

METHODS FOR IDENTIFYING FREE THIOLS IN PROTEINS

TECHNICAL FIELD OF THE INVENTION

The invention is generally related to systems and methods of characterizing free thiol
5 group in proteins, particularly in antibodies.

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims benefit of and priority to US Provisional Patent Application No.
62/792,994 filed January 16, 2019, incorporated herein by reference in its entirety.

BACKGROUND OF THE INVENTION

10 During the development of monoclonal antibodies (mAbs) from drug candidate to
marketed product, issues with stability, post-translational modifications, or other changes to the
antibody can occur. Alterations in antibody structure and function can cause problems such as
poor shelf-life or even immunogenicity in the patient. It is therefore important to properly
characterize antibody structure and monitor it throughout production. Antibody quality control
15 and quality assurance are critical to the purity and safety of mAb products.

Disulfide bonds are important for structural integrity, stability, and biological functions
of mAbs. Non-native disulfide bonds can cause changes in the structure and stability of mAbs.
Binding affinity of mAbs to antigens can be affected by up to 50% if disulfide bonds are
incomplete (Xiang, T., et al., *Anal Chem*, 81:8101-8108 (2009)). The low dissociation energy of
20 disulfide bonds and the high flexibility of the hinge region frequently lead to modifications and
cleavages at the hinge region (Moritz, B., and Stracke, J.O., *Electrophoresis*, 36:769-785
(2017)). In addition, administration of non-native disulfide bonded structures to humans has the
potential to trigger unwanted immune responses. Analysis of disulfide bonds is therefore
important for quality control assessment of mAbs. Current methods of analyzing mAb disulfide
25 bonds are time-consuming and labor intensive.

Therefore, it is an object of the invention to provide systems and methods for
characterizing antibodies, in particular disulfide bonds in monoclonal antibodies.

Another embodiment provides a method for identifying disulfide heterogeneity in a
protein drug product.

30 It is another object of the invention to provide methods and compositions for identifying
free thiols in proteins including but not limited to antibodies.

SUMMARY OF THE INVENTION

Compositions and methods for identifying free thiols (also referred to as free sulfhydryls) are provided. Free thiols can exist in proteins, for example mAbs, as a result of incomplete disulfide bond formation or disulfide bond breakage. The presence of free thiols can lower the thermal stability of mAb products and affect their structural integrity, stability, and biological functions. It is therefore important to characterize free thiols in mAb production. Methods for identifying site specific free thiols in protein drug products are disclosed herein. An exemplary method includes labeling peptides with a tag to identify free thiols and a tag to identify native disulfide bonds and analyzing the tags using targeted MS². In one embodiment, the method provides complete coverage of all 32 cysteine residues in an IgG molecule. In other embodiments the method covers the 16 cysteine residues on the heavy and light chains in an IgG molecule. In another embodiment, the method covers the 5 cysteine residues on each light chain of an IgG molecule. In another embodiment, the method covers the 11 cysteine residues on each heavy chain of an IgG molecule.

One embodiment provides a method for identifying the presence of free thiols in a protein, for example an antibody or fragment thereof. An exemplary method for identifying free thiols in a protein drug product includes the steps of labeling a sample containing the protein drug product with a first label containing a sulfhydryl-reactive iodoacetyl group, an MS-neutral spacer arm, and a first MS/MS reporter having a unique reporter ion mass. Excess label can be removed and the sample is denatured and reduced. The sample is then labeled with a second label having a sulfhydryl-reactive iodoacetyl group, an MS-neutral spacer arm, and a second MS/MS reporter having a unique reporter ion mass. The method includes enzymatically digesting the sample and analyzing the sample using mass spectrometry, for example ultra-performance liquid chromatography tandem mass spectrometry system (UPLC-MS² system) including a charged surface hybrid column and a formic acid buffer mobile phase. Next, the method includes quantifying the first MS/MS reporter and the second MS/MS reporter, wherein the amount of the first MS/MS reporter correlates to the amount of free thiols in the protein drug product and the amount of the second MS/MS reporter correlates to the amount of bound thiols in the protein drug product. In one embodiment the first MS/MS reporter has a mass of 128 and the second MS/MS reporter has a mass of 131. The protein drug product is typically an antibody, for example a monoclonal or chimeric antibody. The presence of free thiols is likely a

result of incomplete disulfide bond formation or disulfide bond degradation. In one embodiment, the method provides complete coverage of all 32 cysteine residues in an IgG molecule. In other embodiments the method covers the 16 cysteine residues on the heavy and light chains in an IgG molecule. In another embodiment, the method covers the 5 cysteine residues on each light chain of an IgG molecule. In another embodiment, the method covers the 11 cysteine residues on each heavy chain of an IgG molecule.

