Dividing To Unveil Protein Microheterogeneities: Traveling Wave Ion Mobility Study

ARTICLE *in* ANALYTICAL CHEMISTRY · JULY 2011

Impact Factor: 5.64 · DOI: 10.1021/ac200994c · Source: PubMed

6 AUTHORS, INCLUDING:

Johnny Habchi
University of Cambridge
33 PUBLICATIONS **322** CITATIONS
SEE PROFILE

Laetitia Cravello
Waters Corporation
10 PUBLICATIONS **127** CITATIONS
SEE PROFILE

Marlène Martinho
Aix-Marseille Université
31 PUBLICATIONS **804** CITATIONS
SEE PROFILE

Sonia Longhi
Architecture et Fonction des Macromolécule…
126 PUBLICATIONS **4,432** CITATIONS
SEE PROFILE

analytical chemistry

ARTICLE

pubs.acs.org/ac

Dividing To Unveil Protein Microheterogeneities: Traveling Wave Ion Mobility Study

F. Halgand,[*,†] Johnny Habchi,[§] Laetitia Cravello,[∥] Marlène Martinho,[‡] Bruno Guigliarelli,[‡] and Sonia Longhi[§]

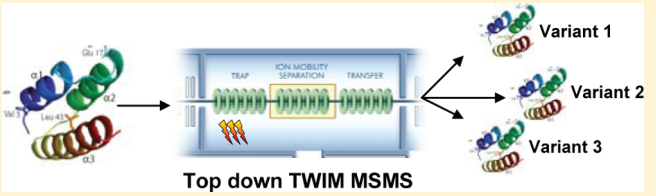
[†]Laboratoire de Bioénergétique et Ingénierie des Protéines, Equipe de Protéomique Fonctionnelle et Dynamique, and
[‡]Laboratoire de Bioénergétique et Ingénierie des Protéines, Equipe de Biophysique des Métalloprotéines, UPR 9036-CNRS, 31 Chemin Joseph Aiguier, 13420 Marseille Cedex, France

[§]Architecture et Fonction des Macromolécules Biologiques (AFMB), UMR 6098, CNRS et Aix-Marseille Université, 163 Avenue de Luminy Case 932, 13288 Marseille Cedex 09, France

[∥]Waters Corporation, Atlas Park, Simonsway, Manchester, Lancashire M22 5PP, United Kingdom

Supporting Information

ABSTRACT: Overexpression of a protein in a foreign host is often the only route toward an exhaustive characterization, especially when purification from the natural source(s) is hardly achievable. The key issue in these studies relies on quality control of the purified recombinant protein to precisely determining its identity as well as any undesirable microheterogeneities. While standard proteomics approaches preclude unbiased search for modifications, the optional technique of top-down tandem mass spectrometry (MSMS) requires the use of highly accurate and highly resolved experiments to reveal subtle sequence modifications. In the present study, the top-down MSMS approach combined with traveling wave ion mobility (TWIM) separation was evaluated for its ability to achieve high sequence coverage and to reveal subtle microheterogeneities that were hitherto only accessible with Fourier-transform ion cyclotron resonance−MS instruments. The power of this approach is herein illustrated in an in-depth analysis of both the wild type and K496C variant of the recombinant X domain (XD; aa's 459−507) of the measles virus phosphoprotein expressed in *Escherichia coli*. Using top-down MSMS combined with TWIM, we show that XD samples occasionally exhibit a microheterogeneity that could not be anticipated from the nucleotide sequence of the encoding constructs and that likely reflects a genetic drift, neutral or not, occurring during expression. In addition, a 1-oxyl-2,2,5,5-tetramethyl-δ3-pyrroline-3-methyl methanethiosulfonate nitroxide probe that was grafted onto the K496C XD variant was shown to undergo oxidation and/or protonation in the electrospray ionization source, leading to artifactual mass increases.



Overexpression of a protein in a foreign host, such as *Escherichia coli*, is frequently the first step toward biochemical, enzymatic, and structural studies and is instrumental when purification from the natural source(s) is hardly achievable. High-level production of functional heterologous proteins in *E. coli* often remains difficult in spite of the improvements achieved in the past decade. The latter involve the production of proteins not only in the cytoplasm but also in the periplasm or the acquisition of a proper protein fold through manipulation of the thioredoxin pathway. However, heterologous protein overexpression in *E. coli* continues to be a challenging task for proteins possessing numerous disulfide bridges and/or being the target of post-translational modifications or when genes enriched in rare codons (i.e., codons that are used with very low frequency in this host) have to be expressed. Despite these limitations, bacterial expression often yields reasonable amounts of eukaryotic and/or viral proteins that can then be extensively studied to get biological and structural insights. The key issue in these studies is to obtain large amounts of the purified recombinant protein with a homogeneity as high as possible prior to proceeding to its biochemical and structural characterization. This requirement is deeply interconnected with the necessity of precisely determining the identity of the recombinant protein and of fully unraveling its primary structure, as well as with the need of unveiling any possible chemical modifications leading to undesirable microheterogeneities (for a review, see ref 1).

Traditional approaches used for quality control of recombinant proteins are based on "bottom-up" proteomics methodologies. Although a wealth of literature reports pointed out the successful use of this approach, the latter suffers from some limitations when it comes to determining the full complexity of a protein sample. For this purpose the top-down MSMS approach has been developed.[2−4] This combines the measurement of the intact experimental mass (intact mass tag, IMT) with the recording of MSMS data on the full-length protein. Such a technique is becoming more and more popular since it allows an extensive description of protein properties.[5−7] It appears from these studies that (i) highly resolved and highly accurate MSMS experiments are necessary to describe in detail subtle changes that apparently occur in a much larger extent than previously expected and that (ii) downstream

Received: May 2, 2011
Accepted: July 29, 2011
Published: July 29, 2011

sample preparation, and data handling still remain critical to achieving meaningful results.[2,3,8,9] Thus, successful and complete protein property characterization and unveiling of potential subtle modifications using the top-down MSMS approach are subordinated to the achievement of good fragmentation yields, good signal-to-noise ratios, and limited isotopic profiles overlapping. For example, the complete description of a 4.3 kDa protein bearing subtle modifications of ~±1 Da mass differences using top-down MSMS was only achieved using a Fourier-transform ion cyclotron resonance (FTICR)−MS instrument, which allowed these species to be distinguished at 750 000 resolution with less than 2 ppm.[8] However, such instruments are not common due to their high cost, and alternative approaches would be of potential interest.

The resurgence of ion mobility in the field of mass spectrometry provides an additional dimension for separation of compounds in the traveling wave ion mobility (TWIM) guide containing an inert gas (usually nitrogen) settled in a commercially available quadrupole time-of-flight mass spectrometer (Synapt HDMS, Waters Corp., Manchester, England). TWIM experiments allow differentiating molecules according to their charge state (CS), collisional cross section (CCS), and size with the determination of drift times, a parameter reflecting gas-phase conformation of proteins and peptides (for further details about TWIM, see refs 10 and 11). The TWIM guide, either alone[12] or in conjunction with hydrogen exchange,[13−15] allows inferring, in the submillisecond time scale, additional insights on ligand−protein interactions. Such a technology also allowed unveiling conformational changes upon ligand binding under non-covalent conditions[16] or revealing structural changes in proteins involved in aggregative neurodegenerative diseases.[17] Proteomics approaches also benefit from this new tool, as exemplified by studies reporting the successful characterization of glycans and the direct identification of proteins in tissue sections.[18,19]

Since a TWIM device implemented on a QToF mass spectrometer allowed gaining another dimensional fractionation of fragment ions, we herein evaluated the suitability of this system to unveil subtle and/or minor changes in protein samples with a precision comparable to that of highly resolved experiments obtained with FTICR−MS.

