Anal. Chem. **2000,** *72,* 574−579

Mass Spectrometric Analysis of Mercury Incorporation into Proteins for X-ray Diffraction Phase Determination

Steven L. Cohen,*,† Júlio C. Padovan, and Brian T. Chait

The Rockefeller University, New York, New York 10021

Heavy-atom incorporation is an essential and often rate-limiting step in the determination of phases for X-ray diffraction studies of protein structures. Until the present, there has been no practical method (short of the X-ray diffraction experiment itself) to judge the success and extent of incorporation. Here we show that mass spectrometry is an effective tool for determining the extent of heavy-atom incorporation in proteins. In particular, we demonstrate the utility of matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) and electrospray ionization mass spectrometry (ESI-MS) for assaying mercury derivatization of cysteinyl thiol groups in proteins. Each of these mass spectrometric methods has advantages and drawbacks. ESI-MS provides a more accurate quantitative measurement of the extent of mercury incorporation, while MALDI-MS provides a useful lower limit to the level of mercury incorporation. Conversely, MALDI-MS does not require removal of excess derivatization reagents, salts and buffers, thus permitting facile analysis of single protein crystals as well as rapid, semiquantitative evaluation of the extent of protein mercuration. The approaches described in the present paper have contributed to the successful X-ray analyses of several noteworthy protein structures.

Protein X-ray crystallography provides an unparalleled means for obtaining atomic-resolution structures of proteins and their complexes. Such determinations are formidable multidisciplinary undertakings, encompassing tasks such as protein cloning, expression and purification procedures, crystallization trials, and computational analyses. Included in these tasks is the phase-determination stage, an obligate step reached once a protein crystal yields a suitable quality X-ray diffraction pattern. Knowledge of the X-ray diffraction phases, along with the diffraction intensities, permits construction of an electron-density map and ultimately a three-dimensional model of the protein. Phase determination is a challenging process that has been formally addressed by several well-known approaches.[1−5] For proteins with completely new folds, the phase problem is generally resolved by the introduction of a heavy atom (e.g., mercury, selenium) into the protein. Unambiguous phase determination requires that the incorporation be made efficiently and with minimal perturbation of the native protein fold. Unfortunately, successful incorporation of heavy atoms into proteins is often a hit or miss procedure, demanding time-consuming screening of multiple combinations of reagents and conditions of derivatization.[6−8] For much of the history of protein X-ray crystallography there has been no practical means by which to measure heavy-atom incorporation into proteins—except, of course, through the X-ray diffraction analysis itself. Certainly, knowing the extent to which derivatization has occurred prior to X-ray analysis would greatly assist in the screening process for identifying suitable heavy-atom derivatized protein candidates. Under such circumstances, X-ray analysis becomes unencumbered by poorly or nonderivatized candidates. In this paper we show that mass spectrometry, a technique that has become an indispensable tool for protein characterization,[9−11] can be applied to the analysis of heavy-atom incorporation into proteins. Using mercury as the heavy atom, we demonstrate that mass spectrometry provides rapid, sensitive information concerning the extent of heavy-atom derivatization. The approaches described in the present paper have already contributed to the successful X-ray analyses of several noteworthy protein structures.[12−14]

EXPERIMENTAL SECTION

Materials. The proteins used in the present study were human hemoglobin (Sigma, St. Louis, MO), enolase I from *Saccharomyces cerevisiae* (Sigma), and three recombinant proteins, i.e., peptidyl-

* Corresponding author: (phone) 215-652-0517; (fax) 215-652-7722; (e-mail) steven_cohen@merck.com.

† Present Address: Merck Research Laboratory, West Point, PA 19486.

(1) Green, D. W.; Ingram, V. M.; Perutz, M. F. *Proc. R. Soc. London, Ser. A* **1954**, *255*, 287−307.
(2) Blow, D. M.; Crick, F. H. C. *Acta Crystallogr.* **1959**, *12*, 794−802.
(3) Watenpaugh, K. D. *Methods Enzymol.* **1985**, *115*, 3−15.
(4) Hendrickson, W. A. *Science (Washington, D.C.)* **1991**, *254*, 51−8.
(5) Ogata, C. M. *Nat. Struct. Biol.* **1998**, (Suppl. 5), 638−40.
(6) Blake, C. C. *Adv. Protein Chem.* **1968**, *23*, 59−120.
(7) Blundell, T. L.; Johnson, L. H. Preparation of heavy atom derivatives. In *Protein crystallography*; Academic Press: New York, 1976; Chapter 8, pp 183−239.
(8) Rould, M. A. *Methods Enzymol.* **1997**, *276*, 461−72.
(9) Mann, M.; Wilm, M. *Trends Biochem. Sci.* **1995**, *20*, 219−24.
(10) Roepstorff, P. *Curr. Opin. Biotechnol.* **1997**, *8*, 6−13.
(11) Fenselau, C. *Anal. Chem.* **1997**, *69*, 661A−5A.
(12) Krishna, T. S.; Fenyö, D.; Kong, X. P.; Gary, S.; Chait, B. T.; Burgers, P.; Kuriyan, J. *J. Mol. Biol.* **1994**, *241*, 265−8.
(13) Xie, X.; Kokubo, T.; Cohen, S. L.; Mirza, U. A.; Hoffmann, A.; Chait, B. T.; Roeder, R. G.; Nakatani, Y.; Burley, S. K. *Nature (London)* **1996**, *380*, 316−22.
(14) Doyle, D. A.; Cabral, J. M.; Pfuetzner, R. A.; Kuo, A.; Gulbis, J. M.; Cohen, S. L.; Chait, B. T.; MacKinnon, R. *Science (Washington, D.C.)* **1998**, *280*, 69−77.

