ELSEVIER

Biochimica et Biophysica Acta 1597 (2002) 252–259

BIOCHIMICA ET BIOPHYSICA ACTA
BBA
www.bba-direct.com

MALDI-TOF mass spectrometry characterization of recombinant hydrophobic mutants containing the GCN4 basic region/leucine zipper motif

Gregory H. Bird, Jumi A. Shin[1,*]

Department of Chemistry, Faculty of Arts and Sciences, University of Pittsburgh, Pittsburgh, PA 15260, USA

Received 14 November 2001; received in revised form 5 February 2002; accepted 1 March 2002

Abstract

We used matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) to characterize hydrophobic, alanine-rich mutants of the basic region/leucine zipper (bZIP) protein GCN4. These bacterially expressed proteins were generated to probe how small, α-helical proteins bind specific DNA sites. Enzymatic digestion mapping combined with MALDI-TOF MS characterization of protein fragments allowed us to resolve mass discrepancies between the expected and observed molecular mass measurements. Changes in mass were attributed to posttranslational modifications (PTMs) by proteolytic cleavage of the initiating methionine residue, carbamylation at the amino terminus, oxidation of histidine side chains, and oxidative addition of β-mercaptoethanol (BME) at the cysteine side chain. Proteins can undergo a wide variety of co-translational modifications and PTMs during growth, isolation, and purification. Such changes in mass can only be detected by a high-resolution technique such as MALDI, which in conjunction with enzymatic digestion mapping, becomes a powerful methodology for characterization of protein structure.

Keywords: bZIP; GCN4; MALDI-TOF; Posttranslational modification; Enzymatic digestion mapping; Protein aggregation

1. Introduction

Matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) has become a powerful technique for the identification of biopolymers. Low sample thresholds for analysis (femtomoles–picomoles of analyte), high mass accuracy (0.01–0.1% mass error), and tolerance to buffers and salts are all attributes of MALDI-MS useful for analysis of proteins [1]. An accurate mass measurement can confirm that a desired protein has been properly expressed and isolated. However, the actual mass of a recombinant protein can differ from the expected mass, which we derive from translation of the gene's DNA sequence. If the DNA sequence is accurate, then inconsistencies in the measured molecular mass are likely to be caused by co- and post-translational modifications (PTMs) [2].

PTMs are typically a consequence of any of the following: the protein expression system (e.g. bacterial, mammalian), isolation and purification procedure, and storage. While some modifications, such as phosphorylation, are reversible, the majority represents permanent changes to the molecule. Nearly all modifications alter the mass of the protein and, as such, can be evaluated by mass spectrometry.

We have focused on alanine-scanning mutants of the α-helical basic region/leucine zipper motif (bZIP) of GCN4, a dimeric yeast transcriptional regulatory protein [3]. The full-length, 281-residue GCN4 monomer comprises the ~60-residue bZIP domain, which is responsible for binding the DNA major groove with high affinity and sequence specificity [3–6]. In order to explore the structural and functional aspects of DNA recognition by α-helical proteins, we substituted alanines into the basic regions of bZIP derivatives comprising the GCN4 basic region (residues 226–252) and C/EBP leucine zipper (residues 310–338) [7]. The wt bZIP (wild type) is the "native" variant comprising the GCN4 basic region and C/EBP zipper (Fig. 1). Mutants **4A**, **11A**, and **18A** contain 4, 11, and 18 Ala replacements in their basic regions, respectively. These proteins are unusual for bacterial expression in that they are short (~100 amino acids),

* Corresponding author. Tel.: +1-412-624-4927; fax: +1-412-624-4255.
E-mail addresses: jumi@imappitt.edu, jshin@utm.utoronto.ca (J.A. Shin).
[1] Present address: Department of Chemistry, University of Toronto at Mississauga, Mississauga, ON L5L 1C6, Canada

