Disulfide Bond Identification of Biotherapeutic Proteins Using Various Fragmentation Techniques Available on an Orbitrap Fusion Tribrid Mass Spectrometer

Stephane Houel, Jennifer Sutton, Seema Sharma, Terry Zhang, Romain Huguet,
Martin Samonig, Vlad Zabrouskov, Jonathan Josephs
Thermo Fisher Scientific, River Oaks Parkway, San Jose, CA, USA, 95134

Poster Note 64808

INTRODUCTION

Disulfide linkages are directly involved in appropriate protein conformations and therefore have a significant influence on protein functions. As a result, regulatory agencies expect disulfide connectivity to be determined as part of product characterization. Disulfide bond mapping is performed in development and production of monoclonal antibodies and not only expected disulfide linkages but also scrambled disulfide bonds have to be identified. Different strategies can be used to identify disulfide bonds like comparison of LC-MS/MS between a reduced and non-reduced protein digest. The use of different fragmentation techniques is also very powerful, especially electron transfer dissociation which cleaves preferentially at the S-S linkages. Here, we describe how multiple fragmentation techniques can be used to identify complex disulfide bond peptides.

MATERIALS AND METHODS

The sample used in this study is Etanercept, a fusion biotherapeutic protein. Etanercept was first diluted in 0.1M Tris/HCl and PNGaseF and neuraminidase were added to the sample to remove N-linked glycans and sialic acids from the o-core1 glycans. The partial deglycosylation was performed at 37ºC for 3h. Next, the sample was denatured in a mixture of guanidine and Tris-HCl containing 10 mM N-ethylmaleimide (NEM) at pH 5 for 3h at 37ºC. The solution was buffer exchanged with 0.1M Tris-HCl. Trypsinization was performed at 37 ºC overnight. An aliquot of the digest was reduced with DTT to generate the reduced sample.

Liquid Chromatography

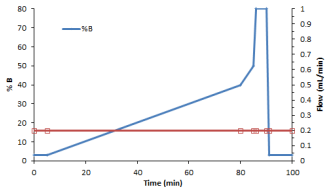
LC: Vanquish UHPLC

Column: Thermo Scientific™ Accucore™, C18 (100mm * 2.1 mm; 1.7 mm)

Mobile phase: 0.1% Formic acid, in H_2O

LC gradient

time	%B	Flow (mL/min)
0	3	0.2
5	3	0.2
80	40	0.2
85	50	0.2
86	80	0.2
90	80	0.2
91	3	0.2
100	3	0.2



Mass Spectrometry

Mass spectrometer: Orbitrap Fusion Tribrid mass spectrometer

Data were acquired in data dependent acquisition (DDA) mode with HCD fragmentation and in targeted mode (inclusion list) for MS2 ETD and MS2 ETD-MS3 HCD (fig. 7).

DDA method: Top speed 2s		ETD targeted method		
MS1 OT res: 120K	MS2 OT res: 30K	MS1 OT res: 120K	MS2 OT res: 30K	MS3 OT or ion trap
AGC MS1: 2E5	AGC MS2: 1E5	AGC MS1: 2E5	AGC MS2: 3E5	AGC MS3: 2.5E5
MS1 Max Inject time: 200 ms	MS2 Max Inject time: 250 ms	MS1 Max Inject time: 200 ms	AGC Reagent : 7E5	HCD CE: 30%
m/z: 300-1500	Intensity threshold 5E4	m/z: 300-1500	reaction time 30 s	

Data Analysis

Raw files were processed with Thermo Scientific™ BioPharma Finder™ Software . The reduced sample was searched with O-core 1 and NEM as variable modifications. For the non-reduced sample, the protease specificity was set to strict. After processing the DDA raw file, some of the identified disulfide bond peptides were selected and targeted for MS2 ETD and MS2 ETD–MS3 HCD. For disulfide bond peptides, only the ETD spectra are shown in the poster.

Figure 1. Data were collected using a Vanquish UHPLC coupled to an Orbitrap Fusion mass spectrometer and processed with Thermo Scientific™ BioPharma Finder™ Software.

RESULTS

A) Peptides in CDR3 domain containing cysteines.

Figure 2. Disulfide bond peptide containing cysteines 104 and 112. a) Targeted ion characteristics, b) annotated MS2 ETD spectra and c) peptide sequence coverage from MS2 ETD.

