Retention of Thiol Protons in Two Classes of Protein Zinc Ion Coordination Centers

Daniele Fabris, Joseph Zaia, Yetrib Hathout, and Catherine Fenselau*

Department of Chemistry and Biochemistry
University of Maryland, Baltimore County
Baltimore Maryland 21250

Received August 15, 1996

The "galvanization of biology"[1] in recent years has promoted the exploration of many important questions about the roles of zinc clusters in proteins. These clusters act to maintain protein structures, regulate zinc homeostasis, assist catalysis, and participate in signal transduction.[1,2] Advances in NMR and mass spectrometry have allowed detailed studies of the primary and tertiary structures of proteins that contain zinc clusters and of ion flux during their interactions with drugs and with other biopolymers.[3,4] Most catalytic centers and some structural examples[5] involve both protein residues and external ligands (usually water). However, zinc ions can also be chelated entirely by functional groups provided by the protein, in tetrahedral ligands comprising only cysteine thiol groups or combinations of two or three cysteine thiols with the imidazole groups in histidine side chains. Examples of the former are found in the metallothionein family, in which 7 divalent zinc ions are chelated by 20 cysteine residues.[6] Examples of the latter are the CCHC zinc finger motifs that occur in the HIV nucleocapsid protein[7] and the gene 32 protein from bacteriophage T4.[8]

One of the structural questions that has not yet been resolved for metallothionein and proteins containing CCHC arrays is whether all of the thiol groups in the ligand field are deprotonated. Because they exchange readily with solvent, these protons have been difficult to resolve by NMR or X-ray crystallography. The determination of molecular mass by mass spectrometry might be expected to provide relevant information (i.e., a proton count). The issue of thiol deprotonation has been discussed for CCCC arrays,[9,10] however mass measurements have thus far been indeterminate.[11,12]

This paper reports accurate mass measurements made using electrospray mass spectrometry of rabbit liver metallothionein 2a (MT),[13] a recombinant analog [4-55] of HIV nucleocapsid p7 (NcP7),[14] a synthetic CCHC zinc finger (F1) comprising the first 18 residues of NcP7,[7] and the recombinant binding domain [16-253] from the gene 32 protein from bacteriophage T4.[15] The molecular weights of these proteins, ranging between 2000 and 27 000 Da, were determined on a JEOL HX110/HX110 double-focusing mass spectrometer equipped with a thermally assisted electrospray ion source.[16] Mass assignments were made by manually interpolating target ion masses from those of reference ions bracketing the unknown on both sides. Optimally, reference ions will fall within 500 m/z units of the unknown peak, and will have a similar charge state. All holo proteins were introduced in 10 mM ammonium acetate pH 7.0. Apo or oxidized (demetalated) samples were introduced at pH 0.5 (MT) or 3.5.

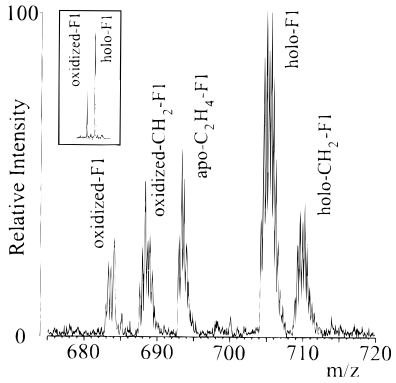

Figure 1. Partial electrospray ionization mass spectrum of the product mixture from the reaction of zinc-complexed F1 with (*p*-nitrobenzene)-sulfonic acid methyl ester. The charge state of all the ions shown is +3. The inset shows the spectrum of the starting material.

The molecular weights determined for this group of zinc-binding proteins are summarized in Table 1 along with molecular weights of apoproteins and oxidized proteins. The table also lists measurements of three standard proteins made to confirm experimental reliability. In all examples studied, the observed molecular masses for holo-metalloproteins exceed those predicted for complete deprotonation. The calculated masses listed in the table are for species that contain two (deprotonated) thiolate groups per zinc ion chelated. In all cases, only enough protons are lost to formally balance the charges of divalent zinc ions (i.e., 14 for MT, which binds 7 metal ions, 4 for NcP7, which chelates 2 zinc ions, and 2 for the synthetic finger and gene 32 protein, which each bind a single divalent zinc ion).