0167-4838/02/$ - see front matter
PII: S0167-4838(02)00303-5

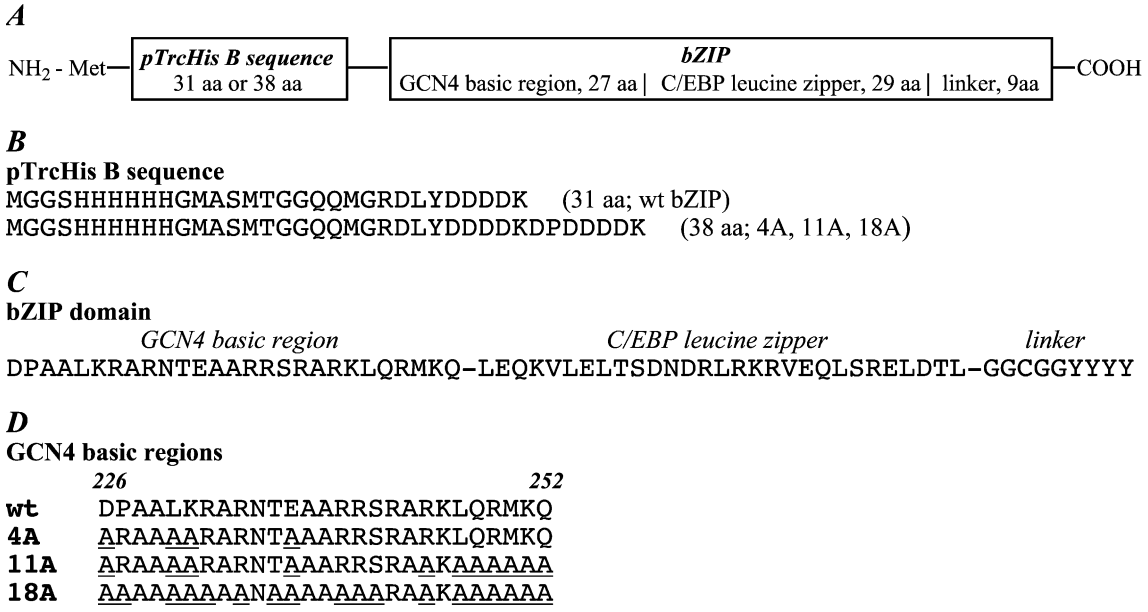

Fig. 1. (A) Schematic of expressed protein. The bZIP proteins were cloned into expression vector pTrcHis B, which contains a six-histidine tag for protein purification and enterokinase cleavage site (DDDDK). Note that the pTrcHis sequences for **4A**, **11A**, and **18A** contain an additional enterokinase site. The bZIP is at the carboxyl termini of the expressed proteins, which is the same positioning of the bZIP domain in native GCN4. The fully expressed bZIP comprises approximately 35 residues from the pTrcHis expression vector, basic region mutants of GCN4 (residues 226–252), leucine zipper from C/EBP (residues 310–338), plus a linker (GGCGGYYYY) for covalent protein dimerization (cysteine) and chemical derivatization (tyrosines) [24]. (B) Sequences of the peptides from the pTrcHis B expression vector. (C) Sequence of the wt bZIP domain. (D) Sequences of the basic regions. The sequences for **4A**, **1A**, and **18A** are shown below wt bZIP; these proteins are the same as wt bZIP, except for the mutated basic regions. Alanine substitutions are underlined. We note that Pro^{227} is arginine in both **4A** and **11A**; this is a cloning artifact, and this residue has no interaction with DNA [5,6,25].

hydrophobic (Ala-rich basic regions, leucine zipper dimerization domains), and prone to aggregation. Previously, we showed that these mutants maintain native structure and function: circular dichroism confirms their α-helical structure, and DNase I footprinting demonstrates that all mutants bind specifically to the AP-1 DNA site, the in vivo target site of GCN4 in yeast [8].

One of the first opportunities for protein modification occurs at the amino terminus when the nascent protein is only 20–40 amino acids in length [2]. Methionine aminopeptidase (MAP) and *N*α-acetyltransferase (NAT), both associated with the ribosome, can act individually or in concert to cleave the initiating Met and add an acetyl group to the N-terminus [2]; during protein overexpression, MAP can become saturated with protein or unable to access the amino termini due to protein aggregation, thereby leading to heterogeneity at the N-terminus [2,9,10]. Therefore, there exist four possibilities for the mature protein's amino terminus: with or without Met (demethioninylation) and with or without *N*-acetyl (deacetylation). For example, Klein et al. [11] expressed recombinant mammalian α interferon (IFN) in *Escherichia coli* and found heterogeneous amino termini: intact IFN, demethioninylated IFN, and demethioninylated IFN with N-terminal acetylation. Bioassay showed all three proteins to be equally active, and circular dichroism showed them to be structurally identical.

Prokaryotic protein synthesis is initiated with formylmethionine t-RNA [12]. The removal of *N*-formylMet by the metallopeptidase MAP is a sequential event [10,12] in which deformylation precedes demethioninylation [13]. Met is readily removed from proteins in which the penultimate amino acid has a small, aliphatic, uncharged side chain. The binding pocket for MAP is small and nonionic; large or charged residues adjacent to Met favor its retention [14]. Therefore, a penultimate proline, valine, cysteine, glycine, alanine, serine, and threonine favor Met removal, and further effect can be made by the residue adjacent to the penultimate amino acid [9]. NAT also prefers substrates with small side chains [2]. Our proteins are constructed with the initiating Met codon followed by two glycines, residues that promote N-terminal modification.

Another type of PTM is oxidation of the histidine imidazole side chain to generate 2-oxohistidine (2-oxoHis), in which the C2 position between the two nitrogens is oxidized to a carbonyl, and the imine nitrogen is reduced to an amine [15]. 2-OxoHis can be thought of as a cyclic urea derivative that does not retain the basicity of an imidazole, making it more hydrophobic [16]. Histidine is particularly susceptible to oxidative modification, presumably because it can chelate metals capable of generation of reactive species [15–17]. Our proteins contain a six-histidine tag at the amino termini that aids in purification by metal ion affinity chromatography, and hence, oxidation can occur.

Because our hydrophobic proteins aggregate readily, we use urea as a denaturant during all stages of preparation and storage. Although urea is considered to be inert, its degra-

dation products, ammonia and isocyanic acid, are not inert [18,19]. Isocyanic acid is an electrophile that can be attacked by the nucleophilic amino terminus and by the lysine amino side chain, depending on pH. After proton transfer and tautomerization, carbamylation of the amine results. Use of purified, deionized urea during protein preparation avoids carbamylation. Also commonly present in buffers used in protein isolation and purification is β-mercaptoethanol (BME), as it reduces cysteine disulfide bonds. Preventing inappropriate disulfide formation is important for optimizing yields and activity of expressed proteins. Near the carboxyl termini, our proteins contain a cysteine, whose thiol side chain is susceptible to disulfide formation with BME.

