

# Disulfide mapping using label-free differential analysis

Michael Athanas<sup>1</sup>, David Sarracino<sup>2</sup>, Scott Peterman<sup>2</sup>, Bryan Krastins<sup>2</sup>, Amol Prakash<sup>2</sup>, Mary F Lopez<sup>2</sup>

<sup>1</sup>VAST Scientific, Cambridge, MA; <sup>2</sup>Thermo Fisher Scientific BRIMS Center, Cambridge, MA;



## Overview

**Purpose:** To discover disulfide bonds in a purified protein sample.

**Methods:** Samples digested with trypsin with and without subsequent reduction and alkylation were compared using LC-MS and label-free analysis software.

**Results:** We demonstrate a workflow for discovery and verification of disulfide bonds using LC-MS and label-free analysis methods. Method robustness appears to be sensitive to sample preparation.

## Introduction

Disulfide cross-linkages formed by the oxidation of the thiol groups of cysteines play an important role in the folding and stability of many proteins. Knowledge of the disulfide linkages in a protein provides insight into protein structure. Determination of the disulfide bonds in recombinant or synthetic proteins is also important, since formation of the correct disulfide cross linkages are an indication of proper folding and function. Experimentally, protein structure can be determined with time consuming methods such as X-ray crystallography or NMR spectroscopy.

In this presentation, we discuss a method of determining disulfide bonds using LC-MS and label-free differential analysis (Thermo Scientific SIEVE V1.3). The method is based upon accurate mass derived from full scan data – no MS/MS fragmentation data were used.

When using accurate mass information for identification assignment, mass degeneracy is often a concern; that is, molecules with the same or similar mass may introduce ambiguity. Because of the mass accuracy of the instrument (Thermo Scientific Exactive) and the simple samples (purified proteins), mass degeneracy is never a problem.

The full scan accurate mass approach is preferred because the data are easier to interpret and are independent from the identification efficiency of fragmentation interpretation software such as Mascot or SEQUEST.

We present a work-in-progress, whereas, the samples we have chosen are pedagogical, but well studied and understood. Our approach in developing this method is to start with simple proteins that may be validated with previously reported results.

## Samples

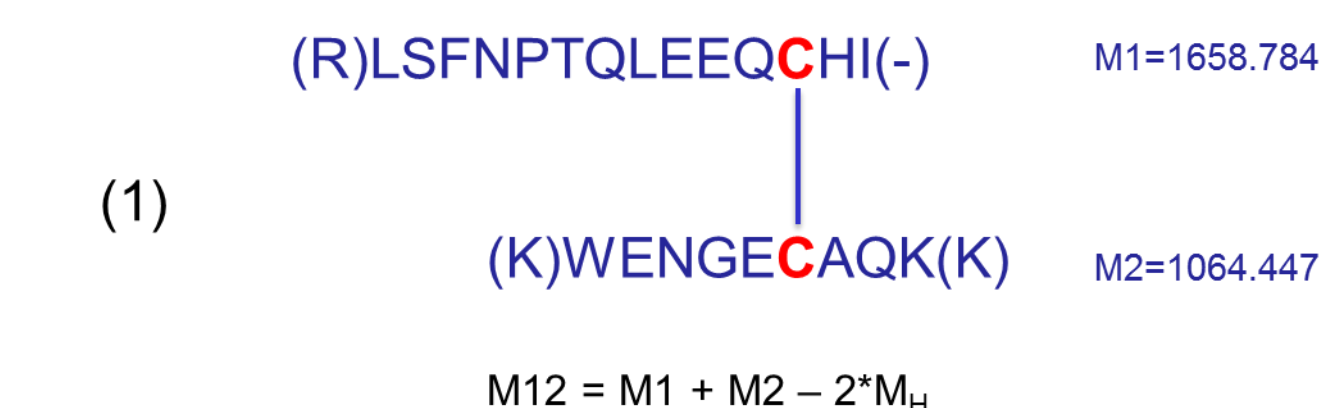
In this presentation analysis methods are emphasized. We focus upon a simple protein sample of beta lactoglobulin of length 178 and a mass of ca 20kDa. Beta lactoglobulin contains six cysteines and three disulfide bonds (Figure 1)

