Published in final edited form as:

Anal Chem. 2011 October 1; 83(19): 7306–7315. doi:10.1021/ac200994c.

Dividing to unveil protein microheterogeneities: A Traveling Wave Ion Mobility study

F. Halgand^{1,*}, Johnny Habchi², Laetitia Cravello³, Marlène Martinho⁴, Bruno Guigliarelli⁴, and Sonia Longhi²

¹Laboratoire de Bioénergétique et Ingénierie des Protéines. Equipe de protéomique fonctionnelle et dynamique. UPR 9036-CNRS; 31 Chemin Joseph Aiguier, 13420, Marseille Cedex. France

²Architecture et Fonction des Macromolécules Biologiques (AFMB), UMR 6098, CNRS et Aix-Marseille Université, 163 Avenue de Luminy Case 932, 13288 Marseille Cedex 09, France

³Waters Corp., Atlas Park, Simonsway, Manchester, Lancashire M22 5PP, United Kingdom

⁴Laboratoire de Bioénergétique et Ingénierie des Protéines. Equipe de Biophysique des Métalloprotéines. UPR 9036-CNRS, 31 Chemin Joseph Aiguier, 13420, Marseille Cedex, France

Abstract

Over-expression 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 micro-heterogeneities. While standard proteomics approaches preclude unbiased search for modifications, the optional technique of top down 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 micro-heterogeneities that were hitherto only accessible with FTICR-MS instruments. The power of this approach is herein illustrated in an in-depth analysis of both *wt* and K496C variant of the recombinant X domain (XD, aa 459-507) of the measles virus phosphoprotein expressed in *E. coli*. Using top down MSMS combined to TWIM, we show that XD samples occasionally exhibit a micro-heterogeneity 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, an MTSN nitroxide probe that was grafted on the K496C XD variant was shown to undergo oxidation and/or protonation in the ESI source leading to artifactual mass increases.

Introduction

Over-expression of a protein in a foreign host, such as *Escherichia coli*, is frequently the first step towards 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 last decade. These 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 over-

* Author for correspondence: Dr Frédéric Halgand. Laboratoire de Bioénergétique et Ingénierie des Protéines. Equipe de protéomique fonctionnelle et dynamique. UPR 9036-CNRS, 31 Chemin Joseph Aiguier, 13420, Marseille Cedex, France. Tel. (33) 4 91 16 46 73; Fax. (33) 4 91 16 40 97; fhalgand@ifr88.cnrs-mrs.fr.

expression 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 so as 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 micro-heterogeneities (for a review see 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, this 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 last combines the measurement of the intact experimental mass (Intact Mass Tag, IMT¹) with the recording of MSMS data on the full-length protein. Such 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 sample preparation, and data handling still remain critical to achieve meaningful results ^{2-3, 8-9}. Thus, successful and complete protein properties 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 as well as to limited isotopic profiles overlapping. For example, the complete description of a 4.3 kDa protein bearing subtle modifications of $\pm \sim 1$ Da mass differences using top down MSMS was only achieved using a 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)¹ 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)¹, collisional cross section (CCS)¹ and size with the determination of drift times, a parameter reflecting gas phase conformation of proteins and peptides (for further details about TWIM, see ¹⁰⁻¹¹). The TWIM guide, either alone ¹² or in conjunction with hydrogen exchange ¹³⁻¹⁵, allows inferring, in the sub-millisecond 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 a 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 (XD¹, aa 459-507) of the measles virus (MV)¹ phosphoprotein (P)¹. Recent studies allowed the determination of the structure of XD by both X-ray crystallography²⁰ and NMR²¹. 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}¹, aa 401-525) of the MV nucleoprotein (N)^{1 22-30}. Those studies provided structural details on the N_{TAIL} binding interface of XD³¹⁻³⁴ and designated residue K496 of XD as an interesting target for cysteine substitution and ensuing labeling with an electron paramagnetic resonance (EPR)¹ label, such as the (1-oxyl-2,2,5,5-tetramethyl- δ 3-pyrroline-3-methyl) methanethiosulfonate (MTSL)¹, in view of studies aimed at inferring information in terms of spin probe mobility under various conditions.

