BIOORGANIC CHEMISTRY 23, 380–396 (1995)

Structural Characterization of the Recombinant P40 Heavy Chain Subunit Monomer and Homodimer of Murine IL-12

ELLIOTT B. NICKBARG,[1] JAMES E. VATH, DEBRA D. PITTMAN, JOHN E. LEONARD, KRISTINE E. WALDBURGER, AND MICHAEL D. BOND

Genetics Institute, Inc., 87 Cambridgepark Drive, Cambridge, Massachusetts 02140

Received May 12, 1995

Interleukin-12 (IL-12) is a heterodimeric cytokine that consists of two structurally unrelated subunits, P35 and P40. However, when expressed alone in Chinese hamster ovary (CHO) cells, murine P40 showed two species of different molecular weights under nonreducing conditions, a monomeric form of 45 kDa and a homodimer of >97 kDa. Under reducing conditions the two forms migrated as an identical array of species of 40–45 kDa. The monomer was separated from the homodimer under nonreducing conditions by heparin affinity chromatography and the disulfide bond structures of both species were determined by peptide mapping, Edman sequencing, and mass spectrometry. The peptide maps of the two species were identical except for a single peak that changed retention time. Sequencing showed that this peak contained two peptides of identical sequences in both maps. Mass spectrometric analysis of the peak from the >97-kDa species revealed an ion of double the expected mass, thus indicating that the peptide pair had dimerized. Mass analysis of the peak from the 40- to 45-kDa species showed that the peptide pair contained a mass difference that corresponded to that of an extra cysteine and which disappeared upon reduction. Amino acid analysis confirmed that the monomeric form of rmP40 is modified by a reducible cysteine. Structural analysis of the remainder of the cysteine-containing peaks showed that both species of rmP40 contained the same set of intramolecular disulfide bonds. The murine P40 homodimer arises from formation of a single intermolecular disulfide bond at Cys^{175}. In the monomeric P40, however, this cysteine is capped by an additional cysteine. Purified rmP40 monomer and homodimer inhibited the IL-12-dependent induction of interferon-γ, but neither appeared capable of inducing IL-12-like biological activity.

INTRODUCTION

Interleukin 12 (IL-12)[2] is a heterodimeric glycoprotein that is produced by macrophages and B lymphocytes, and its biological activities include the production of interferon-γ, stimulation of T-cell proliferation, and the enhancement of natural

[1] To whom correspondence should be addressed.

[2] Abbreviations used: CHO, Chinese hamster ovary; DTT, dithiothreitol; IFN-γ, interferon-γ; IL-12, interleukin-12; LSIMS, liquid secondary ionization mass spectrometry; Lys-C, endoproteinase Lys-C from *Achromobacter lyticus*; MALDI-TOF MS, matrix-assisted laser desorption time-of-flight mass spectrometry; MEM, minimal essential medium; $(M + H)^+$, mass of protonated molecular ion; *m/z*, mass-to-charge ratio; P35, P35 light chain subunit of IL-12; P40, P40 heavy chain subunit of IL-12; PITC, phenyl isothiocyanate; PLP, myelin proteolipid protein; PMSF, phenylmethylsulfonyl fluoride; PTC, phenylthiohydantoin; rmP40, recombinant murine interleukin 12 heavy chain; RP/HPLC, reversed-phase high performance liquid chromatography; SDS-PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis.

0045-2068/95 $12.00

killer cell-mediated cytotoxicity (*20, 11, 28, 32*). IL-12 is unique among cytokines in that it consists of two structurally unrelated polypeptide chains, light chain (P35) and heavy chain (P40), that are biologically inactive when expressed alone (*33*). IL-12 rapidly loses activity upon reduction with DTT, thus suggesting that disulfide bonding is necessary for maintenance of biological activity (*26*). Human P40 contains 10 cysteines and the murine P40 chain contains 13 cysteines, nine of which are conserved in human P40 (*27*). The P40 chain exhibits sequence homology to the extracellular domain of the interleukin 6 receptor (*12*) and the P35 chain bears a lesser degree of similarity to interleukin 6 (*25*). These observations suggest that IL-12 is structurally related to a soluble cytokine–receptor complex (*12*). Although the P40 subunit has no apparent biological activity or ability to bind the IL-12 receptor alone, it may interact with the receptor or stabilize P35 in a conformation that allows binding to the receptor (*26*). A study using COS cell culture supernatants transfected with mouse P40 cDNA shows that these supernatants could inhibit IL-12-dependent biological activities (*24*). More recent work, using purified human P40 derived from COS cells, has shown that a homodimer of P40 is responsible for this inhibition (*23*).