For disulfide bond peptide, Biopharma Finder will create two different annotated spectra of the MS2 spectrum. For each spectrum, fragment ions specific to each peptide will be annotated. In fig. 2b, ions at m/z 218.1 and m/z 1285.5 represent respectively the z_2 ion of PGWYCALSK and the c_6 ion of PGWYCALSK with QEGCR still linked to it.

Figure 3. Disulfide bond peptide containing cysteines 112 and 115. a) Targeted ion characteristics, b) annotated MS2 ETD spectra and c) peptide sequence coverage from MS2 ETD.

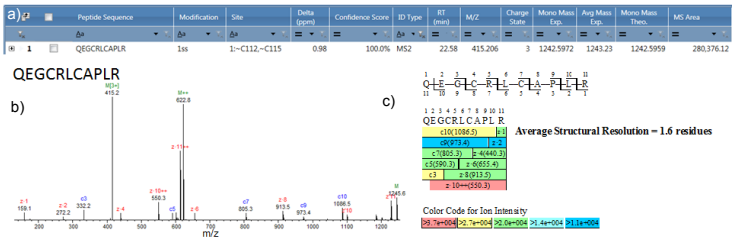
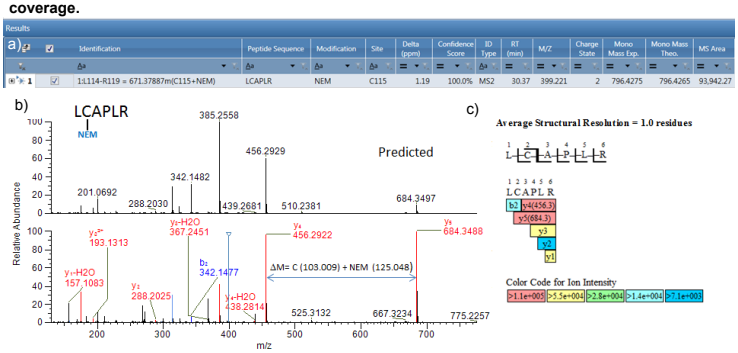


Figure 4. Peptide containing cysteines 115 identified with N-ethylmaleimide (NEM) modification. a) Targeted ion characteristics, b) annotated MS2 HCD spectrum and c) peptide sequence coverage.

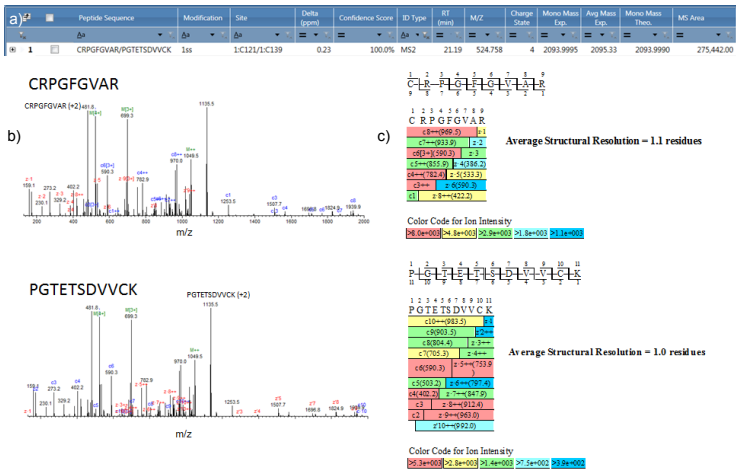


The reduced and non-reduced samples were first analyzed by LC-MS using a data dependent mode of acquisition and HCD fragmentation. After processing the raw files with Biopharma Finder, disulfide bond peptides but also peptides containing cysteine modified with N-ethylmaleimide were identified.

Peptides from the CDR3 domain are shown in figures 2,3,4 and 5. Peptides containing disulfide linkages C104-C112, C112-C115, C121-C139 were targeted in ETD experiments to confirm the initial identification. ETD fragmentation favors disulfide bond cleavage generating ions of the individual peptides but also c and z fragments ions. After processing the ETD MS2 targeted experiment, for most disulfide bond peptides, the average structural resolution is close to 1 confirming without ambiguity the initial identification.