Another embodiment provides a method of identifying disulfide heterogeneity in a protein drug product including the steps of labeling a sample containing the protein drug product with a first label containing a sulfhydryl-reactive iodoacetyl group, an MS-neutral spacer arm, and a first MS/MS reporter having a unique reporter ion mass. Excess label can be removed and the sample is denatured and reduced. The sample is then labeled with a second label having a sulfhydryl-reactive iodoacetyl group, an MS-neutral spacer arm, and a second MS/MS reporter having a unique reporter ion mass. The method includes enzymatically digesting the sample and analyzing the sample using mass spectrometry, for example an ultra-performance liquid chromatography tandem mass spectrometry system (UPLC-MS² system) including a charged surface hybrid column and a formic acid buffer mobile phase. Next, the method includes quantifying the first MS/MS reporter and the second MS/MS reporter, wherein the amount of the first MS/MS reporter correlates to the amount of free thiols in the protein drug product and the amount of the second MS/MS reporter correlates to the amount of bound thiols in the protein drug product. If the assay detects free thiols in the protein drug product, the protein drug product contains disulfide heterogeneity. The presence of these free thiols is likely a result of incomplete disulfide bond formation or disulfide bond degradation. In one embodiment the first MS/MS reporter has a mass of 128 and the second MS/MS reporter has a mass of 131. The protein drug product is typically an antibody, for example a monoclonal, chimeric antibody, bispecific antibody, or antigen binding fragments thereof. In one embodiment, the method provides complete coverage of all 32 cysteine residues in an IgG molecule. In other embodiments the method covers the 16 cysteine residues on the heavy and light chains in an IgG molecule. In another embodiment, the method covers the 5 cysteine residues on each light chain of an IgG molecule. In another embodiment, the method covers the 11 cysteine residues on each heavy chain of an IgG molecule.

A method for selecting a protein drug product including the steps of labeling a sample containing the protein drug product with a first label containing a sulfhydryl-reactive iodoacetyl group, an MS-neutral spacer arm, and a first MS/MS reporter having a unique reporter ion mass. Excess label can be removed and the sample is denatured and reduced. The sample is then
5 labeled with a second label having a sulfhydryl-reactive iodoacetyl group, an MS-neutral spacer arm, and a second MS/MS reporter having a unique reporter ion mass. The method includes enzymatically digesting the sample and analyzing the sample using an ultra-performance liquid chromatography tandem mass spectrometry system (UPLC-MS² system) including a charged surface hybrid column and a formic acid buffer mobile phase. Next, the method includes
10 quantifying the first MS/MS reporter and the second MS/MS reporter, wherein the amount of the first MS/MS reporter correlates to the amount of free thiols in the protein drug product and the amount of the second MS/MS reporter correlates to the amount of bound thiols in the protein drug product. If a free thiol is detected, the protein drug product has disulfide heterogeneity. The method includes selecting the protein drug product that does not display disulfide
15 heterogeneity. In one embodiment the first MS/MS reporter has a mass of 128 and the second MS/MS reporter has a mass of 131. The protein drug product is typically an antibody, for example a monoclonal, chimeric antibody, or a bispecific antibody, or antigen binding fragments thereof. Another embodiment provides a pharmaceutical composition containing the protein drug product selected using the method described above. In one embodiment, the method
20 provides complete coverage of all 32 cysteine residues in an IgG molecule. In other embodiments the method covers the 16 cysteine residues on the heavy and light chains in an IgG molecule. In another embodiment, the method covers the 5 cysteine residues on each light chain of an IgG molecule. In another embodiment, the method covers the 11 cysteine residues on each heavy chain of an IgG molecule.

25 BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1A is a schematic illustration of four IodoTMTsixplex reagents. Figure 1B is a schematic illustration of an exemplary workflow for labeling peptides with iodoTMT labels.

Figure 2 is a schematic illustration of an exemplary workflow for analyzing iodoTMT labeled peptides using targeted MS².

Figures 3A-3B are chromatogram results from MS² analysis of iodoTMT labeled peptides run with data dependent run (Fig. 3A) or a targeted MS² with inclusion list (Fig. 3B). The X-axis represents time and the Y-axis represents relative abundance.

Figures 4A-4B are chromatogram results for MS² analysis of iodoTMT labeled peptides, specifically looking at TMT-tags 128/131. Figure 4A is the complete chromatogram and Figure 4B is a zoomed in on the TMT-tags 128/131. The X-axis represents time and the Y-axis represents relative abundance.

Figures 5A-5G are chromatogram results from MS² analysis of iodoTMT labeled peptides from different lots of an antibody. The X-axis represents time and the Y-axis represents relative abundance.

Figure 6 is a bar graph showing relative abundance of various cysteine residues from a spike-in study in which 10%, 5%, 1%, 0.5%, or 0.1% of protein sample 1 was spiked into protein sample 2.

Figure 7 is a bar graph showing relative abundance of various cysteine disulfide bond pairs in antibody Lots 1-3 of production process A and antibody Lots 4-6 of production process B. Disulfide bond pairs are indicated by curved lines connecting the two cysteines.

DETAILED DESCRIPTION OF THE INVENTION

I. Definitions

It should be appreciated that this disclosure is not limited to the compositions and methods described herein as well as the experimental conditions described, as such may vary. It is also to be understood that the terminology used herein is for the purpose of describing certain embodiments only, and is not intended to be limiting, since the scope of the present disclosure will be limited only by the appended claims.

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this disclosure belongs. Although any compositions, methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention. All publications mentioned are incorporated herein by reference in their entirety.

The use of the terms “a,” “an,” “the,” and similar referents in the context of describing the presently claimed invention (especially in the context of the claims) are to be construed to cover

both the singular and the plural, unless otherwise indicated herein or clearly contradicted by context.

Recitation of ranges of values herein are merely intended to serve as a shorthand method of referring individually to each separate value falling within the range, unless otherwise indicated herein, and each separate value is incorporated into the specification as if it were individually recited herein.