All of the modifications discussed above were found to be present in our bZIP proteins. Using mass spectrometry combined with enzymatic digestion mapping, we were able to characterize these modifications in wt bZIP, **4A**, **11A**, and **18A**. Although they changed the expected molecular masses of our proteins, these modifications did not affect their α-helical structure or DNA-binding function, as these changes were not in the bZIP regions of our proteins [8].

2. Materials and methods

2.1. Materials

All solvents were of HPLC grade and were used without further purification. Water was purified through a Milli Q filtration system (Millipore). In later experiments (see Results and discussion), urea used in all solutions was deionized by the following procedure: a column containing the mixed bed resin DOWEX MR3C (Supelco/Dow Chemicals) was sequentially equilibrated with 5% NaOH, water, and 1 M HCl, before the stock solution of 8 M urea in water was passed through it. This deionized urea solution was used to make other urea-containing solutions.

2.2. Protein preparation

A brief summary of the previously described protocols for gene construction and cloning, and protein overexpression and purification follows [7]. Genes for expression of bZIP mutants were constructed by mutually primed synthesis, cloned into protein expression vector pTrcHis B (Invitrogen), and transformed into *E. coli* strain BL21(DE3) (Stratagene) by electroporation (Bio-Rad). These His_6 tagged proteins were purified first on TALON cobalt metal-ion affinity resin (Clontech), followed by further purification either by size-exclusion chromatography (Superdex 75 HR column, Pharmacia) or reverse-phase chromatography (C4 column, Vydac) on a Beckman System Gold HPLC. Purification was monitored by SDS polyacrylamide gel electrophoresis (SDS-PAGE) and Western immunoblot assay. Purified stocks of wt bZIP, **4A**, **11A**, and **18A** were stored in SEC buffer (50 mM phosphate, pH 6.8, 10% acetonitrile, 150 mM NaCl, 4 M urea) at final concentrations of 28.4, 24.6, 28.5, and 23.5 μM, respectively, at −80 °C with protease inhibitors (1 mM PMSF and 1 μg/ml pepstatin).

2.3. MALDI sample preparation

The matrix solutions were prepared as described previously [20]. Protein stocks (250 μl) in SEC buffer were pipetted into Ultrafree-15 Amicon centrifuge concentrators (Millipore). Washes (3×10 ml) with Endo-Lys buffer (0.1 M Tris–HCl, pH 9.0, 4 M urea; we note that Endo-Lys is inhibited by monovalent cations, so no basic adjustments were made to the pH) were used to equilibrate protein buffer, remove protease inhibitors, and concentrate protein stocks.

2.4. On-plate digestion with Endo-Lys carboxypeptidase

One microliter of saturated sinapinic acid solution was spotted onto a 100-well gold MALDI-time-of-flight (TOF) MS sample plate (Applied Biosystems) and allowed to dry completely at room temperature. Then, 2–4 μl protein solution (~40 μM monomeric bZIP) was added on top of the sinapinic acid spot and allowed to dry completely at room temperature. To this dried droplet, 5 μl of a 20 nM solution of Endo-Lys carboxypeptidase (Wako) was added, and the plate was placed on a crystallizing dish in a 37 °C water bath for 1–2 h: **11A** and **18A** were digested with Endo-Lys to give spectra with intense peaks in the expected mass/charge (m/z) range by using 2 μl protein solution with a 1-h reaction time, while wt bZIP and **4A** required 4 μl protein solution and a 2-h reaction time.

Two microliters of water containing 5% TFA was added to the droplet to stop the enzyme reaction, and the droplet was allowed to dry at room temperature. Ten microliters ice-cold 0.1% TFA in water was placed on the dried spot and allowed to sit for 3 s and then pipetted off, and the spot was allowed to dry at room temperature. Two microliters saturated sinapinic acid was added to the spot and allowed to dry at room temperature. Finally, 0.5 μl of a 10 μM solution of thyrocalcitonin calibrant (3417.9 Da, Sigma) was added to the dried spot and allowed to dry at room temperature.

2.5. Mass spectrometry

Mass spectra were obtained on a PerSeptive Voyager-DE STR MALDI-TOF (Applied Biosystems) at the Center for Molecular Analysis, Carnegie Mellon University, as described previously [20]. Data Explorer software (version 3.4) from Applied Biosystems was used to work up data.

3. Results and discussion

Our alanine-rich bZIP mutants display strong propensity toward aggregation [7,8]; as such, they are very difficult to use in experimental manipulations, including MALDI-TOF

MS. We were able to obtain superior quality mass spectra for each of our four intact hydrophobic proteins by careful attention to concentrations of salts and denaturants, use of high matrix/protein molecular ratios (50,000:1 or 100,000:1), and utilization of the temperature-leap tactic (T-leap) [20]. T-leap is employed to prevent protein aggregation, occurring when the concentration of denaturant is reduced upon dilution of aliquots of protein stock solutions, containing 4 M urea, with matrix solution. Upon dilution at room temperature, protein aggregation is favored over proper refolding. However, by allowing slower refolding at 4 °C, protein is shunted toward the pathway leading to proper refolding; this accumulated nonaggregating intermediate is converted to the properly folded protein by warming to 37 °C [21].