In the present study, we combined the top-down approach with TWIM separation to achieve the best possible separation of fragment ions in view of assessing the homogeneity of a recombinant protein purified from the soluble fraction of *E. coli*, namely, the C-terminal X domain (XD; aa's 459−507) of the measles virus (MV) phosphoprotein (P). Recent studies allowed the determination of the structure of XD by both X-ray crystallography[20] and NMR.[21] The XD protein was thereafter extensively used in numerous biochemical and biophysical studies that unveiled its role in triggering α-helical folding of the intrinsically disordered C-terminal domain (N_{TAIL}; aa's 401−525) of the MV nucleoprotein (N).[22−30] Those studies provided structural details on the N_{TAIL} binding interface of XD[31−34] and designated residue K496 of XD as an interesting target for cysteine substitution and ensuing labeling with an electron paramagnetic resonance (EPR) label, such as 1-oxyl-2,2,5,5-tetramethyl-δ3-pyrroline-3-methyl methanethiosulfonate (MTSL), in view of studies aimed at inferring information in terms of spin probe mobility under various conditions.

MATERIALS AND METHODS

Materials. All chemical reagents, such as methanol, 2-propanol, acetonitrile, formic acid, and cesium iodide, were purchased from Sigma-Aldrich at the highest HPLC purity grade.

Methods. *Construction of Protein Expression Plasmids.* The XD gene construct (pDest14/XD_{HC}), encoding the XD (residues 459−507) of the MV P protein (strain Edmonston B) with a hexahistidine tag fused to its C-terminus and under the control of the T7 promoter, has already been described.[2] The pDest14/XD_{HC} K496C construct, encoding C-terminally hexahistidine-tagged XD bearing the K496C substitution, was obtained by polymerase chain reaction (PCR) using pDest14/XD_{HC} as a template, Turbo-Pfu polymerase (Stratagene), and a pair of complementary mutagenic primers of 39 nucleotides in length (Operon). After digestion with *Dpn*I to remove the methylated DNA template, transformation of *E. coli* strain DH10β (Stratagene) with the amplified PCR product was carried out. The nucleotide sequence of both strands of the open reading frame (ORF) of both native and K496C XD constructs was determined (GATC Biotech, Germany)

Expression and Purification of XD Proteins. The *E. coli* strain Rosetta [DE3] [pLysS] (Novagen) was used for the expression of both native and K496C XD. Expression and induction were carried out as already described.[2] The induced cells were harvested, washed, and collected by centrifugation. The resulting pellets were frozen at −20 °C. Purification of histidine-tagged wt and K496C XD was carried out as described in ref 2. The proteins (in 10 mM sodium phosphate buffer, pH 7) were stored at −20 °C. The apparent molecular mass of proteins eluted from gel filtration columns was deduced from a calibration curve established with low and high molecular weight calibration kits (GE Healthcare). The theoretical Stokes radius (R_S) of a folded protein of the same molecular mass as XD was calculated according to ref 35. Protein concentrations were estimated using the Biorad protein assay (Bio-Rad).

Circular Dichroism. Circular dichroism (CD) spectra were recorded on a Jasco 810 dichrograph using 1 mm thick quartz cells in 10 mM sodium phosphate, pH 7, at 20 °C. CD spectra were measured between 185 and 260 nm at 0.2 nm/min and were averaged from three scans. The spectra were corrected for buffer signal and smoothed by using a third-order least-squares polynomial fit. Finally, mean ellipticity values per residue ([Θ]) were calculated as $[\Theta] = 3300m\Delta A/lcn$, where l (path length) = 0.1 cm, n = number of residues, m = molecular mass (Da), and c = protein concentration (mg/mL). The number of residues (n) is 56 for both XD proteins, and m is 6686 Da (wt XD) or 6661 Da (XD K496C). Protein concentrations of 0.1 mg/mL were used.

Spin-Labeling of the XD K496C Variant. Before spin-labeling, dithiothreitol (DTT) was added (to yield a final concentration of 20 mM) to the K496C XD protein (125 μL at 459 μM). The mixture was incubated for 30 min in an ice bath to reduce the unique free cysteine residue. DTT was removed by gel filtration using a PD-10 column (GE Healthcare) and 10 mM 2-morpholinoethanesulfonic acid (MES), 150 mM NaCl at pH 6.5 as the elution buffer. The fractions containing the protein were pooled. Prior to utilization, the spin label MTSL (Toronto Research Chemicals Inc., Toronto, Canada) was solubilized in acetonitrile to yield a final concentration of 10 mg/mL. Then it was immediately incorporated into the protein at a molar excess of 20 every hour for 4 h, leading to a total molar excess of 80. The reaction was carried out for 4 h in the dark and in an ice bath under gentle stirring and a continuous flow of argon to avoid oxidation. The excess of unbound spin label was removed by gel filtration as described above, except that 10 mM sodium phosphate at pH 7 was used as the elution buffer. The fractions containing the labeled protein were pooled, and the labeling of

the XD protein was controlled by EPR spectroscopy at 296 K on an ESP 300E Bruker spectrometer equipped with an ELEXSYS super-high-sensitivity resonator operating at 9.9 GHz. The spin-labeled K496C XD variant is hereafter referred to as K496C* XD.

Sample Preparation Prior to MS Experiments. Samples were desalted through cold acetone precipitation prior to analysis. To that purpose, 5 μL of concentrated XD protein solution (~150 μM) was diluted to a final volume of 100 μL using a 50 mM ammonium acetate buffer at pH 7.5. Then 1 mL of 80% cold acetone was added. The mixture was vigorously shaken and incubated for 24 h at −20 °C. After a 10 min centrifugation at 14 000 rpm, protein pellets were recovered and dissolved in a MeOH/H_2O (50/50, v/v) solution containing 1% formic acid to a final concentration of 2 μM. In the case of the K496C XD variant, which was shown to consist of a mixture of monomers and covalent dimers in solution, we introduced an additional step of reduction and alkylation of cysteine residues. For that purpose, DTT was added to a final concentration of 500 μM (10-fold excess) and the resulting solution incubated for 1 h at room temperature. Then cysteine alkylation was achieved by adding iodoacetamide at a final concentration of 1.5 mM (3-fold excess with respect to DTT) and incubating the resulting solution for 30 min in the dark at room temperature.

Electrospray Ionization Mass Spectrometry (ESI-MS) Experiments. ESI-MS spectra of full-length XD proteins were recorded in positive ion mode using a "nanoflow" electrospray ion source fitted on a quadrupole time-of-flight hybrid mass spectrometer equipped with a TWIM cell (Synapt HDMS G1, Waters Corp., Manchester, U.K.). Protein solutions were introduced at a 0.5 μL/min flow rate, and mass spectra were recorded for 2 min for full-length proteins and 15 min for top-down TWIM MSMS experiments. Instrumental parameters were optimized to attain the best signal-to-noise ratio. Calibration was achieved using a 1 mg/mL cesium iodide solution dissolved in 70% 2-propanol. Typical mass errors on MSMS fragments ranged from 11 to 25 ppm depending on the sample analyzed.

Traveling Wave Ion Mobility Experiments on Full-Length Proteins. Instrumental parameters were optimized to achieve the best possible separation of ions displaying different mobilities. For mobility separation of full-length wt and K496C variants either treated with iodoacetamide (IAA) or not, or labeled with MTSL or not, the wave height, the wave velocity, and the nitrogen pressure in the trap T-wave were 0.2 V, 300 m/s, and 4.5 × 10^{-3} mBar, respectively. The wave height, the wave velocity, and the nitrogen pressure in the ion mobility cell were 8 V, 300 m/s, and 0.5 mbar, respectively. The wave height and the wave velocity in the ion transfer cell were 3 V and 248 m/s, respectively. TWIM experimental conditions identical to those used for wt XD were used to record drift times of peptides from an enolase digest and to relate drift times to their CCS. This step was achieved using the CCS calibration tools within the Driftscope software.

