Defining the Disulfide Bonds of Insulin-like Growth Factor-binding Protein-5 by Tandem Mass Spectrometry with Electron Transfer Dissociation and Collision-induced Dissociation**\[5\]

Mahta Nili,1 Aditi Mukherjee, Ujwal Shinde, Larry David, and Peter Rotwein2
From the Department of Biochemistry and Molecular Biology, Oregon Health & Science University, Portland, Oregon 97239-3098

**Background:** Only 2 of 9 putative disulfide bonds have been mapped for IGFBP-5.

**Results:** Using a MS-based strategy combining ETD and CID, and ab initio molecular modeling, we have mapped all 9 disulfide bonds in IGFBP-5.

**Conclusion:** Our results provide new insights into the IGFBP-5 structure.

**Significance:** We define an approach using tandem MS and ab initio molecular modeling to characterize unknown disulfide linkages in proteins.

The six high-affinity insulin-like growth factor-binding proteins (IGFBPs) comprise a conserved family of secreted molecules that modulate IGF actions by regulating their half-life and access to signaling receptors, and also exert biological effects that are independent of IGF binding. IGFBPs are composed of cysteine-rich amino- (N-) and carboxyl- (C-) terminal domains, along with a cysteine-poor central linker segment. IGFBP-5 is the most conserved IGFBP, and contains 18 cysteines, but only 2 of 9 putative disulfide bonds have been mapped to date. Using a mass spectrometry (MS)-based strategy combining sequential electron transfer dissociation (ETD) and collision-induced dissociation (CID) steps, in which ETD fragmentation preferentially induces cleavage of disulfide bonds, and CID provides exact disulfide linkage assignments between liberated peptides, we now have definitively mapped 5 disulfide bonds in IGFBP-5. In addition, in conjunction with ab initio molecular modeling we are able to assign the other 4 disulfide linkages to within a GCGCCXXC motif that is conserved in five IGFBPs. Because of the nature of ETD fragmentation MS experiments were performed without chemical reduction of IGFBP-5. Our results not only establish a disulfide bond map of IGFBP-5 but also define a general approach that takes advantage of the specificity of ETD and the scalability of tandem MS, and the predictive power of ab initio molecular modeling to characterize unknown disulfide linkages in proteins.

**The two closely related peptide growth factors**, insulin-like growth factor-I and -II (IGF-I and IGF-II), are necessary for normal growth and development in mammals and other vertebrates, and exert biological effects that promote proliferation, differentiation, and/or survival of a variety of cell and tissue types (1–3). In the circulation and in the extracellular space, IGFs are normally bound to one of six members of a conserved family of IGF-binding proteins (IGFBPs), which modulate IGF actions by regulating IGF half-life and access to cell surface signaling receptors (4). Several studies also suggest that IGFBPs control other biological processes that are independent of their IGF binding properties (5, 6). Each IGFBP mediates both unique and overlapping actions based in part on tissue- and developmental-stage specific patterns of expression, and on different affinities for each IGF and for other bioactive molecules (6, 7).

The six IGFBPs are secreted proteins of 201–289 amino acids in length (6, 8) and share ~36% sequence identity (8). Each IGFBP contains highly conserved N- and C-terminal domains, along with a less conserved central linker segment (6, 9). Most IGFBPs have 12 and 6 cysteine residues in their N- and C-terminal domains, respectively, but lack cysteines in the linker region. Exceptions include IGFBP-4, with two cysteines in its linker segment (10), and IGFBP-6, with only 10 cysteines in its N-terminal domain (11). In addition, IGFBPs 1–5 share a cysteine-rich motif, GCGCCXXC (where X is any amino acid), within the N-terminal domain (6, 8, 12). Limited insights into the three-dimensional organization of IGFBPs have come from results of high-resolution x-ray crystallographic analyses of the isolated N-terminal domain of IGFBP-4 and the C-terminal segments of IGFBP-1 and IGFBP-4 (13, 14). One consistent observation from these data is that IGFBPs lack inter-domain disulfide bonds (12–15). However, as the structure of a full-length IGFBP has not been solved, possibly because of the dis-
ordered nature of the linker segment, this conclusion remains provisional.

IGFBP-5, a 252-amino acid mature protein with 18 cysteine residues (16), is the most conserved IGFBP in mammals (17); for example, human and mouse IGFBP-5 are 97% identical to one another (18). IGFBP-5 has been found to be a key component of the IGF signaling axis in tissue repair and regeneration, and is able to regulate osteogenesis (19–22), muscle differentiation (23–27), and kidney development (28), among other processes (16). Furthermore, its deficiency in mice has led to increased growth but diminished glucose tolerance (29). IGFBP-5 also has been shown to exert IGF independent actions (5, 30–33). As with other IGFBPs, the N-terminal domain (residues 1–84) of IGFBP-5 encodes the primary IGF-binding site, with the C-terminal region (residues 165–252) contributing in a secondary way to binding stability and affinity (34–36). NMR and protein crystallographic studies of a portion of the N-terminal segment of IGFBP-5 (amino acids 40–92) have demonstrated that this part of the protein appears to be organized into a tight globular structure that contains an anti-parallel β-sheet stabilized by two disulfide bonds linking Cys57 to Cys60 and Cys54 to Cys68 (34, 36). Five residues located near these disulfide-linked cysteines (K68, P69, L70, L73, and L74) have been shown via mutagenesis studies to be major contributors to high affinity binding of IGF-I and IGF-II (34–37).