In order to further investigate the structural and biological properties of the P40 heavy chain of IL-12, we have expressed the murine P40 subunit in CHO cells. The CHO rmP40 subunit, like the COS-derived subunit, is expressed as a monomer of molecular weight 40–45 kDa, and a homodimer of >97 kDa under nonreducing conditions. Both the monomer and homodimer migrate as two identical species of 40–45 kDa molecular weight under reducing conditions. In order to determine the basis of this structural difference, we have purified and separated the two species and have investigated their disulfide bond structures using peptide mapping, Edman sequencing, amino acid analysis, and mass spectrometry, which has the capability to analyze in detail the disulfide bonding patterns present in natural and recombinant proteins (*29*). The ability of both the monomer and homodimer to affect the antigen-stimulated induction of IFN-γ by IL-12 *in vitro* was also measured. Our results outline the structural basis for dimerization of the murine P40 subunit and confirm its ability to act as an IL-12 antagonist.

MATERIALS AND METHODS

Cell line development. The cDNAs encoding the mouse IL-12 P40 subunit and IL-12 P35 subunit were cloned into the *Eco*R1 site of pED and the resultant plasmids were designated EMCmp40 and EMCmp35, respectively (*19, 27*). The cell clone expressing the rmP40 subunit was obtained by cotransfection using lipofection of equal quantities (10 μg) of EMCmp35 and EMCmp40 DNA, linearized with the restriction enzyme *Nde*I into Chinese hamster ovary cells deficient in dihydrofolate reductase (*30, 10*). Cells were selected for growth in the absence of nucleosides and increasing concentrations of methotrexate. Several methods were used to identify cell lines which expressed only the P40 subunit. Conditioned medium was analyzed by Western blotting using anti-P40 monoclonal antibody or anti-P35 rabbit polyclonal antibody and horseradish peroxidase-coupled secondary antibody (*28*).

The antibody complexes were detected using chemiluminescence (ECL, Amersham, Arlington Heights, IL). Total RNA was isolated by the guanidine thiocyanate method, fractionated on a 0.8% agarose formaldehyde gel, and subsequently transferred to nitrocellulose (*2, 7*). Probes were prepared by random priming of the *Eco*RI fragment of the P40 cDNA and the *Eco*RI fragment of the P35 cDNA. The RNA was analyzed by polymerase chain reaction using P40- or P35-specific oligonucleotides to confirm the absence of P35 mRNA.One cell line expressed only the P40 subunit and was selected for growth in 0.5 mM methotrexate. Although this line was produced from a series of transfections that was originally intended to yield IL-12 heterodimer, we feel that these tests conclusively demonstrate that only P40 was expressed.

The cells expressing the P40 subunit were seeded into 1700-ml roller bottles. Logarithmically growing cells were fed defined serum free containing medium (Dulbecco's Modified Eagle Medium/Ham's F-12, 1 : 1) containing 20 μg/ml aprotinin (Sigma, St. Louis, MO). Conditioned medium was harvested daily for 4 days, centrifuged at a low speed to remove cellular debris, and stored at −80°C until purified.

Purification of recombinant murine IL-12 and rmP40. CHO recombinant murine IL-12 was purified as previously described (*15*). rmP40 was purified in the same fashion as murine IL-12 with the following modifications. Serum-free culture supernatant (10 liters) was diluted with an equal volume of 50 mM Tris–Cl, pH 8.0, and then applied to a column (7 × 6.5 cm) of Q-Sepharose fast flow (Pharmacia, Piscataway, NJ) at a flow rate of 150 ml/min at 4°C. The column was washed with 1 liter of 0.1 M Tris–Cl, pH 8.0, and then rmP40 was eluted with 0.5 M NaCl in 0.1 M Tris–Cl, pH 8.0, in six 50-ml fractions. After purification by phenyl-superose chromatography as previously described for murine IL-12 (*15*), the pooled phenyl superose fractions were dialyzed against 50 mM NaP_i buffer, pH 7.2, containing 50 mM NaCl and applied to a column (1.0 × 8.0 cm) of heparin toyopearl (TosoHaas, Philadelphia, PA) connected to an FPLC apparatus (Pharmacia) at a flow rate of 1.0 ml/min and then was eluted with a linear gradient of NaCl (50 mM to 1 M) in the same buffer. Fractions containing rmP40 were identified by SDS–PAGE using Tris–glycine gels containing 12% polyacrylamide (Novex, San Diego, CA) and Western blot analysis (*28*). Apparent molecular weight was determined using the prestained protein molecular weight markers (high range) from Amersham. Protein was determined using the Bio-Rad protein assay (Bio-Rad, Hercules, CA) using bovine serum albumin (Pierce, Rockford, IL) as a standard.

