (250 μ l resin) to remove any remnants of tagged protein. The protein was then concentrated to 6.5 mg/ml for subsequent analysis.

Mass Spectrometer Instrument Parameters. Q-ToF: The Synapt G2 HDMS instrument was operated in positive nanoflow-ESI mode. All critical instrument voltages and pressures are as follows: capillary voltage 0.5-1kV; sample cone 40V, extraction cone 1V; source block temperature 30 °C; trap collision energy 4.0 to 110V; transfer collision energy 3V; trap entrance 0V; trap bias 2V; trap DC -2.0V; trap exit 0.0V; IMS entrance 25V; IMS helium cell DC 30V; IMS helium exit -20.0V; IMS Bias 2.0V; IMS exit 0.0V; transfer entrance 4.0V; transfer exit 5.0V; IMS wave velocity 250 m/sec; IMS wave amplitude 22.0V; transfer velocity 47 m/sec; transfer wave amplitude 4.0V; mobility trapping release time 100 μ sec; trap height 20.0V; extract height 0.0V; source RF-amplitude (peak-to-peak) 350V; triwave RF-amplitudes (peak-to-peak) trap 320V, IMS 250V, transfer 250V; source backing pressure 6.0 mbar; trap/transfer pressure SF₆, 2.93e⁻² mbar (pirani gauge indicated; flow rate 4.0 mL/min); Instrument control and data acquisition was carried out through MassLynx 4.1 SCN 781.

FT-ICR: The experiments were performed using a 15 Tesla Bruker Solarix FT-ICR-MS instrument possessing an ICR infinity cell. The nESI capillary voltage was set to 0.7~0.95 kV. The temperature of dry gas was 100 °C and the flow rate was 3.0 L/min. The RF amplitude of the ion-funnels was 300 V_{pp}, and the applied voltages were 150 V and 6 V for funnels 1 and 2, respectively. The voltage of skimmer 1 ranged from 50 to 125 V, skimmer 2 voltage was kept at 5 V. The lowest values of RF frequencies were used in all ion-transmission regions: multipole 1 (2 MHz), quadrupole (1.4 MHz), and transfer hexapole (1MHz). Ions were accumulated for 500 ms in the hexapole collision cell before being transmitted to the infinity ICR cell. The time-of-flight of 2.5 ms was used. Vacuum pressures for different regions were ~3 mbar for the source region, ~3×10⁻⁶ mbar for the quadrupole region, and ~2×10⁻⁹ mbar for the UHV-chamber pressure. Various types of activation techniques were performed to liberate membrane proteins from detergent micelles. In-source dissociation (ISD) was performed by varying the voltage of skimmer 1 from 50 to 125 V. Collision induced dissociation (CID) was performed in the hexapole collision cell by colliding ions with SF₆. The mass

spectrometer was externally calibrated with cesium iodide. Data was collected in magnitude mode with 300 scans from m/z 153–10,000 averaged for each spectrum and recorded at 512k data points.

Both instruments were externally calibrated using a 50 $\mu\text{g}/\mu\text{L}$ solution of cesium iodide in 50% acetonitrile/water solution.

References

1. Savage, D.F.E., Pascal F.; Robles-Colmenares, Yaneth; O'Connell III, Joseph D. and Stroud, Robert M., *Architecture and Selectivity in Aquaporins 2.5: A X-Ray Structure of Aquaporin Z*. PLOS Biology, 2003. **1**(3): p. 334-340.

