Pei Lun Tsai, Sung-Fang Chen and Sheng Yu Huang*

Mass spectrometry-based strategies for protein disulfide bond identification

Abstract: The formation of disulfide bonds is critical for stabilizing protein structures and maintaining protein functions. It is important to understand the linkages between multiple cysteine residues within a protein. In this review, the analytical approaches using mass spectrometry (MS) for disulfide linkage assignment are classified and discussed. Enzymatic digestion under appropriate conditions followed by various MS detection strategies remains the primary method for cysteine linkage analysis. In-source decay (ISD) and electron transfer dissociation (ETD) have been used to generate significant peptide signals that indicate the identities of peptides involved in disulfide bonds. In addition, chemical labeling and software algorithms were also developed to facilitate the automation of disulfide bond analysis. For proteins with complex disulfide structure, methods involving partial reduction coupled with differential alkylation were demonstrated to be useful. In the past two decades, MS has become one of the most valuable tools for protein disulfide bond analysis. It provides irreplaceable information including the peptide backbone sequences as well as the cysteine connection pattern when coupling with appropriate sample preparations. The related approaches with their unique features can be applied for different aims such as structural characterization or functional studies of proteins.

Keywords: disulfide bond; mass spectrometry; protein pharmaceuticals.

Introduction

Correct folding featured by disulfide linkages has significant influence on protein functions.

As the connection between cysteines cannot be predicted by amino acid sequence, detailed disulfide bond information becomes important for functional studies of proteins and quality assessment of biopharmaceuticals (Yano et al. 2002, Srebalus Barnes and Lim 2007). To analyze disulfide connections, conventional methods such as NMR (Sharma and Rajarathnam 2000, von Ossowski et al. 2006, Walewska et al. 2008) and Edman degradation (Haniu et al. 1994, John and Forssmann 2001) were widely used. However, both methods require a high concentration or large amount of homogeneous proteins and well-purified peptides after enzymatic digestion are needed for Edman sequencing. Mass spectrometry (MS) with soft ionization sources developed for biomolecules (Andersen et al. 1996, Winston and Fitzgerald 1997) serves as an alternative option for disulfide bond analysis. The formation of one disulfide bond results in a 2-Da reduction of molecular weight, which can be distinguished by most mass spectrometers. Albeit LC-MS/MS coupled with database search has become a popular tool for high-throughput protein identification and posttranslational modification analysis, it is still challenging for disulfide bond determination because database search engines are primarily designed for linear peptides only. Disulfide assignment from raw mass spectra still significantly relies on manual interpretation. It is a complicated task if the target protein is cysteine rich, or it contains unexpected folding pattern with unknown disulfide bonds.

Remarkable progress in high-resolution MS allows one to analyze proteins without digestion (Reid and McLuckey 2002, Kelleher 2004). Compared with the peptide-based bottom-up proteomics, the top-down approaches are more suitable for comprehensive analysis of purified proteins/recombinant proteins with posttranslational modifications because a higher sequence coverage can usually be obtained (Zhang and Ge 2011). It also preserves the labile modifications that are likely destroyed by bottom-up MS and introduce minimal artificial modifications, which result from lengthy sample preparation processes such
as overnight digestion. These distinct features make top-down MS ideal for analyzing protein pharmaceuticals such as monoclonal antibodies (Bondarenko et al. 2009, Tsybin et al. 2011). For disulfide bond assignment, top-down MS can be useful because it avoids possible disulfide rearrangement that usually occurs in mild alkaline conditions during trypsin digestion. Collision-activated dissociation (CAD) and electron-captured dissociation (ECD) equipped with liner ion trap/Fourier transform ion cyclotron resonance (FTICR) were reported to identify the five disulfide bonds in human salivary α-amylase without digestion (Peng et al. 2012). However, there are still technical challenges yet to be solved for top-down MS such as protein solubility, sensitivity, throughput, and the requirement of high-end instrument (Zhang and Ge 2011), which prohibits its universal use. For a protein-containing interlocked disulfide bonds, limited information can be acquired from top-down analysis. Therefore, disulfide bonds were mostly cleaved in top-down analysis to increase the protein backbone sequence coverage (Quinton et al. 2007, Zhang et al. 2012); consequently, the data regarding disulfide linkages were missing. So far, only a handful of articles demonstrated the use of top-down MS to solve cysteine connections.