Lys-C proteolytic digestion and mapping. Purified rmP40 (350 μg) was digested with endoproteinase Lys-C (EC 3.4.21.50; Wako, Dallas, TX, 1/10 w/w enzyme/substrate) in phosphate-buffered saline with 4 M urea at 30°C for 20 h. The mixture of peptides was acidified with acetic acid and separated by RP/HPLC on a 250 × 4.6 mm Vydac protein C4 column (The Separations Group, Hesperia, CA) connected to a Waters (Milford, MA) 600E separation system and 490 UV/vis detector as described in Fig. 2. Fractions were collected manually.

Subdigestion of the endoproteinase Lys-C fragments. In order to identify disulfide pairings in peptides containing multiple cysteine residues (fractions **a, c,** and **f** in Fig. 2), subdigestion of the Lys-C-derived peptides was performed. A 20-μl aliquot

of each RP/HPLC fraction was concentrated to approximately 5 μl using vacuum centrifugation and was then mixed with 5 μl of 100 mM Tris–Cl buffer, pH 7.5. For fraction **a,** a 1-μl aliquot of a 0.1 μg/μl solution of sequenal grade trypsin (Boehringer Mannheim, Indianapolis, IN) was added and the sample incubated for 2 h at 37°C under argon. For fractions **c** and **f,** a 1-μl aliquot of a 0.1 μg/μl solution of endoproteinase Asp-N (Boehringer Mannheim) was added and the sample was incubated for 16 hr at 37°C under argon. A 1-μl aliquot of each of these digest solutions was mixed with 2 μl of matrix solution and analyzed directly by MALDI-TOF MS as described below.

Peptide analysis. Peptides were analyzed by both N-terminal Edman sequencing using an Applied Biosystems (Foster City, CA) 470A sequenator coupled to an on-line Applied Biosystems 120A PTH analyzer as previously described (*14, 17*) and by mass spectrometry using MALDI-TOF MS and LSIMS as described below.

MALDI-TOF MS was performed on a Bruker (Billerica, MA) REFLEX mass spectrometer equipped with a nitrogen laser emitting radiation at 337 nm and an ion mirror reflector. The accelerating voltage was set to 28.5 kV during the measurements. Spectra were obtained in either the linear mode (no potential on ion mirror, flight path 130 cm) or the reflector mode (30 kV applied to mirror, flight path 210 cm). The instrument was calibrated using a mixture of adrenocorticotropic hormone 18–39 (BACHEM, LaJolla,CA) and *Escherichia coli* thioredoxin (Calbiochem, Torrance, CA). Based on repeated analysis of the standards, the typical resolution and mass accuracies are 400 and 0.1%, respectively, for the linear mode, and 1000 and 0.05% for reflector mode. The laser flux is adjusted via a computer controlled variable attenuator at a repetition rate of 3 Hz. Spectra displayed represent the summation of 10 laser shots. The signal from the detectors was digitized at a sampling rate of 100 MHz using a transient recorder. The instrument control and data processing was accomplished with software supplied by Bruker using a Sun Microsystem (Mountain View, CA) IPX workstation. A 1-μl aliquot of each RP/HPLC fraction (about 1 μg/μl) was mixed with 2 μl of matrix solution (6 mg of sinapinic acid in 0.5 ml of 30:0.1:69.9 acetonitrile/trifluoroacetic acid/water) and air dried at room temperature on a stainless steel MALDI probe. For MALDI analysis of reduced peptides, the sample (approximately 0.1–2 pmol/μl) was reacted in 20 mM Tris–Cl (pH 8.0), 10 mM DTT for 30 min at 50°C, and 1 μl of the mixture was mixed with 2 μl of matrix solution and air dried on the probe for analysis.