Skimmer 1 at 125 V

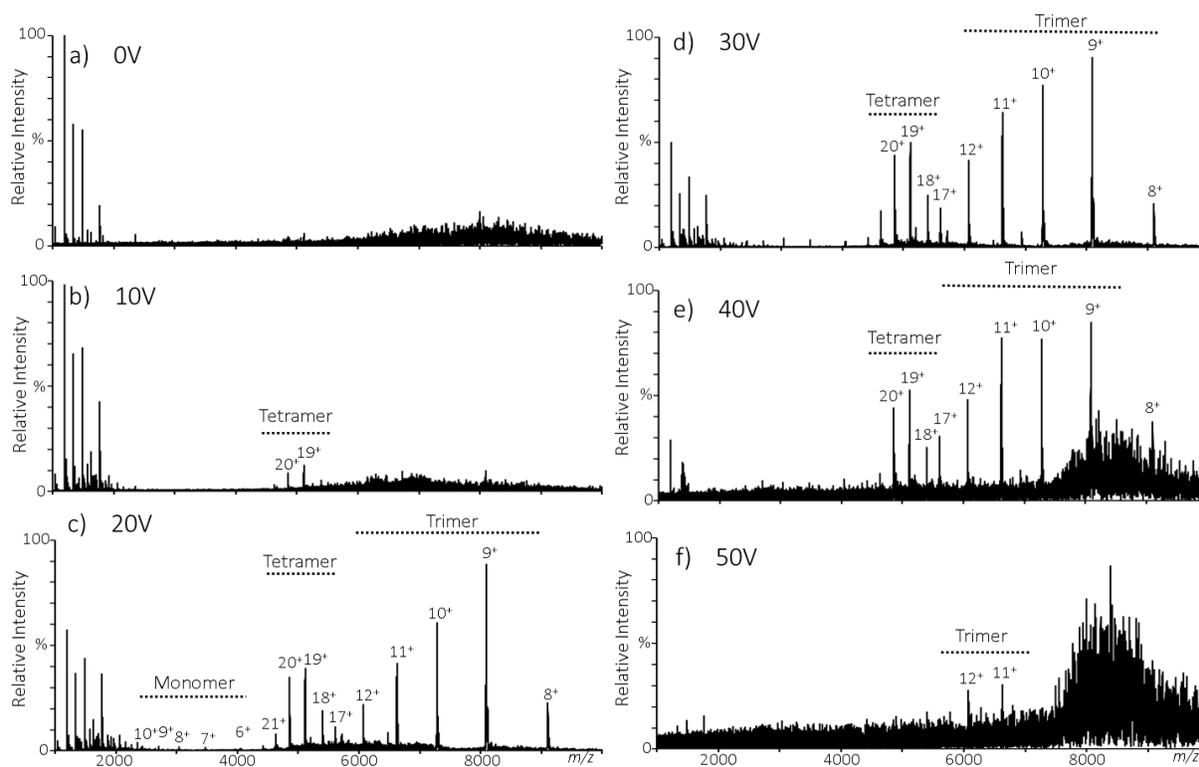


Figure S1. Ejection of AqpZ from OG micelles as a function of collision energy at Skimmer 1 of 125 V on the FT-ICR. At this higher skimmer 1 setting, shown here, more intense fragmentation is observed as the amount of CID voltage is increased. a) No collision energy shows only a detergent micelle distribution; while b) low tetramer signal can be observed at 10 V CE. Gradual increases of CE (10 V increments) from c) to f) results intense tetramer signal co-observed with trimer signal which remains in equal or higher abundance than the intact tetramer.