Disulfide bonds contribute to the proper folding of proteins and to the integrity and stability of their three-dimensional structures (38). Mass spectrometric (MS) methods for identifying disulfide bonds have improved over the last few years, and several proteins recently have been mapped using a tandem MS approach in which peptides, including disulfide-linked species, are selected in the MS1 scan and then subjected to ETD (ETD-MS2) followed by CID (CID-MS3) (39, 40). ETD fragments peptides via electron transfer from a radical anion to a proto-nated peptide, causing cleavage between Co-N bonds which results in c and z ions (41). However, in a disulfide-linked peptide, ETD has been demonstrated to preferentially cleave the disulfide bond rather than the peptide backbone (42). In contrast, CID rarely dissociates disulfide bonds, and generally fragments peptide backbones at the amide bond generating a series of y and b ions (43). Traditional MS disulfide mapping methodologies have employed CID, but only to compare protease-digested peptides in proteins treated with or without reducing agents (44). However, because ETD preferentially cleaves disulfide bonds, this approach may be applied to protein samples without prior chemical reduction (40), as subsequent CID fragmentation of the peptides liberated by ETD will then identify the cysteines involved in disulfide bonds (40).

Here we have characterized the disulfide linkage map for mouse IGFBP-5 by using a tandem MS approach combining ETD and CID. Our results definitively identify 5 of 9 disulfide bonds in the protein, and determine that the other 4 linkages involve the four cysteines within the conserved GCGCCMTC motif (residues 32–39) in the N-terminal region of the protein. As CID spectra of peptides containing this cysteine-rich motif could not precisely assign the disulfide linkages within the N-terminal segment of IGFBP-5, we employed unconstrained ab initio modeling to further refine the map. Overall, our results demonstrate the power of a combined approach employing both sequential MS and ab initio molecular modeling to identify and characterize disulfide bonds in a protein, and define a complete disulfide linkage map for IGFBP-5. In addition, we find that amino acid substitution mutations in N-terminal domain residues that are critical for maintaining ligand binding affinity (K68, P69, L70, L73, and L74) have a minimal impact on the global tertiary structure of IGFBP-5.

EXPERIMENTAL PROCEDURES

**Materials**—Fetal bovine serum (FBS), Dulbecco’s modified Eagle’s medium (DMEM), phosphate-buffered saline (PBS), and trypsin/EDTA were purchased from Invitrogen (Carlsbad, CA). Sequencing grade chymotrypsin was purchased from Roche (Indianapolis, IN). Proteomics grade trypsin and heparin agarose were from Sigma-Aldrich; Criterion precast gels were purchased from Bio-Rad. GelCode Blue Stain Reagent was purchased from Pierce Biotechnologies. NitroBind nitrocellulose was from GE Water & Process Technologies (Trevose, PA). Biotinylated human IGF-II was from GroPep (Adelaide, Australia). Polyclonal anti-IGFBP-5 antibody was purchased from Santa Cruz Biotechnology (Santa Cruz, CA). Secondary antibodies, AlexaFluor 680-conjugated goat-anti-mouse IgG and IR800-conjugated streptavidin were from Invitrogen and Rockland Immunochemical (Gilbertsville, PA), respectively. Other chemicals and reagents were purchased from commercial suppliers.

**Expression and Purification of IGFBP-5**—Recombinant adenoviruses encoding the tetracycline transactivator protein (Ad-tTA), mouse IGFBP-5, and a modified mouse IGFBP-5 with amino acid substitutions within the IGF binding region (N mutant) have been described (24). C3H10T1/2 mouse embryonic fibroblasts (CCL226; ATCC, Rockville, MD), incubated at 37 °C in humidified air with 5% CO₂ in DMEM with 10% fetal calf serum, were infected at ~50% of confluent density with Ad-tTA plus either wild-type or N-terminal mutant IGFBP-5 at a multiplicity of infection of 500. The following day medium was replaced with DMEM plus 2% fetal bovine serum. IGFBP-5 was purified from medium conditioned for 48 h using heparin-affinity chromatography, as described (24), and was stored in aliquots at ~80 °C until use.