Figure 1

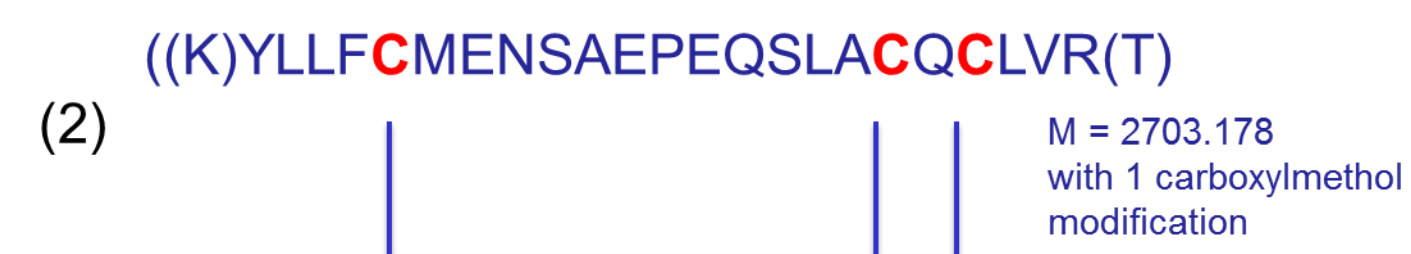
Disulfide bond	82 ↔ 176	Alternate (Ref.13) (Ref.19)	— + —
Disulfide bond	122 ↔ 137	Alternate (Ref.13)	— + +
Disulfide bond	122 ↔ 135	(Ref.13) (Ref.19)	— + +

## Sample Preparation and Acquisition

Purified protein samples were weighed and dissolved in a buffer of 8M Urea 150mM Tris-HCl 2.5% n-propanol, pH 8.5 to denature the proteins.



Two equal fractions were created. The first fraction was reduced with DTT to break disulfide bonds and create free cysteine groups. The sample was then alkylated with iodoacetic acid to modify all cysteines in a way that prevented disulfide bonds from forming. Trypsin was then added to the first fraction for protein digestion into peptides. The second fraction was treated with trypsin only (no digestion alkylation).



The fractions were then divided into four equal aliquots and placed in a 96 well plate for LC-MS acquisition.

Samples were injected at 150uL/min onto a Hypersil gold 5micron 1mmx150mm column using a Thermo Scientific Surveyor Autosampler with a pump. Solvents were Fisher Optima LC-MS grade Water with 0.2% formic acid buffer A, acetonitrile 0.2% formic acid buffer B. A gradient from 3% B to 30% B was used to separate the polypeptides. The source was an Ion Max with a standard ESI probe fitted with a 32 gauge metal needle.

The mass spectrometer was tuned on Angiotensin I, T'ed into a 150uL min flow of 15% B. Tune parameters were 20 on sheath, 3 on aux, with a spray voltage of 3.8kV and a capillary temp of 300degC. Scan parameters, 60K resolution with 3 microscans and lock mass were enabled. The divert valve was used during the first 5 min to prevent salts and urea in the digests from contaminating the source.

## Sample Preparation and Acquisition

The acquired data were processed using SIEVE V1.3 in an "A vs B" differential analysis experiment with technical replicates. The SIEVE workflow first performed a chromatographic alignment using full scan spectral shape (no peaks). Frames (a.k.a. features) were constructed based upon prominent peaks throughout the full data set. Identification assignment was based upon the accurate mass database described above using the SIEVE MZLookup tool. Replicates were required to have a CV<20%.