This review focuses on MS-based strategies, which aim to solve the cysteine-connecting manner within a protein. As summarized in Figure 1, several approaches are presented: (A) peptide profile comparison before and after reduction, (B) partial reduction and stepwise alkylation, (C) analyses solely rely on computational algorithms, (D) utilizing in situ or on-line reduction to help identify peptide components during detection, and (E) chemical labeling designed specifically for disulfide-linked peptides. Various strategies were proposed with respect to different types of mass spectrometers equipped with MS/MS modes. Although we tried to classify these methods, it should be noted that one method may belong to more than one group due to its multiple features. Rather than providing an absolute classification, our intention is to give a general idea of the reported methods as well as their strengths and limitations. Enzymatic digestion is usually involved in these approaches as top-down MS for complex disulfide bond determination has not yet been popular and successful. In this case, one should be cautious to ascertain that the obtained disulfide structure truly reflects the native protein because disulfide rearrangement can happen during digestion.

**Figure 1** An overview of MS-based strategies for disulfide bond analysis. The flowchart in each group is just for illustration. Details may vary with methods.
General considerations in sample preparation

It is known that disulfide bonds are prone to rearrange in the presence of free thiols or under neutral to alkaline conditions (Sanger 1953, Ryle and Sanger 1955, Gorman et al. 2002). Therefore, the common procedure among various strategies is to block free cysteines followed by enzymatic digestion with a well-controlled pH condition before detection. The most popular alkylation reagents in use before typical trypsin digestions are iodoacetamide (IAM) and iodoacetic acid (IAA). Unfortunately, pH 8 is their optimal reaction environment, which likely results in disulfide shuffling, and over-alkylation toward the amine groups of lysines (Gundlach et al. 1959, Boja and Fales 2001) and protein N-terminus (Yang and Attygalle 2007) was reported in situation where the incubation time was longer than required. N-ethylmaleimide (NEM) is a widely used alkylation reagent to quench free thiols in disulfide bond analysis due to its thiol specificity and stability below pH 7 (Brewer and Riehm 1967, Hansen and Winther 2009). It was also found to be more efficient than IAA or IAM in terms of reaction time and reagent amount needed (Rogers et al. 2006, Hansen and Winther 2009). However, the three-dimensional structure of proteins may hinder the access of NEM and decrease the blocking efficiency (Lind et al. 2002). Thereby, denaturing reagents such as urea and guanidine hydrochloride are sometimes added to help unfold proteins (Yen et al. 2000). The disadvantage of urea is it contains slight amount of cyanate, which reacts with thiol groups to form thiocarbamates (Stark et al. 1960). Guanidine hydrochloride has, thus, been used as an alternative. Surfactant-based denaturation such as the use of 1–5% sodium dodecyl sulfate (SDS) is another option, but it was suggested to decrease the alkylation rates (Galvani et al. 2001). If the addition of a chaotrophic reagent or surfactant is inevitable, a desalting or a detergent-removal step before MS detection is generally required to prevent possible interferences.

The choice of protease depends on its optimal pH condition and cleavage capabilities. Trypsin is the most used enzyme to hydrolyze protein into peptides prior to MS detection as it offers adequate charge states and better b/y signals under collision-induced dissociation (CID) fragmentation by leaving basic amino acids on the C-terminals of peptides. However, the alkaline condition optimized for trypsin digestion was reported to cause disulfide bond rearrangement (Sanger 1953, Ryle and Sanger 1955). The problem can be solved by performing the digestion at suboptimal pH conditions or, alternatively, using an enzyme or chemical cleavage reagent that works well under acidic conditions. Pepsin, which prefers low pH, is a proper choice for disulfide bond analysis (Bures et al. 1998), but its broad specificity can complicate the identification. Endoproteinase Glu-C that works under both pH 4 and 8 is another favored choice (Lippincott and Apostol 1999). CNBr, a chemical reagent that cleaves at methionine in acidic pH, can be used to provide initial fragmentation (Chong et al. 2002). In addition to pH consideration, an ideal enzyme should be capable of hydrolyzing proteins into appropriate length and separating the cysteine residues onto different peptides to facilitate MS detection. Seldom can enzymes work alone to achieve this goal. Therefore, consecutive or simultaneous use of more than one enzyme is often required for proteins with a complicated disulfide structure. Otherwise, partial reduction has to be included as discussed later in this review. For a highly glycosylated protein, an extra deglycosylation step under appropriate condition that eludes disulfide exchange is usually necessary (Pitt et al. 2000, Go et al. 2011). Sample preparation methods to avoid disulfide scrambling before disulfide analysis have been well summarized and discussed (Tang and Speicher 2001).