Use of the term “about” is intended to describe values either above or below the stated value in a range of approx. +/- 10%; in other embodiments the values may range in value either above or below the stated value in a range of approx. +/- 5%; in other embodiments the values may range in value either above or below the stated value in a range of approx. +/- 2%; in other embodiments the values may range in value either above or below the stated value in a range of approx. +/- 1%. The preceding ranges are intended to be made clear by context, and no further limitation is implied. All methods described herein can be performed in any suitable order unless otherwise indicated herein or otherwise clearly contradicted by context. The use of any and all examples, or exemplary language (e.g., “such as”) provided herein, is intended merely to better illuminate the invention and does not pose a limitation on the scope of the invention unless otherwise claimed. No language in the specification should be construed as indicating any non-claimed element as essential to the practice of the invention.

“Protein” refers to a molecule comprising two or more amino acid residues joined to each other by a peptide bond. Protein includes polypeptides and peptides and may also include modifications such as glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues, alkylation, hydroxylation and ADP-ribosylation. Proteins can be of scientific or commercial interest, including protein-based drugs, and proteins include, among other things, enzymes, ligands, receptors, antibodies and chimeric or fusion proteins. Proteins are produced by various types of recombinant cells using well-known cell culture methods, and are generally introduced into the cell by genetic engineering techniques (e.g., such as a sequence encoding a chimeric protein, or a codon-optimized sequence, an intronless sequence, etc.) where it may reside as an episome or be integrated into the genome of the cell.

“Antibody” refers to an immunoglobulin molecule consisting of four polypeptide chains, two heavy (H) chains and two light (L) chains inter-connected by disulfide bonds. Each heavy chain has a heavy chain variable region (HCVR or VH) and a heavy chain constant region. The

heavy chain constant region contains three domains, CH1, CH2 and CH3. Each light chain has a light chain variable region and a light chain constant region. The light chain constant region consists of one domain (CL). The VH and VL regions can be further subdivided into regions of hypervariability, termed complementarity determining regions (CDR), interspersed with regions that are more conserved, termed framework regions (FR). Each VH and VL is composed of three CDRs and four FRs, arranged from amino-terminus to carboxy-terminus in the following order: FR1, CDR1, FR2, CDR2, FR3, CDR3, FR4. The term "antibody" includes reference to both glycosylated and non-glycosylated immunoglobulins of any isotype or subclass. The term "antibody" includes antibody molecules prepared, expressed, created or isolated by recombinant means, such as antibodies isolated from a host cell transfected to express the antibody. The term antibody also includes bispecific antibody, which includes a heterotetrameric immunoglobulin that can bind to more than one different epitope. Bispecific antibodies are generally described in US Patent No. 8,586,713, which is incorporated by reference into this application.

"Hinge region" refers to the flexible amino acid stretch in the central part of the heavy chains of the IgG and IgA immunoglobulin classes, which links these 2 chains by disulfide bonds. In IgG immunoglobulins the hinge region is located between the CH1 and CH3 constant domains. The hinge region affords flexibility to the antibody, and allows easier binding to the antigen.

"Fc fusion proteins" comprise part or all of two or more proteins, one of which is an Fc portion of an immunoglobulin molecule, which are not otherwise found together in nature. Preparation of fusion proteins comprising certain heterologous polypeptides fused to various portions of antibody-derived polypeptides (including the Fc domain) has been described, e.g., by Rath, T., et al., *Crit Rev Biotech*, 35(2): 235-254 (2015), Levin, D., et al., *Trends Biotechnol*, 33(1): 27-34 (2015)) "Receptor Fc fusion proteins" comprise one or more extracellular domain(s) of a receptor coupled to an Fc moiety, which in some embodiments comprises a hinge region followed by a CH2 and CH3 domain of an immunoglobulin. In some embodiments, the Fc-fusion protein comprises two or more distinct receptor chains that bind to a one or more ligand(s). For example, an Fc-fusion protein is a trap, such as for example an IL-1 trap or VEGF trap.

The term "disulfide bond" refers to the linkage formed by the oxidation of two SH groups, each attached to a cysteine. Disulfide bonds play an important role in the folding and

stability of many proteins. IgGs include two heavy chains (HC) and two light chains (LC) covalently linked by a total of 16 inter- or intra-molecular disulfide bonds. IgG mAbs contain 32 cysteine residues, 5 cysteine residues on each LC and 11 cysteine residues on each HC. Each LC contains one variable domain and one constant domain with a disulfide bond connection. The 5th cysteine on the LC is linked to either the 3rd or 5th cysteine of the HC to form an interchain disulfide bond. The heavy chains include an N-terminal variable domain (VH) and three constant domains (CH1, CH2, and CH3) with a hinge region between CH1 and CH2 (Vidarsson, G., et al., *Front Immunol*, 5:520 (2014)). The 6th and 7th cysteine on each HC are bonded forming the hinge region. The hinge region of an immunoglobulin helps form the Y-shaped structure of the immunoglobulin molecule. The Y shape makes possible the flexibility of the immunoglobulin molecules required in antigen binding.

The term “LC-MS” refers to liquid chromatography–mass spectrometry which is an analytical chemistry technique that combines the physical separation capabilities of liquid chromatography (or HPLC) with the mass analysis capabilities of mass spectrometry (MS). The term MS/MS or MS² refers to tandem mass spectrometry.

The term “free thiols” and “free sulfhydryls” are used interchangeably.