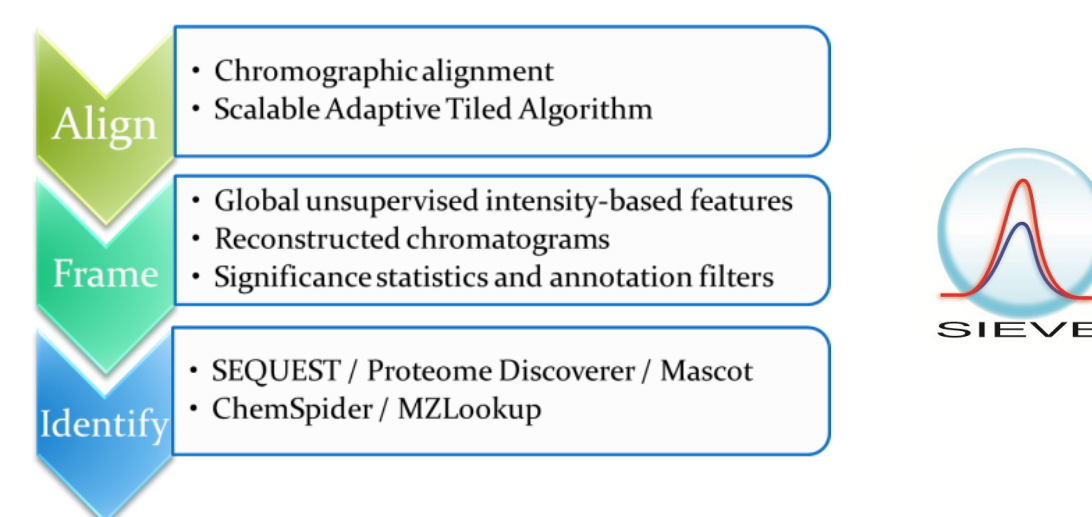


Figure 1. A disulfide folded peptide displayed in sequence (2) is shown in this view of SIEVE where the red trace is from the digested fraction.