**Immunoblotting and Ligand Blotting**—For immunoblotting, purified IGFBP-5 (500 ng/lane) was separated by reducing SDS-PAGE and transferred to nitrocellulose membranes. After blocking with 50% AquaBlock solution for 1 h at 20 °C, membranes were incubated sequentially with α-IGFBP-5 (1:500 dilution) and AlexaFluor 680-conjugated goat-anti-mouse IgG (1:10,000). For ligand blotting, purified IGFBP-5 (100 ng/lane) was resolved by non-reducing SDS-PAGE and transferred to nitrocellulose membranes. After blocking in 50% AquaBlock, membranes were incubated with biotin-conjugated IGF-II (100 ng/ml) in 50% AquaBlock in PBS with 0.1% Tween-20 for 16 h at 4 °C. Membranes were then washed with PBS plus 0.1% Tween-20, followed by incubation with IR800-conjugated streptavidin (1:10,000). Results were visualized and images captured using the Odyssey® Infrared Imaging System and version 3.0 analysis software (LI-COR, Lincoln, NE).
Disulfide Mapping of IGFBP-5

Digestion of IGFBP-5—Thawed IGFBP-5 (6 μg) was incubated protected from light with iodoacetamide (5 mM) with shaking for 30 min at 20 °C in buffer with 4 m urea, and then was separated by non-reducing SDS-PAGE using Criterion precast gels. Alternatively, IGFBP-5 was incubated in the gel after electrophoresis was completed with or without iodoacetamide. Protein bands were stained with GelCode Blue, de-stained with double deionized water, excised, and incubated twice in 500 μl of 50 mM ammonium bicarbonate, 50% (v/v) acetonitrile while shaking for 30 min at 20 °C. Samples were dehydrated in 100% acetonitrile for 2 min, dried by vacuum centrifugation, and rehydrated with 10 mg/ml of trypsin or chymotrypsin in buffer containing 50 mM ammonium bicarbonate and 5 mM calcium chloride for 15 min on ice. Excess buffer was removed and replaced with 50 μl of the same buffer without enzyme, followed by incubation for 16 h at 37 °C or 20 °C, for trypsin and chymotrypsin, respectively, with shaking. Digestions were stopped by addition of 3 μl of 88% formic acid, and after brief vortexing, the supernatant was removed and stored at −20 °C until analysis.

Localization of Disulfide Bonds by Mass Spectrometry—Peptides were injected onto a 1 mm × 8 mm trap column (Michrom BioResources, Inc., Auburn, CA) at 20 μl/min in a mobile phase containing 0.1% formic acid. The trap cartridge was then placed in-line with a 0.5 mm × 250 mm column containing 5 mm Zorbax SB-C18 stationary phase (Agilent Technologies Inc., Santa Clara, CA), and peptides separated by a 2–30% acetonitrile gradient over 90 min at 10 μl/min with a 1100 series capillary HPLC (Agilent Technologies). Peptides were analyzed using a LTQ Velos linear ion trap with an electron transfer dissociation (ETD) source (Thermo Scientific, San Jose, CA). Electrospray ionization was performed using a Captive Spray source (Michrom Bioresources, Inc.). Survey MS scans were followed by 7 data-dependant scans consisting of collision-induced-dissociation (CID) and ETD MS2 scans on the most intense ion in the survey scan, followed by 5 MS3 CID scans on the 1st to 5th most intense ions in the ETD MS2 scan. CID scans used normalized collision energy of 35, and ETD scans used a 100 ms activation time with supplemental activation enabled. Minimum signals to initiate MS2 CID and ETD scans were 10,000, minimum signals for initiation of MS3 CID scans were 1000, and isolation widths for all MS2 and MS3 scans were 3.0 m/z. The dynamic exclusion feature of the software was enabled with a repeat count of 1, exclusion list size of 100, and exclusion duration of 30 s. These experiments used inclusion lists to target specific cross-linked species for collection of ETD MS2 scans. Separate dta files for MS2 and MS3 scans were created by Bioworks 3.3 (Thermo Scientific) using ZSA charge state analysis. Matching of MS2 and MS3 scans to peptide sequences was performed by Sequest (V27, Rev 12, Thermo Scientific), using a database consisting of reversed yeast sequence entries, supplemented with the sequences of common contaminants, and the sequence of mouse IGFBP-5 (6182 entries total). The analysis was performed without enzyme specificity, a parent ion mass tolerance of 2.5, fragment mass tolerance of 1.0, and a variable mass of +16 for oxidized methionine residues. Searches of CID MS2 and MS3 data both specified matches to y and b ions. Results then were analyzed using the program Scaffold (V3.00.08, Proteome Software, Portland, OR) (45, 46) with minimum peptide and protein probabilities of 95 and 99% being used. IGFBP-5 peptides from MS3 results were sorted by scan number, and cysteine containing peptides were identified from groups of MS3 scans produced from the 5 most intense ions observed in ETD MS2 scans. The identities of cysteine peptides participating in disulfide-linked species were further confirmed by manual examination of the parent ion masses observed in the survey scan and the ETD MS2 scan.

IGFBP-5 ab initio and Homology Modeling—Structural models for an N-terminal segment, consisting of amino acids 5–41 and 1–84 of mouse IGFBP-5, were constructed using Rosetta ab initio modeling (47). Structures were generated using the standard Rosetta fragment server. The fragment selection procedure was performed de novo (without templates from existing structural homologues in the Protein Data Bank (PDB)) and also with access to PDB homologues. The 5,000 independent predicted structures from each search were subjected to clustering analysis. The centers of the five largest clusters were chosen as the best models, defined as having the lowest standard deviation of the mean among positions of α-carbon atoms of all residues when compared with all other simulations in a cluster. A homology model of amino acids 5–84 of IGFBP-5 also was built using the alignment interface of SwissModel, which predicts structures reliably with a root mean square deviation <2 Å for sequences with 50–60% identity (48), and using as a template the x-ray structure of the IGFBP-4 N-terminal domain (PDB Code 2DSR).