10.1021/ac990938e CCC: $19.00
Published on Web 01/05/2000

prolyl cis−trans isomerase FK506-binding protein 12 (FKBP12), neuronal ventral antigen protein-1 (Nova1), and the potassium channel from *Streptomyces lividans* (KcsA). Constructs of the three recombinant proteins were kindly provided by Drs. John Kuriyan (FKBP12), Stephen K. Burley (Nova1), and Roderick MacKinnon (KcsA). FKBP12 is the full-length protein without the N-terminal start methionine (amino acid residues 2−108; 107 residues). The Nova1 construct consists of the KH3 domain (amino acid residues 430−501 with four additional residues, GAHM, at the N-terminus; 76 residues). KcsA consists of the membrane-spanning portion of the protein (amino acid residues 1−125). Hemoglobin contains one cysteinyl residue in the α chain and two in the β chain, and enolase has only one cysteinyl residue. The three recombinant proteins each contain a single cysteinyl residue—Cys 23 (wild type) in FKBP12, Ala37Cys in Nova1, and Ala32Cys in KcsA, the latter two by site-directed mutagenesis. (Throughout this paper KcsA and Nova1 will refer to the respective cysteine-containing protein mutants). The mercurating reagents were ethylmercury phosphate (EMP; $C_2H_5HgH_2PO_4$, (CAS no. 2235-25-8) Noah Technologies, San Antonio, TX), *p*-(chloromercuri)phenylsulfonate, monosodium salt (*p*CMPS; $C_6H_4ClHgSO_3Na$ (CAS no. 14110-97-5) Sigma), and methylmercury chloride (MMC; CH_3HgCl, (CAS no. 115-09-3) Alfa Aesar, Ward Hill, MA). The MALDI matrixes were α-cyano-4-hydroxy-cinnamic acid (4hcca; Aldrich) and sinapinic acid (SA; Aldrich). The matrix 4hcca was recrystallized from ethanol, whereas SA was used without further purification. All organic solvents were HPLC grade or better. Water was supplied from a MilliQ UV Plus water purification system (Millipore, Bedford, MA). Formic acid (88%) was from Fisher Scientific (Springfield, NJ), trifluoroacetic acid (TFA; sequencing grade) was from Pierce (Rockford, IL), and acetic acid was from Aldrich. Measurements of the pH of matrix solutions were performed at room temperature with a combined pH microelectrode (model MI-415; Microelectrodes, Inc., Bedford, NH) and pH meter (model PHM 95; Radiometer, Copenhagen, Denmark).

Preparation of Mercury-Derivatized Proteins. Mercury derivatization reactions carried out in solution were performed in 0.65-mL microcentrifuge tubes. *Mercurating compounds are highly toxic and should be handled carefully with gloves at all times. All reactions were performed in a hood.* FKBP12 and Nova1 were derivatized by combining small aliquots (∼20 μL) of protein solutions (30−50 μM) with excess mercurating reagent [usually 15−20-fold (molar) over cysteine] prepared in 50 mM HEPES buffer (pH 8) and 100 mM NaCl for approximately 1 h at room temperature. Longer reaction times did not alter the mass spectrometric results, and thus we assumed that the protein was fully derivatized. Hemoglobin, enolase I, FKBP12, and Nova1 remained soluble following derivatization, although mercuration has been observed to cause precipitation of other proteins. Adjusting the solution pH, changing the choice of buffer, or slowly diffusing the mercurating reagent to the protein solution through dialysis are common ways to minimize sample-handling difficulties.[6,7,15] Mercuration of the potassium channel protein (KcsA) was performed by protein crystal soaking methods.[7,14] Solid methylmercury chloride was added to mother liquor crystallization solutions containing KcsA crystals (5−10 mg/mL KcsA in 45% poly(ethylene glycol) 400, 50 mM Tris-HCl (pH 7.5), 200 mM $CaCl_2$, 150 mM KCl, 2 mM dithiothreitol, 5 mM lauryldimethylamine *N*-oxide). Soaking proceeded at room temperature for 2 to 4 h.