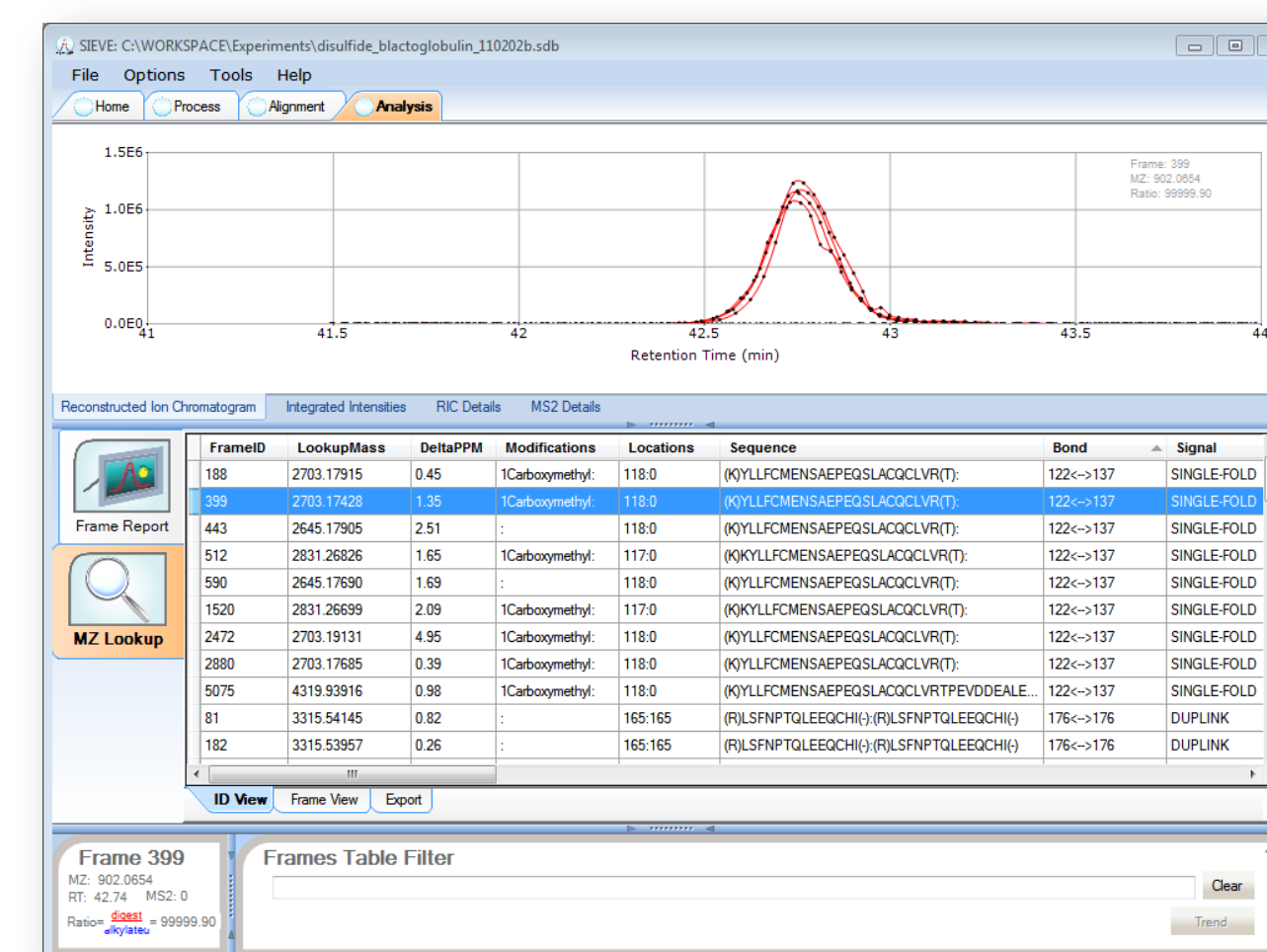
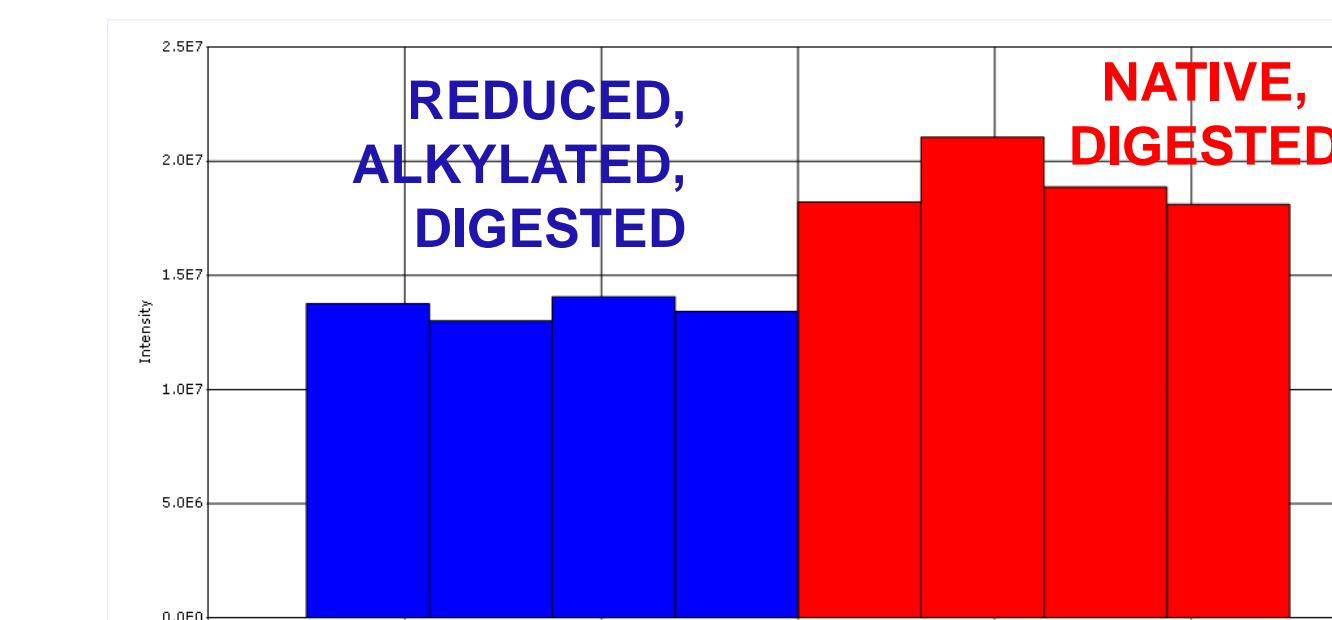