Once high-quality, calibrated mass spectra were generated on the full intact bZIP proteins, it became evident that these proteins were truncated. The wt bZIP and **18A** were 10.31 and 10.37 Da less than expected, respectively, while **4A** and **11A** were 94.65 and 127.24 Da less than expected, respectively [20]. Protein translation is initiated by **N**-formyl methionine, which can be removed by MAP if the adjacent amino acid possesses a small side chain, as in our proteins [2,9,10,12]. At first glance, it appeared that the initiating Met on only two of the four proteins was removed, but this conclusion is unlikely given that the sequence for all four proteins is identical for the first 31 amino acids. To investigate the source(s) of these mass discrepancies, enzymatic digestion was performed on each protein followed by MALDI-TOF MS analysis. We used the endopeptidase Endo-Lys, which cleaves after lysine. Other commonly used carboxypeptidases, such as trypsin, cleave after both Lys and Arg, thereby creating more ambiguity in the identities of the generated fragments. Calculated masses for each digestion fragment and potential PTMs at the amino termini are listed in Tables 1 and 2.

By using an Endo-Lys carboxypeptidase on-plate digestion reaction reported by Kussmann et al. [22], we were able to vary systematically the amounts of both protein and enzyme, as well as reaction times, to optimize the intensities of the protein fragments. The enzymatic digestion reactions combined with MALDI proved that all four proteins were missing their initiating methionines. The spectra in Figs. 2 and 3 show N-terminal fragments of masses ranging from 3316 to 3320 Da. The expected mass of fragment 1 minus methionine is 3313.33 Da (Table 2), and therefore, removal of methionine is most in agreement with the results.

Also evident in Fig. 2 is adduct formation, which is adding masses of 43.10, 43.07, 43.56, and 41.61 Da to the N-terminal fragments of wt bZIP, **4A**, **11A**, and **18A**, respectively. Since the penultimate amino acid is glycine, one would expect the initiating Met to be cleaved. However, MALDI is unable to distinguish between the addition of an acetyl group (+42.037 Da) or a carbamyl group (+43.025 Da). The proteins were in contact with urea, beginning with isolation from cell debris through chromatographic purification. To eliminate or at least suppress carbamylation, new urea solutions were purified by deionization on mixed-bed ion-exchange resin. Fresh proteins were expressed and purified exclusively with deionized urea and digested on plate. As seen in the simpler, cleaner spectra shown in Fig. 3, adduct formation at the N-terminus is not present. From this, we conclude that carbamylation was the cause of adduct formation in Fig. 2, not acetylation.

During isolation and purification of expressed prokaryotic and eukaryotic proteins, high concentrations of denaturant, up to 8 M, are often necessary. Urea is typically used during protein preparation, because it is much less expensive than guanidine, another effective denaturant. To minimize the problems associated with urea degradation products, a simple procedure is to use purified, deionized urea solutions (see Materials and methods); an additional safeguard is to substitute guanidine for urea in the final storage buffer.

Table 1
Expected masses for each Endo-Lys digestion fragment

Peak	*m/z*	wt bZIP
	614.35	DPAALK[a]
	628.33	QLEQK[b] (pyroglutamate = loss of NH_3 to cyclize)
	645.36	QLEQK[b]
	675.40	LQRMK[a]
	1558.86	VLELTSDNDRLRK[b]
	1799.03	RARNTEAARRSRARK[a]
7	2442.14	RVEQLSRELDTLGGCGGYYYY[b]
1	3445.41	MGGSHHHHHHGMASMTGGQQMGRDLYDDDDK[c]

Peak	*m/z*	**4A**
	628.33	QLEQK[b] (pyroglutamate = loss of NH_3 to cyclize)
	645.36	QLEQK[b]
	675.40	LQRMK[a]
	819.30	DPDDDDK[c]
	1558.86	VLELTSDNDRLRK[b]
	2252.31	ARAAAARARNTAAARRSRARK[a]
7	2442.14	RVEQLSRELDTLGGCGGYYYY[b]
1	3445.41	MGGSHHHHHHGMASMTGGQQMGRDLYDDDDK[c]

Peak	*m/z*	**11A**
	819.30	DPDDDDK[c]
	943.52	AAAAAALEQK[a]
	1558.86	VLELTSDNDRLRK[b]
	2167.25	ARAAAARARNTAAARRSRAAK[a]
7	2442.14	RVEQLSRELDTLGGCGGYYYY[b]
1	3445.41	MGGSHHHHHHGMASMTGGQQMGRDLYDDDDK[c]

Peak	*m/z*	**18A**
	819.30	DPDDDDK[c]
	943.52	AAAAAALEQK[a]
	1558.86	VLELTSDNDRLRK[b]
	1695.93	AAAAAAAAANAAAAAAARAAK[a]
7	2442.14	RVEQLSRELDTLGGCGGYYYY[b]
1	3445.41	MGGSHHHHHHGMASMTGGQQMGRDLYDDDDK[c]

Peak numbers correlate with Figs. 2 and 3.
[a] From basic region sequence (see Fig. 1).
[b] From leucine zipper/linker sequence (see Fig. 1).
[c] From pTrcHis B sequence (see Fig. 1).