RESULTS

An MS Approach to Mapping Disulfide Bonds in IGFBP-5—There is little information on the disulfide bonding pattern of IGFBP-5, as only 4 of 18 cysteines in the protein have been mapped to date (34). Here we have applied a tandem MS approach combining ETD and CID to identify the linked cysteine residues in wild type mouse IGFBP-5 and in a N-terminal amino acid substitution mutant (Fig. 1A). Both proteins were purified after overexpression of recombinant adenoviruses in cultured mammalian cells (Fig. 1B). The 18 cysteines of IGFBP-5 potentially reside in 11 tryptic peptides, which could form up to 9 disulfide bonds (Fig. 1C). Purified IGFBP-5 was digested with trypsin under non-reducing conditions and subjected to a MS3 protocol with ETD (ETD-MS2) followed by CID (CID-MS3) (40). CID-MS2 also was employed to help identify disulfide-linked peptides (Fig. 1A). The tryptic peptides analyzed by these methods covered 87% of IGFBP-5, and we could account for each of the 18 cysteines in the protein (data not shown). The m/z values for cysteine-containing tryptic peptides are found in supplemental Table S1.

IGFBP-5 used in our assays was prepared by overexpression in mammalian cells, and there was the possibility that these proteins underwent disulfide scrambling during or after biosynthesis. Although we could not prevent any re-arrangements that occurred during protein maturation, to limit any subsequent scrambling we alkylated IGFBP-5 with iodoacetamide. Therefore, any potentially free cysteine residues were blocked prior to SDS-PAGE and in-gel protease digestion of purified
IGFBP-5. We obtained identical disulfide linkage assignments with and without alkylation (data not shown), and also observed highly similar chromatographic elution profiles for trypsin-digested IGFBP-5 and its N-terminal amino acid substitution mutant in multiple experiments (Fig. 1D). Therefore, based on the reproducibility of these results, we believe that our data accurately reflect the real disulfide map.

**Mapping Disulfide Bonds in the C-terminal Domain of IGFBP-5**—ETD was used to fragment the linked tryptic peptides, P9, P10, and P11 within the C-terminal domain of IGFBP-5 (Fig. 2A). Cysteine residues 221 and 223 in P10 are separated by a single amino acid, making it initially difficult to identify specific fragment ions in the CID-MS2 step to assign exact disulfide-linkages involving these two cysteines. However, in the CID-MS3 step following ETD-MS2 fragmentation in which P9 was released, we recovered an abundant fragment ion, y^4, allowing us to assign the linkage between Cys^{223} in P10 and Cys^{243} in P11 (Fig. 2B). Based on these results we then can deduce that Cys^{210} in P9 is bonded to Cys^{221} in P10. The y and b ions from the CID-MS3 fragmentation step supporting the latter assignment are found in supplemental Table S2.

We next used CID-MS2 and ETD-MS2 to identify the remaining disulfide bond in the C-terminal domain, which we find linked Cys^{172} and Cys^{199} (Fig. 2C and D). CID-MS2 y and b ions are listed in supplemental Table S3. Taken together, results in Fig. 2 show that there are three disulfide bridges within the C-terminal segment of IGFBP-5, and also indicate that there are no cysteine linkages that join the C- and N-terminal regions of the protein.

**Mapping Disulfide Linkages in the N-terminal Domain of IGFBP-5**—By CID-MS2 we detected 10 of the 12 cysteines in the N-terminal domain of IGFBP-5 within 4 linked
peptides connecting tryptic fragments P1, P2, P3, and P5 (Fig. 3). To resolve these potentially highly intertwined disulfide bonds, we first analyzed the previously identified disulfide-linkage between Cys47 in P3 and Cys60 in P5 (34, 36) by ETD-MS2 followed by CID-MS3 of the precursor ion, m/z 604.8 (+7), corresponding to P9-P10-P11, and containing cysteine residues 210, 221, 223, and 243. B. CID-MS3 of precursor ion, m/z 1128.5 (+3), of disulfide-linked peptides P10 and P11. The y ions marked with an asterisk are from P10; other y ions are from P11. C. CID-MS2 of precursor ion, m/z 521.5 (+4) of disulfide-linked peptides containing cysteine residues 172 and 199. The y ions are from P8, D. ETD-MS2 of precursor ion m/z 521.5 (+4). In A, D [M] denotes the parent mass. Additional y ions were identified in B and C, but have been removed from the graphs for clarity.

To provide additional support for these findings IGFBP-5 was digested with chymotrypsin and subjected to the MS3 protocol. We identified two linked peptides (C4 and C5) that confirmed disulfide bonds between Cys47 and Cys60 (Fig. 4A, supplemental Table S1). Three other linked peptides containing six cysteine residues (Cys 7, 10, 18, 33, 35, 36) also were detected after CID-MS2 (C1, C2, C3, Fig. 4B). Digestion of IGFBP5 with chymotrypsin also should generate two cysteine-containing peptides of 4 amino acids each, GC25EL and TC39AL. The fact that neither of these peptides were found to be associated with the larger group of N-terminal chymotrypsin peptides (Fig. 4B), suggests that Cys25 and Cys39 are disulfide bonded to one another, in agreement with analysis of the tryptic peptides (Fig. 4A).
However, as we did not recover this small putative disulfide-linked peptide, we cannot definitively reach this conclusion.