Material and Methods

Materials

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

Methods

Construction of protein expression plasmids—The XD gene construct (pDest14/XD_{HC}), encoding the X domain (XD, residues 459-507) of the MV P protein (strain Edmonston B) with an hexahistidine tag fused to its C-terminus and under the control of the T7 promoter, has already been described². The pDest14/XD_{HC} K496C construct, encoding C-terminally hexahistidine tagged XD bearing the K496C substitution, was obtained by PCR using pDest14/XD_{HC} as template, Turbo-Pfu polymerase (Stratagene) and a pair of complementary mutagenic primers of 39 nucleotides in length (Operon). After digestion with *DpnI* to remove the methylated DNA template, transformation of *E. coli* strain DH10 β (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 described². 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 in². The proteins (in 10 mM sodium phosphate buffer pH 7) were stored at -20°C . 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)¹ of a folded protein of the same molecular mass as XD was calculated according to³⁵. Protein concentrations were estimated using the Biorad protein assay (Bio-Rad).

Circular Dichroism—Circular dichroism (CD)¹ 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 square polynomial fit. Finally, mean ellipticity values per residue ($[\Theta]$) were calculated as $[\Theta] = 3300 \text{ m } \Delta A / (l \text{ c } n)$, where l (path length) = 0.1 cm, n = number of residues, m = molecular mass in daltons and c = protein concentration expressed in mg/ml.

Number of residues (n) is 56 for both XD proteins, and m is 6 686 Da (*wt* XD) or 6 661 Da (XD K496C). Protein concentrations of 0.1 mg/ml were used.

Spin-labeling of the XD K496C variant—Before spin-labeling, DTT was added (to yield a final concentration of 20 mM) to the K496C XD protein (125 μ L at 459 μ M). The mixture was incubated for 30 min in an ice bath in order to reduce the unique free cysteine residue. DTT was removed by gel filtration using a PD-10 column (GE Healthcare) and 10 mM MES, 150 mM NaCl at pH 6.5 as elution buffer. The fractions containing the protein were pooled. Prior utilization, the spin label (1-oxy-2,2,5,5-tetramethyl δ 3-pyrroline-3-methyl) methane-thiosulfonate, MTSL¹, (Toronto Research Chemicals Inc., Toronto, Canada) was solubilized in acetonitrile to yield a final concentration of 10 mg/mL. Then it was immediately incorporated to the protein at a molar excess of 20 every hour for 4 hours, leading to a total molar excess of 80. The reaction was carried out for 4 hours 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 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 a 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 μ L of concentrated XD protein solutions ($\sim$ 150 μ M) were diluted to a final volume of 100 μ L 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 14000 rpm, protein pellets were recovered and dissolved in a MeOH/H₂O (50/50 V/V) solution containing 1% formic acid, to a final concentration of 2 μ M. 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 μ M final concentration (10 fold excess) and incubated 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 incubated for 30 min in the dark at room temperature.

Electrospray ionization mass spectrometry (ESI-MS)¹ experiments—ESI-MS spectra of full-length XD proteins were recorded in positive ion mode using a “nanoflow”TM electrospray ion source fitted on a quadrupole-time of flight hybrid mass spectrometer equipped with a TWIM cell (Synapt HDMS G1, Waters Corp., Manchester). Protein solutions were introduced at a 0.5 μ L/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*, *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 collision gas (99.999% argon). Collision energy was optimized to achieve the best fragmentation yield of the selected parent ion. Typically, collision energy was about 35 to 90 eV, depending on protein CS. Proteins were fragmented using various CS 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.* Fig. 3). The DriftScope™ software was used to visualize and reprocess TWIM data.

Semi-automated 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, UK). 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 Supplemental materials. 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 expectations² (see supplementary Fig. S1 and S2), the two plasmid constructs were transformed in the Rosetta pLysS strain of *E. coli*. Fig. 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%) (Fig. 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 (Fig. 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 ($13 \pm 2 \text{ \AA}$)² and consistent with the theoretical value expected for a folded protein form³⁵. Thus, *wt* and K496C XD share similar hydrodynamic properties, adopting both 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 (Fig. 1B). These data indicate that the cysteine substitution induces little, 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 manuscript. 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 homogenous, 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.* supplementary **Fig. S8** 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. **Figs. 2A** and **2B** show such a comparison for the +8 CS of the *wt* XD and +7 CS of K496C* XD proteins, respectively.