Top-Down ESI-TWIM-MSMS. MSMS spectra were recorded from 50 to 2500 m/z mass range using argon as the collision gas (99.999% argon). The collision energy was optimized to achieve the best fragmentation yield of the selected parent ion. Typically, the collision energy was about 35−90 eV, depending on the protein CS. Proteins were fragmented using various CSs to obtain complementary fragmentation information. Ion mobility separation of fragments was performed to reduce MSMS complexity. This led to a better signal-to-noise ratio while avoiding overlapping of isotopic profiles of fragment ions (cf. Figure 3). The DriftScope software was used to visualize and reprocess TWIM data.

Semiautomated Analysis with Prosight PTM and Manual Scrutiny of MSMS Data. MSMS spectra were examined with a combination of manual and automatic procedures. Monoisotopic mass lists were obtained using the Maxent 3 deconvolution software (MassLynx 4.1, Waters Corp., Manchester, U.K.). The Prosight PTM (https://prosightptm.scs.uiuc.edu) software suite was used with a threshold of 35 ppm and the deltamass feature deactivated, with custom post-translational modifications as required. Interpretation was a manual, iterative process, as different sequences and post-translational modifications were independently tested to maximize the number of product ions matched. Nomenclature for assignment of peptide/protein ions was according to Roepstorff and Fohlman.[36,37] We also used a manual Pscore that similarly reflects the confidence in data interpretation, but which relies on masses of product ions that best matched the sequence and is updated with product ions manually identified in the MSMS spectra. Unique ions were then searched for unambiguous assessment of protein sequences and modifications proposed using 25 ppm error corresponding to the rounded highest rms value calculated. The extracted peak lists for each experiment are provided in the Supporting Information. Systematic theoretical isotopic profile superimposition was used to reveal hidden peaks or confirm the presence of modifications in fragment ions using the Isotope Model tool in the MassLynx software.

RESULTS

Expression, Purification, and Biophysical Characterization of XD Proteins. After having checked that the nucleotide sequence of the ORFs of both native and K496C XD constructs did conform to expectations[2] (see the Supporting Information, Figures S1 and S2), the two plasmid constructs were transformed in the Rosetta pLysS strain of *E. coli*. Figure 1A shows the expected amino acid sequence of XD, as well as the ribbon representation of its crystal structure. The XD proteins were purified to homogeneity (>95%) (Figure 1B). Contrary to wt XD that is eluted from the gel filtration column as a monomer, the elution profile of the K496C XD variant points out the presence of an additional peak corresponding to a dimeric form. Such a dimeric form can also be observed as a minor band in SDS−PAGE analyses (Figure 1B). Addition of a 100-fold molar excess of DTT prior to SDS−PAGE or gel filtration led to a single band at approximately 6 kDa and resulted in an elution profile quite comparable to that observed with the wt protein (data not shown). Consistent with this similar elution profile, the Stokes radius (R_S) of the monomeric form of the K496C XD variant, as inferred from gel filtration, is close to that observed for wt XD (13 ± 2 Å)[2] and consistent with the theoretical value expected for a folded protein form.[35] Thus, wt and K496C XD share similar hydrodynamic properties, both adopting a folded conformation. In further support of this observation, the far-UV CD spectrum of K496C XD at neutral pH quite well superimposes onto that of wt XD and is typical of a structured protein with a predominant α-helical content, as seen by the positive ellipticity between 185 and 200 nm and by the two minima at 208 and 222 nm (Figure 1B). These data indicate that the cysteine substitution induces few, if any, structural perturbations.

ESI-MS Mass Measurements of Full-Length XD Proteins. The top-down approach first requires the experimental determination of the molecular masses of the proteins under study prior to recording top-down MSMS data. To that endeavor, ESI-MS spectra of wt XD and of the K496C XD variant, either reduced and alkylated with IAA or spin-labeled with MTSL, were recorded (see Table 1). The obtained results showed that while

the wt XD sample was monomeric, the K496C XD variant, either labeled or not by the MTSL, were found to consist in a mixture of monomer and dimer. In agreement with the biochemical data described above (i.e., SDS−PAGE and gel filtration), the disappearance of the dimer for both K496C and K496C* XD variants after DTT treatment demonstrated that dimers arose from association of two unlabeled monomers of K496C XD through a disulfide bridge. Data also revealed that all samples were mainly composed of monomers with averaged molecular masses consistent with the removal of the initial methionine. Accordingly, the alanine residue following the initial methionine is taken as position number 1 throughout the whole paper. The observed mass difference (186 Da) between K496C XD and K496C* XD monomers is in agreement with the expected mass increment resulting from chemical labeling of the unique cysteine residue by MTSL. However, mass discrepancies observed between monoisotopic and average masses (see Table 1) suggested that all XD samples were probably a mixture of unexpected XD variants. Indeed, if the XD protein samples were homogeneous, the differences between experimental and theoretical monoisotopic masses would be the same (or fall in the same range) as those observed between experimental and theoretical averaged masses (cf. Figure S8 in the Supporting Information for an illustrative example). The presence of a mixture was further confirmed by the comparison between the experimental and the theoretical isotopic profiles of protein CSs. Parts A and B of Figure 2 show such a comparison for the +8 CS of the wt XD and +7 CS of K496C* XD proteins, respectively.

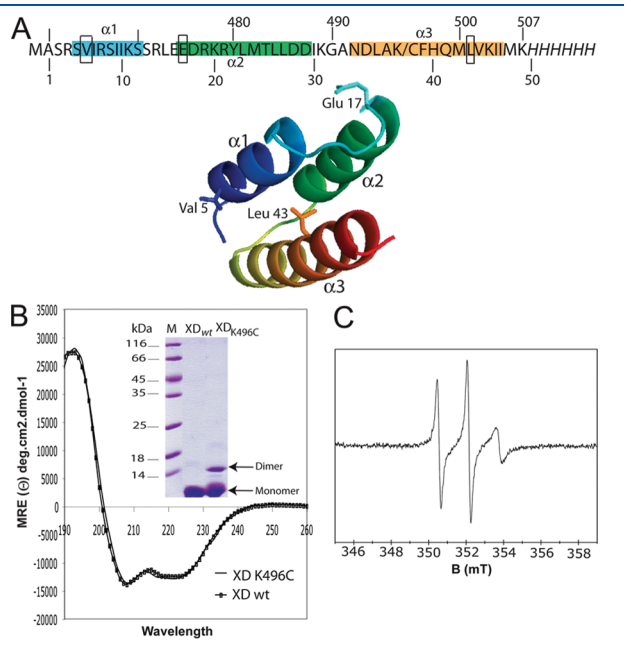


Figure 1. (A) Amino acid sequence and ribbon representation of the crystal structure of MeV XD (PDB code 1OKS) as obtained using Pymol.[43] Position 496 is either a K or a C depending on whether the native or the K496C form is considered. Numbers above the sequence indicate amino acid positions within the full-length P protein. Numbers below the sequence correspond to amino acid positions within XD. Amino acids within XD are numbered taking the alanine residue following the initial methionine as position number 1 to reflect the fact that the first methionine is cleaved off. The α-helices, as seen in both the crystal and solution structures of XD (PDB codes 1OKS and 2K9D, respectively) are highlighted in color. The positions found to be substituted in the purified protein samples of either native or K496C MeV XD are framed. The side chains of these latter residues are also shown in sticks in the ribbon representation of the structure of XD. (B) Far-UV CD spectra of wt and K496C XD, each at 0.1 mg/mL in 10 mM sodium phosphate buffer, pH 7, at 20 °C. The inset shows a Coomassie blue staining of an 18% SDS−PAGE loaded with approximately 5 μg of wt and K496C XD proteins purified from the soluble fraction of *E. coli*. M = molecular mass markers. The dimeric and monomeric forms of XD are highlighted by arrows. The faint bands of higher molecular mass correspond to minor contaminants. (C) EPR spectrum of MeV K496C* XD recorded at 296 K (±1 K): microwave power 8 mW, amplitude modulation 0.1 mT, sweep width 15 mT, number of accumulations 8.