MALDI-MS − Sample Preparation. MALDI matrix solutions were prepared as described previously[16,17] and as indicated in the figure captions. For experiments with FKBP12 and Nova1, 0.5−1-μL aliquots of derivatized protein were withdrawn directly from the derivatization reaction (see above) and added to 15−30 μL of matrix solution in which the final protein concentration was approximately 1 μM. Unreacted mercurating reagent was not removed, and its presence was determined not to influence the MALDI results (data not shown). The mixture was briefly vortexed and spotted (0.5−1 μL) onto the MALDI probe. For the crystal-soaking experiments, protein crystals were retrieved from the mother liquor with a microloop under a microscope. The crystals were washed in a drop of protein-free mother liquor solution placed on a glass coverslip. A single washed protein crystal (∼1 pmol protein) was dropped into a small aliquot (∼0.75 μL) of 0.1% aqueous TFA/acetonitrile (1:1 v/v) into which the crystal promptly dissolved. The solution was immediately retrieved with a micropipettor and mixed with an equal volume of matrix solution (4hcca in 0.1% aqueous TFA/acetonitrile (2:1 v/v)) placed on the MALDI probe. (The protein crystal can also be dissolved directly into matrix solution). Solvent was allowed to evaporate at room temperature, leaving a bed of matrix crystals ready for MALDI-MS analysis. To remove excess involatile salts and biochemical additives, the matrix crystals were washed on the probe with cold 0.1% aqueous TFA solution prior to MALDI analysis.

MALDI-MS − Instrumentation and Data Collection. Several different MALDI-MS instruments were used in the experiments. The principal device was a MALDI time-of-flight (TOF) mass spectrometer (model STR, PE Biosystems, Foster City, CA) operating in linear delayed extraction mode. This instrument uses a nitrogen laser that delivers pulses of ultraviolet light (wavelength = 337 nm) at 2 Hz to the matrix spots, each pulse yielding a full mass spectrum. Spectra from one hundred individual laser shots were averaged (using 10-ns data channel widths) with software provided by the manufacturer. The spectra were smoothed and further analyzed using the software program M-over-Z. (http://www.proteometrics.com and http://prowl.rockefeller.edu). Two other linear delayed-extraction MALDI-TOF instruments, constructed in the laboratory,[18,19] were also used. Details of their operation can be found elsewhere.[16] These devices were equipped with UV lasers (Neodymium-YAG, wavelength = 355 nm). One of the instruments was also equipped with an infrared laser (wavelength = 2.94 μm), providing a means to compare IR- and UV-MALDI results from the same matrix crystals.[20,21]

ESI-MS − Sample Preparation. In contrast to the sample preparation of MALDI-MS, ESI-MS required thorough removal of the mercury-derivatization compounds as well as salts and buffers. One nanomole of mercury-derivatized protein was injected into a reverse-phase C_4 cartridge (Michrom Bioresources, Inc., Auburn, CA) previously equilibrated in 0.1% aqueous acetic acid.

(15) Petsko, G. A. *Methods Enzymol.* **1985**, *114*, 147−56.
(16) Cohen, S. L.; Chait, B. T. *Anal. Chem.* **1996**, *68*, 31−7.
(17) Beavis, R. C.; Chait, B. T. *Methods Enzymol.* **1996**, *270*, 519−51.
(18) Beavis, R. C.; Chait, B. T. *Rapid Commun. Mass Spectrom.* **1989**, *3*, 233−7.
(19) Beavis, R. C.; Chait, B. T. *Anal. Chem.* **1990**, *62*, 1836−40.
(20) Zhang, W.; Niu, S.; Chait, B. T. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 879−84.
(21) Niu, S.; Zhang, W.; Chait, B. T. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 1−7.

Excess reagents were removed by washing the cartridge with 2 mL of deionized water. The purified derivatized protein was eluted with 100 μL of methanol to yield a 10 μM stock solution. We recommend not concentrating or drying, since these procedures produced dimerization of the underivatized proteins. For ESI-MS analysis, stock solutions were diluted either in deionized water (pH 5.6), 5% (v/v) aqueous acetic acid solution (pH 2.4), or 0.1% (v/v) aqueous TFA solution (pH 1.6) to a final concentration of 0.2 μM. Underivatized proteins were treated exactly as described above.

ESI-MS − Instrumentation and Data Collection. ESI-MS analysis was performed using an electrospray triple quadrupole mass spectrometer, model TSQ 700 (Finnigan-MAT, San Jose, CA), previously calibrated and tuned with a solution of horse myoglobin in the mass-to-charge range 500−2000 according to the manufacturer's instructions. Protein solutions (0.2 μM) were infused at a constant flow rate of 3 μL·min^{-1} by an infusion pump, model 22 (Harvard, South Natick, MA) through a 50-μm i.d. fused silica capillary directly into the electrospray ion source (+3.6 kV; nebulization gas, N_2, at 40 psi) of the mass spectrometer. Desolvation conditions for protein ions included maintaining the heated capillary at ≥150 °C and declustering potentials of at least +44 V. Data were collected in profile mode, and 100 spectra were acquired and integrated to produce a single average spectrum. Processed spectra were obtained using a deconvolution program developed in our laboratory at Rockefeller University.