Profile comparison

A conventional way to solve disulfide linkages is to analyze the molecular weights of peptides from nonreduced protein digest and derive the peptide components from possible combinations of cysteinyl peptides. It is a huge task to process all the signals without knowing which ones come from disulfide-linked peptides. Generally, a part of nonreduced protein digest is subjected to reduction, and the LC profiles between reduced and non-reduced peptides are compared. As shown in Figure 1A, the signals that are present in one run, but not in the other, are considered to be disulfide bond related and can be collected for subsequent identification. RP-HPLC/MS was used to analyze the native and reduced Lys-C digest from an IgG4 monoclonal antibody (mAb) (Zhang et al. 2002). By comparing the HPLC-UV chromatograms, the disulfide-linked peptides were localized, collected, and sequenced by offline nanoLC-MS/MS or Edman degradation. They also demonstrated that the two closely spaced and symmetrical disulfide bonds at the hinge region of mAb can be characterized by N-terminal Edman sequencing. A similar approach was demonstrated with nonreduced Glu-C digest of insulin. By peptide mapping between reduced and nonreduced samples, the peptide containing
two disulfide bonds was collected and manually derivatized/cleaved with Edman sequencing chemistry. The product from each cycle was analyzed by MS, and the one disulfide structure among the three possible connections could be determined (Zhang and Cockrill 2009). The same method was further applied to identify the Fab-hinge disulfide linkage of IgG2 (Zhang et al. 2010). MALDI-TOF was reported to analyze disulfide bonds of monoclonal antibodies. The protein was digested by endoproteinase Lys-C without reduction and detected by MS directly. On-target reduction was then performed by depositing dithiothreitol (DTT) on the same spot, and the sample was reanalyzed for the reduced peptide products (Mhatre et al. 1999). Owing to the ion suppression of MALDI, LC/ESI-MS was used complementarily to increase the number of observed nonreduced peptides. Pepsin digestion without reduction followed by HPLC-UV separations before and after reduction were utilized to determine the disulfide bond arrangement of Newcastle disease virus hemagglutinin-neuraminidase (NDV HN) (Pitt et al. 2000). Those peaks found in the nonreduced sample, but absent in the reduced one, were collected and analyzed by MALDI-TOF with and without reduction. Subsequent digestion with trypsin or CNBr or deglycosylation with PNGase F was performed for several fractions to decrease the complexity of disulfide-linked peptides with multiple chains or glycan attached and facilitate the following Edman sequencing or MS/MS detection. A similar approach involving the use of HPLC-UV, MADLI-MS, and ESI-MS was applied to solve the disulfide bonds of human signal transducer gp130 (Moritz et al. 2001).

To determine the disulfide structure by peptide mapping, in the absence and presence of a reducing agent, followed by MS is a straightforward way to locate disulfide-associated peptides from HPLC fractions. However, a larger sample amount is required for these iterative reduction fraction collection steps. Besides, the LC-UV profile or mass spectra may be complicated when dealing with a large protein containing numerous disulfide bonds especially when missed cleavage peptide products are involved due to nonreduced digestion condition.