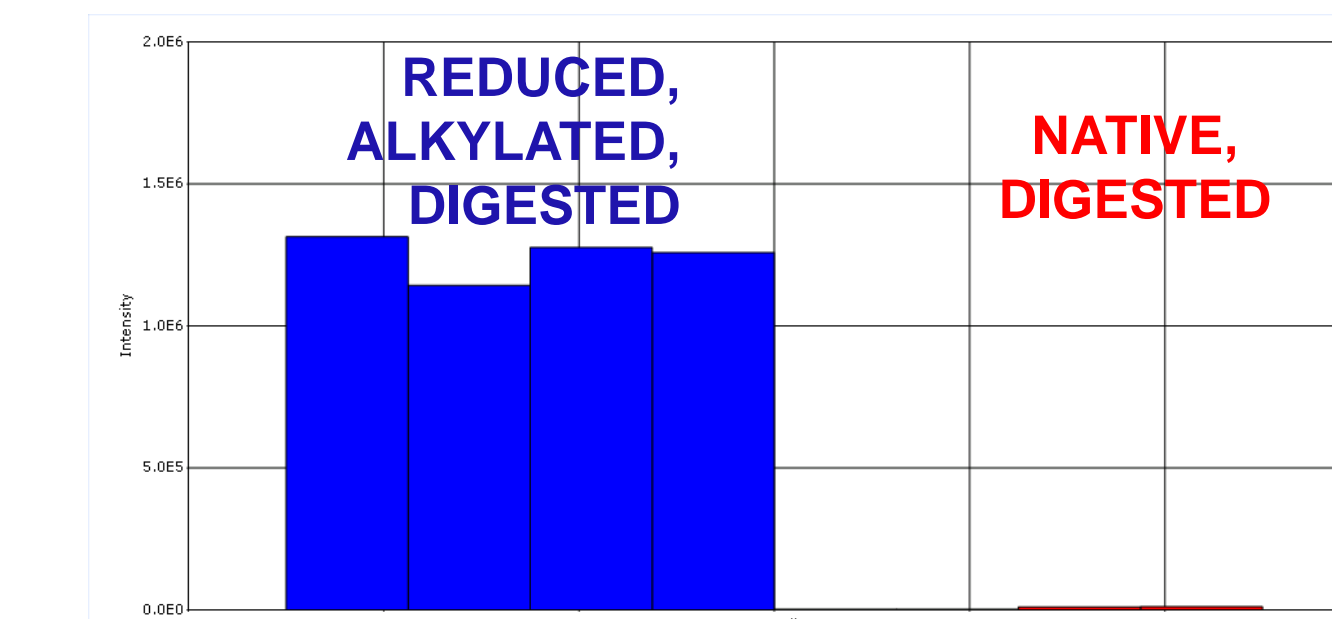