As described above, disulfide bonds involving the N-terminal \textsuperscript{32}GCGCCMT\textsuperscript{39} motif were not definitively established, because mapping the linkages between peptides with multiple cysteines by CID-MS3 requires ample spacing between individual cysteine residues to assign y and b ions. Unfortunately, neither protease digestion nor chemical cleavage strategies were able to separate these cysteines from one another. Because of these difficulties, we employed \textit{ab initio} molecular modeling to collect information on all possible combinations of disulfide bonds within this region. For these experiments we limited our analyses to amino acids 5–41 of mouse IGFBP-5, which contained the 8 cysteines whose linkages could not be resolved completely by our tandem MS approach. We first employed \textit{de}
novo ab initio modeling using Rosetta (47), in which homologous structures in PDB, such as other IGFBPs, are not used as templates to guide predictions. Using this approach, we were able to generate multiple highly related predicted structures (Fig. 5A) that were remarkably similar to the structures obtained using a homology-based search (Fig. 5B). Both models also identified identical disulfide linkages involving the GCG-CMTC motif: Cys33-Cys7, Cys35-Cys10, Cys36-Cys18, and Cys39-Cys25 (Fig. 5C). Moreover, a de novo ab initio model of the complete N-terminal domain of IGFBP-5 (amino acids 1–84) aligned closely with predictions for IGFBP-5 based on the x-ray crystallographic structure of the N-terminal segment of IGFBP-4 (Fig. 5D). Taken together, these results support the validity of using molecular modeling as part of a combined experimental approach with the MS3 protocol described here for delineating previously undefined disulfide linkages in proteins.

Amino Acid Substitution Mutations in the N-terminal Domain of IGFBP-5 that Reduce IGF Binding Do Not Alter Disulfide Bonds—Previous NMR studies had mapped the disulfide bond between Cys54 in P4 and Cys80 in P6, using a mini-IGFBP-5 protein as the starting material (36). We analyzed this linkage in full-length IGFBP-5 and compared it to results obtained with a N-terminal domain amino acid substitution mutant involving residues K68N, P69Q, L70Q, L73Q, and L74Q to determine whether disruption of this disulfide bond might account for the diminished IGF binding affinity of the latter protein (35, 37). Both wild type and N-mutant IGFBP-5 were digested with trypsin and subjected to the MS3 protocol (Fig. 1A). The Cys54-Cys80 linkage was identified between peptides P4 and P6, precursor ion of m/z 352.8 (+4), by CID-MS2 and ETD-MS2 in both wild type and N-mutant IGFBP-5 (Fig. 6). In addition, since the complete elution profile of cysteine-containing peptides of the N-terminal IGFBP-5 mutant matches that of the wild type protein (Fig. 1D), these results imply that the overall tertiary structure of the N-terminal mutant is not perturbed.

DISCUSSION

In this study we have shown that the 18 cysteines in mouse IGFBP-5 form 9 disulfide bonds. Application of a tandem MS approach employing ETD and CID directly identified 5 disulfide linkages: Cys47-Cys60, Cys54-Cys80, Cys172-Cys199, Cys210-Cys221, and Cys223-Cys243, and the combination of MS with ab initio molecular modeling established the most likely arrangement of the other 4 disulfide pairs: Cys7-Cys33, Cys10-Cys36, Cys18-Cys39, and Cys25-Cys39. Taken together, our studies show that IGFBP-5 is composed of structurally independent N- and C-terminal domains, containing 6 and 3 disulfide bonds, respectively.

To date there has been no x-ray crystal structure reported for full-length IGFBP-5, nor for any other full-length IGFBP, although the complete C-terminal domains of IGFBP-1 and IGFBP-4 have been characterized (12–14, 34). Based on these structural data, on amino acid sequence similarity with
IGFBP-1 and IGFBP-4, and on concordance in the location of cysteine residues between the two proteins, it is likely that the C-terminal segment of IGFBP-5 also is composed of a thyroglobulintype-1 fold consisting of an $\alpha$-helix and three-stranded antiparallel $\beta$-sheets held in a compact formation through the 3 disulfide bonds.