LSIMS was performed on the first two sectors of a JEOL (Peabody, MA) HX110/HX110 high performance tandem mass spectrometer. The instrument was scanned from *m/z* 0–8000 at a resolution of 1:1000 or 1:3000 using a 10 kV accelerating voltage. The ions were generated using a JEOL cesium gun operating at 22 kV with 2.2 A. The data was stored and processed using the JEOL Complement software. The instrument was calibrated using cesium iodide cluster ions from *m/z* 133–7927. A 0.5-ml aliquot of the RP/HPLC fraction was mixed with 100 μl of 2% glycerol and then concentrated *in vacuo*. *In situ* reduction was performed with a 0.5-μl aliquot of the sample solution in glycerol mixed on the LSIMS probe with 0.5 μl of mono-thioglycerol and 0.5 μl of aqueous ammonia (30% by weight). The solution was gently warmed for approximately 1 min with a heat gun followed

by addition of 0.5 μl of aqueous ammonia and another application of heat. It was then inserted directly into the instrument for measurement.

Amino acid analysis. Amino acid analyses employed the PicoTag procedure as described elsewhere (Waters Chromatography PicoTag Amino Acid Analysis manual; and Ref. 4). Briefly, aliquots of 300–500 pmol peptide were hydrolyzed at 120°C for 18 h in the presence of HCl and phenol in the vapor phase and then derivatized by incubation with 7:1:1:0.8 ethanol:water:triethylamine:PITC for 25 min. After removal of excess reactants by vacuum centrifugation, the samples were redissolved in 5 mM sodium phosphate, 5% acetonitrile, pH 7.4, and analyzed by RP/HPLC using a 3.9 × 300 mm PicoTag column (Waters).

Detection of Cys from cysteinylation employed a modification of the above procedure. Samples were derivatized without prior hydrolysis in methanol:water:triethylamine:PITC (7:1:1:0.8) and then analyzed both with and without an intervening reduction step. Reduction served to release any disulfide-bonded PTC-Cys for analysis, and was performed by incubation in 20 μl of 5 mM DTT, 5 mM sodium phosphate, 5% acetonitrile, pH 7.4, for 10 min at 60°C. Quantitation of PTC-Cys was based on the peak area obtained after reduction of derivatized amino acid standards including PTC-cystine. All attempts to derivatize free Cys, rather than cystine, were unsuccessful.

Effect of murine P40 on IFN-γ production. The ability of P40 to inhibit IL-12-stimulated IFN-γ production was tested *in vitro*, following stimulation of antigen-primed splenocytes with IL-12. SJL/J mice were immunized with myelin proteolipid protein (PLP) in complete freunds adjuvant (CFA) as described previously (*22*). After 10 days, spleens were removed and single cell suspensions (1×10^6/ml) stimulated with either PLP alone (5 μg/ml) or PLP and increasing concentrations of murine IL-12 (0.2 to 2000 pg/ml). The P40 (either homodimer or monomer) was added to all wells at a concentration of 200 ng/ml, which represents a 100-fold excess compared to the highest concentration of IL-12 added. IFN-γ production after 48 h in culture was measured using commercially available ELISA assay (Endogen).

RESULTS

Expression and purification of murine P40. The mature processed translation product of the murine P40 subunit cDNA is a polypeptide of 313 amino acids with a calculated protein molecular weight of 35801.5 Da. It contains five potential N-linked glycosylation sites. SDS–PAGE analysis of the conditioned media under nonreducing conditions shows that approximately half of the rmP40 migrates as a set of bands between 40 and 45 kDa and half as a series of bands at greater than 97 kDa. The relative proportions of the two species remained constant throughout the purification procedure. The two species were separated by heparin affinity chromatography using a linear gradient of NaCl. SDS–PAGE analysis of both peaks showed that the first peak corresponded to the 40- to 45-kDa species (Fig.