ESI-MS mass measurements of full-length XD proteins using TWIM

In order 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 CSD of XD samples and in 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 under detection of isotopic profiles of XD *wt* variants (see supplementary **Fig. S3**). Indeed, the dynamic range was shown to be smaller under TWIM experiment due to the fact that ions arrived in packets, then 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 supplementary **Fig. 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 profiles overlap, using the ability of the TWIM guide to segregate MSMS fragments according to their size, CCSs and CSs (see Fig. 3). The propensity of TWIM to separate fragment ions was our voucher to achieve reliable top down MSMS data interpretation. MSMS datasets were analyzed in detail using both semi-automated and manual approaches (see Materials and Methods). Combined MSMS datasets recorded on several CS 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 supplementary **Fig. S4**). However, experimental mass differences observed for *wt* XD between averaged (–2.1 Da) and monoisotopic molecular weights (–0.61 Da) still suggests 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 for native XD regular sequence and for *wt* XD variants (V5P, E17K/Q and L43N variants) is provided in Fig. 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 supplementary **Fig. S5**. The experimental masses of the b and y ions, the mass errors calculated in Da and PPM, as well as the identification of unique fragment ions providing further support to our interpretation of the observed heterogeneities are given in supplementary **Tables S1 to S3**.

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 (Fig. 4). A summary of our findings for XD K496C IAA variant with sequence charts, probability scores and RMS is given in supplementary Fig. S6 and Table S4. Finally, analysis of top down MSMS data recorded for K496C* XD revealed that this latter bears the same modification as K496C IAA XD at position 43. However, the measured mass increment for K496C* XD indicates that this 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 supplementary **Fig. 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 these latter conditions (data not shown), revealed that the MTSL was neither protonated on the nitroxile nor did it retain any proton in the surrounding of the nitroxile 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 this 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 Fig. 5. Further support to this conclusion can be found in the supplementary **Table S5**, which shows that, K496C* XD

variant mass was still 2 Da heavier than expected when MTSL was the only added modification.....

Discussion

In this article we evaluated the power of the TWIM guide for reducing top down MSMS spectra complexity and for elucidating the origin of subtle sample heterogeneities. Indeed, it is critical to accurately identify fragment ions for correct assignment of 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 these 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 these latter were homogeneous, then the observed mass errors for monoisotopic and averaged masses would be in the same range (see **Table** in Figure S8).

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 achieving an efficient separation of fragment ions according to their drift times related to their respective sizes, CSs and CCSs (Fig. 3). Ion separation allowed large b and y series ions to be distinguished with better signal-to-noise ratios and better statistics compared to full MSMS spectrum. A similar achievement could only have been accomplished with an FTICR-MS instrument. These datasets allowed achieving 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 ions 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 supplementary **Tables S1 to S5**). 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 based on isotopic profiles modeling (see supplementary **Figs. S4 to 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 monomers. This observation implies that spin labeling of the XD K496C is partial, probably as a result of uncomplete 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 proline³⁸.

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 V/V/ 1%) used in the ESI experiments. The mass spectrum revealed that the MTSL group is represented by 3 peaks of 264.07, 265.08 and 266.09 Da, respectively (Figure 5). The structures proposed for these forms are displayed in Fig. 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 1 and/or 2 Da mass increase in the full-length K496C* variant (see supplementary **Fig. 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 get further insights into the structural changes related to the amino acid substitutions observed in the native XD sample, we attempted at calculating 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 having no effect on the global conformation to lead to any distinguishable features in 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 of 1371.5 and 1491.0 Å², respectively (see supplementary **Fig. S4**). Notably, the XD V5P variant was found to have a lower CCS value (minus 213 Å²) 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 metaserver³⁹, indicate that the V5P substitution is predicted to break helix α 1. Since we estimated that the CCSs of the two species differed by ~ 200 Å² (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 miss-insertion nor any nucleotide substitutions, the observed micro-heterogeneity 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 not only found 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 non-random 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. Miss-incorporation of an Asn during translation as a result of a three miss-match 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%) codons⁴⁰, 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 Fig. 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 substitution nor the E17Q/K one are 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 metaserver³⁹ (data not shown).