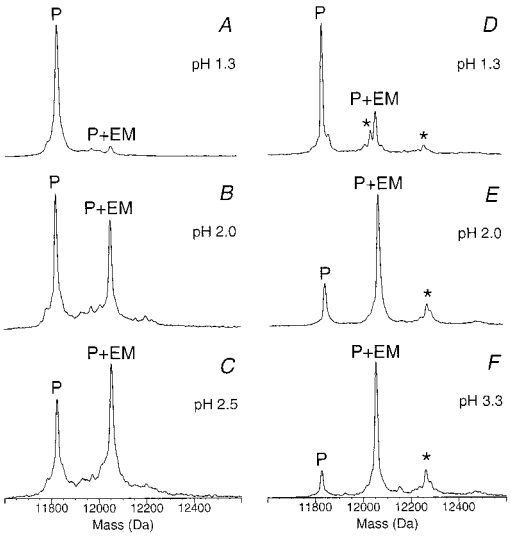

Figure 1. MALDI-MS of ethylmercury-derivatized FKBP12 using two MALDI matrixes (A−C) α-cyano-4-hydroxy-cinnamic acid (4hcca) and (D−F) sinapinic acid (SA). The pH of the matrix solutions is indicated in each spectrum. The matrix solutions consisted of (A and D) formic acid/water/2-propanol (1:3:2 v/v/v), (B and E) 0.1% trifluoroacetic acid/acetonitrile (2:1 v/v), and (C and F) water/acetonitrile (2:1 v/v). For clarity, only the portion of the spectra encompassing the singly charged region is shown. Peaks labeled 'P' designate the singly charged underivatized protein and 'P+EM' the singly charged ethylmercury derivative. Peaks labeled with an asterisk (*) represent matrix-related adduct peaks.[23,24]

RESULTS AND DISCUSSION

In the present study we compare the matrix-assisted laser desorption/ionization (MALDI) and electrospray ionization (ESI) mass spectrometry of mercury-derivatized proteins. An in-depth investigation of the MALDI-MS of proteins mercurated in solution will be shown, followed by a parallel study using ESI-MS. Finally, application of MALDI-MS to the analysis of protein crystals mercurated using traditional crystal-soaking methods will be presented.

MALDI-MS of Mercury-Derivatized Proteins. MALDI-MS normally requires the use of an organic matrix compound that is dissolved in an aqueous solution containing a water-miscible organic solvent. The two most commonly used matrixes for protein analysis are α-cyano-4-hydroxy-cinnamic acid (4hcca) and sinapinic acid (SA). The 4hcca matrix yields good-quality MALDI spectra for peptides and small proteins (below 30 kDa),[16,22] whereas SA is more suitable for protein analysis.[19,23,24] Figure 1 shows the MALDI mass spectra of ethylmercurated FKBP12 using 4hcca and SA matrixes and several matrix solution combinations. FKBP12 was derivatized in solution with ethylmercury phosphate (EMP), a reagent that selectively reacts with accessible sulfhydryl groups in the protein.[25] After EMP-derivatization of FKBP12 (one cysteine, M_r = 11 820), two major peaks were observed in the mass spectrum (Figure 1A)—one corresponding to underivatized protein (labeled P), the other to its ethylmercury derivative (labeled P + EM) with molecular mass 229 ± 2 Da higher than P. This shift in mass agreed well with the calculated shift in average mass of 229 Da produced by covalent attachment of one ethylmercury group to the cysteinyl residue of the protein. Surprisingly, the same EMP-derivatized protein sample yielded different results when analyzed in matrix solutions having different pH values (Figure 1A−C). The mass spectrum obtained from a 4hcca/pH 1.3 matrix solution indicates that only a small percentage of protein appeared to be ethylmercury-derivatized (~6%); the rest appeared to be underivatized (94%) (Figure 1A). Note that the apparent fraction of derivatized protein was calculated as (P + EM)/{P + (P + EM)} where P and (P + EM) correspond to peak heights, assuming that underivatized and derivatized protein had the same MS response. The use of matrix solutions with higher pH values led to dramatically different results. Thus, at pH 2.0 and 2.5 the extent of apparent derivatization was 42 and 57%, respectively (Figure 1 parts B and C). The MALDI-MS data show that the apparent fraction of protein derivatized by EMP is highly dependent on the pH of the matrix solution.

In addition to pH, we have observed that the choice of MALDI matrix, itself, influences the results. Figure 1 parts D−F show the mass spectra of EMP-derivatized FKBP12 using sinapinic acid (SA) matrix solutions prepared at three different pH values. At pH 1.3, 2.0, and 3.3, the apparent extent of derivatization was determined to be 20, 74, and 83%, respectively (Figure 1 parts D, E, and F). These findings show a direct relationship between matrix solution pH and the apparent extent of derivatization, a trend that was also observed with the matrix 4hcca (Figure 1 parts A−C). Inspection of Figure 1 also shows that analysis with SA yielded a higher apparent level of mercury incorporation than was observed that with 4hcca.

(22) Beavis, R. C.; Chaudhary, T.; Chait, B. T. *Org. Mass Spectrom.* **1992**, *27*, 156−8.
(23) Beavis, R. C.; Chait, B. T. *Rapid Commun. Mass Spectrom.* **1989**, *3*, 436−9.
(24) Beavis, R. C.; Chait, B. T. *Rapid Commun. Mass Spectrom.* **1989**, *3*, 432−5.
(25) Wong, S. S. *Chemistry of Protein Conjugation and Cross-linking*; CRC Press: Boca Raton, FL, 1993; p 32.