**Partial reduction and differential alkylation**

The common problem encountered in disulfide bond analysis is either the inability to find a suitable enzyme to cut between highly bridged cysteines or that the sizes of disulfide-linked peptides resulted from nonreduced digestion are too big and, thus, beyond the MS detection range. Partial reduction with low concentration of tris(2-carboxyethyl)phosphine (TCEP) has been introduced to simplify the situation (Gray 1993). Compared to DTT and 2-mercaptoethanol, TCEP is more stable and works well at low pH, which is important to prevent disulfide bond rearrangement. The application of partial reduction after digestion can assist MS detection especially for multichain peptides with several disulfide bonds. Lu’s group used partial reduction followed by IAA alkylation for one collected HPLC fraction from nonreduced tumor necrosis factor binding protein digest. The disulfide-linked peptide consisting of four peptide chains becomes several carboxymethylated peptides with different chain numbers varying from 1 to 4. The three-disulfide bond linkage pattern can be solved from the molecular weight analysis of these products (Jones et al. 1997). A similar approach using partial reduction and NEM alkylation followed by LC-MS/MS detection was applied for the disulfide mapping of murine TIM-1 IgV domain protein (Foley et al. 2008). Stepwise alkylation was described to be coupled with partial reduction to determine closely spaced disulfide linkages such as the hinge region of IgG or peptides with multiple disulfide bonds (Yen et al. 2002). TCEP reaction followed by the first step alkylation with maleimide-biotin (M-biotin) or NEM was used to target more accessible disulfide bonds. Afterward, the complete reduction with DTT and alkylation by IAM was conducted to target the rest. These cysteinyl peptides were further identified by LC-MS/MS, as illustrated in Figure 1B. Only those cysteines modified by the same alkylation reagent could have connected to each other so that the close disulfide bond connection can be distinguished with this approach as long as they have different reduction rates. The crucial step was the TCEP concentration, which has been optimized to be 0.1–0.5 mM at 65°C for 10–15 min instead of 10–20 mM that can result in full reduction. The partially reduced products were reported to be stable at room temperature and could be stored for more than 4 weeks (Yen et al. 2002).

Partial reduction followed by differential alkylation at protein level was reported to assess the susceptibility of intrachain and interchain disulfide bonds of IgG1 molecules, in which 13C-iodoacetic acid and 13C-iodoacetic acid were used as the first and the second alkylation reagents, respectively, before enzyme digestion and LC-MS analysis (Liu et al. 2010). In addition to confirming that interchain disulfide bonds were more susceptible than intrachain ones, Liu et al. further determined that the disulfide bonds in CH2 domain were the most susceptible, and those in CH3 domain were the least susceptible. It was also noticed that the intrachain disulfide bonds were only reduced in the presence of guanidine hydrochloride. Cyanylation with
cyanodiaminopyridinium tetrafluoroborate (CDAP) after partial reduction was reported for disulfide bond analysis (Wu and Watson 1997, Yang et al. 1998). The N-terminal of cyanylated cysteine can be cleaved in alkaline condition leaving an iminothiazolidine group tag on the cysteine as an indicator of the original disulfide linkage. Afterward, a full reduction was performed to break the rest of the disulfide bonds and make peptides detectable by the following MALDI-MS analysis. It is especially useful for peptides containing sequential cysteines, and no enzyme can cut in between, thereby, hampering the disulfide bond identification. The concept of “negative signature mass algorithm” (NSMA) was introduced: The obtained mass spectrum contained information of all related products as indirect evidence that was used to rule out enough disulfide linkages so that only one disulfide structure can be constructed (Qi et al. 2003, Wu et al. 2004). The difficult part of cyanylation is that the fragments after cleavage may not be detected due to either the variation of reaction efficiency at different locations on peptides or the chemical background interference during MALDI analysis (Wu et al. 2004). The method was later modified by adding a second alkylation step with NEM and using enzyme digestion instead of chemical cleavage prior to detection with MALDI-ISD/PSD (Schnaible et al. 2002a).

The prerequisite of partial reduction-based disulfide bond analysis is to critically optimize the reduction condition, which varies with different proteins. If differential alkylation (stepwise alkylation) is used to distinguish closely located disulfide bonds, a desalting step or a buffer exchange is necessary in most approaches in order to remove the first alkylation reagent used. Ambiguous results may be obtained when the disulfide bonds have similar reduction rates (Jones et al. 1997). Besides, the partial reduction step plus denaturation may lead to disulfide scrambling. Despite these potential problems, partial reduction serves as a powerful tool when a protein with numerous cysteines in close proximity to one another is being analyzed, and there is no suitable enzyme to cut in between under non-reduced conditions. Partial reduction-incorporated sample preparation can be used in conjunction with other detection methods for solving complicated disulfide bonds.