II. Methods of Identifying Free Thiols

Methods for identifying free thiols in proteins, including but not limited to antibodies are provided. A thiol or sulfhydryl generally refers to an organic compound containing the group —SH. In proteins any two cysteines in close proximity will form a covalent bond, even cysteines that do not naturally pair together. This covalent bond between two cysteines is referred to as a disulfide bond. Disulfide bonds are critical for IgG tertiary structure, stability, and biological function. A free thiol or free sulfhydryl refers to a cysteine in a protein that is not part of a disulfide bond and can be indicative of improper structure formation in the protein which could result in adverse effects on protein drug potency, half-life, stability, or render the protein drug ineffective. Free sulfhydryls (also referred to as free thiols) can also occur as a result of incomplete disulfide bond formation or disulfide bond degradation. Increased free sulfhydryls can result in lower thermal stability and can affect binding affinity of antibodies to antigens by up to 50%. Disclosed herein are methods for identifying free sulfhydryls or free thiols.

A. Methods for Identifying Site Specific Free Thiols

Free thiols can exist in mAbs as a result of incomplete disulfide bond formation or disulfide bond breakage. The presence of free thiols can lower the thermal stability of mAb products and affect their structural integrity, stability, and biological functions. It is therefore important to characterize free thiols in mAb production. Methods for identifying site specific free thiols in protein drug products are disclosed herein. An exemplary method includes labeling peptides with a tag to identify free thiols and a tag to identify native disulfide bonds and analyzing the tags using targeted MS². In one embodiment, the method provides complete coverage of all 32 cysteine residues in an IgG molecule. In one embodiment, the method provides complete coverage of all 32 cysteine residues in an IgG molecule. In other embodiments the method covers the 16 cysteine residues on the heavy and light chains in an IgG molecule. In another embodiment, the method covers the 5 cysteine residues on each light chain of an IgG molecule. In another embodiment, the method covers the 11 cysteine residues on each heavy chain of an IgG molecule.

One embodiment provides a method for identifying the presence of free thiols in a protein, for example an antibody or fragment thereof. An exemplary method for identifying free thiols in a protein drug product includes the steps of labeling a sample containing the protein drug product with a first label containing a sulfhydryl-reactive iodoacetyl group, an MS-neutral spacer arm, and a first MS/MS reporter having a unique reporter ion mass. Excess label can be removed and the sample is denatured and reduced. The sample is then labeled with a second label having a sulfhydryl-reactive iodoacetyl group, an MS-neutral spacer arm, and a second MS/MS reporter having a unique reporter ion mass. The method includes enzymatically digesting the sample and analyzing the sample using mass spectrometry, for example ultra-performance liquid chromatography tandem mass spectrometry system (UPLC-MS² system) including a charged surface hybrid column and a formic acid buffer mobile phase. Next, the method includes quantifying the first MS/MS reporter and the second MS/MS reporter, wherein the amount of the first MS/MS reporter correlates to the amount of free thiols in the protein drug product and the amount of the second MS/MS reporter correlates to the amount of bound thiols in the protein drug product. In one embodiment the first MS/MS reporter has a mass of 128 and the second MS/MS reporter has a mass of 131. The protein drug product is typically an antibody, for example a monoclonal or chimeric antibody. The presence of free thiols is likely a

result of incomplete disulfide bond formation or disulfide bond degradation. In one embodiment, the method provides complete coverage of all 32 cysteine residues in an IgG molecule. In other embodiments the method covers the 16 cysteine residues on the heavy and light chains in an IgG molecule. In another embodiment, the method covers the 5 cysteine residues on each light chain of an IgG molecule. In another embodiment, the method covers the 11 cysteine residues on each heavy chain of an IgG molecule.

Another embodiment provides a method of identifying disulfide heterogeneity in a protein drug product including the steps of labeling a sample containing the protein drug product with a first label containing a sulfhydryl-reactive iodoacetyl group, an MS-neutral spacer arm, and a first MS/MS reporter having a unique reporter ion mass. Excess label can be removed and the sample is denatured and reduced. The sample is then labeled with a second label having a sulfhydryl-reactive iodoacetyl group, an MS-neutral spacer arm, and a second MS/MS reporter having a unique reporter ion mass. The method includes enzymatically digesting the sample and analyzing the sample using mass spectrometry, for example an ultra-performance liquid chromatography tandem mass spectrometry system (UPLC-MS² system) including a charged surface hybrid column and a formic acid buffer mobile phase. Next, the method includes quantifying the first MS/MS reporter and the second MS/MS reporter, wherein the amount of the first MS/MS reporter correlates to the amount of free thiols in the protein drug product and the amount of the second MS/MS reporter correlates to the amount of bound thiols in the protein drug product. If the assay detects free thiols in the protein drug product, the protein drug product contains disulfide heterogeneity. The presence of these free thiols is likely a result of incomplete disulfide bond formation or disulfide bond degradation. In one embodiment the first MS/MS reporter has a mass of 128 and the second MS/MS reporter has a mass of 131. The protein drug product is typically an antibody, for example a monoclonal, chimeric antibody, bispecific antibody, or antigen binding fragments thereof. In one embodiment, the method provides complete coverage of all 32 cysteine residues in an IgG molecule. In other embodiments the method covers the 16 cysteine residues on the heavy and light chains in an IgG molecule. In another embodiment, the method covers the 5 cysteine residues on each light chain of an IgG molecule. In another embodiment, the method covers the 11 cysteine residues on each heavy chain of an IgG molecule.