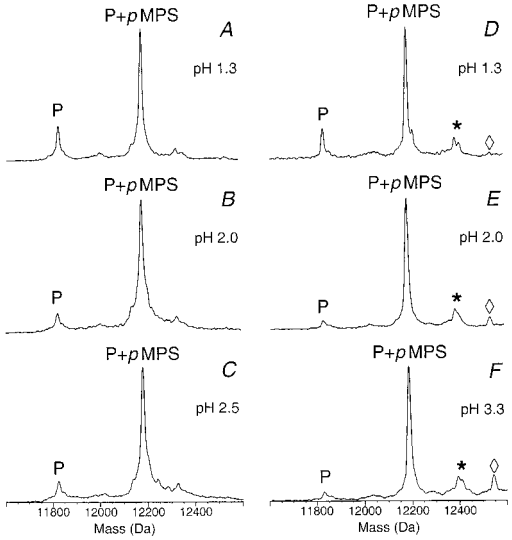

Figure 2. MALDI-MS of *p*-(chloromercuri)phenylsulfonate-derivatized FKBP12 using two matrixes, (A−C) 4hcca and (D−F) SA. The pH of the matrix solutions is indicated in each spectrum. Matrix solutions used in A−F were prepared as described in Figure 1. Peaks labeled 'P' designate the singly charged underivatized protein and 'P+*p*MPS' the singly charged *p*-(mercuriphenyl)sulfonate derivative. Peaks labeled with an asterisk (*) represent matrix-related adduct peaks. Attachment of a second *p*-(mercuriphenyl)sulfonate group is also indicated (◊) for the sinapinic acid matrix spectra.

The identity of the organic substituent bound to the mercury was a third factor that influenced the MALDI data. Figure 2 shows MALDI mass spectra of FKBP12 derivatized with *p*-(chloromercuri)phenylsulfonate (*p*CMPS). Derivatization by *p*CMPS replaces the sulfhydryl hydrogen of cysteine with a *p*-(mercuriphenyl)-sulfonate group (*p*MPS), increasing the average molecular mass of the protein by 357 Da. The matrix solutions used to obtain the spectra in Figure 2 (*p*CMPS derivatization) paralleled those used to generate Figure 1 (EMP derivatization). Again, pairs of peaks were observed (here separated by 357 ± 2 Da) which corresponded to underivatized (P) and derivatized (P + *p*MPS) FKBP12. A matrix solution pH-dependence was also observed with *p*CMPS-treated protein, albeit to a lower extent than that observed with EMP-derivatized protein. Thus, the apparent extent of *p*CMPS derivatization using 4hcca matrix solutions at pH 1.3, 2.0, and 2.5 was 79%, 89%, and 89%, respectively (Figure 2 parts A, B, and C). This narrow range (79−89%) contrasted with the much wider range (6−57%) observed using the same 4hcca matrix solutions for EMP-derivatized FKBP12 (Figure 1 parts A−C). Results obtained using sinapinic acid matrix exhibited similar behavior to that obtained with 4hcca. Analysis of *p*CMPS-derivatized FKBP12 using SA matrix solutions at pH 1.3, 2.0, and 3.3 yielded apparent *p*MPS incorporation levels of 84, 94, and 95%, respectively (Figure 2 parts D−F). Interestingly, spectra from sinapinic acid matrix solutions (Figure 2 parts D−F) showed an extra peak 357 ± 2 Da higher in mass than the *p*CMPS-derivatized protein. This second peak suggests an unexpected attachment of a second *p*MPS group to the protein. Note that addition of a second *p*MPS group was not observed to occur from any of the 4hcca matrix solutions (Figure 2 parts A−C). The multiple-attachment behavior was also specific for *p*CMPS since it was not observed to occur with EMP-derivatized FKBP12 ran in any of the matrix solution combinations used (see Figure 1).

The trends illustrated in Figures 1 and 2 suggest that the mercury substituent bound to the protein is labile when analyzed by MALDI-MS. To better understand these findings, we explored various experimental conditions that could influence these results. We summarize our findings as follows (data not shown).

Since the mercury lability appeared to be related to the matrix solution acidity, we explored the effects of elevating the matrix solution pH. This required the use of a neutral matrix (e.g., 2′,4′,6′-trihydroxyacetophenone) prepared in ammonium acetate solutions (pH range 4−7). However, the protein response by MALDI using pH 4−7 matrix solutions was very weak and the results were inconclusive.

In all of the experiments described so far, we used the standard dried-droplet method of matrix-sample preparation.[16,17,26] Derivatized protein was first combined with matrix solution, the mixture was immediately spotted onto the MALDI probe, and the solvent was allowed to evaporate at ambient temperature, leaving matrix crystals for MS analysis. The entire sample preparation process took no more than 10 minutes. We found no significant changes in the results shown in Figures 1 and 2 if the mercury-derivatized proteins were left in the acidic matrix solutions for an extended period (e.g., hours to days) prior to spotting on the probe or if different organic solvents (e.g., acetonitrile or 2-propanol) were used to prepare the matrix solution.