Skimmer 1 at 100 V

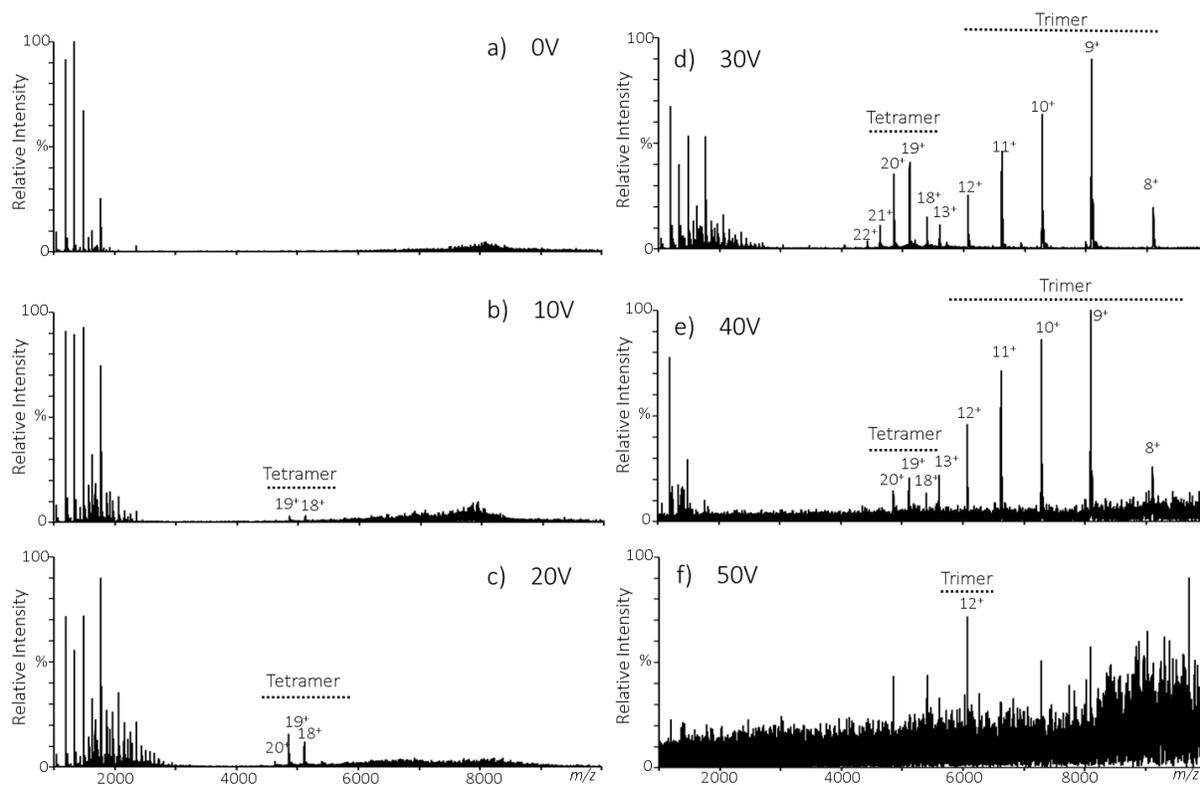


Figure S2. Ejection of AqpZ from OG micelles as a function of collision energy at Skimmer 1 of 100 V on the FT-ICR. a) No collision energy shows only a detergent micelle distribution while low tetramer signal can be observed at 10 and 20 V CE (b and c). Gradual increases of CE (10 V increments) from d) to f) results intense tetramer signal co-observed with trimer signal which remains in equal or higher abundance than the intact tetramer.