FIGURE 2. SIEVE results demonstrating characteristic intensity signal shapes of four technical replicate injections for A) Non-cysteine containing peptides, B) Carboxymethyl-modified cysteine containing peptides, and C) Unmodified peptides.

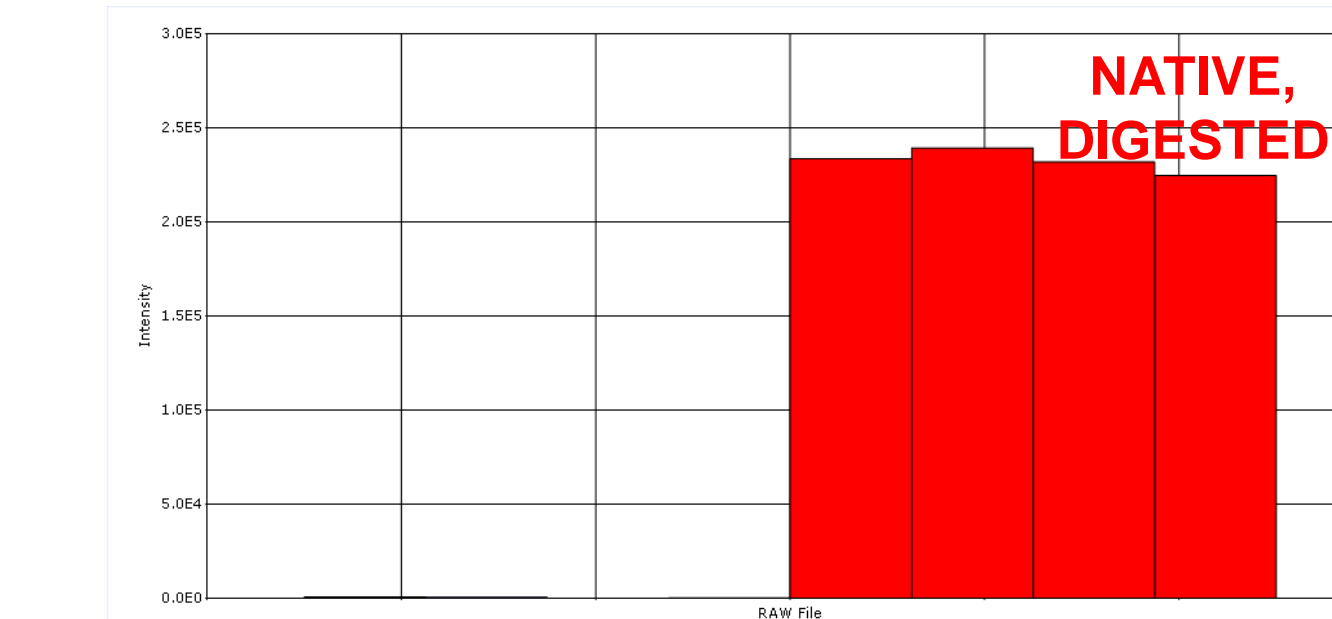
**A. Non-cysteine containing peptides** – Detection levels of non-cysteine containing peptides were expected to be similar. The peptide sequence shown here is (R)TPEVDDEALEK(F) and is located at position 139.



**B. Reduced cysteine containing peptide** – Peptides containing a reduced and alkylated cysteine will generally not appear in the sample digested under native (=non-reducing) conditions. The peptide shown here is (K)YLLFCMENSAPPEQSLACQCLVR(T), located at position 117.