A method for selecting a protein drug product including the steps of labeling a sample containing the protein drug product with a first label containing a sulfhydryl-reactive iodoacetyl group, an MS-neutral spacer arm, and a first MS/MS reporter having a unique reporter ion mass. Excess label can be removed and the sample is denatured and reduced. The sample is then

5 labeled with a second label having a sulfhydryl-reactive iodoacetyl group, an MS-neutral spacer arm, and a second MS/MS reporter having a unique reporter ion mass. The method includes enzymatically digesting the sample and analyzing the sample using an ultra-performance liquid chromatography tandem mass spectrometry system (UPLC-MS² system) including a charged surface hybrid column and a formic acid buffer mobile phase. Next, the method includes

10 quantifying the first MS/MS reporter and the second MS/MS reporter, wherein the amount of the first MS/MS reporter correlates to the amount of free thiols in the protein drug product and the amount of the second MS/MS reporter correlates to the amount of bound thiols in the protein drug product. If a free thiol is detected, the protein drug products has disulfide heterogeneity. The method includes selecting the protein drug product that does not display disulfide

15 heterogeneity. In another embodiment the selected protein drug product is an IgG that contains less than 5 free thiols, less than 4 free thiols, less than 3 free thiols, less than 2 free thiols. In some embodiments the free thiol is in a light chain of the IgG molecule. In some embodiments, the free thiol is in a heavy chain of the IgG molecule. In some embodiments, there is at least one free thiol in a light chain and at least one free thiol in a heavy chain of the IgG molecule. In one

20 embodiment the first MS/MS reporter has a mass of 128 and the second MS/MS reporter has a mass of 131. The protein drug product is typically an antibody, for example a monoclonal, chimeric antibody, or a bispecific antibody, or antigen binding fragments thereof. Another embodiment provides a pharmaceutical composition containing the protein drug product selected using the method described above. In one embodiment, the method provides complete coverage

25 of all 32 cysteine residues in an IgG molecule. In other embodiments the method covers the 16 cysteine residues on the heavy and light chains in an IgG molecule. In another embodiment, the method covers the 5 cysteine residues on each light chain of an IgG molecule. In another embodiment, the method covers the 11 cysteine residues on each heavy chain of an IgG molecule.

1. Iodoacetyl Tags

In one embodiment, peptides can be labeled with tags to identify native disulfide bonds and free thiol residues. The tag can be an iodoacetyl tag that labels cysteine-containing peptides. Exemplary tags include but are not limited to Thermo Scientific™ Iodoacetyl Tandem Mass Tag™ (iodoTMT™). Iodoacetyl tags irreversibly label free sulfhydryl groups on cysteine residues. The tags include a sulfhydryl-reactive iodoacetyl group, a mass normalizer arm, and a MS/MS reporter. A total of five isotope atoms made of a combination of ¹³C and ¹⁵N are incorporated into the mass normalizer arm and the MS/MS reporter region in each reagent but distributed at different locations as labeled by asterisks in Figure 1A. As a result, each reagent has the same nominal parent mass but a unique mass for MS/MS reporter region that could be differentiated on MS/MS spectra: 127 Da, 128 Da, 130 Da and 131 Da. In one embodiment, free thiol is tagged with a label having a mass of 126, 127, 128, 129, 130, or 131 Da. In another embodiment, native disulfide bonds are tagged with a label having a mass that is different from the mass label used to detect the free thiol. A non-limiting example includes a free thiol tagged with a label having a mass of 127 Da, and a native disulfide bond tagged with a label having a mass of 130 Da.

An exemplary workflow is provided in Figure 1B. In one embodiment, an aliquot of the peptide sample is mixed with the first tag and incubated. Excess reagent can be removed from the sample and the sample can be denatured and reduced. In one embodiment, the second mass tag is added to the sample after denaturing the sample. The sample can be enzymatically digested after addition of the second tag. Exemplary methods for enzymatic digestion were discussed above. In another embodiment, the two samples can be added together before further analysis. In some embodiments, up to three pairs of tags can be analyzed in a single run.

2. MS/MS

In one embodiment, the iodoacetyl tag labeled peptides are analyzed using mass spectrometry, for example ultra performance liquid chromatography (UPLC)-MS/MS. In one embodiment, the UPLC is performed on an ethylene bridged hybrid (BEH) column or a charged surface hybrid (CSH) column. The buffers can be trifluoroacetic acid (TFA) buffers or formic acid (FA) buffers. The liquid chromatography can be run for about 90 minutes to about 150 minutes. It will be appreciated that buffers and run times can be optimized for the particular sample being analyzed. In some embodiments, the run includes a re-equilibration step of about

40 minutes. In a preferred embodiment, the UPLC is run on a 150mm CSH column using FA buffers for 150 minutes with a 40 minute re-equilibration step.

After the peptides are separated using UPLC, they can be analyzed using mass spectroscopy. In a preferred embodiment, targeted MS² spectrometry is used. The mass spectrometer can be, for example, a Thermo Scientific Q Exactive hybrid quadrupole Orbitrap® mass spectrometer. An exemplary workflow for targeted MS² is shown in Figure 10. The peptides can be introduced into the mass spectrometer and analyzed using parallel reaction monitoring. In one embodiment, precursor mass filtering is performed in Q1 followed by fragmentation in the HCD cell and high resolution/high mass accuracy (HR/HA) fragment ion detection in the Orbitrap® mass analyzer. In some embodiments, the peptides are captured in MS¹ to verify the correct peptide at the correct mass. The reporter label is broken off of the peptide in the Orbitrap® of MS² and is then quantified in the Orbitrap. The relative abundance of each cysteine corresponding to the detected mass tag can be calculated.

In one embodiment, inclusion masses are included in the MS² scan. For example, the masses of specific cysteines of interest or concern can be programmed into an inclusion list. In one embodiment, the use of an inclusion list can more intensely quantify the residues of interest. In another embodiment, exclusion masses are excluded from the MS² scan. For example, the masses of specific cysteines that are known to be stable or rarely form scrambled bonds can be excluded from the scan.