**CID fragmentation of disulfide-linked peptides and software development**

The most straightforward way to determine disulfide bond arrangement is to analyze nonreduced peptides with minimal sample preparation. Tandem mass spectrometry, which provides peptide sequence information in addition to molecular weight values, has been proven useful for peptide identification at a higher confidence level. CID, the most popular fragmentation mode used in peptide sequence analysis, causes specific cleavage at amide bonds to yield b/y ion series. For disulfide-linked peptides, ideally, the b and y ions for each peptide chain can be observed, and then, the connection pattern can be identified. However, disulfide bonds under CID fragmentation can be either intact or broken. In addition to the typical b/y ions from individual peptide chains, the b/y ions containing cysteine residue, cysteine thioaldehyde (-2 Da), cysteine persulfide (+32 Da), or dehydroalanine (-34 Da) due to the breakage of S-S or C-S bond can also be observed (Zhang and Kaltashov 2006, Choi et al. 2010), as shown in Figure 2, and so are those fragment ions linked by intact disulfide bonds. Similar phenomenon was reported in the high-energy CID mode in MALDI TOF/TOF instrument (Janecki and Nemeth 2011). It is labor intensive and time consuming to examine all the MS/MS spectra for all possible cysteinyl peptide combinations due to spectra complexity. It is especially not practical when the disulfide linkage is unknown for a protein with significant number of cysteines. Therefore, several programs were designed to automatically interpret the CID spectra of disulfide-linked peptides. MassMatrix, which was originally developed for general protein database search (Xu and Freitas 2007), has included the function of analyzing disulfide-linked peptides by comparing the CID products of each peptide backbone. Peptides with up to two disulfide bonds were considered, and a statistical scoring model was employed. Certain types of disulfide-linked peptides, such as peptides with more than one intra-chain disulfide bond, may

---

**Figure 2** Fragmentation patterns of disulfide-linked peptides. For discussion of each fragmentation mechanism, please refer to the references quoted in the figure. *Under CID, disulfide bonds can be either intact or broken as discussed in the text.*
not be found by the program due to the lack of sufficient product ions (Xu et al. 2008). DBond further takes the four disulfide bond-specific fragments into account as shown in Figure 2 (Choi et al. 2010). Another software named MS2DB+ was developed to allow the consideration of multiple ion types such as a, a*, a0, b, b*, c, x, y, y*, and z ions in addition to the commonly observed b and y ions (Murad et al. 2011). None of these algorithms have interpreted the notable products caused from two peptide bond cleavages (or a double cleavage), which results in the so-called internal ions. The double cleavage under CID fragmentation was found to be common for disulfide-linked peptides and independent of charge state or peptide size (Clark et al. 2011). For peptides with more than one disulfide bond, double cleavage products involving only one disulfide are especially useful as they provide information on cysteine connections.

Figure 3 illustrates the MS² spectrum of a peptide pair derived from trypsin+GluC digestion of a monoclonal antibody. In addition to typical b/y ions from individual peptide chains, fragments such as B6+32 and y2+34 coming from C-S bond breakage were also observed. On the other hand, a variety of internal ions such as PEVTCV, PEVTCVV, and PEVTCVVVD, etc., resulting from the double cleavage, were found to be in association with the y2 ion (the *C*K molecule) via an intact disulfide bond. These ions are so prominent because the fragmentation tends to occur next to proline residues. Another ion, which contains an intact disulfide, is Y14y2. The MS² spectrum demonstrates that various types of ions can be observed under CID fragmentation for disulfide-linked peptides. A significant number of peaks are disulfide specific, but only part of them can be interpreted by developed algorithms. In such case, manual assignment to explain all the dominant peaks is still in great demand.

Using CID solely for disulfide bond analysis can be difficult due to the complicated spectra that come from various fragmentation possibilities. Even if a program covering all possible fragments is developed, a long processing time and a certain amount of false positives are still to be expected. Nevertheless, the approach involving nonreduced protein digestion followed by CID and search algorithm as illustrated in Figure 1C is automatic without the need of complicated sample preparation or a high-end mass spectrometer equipped with special MS/MS modes. It provides a good starting point for disulfide bond analysis.