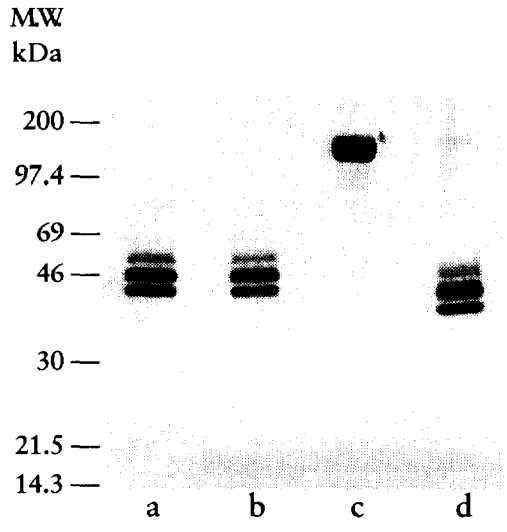

FIG. 1. SDS–PAGE of heparin-purified rmP40. Samples (2.5 μg) of the separated rmP40 peaks were analyzed under nonreducing and reducing (+50 mM DTT) conditions on a 12% slab gel. Visualization was by Coomasie blue. The molecular masses in kDa were estimated from standards electrophoresed in a parallel lane. Lane a, rmP40 monomer reduced with DTT; lane b, rmP40 aggregate reduced with DTT; lane c, rmP40 aggregate nonreduced; lane d, rmP40 monomer nonreduced.

1, lane d) and the second to the high molecular weight form (Fig. 1, lane c). However, electrophoresis of either species under reducing conditions gave the same array of bands (at 40–45 kDa; Fig. 1, lanes a and b). These results imply that rmP40 is expressed both as a monomeric species and as a disulfide-linked covalent aggregate. The variety of bands observed in the 40- to 45-kDa region of the gel most likely represent different glycosylation variants of the monomeric species, although this remains to be proven.

Peptide mapping of rmP40. In order to determine the reason for the molecular weight differences in rmP40, the low and high molecular weight species each were digested with the lysine-specific endoprotease Lys-C under nonreducing conditions, and the peptide mixtures from each digestion were purified by RP/HPLC (Fig. 2). Cleavage after lysine was expected to yield nine cysteine-containing peptides (Table 1). Digestion was conducted under neutral or near-neutral conditions in order to minimize the possibility of disulfide bond scrambling. Peptide-containing peaks were identified by Edman sequencing and further characterized under reducing and nonreducing conditions by LSIMS and/or MALDI-TOF mass spectrometry depending on the resolution, mass accuracy and mass range requirements of the measurement. In general, MALDI-TOF was the most useful for the nonreducing measurements because of its extended mass range capabilities and compatibility with a variety of protein modifications. LSIMS was primarily used for the reducing measurements for its resolution and mass accuracy. Under nonreducing conditions, ideally, one expects to observe an m/z value consistent with the $(M + H)^+$ of the

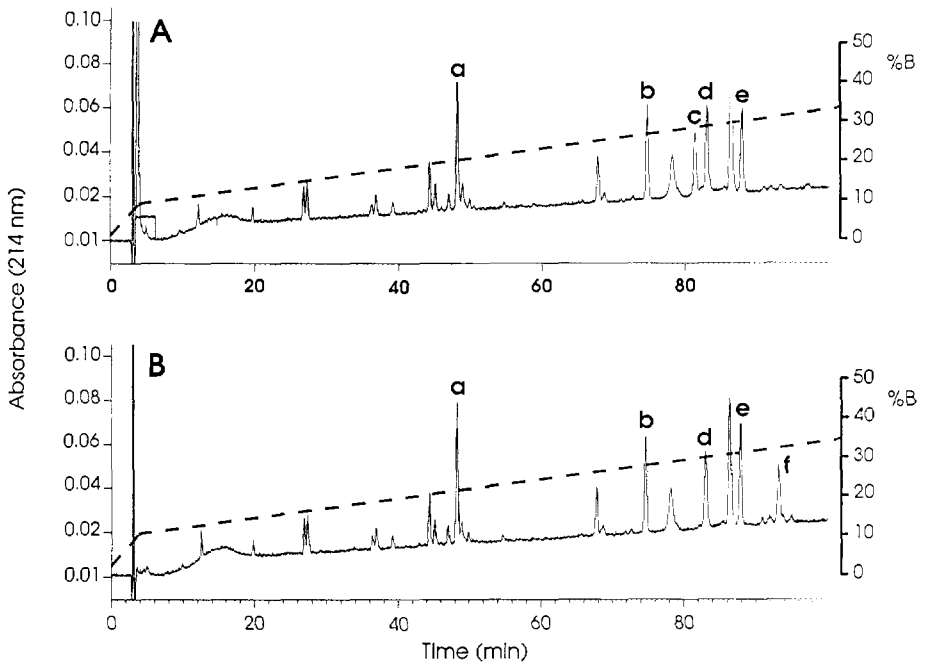

FIG. 2. RP/HPLC separation of endoproteinase Lys-C peptides of rmP40 low- and high-molecular weight forms. (A) Low-molecular weight rmP40; (B) high-molecular weight rmP40. Buffer A: 0.1% trifluroacetic acid in water; buffer B: acetonitrile/water/trifluroacetic acid (95/4.9/0.1). The column was eluted at a flow rate of 1.0 ml/min with an increasing gradient (dashed line) of buffer B. Protein was monitored by absorbance at 214 nm (solid line). The cysteine-containing peptides identified by Edman sequence analysis are noted as fractions **a–f.**