Amino acid sequencing and MS methods have been used previously to map some of the disulfide linkages in several IGFBPs. Protease cleavage followed by N-terminal sequencing was employed to identify the 3 disulfide bonds in the C-terminal segment of IGFBP-2 (49), and electrospray ionization (ESI)-MS was used to identify all 8 disulfide linkages in IGFBP-6 (11). In both of these cases the characterized disulfide bonds in the C-terminal domain match the results we have established here for IGFBP-5. Thus, in conjunction with x-ray crystallographic data for the C-terminal segments of IGFBP-1 and IGFBP-4, it is likely that the C-terminal domains of all five IGFBPs adopt a very similar overall

**FIGURE 6. Amino acid substitution mutations in the N-terminal domain of IGFBP-5 do not alter disulfide-linkages.**

A. CID-MS2 m/z 352.8 (+4), of disulfide-linked peptides P4-P6 from wild type (top) and N-terminal mutant (bottom) IGFBP-5 containing cysteine residues 54 and 80. The y and b ions marked with an asterisk are from P4; other y and b ions are from P6. Additional y and b ions were identified, but have been removed from the graph for clarity. B. ETD-MS2 spectra of precursor ions, m/z 352.8 (+4) from wild type (top) and N-terminal mutant (bottom) IGFBP-5 containing cysteine residues 54 and 80. [M] represents parent mass.
Disulfide Mapping of IGFBP-5

conformation with only slight differences in secondary structural features.

Recently in a search for new antimicrobial peptides, Osaki et al. discovered in cell-conditioned tissue culture medium a disulfide-linked amidated peptide containing amino acids 193–214 derived from the C-terminal portion of IGFBP-5, in which Cys" was bonded to Cys" (50). Although surprisingly, we also identified this Cys"–Cys" linkage in our analyses, but it was present as a very minor peptide species and was not detected in all protein samples evaluated (data not shown). In contrast, disulfide linkages Cys"–Cys", Cys"–Cys", and Cys"–Cys" were the dominant pairings found in every purified IGFBP-5 protein sample that we analyzed. Clearly, further studies will be needed to elucidate the biochemical mechanisms responsible for generation of this potentially alternative peptide from full-length IGFBP-5, and to define the structural features responsible for its novel biological properties.

The N-terminal domain of IGFBP-4 consists of a series of disulfide bridges that leads to a globular base, structural features that may define the IGF binding motif (12). In the N-terminal segment of IGFBP-5, two disulfide bonds analogous to two of the six disulfides in IGFBP-4, Cys–Cys and Cys–Cys, had been identified previously using solution-based NMR and x-ray crystallography of a mini IGFBP-5 N-terminal domain protein (amino acids 40–92) (34, 36). We now confirm these assignments in full-length IGFBP-5. Amino acids 32–39 within the N-terminal part of IGFBP-5 comprise a conserved motif of GCGCCMTC that is found (as GCGCCX) in IGFBPs 1–4 (6, 8). We establish here that the four cysteine residues in this motif form the disulfide bonds that connect with cysteines 7, 10, 18, and 25 at the extreme N terminus of IGFBP-5, a conclusion reached in conjunction with the application of de novo ab initio modeling. Overall, as depicted in Fig. 5C, it is likely that the three-dimensional structure of the N-terminal domain of IGFBP-5 is very similar to IGFBP4, and we predict that IGFBP-1–3 (6) will exhibit analogous structural features.

A series of engineered amino acid substitutions within the N-domain of IGFBP-5 (K68N, P69Q, L70Q, L73Q, L74Q) results in a nearly 100-fold decline in binding affinity for IGF-I and IGF-II (35, 37). Despite this major perturbation in IGF binding capability, our results show that the disulfide-binding pattern of the cysteines flanking these mutations is not compromised. Thus, lower affinity binding of IGFs to this mutant IGFBP-5 does not reflect lack of structural integrity, but rather represents a loss of key interactions between the two molecules.

Traditional approaches for mapping disulfide bonds have relied on a strategy comparing data generated with and without reductive alkylation, in which peptides isolated from the protein of interest after single or multiple proteolytic digestions were subjected to MS or other analytical methods (44). As ETD causes preferential cleavage of disulfide bonds rather than the peptide backbone it can obviate the need for reducing agents (40, 43). A subsequent CID step then can facilitate identification of individual peptides (39, 40). Based on our current experience, we can envision the development of more optimized approaches for determining the location of disulfide bonds in proteins in which no structural data are available.

In summary, we have used an MS-based strategy combining ETD and CID steps coupled with ab initio molecular modeling to elucidate the disulfide-bond map for IGFBP-5. Our results represent an extension of recent observations employing tandem MS to identify disulfide linkages in an immunoglobulin light chain (40), in human growth hormone (40), and in tissue plasminogen activator (39), three proteins in which the disulfide map had been known previously. Similar combinatorial approaches that also take advantage of the rapidly improving computational landscape of molecular modeling (51, 52) should be applicable to other proteins in which the number or pattern of disulfide bonds is unknown.

Acknowledgments—We thank Tony Capps, Debra McMillen, and John Klimek for experimental and technical advice, and the assistance of the OHSU Shared Proteomics Resource.