**Figure 3** MS/MS spectrum of m/z 733.38 (3+) derived from trypsin+GluC digestion of bevacizumab, a monoclonal antibody, under non-reduced condition. The peptide pair **“TPeVTCVVVDVSHEDPE, *C*K** linked by the disulfide bond (C267–C327) was labeled with dimethyl group at the N-terminals and the lysine. Capital B/Y denotes the fragments from the upper peptide. Stars indicate the sites of dimethyl labeling. The m/z range above the a1 ion is increased by four times to make a clear annotation. The MS/MS spectrum illustrates various fragment types of disulfide-linked peptides under CID as well as the a1 ion enhancement feature that can be used to identify disulfide-linked peptides as discussed in the text.
Gas phase reduction for disulfide bond analysis

Disulfide-linked peptides behave differently under various types of fragmentation modes, as summarized in Figure 2. Some of them prefer the disulfide bond cleavage over peptide backbone fragmentation, which leads to the observation of individual peptides, and such information can be useful to determine the identities of peptides connected by disulfide bonds. In addition to these fragmentation modes such as ISD, ECD, ETD, negative mode detection and some other online reduction methods also allow simultaneous observation of disulfide-linked peptides and disulfide-dissociated peptides. The advantage of these gas phase reduction approaches is that they do not require performing two independent experiments (with and without chemical reduction) as described in the previous section.

Laser-based ionization was shown to induce disulfide bond cleavage (Qiu et al. 2007, Yang et al. 2009). MALDI-ISD generates \( R_1\text{SH} \) and \( R_2\text{SH} \) for a disulfide-linked peptide, \( R_1\text{S-S-R}_2 \) (Patterson and Katta 1994), as shown in Figure 2. A strategy for disulfide bond analysis based on ISD followed by peptide identification with other MS/MS fragmentation was developed. Once the signals on a mass spectrum was found to fulfill the equation \( m/z (\text{Peak A}) + m/z (\text{Peak B}) - m/z (\text{H}_2^+ + \text{H}^+) = m/z (\text{Peak C}) \), the putatively disulfide-linked peptides were subjected to PSD or LIFT-TOF/TOF analysis to confirm its identity via the fragment ions from individual peptide chains (Schnable et al. 2002b).

The computer program SearchXLinks was used to link up the ISD data with the tandem mass spectra (Wefing et al. 2006). In situ reduction methods using various MALDI matrices were reported to cleave disulfide-bridged peptides (Crimmins et al. 1995). In addition to typical chemical reduction on a target plate (Fischer et al. 1993, Pitt and Gorman 1996), 1,5-diaminonaphthalene (1,5-DAN) was applied as the MALDI matrix, and the reduction reaction of 1,5-DAN was found to occur in vacuum in a mass spectrometer and not in the liquid phase (Fukuyama et al. 2006). For peptides containing intrachain disulfide bonds, a mass shift indicating the disulfide bond numbers can be observed by comparing the mass spectra of using DHB and 1,5-DAN. MS/MS spectra of the reduced ions provide a series of product ions that facilitate amino acid sequencing. The reductive matrix was also found to enhance the ISD efficiency (Demeure et al. 2007). The combination of using on-target 1,5-DAN reduction and ISD was later used for top-down analysis of venom proteins (Quinton et al. 2007). One drawback of MALDI is that the offline HPLC separation is generally required to reduce ion suppression to allow all peptides to be observed.

The applications of ECD and ETD for protein modification analysis have been described in a great number of literatures and comprehensively reviewed as well (Meng et al. 2005, Bakhtiar and Guan 2006, Wiesner et al. 2008, Zhou et al. 2011). In addition to obtaining \( c \) and \( z \) ion types generated from N-Cα bond breakage, these electron-based fragmentations also prefer the cleavage of disulfide bonds (Zubarev et al. 1999, Chrisman et al. 2005), which leads to the observation of individual molecular weights of peptides like ISD. For intrachain disulfide-bonded peptides, ions including \( c-33, z+33, c+32, \) and \( z-32 \) can also be observed in ETD (Cole et al. 2012). A strategy using CID-MS\(^2\), ETD, and CID-MS\(^3\) of an isolated product ion derived from ETD was designed to identify the disulfide-linked peptides (Wu et al. 2009). For a peptide in which two disulfide bonds link up three peptide backbones (P1, P2, and P3), ETD provides not only the molecular weight of each peptide but also information about partially disulfide-dissociated peptide such as P1-P2 and P2-P3. The exact cysteine connection can be obtained by performing MS\(^3\) on P1-P2 and P2-P3. With the aid of SDS-PAGE separation, this approach was later used to identify the disulfide bonds of several therapeutic proteins including monoclonal antibodies as well as their scrambled forms under heat stress (Wang et al. 2011). Low-energy electron-based fragmentation methods provide valuable information on disulfide-dissociated peptides; however, it requires much more effort to interpret the spectra that contain abundant peptide backbone fragments.