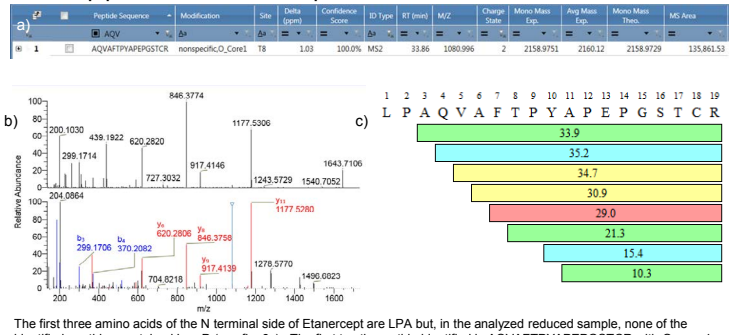
Sample denaturation was done at pH 5 in presence of N-Ethylmaleimide to minimize disulfide bond scrambling therefore the identification of two different disulfide bond peptides containing C112 and the peptide LCAPLR modified with NEM suggests the presence of different isoforrms of Etanercept.

Figure 5. Disulfide bond peptide containing cysteines 121 and 139. a) Targeted ion characteristics, b) annotated MS2 ETD spectra and c) peptide sequence coverage from MS2 ETD.



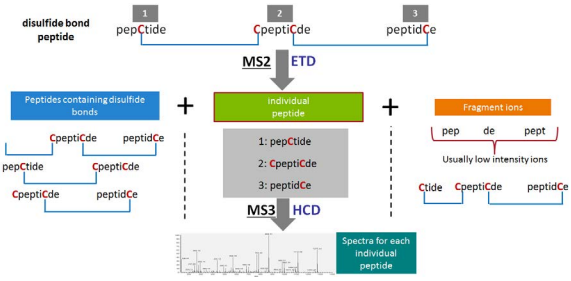
B) N-terminal discrepancy of the analyzed Etanercept sample.

Figure 6. Identification of a non specific peptide AQVAFTPYAPEPGSTCR with O-core1 modification a) Targeted ion characteristics b) experimental and predicted MS2 HCD spectra and c) N-terminal identified peptides from the reduced sample.



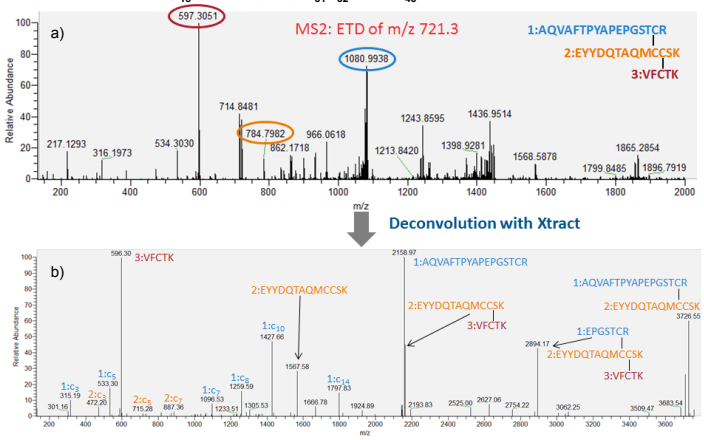
The first three amino acids of the N terminal side of Etanercept are LPA but, in the analyzed reduced sample, none of the identified peptides contained L or P (see fig. 6c). The first tryptic peptide identified is AQVAFTPYAPEPGSTCR with O-core 1 modification (fig. 6a and 6b) and it contains a cysteine at position 18. The modified sequence will then be used to identify disulfide bond linkages in the non-reduced sample. Also, to add an extra level of identification, after the MS2 ETD event of the disulfide bond peptide parent ion, ions representing each individual peptide will be selected for a MS3 HCD event (see fig. 7).

Figure 7. Workflow of the MS2 ETD – MS3 HCD for disulfide bond peptide identification



ETD fragmentation favors disulfide bond cleavages, so for 3 peptides linked though disulfide bonds (fig. 7) the most intense ions in MS2 ETD spectra are usually representing peptides still linked through one disulfide bond and each individual peptide. For complex disulfide bond linkages, few fragments from the backbone of each peptide are observed in MS2 ETD spectra. To gain specificity and more insight of the peptide backbone fragments of each peptide linked through disulfide bonds, an MS3 event on ions representing the individual peptide can be triggered. The fragmentation for the MS3 scan could be ETD, CID or HCD but in our case, for all of MS3 events, we used HCD fragmentation.