We carried out MALDI analysis on a second protein—EMP- and *p*CMPS-derivatized Nova1 (one cysteine, $M_r = 8137$)—using 4hcca and SA matrixes and found trends similar to those presented by FKBP12. In addition, our findings were essentially independent of the MALDI instrumentation used to acquire the mass spectra, regardless of laser wavelength (UV $\lambda = 337$ or $\lambda = 355$ nm, or IR $\lambda = 2.94$ μm) or time-of-flight analyzer type (linear or reflector).

ESI-MS of Mercury-Derivatized Protein. The results described above showed that the mercury−sulfur bond in mercury-derivatized proteins is labile under MALDI-MS conditions. To further understand these observations, we turned to ESI-MS, which offers good control of the electrospray solution conditions. For example, the solution pH can be freely adjusted, in contrast to MALDI matrix solutions, which are usually maintained at acidic pH values for protein analysis. In addition, the electrospray process produces "cooler" ions (i.e., ions having lower internal energy) than those produced by MALDI, thus enabling the measurement of relatively labile ionic species.

Figure 3 shows convoluted ESI-MS spectra of mercury-derivatized protein. FKBP12 that had been derivatized with EMP and electrosprayed in 0.1% aqueous TFA solution (pH 1.6) showed two peaks corresponding to underivatized (P) and derivatized protein (P + EM) (Figure 3A). Assuming no significant difference in mass-spectrometric response for underivatized and derivatized FKBP12, we calculated the extent of derivatization to be 95% on the basis of the relative heights of the two peaks. Identical patterns were observed when EMP-treated FKBP12 was electrosprayed from an aqueous solution of 5% (v/v) acetic acid (pH 2.4, Figure 3B) or from deionized water (pH 5.6, Figure 3C). In all three cases, the extent of derivatization was ≥95%. Similar results were obtained for FKBP12 treated with *p*CMPS (Figure 3 parts D−F) where the extent of derivatization was also found to be at least 95%. The electrospray results are distinct from those obtained by

(26) Karas, M.; Hillenkamp, F. *Anal. Chem.* **1988**, *60*, 2299−301.

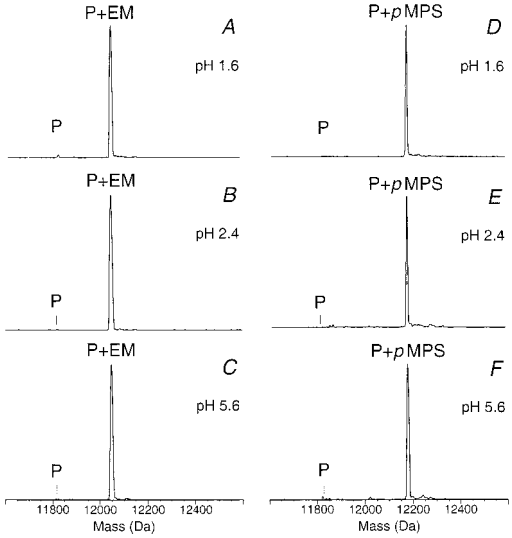

Figure 3. Convoluted ESI-MS spectra of FKBP12 derivatized with ethylmercury phosphate (A−C) and *p*-(chloromercuri)phenylsulfonate (D−F). The pH of the electrospray solutions is indicated in each spectrum. The solutions consisted of 0.1% aqueous trifluoroacetic acid (A and D), 5% aqueous acetic acid (B and E), and water (C and F). Two peaks were observed in each convoluted spectrum—one peak corresponding to underivatized FKBP12 (identified as P) and the other peak to either EMP-derivatized FKBP12 (P + EM) (panels A−C) or *p*CMPS-derivatized FKBP12 (P + *p*MPS) (panels D−F). (The raw ESI-MS spectra of derivatized FKBP12 yielded an envelope of multiply charged peaks from 8+ to 17+ for each component present (data not shown). The mass-to-charge spectra were converted into spectra of molecular mass using a deconvolution program developed in our laboratory at the Rockefeller University).

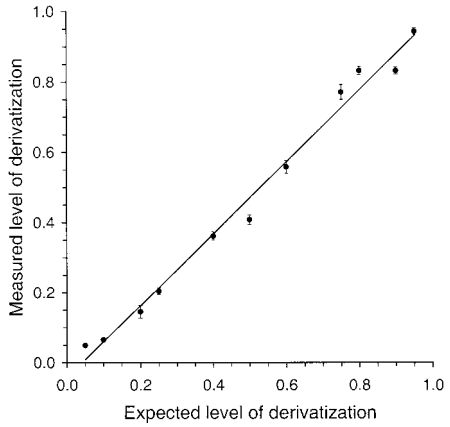

Figure 4. Standard curve for quantitation of *p*CMPS-derivatized FKBP12. The plot was obtained from ESI-MS of solutions with known ratios of FKBP12 and *p*CMPS-derivatized FKBP12. Known ratios of derivatized and underivatized protein were mixed to provide the "expected level of derivatization". Linear regression yielded a line with slope of 1.03 ± 0.04 and a correlation coefficient (r) of 0.994. Bars represent standard deviation, $n = 3$.