C. Proteins of Interest

In one embodiment the protein of interest is a protein drug product or is a protein of interest suitable for expression in prokaryotic or eukaryotic cells. For example, the protein can be an antibody or antigen-binding fragment thereof, a chimeric antibody or antigen-binding fragment thereof, an ScFv or fragment thereof, an Fc-fusion protein or fragment thereof, a growth factor or a fragment thereof, a cytokine or a fragment thereof, or an extracellular domain of a cell surface receptor or a fragment thereof. Proteins in the complexes may be simple polypeptides consisting of a single subunit, or complex multisubunit proteins comprising two or more subunits. The protein of interest may be a biopharmaceutical product, food additive or preservative, or any protein product subject to purification and quality standards

In some embodiments, the protein of interest is an antibody, a human antibody, a humanized antibody, a chimeric antibody, a monoclonal antibody, a multispecific antibody, a

bispecific antibody, an antigen binding antibody fragment, a single chain antibody, a diabody, triabody or tetrabody, a dual-specific, tetravalent immunoglobulin G-like molecule, termed dual variable domain immunoglobulin (DVD-IG), an IgD antibody, an IgE antibody, an IgM antibody, an IgG antibody, an IgG1 antibody, an IgG2 antibody, an IgG3 antibody, or an IgG4 antibody. In one embodiment, the antibody is an IgG1 antibody. In one embodiment, the antibody is an IgG2 antibody. In one embodiment, the antibody is an IgG4 antibody. In another embodiment, the antibody comprises a chimeric hinge. In still other embodiments, the antibody comprises a chimeric Fc. In one embodiment, the antibody is a chimeric IgG2/IgG4 antibody. In one embodiment, the antibody is a chimeric IgG2/IgG1 antibody. In one embodiment, the antibody is a chimeric IgG2/IgG1/IgG4 antibody.

In some embodiments, the antibody is selected from the group consisting of an anti-Programmed Cell Death 1 antibody (e.g. an anti-PD1 antibody as described in U.S. Pat. Appln. Pub. No. US2015/0203579A1), an anti-Programmed Cell Death Ligand-1 (e.g., an anti-PD-L1 antibody as described in in U.S. Pat. Appln. Pub. No. US2015/0203580A1), an anti-Dll4 antibody, an anti-Angiopoietin-2 antibody (e.g., an anti-ANG2 antibody as described in U.S. Pat. No. 9,402,898), an anti-Angiopoietin-Like 3 antibody (e.g., an anti-AngPt13 antibody as described in U.S. Pat. No. 9,018,356), an anti-platelet derived growth factor receptor antibody (e.g., an anti-PDGFR antibody as described in U.S. Pat. No. 9,265,827), an anti-Erb3 antibody, an anti-Prolactin Receptor antibody (e.g., anti-PRLR antibody as described in U.S. Pat. No. 9,302,015), an anti-Complement 5 antibody (e.g., an anti-C5 antibody as described in U.S. Pat. Appln. Pub. No US2015/0313194A1), an anti-TNF antibody, an anti-epidermal growth factor receptor antibody (e.g., an anti-EGFR antibody as described in U.S. Pat. No. 9,132,192 or an anti-EGFRvIII antibody as described in U.S. Pat. Appln. Pub. No. US2015/0259423A1), an anti-Proprotein Convertase Subtilisin Kexin-9 antibody (e.g., an anti-PCSK9 antibody as described in U.S. Pat. No. 8,062,640 or U.S. Pat. No. 9,540,449), an Anti-Growth and Differentiation Factor-8 antibody (e.g. an anti-GDF8 antibody, also known as anti-myostatin antibody, as described in U.S. Pat Nos. 8,871,209 or 9,260,515), an anti-Glucagon Receptor (e.g. anti-GCGR antibody as described in U.S. Pat. Appln. Pub. Nos. US2015/0337045A1 or US2016/0075778A1), an anti-VEGF antibody, an anti-IL1R antibody, an interleukin 4 receptor antibody (e.g., an anti-IL4R antibody as described in U.S. Pat. Appln. Pub. No. US2014/0271681A1 or U.S. Pat Nos. 8,735,095 or 8,945,559), an anti-interleukin 6 receptor antibody (e.g., an anti-IL6R antibody as