Figure 8. (a) MS2 ETD and (b) deconvoluted spectra of disulfide bond peptide AQVAFTPYAPEPGSTC$_{18}$R-EYYDQTAQMC$_{31}$C$_{32}$SK-VFC$_{45}$TK.

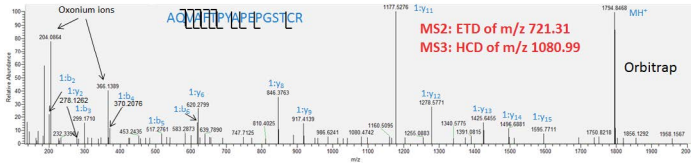


In the deconvoluted spectra (fig. 8b) of the disulfide bond peptide containing cysteines 18,31, 32 and 45, the major peaks represent peptides still containing disulfide bonds or the individual peptides. Fragment ions are at lower intensity, six and three fragments ions can be respectively assigned to AQVAFTPYAPEPGSTC$_{18}$R and EYYDQTAQMC$_{31}$C$_{32}$SK.

Fragments ions 1:c_7 and 1:c_8 are containing the O-core1 modification (+365.1322).

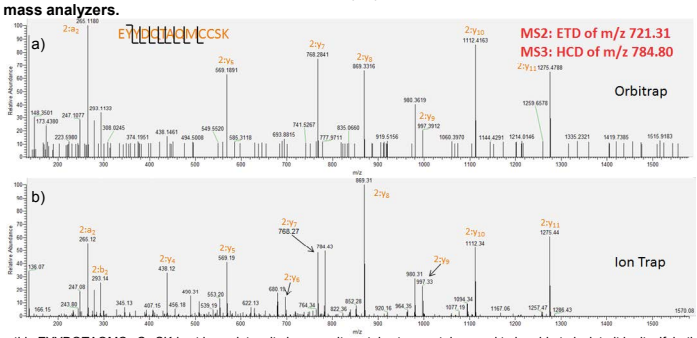
In the MS2 ETD spectrum (Fig. 8a), the ions circled are the ones representing individual peptides and will be isolated for MS3 HCD.

Figure 9. MS3 HCD spectrum of AQVAFTPYAPEPGSTC$_{18}$R.



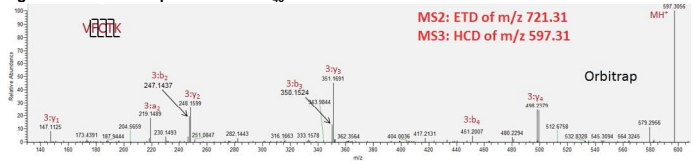
The oxonium ions as well as fourteen b/y ions were identified in the MS3 spectrum confirming the truncated amino acid sequence of the N-terminal peptide.

Figure 10. MS3 HCD spectrum of EYYDQTAQMC$_{31}$C$_{32}$SK using a) the orbitrap or b) the ion trap mass analyzers.



Peptide EYYDQTAQMC$_{31}$C$_{32}$SK is at lower intensity because it contains two cysteines and to be able to isolate it by itself, both disulfide bonds have to be cleaved at the same time. Nevertheless, spectrum containing identifiable fragment ions were recorded in the Orbitrap (fig. 10a) and Ion Trap (fig. 10b) mass analyzers. Even for this low intensity ETD fragment, the MS3 HCD spectrum is rich enough to identify the peptide without ambiguity. In figure 10b, HCD fragmentation was applied but the MS3 spectrum was recorded in the Ion Trap to increase sensitivity. In comparison to the spectrum recorded in the Orbitrap (Fig.10a), three more fragments ions were identified (2:b_2, 2y_4 and 2y_6) in the MS3 HCD spectrum recorded in the Ion trap (Fig 10b)

Figure 11. MS3 HCD spectrum of VFC$_{45}$TK.