Several more reduction-associated detection methods have been reported to preclude the use of chemical reductions. In negative CID, disulfide-linked peptides were known to undergo gas phase reduction (Chrisman and McLuckey 2002) due to the tendency of collision-activated peptide ions to lose preferentially the side chain of select amino acids such as cysteine. The in-source fragmentation was utilized for solving the IgG2 disulfide bonds except for the hinge region (Chelius et al. 2006). A modified method using broadband selection in a Q-TOF instrument was also reported (Zhang and Kaltashov 2006). Electrolytic reduction, another option for reducing disulfide bonds (Honeychurch 1997), has been online coupled with desorption electrospray ionization mass spectrometry (DESI) for disulfide bond study (Zhang et al. 2011). Before and after electrolyses, disulfide-linked peptides and dissociated peptide components can be observed, respectively. CID spectra for the individual peptides are then used for peptide identification to establish the disulfide linkage.
pattern. The online reduction was also applied in the top-down approach to assist the disulfide bond breakage and increase the protein sequence coverage (Zhang et al. 2012). Another MS-based method involving gold cationization was reported to facilitate disulfide cleavage. The incorporation of two gold cations into a disulfide-bonded protein was found to result in efficient cleavage of disulfide bonds under CID, and more sequence information can be generated (Mentinova and McLuckey 2011).

Substitution of chemical reduction in solution with gas phase dissociation of disulfide bonds results in shorter analysis time because lengthy sample-processing steps can be avoided. However, these detection-based strategies require mass spectrometers equipped with specific fragmentation modes, which might potentially prohibit them from becoming popular.

Chemical labeling for disulfide bond analysis

To locate the MS/MS spectra related to disulfide bonds can be difficult as one LC-MS/MS can generate ten to hundreds of thousands of spectra with an advanced instrument. If the disulfide structure of a protein with many cysteines is unknown, a long list of m/z, which came from possible molecular weight combinations from cysteiny1 peptides, has to be checked thoroughly. To find the needle in a haystack, there were approaches aiming to pick up the disulfide bond-associated spectra prior to data analysis. The most noticeable character of disulfide-linked peptides is its multiple N- or C-terminals. Chemical labeling techniques during or after digestion without reduction were developed in order to “mark” these peptides with multiple chains. Consequently, these peptides can be distinguished from linear peptides. The incorporation of $^{18}$O into the carboxyl group at the peptide C-terminal was achieved by performing trypsin digestion in $H_2^{18}$O-containing buffer, and the enzymatic labeling approach has been applied to various quantitation experiments (Mirza et al. 2008, Elliott et al. 2009, White et al. 2009). For disulfide bond analysis, pepsin digestion at pH 3–4 was used instead to avoid possible disulfide rearrangement despite of its lack of specificity (Wallis et al. 2001). Enzymatic digestion in 50% $H_2^{18}$O (v/v) in $H_2^{16}$O can generate unique isotopic profiles on mass spectra for proteolytically derived peptides (Rose et al. 1983, Whaley and Caprioli 1991). For multiple peptide chains linked by crosslinking reagents or disulfide bonds, the isotope pattern can be used to distinguish them from single-chain peptides. Different isotopic profiles suggest different linkage type, and thereby, a two- or three-chain constitution can be defined (Gorman et al. 2002).

One disadvantage of this method is that the mass spectra as well as the MS$^2$ spectra can be very complicated due to the increasing number of species. Another approach regarding the use of $^{18}$O/$^{16}$O is to divide the sample into two, and one of them is labeled with $^{16}$O, while the other is labeled with $^{18}$O. Compared to peptides with a single chain, those with multiple C-terminals exhibit a higher mass shift between the spectra acquired from $^{18}$O- and $^{16}$O-labeled samples and, thereby, can be picked out. A program called Pro-CrossLink is designed for crosslinking analysis. It includes the function of mass shift detection to locate the cross-linked candidates, identification of peptide pairs followed by the assignment of b/y ions on MS2 spectra (Gao et al. 2006).