As for the V5P and E17Q/K substitutions, at least two or one nucleotide substitutions, respectively, are 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 recombination⁴¹ and are also more sensitive to UV mutagenesis⁴² than *recA*⁻ strains, i.e. strains bearing a mutated, non functional *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 to undesired recombination events and/or mutations. Therefore, it is conceivable that the protein micro-heterogeneities observed upon expression of XD proteins in the Rosetta pLysS strain may result from this kind of events.

On the other hand, we would like to underscore that such protein micro-heterogeneity 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 supplementary **Fig. S8**). 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 micro-heterogeneous *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 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 micro-heterogeneities, presumably due to genetic drift events (neutral or not), has important implications in terms of both biochemical and structural characterization of proteins. Indeed, such micro-heterogeneities may impair crystallization of pure and monodisperse protein samples and may lead to an unexpected behavior such as altered interaction ability towards a partner or ligand. As such, it is advised to keep in mind

the possibility of protein micro-heterogeneity when protein samples do not behave as expected and to submit them to extensive analyses such 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 micro-heterogeneity was achieved by a reliable identification of fragment ions and by methodical comparison of isotopic profiles of fragment ions. In particular, TWIM allowed calculating 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.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors are grateful to the proteomics core facility of the Institut de Microbiologie de la Méditerranée for their technical assistance. SL wishes to thank Aissette Baanannou and Stéphanie Costanzo for their contribution to cloning and purification procedures. The authors 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 (R01 NS031693-11A2). The authors are also grateful to the Conseil Régional of Provence Alpes Cotes d’Azur, to the city of Marseille as well as to the University of Provence for financial support in the acquisition of instrumentation. Sylvain Marque is acknowledged for its scientific support.

Abbreviations

CD	Circular Dichroism
CS	Charge State
CSD	Charge State Distribution
CCS	Collisional Cross Section
DC	Direct Current
EPR	Electron Paramagnetic Resonance
ESI-MS	Electrospray Ionization Mass Spectrometry
HEX	Hydrogen Exchange
IAA	IodoAcetamide Alkylation
IMT	Intact Mass Tag
MV	Measles virus
ORF	Open Reading Frame
P	phosphoprotein

QTof	Quadrupole Time-of-flight (mass spectrometer)
MTSL	(1-oxyl-2,2,5,5-tetramethyl- δ 3-pyrroline-3-methyl) methanethiosulfonate
MSMS	tandem mass spectrometry
N_{TAIL}	C-terminal region of measles virus nucleoprotein spanning residues 401 to 525
N	measles virus nucleoprotein
Rf	radiofrequency
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 to 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 JK. Strategies for efficient production of heterologous proteins in *Escherichia coli*. *Appl Microbiol Biotechnol*. 2005; 67:289–298. [PubMed: 15635462]
2. Kelleher NL, Lin HY, Valaskovic GA, Aaserud DJ, Fridriksson EK, M FW. Top down versus bottom up protein characterization by tandem high-resolution mass spectrometry. *J. Am. Chem. Soc.* 1999; 121:806–807.
3. Pesavento JJ, Kim YB, Taylor GK, Kelleher NL. Shotgun annotation of histone modifications: a new approach for streamlined characterization of proteins by top down mass spectrometry. *J Am Chem Soc.* 2004; 126:3386–3387. [PubMed: 15025441]
4. Collier TS, Sarkar P, Rao B, Muddiman DC. Quantitative top-down proteomics of SILAC labeled human embryonic stem cells. *J Am Soc Mass Spectrom.* 21:879–889. [PubMed: 20199872]
5. Wu S, Lourette NM, Tolic N, Zhao R, Robinson EW, Tolmachev AV, Smith RD, Pasa-Tolic L. An integrated top-down and bottom-up strategy for broadly characterizing protein isoforms and modifications. *J Proteome Res.* 2009; 8:1347–1357. [PubMed: 19206473]
6. Ouvry-Patat SA, Torres MP, Gelfand CA, Quek HH, Easterling M, Speir JP, Borchers CH. Top-down proteomics on a high-field Fourier transform ion cyclotron resonance mass spectrometer. *Methods Mol Biol.* 2009; 492:215–231. [PubMed: 19241035]
7. Lee JE, Kellie JF, Tran JC, Tipton JD, Catherman AD, Thomas HM, Ahlf DR, Durbin KR, Vellaichamy A, Ntai I, Marshall AG, Kelleher NL. A Robust Two-Dimensional Separation for Top-Down Tandem Mass Spectrometry of the Low-Mass Proteome. *J Am Soc Mass Spectrom.* 2009
8. Halgand F, Zabrouskov V, Bassilian S, Souda P, Wong DT, Loo JA, Faull KF, Whitelegge JP. Micro-heterogeneity of human saliva Peptide P-C characterized by high-resolution top-down Fourier-transform mass spectrometry. *J Am Soc Mass Spectrom.* 2010; 21:868–877. [PubMed: 20185333]
9. Ryan CM, Souda P, Halgand F, Wong DT, Loo JA, Faull KF, Whitelegge JP. Confident assignment of intact mass tags to human salivary cystatins using top-down Fourier-transform ion cyclotron resonance mass spectrometry. *J Am Soc Mass Spectrom.* 21:908–917. [PubMed: 20189825]