described in U.S. Pat. Nos. 7,582,298, 8,043,617 or 9,173,880), an anti-IL1 antibody, an anti-IL2 antibody, an anti-IL3 antibody, an anti-IL4 antibody, an anti-IL5 antibody, an anti-IL6 antibody, an anti-IL7 antibody, an anti-interleukin 33 (e.g., anti-IL33 antibody as described in U.S. Pat. Nos. 9,453,072 or 9,637,535), an anti-Respiratory syncytial virus antibody (e.g., anti-RSV antibody as described in U.S. Pat. Appln. Pub. No. 9,447,173), an anti-Cluster of differentiation 3 (e.g., an anti-CD3 antibody, as described in U.S. Pat. Nos. 9,447,173 and 9,447,173, and in U.S. Application No. 62/222,605), an anti-Cluster of differentiation 20 (e.g., an anti-CD20 antibody as described in U.S. Pat. Nos. 9,657,102 and US20150266966A1, and in U.S. Pat. No. 7,879,984), an anti-CD19 antibody, an anti-CD28 antibody, an anti-Cluster of Differentiation-48 (e.g. anti-CD48 antibody as described in U.S. Pat. No. 9,228,014), an anti-Fcγd1 antibody (e.g. as described in U.S. Pat. No. 9,079,948), an anti-Middle East Respiratory Syndrome virus (e.g. an anti-MERS antibody as described in U.S. Pat. Appln. Pub. No. US2015/0337029A1), an anti-Ebola virus antibody (e.g. as described in U.S. Pat. Appln. Pub. No. US2016/0215040), an anti-Zika virus antibody, an anti-Lymphocyte Activation Gene 3 antibody (e.g. an anti-LAG3 antibody, or an anti-CD223 antibody), an anti-Nerve Growth Factor antibody (e.g. an anti-NGF antibody as described in U.S. Pat. Appln. Pub. No. US2016/0017029 and U.S. Pat. Nos. 8,309,088 and 9,353,176) and an anti-Protein Y antibody. In some embodiments, the bispecific antibody is selected from the group consisting of an anti-CD3 x anti-CD20 bispecific antibody (as described in U.S. Pat. Appln. Pub. Nos. US2014/0088295A1 and US20150266966A1), an anti-CD3 x anti-Mucin 16 bispecific antibody (e.g., an anti-CD3 x anti-Muc16 bispecific antibody), and an anti-CD3 x anti-Prostate-specific membrane antigen bispecific antibody (e.g., an anti-CD3 x anti-PSMA bispecific antibody). In some embodiments, the protein of interest is selected from the group consisting of abciximab, adalimumab, adalimumab-atto, ado-trastuzumab, alemtuzumab, alirocumab, atezolizumab, avelumab, basiliximab, belimumab, benralizumab, bevacizumab, bezlotoxumab, blinatumomab, brentuximab vedotin, brodalumab, canakinumab, capromab pendetide, certolizumab pegol, cemiplimab, cetuximab, denosumab, dinutuximab, dupilumab, durvalumab, eculizumab, elotuzumab, emicizumab-kxwh, emtansinealirocumab, evinacumab, evolocumab, fasinumab, golimumab, guselkumab, ibritumomab tiuxetan, idarucizumab, infliximab, infliximab-abda, infliximab-dyyb, ipilimumab, ixekizumab, mepolizumab, necitumumab, nesvacumab, nivolumab, obiltoxaximab, obinutuzumab, ocrelizumab, ofatumumab, olaratumab, omalizumab, panitumumab,

pembrolizumab, pertuzumab, ramucirumab, ranibizumab, raxibacumab, reslizumab, rinucumab, rituximab, sarilumab, secukinumab, siltuximab, tocilizumab, tocilizumab, trastuzumab, trevogrumab, ustekinumab, and vedolizumab.

In some embodiments, the protein of interest is a recombinant protein that contains an Fc moiety and another domain, (e.g., an Fc-fusion protein). In some embodiments, an Fc-fusion protein is a receptor Fc-fusion protein, which contains one or more extracellular domain(s) of a receptor coupled to an Fc moiety. In some embodiments, the Fc moiety comprises a hinge region followed by a CH2 and CH3 domain of an IgG. In some embodiments, the receptor Fc-fusion protein contains two or more distinct receptor chains that bind to either a single ligand or multiple ligands. For example, an Fc-fusion protein is a TRAP protein, such as for example an IL-1 trap (e.g., rilonacept, which contains the IL-1RAcP ligand binding region fused to the IL-1R1 extracellular region fused to Fc of hIgG1; see U.S. Pat. No. 6,927,004, which is herein incorporated by reference in its entirety), or a VEGF trap (e.g., aflibercept or ziv-aflibercept, which comprises the Ig domain 2 of the VEGF receptor Flt1 fused to the Ig domain 3 of the VEGF receptor Flk1 fused to Fc of hIgG1; see U.S. Pat. Nos. 7,087,411 and 7,279,159). In other embodiments, an Fc-fusion protein is a ScFv-Fc-fusion protein, which contains one or more of one or more antigen-binding domain(s), such as a variable heavy chain fragment and a variable light chain fragment, of an antibody coupled to an Fc moiety.

D. Producing mAb with Little or No Disulfide Heterogeneity

One embodiment provides methods of producing a protein drug product containing little or no free thiols. An exemplary method includes culturing cells producing the antibody in a cell culture under suitable conditions to produce the antibody, purifying the antibody under suitable conditions to extract the antibody, admixing the antibody with excipients under suitable conditions to stabilize the antibody, obtaining a sample of the antibody from the cell culture, following purification of the antibody from the cell culture, or following the addition of excipients to the purified antibody, characterizing disulfide bonds of the antibody according to the disclosed methods, and modifying one or more cell culture, purification or excipient conditions to reduce the amount of crossed hinge disulfide bonds of the antibody.

The one or more cell culture, purification, or excipient conditions that are changed to reduce the amount of free thiols in the antibody include but are not limited to temperature, pH,

oxygen levels, reactive oxygen species, surfactants, or combinations thereof. In one embodiment, an amino acid free strategy of cell culture could affect disulfide bond formation.

In one embodiment, the cells producing the antibody are Chinese hamster ovary cells. In another embodiment, the cells are hybridoma cells. In embodiment, the protein drug product is an IgG monoclonal antibody are a fragment thereof.

In one embodiment, the protein drug product contains no free thiols. In another embodiment the protein drug product contains less than 5 free thiols, less than 4 free thiols, less than 3 free thiols, less than 2 free thiols in an IgG molecule. In some embodiments the free thiol is in a light chain of an IgG molecule. In some embodiments, the free thiol is in a heavy chain of an IgG molecule. In some embodiments, there is at least one free thiol in a light chain and at least one free thiol in a heavy chain of an IgG molecule.