Skimmer 1 at 75 V

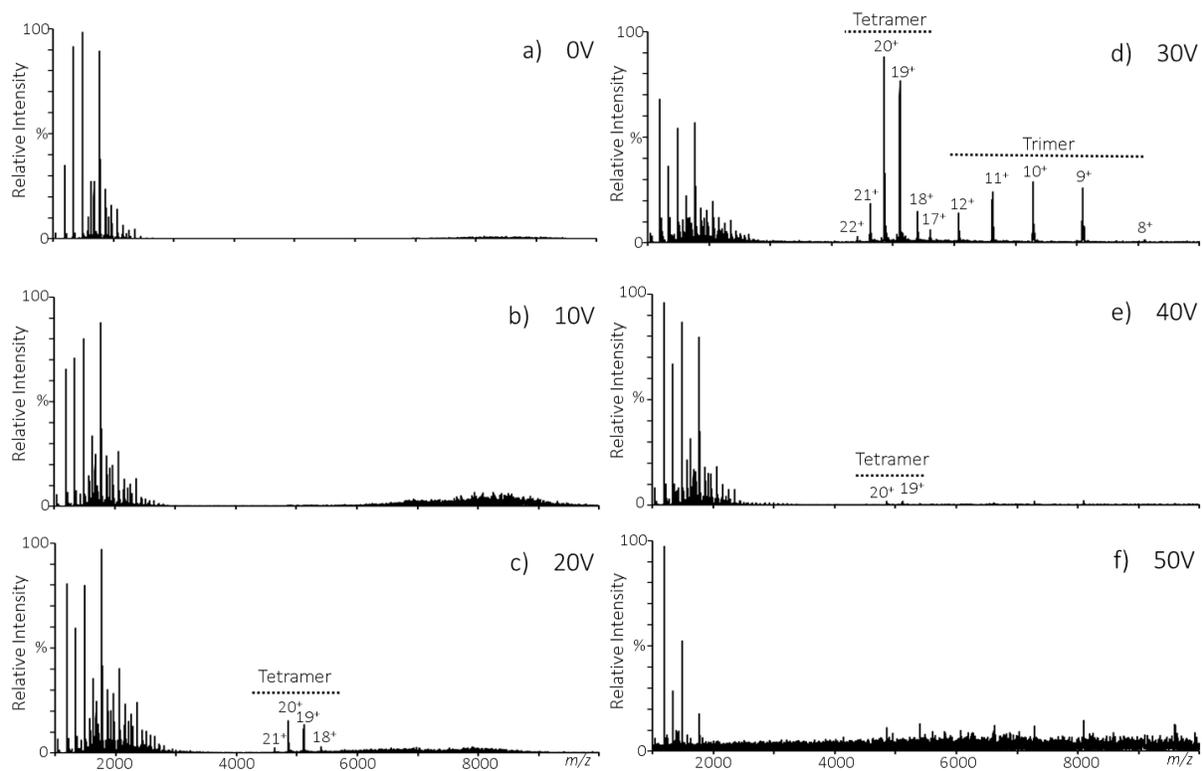


Figure S3. Ejection of AqpZ from OG micelles as a function of collision energy at Skimmer 1 of 75 V on the FT-ICR. This skimmer setting in combination with 20 V CE (c) showed only intact tetramer signal, while 30 V (d) generated fragmentation into the trimer. Minimal tetramer signal was achieved at the lower and highest voltage settings (a, b and f).