MALDI-MS in that ESI-MS yielded accurate levels of mercury incorporation for the fully derivatized FKBP12. The standard quantitation curve shown in Figure 4 confirms this accuracy over a wide range of derivatization levels.

The high levels of mercury incorporation (≥95%) were also observed under different solution and instrumental electrospray conditions. For example, electrospraying from strongly acidic solutions (e.g., 17% aqueous formic acid) resulted in no change in the level of mercury incorporation even after prolonged exposure (up to 5 days) of the protein to the strongly acidic conditions. In addition, increasing the internal energy of electrosprayed ions by varying the transport capillary temperatures from 150 to 300 °C and declustering potentials from +40 to +140 V did not influence the apparent level of mercury incorporation (data not shown; J. C. Padovan, 1998, unpublished results).

Analysis of Proteins with Higher Molecular Mass and Proteins Containing Multiple Cysteines. We have applied the MALDI and ESI methods to proteins with molecular masses as high as 47 kDa. For example, *p*CMPS-treated enolase (one cysteine, $M_r = 46\,671$) yielded results analogous to those shown above for FKBP12. MALDI-MS analysis of derivatized enolase using 4hcca/pH 2.5 matrix solution yielded a mass shift of 361 ± 10 Da (expected, 357 Da) and a lower limit of derivatization of 72%. ESI-MS analysis of the same preparation of enolase yielded a mass shift of 361 ± 5 Da and showed results consistent with 100% derivatization (data not shown).

Likewise, proteins with multiple cysteines can also be analyzed. Human hemoglobin is composed of an α chain (one cysteine, $M_r = 15\,126$) and a β chain (two cysteines, $M_r = 15\,867$). MALDI-MS analysis of *p*CMPS-derivatized hemoglobin in SA/pH 3.3 matrix solution showed a singly derivatized peak for the α chain with a lower limit of derivatization of 89%. For the β chain, the same MALDI spectrum showed two mass-shifted peaks corresponding to a singly and a doubly derivatized β chain with lower limits of derivatization of 30 and 45%, respectively. In contrast, ESI-MS analysis of the same hemoglobin solution showed one peak for the α chain corresponding to 100% derivatization of its single cysteinyl residue, and only one peak for the β chain, in accordance with the full derivatization of its two cysteinyl residues.

MALDI-MS of Protein Crystals. Protein mercuration can be carried out either in solution or in the protein crystal itself. This latter method is usually attempted first for X-ray crystallographic phase determination purposes. If soaking methods fail to produce a suitably derivatized crystal, solution mercuration is undertaken followed by new crystallization trials.

We have already described the MALDI- and ESI-MS of protein derivatized in solution. Our results have shown that ESI-MS yields accurate information regarding the extent of mercury derivatization of proteins whereas MALDI-MS yields a lower estimate of mercury incorporation. However, we have found MALDI-MS invaluable to the analysis of single protein crystals. Because MALDI-MS is highly tolerant of additives used in protein crystallization, it is the method of choice for direct analysis of protein crystals since ESI-MS would require thorough sample cleanup prior to analysis.

Figure 5 shows the MALDI mass spectrum of a single protein crystal of the potassium channel from *Streptomyces lividans* (KcsA) following crystal soaking with the mercurating reagent methylmercury chloride. This reagent was observed to yield high levels of methylmercury incorporation in KcsA by MALDI-MS in earlier screening experiments (data not shown; S. L. Cohen, 1998, unpublished results). Two principal peaks were observed in the mass spectrum of the protein crystal: underivatized KcsA (Ala32Cys mutant, $M_r = 13\,388$) and methylmercurated KcsA (shifted 214 Da higher in mass). From the peak heights we determined the apparent level of derivatization to be 60%, representing a lower limit of the extent of methylmercury incorporation into the KcsA

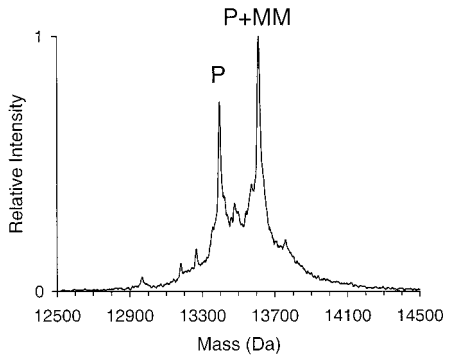

Figure 5. MALDI-MS of a single crystal of the *Streptomyces lividans* potassium channel protein, KcsA. The crystal was soaked in mother-liquor solution spiked with methylmercury chloride prior to MS analysis (see Experimental Section). The spectrum was obtained from a saturated solution of 4hcca in 0.1% trifluoroacetic acid/acetonitrile (2:1 v/v) at pH 2.0. (Although sinapinic acid matrix would have been a better choice of matrix (see text), we used 4hcca because it gave a superior protein response.) Peaks corresponding to underivatized (P) and methylmercury-derivatized (P + MM) KcsA are labeled. The smaller peaks on the left side of the peak labeled 'P' correspond to N-terminal heterogeneity of the expressed recombinant KcsA protein.