Although partial retention of protons within the metal clusters is the simplest explanation, molecular mass determinations do not exclude the possibility of charge adjustments elsewhere in each protein. Others have suggested that the electrospray process might reprotonate the zinc arrays[11] (6 protons for metallothionein). In order to test both of these possibilities the zinc-complexed octadecamer F1 was treated with (*p*-nitrobenzene)sulfonic acid methyl ester under conditions that produced primarily monoalkylation (30 min, pH 7.0). The mass spectrum (averaged from 100 scans) revealed an increase in mass of 14 Da, with retention of the zinc ion (Table 1). When the reaction time was extended so that dimethylated protein could also be observed, the product mixture contained the monomethylated zinc cluster; however, the dimethylpeptide was observed only as a demetalated species, the reduced apopeptide plus C_2H_4 (Table 1). Figure 1 shows the electrospray mass spectrum of the product mixture from this reaction. The inset in Figure 1 is the spectrum of the starting material, which contained both holo-F1 and oxidized (demetalated) F1. Both of these peptides were found to be partially monomethylated. The monoisotopic mass (Table 1) of the dimethylated peptide characterizes it as reduced and thus derived from the holopeptide by demetalation.

The monomethylpeptide was sequenced by high-energy collisional activation[17] and methylation was found to have taken

(1) Berg, J. M.; Shi, Y. *Science* **1996**, *271*, 1081−1085.
(2) Vallee, B. L.; Auld, D. S. *Biochemistry* **1990**, *29*, 5647−5659.
(3) Blake, P. R.; Summers, M. F. *Adv. Biophys. Chem.* **1994**, *4*, 1−30.
(4) Zaia, J.; Jiang, L.; Han, M. S.; Tabb, J. R.; Wu, Z.; Fabris, D.; Fenselau, C. *Biochemistry* **1996**, *35*, 2830−2835.
(5) Brader, M. L.; Dunn, M. F. *Trends Biol. Sci.* **1991**, *16*, 341−345.
(6) Stillman, M. J. *Coord. Chem. Rev.* **1995**, *144*, 461−511.
(7) South, T. L.; Blake, P. R.; Sowder, R. C.; Arthur, L. O.; Henderson, L. E.; Summers, M. F. *Biochemistry* **1990**, *29*, 7786−7789.
(8) Karpel, R. L. In *The Biology of Nonspecific DNA-Protein interactions*; Revson, A., Ed.; CRC Press, Inc.: Boca Raton, FL, 1990; pp 103−130.
(9) Myers, L. C.; Wagner, G.; Verdine, G. L. *J. Am. Chem. Soc.* **1995**, *117*, 10749−10750.
(10) Wilker, J. J.; Lippard, S. J. *J. Am. Chem. Soc.* **1995**, *117*, 8682−8683.
(11) Allen, M. H.; Hutchens, T. W. *Rapid Commun. Mass Spectrom.* **1992**, *6*, 308−312.
(12) Witkowska, H. E.; Shackleton, C. H.; Dahlman-Wright, K.; Kim, J. Y.; Gustafsson, J. A. *J. Am. Chem. Soc.* **1995**, *1177*, 3319−3325.
(13) Yu, X.; Wojciechowski, M.; Fenselau, C. *Anal. Chem.* **1993**, *65*, 1355−1359.
(14) Hathout, Y.; Fabris, D.; Sowder, R. C., II; Henderson, L. E.; Fenselau, C. *Drug Metab. Dispos.* **1996**, *24*. In press.
(15) Zaia, J.; Fabris, D.; Karpel, R; Fenselau, C. Submitted.
(16) Cody, R. B.; Tamura, J.; Musselman, B. D. *Anal. Chem.* **1992**, *64*, 1561−1570.