Another labeling method targeting the multiple N-terminals of disulfide-linked peptides is dimethyl labeling, a fast and complete chemical reaction originally designed for quantitative proteomics analysis (Hsu et al. 2003). Formaldehyde was used to produce a 28-Da chemical tag (32 Da for formaldehyde-D$_2$) at each N-terminal or lysine, and enhanced a$_1$ ions, which offer the identities of N-terminal amino acids can be observed in the MS/MS spectra (Hsu et al. 2005). Multiple a$_1$ ions can be observed for peptides containing multiple N-terminals such as disulfide-linked peptides. Figure 3 illustrates the MS$^2$ spectrum showing that the dominant a$_1$ ions, 106.11 and 106.07 (resolved spectrum not shown), suggest Thr and Cys as the N-terminal amino acids. A customized program named RADAR was developed to automatically screen out the a$_1$ ion match from the peak list file and then search for molecular weight combination match (Huang et al. 2008). With the b and y ions of individual peptides as further evidence, the disulfide bonds of monoclonal antibodies can be solved with minimal labor effort (Huang et al. 2012).

Chemical labeling using either $H_2^{18}$O or formaldehyde can be useful for disulfide bond analysis due to its specificity. A$_1$ ion screening with RADAR has great potential for full automation as a routine quality control step for biopharmaceutical manufacturing. With such methods, it is not necessary to inspect all the spectra, which is usually thought to be a labor-intensive process. However, for nearby cysteines that cannot be separated by enzyme digestion, it still requires informative MS$^2$ spectra to distinguish close disulfide bonds. Complementary methods such as partial reduction or the addition of a second enzyme may be included to assist unambiguous disulfide determination.
Summary

MS-based approaches have been proven extremely useful for disulfide bond analysis. Various methods can be selected for different purposes based on instrument availability as well as the complexity of a protein in terms of protein size and the number of cysteine residues. In general, all methods described in this review can be applied for proteins containing up to four disulfide bonds. In the case of a larger protein with more cysteines such as monoclonal antibodies being the target, ETD-based approaches, dimethyl labeling coupled with RADAR, and MALDI with ISD were shown to be successful. To distinguish between cysteines in close proximity, partial reduction followed by differential alkylation can be involved. If it is a protein with complicated disulfide structure, which might result in hindered cleavage sites, digestion efficiency has to be optimized to make disulfide-linked peptides fall into a detectable range. Partial reduction in the presence of denaturing reagents or multiple enzyme digestions can be helpful in this case. If a disulfide structure of a protein is not available at all, or degraded forms with unexpected disulfide bonds are desired, using software-based methods is a good starting point as trying all possible combinations from ten to hundreds of thousands of MS² spectra can be a tedious and impractical job. The number of cysteines in a protein can be used as a reference to decide if prefiltering algorithm should be involved to reduce the data complexity and relieve labor work. In summary, MS serves as an efficient tool for disulfide linkage analysis. In conjunction with different sample preparations, it provides significant support for protein functional studies and quality assessment of protein products for pharmaceutical industry.

Received March 9, 2013; accepted June 27, 2013; previously published online August 9, 2013

References


Pei-Lun Tsai, MSc, graduated from the Department of Chemistry at the National Taiwan University (1999). He has been a PhD Candidate at the Department of Chemistry, National Taiwan Normal University since 2011. He has also been a supervisor at Mass Solutions Technology Co., Ltd, Taiwan since 2004. His research interests include peptide mapping, protein disulfide bridge analysis, protein glycosylation site analysis, and carbohydrate structure analysis.

Sung-Fang Chen received his PhD degree in Analytical Chemistry from the National Cheng Kung University in 2006. For her PhD thesis, she received the Excellent Thesis Award of the Chinese Chemistry Society in 2007. She was granted scholarships from both the National Science Council and the DAAD Germany for her postdoctoral research in proteomics at the Max Planck Institute for Molecular Genetics in Berlin. As a protein chemist, she has accumulated extensive experience in mass spectrometry-based quantitative proteomics and PTM analysis for various biological samples and therapeutic protein products. She is currently the head of the R&D department as well as the GLP lab manager at Mass Solutions Technology, part of Mithra Biotechnology.

Sheng Yu Huang received her PhD in Analytical Chemistry from the National Cheng Kung University in 2006. After conducting postdoctoral research at Academia Sinica in Taiwan, she joined and was in charge of the mass spectrometry facility at the Biomedical Engineering Research Laboratories, Industrial Technology Research Institute until 2009. Dr. Chen is currently an Associate Professor in the Department of Chemistry at the National Taiwan Normal University. His expertise includes mass spectrometry-based quantitative proteomics and posttranslational modification analysis.