10. Giles K, Pringle SD, Worthington KR, Little D, Wildgoose JL, Bateman RH. Applications of a travelling wave-based radio-frequency-only stacked ring ion guide. *Rapid Commun Mass Spectrom.* 2004; 18:2401–2414. [PubMed: 15386629]
11. Smith DP, Knapman TW, Campuzano I, Malham RW, Berryman JT, Radford SE, Ashcroft AE. Deciphering drift time measurements from travelling wave ion mobility spectrometry-mass spectrometry studies. *Eur J Mass Spectrom (Chichester, Eng).* 2009; 15:113–130.
12. Williams JP, Phillips HI, Campuzano I, Sadler PJ. Shape changes induced by N-terminal platination of ubiquitin by cisplatin. *J Am Soc Mass Spectrom.* 2010; 21:1097–1106. [PubMed: 20227292]
13. Hopper JT, Oldham NJ. Collision induced unfolding of protein ions in the gas phase studied by ion mobility-mass spectrometry: the effect of ligand binding on conformational stability. *J Am Soc Mass Spectrom.* 2009; 20:1851–1858. [PubMed: 19643633]
14. Rand KD, Pringle SD, Murphy JP 3rd, Fadgen KE, Brown J, Engen JR. Gas-phase hydrogen/deuterium exchange in a traveling wave ion guide for the examination of protein conformations. *Anal Chem.* 2009; 81:10019–10028. [PubMed: 19921790]
15. Abzalimov RR, Kaplan DA, Easterling ML, Kaltashov IA. Protein conformations can be probed in top-down HDX MS experiments utilizing electron transfer dissociation of protein ions without hydrogen scrambling. *J Am Soc Mass Spectrom.* 2009; 20:1514–1517. [PubMed: 19467606]
16. Atmanene C, Wagner-Rousset E, Malissard M, Chol B, Robert A, Corvaia N, Van Dorsselaer A, Beck A, Sanglier-Cianferani S. Extending mass spectrometry contribution to therapeutic monoclonal antibody lead optimization: characterization of immune complexes using noncovalent ESI-MS. *Anal Chem.* 2009; 81:6364–6373. [PubMed: 19588976]
17. Hilton GR, Thalassinou K, Grabenauer M, Sanghera N, Slade SE, Wytenbach T, Robinson PJ, Pinheiro TJ, Bowers MT, Scrivens JH. Structural analysis of prion proteins by means of drift cell and traveling wave ion mobility mass spectrometry. *J Am Soc Mass Spectrom.* 21:845–854. [PubMed: 20206551]
18. Fenn LS, McLean JA. Simultaneous glycoproteomics on the basis of structure using ion mobility-mass spectrometry. *Mol Biosyst.* 2009; 5:1298–1302. [PubMed: 19823744]
19. Djidja MC, Francese S, Loadman PM, Sutton CW, Scriven P, Claude E, Snel MF, Franck J, Salzet M, Clench MR. Detergent addition to tryptic digests and ion mobility separation prior to MS/MS improves peptide yield and protein identification for in situ proteomic investigation of frozen and formalin-fixed paraffin-embedded adenocarcinoma tissue sections. *Proteomics.* 2009; 9:2750–2763. [PubMed: 19405023]
20. Johansson K, Bourhis JM, Campanacci V, Cambillau C, Canard B, Longhi S. Crystal structure of the measles virus phosphoprotein domain responsible for the induced folding of the C-terminal domain of the nucleoprotein. *J Biol Chem.* 2003; 278:44567–44573. [PubMed: 12944395]
21. Bernard C, Gely S, Bourhis JM, Morelli X, Longhi S, Darbon H. Interaction between the C-terminal domains of N and P proteins of measles virus investigated by NMR. *FEBS Lett.* 2009; 583:1084–1089. [PubMed: 19275899]
22. Longhi S, Receveur-Brechot V, Karlin D, Johansson K, Darbon H, Bhella D, Yeo R, Finet S, Canard B. The C-terminal domain of the measles virus nucleoprotein is intrinsically disordered and folds upon binding to the C-terminal moiety of the phosphoprotein. *J Biol Chem.* 2003; 278:18638–18648. [PubMed: 12621042]
23. Bourhis J, Johansson K, Receveur-Brechot V, Oldfield CJ, Dunker AK, Canard B, Longhi S. The C-terminal domain of measles virus nucleoprotein belongs to the class of intrinsically disordered proteins that fold upon binding to their physiological partner. *Virus Research.* 2004; 99:157–167. [PubMed: 14749181]
24. Bourhis JM, Receveur-Brechot V, Oglesbee M, Zhang X, Buccellato M, Darbon H, Canard B, Finet S, Longhi S. The intrinsically disordered C-terminal domain of the measles virus nucleoprotein interacts with the C-terminal domain of the phosphoprotein via two distinct sites and remains predominantly unfolded. *Protein Sci.* 2005; 14:1975–1992. [PubMed: 16046624]
25. Morin B, Bourhis JM, Belle V, Woudstra M, Carrière F, BGuigliarelli B, Fournel A, Longhi S. Assessing induced folding of an intrinsically disordered protein by site-directed spin-labeling EPR spectroscopy. *J. Phys. Chem. B.* 2006; 110:20596–20608. [PubMed: 17034249]