oxidized peptide(s) that provides direct evidence that a disulfide bond exists. The observation is further substantiated by disappearance of the signal after chemical reduction of the sample and the detection of the thiol form of the peptides involved. Both mass spectrometric and sequence analysis showed that only five peaks, labeled **a–f** (Fig. 2) had cysteine-containing peptides. Repeat digestions (data not shown) showed no evidence of disulfide containing peptides in other peaks, thus indicating that disulfide bond scrambling was not occurring. The results from the Edman sequencing and mass spectrometric analyses are described below and summarized in Table 2.

Four of the cysteine-containing RP/HPLC fractions **a, b, d,** and **e** (Table 2), gave identical results for the low and high molecular weight forms of rmP40. Edman sequence analysis showed that fraction **a** consists of four peptides containing a total of six cysteines, and MALDI-TOF analysis under nonreducing conditions gave a species of *m/z* 5183 (Fig. 3) that disappeared under reducing conditions. These results indicate that the four peptides, one of which is the C-terminus of rmP40 (Trp^{304}-Ser^{313}), exist as a disulfide bonded cluster. Further digestion of this cluster by trypsin, followed by MALDI-TOF mass spectrometric analysis without further

TABLE 1
Expected Cleavage Products Derived from Digestion of rMp40 by Endoproteinase Lys-C from *Achromobacter lyticus*

No. Cys	Fragment	Position	Amino acid sequence
	1	1–6	MWELEK
1	**2**	7–51	DVYVVEVDWTPDAPGETVNLTCDTPEEDDITWTSDQRHGVIGSGK
	3	52–58	TLTITVK
1	**4**	59–70	EFLDAGQYTCHK
	5	71–84	GGETLSHSHLLLHK
	6	85–85	K
	7	86–96	ENGIWSTEILK
	8	97–99	NFK
	9	100–101	NK
	10	102–105	TFLK
2	**11**	106–128	CEAPNYSGRFTCSWLVQRNMDLK
	12	129–132	FNIK
1	**13**	133–154	SSSSPPDSRAVTCGMASLSAEK
	14	155–164	VTLDQRDYEK
2	**15**	165–195	YSVSCQEDVTCPTAEETLPIELALEARQQNK
	16	196–210	YENYSTSFFIRDIIK
	17	211–215	PDPPK
	18	216–220	NLQMK
	19	221–223	PLK
	20	224–247	NSQVEVSWEYPDSWSTPHSYFSLK
	21	248–255	FFVRIQRK
	22	256–256	K
	23	257–258	EK
	24	259–260	MK
1	**25**	261–269	ETEEGCNQK
	26	270–276	GAFLVEK
1	**27**	277–284	TSTEVQCK
2	**28**	285–303	GGNVCVQAQDRYYNSSCSK
2	**29**	304–313	WACVPCRVRS

purification from the enzyme or buffer, revealed an ion at m/z 1986.5. This mass corresponds to a disulfide coupling of peptides Glu^{261}-Lys^{269} and Tyr^{296}-Lys^{303}, $(M + H)^+$ m/z calculated 1987.1, and indicates the presence of a disulfide bond between Cys^{266} and Cys^{301}. This assignment is in agreement with the prompt fragmentation occurring during the measurement in Fig. 3 at m/z 2070 and 3115 that is consistent with the cluster breaking up into pairs of peptides **27–29** (Cys^{283}, Cys^{306}, and Cys^{309}) and peptides **25–28** (Cys^{266}, Cys^{289}, and Cys^{301}), respectively. The remainder of the disulfide pairings in this fraction are undetermined due to the inability to cleave at either the valine or proline residue between Cys^{306} and Cys^{309}. Digestion with prolyl endopeptidase was attempted, but did not yield any peptide bond cleavage.

Sequencing of fraction **b** showed only a single peptide, Cys^{106}-Lys^{128}, that contains both Cys^{106} and Cys^{117}. LSIMS analysis at a resolution of approximately $m/\Delta m$