ESI-MS Mass Measurements of Full-Length XD Proteins Using TWIM. To gain further insights into the nature and origin of the observed protein heterogeneities, the ESI mass spectra of full-length XD samples were recorded using the TWIM mode. If the use of the TWIM mode allowed "purification" of MS spectra with observation of cleaner CSD, we noticed changes in both the CSD of XD samples and the isotopic profiles of higher CS ions. These results were ascribed to an artifactual saturation of the detector in TWIM mode. In that case abnormal peak ratios of CSD and isotopic profiles arose from over- and underdetection of isotopic profiles of XD wt variants (see the Supporting Information, Figure S3). Indeed, the dynamic range was shown to be smaller under the TWIM experiment due to the fact that ions arrived in packets, leading to faster detector saturation. Consequently, we paid a lot of attention so as to avoid saturation of the ToF detector when ESI-TWIM-MS and ESI-TWIM-MSMS experiments were performed. ESI-TWIM-MS experiments did not show resolution of new precursors of XD variants except for one XD protein variant that was found to be more retained in the ion mobility cell. In that case, no signal saturation was observed, as exemplified by the isotopic profile ratios of wt XD and XD variants (see the Supporting Information, Figure S4). These data further corroborate the presence of some unexpected heterogeneity.

Top-Down ESI-TWIM-MSMS on Full-Length XD Proteins. As a further step in the investigation of XD samples, top-down MSMS spectra were recorded under TWIM mode. Collision of XD CS was optimized in terms of fragmentation yield and on several CSs to obtain complementary fragment ions. TWIM parameters were adjusted to achieve the best possible signal-to-noise ratios and limitation of isotopic profile overlap using the

Table 1. Experimental Monoisotopic and Averaged Masses (Da) of XD Monomers[a]

	experimental mass		theoretical mass		mass error	
sample	monoisotopic mass	av mass	monoisotopic mass	av mass	monoisotopic	av
XD wt	6550.89	6553.60	6551.54	6555.76	−0.65000	−2.16
XD K496C+ IAA	6584.33	6587.76	6583.47	6587.78	0.86300	−0.02
XD K496C+ MTSL	6713.44	6715.06	6711.54	6715.01	1.90000	0.05

[a] Calculated errors between experimental and theoretical monoisotopic and averaged masses are given in daltons and parts per million in the last two columns. For MTSL, monoisotopic mass +185.09, averaged mass +186.3 Da. For IAA, monoisotopic mass +57.02 Da, averaged mass +57.05 Da.

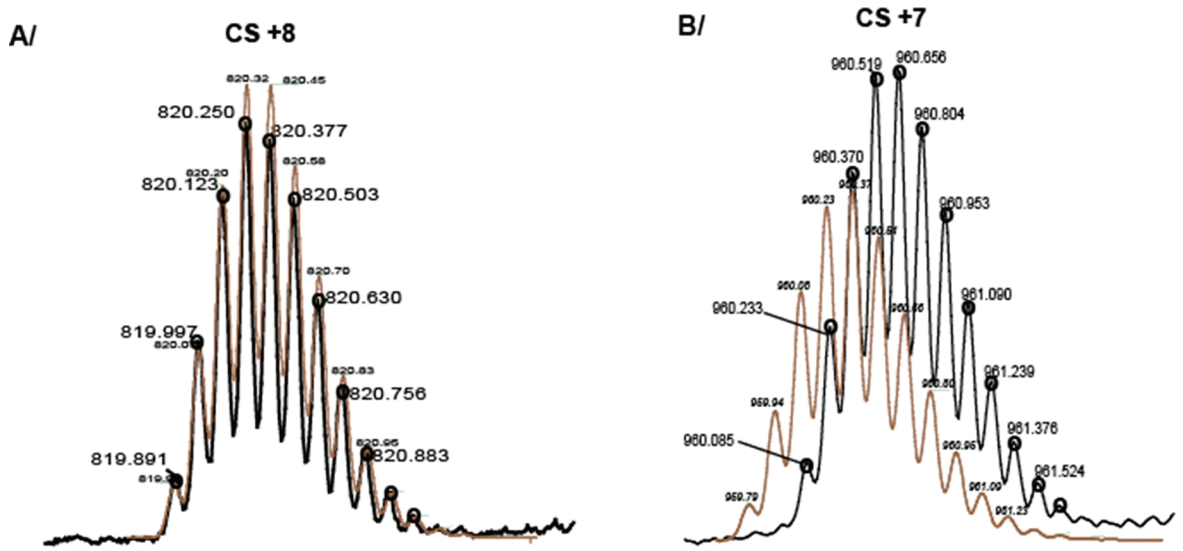


Figure 2. Comparison of experimental (black) and theoretical (gray) isotopic profiles of the (A) +8 CS of wt XD protein and (B) +7 CS of K496C* XD monomer.

ability of the TWIM guide to segregate MSMS fragments according to their sizes, CCSs, and CSs (see Figure 3). The propensity of TWIM to separate fragment ions was our voucher to achieve reliable top-down MSMS data interpretation. MSMS data sets were analyzed in detail using both semiautomated and manual approaches (see the Materials and Methods). Combined MSMS data sets recorded on several CSs on both wt and K496C XD samples confirmed the expected sequence and expected lysine to cysteine substitution at position 38 (position 496 within the P protein) in the K496C XD variant, as well as the presence of the EPR probe grafted at position 38 in the spin-labeled K496C* XD variant. Methodical analyses of MSMS spectra and isotopic profiles of fragment ions revealed the presence of subtle amino acid changes. In particular, the wt XD protein was found to be a mixture of at least two species, one corresponding to the expected sequence and one corresponding to a species bearing a valine to proline substitution at position 5 (−2 Da). This latter result is in agreement with the observation of the monoisotopic ion of the V5P variant on ion mobility chart and with calculation of a lower CCS (see the Supporting Information, Figure S4). However, experimental mass differences observed for wt XD between averaged (−2.1 Da) and monoisotopic (−0.61 Da) molecular weights still suggest the occurrence of at least another modification counterbalancing the effect of the above-mentioned substitution. Similar scrutiny of MSMS spectra allowed us to infer the presence of two additional sequence changes, namely, replacement of Glu17 by a Gln or a Lys (−0.984 Da) and of Leu43 by an Asn (0.958 Da). A summary of our findings for wt XD with sequence charts, probability scores, and rms values for the native XD regular sequence and for wt XD variants (V5P, E17K/Q, and L43N variants) is provided in Figure 5. The superimposition of theoretical and experimental isotopic profiles of the b29 ion for wt XD and for the wt XD V5P, E17K/Q, and L43N variants are shown in the Supporting Information, Figure S5. The experimental masses of the b and y ions, the mass errors calculated in daltons and parts per million, as well as the identification of unique fragment ions providing further support to our interpretation of the observed heterogeneities are given in Tables S1−S3, Supporting Information.