Skimmer 1 at 50 V

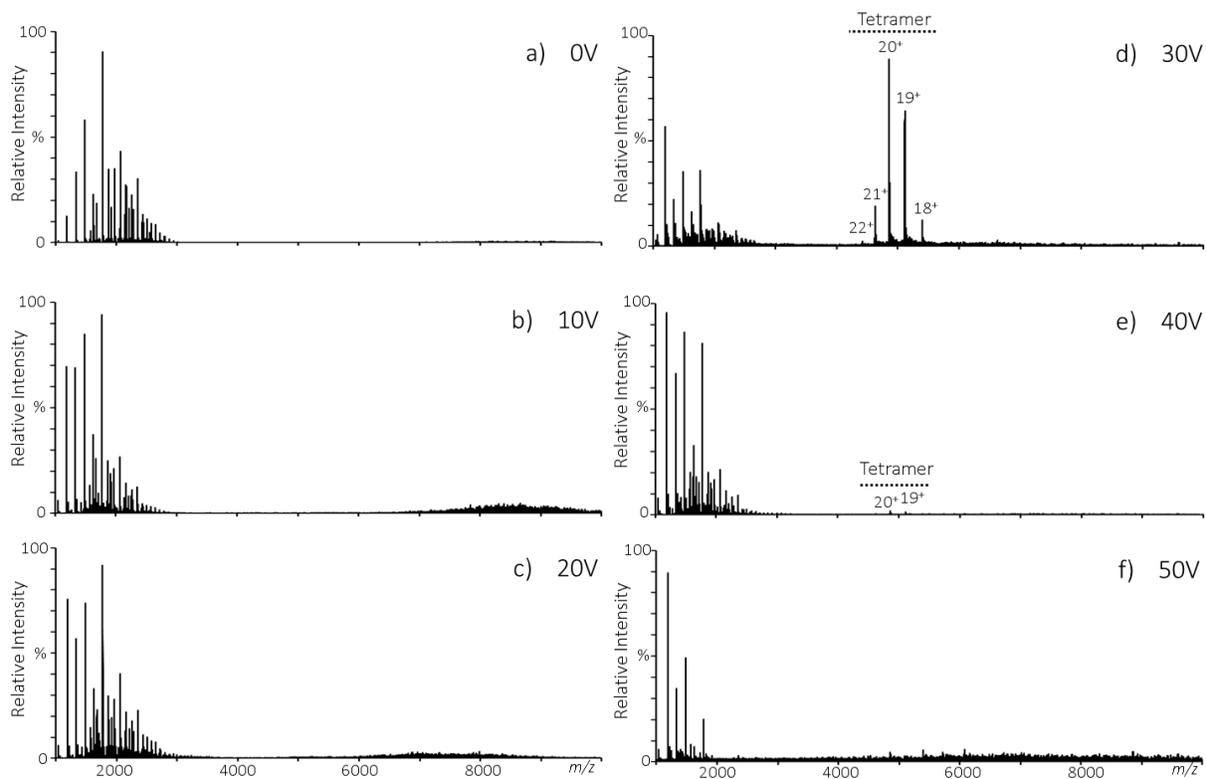


Figure S4. Ejection of AqpZ from OG micelles as a function of collision energy at Skimmer 1 of 50 V on the FT-ICR, a) through f). AqpZ was observed solely with c) a CID of 30 V at this skimmer 1 setting. Unlike the other Skim1 settings this low setting provided solely intact tetramer within a very narrow CE range. However, no fragmentation of the tetramer was observed, even at the highest CE investigated.