protein crystal. However, knowledge of this level of mercuration was sufficient to proceed to the X-ray analysis and phase determination of the KcsA protein.[14]

SUMMARY AND CONCLUSIONS

Mass spectrometry has matured into a useful tool for structural biology. MALDI- and ESI-MS allow ready assessment of the correctness of expression and purity of recombinant proteins[27−29] as well as provide a facile means to determine precise compact folding domains.[30] In this paper, we have shown that mass spectrometry also provides useful information concerning the incorporation of heavy atoms into proteins for X-ray phase determination. In particular, our study shows that both MALDI- and ESI-MS can be used to analyze the level of mercury incorporation in mercury-derivatized proteins, although the two methods exhibit significant differences (see below).

ESI-MS. The linear standard curve shown in Figure 4 demonstrates that ESI-MS provides an accurate determination of the level of mercury incorporation in proteins. The measured level of incorporation as assessed by electrospray mass spectrometry was independent of the ESI solution pH as well as the desolvation conditions. Thus, ESI-MS is the method of choice for accurate determination of mercury incorporation in proteins. For similar reasons, ESI-MS is likely to be the method of choice for measuring the incorporation of other heavy metals frequently used for phase determination. However, ESI-MS has a considerable drawback—the necessity for thorough sample cleanup (of salts, buffers, and other additives) prior to analysis. Because of this need, sample turnaround and throughput is considerably lower compared with that observed by MALDI-MS. Thus, the ideal heavy-atom MS analysis would involve a combination of MALDI (for rapid screening; see below) and ESI (for high accuracy of derivatization).

MALDI-MS. In contrast to the ESI-MS results of fully mercurated proteins, which yielded measured incorporation levels ≥95%, MALDI-MS of the same fully mercurated proteins exhibited a broad range of apparent level of mercuration (between 6 and 95%). The MALDI results showed a dependence on the choice of matrix, matrix solution acidity, and the identity of the mercury substituent. For similar reasons, caution should be taken in analyzing protein derivatized by reagents containing heavy atoms other than mercury. (Note: an exception is the analysis of proteins containing the nonlabile selenomethionyl group, often used for phasing.[4] Both MALDI- and ESI-MS provide accurate levels of selenomethionine incorporation[31] (S. L. Cohen, 1998, unpublished results)). Although not as accurate as ESI in determining the extent of derivatization for heavy atoms such as mercury, MALDI does provide a useful lower limit to the level of metal incorporation through the use of an appropriate matrix and matrix solution combination. We have consistently found that the high throughput capabilities and quick turnaround of information make MALDI-MS a more practical choice over ESI-MS in the analysis of heavy-atom derivatized proteins. A good example is the structural determination of the potassium channel from *Streptomyces lividans*.[14] MALDI-MS provided useful information in a timely manner—first guiding in the selection of suitable heavy-atom reagents and derivatization conditions and ultimately in identifying protein crystals that were likely to yield useful X-ray diffraction phase data.

Glossary. EM, ethylmercury; EMP, ethylmercury phosphate; ESI, electrospray ionization; FKBP12, FK506-binding protein 12; 4hcca, α-cyano-4-hydroxy-cinnamic acid; IR, infrared; KcsA, the potassium channel from *Streptomyces lividans*; MALDI, matrix-assisted laser desorption/ionization; MMC, methylmercury chloride; MS, mass spectrometry; *p*MPS, *p*-(mercuri)phenylsulfonate; *p*CMPS, *p*-(chloromercuri)phenylsulfonate, monosodium salt; TFA, trifluoroacetic acid; SA, sinapinic acid; UV, ultraviolet.

ACKNOWLEDGMENT

We thank Roderick MacKinnon, Stephen Burley, and John Kuriyan for their insightful discussions concerning X-ray crystallography and phasing methods. We would also like to thank Morgan Huse (FKBP12), Hal Lewis (Nova1), Declan Doyle (KcsA), João Cabral (KcsA), and Joe Marcotrigiano for supplying protein and reagents and for their guidance concerning heavy-atom derivatization. This work was supported in part by a Grant from the National Institutes of Health (RR00862). J.C.P. was supported by a postdoctoral fellowship from Fundação de Amparo à Pesquisa do Estado de São Paulo—FAPESP, Brazil.

Received for review August 17, 1999. Accepted November 16, 1999.

AC990938E

(27) Chait, B. T. *Structure* **1994**, *2*, 465−7.
(28) Jensen, C.; Haebel, S.; Andersen, S. O.; Roepstorff, P. *Int. J. Mass Spectrom. Ion Processes* **1997**, *160*, 339−56.
(29) Easterling, M. L.; Colangelo, C. M.; Scott, R. A.; Amster, I. J. *Anal. Chem.* **1998**, *70*, 2704−9.
(30) Cohen, S. L. *Structure* **1996**, *4*, 1013−6.
(31) Budisa, N.; Steipe, B.; Demange, P.; Eckerskorn, C.; Kellermann, J.; Huber, R. *Eur. J. Biochem.* **1995**, *230*, 788−96.