Likewise, this approach allowed confirmation of the expected sequence of K496C IAA XD and also demonstrated the presence of another form bearing a leucine to asparagine substitution at position 43 (Figure 4). A summary of our findings for XD K496C IAA variant with sequence charts, probability scores, and rms values is given in the Supporting Information, Figure S6 and Table S4. Finally, analysis of top-down MSMS data recorded for K496C* XD revealed that the latter bears the same modification as K496C IAA XD at position 43. However, the measured mass increment for K496C* XD indicates that the latter is still 1 Da heavier. Comparison of experimental isotopic profiles of b and y ions with their calculated counterparts proved that the modification was brought by the cysteine labeled with MTSL (see the Supporting Information, Figure S7). This suggests that the MTSL may have captured a proton under our experimental conditions. To assess this hypothesis, the EPR and MS spectra of MTSL were recorded in water/methanol and water/methanol/formic acid solutions. The EPR spectrum of MTSL recorded under the latter conditions (data not shown) revealed that the MTSL neither was protonated on the nitroxide nor retained any proton in the surroundings of the nitroxide group, since the EPR spectrum was superimposable onto that recorded in sodium phosphate buffer. In contrast, ESI-MS spectra of MTSL revealed the "protonation" of the latter with the detection of an ion at m/z 266.08 which was 1 Da heavier than the $[M + H]^+$ ion (m/z 265.08). This result led us to conclude that a protonation of MTSL occurred in the gas phase, accounting for an artifactual mass increase of 1 Da. The ESI-MS spectrum of MTSL recorded in water/methanol solution containing 1% formic acid and the proposed chemical structures for the MTSL masses detected are shown in Figure 5. Further support to this conclusion can be found in the Supporting Information, Table S5, which shows that the K496C* XD variant mass was still 2 Da heavier than expected when MTSL was the only added modification.

DISCUSSION

In this paper we evaluated the power of the TWIM guide for reducing top-down MSMS spectrum complexity and for elucidating the origin of subtle sample heterogeneities. Indeed, it is

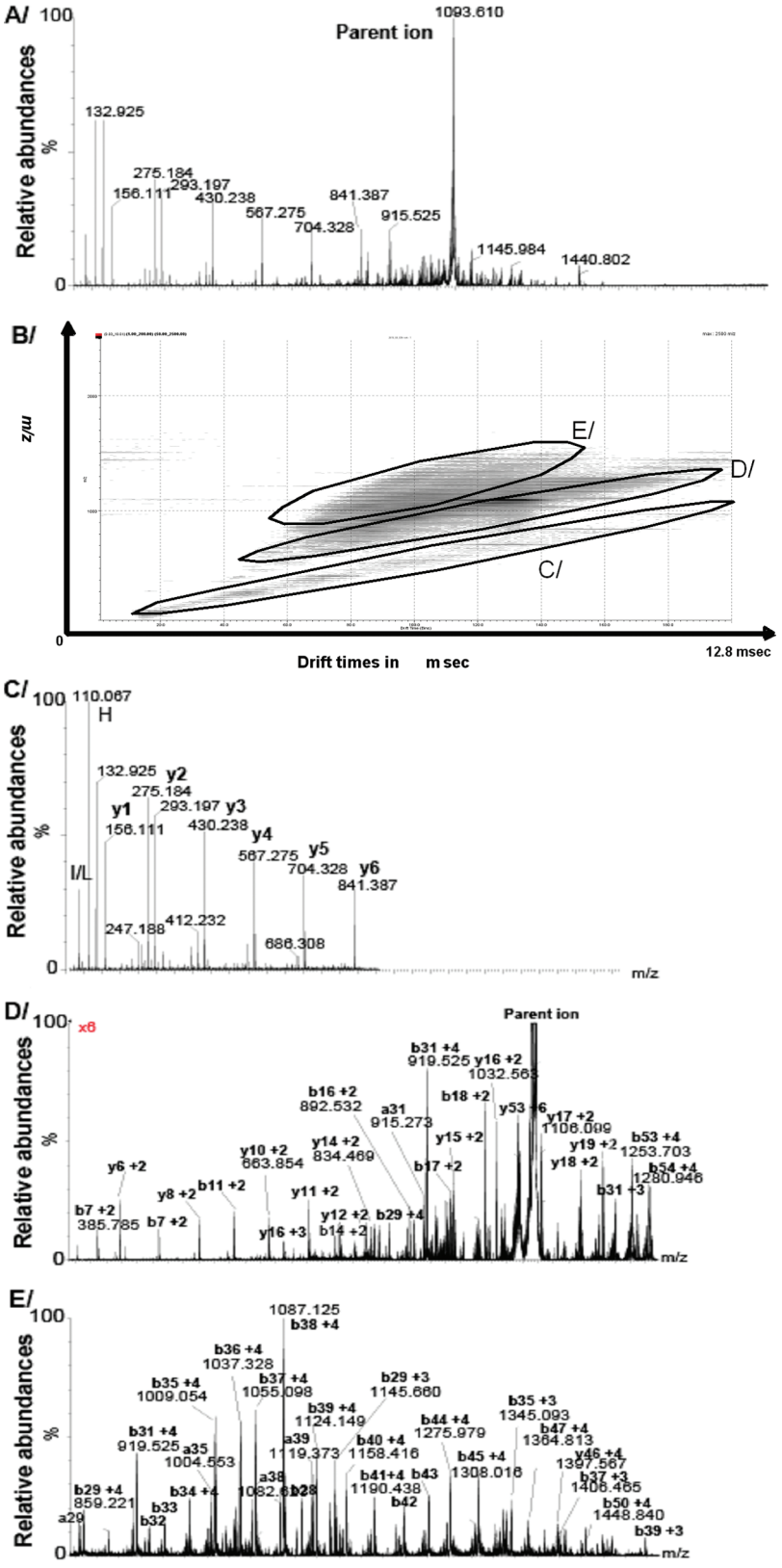


Figure 3. Illustration of the segregation of fragment ions achieved using TWIM and leading to a better signal-to-noise ratio, while avoiding overlapping of isotopic profiles of fragments ions. (A) Full top-down ESI-MSMS spectrum of wt XD. (B) Ion mobility chart showing separation of fragment ions according to their sizes, CSs, and CCSs. (C−E) Top-down ESI-MSMS spectra of wt XD displaying singly, doubly and triply, and triply and quadruply charged ions separated using TWIM, respectively.

dx.doi.org/10.1021/ac200994c |*Anal. Chem.* 2011, 83, 7306–7315

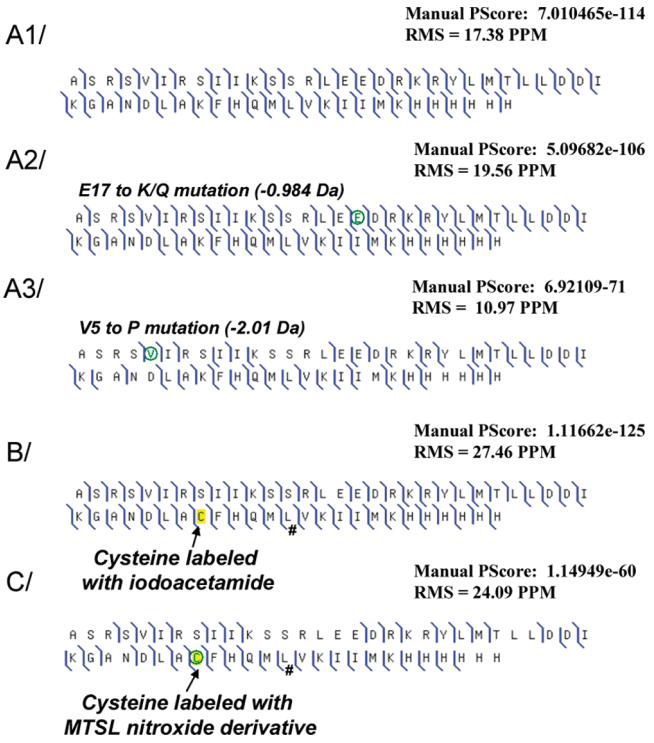


Figure 4. Sequence charts showing fragment ions matching the sequences of native and XD variants. Manual Pscores and rms values are given for each experiment. (A1) Regular sequence of wt XD. (A2) Sequence of wt XD containing the E17 to K/Q substitution. (A3) Sequence of wt XD containing the V5 to P substitution. (B) Sequence of XD K496C IAA. (C) Sequence of XD K496C*.