Table 2
Expected masses for potential modifications of N-terminal fragments

Modification	*m/z*	N-terminal fragment
Met	3445.41	MGGSHHHHHHGMASMTGGQQMGRDLYDDDDK
Acetyl-Met	3487.44	(acetyl)MGGSHHHHHHGMASMTGGQQMGRDLYDDDDK
Formyl-Met	3474.42	(formyl)MGGSHHHHHHGMASMTGGQQMGRDLYDDDDK
Carbamyl-Met	3488.43	(carbamyl)MGGSHHHHHHGMASMTGGQQMGRDLYDDDDK
DeMet	3313.33	GGSHHHHHHGMASMTGGQQMGRDLYDDDDK
Acetyl-DeMet	3355.36	(acetyl)GGSHHHHHHGMASMTGGQQMGRDLYDDDDK
Formyl-DeMet	3342.34	(formyl)GGSHHHHHHGMASMTGGQQMGRDLYDDDDK
Carbamyl-DeMet	3356.35	(carbamyl)GGSHHHHHHGMASMTGGQQMGRDLYDDDDK

We also observe up to four or five sequential additions of approximately 16 u, as shown in the insets in Figs. 2 and 3. This was attributed to formation of 2-oxoHis. Because our proteins possess a six-histidine tag that aids in purification, reactive species, such as metal-generated radicals, can oxidize the histidine side chains, thus producing a series of descending peaks varying in mass increments of 16 u. Indeed, Uchida and Kawakishi [23] found that 2-oxoHis was the main product in oxidative modification of proteins by free radicals, and that 2-oxoHis may serve as a biological marker for oxidative stress: for example, inactivation of Cu,Zn-superoxide dismutase occurs readily upon radical oxidation of His^{118} to 2-oxoHis in the active site. Because radicals are often generated in the presence of metal ions, and because the His_6 tag on our proteins is utilized to chelate a metal-ion affinity column during purification, generation of 2-oxoHis under these conditions is likely.

With the positive identification of demethioninylation and carbamylation by enzymatic digestion, we can rationalize the mass discrepancies for wt bZIP, **4A**, **11A**, and **18A**.

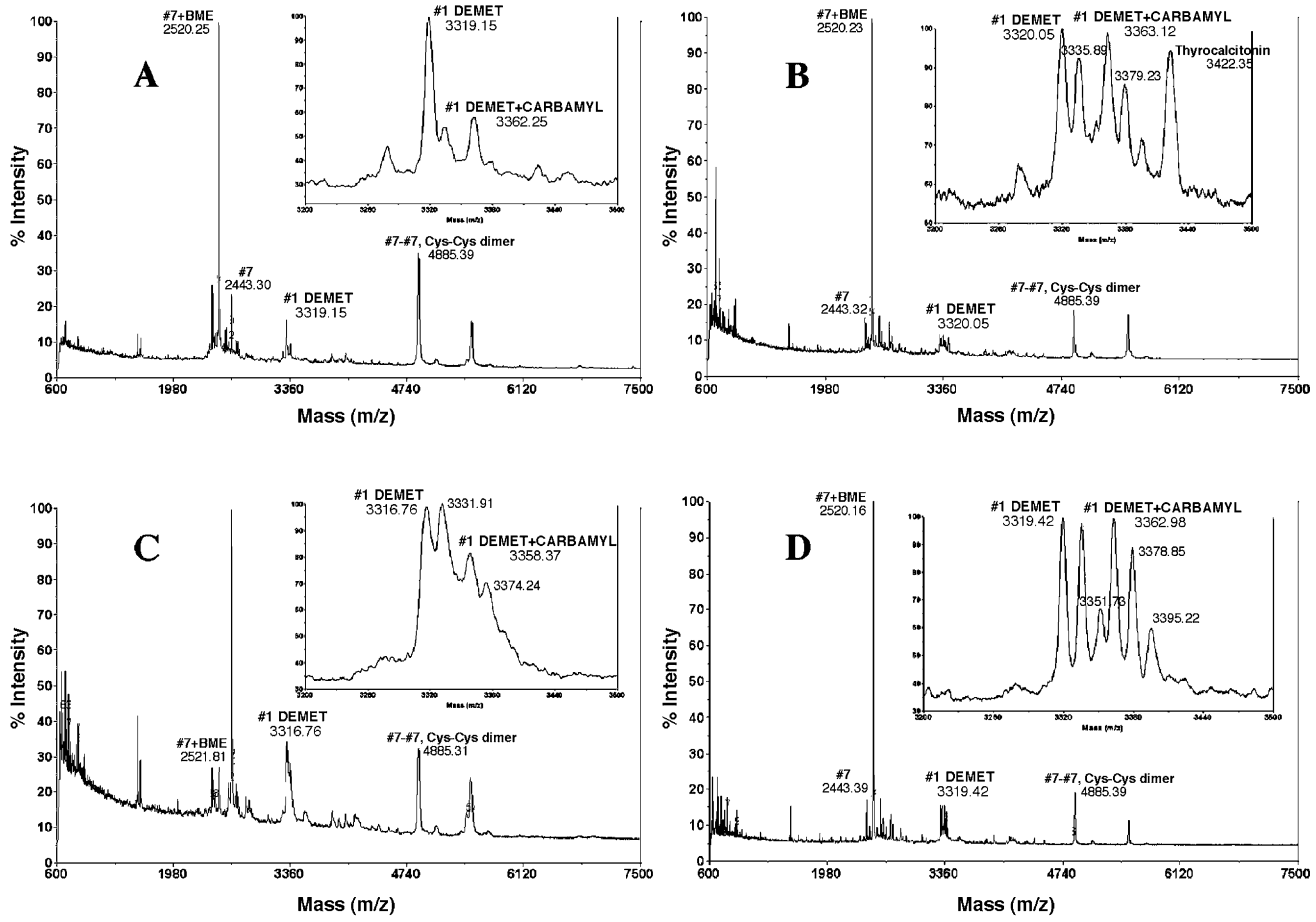

Fig. 2. MALDI-TOF spectra of Endo-Lys carboxypeptidase digestions of bZIP mutants showing presence of carbamylation. Insets shows expansion around peak 1 (see Table 1). Peaks are labeled according to Table 1. (A) wt bZIP. (B) **4A**. (C) **11A**. (D) **18A**.