26. Belle V, Rouger S, Costanzo S, Liquiere E, Strancar J, Guigliarelli B, Fournel A, Longhi S. Mapping alpha-helical induced folding within the intrinsically disordered C-terminal domain of the measles virus nucleoprotein by site-directed spin-labeling EPR spectroscopy. *Proteins*. 2008; 73:973–988. [PubMed: 18536007]
27. Gely S, Lowry DF, Bernard C, Ringkjober-Jensen M, Blackledge M, Costanzo S, Darbon H, Daughdrill GW, Longhi S. Solution structure of the C-terminal X domain of the measles virus phosphoprotein and interaction with the intrinsically disordered C-terminal domain of the nucleoprotein. *J Mol Recognit*. 2010; 23:435–447. [PubMed: 20058326]
28. Kavalenka A, Urbancic I, Belle V, Rouger S, Costanzo S, Kure S, Fournel A, Longhi S, Guigliarelli B, Strancar J. Conformational analysis of the partially disordered measles virus NTAIL-XD complex by SDSL EPR spectroscopy. *Biophys J*. 2010; 98:1055–1064. [PubMed: 20303863]
29. Bischak CG, Longhi S, Snead DM, Costanzo S, Terrer E, Londergan CH. Probing structural transitions in the intrinsically disordered C-terminal domain of the measles virus nucleoprotein by vibrational spectroscopy of cyanylated cysteines. *Biophys J*. 2010; 99:1676–1683. [PubMed: 20816082]
30. Jensen MR, Communie G, Ribeiro EA Jr, Martinez N, Desfosses A, Salmon L, Mollica L, Gabel F, Jamin M, Longhi S, Ruigrok RW, Blackledge M. Intrinsic disorder in measles virus nucleocapsids. *Proc Natl Acad Sci U S A*. 2011; 108:9839–9844. [PubMed: 21613569]
31. Bourhis JM, Canard B, Longhi S. Structural disorder within the replicative complex of measles virus: functional implications. *Virology*. 2006; 344:94–110. [PubMed: 16364741]
32. Bourhis JM, Canard B, Longhi S. Predicting protein disorder and induced folding: from theoretical principles to practical applications. *Curr Protein Pept Sci*. 2007; 8:135–149. [PubMed: 17430195]
33. Longhi, S. Measles virus nucleoprotein. Nova Publishers Inc.; Hauppauge, NY: 2007.
34. Longhi S, Oglesbee M. Structural disorder within the measles virus nucleoprotein and phosphoprotein. *Protein and Peptide Letters*. 2010; 17:961–978. [PubMed: 20450481]
35. Uversky VN. Use of fast protein size-exclusion liquid chromatography to study the unfolding of proteins which denature through the molten globule. *Biochemistry*. 1993; 32:13288–13298. [PubMed: 8241185]
36. Roepstorff P, Fohlman J. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed Mass Spectrom*. 1984; 11:601. [PubMed: 6525415]
37. LeDuc RD, Taylor GK, Kim YB, Januszyk TE, Bynum LH, Sola JV, Garavelli JS, Kelleher NL. ProSight PTM: an integrated environment for protein identification and characterization by top-down mass spectrometry. *Nucleic Acids Res*. 2004; 32:W340–345. [PubMed: 15215407]
38. Frottin F, Martinez A, Peynot P, Mitra S, Holz RC, Giglione C, Meinnel T. The proteomics of N-terminal methionine cleavage. *Mol Cell Proteomics*. 2006; 5:2336–2349. [PubMed: 16963780]
39. Lieutaud P, Canard B, Longhi S. MeDor: a metasever for predicting protein disorder. *BMC Genomics*. 2008; 9:S25. [PubMed: 18831791]
40. Maloy, S.; Stewart, V.; Taylor, R. Genetic analysis of pathogenic bacteria. Cold Spring Harbor Laboratory Press; NY: 1996.
41. Voloshin ON, Wang L, Camerini-Otero RD. Homologous DNA pairing promoted by a 20-amino acid peptide derived from RecA. *Science*. 1996; 272:868–872. [PubMed: 8629021]
42. Dutreix M, Moreau PL, Bailone A, Galibert F, Battista JR, Walker GC, Devoret R. New recA mutations that dissociate the various RecA protein activities in Escherichia coli provide evidence for an additional role for RecA protein in UV mutagenesis. *J Bacteriol*. 1989; 171:2415–2423. [PubMed: 2651400]
43. DeLano WL. The PyMOL molecular graphics system. *Proteins: Structure, Function and Bioinformatics*. 2002; 30:442–454.