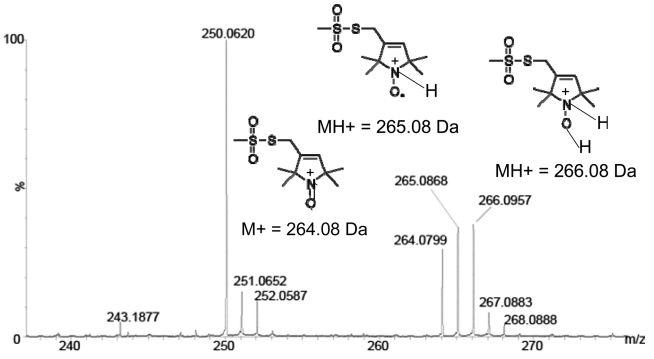


Figure 5. ESI mass spectrum of MTSL recorded in H_2O/MeOH/formic acid (50/50/1) showing both oxidation and protonation of the nitroxide derivative group. The proposed chemical structures for the three peaks at m/z values of 264.07, 265.08, and 266.09 Da are also displayed.

critical to accurately identify fragment ions for correct assignment of the sequence and chemical properties. Dubious peak assignment can arise from overlapping of isotopic profiles of regular fragment ions, from fragments resulting from multiple collisional events or from chemical losses such as water and ammonia. Another feature that can preclude precise fragment ion identification is related to desorption/ionization competitions of species produced in different yields, leading to poor statistics on some ions. For all these reasons, TWIM can be viewed as an alternative to fully or partially overcome the limitations mentioned above. The model protein used in this study is the C-terminal X domain of the measles viral phosphoprotein produced in *E. coli*.

The first requirement in a top-down MSMS experiment is the measurement of an intact mass tag. Indeed, determination of monoisotopic and average experimental masses is critical since primary structure data revealed by MSMS must correlate with the latter. In our case, these measurements revealed discrepancies in mass increments between experimental and theoretical monoisotopic and average molecular masses that argued for the presence of heterogeneities in XD samples (see Table 1). If the latter were homogeneous, then the observed mass errors for monoisotopic and averaged masses would be in the same range (see Table in Figure S8, Supporting Information).

To thoroughly document the structural properties of XD proteins, top-down MSMS spectra were then recorded under TWIM mode. The use of ion mobility separation allowed achievement of an efficient separation of fragment ions according to their drift times related to their respective sizes, CSs, and CCSs (Figure 3). Ion separation allowed large b and y series ions to be distinguished with better signal-to-noise ratios and better statistics compared to those of the full MSMS spectrum. A similar achievement could only have been accomplished with an FTICR−MS instrument. These data sets allowed achievement of complete sequence coverage of XD proteins. In particular, TWIM turned out to be a highly efficient approach for the detection of XD variants, through methodical comparison of theoretical and experimental isotopic profiles of b and y fragment ions. This approach proved to be sensitive enough to reveal the presence of other XD forms and to locate regions and amino acid residues bearing subtle modifications. Indeed, except for the V5P XD variant, where a monoisotopic ion precursor was resolved from other variants in the TWIM cell, the isotopic profiles of other forms (E17Q/K, L43N, MTSL protonation) were superimposed to "regular" ones due to the limited resolution of the TWIM guide. Such superimposition of isotopic profiles explains the increase of rms values of fragment ion masses observed for the variants bearing the E17Q/K and L43N substitutions, as well as MTSL protonation. Finally, all the proposed amino acid substitutions were confirmed by the search for unique fragment ions (cf. masses labeled in bold in Tables S1−S5, Supporting Information). This workflow allowed us to demonstrate that the native XD protein did in fact consist in a mixture of protein species comprising the regular form, a variant bearing a V5P substitution, a variant bearing the E17 to Q/K substitution, and a variant with the L43 to N substitution. Similarly, both K496C samples, either labeled with MTSL or not, were found to be represented by a form with the expected sequence and by a variant bearing the L43 to N substitution. All these variants were found to be minor species on the basis of isotopic profile modeling (see the Supporting Information, Figures S4−S7).

The presence of dimers in XD K496C samples, either labeled or not with MTSL, was unambiguously ascribed to the occurrence of a disulfide bridge formation between the free cysteine residues of the monomers. This observation implies that spin labeling of the XD K496C is partial, probably as a result of incomplete DTT reduction, whereas the totality of the XD K496C monomer was found to bear the MTSL nitroxide group. Finally, the N-terminal methionine was found to be cleaved in all XD proteins, in agreement with previous reports showing that the N-terminal methionine of proteins expressed in *E. coli* is cleaved off if followed by a residue with a small side chain, such as glycine, alanine, serine, and proline.[38]

To check whether the MTSL group could be oxidized and/or protonated under the experimental conditions used in the ESI source, the mass spectrum of the MTSL group alone was recorded

under the same conditions (water/methanol/formic acid solution, 50/50/1, v/v/v) used in the ESI experiments. The mass spectrum revealed that the MTSL group is represented by three peaks of 264.07, 265.08, and 266.09 Da (Figure 5). The structures proposed for these forms are displayed in Figure 5. From this spectrum it appears that the MTSL group could undergo oxidation due to the presence of free electrons in the ESI source, as well as protonations due to the acidic condition used. This result allowed us to conclude that not only one but two protonations could occur in the MTSL group, thus accounting for the 1 and/or 2 Da mass increase in the full-length K496C* variant (see the Supporting Information, Figure S7, for further support of the protonation of the MTSL group on fragment b39 of the XD K496C* variant). The experimental evidence arguing for the artifactual character of such a protonation points out that we should be extremely cautious during MS and MSMS data interpretations.

To gain further insight into the structural changes related to the amino acid substitutions observed in the native XD sample, we attempted to calculate the CCSs of wt XD and of the wt XD V5P variant for +9 CS since TWIM allowed resolving the segregation of their corresponding monoisotopic peaks. Calculation of CCS values for other substitutions (e.g., E17 to Q/K and L43 to N) were impossible since their isotopic profiles were merged together with the isotopic profile of native XD. In addition, the structural changes related to other proposed amino acid substitutions were thought to be too subtle or to have no effect on the global conformation to lead to any distinguishable features in the ion mobility mode. Calibration of CCSs to drift times was obtained using a tryptic digest of enolase. Rigorous identical instrumental conditions were used to ensure data consistence. Calibration of CCSs was successfully achieved after peak detection with DriftScope, and calculation of CCSs was made using the CCS calculator extension of this software. Calibration was saved and applied to TWIM data of native XD. CCS values determined for native XD and for the XD V5P variant +9 CS were 1371.5 and 1491.0 $Å^2$, respectively (see the Supporting Information, Figure S4). Notably, the XD V5P variant was found to have a lower CCS value (minus 213 $Å^2$) than native XD, a result that could be explained by a more compact conformation of the former species. Structure predictions, as provided by the MeDor metaserver,[39] indicate that the V5P substitution is predicted to break helix α1. Since we estimated that the CCSs of the two species differed by ∼200 $Å^2$ (corresponding to approximately a 15% difference), we assumed that the breaking of the N-terminal α-helix could induce a collapse of the core of the protein, leading to a more compact structure.