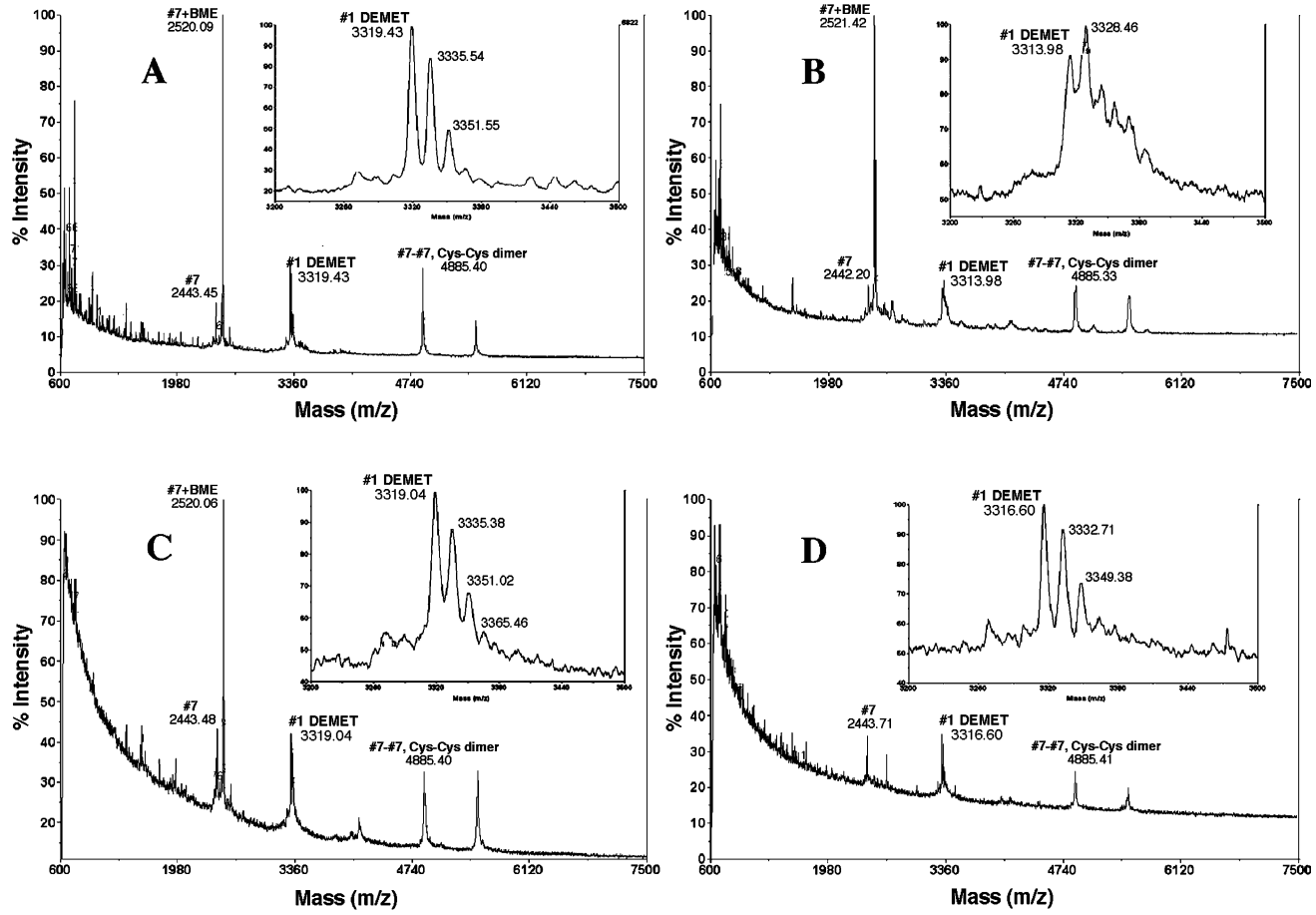

Fig. 3. MALDI-TOF spectrum of Endo-Lys carboxypeptidase digestions of bZIP mutants with no carbamylation present. Insets shows expansion around peak 1 (see Table 1). Peaks are labeled according to Table 1. (A) wt bZIP. (B) **4A**. (C) **11A**. (D) **18A**.

The observed mass of intact wt bZIP was 11,062.85 Da [20]. The expected mass from the DNA sequence is 11,073.16 Da. This difference has been shown to be due to demethioninylation, subtracting 132.08 Da from the calculated mass to give a mass of 10,941.08 Da. Additionally, wt bZIP showed addition of BME, adding 76.11 Da to give a mass of 11,017.19 Da. High concentrations of BME were used as a reductant in all protein isolation buffers. Because our proteins contain a cysteine near the carboxyl terminus, BME modification of the Cys thiol side chain is observed (#7 + BME in Figs. 2 and 3). We also observe disulfide dimers between cysteines (#7–#7, Cys–Cys dimer in Figs. 2 and 3). Our protein samples are maintained under reducing conditions until the second chromatographic purification by HPLC, which contains no reductant in the liquid phase. Although only the monomeric peak is collected during HPLC, no reductant is present during enzymatic digestion and MALDI-TOF MS, and therefore, equilibration occurs between free monomer, monomer plus BME adduct (#7 + BME), and dimer (#7–#7, Cys–Cys dimer). Carbamylation at the N-terminus, consistent with Fig. 2, adds 43.03 Da, yielding a mass of 11,060.22 Da. Therefore, the difference between the calculated and observed molecular masses is 0.023% (2.63 Da), well within the 0.01–0.1% range of mass error for MALDI-TOF MS.