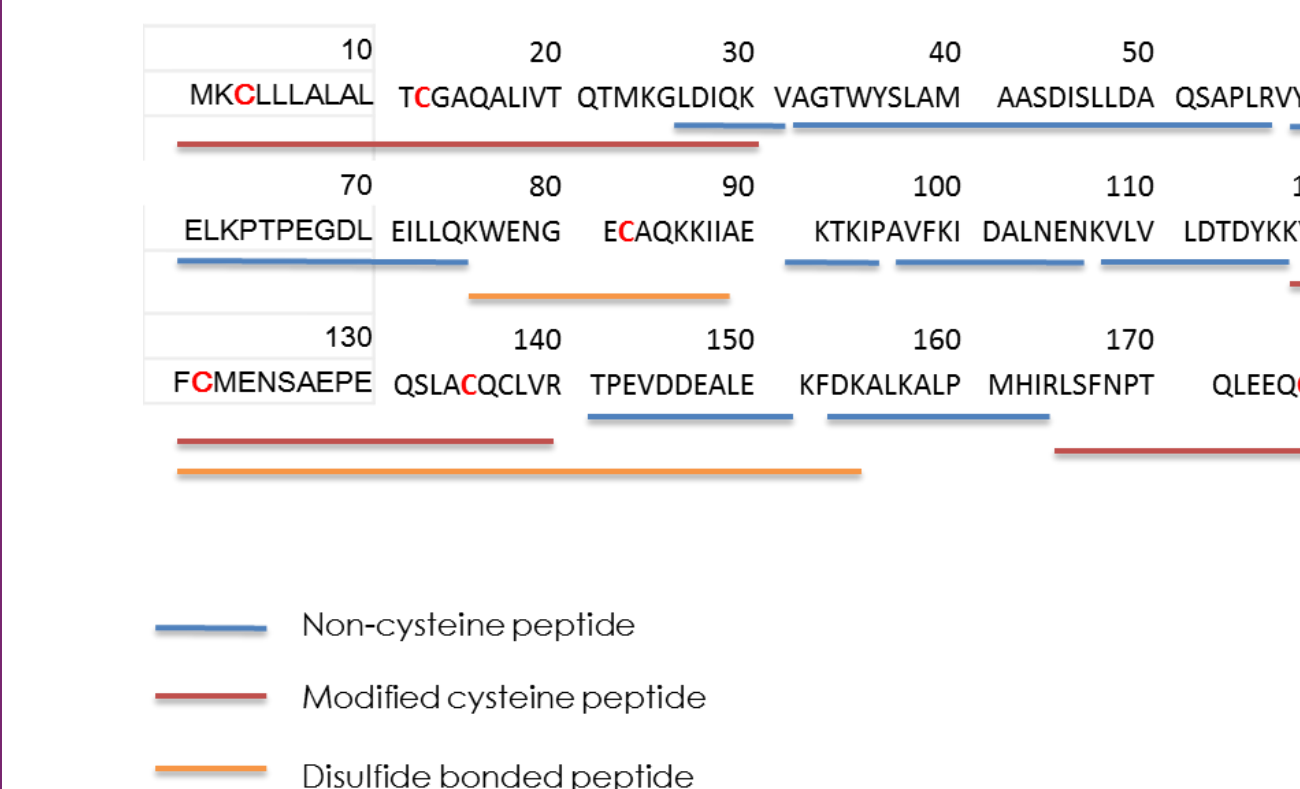


**C. Native cysteine containing peptides** – a dimer of the peptide (R)LSNPTQLEEQCHI(-) located at position 165 was identified and can be attributed to a side reaction during sample preparation.

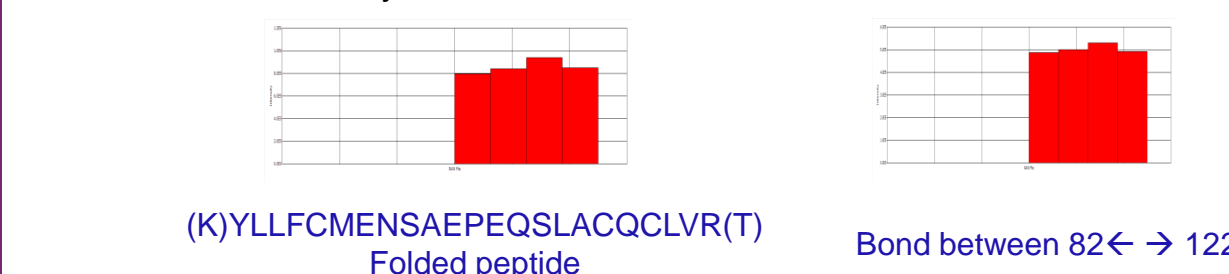


## Results

Figure 2. The protein coverage for alkylated cysteine containing peptides (red), non-cysteine containing peptides (blue), and disulfide bonded peptides (yellow) for beta lactoglobulin are shown below.



As is evident from the data, the method provided complete sequence coverage. The bond between cysteines 82 ↔ 122 (not previously reported in the literature) was confirmed. The bonds between 122 ↔ 137, 122 ↔ 135 and 135 ↔ 137 were generally confirmed but cannot be distinguished since this peptide was not cleaved between cysteines. The bond between 176 ↔ 82 was also confirmed.



## Conclusion

We have demonstrated a workflow for the discovery and verification of disulfide bonds using label-free LC-MS methods. The method worked well for peptides that were well separated by an enzymatic cleavage point; however, it was not possible to distinguish intra-molecular bonds in peptides bonds in peptides with three or more cysteines.

One caveat of the method was the observation of prominent signals for peptide dimers. This may be an indication of insufficient sample preparation chemistry. Although the samples were treated with additional DTT post reduction/alkylation, (in order to quench the alkylation reaction and return the sampled to a reduced state), free sulfhydryl containing cysteine peptides apparently remained.