REFERENCES
1. Nakae, J., Kido, Y., and Accili, D. (2001) Distinct and overlapping functions of insulin and IGF-1 receptors. Endocr. Rev. 22, 818–835
2. Samani, A. A., Yakar, S., LeRoith, D., and Brodt, P. (2007) The role of the IGF system in cancer growth and metastasis: overview and recent insights. Endocr. Rev. 28, 20–47
3. Maki, R. G. (2010) Small is beautiful: insulin-like growth factors and their role in growth, development, and cancer. J. Clin. Oncol. 28, 4985–4995
4. Clemmons, D. R. (1998) Role of insulin-like growth factor binding proteins in controlling IGF actions. Mol. Cell Endocrinol. 140, 19–24
5. Mohan, S., and Baylink, D. J. (2002) IGF-binding proteins are multifunctional and act via IGF-dependent and -independent mechanisms. J. Endocrinol. 175, 19–31
6. Daza, D. O., Sundström, G., Bergqvist, C. A., Duan, C., and Larhammar, D. (2011) Evolution of the insulin-like growth factor binding protein (IGFBP) family. Endocr. Rev. 152, 2278–2289
7. Bach, L. A., Headley, S. J., and Norton, R. S. (2005) IGF-binding proteins—the pieces are falling into place. Trends Endocrinol. Metab. 16, 228–234
8. Hwa, V., Oh, Y., and Rosenfeld, R. G. (1999) The insulin-like growth factor–binding protein (IGFBP) superfamily. Endocr. Rev. 20, 761–787
9. Firth, S. M., and Baxter, R. C. (2002) Cellular actions of the insulin-like growth factor binding proteins. Endocr. Rev. 23, 824–854
10. Zhou, R., Diehl, D., Hoeflich, A., Lahm, H., and Wolf, E. (2003) IGF-binding protein-4: biochemical characteristics and functional consequences. J. Endocrinol. 178, 177–193
11. Neumann, G. M., and Bach, L. A. (1999) The N-terminal disulfide linkages of human insulin-like growth factor-binding protein-6 (hIGFBP-6) and hIGFBP-1 are different as determined by mass spectrometry. J. Biol. Chem. 274, 14387–14594
12. Sitar, T., Popowicz, G. M., Siwanowicz, L., Huber, R., and Holak, T. A. (2006) Structural basis for the inhibition of insulin-like growth factors by insulin-like growth factor-binding proteins. Proc. Natl. Acad. Sci. U.S.A. 103, 13028–13033
13. Sala, A., Capaldi, S., Campagnoli, M., Faggion, B., Labò, S., Perduca, M., Romano, A., Carrizo, M. E., Valli, M., Visai, L., Minchietti, L., Galliano, M., and Monaco, H. L. (2005) Structure and properties of the C-terminal domain of insulin-like growth factor-binding protein-1 isolated from human amniotic fluid. J. Biol. Chem. 280, 29812–29819
14. Siwanowicz, L., Popowicz, G. M., Wisniewska, M., Huber, R., Kuenke, K. P., Lang, K., Engh, R. A., and Holak, T. A. (2005) Structural basis for the regulation of insulin-like growth factors by IGF binding proteins. Structure 13, 155–167
15. Fernández-Tornero, C., Lozano, R. M., Rivas, G., Jiménez, M. A., Stämd- ter, L., Díaz-González, D., Forssmann, W. G., Cuevas, P., Romero, A., and Giménez-Gallego, G. (2005) Synthesis of the blood circulating C-terminal fragment of insulin-like growth factor (IGF)-binding protein-4 in its na-
Disulfide Mapping of IGFBP-5