The intact bZIP protein **4A** has an observed mass of 11,637.20 Da [20]. The expected mass is 11,731.85 Da. Removal of methionine gives mass of 11,599.77 Da. As seen in Fig. 2, carbamylation at the N-terminus further adjusts the mass to 11,643.07 Da. The percent error between the observed and calculated masses of **4A** is 0.050%. We note that in Fig. 2, one of the peaks corresponds to the C-terminal fragment released by Endo-Lys digestion containing the BME adduct at Cys (#7 + BME). In the MALDI-TOF spectrum of the *intact* **4A** mutant, no peak was observed corresponding to the BME adduct [20]. We note that achieving superior MALDI spectra of our intact bZIP proteins was a difficult challenge [20], involving close attention to salt and denaturant concentrations, matrix/analyte ratios, and the T-leap [21], all of which were critical to maintaining protein solubility. However, the smaller enzymatic digestion fragments were less prone to aggregation than the full proteins. In samples of **4A**, there is a mix of **4A** with or without the BME adduct at Cys, as well as disulfide exchange to give Cys–Cys dimers (Fig. 2). In the MALDI

MS of the intact protein, only those **4A** molecules with no BME adduct provided a strong signal under our conditions promoting protein solubility [20]. However, once **4A** is digested with Endo-Lys, the released fragments are more soluble, and we see both C-terminal fragments with or without the BME adduct.

The observed mass of intact **11A** is 11,158.61 Da [20]. The expected mass is 11,288.27 Da. Demethioninylation adjusts this mass to 11,156.19 Da, and, therefore, the difference between the observed and calculated masses is 0.022%. Note that in the spectrum in Fig. 2, there is a weak peak corresponding to the BME adduct on the **11A** C-terminal fragment, as well as a Cys–Cys dimer. In comparison, the BME adduct does not give a strong signal in the MALDI-TOF spectrum of the *intact* **11A** mutant [20]. Likewise, although the spectrum in Fig. 2 shows a peak corresponding to the carbamylated N-terminus of **11A** being released upon Endo-Lys digestion, the MALDI MS of intact **11A** does not show a peak corresponding to the carbamylated protein [20].

The measured mass of the intact bZIP **18A** is 10,806.39 Da [20]. The expected mass is 10,816.76 Da. The MALDI-TOF spectrum in Fig. 2 demonstrates that demethioninylation is occurring, modifying the mass to 10,684.68 Da. The signal from the BME adduct on **18A** is strong, thereby adjusting the mass to 10,760.79 Da. Carbamylation at the N-terminus gives an expected mass of 10,803.82 Da. The percent mass error is 0.024%. Similar to the discussion above, the intact **18A** mutant gives a strong signal corresponding to carbamylation at the N-terminus and BME adduct at Cys near the C-terminus [20]. However, after Endo-Lys digestion on **18A**, peaks corresponding to fragments without carbamylation and without BME adduct are observed.

This MALDI analysis suggests two noteworthy observations: the first, more obvious observation is that detailed information about PTMs of proteins can be gained by MALDI MS in conjunction with enzymatic digestion, thereby providing a powerful tool for confirming the identity of a protein and establishing the presence of isoforms. When PTMs block the amino terminus, sequencing information can be gained by MALDI MS, for Edman degradation cannot be performed in these cases. A second observation is that although the MALDI-TOF spectra of the full, intact bZIP proteins showed one strong peak for each mutant [20], leading us to conclude that our samples were homogeneous as far as the PTMs were concerned, the Endo-Lys digestion data show that, in fact, disulfide exchange had occurred to give the free Cys, the BME–Cys adduct, and the Cys–Cys dimer forms. In addition, the Endo-Lys digestions show that both carbamylated and decarbamylated N-termini exist in the same sample. For comparison, we provide MALDI-TOF spectra in Fig. 3 demonstrating that by using purified, deionized urea in all protein preparation steps, carbamylation at the N-termini no longer occurs; in particular, the insets in Fig. 3 show the absence of carbamylated N-terminal fragments, and the spectra in Fig. 3 are cleaner than the corresponding spectra in Fig. 2. In addition, this experiment proves that acetylation at the N-termini is not occurring.

We emphasize that gaining MALDI MS for the *intact* bZIP mutants was extremely challenging, and under those conditions, promoting the intact proteins' solubility and nonaggregation discussed above, one isoform of each protein provided an intense signal peak [20]. However, the enzymatic digestion fragments were more soluble and amenable to MALDI-TOF analysis than the full proteins, and therefore, detection of all fragments and modifications was feasible. MALDI-TOF MS analysis coupled with the on-plate Endo-Lys digestion allowed mass discrepancies to be assigned and PTMs to be ascertained. As research on proteins—even especially intractable, hydrophobic proteins—increases, MALDI MS will become a more common and useful technique in physical characterization of protein samples and mixtures.

Acknowledgements

We are grateful to Jim Noll for technical expertise and to Steve Weber for critical reading of the manuscript. We acknowledge the Center for Molecular Analysis, Carnegie Mellon University, for use of the MALDI-TOF mass spectrometer (NSF CHE-9808188). This work was supported by a grant from the National Science Foundation (CAREER MCB-9733410) to J.A.S. and the University of Pittsburgh.