Finally, since standard nucleotide sequencing of plasmid constructs encoding native or K496C XD did not reveal any misinsertion nor any nucleotide substitutions, the observed microheterogeneity of both XD proteins is thought to arise either from extremely rare nucleotide substitutions (below the threshold of detection of standard sequencing) or from errors in DNA replication occurring during the bacterial expression of the constructs. Errors in transcription and translation (either alone or in combination) are unlikely to be the source of the observed heterogeneity, since such occasional events would lead to substitutions that would expectedly escape detection. Strikingly, the L43N substitution was found not only in the wt XD sample but also in the K496C variant sample. The fact that the L43N substitution has been found in various samples argues for a nonrandom event, implying either an inherent propensity to the substitution or a selective evolutionary advantage in terms of stability of the protein. Notably, changing the CTG leucine-encoding codon occurring at position 43 into an asparagine-encoding codon (encoded by either an AAT or an AAC codon) requires three nucleotide substitutions, an event that does not occur with a high frequency in the absence of a selective pressure. Misincorporation of an Asn during translation as a result of a three-mismatch codon−anticodon pairing is unlikely either. Interestingly, since the CTG codon is used with a higher frequency (5.2%) in *E. coli* than the AAT (1.6%) and AAC (2.6%) codons,[40] the occurrence of the L43N substitution in both native and K496C XD likely reflects an evolutionary advantage possibly related to a higher stability of the L43N variant. However, analysis of the 3D structure of native XD (PDB code 1OKS), in which the side chain of Leu43 is fully buried and not accessible to the solvent (see Figure 1A), failed to provide any hints about the supposed increased stability of the L43N variant. In the same vein, analysis of the predicted secondary structure content does not point out an increased stabilization of the α-helices in the L43N variant nor an increased compactness. Indeed, neither the L43N nor the E17Q/K substitution is predicted to affect the secondary structure content of the protein, contrary to the V5P substitution, as judged using the secondary structure predictor implemented in the MeDor metaserver[39] (data not shown).

As for the V5P and E17Q/K substitutions, at least two or one nucleotide substitution, respectively, is required to occur in the respective parental codons. It should be pointed out that the *E. coli* strains that are used for protein expression, including the Rosetta pLysS used in this study, possess a functional *recA* gene, which means that they can undergo recombination[41] and are also more sensitive to UV mutagenesis[42] than *recA*$^-$ strains, i.e., strains bearing a mutated, nonfunctional *recA* gene. As such, these strains, while being optimized for the expression of recombinant proteins, are not recommended for selection and amplification of plasmids as they can undergo undesired recombination events and/or mutations. Therefore, it is conceivable that the protein microheterogeneities observed upon expression of XD proteins in the Rosetta pLysS strain may result from this kind of event.

On the other hand, we underscore that such protein microheterogeneity was not systematically found in all XD samples. In fact, two other XD samples (i.e., wt and K496C XD), which were purified from two independent bacterial cultures (referred to as batch 2), were shown to be devoid of any unexpected substitution (see Table and sequence charts in Figure S8, Supporting Information). Note that the bacterial cultures of batch 2 were grown under the same conditions as those of batch 1, i.e., of those cultures that led to the microheterogeneous wt and K496C XD samples. Indeed, although we systematically searched for any modification in all XD samples, the monoisotopic and averaged experimental masses of batch 2 XD samples were found to be in agreement with the expected ones and within the same error range. In addition, neither the V5P, E17K/Q, and L43N substitutions observed in batch 1 XD samples nor any other unexpected modifications were found, and the amino acid sequence was found to perfectly match the expected one.

Whatever the underlying mechanism, the finding that the expression in *E. coli* can occasionally result in protein microheterogeneities, presumably due to genetic drift events (neutral or not), has important implications in terms of both biochemical and structural characterization of proteins. Indeed, such microheterogeneities may impair crystallization of pure and monodisperse protein samples and may lead to an unexpected behavior such as altered interaction ability toward a partner or ligand. As such, it is advised to keep in mind the possibility of protein microheterogeneity when protein samples do not behave as expected and to submit

them to extensive analyses such as the ones described in the present study.

CONCLUSION

In this study, we showed that TWIM can advantageously be used to unveil subtle changes in proteins, whatever the biological mechanism underlying the generation of these amino acid substitutions. Experimental evidence for such a microheterogeneity was achieved by a reliable identification of fragment ions and by methodical comparison of isotopic profiles of fragment ions. In particular, TWIM allowed calculation of differential CCSs for two XD variant forms differing by only ~15%. TWIM was shown here to be able to document minor changes in the primary structure of XD and, to some extent, in its tertiary structure too. It appears that ion mobility is a valuable alternative approach to highly resolved top-down tandem experiments: its unique ability of distinguishing different drift times and different isotopic profiles renders it extremely powerful in unveiling minor amino acid substitutions in protein samples.

ASSOCIATED CONTENT

Supporting Information. Additional information as noted in text. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Phone: (33) 4 91 16 46 73. Fax: (33) 4 91 16 40 97. E-mail: fhalgand@ifr88.cnrs-mrs.fr.

ACKNOWLEDGMENT

We grateful to the proteomics core facility of the Institut de Microbiologie de la Méditerranée for their technical assistance. S.L. thanks Aissette Baanannou and Stéphanie Costanzo for their contribution to cloning and purification procedures. We warmly thank Denis Gerlier for helpful scientific contribution. This work was carried out with the financial support of the Agence Nationale de la Recherche, program "Blanc", ANR-09-BLAN-0100 (B.G.), and with the partial support of the ANR specific program "Physico-Chimie du Vivant", ANR-08-PCVI-0020-01 (S.L.), and of the CNRS. It was also partly supported by the National Institute of Neurological Disorders and Stroke (Grant R01 NS031693-11A2). We are also grateful to the Conseil Regional of Provence Alpes Cotes d'Azur, to the city of Marseille, and to the University of Provence for financial support in the acquisition of instrumentation. Sylvain Marque is acknowledged for its scientific support.

ABBREVIATIONS

CCS	collisional cross section
CD	circular dichroism
CS	charge state
CSD	charge state distribution
dc	direct current
EPR	electron paramagnetic resonance
ESI-MS	electrospray ionization mass spectrometry
HEX	hydrogen exchange
IAA	iodoacetamide alkylation
IMT	intact mass tag
MSMS	tandem mass spectrometry
MTSL	1-oxyl-2,2,5,5-tetramethyl-δ3-pyrroline-3-methyl methanethiosulfonate
MV	measles virus
N	measles virus nucleoprotein
N_{TAIL}	C-terminal region of measles virus nucleoprotein spanning residues 401−525
ORF	open reading frame
P	phosphoprotein
QTof	quadrupole time-of-flight (mass spectrometer)
rf	radio frequency
R_S	Stokes radius
SDS−PAGE	sodium dodecyl sulfate−polyacrylamide gel electrophoresis
TWIM	traveling wave ion mobility
XD	X domain of measles virus phosphoprotein spanning residues 459−507
XD K496C IAA	XD K496C variant reduced and alkylated with iodoacetamide
XD K496C*	XD K496C variant labeled with the MTSL nitroxide derivative
wt	wild type