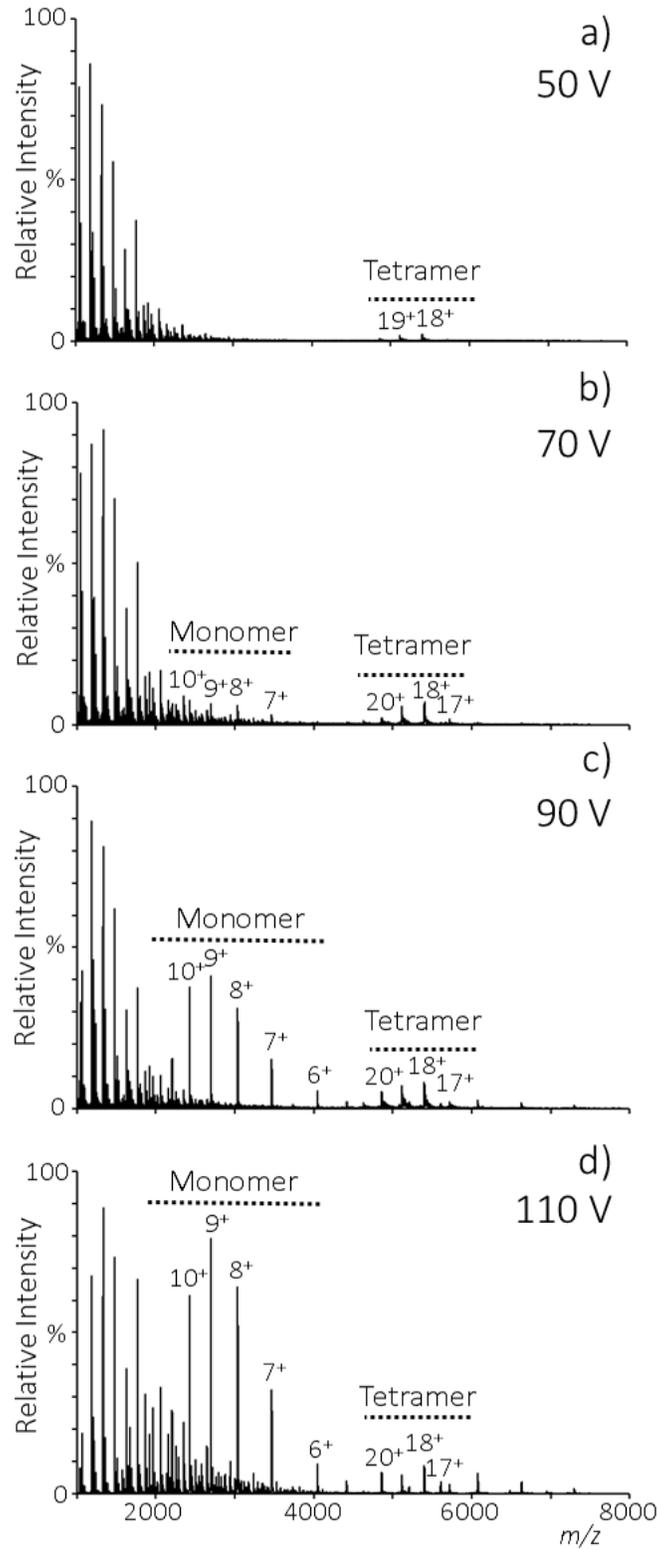


Figure S5. Ejection of AqpZ from OG micelles on the ion mobility Q-ToF as a function of Trap CE, increasing from a) 50 V; b) 70 V; c) 90 V to d) 110 V observed with a sampling cone setting of 40 V on the Q-ToF. Increasing monomer signal is observed with increasing TrapCE.