References

[1] F. Hillenkamp, M. Karas, R.C. Beavis, B.T. Chait, Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers, Anal. Chem. 63 (1991) 1193–1203.
[2] S.M. Arfin, R.A. Bradshaw, Cotranslational processing and protein turnover in eukaryotic cells, Biochemistry 27 (1988) 7979–7984.
[3] D.E. Hill, I.A. Hope, J.P. Macke, K. Struhl, Saturation mutagenesis of the yeast *his3* regulatory site: requirements for transcriptional induction and for binding by GCN4 activator protein, Science 234 (1986) 451–457.
[4] W.H. Landschulz, P.F. Johnson, S.L. McKnight, The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins, Science 240 (1988) 1759–1764.
[5] T.E. Ellenberger, C.J. Brandl, K. Struhl, S.C. Harrison, The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted α helices: crystal structure of the protein–DNA complex, Cell 71 (1992) 1223–1237.
[6] P. König, T.J. Richmond, The X-ray structure of the GCN4-bZIP bound to ATF/CREB site DNA shows the complex depends on DNA flexibility, J. Mol. Biol. 233 (1993) 139–154.
[7] A.R. Lajmi, T.R. Wallace, J.A. Shin, Short, hydrophobic, alanine-based proteins based on the bZIP motif: overcoming inclusion body formation and protein aggregation during overexpression, purification, and renaturation, Protein Exp. Purif. 18 (2000) 394–403.
[8] A.R. Lajmi, M.E. Lovrencic, T.R. Wallace, R.R. Thomlinson, J.A. Shin, Minimalist, alanine-based, helical protein dimers bind to specific DNA sites, J. Am. Chem. Soc. 122 (2000) 5638–5639.

[9] S. Huang, R.C. Elliott, P.-S. Liu, R.K. Koduri, J.L. Weickmann, H.-H. Lee, L.C. Blair, P. Ghosh-Dastidar, R.A. Bradshaw, K.M. Bryan, B. Einarson, R.L. Kendall, K.H. Kolacz, K. Saito, Specificity of cotranslational amino-terminal processing of proteins in yeast, Biochemistry 26 (1987) 8242–8246.
[10] P. Wingfield, P. Graber, G. Turcatti, N.R. Movva, M. Pelletier, S. Craig, K. Rose, C.G. Miller, Purification and characterization of a methionine-specific aminopeptidase from *Salmonella typhimurium*, Eur. J. Biochem. 180 (1989) 23–32.
[11] M.L. Klein, T.D. Bartley, J.M. Davis, D.W. Whiteley, H.S. Lu, Isolation and structural characterization of three isoforms of recombinant consensus α interferon, Arch. Biochem. Biophys. 276 (1990) 531–537.
[12] C. Miller, K.L. Strauch, A.M. Kukral, J.L. Miller, P.T. Wingfield, G.J. Mazzei, R.C. Werlen, P. Graber, N.R. Movva, N-terminal methionine-specific peptidase in *Salmonella typhimurium*, Proc. Natl. Acad. Sci. U. S. A. 84 (1987).
[13] J.M. Adams, On the release of the formyl group from nascent protein, J. Mol. Biol. 33 (1968) 571–589.
[14] T.E. Creighton, Proteins: Structures and Molecular Properties, 2nd edn., Freeman, New York, 1993.
[15] C. Schöneich, Mechanisms of metal-catalyzed oxidation of histidine to 2-oxo-histidine in peptides and proteins, J. Pharm. Biomed. Anal. 21 (2000) 1093–1097.
[16] S.A. Lewisch, R.L. Levine, Determination of 2-oxohistidine by amino acid analysis, Anal. Biochem. 231 (1995) 440–446.
[17] E. Shacter, Quantification and significance of protein oxidation in biological samples, Drug Metab. Rev. 32 (2000) 307–326.
[18] G.R. Stark, Modification of proteins with cyanate, Methods Enzymol. 11 (1967) 590–594.
[19] G.R. Stark, Use of cyanate for determining NH_2-terminal residues in protein, Methods Enzymol. 11 (1967) 125–138.
[20] G.H. Bird, A.R. Lajmi, J.A. Shin, Manipulation of temperature improves solubility of hydrophobic proteins and cocrystallization with matrix for analysis by MALDI-TOF mass spectrometry, Anal. Chem. 74 (2002) 219–225.
[21] Y. Xie, D.B. Wetlaufer, Control of aggregation in protein refolding: the temperature-leap tactic, Protein Sci. 5 (1996) 517–523.
[22] M. Kussmann, E. Nordhoff, H. Rahbek-Nielsen, S. Haebel, M. Rossel-Larsen, L. Jakobsen, J. Gobom, K. Mirgorodskaya, A. Kroll-Kristensen, L. Palm, P. Roepstorff, Matrix-assisted laser desorption/ionization mass spectrometry sample preparation techniques designed for various peptide and protein analytes, J. Mass Spectrom. 32 (1997) 593–601.
[23] K. Uchida, S. Kawakishi, Identification of oxidized histidine generated at the active site of Cu,Zn-superoxide dismutase exposed to H_2O_2, J. Biol. Chem. 269 (1994) 2405–2410.
[24] J.A. Shin, Specific DNA binding peptide derivatized solid support, Bioorg. Med. Chem. Lett. 7 (1997) 2367–2372.
[25] W. Keller, P. König, T.J. Richmond, Crystal structure of a bZIP/DNA complex at 2.2 Å: determinants of DNA specific recognition, J. Mol. Biol. 254 (1995) 657–667.