S0002-7863(96)02863-6 CCC: $12.00

Table 1. Accurate Mass Measurements on Zinc-Binding Proteins

protein	formula	calcd[a]	obsd[b] (Da)
aprotinin control	$C_{284}H_{432}N_{84}O_{79}S_7$	6511.5	6511.5 ± 0.1[c]
apo-MT	$C_{226}H_{382}N_{72}O_{84}S_{21}$	6125.3	6125.4 ± 0.1
Zn_7-MT	$C_{226}H_{368}N_{72}O_{84}S_{21}Zn_7$	6568.9	6568.7 ± 0.1
apo-NcP7 [4−55]	$C_{250}H_{411}N_{91}O_{70}S_7$	6036.0	6036.2 ± 0.2
holo-NcP7 [4−55]	$C_{250}H_{407}N_{91}O_{70}S_7Zn_2$	6162.8	6162.9 ± 0.1
oxid.NcP7 [4−55] (three S−S)	$C_{250}H_{405}N_{91}O_{70}S_7$	6030.0	6030.3 ± 0.2
gene 32 protein [16−253]	$C_{1192}H_{1829}N_{309}O_{366}S_{11}Zn$	26762.6	26762.1 ± 0.4
dynorphin A [1−13] control	$C_{75}H_{126}N_{24}O_{15}$	1603.0	1602.9 ± 0.1
α-endorphin control	$C_{77}H_{120}N_{18}O_{26}S$	1744.8	1744.8 ± 0.1
holo-F1	$C_{84}H_{137}N_{31}O_{23}S_3Zn$	2107.9	2107.7 ± 0.1
apo-F1	$C_{84}H_{139}N_{31}O_{23}S_3$	2046.0	2045.8 ± 0.1
oxid-F1 (one S−S)	$C_{84}H_{137}N_{31}O_{23}S_3$	2044.0	2043.8 ± 0.1
holo-CH_2-F1	$C_{85}H_{139}N_{31}O_{23}S_3Zn$	2121.9	2121.8 ± 0.1
apo-C_2H_4-F1	$C_{86}H_{143}N_{31}O_{23}S_3$	2074.0	2073.8 ± 0.1
oxid-CH_2-F1 (one S−S)	$C_{85}H_{139}N_{31}O_{23}S_3$	2058.0	2057.7 ± 0.2

[a] The neutral masses of all metalloproteins are calculated with loss of two protons per zinc ion chelated. [b] The first six average mass measurements were made using bovine insulin to provide bracketing reference ions. Ions formed from α-chymotrypsinogen at three charge states were used to bracket two charge states of gene 32 protein [16−253]. The final eight monoisotopic mass measurements used porcine renin substrate tetradecapeptide [1−14] to provide bracketing reference ions. [c] Standard deviations are calculated from the multiple charge states in each spectrum.

place on Cys-6 and Cys-16, (see footnote on Supporting Information) in a ratio estimated from fragment ion intensities as 1.8:1. Any methylation at Cys-3 was too limited to be detected. We suggest that the first methyl group is bound in the zinc center by a thiolate group that also has a high equilibrium constant for protonation. We also conclude that one, but not two, thiol groups or thiol ethers may be ligands in the CCHC array.

It has been previously shown that 1 and 2 equiv of a therapeutic mustard can be covalently bonded to cysteine side chains in metallothionein with retention of all seven metal ions.[4] Although an accurate average mass was not determined for this product, participation by thiol ethers in the ligand field is consistent with the observations reported here for the F1 CCHC array. It is also relevant that methylation of CCCC arrays occurs autocatalytically with retention of zinc ions, in the *Escherichia coli* Ada protein[18] and chemically in a 17 kDa Ada fragment.[9] Proton sharing has also been reported between thiolate groups in zinc fingers and nearby protons on the amide backbone.[3]

In summary, accurate mass determinations of rabbit liver metallothionein 2a, of analogs of the nucleocapsid protein from HIV, and of the binding domain of the gene 32 protein from bacteriophage T4 show that two protons are lost for each zinc ion complexed. This and other[4,9] alkylation studies indicate that thiol groups in a zinc ion coordination center may be alkylated and still retain chelating capability. In the CCHC array, zinc ions are complexed as long as sufficient thiolate anions are available to match the formal charge of each divalent metal ion. Further studies will address protonation of CCCC and CCHH fingers and examine the effects of different metal ions.

Acknowledgment. The authors thank M. F. Summers for discussions and for a sample of F1, R. L. Karpel for a sample of gene 32 protein core domain, and L. E. Henderson for a sample of NcP7 [4-55]. The work was supported in part by a grant from National Institutes for Health GM21248. Salary support for Joseph Zaia was provided by NIH Grant T32ES07263.

Supporting Information Available: The high-energy collisionally activated tandem mass spectrum of methyl-F1 (1 page). See any current masthead page for ordering and Internet access instructions.

JA962863R

(17) Biemann, K. *Biomed. Mass Spectrom.* **1988**, *16*, 99−111.
(18) Myers, L. C.; Terranova, M. P.; Ferentz, A.; Wagner, G.; Verdine, G. L. *Science* **1993**, *261*, 1164−1167.