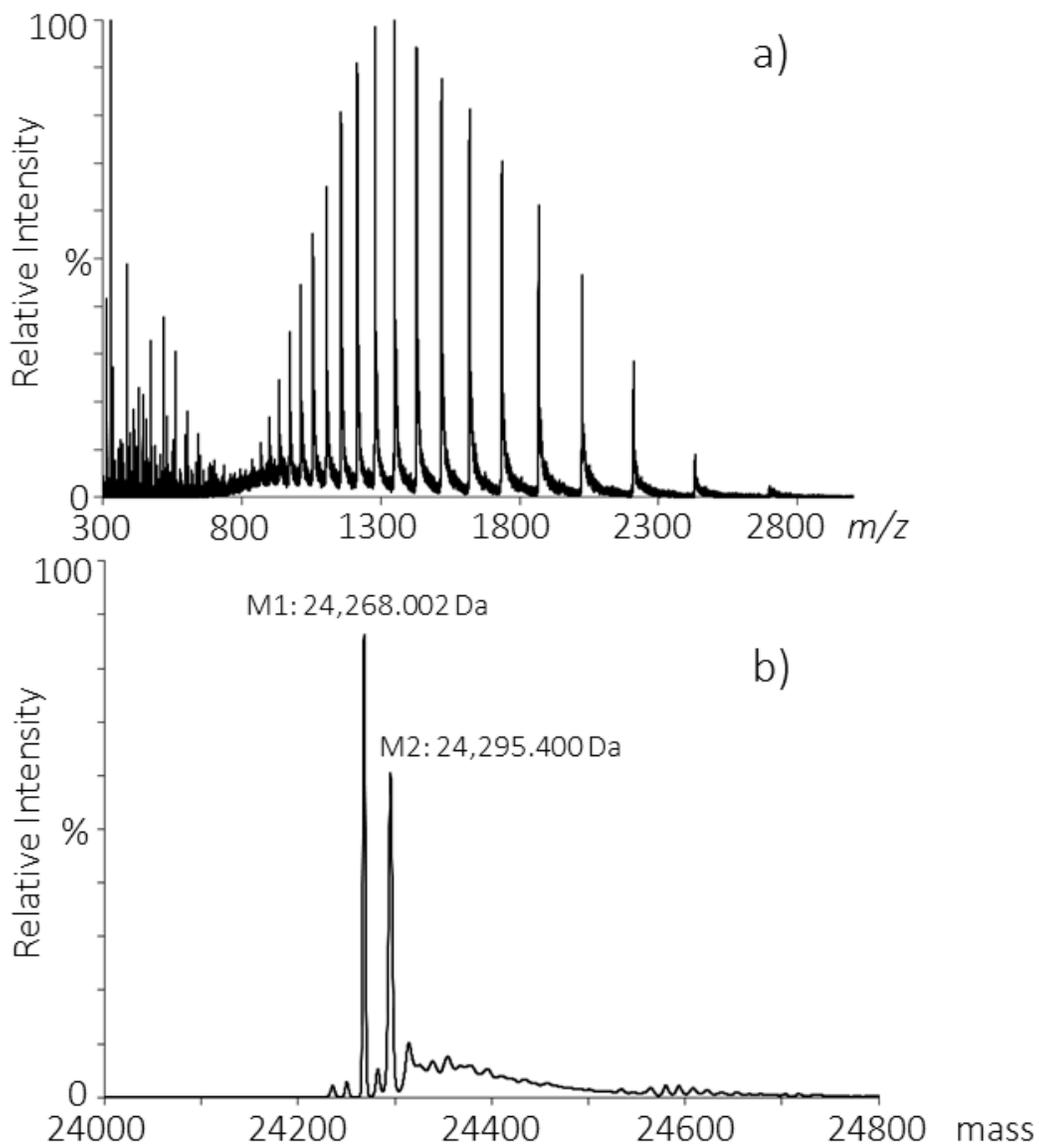


Figure S6. Denaturing LC-MS results from the ion-mobility Q-ToF a) Full mass spectrum from m/z 300-3000 and b) Deconvoluted zero-charge masses of monomer species. Two monomeric species detected are denoted as M1 and M2, respectively.

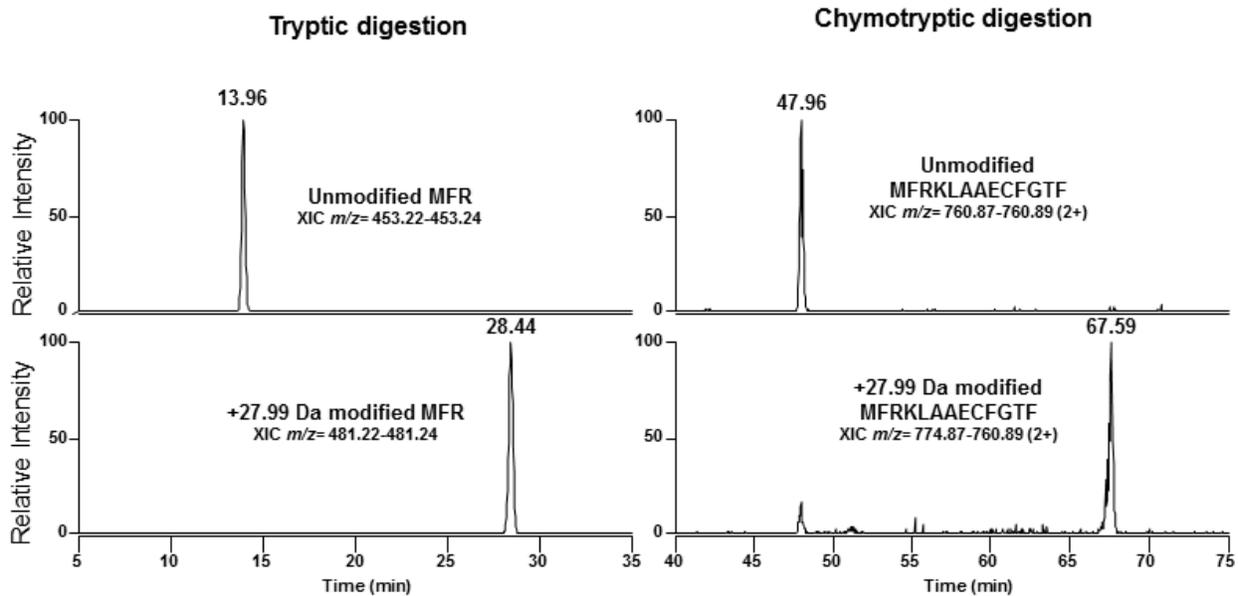


Figure S7. Chromatography results from both the chymotrypsin (left) and trypsin (right) digests. As shown, the unmodified peptides maintain a shorter retention time as compared to the 27.99 Da modified peptides in both digests. This shift in retention time, helped to support our hypothesis that the existing modification on the monomer was in fact present in the N-terminal region.