REFERENCES

(1) Jana, S.; Deb, J. K. *Appl. Microbiol. Biotechnol.* **2005**, *67*, 289–298.
(2) Kelleher, N. L.; Lin, H. Y.; Valaskovic, G. A.; Aaserud, D. J.; Fridriksson, E. K.; FW., M. *J. Am. Chem. Soc.* **1999**, *121*, 806–807.
(3) Pesavento, J. J.; Kim, Y. B.; Taylor, G. K.; Kelleher, N. L. *J. Am. Chem. Soc.* **2004**, *126*, 3386–3387.
(4) Collier, T. S.; Sarkar, P.; Rao, B.; Muddiman, D. C. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 879–889.
(5) Wu, S.; Lourette, N. M.; Tolic, N.; Zhao, R.; Robinson, E. W.; Tolmachev, A. V.; Smith, R. D.; Pasa-Tolic, L. *J. Proteome Res.* **2009**, *8*, 1347–1357.
(6) Ouvry-Patat, S. A.; Torres, M. P.; Gelfand, C. A.; Quek, H. H.; Easterling, M.; Speir, J. P.; Borchers, C. H. *Methods Mol. Biol.* **2009**, *492*, 215–231.
(7) Lee, J. E.; Kellie, J. F.; Tran, J. C.; Tipton, J. D.; Catherman, A. D.; Thomas, H. M.; Ahlf, D. R.; Durbin, K. R.; Vellaichamy, A.; Ntai, I.; Marshall, A. G.; Kelleher, N. L. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 2183–2191.
(8) Halgand, F.; Zabrouskov, V.; Bassilian, S.; Souda, P.; Wong, D. T.; Loo, J. A.; Faull, K. F.; Whitelegge, J. P. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 868–877.
(9) Ryan, C. M.; Souda, P.; Halgand, F.; Wong, D. T.; Loo, J. A.; Faull, K. F.; Whitelegge, J. P. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 908–917.
(10) Giles, K.; Pringle, S. D.; Worthington, K. R.; Little, D.; Wildgoose, J. L.; Bateman, R. H. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 2401–2414.
(11) Smith, D. P.; Knapman, T. W.; Campuzano, I.; Malham, R. W.; Berryman, J. T.; Radford, S. E.; Ashcroft, A. E. *Eur. J. Mass Spectrom. (Chichester, Eng.)* **2009**, *15*, 113–130.
(12) Williams, J. P.; Phillips, H. I.; Campuzano, I.; Sadler, P. J. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 1097–1106.
(13) Hopper, J. T.; Oldham, N. J. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 1851–1858.
(14) Rand, K. D.; Pringle, S. D.; Murphy, J. P., III; Fadgen, K. E.; Brown, J.; Engen, J. R. *Anal. Chem.* **2009**, *81*, 10019–10028.
(15) Abzalimov, R. R.; Kaplan, D. A.; Easterling, M. L.; Kaltashov, I. A. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 1514–1517.
(16) Atmanene, C.; Wagner-Rousset, E.; Malissard, M.; Chol, B.; Robert, A.; Corvaia, N.; Van Dorsselaer, A.; Beck, A.; Sanglier-Cianferani, S. *Anal. Chem.* **2009**, *81*, 6364–6373.
(17) Hilton, G. R.; Thalassinos, K.; Grabenauer, M.; Sanghera, N.; Slade, S. E.; Wyttenbach, T.; Robinson, P. J.; Pinheiro, T. J.; Bowers, M. T.; Scrivens, J. H. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 845–854.
(18) Fenn, L. S.; McLean, J. A. *Mol. Biosyst.* **2009**, *5*, 1298–1302.

(19) Djidja, M. C.; Francese, S.; Loadman, P. M.; Sutton, C. W.; Scriven, P.; Claude, E.; Snel, M. F.; Franck, J.; Salzet, M.; Clench, M. R. *Proteomics* **2009**, *9*, 2750–2763.
(20) Johansson, K.; Bourhis, J. M.; Campanacci, V.; Cambillau, C.; Canard, B.; Longhi, S. *J. Biol. Chem.* **2003**, *278*, 44567–44573.
(21) Bernard, C.; Gely, S.; Bourhis, J. M.; Morelli, X.; Longhi, S.; Darbon, H. *FEBS Lett.* **2009**, *583*, 1084–1089.
(22) Longhi, S.; Receveur-Brechot, V.; Karlin, D.; Johansson, K.; Darbon, H.; Bhella, D.; Yeo, R.; Finet, S.; Canard, B. *J. Biol. Chem.* **2003**, *278*, 18638–18648.
(23) Bourhis, J.; Johansson, K.; Receveur-Bréchot, V.; Oldfield, C. J.; Dunker, A. K.; Canard, B.; Longhi, S. *Virus Res.* **2004**, *99*, 157–167.
(24) Bourhis, J. M.; Receveur-Brechot, V.; Oglesbee, M.; Zhang, X.; Buccellato, M.; Darbon, H.; Canard, B.; Finet, S.; Longhi, S. *Protein Sci.* **2005**, *14*, 1975–1992.
(25) Morin, B.; Bourhis, J. M.; Belle, V.; Woudstra, M.; Carrière, F.; Guigliarelli, B.; Fournel, A.; Longhi, S. *J. Phys. Chem. B* **2006**, *110*, 20596–20608.
(26) Belle, V.; Rouger, S.; Costanzo, S.; Liquiere, E.; Strancar, J.; Guigliarelli, B.; Fournel, A.; Longhi, S. *Proteins* **2008**, *73*, 973–988.
(27) Gely, S.; Lowry, D. F.; Bernard, C.; Ringkjobing-Jensen, M.; Blackledge, M.; Costanzo, S.; Darbon, H.; Daughdrill, G. W.; Longhi, S. *J. Mol. Recognit.* **2010**, *23*, 435–447.
(28) Kavalenka, A.; Urbancic, I.; Belle, V.; Rouger, S.; Costanzo, S.; Kure, S.; Fournel, A.; Longhi, S.; Guigliarelli, B.; Strancar, J. *Biophys. J.* **2010**, *98*, 1055–1064.
(29) Bischak, C. G.; Longhi, S.; Snead, D. M.; Costanzo, S.; Terrer, E.; Londergan, C. H. *Biophys. J.* **2010**, *99*, 1676–1683.
(30) Jensen, M. R.; Communie, G.; Ribeiro, E. A., Jr.; Martinez, N.; Desfosses, A.; Salmon, L.; Mollica, L.; Gabel, F.; Jamin, M.; Longhi, S.; Ruigrok, R. W.; Blackledge, M. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 9839–9844.
(31) Bourhis, J. M.; Canard, B.; Longhi, S. *Virology* **2006**, *344*, 94–110.
(32) Bourhis, J. M.; Canard, B.; Longhi, S. *Curr. Protein Pept. Sci.* **2007**, *8*, 135–149.
(33) Longhi, S. *Measles Virus Nucleoprotein*; Nova Publishers Inc.: Hauppage, NY, 2007.
(34) Longhi, S.; Oglesbee, M. *Protein Pept. Lett.* **2010**, *17*, 961–978.
(35) Uversky, V. N. *Biochemistry* **1993**, *32*, 13288–13298.
(36) Roepstorff, P.; Fohlman, J. *Biomed. Mass Spectrom.* **1984**, *11*, 601.
(37) LeDuc, R. D.; Taylor, G. K.; Kim, Y. B.; Januszyk, T. E.; Bynum, L. H.; Sola, J. V.; Garavelli, J. S.; Kelleher, N. L. *Nucleic Acids Res.* **2004**, *32*, W340–345.
(38) Frottin, F.; Martinez, A.; Peynot, P.; Mitra, S.; Holz, R. C.; Giglione, C.; Meinnel, T. *Mol. Cell. Proteomics* **2006**, *5*, 2336–2349.
(39) Lieutaud, P.; Canard, B.; Longhi, S. *BMC Genomics* **2008**, *9*, S25.
(40) Maloy, S.; Stewart,V. Taylor., R. *Genetic Analysis of Pathogenic Bacteria*; Cold Spring Harbor Laboratory Press: Plainview, NY, 1996.
(41) Voloshin, O. N.; Wang, L.; Camerini-Otero, R. D. *Science* **1996**, *272*, 868–872.
(42) Dutreix, M.; Moreau, P. L.; Bailone, A.; Galibert, F.; Battista, J. R.; Walker, G. C.; Devoret, R. *J. Bacteriol.* **1989**, *171*, 2415–2423.
(43) DeLano, W. L. *Proteins: Struct., Funct., Bioinf.* **2002**, *30*, 442–454.