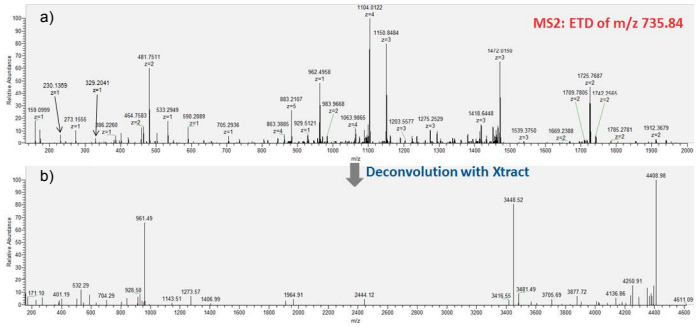


In the MS2 ETD deconvoluted spectra of the of the disulfide bond peptide containing cysteines 18,31, 32 and 45 (fig 8b), none of the fragment ions of peptide VFCTK were identified. In the MS3 HCD spectrum (fig 11), all of the b/y ions that can be recorded (b1 is too small) were identified.

When a disulfide bond peptide is composed by only two peptides, MS2 ETD spectra provides enough fragment ions from both peptides to obtain high peptide sequence coverage (fig. 2 and fig. 5). For more complex disulfide linkages, MS2 ETD followed by MS3 HCD provides high quality MS/MS spectra for each individual peptide.

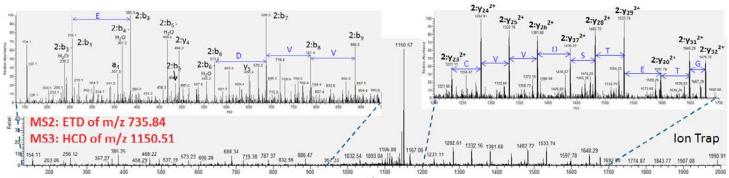
C) Using MS2 ETD – MS3 HCD as a discovery tool for unknown disulfide bond peptide.

Figure 12. (a) MS2 ETD and (b) deconvoluted spectra of targeted m/z 735.84.



The ion at m/z 735.84 eluting around 31 min was unique to the non-reduced sample suggesting a disulfide bond peptide but unfortunately no disulfide bond peptide was identified. The MS2 ETD spectrum (fig. 12a) is rich but it can be difficult to identify patterns for manual interpretation. The deconvoluted spectrum contains three majors fragments (fig. 12b): the parent mass and two masses at 961.49 and 3448.52 suggesting a disulfide bond peptide from only two peptides. The first mass at 961.49 could correspond to CRPGFGVAR (see fig. 5b) and the presence of ions at m/z 159.0999, m/z 230.1369 and m/z 329.2041 (Fig. 12a) seems to confirm it. An MS3 HCD experiment was performed on the doubly charged ion m/z 481.7 (data not shown) and the identification was confirmed. On the other hand, mass 3448.52 does not correspond to any reduced tryptic peptide so this part of the disulfide bond peptide is still unknown.

Figure 13. Spectrum acquired from MS2 ETD of m/z 735.84 ion followed by MS3 HCD of the m/z 1150.51 ion.



The MS3 HCD spectrum recorded in the ion trap contains a major peak at 1150.57 (Fig. 13), the unfragmented parent ion. Nevertheless, in the mass ranges below and above m/z 1150, fragment ions are present. The delta mass between major peaks can be quickly assigned to amino acids (annotated in blue) and the spectrum can be assigned to PGTETSDVVC$_{139}$KPC$_{142}$APGTFSNTTSSTDIC$_{157}$RPHQI with the asparagine chemically converting to aspartic acid due to the deglycosylation and two cysteines still linked through disulfide bond. The amino acid corresponds to a non tryptic peptide explaining why the spectrum was not matched initially. After identification, b and y ions were annotated on the spectrum.

CONCLUSIONS

For the CDR3 domain of Etanercept, peptides containing disulfide linkages C104-C112, C112-C115, C121-C139 were identified with HCD and ETD fragmentation after being processed with BioPharma Finder. The peptide containing C115 was also identified and was modified with NEM suggesting that cysteine 115 can be free. These results reveal the cysteine linkage heterogeneity of Etanercept.

For the analyzed sample of Etanercept, the amino acid sequence was truncated on the N terminal side. The N terminal side tryptic peptide contains a cysteine, and the disulfide bond linkages are C_{18}-C_{31} C_{32}-C_{45}.

MS2 ETD combined with MS3 HCD is a powerful mode of acquisition for unknown and complex disulfide linkages.

PN64808-EN 0716S