Figure 1.

(A) Amino acid sequence and ribbon representation of the crystal structure of MeV XD (pdb code 1OKS) as obtained using Pymol⁴³. 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 structure of XD (pdb codes 1OKS and 2K9D, respectively) are highlighted in colour. 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 a 18% SDS-PAGE loaded with approximately 5 μ g 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 ($\pm$ 1 K): Microwave power = 8 mW, Amplitude modulation = 0.1 mT, Sweep width = 15 mT, Number of accumulations = 8.

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

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 size, CSs and CCSs. **(C), (D) and (E)** Top down ESI-MSMS spectra of *wt* XD displaying mono, doubly and triply and triply and quadruply charged ions separated using TWIM, respectively.

Figure 4.

Sequence charts showing fragment ions matching the sequences of native and XD variants. Manual Pcores 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 3 peaks @ m/z values of 264.07 Da, 265.08 and 266.09 Da, are also displayed.

Table 1

Experimental monoisotopic and averaged masses of XD monomers. Calculated errors between experimental and theoretical monoisotopic and averaged masses are given in Da and PPM, in the two last columns.

	Experimental masses		Theoretical masses		Mass errors	
	Monomer		Monomer		Mono. Da	Av. Da
Samples	Mono. mass	Av. mass	Mono. mass	Av. mass		
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

MTSL = monoisotopic mass +185.09; averaged mass = +186.3 Da

IAA = monoisotopic mass +57.02 Da; averaged mass = +57.05 Da