JANUARY 6, 2012 • VOLUME 287 • NUMBER 2

Disulfide Mapping of IGFBP-5

factor-I with the N-terminal domain of IGFBP-5. EMBO J. 20, 3638–3644
35. Imai, Y., Moralez, A., Andag, U., Clarke, J. B., Busby, W. H., Jr., and Cler-
mons, D. R. (2000) Substitutions for hydrophobic amino acids in the N-
terminal domains of IGFBP-3 and -5 markedly reduce IGF-1 binding and alter
their biologic actions. J. Biol. Chem. 275, 18188–18194
36. Kalus, W., Zweckstetter, M., Renner, C., Sanchez, Y., Georgescu, J., Grol,
M., Demuth, D., Schumacher, R., Dony, C., Lang, K., and Holak, T. A.
(1998) Structure of the IGF-binding domain of the insulin-like growth
factor-binding protein-5 (IGFBP-5) implications for IGF and IGF-I re-
ceptor interactions. EMBO J. 17, 6558–6572
37. Shand, J. H., Beattie, J., Song, H., Phillips, K., Kelly, S. M., Flint, D. J., and
Allan, G. J. (2003) Specific amino acid substitutions determine the differ-
ential contribution of the N- and C-terminal domains of insulin-like growth
factor (IGF)-binding protein-5 in binding IGF-1. J. Biol. Chem. 278, 17859–17866
38. Mamathambika, B. S., and Bardwell, J. C. (2008) Disulfide-linked protein
folding pathways. Annu. Rev. Cell Dev. Biol. 24, 211–235
39. Wu, S. L., Jiang, H., Hancock, W. S., and Karger, B. L. (2010) Identification
of the unpaired cysteine status and complete mapping of the 17 disulfides
of recombinant tissue plasminogen activator using LC-MS with electron
transfer dissociation/collision induced dissociation. Anal. Chem. 82,
5296–5303
40. Wu, S. L., Jiang, H., Lu, Q., Dai, S., Hancock, W. S., and Karger, B. L. (2009)
Mass spectrometric determination of disulfide linkages in recombinant
therapeutic proteins using online LC-MS with electron-transfer dissocia-
tion. Anal. Chem. 81, 112–122
41. Wiesner, J., Premslter, S., and Sickmann, A. (2008) Application of electron
transfer dissociation (ETD) for the analysis of posttranslational modifica-
tions. Proteomics 8, 4466–4483
42. Chrisman, P. A., Pitteri, S. J., Hogan, J. M., and McLuckey, S. A. (2005)
SO2- transfer electron ion-ion reactions with disulfide linked polypep-
side ions. J. Am. Soc. Mass Spectrom 16, 1020–1030
43. Mikesh, L. M., Ueberheide, B., Chi, A., Coon, J. J., Syka, J. E., Shabanowitz,
J., and Hunt, D. F. (2006) The utility of ETD mass spectrometry in pro-
etomic analysis. Biochem. Biophys. Acta 1764, 1811–1822
44. Gorman, J. J., Wallis, T. P., and Pitt, J. J. (2002) Protein disulfide bond
determination by mass spectrometry. Mass Spectrom Rev. 21, 183–216
45. Humphries, J. D., Byron, A., Bass, M. D., Craig, S. E., Pinney, J. W., Knight,
D., and Humphries, M. J. (2009) Proteomic analysis of integrin-associated
complexes identifies RCC2 as a dual regulator of Rac1 and Arf6. Sci. Signal
2, ra51
46. Keshamouni, V. G., Jagtap, P., Michailidis, G., Strahler, J. R., Kuick, R.,
Reka, A. K., Papoulas, P., Krishnapuram, R., Srinagaram, A., Standiford,
T. J., Andrews, P. C., and Omenn, G. S. (2009) Temporal quantitative
proteomics by iTRAQ 2D-LC-MS/MS and corresponding mRNA expres-
sion analysis identify post-transcriptional modulation of actin-cytoskele-
ton regulators during TGF-beta-Induced epithelial-mesenchymal transi-
tion. J. Proteome Res 8, 35–47
47. Bonneau, R., Strauss, C. E., Rohl, C. A., Chivian, D., Bradley, P., Malm-
strom, L., Robertson, T., and Baker, D. (2002) De novo prediction of three-
dimensional structures for major protein families. J. Mol. Biol. 322, 65–78
48. Schwede, T., Kopp, J., Guex, N., and Peitsch, M. C. (2003) SWISS-MODEL:
An automated protein homology-modeling server. Nucleic Acids Res.
31, 3381–3385
49. Forbes, B. E., Turner, D., Hodge, S. J., McNeil, K. A., Forsberg, G., and
Wallace, J. C. (1998) Localization of an insulin-like growth factor (IGF)
binding site of bovine IGFBP binding protein-2 using disulfide mapping and
deletion mutation analysis of the C-terminal domain. J. Biol. Chem. 273,
4647–4652
50. Osaki, T., Sasaki, K., and Minamino, N. (2011) Peptidomics-based discovery
of an antimicrobial peptide derived from insulin-like growth factor-binding
protein 5 J. Proteome Res 10, 1870–1880
51. Fleishman, S. J., Corn, E. J., Strauch, E. M., Whitehead, T. A., Karanicolos,
J., and Baker, D. (2011) J. Mol. Biol.
52. Khadb, S. A. (2009) The pro-
53. Forbes, B. E., Turner, D., Hodge, S. J., McNeil, K. A., Forsberg, G., and
Wallace, J. C. (1998) Localization of an insulin-like growth factor (IGF)
binding site of bovine IGFBP binding protein-2 using disulfide mapping and
deletion mutation analysis of the C-terminal domain. J. Biol. Chem. 273,
4647–4652
50. Osaki, T., Sasaki, K., and Minamino, N. (2011) Peptidomics-based discovery
of an antimicrobial peptide derived from insulin-like growth factor-binding
protein 5 J. Proteome Res 10, 1870–1880
51. Fleishman, S. J., Corn, E. J., Strauch, E. M., Whitehead, T. A., Karanicolos,
J., and Baker, D. (2011) J. Mol. Biol.
52. Khadb, S. A. (2009) The pro-
53. Forbes, B. E., Turner, D., Hodge, S. J., McNeil, K. A., Forsberg, G., and
Wallace, J. C. (1998) Localization of an insulin-like growth factor (IGF)
binding site of bovine IGFBP binding protein-2 using disulfide mapping and
deletion mutation analysis of the C-terminal domain. J. Biol. Chem. 273,
4647–4652
50. Osaki, T., Sasaki, K., and Minamino, N. (2011) Peptidomics-based discovery
of an antimicrobial peptide derived from insulin-like growth factor-binding
protein 5 J. Proteome Res 10, 1870–1880
51. Fleishman, S. J., Corn, E. J., Strauch, E. M., Whitehead, T. A., Karanicolos,
J., and Baker, D. (2011) J. Mol. Biol.
52. Khadb, S. A. (2009) The pro-
53. Forbes, B. E., Turner, D., Hodge, S. J., McNeil, K. A., Forsberg, G., and
Wallace, J. C. (1998) Localization of an insulin-like growth factor (IGF)
binding site of bovine IGFBP binding protein-2 using disulfide mapping and
deletion mutation analysis of the C-terminal domain. J. Biol. Chem. 273,