EXAMPLES

Example 1: Analysis of Site-Specific Free Thiols

Methods

Cysteine Labeling

The workflow from Figure 1B was followed. In brief, a control sample and an antibody sample were labeled with iodo-TMT reagents (Figure 1A). The excess reagents were removed and the samples were denatured and reduced using 0.5M TCEP and incubated for 1 hour. After denaturation and reduction, a second label was added to the samples. The samples were then combined, enzymatically digested, and analyzed using UPLC-MS/MS. UPLC conditions were as follows: Samples were run on a 150mm CSH column for a 150 minutes run, including a 40 minute re-equilibration step. FA buffers were used.

Spike-in Study

For the spike-in study, aliquots of two different samples of the same protein were prepared. The proteins were denatured using GuanHCl and reduced. The samples were labeled with two different IodoTMT tags. A spike-in concentration gradient was then prepared. 10%, 5%, 1%, 0.5%, or 0.1% of protein sample 1 was added to aliquots of protein sample 2. The samples were then enzymatically digested and analyzed using UPLC/MS-MS (Figure 2).

Results:

Table 1 shows the results from targeted MS² analysis of a Regeneron antibody labeled with iodo-TMT tags to identify site-specific free thiols. As can be seen in the table, all sixteen cysteine residues were identified.

5 Figures 3A-3B show a comparison of data dependent MS² run and a targeted MS² run (with an inclusion list) for the same sample. As shown in the figures, targeted MS² run with an inclusion list of specific cysteines can help to more intensely quantify specific cysteine residues. In addition, the use of the iodo-TMT tags coupled with targeted MS² allows for better separation of labeled cysteines. Figure 4A-4B shows clear separation of 128 Da labeled cysteine and a 131
10 Da labeled cysteine. The disclosed method has very clear separation with no concern of overlapping isotope patterns as is seen in current methods such as labeled IAA methods.

To further show the utility of the disclosed method, different lots of the same antibody product were compared. Figures 5A-5G show the relative abundance of free thiols and disulfide bonded cysteines in different lots of the same antibody.

15 Additional validation of the method was performed using a spike-in study. Various percentages of labeled protein sample 1 were added into labeled protein sample 2. Figure 6 shows the results of the spike-in study. The majority of the cysteines showed the expected amount of free thiols that would be predicted by the amount of protein sample 1 spiked into protein sample 2, indicating that the method is effective at predicting relative abundance of free
20 thiols.

Table 1. Targeted MS² Results

Mass [m/z]	Polarity	Start [min]	End [min]	NCE	CS [z]	Comment	Retention Time
827.7385	Positive	57	63	30	3	C22H- 2481.2010	60.57
781.8890	Positive	30	36	30	2	C96H- 1562.7694	33.72
797.4441	Positive	40	46	30	2	C145H- 1593.8809	43.90
1748.1358	Positive	72	78	30	4	C201H- 6985.5196	75.62
391.2077	Positive	10	16	30	2	C221H- 781.4062	14.70
1130.2953	Positive	58	64	30	3	C227H/C230H- 3388.8659	62.23
804.4156	Positive	45	51	30	3	C262H- 2411.2322	48.67
500.2775	Positive	16	22	30	2	C322H- 999.5477	19.64
717.4208	Positive	42	48	30	2	C368H- 1433.8334	46.23
1025.4963	Positive	46	52	30	3	C426H- 3073.4685	50.15
489.7835	Positive	20	26	30	2	C23L- 978.5581	23.53
1637.7747	Positive	64	70	30	3	C94L- 4911.3185	67.71
690.7049	Positive	60	66	30	3	C140L- 2070.1001	63.31
716.7147	Positive	35	41	30	3	C200L- 2148.1294	38.98
571.2846	Positive	22	28	30	2	C220L- 1141.5597	25.24

Table 2 summarizes the results of free thiol and disulfide bond abundance for a Regeneron antibody product. Site specific free thiol percentages were calculated for each cysteine site. All sixteen cysteines were identified and quantified. A threshold of 5-6% free thiols was set, anything below the threshold does not show a potential for disulfide bond breakage. Figure 7 presents the relative abundance of disulfide bond pairs in different lots of a Regeneron antibody product.

The disclosed methods and results can be used for site-specific total free thiol quantitation and as a potential indicator of disulfide scrambling locations. In addition it can be used to identify intrachain disulfides leading to half molecule or aggregation.

Table 2. Process comparability.

Disulfide Bond		Site Specific Free Thiol Content (%)										Avg	STD Dev
Disulfide Bond	Cys Site	Process A					Process B						
		Lot 1	Lot 2	Lot 3	Lot 4	Lot 5	Lot 6						
C22H-C96H	C22H	0.09	0.10	0.08	0.08	0.07	0.10					0.09	0.01
	C96H	0.03	0.04	0.04	0.05	0.05	0.05					0.04	0.01
C144H-C200H	C144H	0.62	0.54	0.47	1.02	1.13	0.96					0.79	0.21
	C200H	0.99	0.87	0.87	1.44	1.37	1.34					1.15	0.24
C258H-C318H	C258H	0.48	0.42	0.42	0.78	0.86	0.88					0.64	0.15
	C318H	0.53	0.57	0.50	1.02	1.03	1.04					0.78	0.21
C364H-C422H	C364H	1.64	1.54	1.64	2.23	2.63	2.65					2.05	0.27
	C422H	2.22	2.05	1.96	2.34	3.03	2.97					2.43	0.14
C23L-C88L	C23L	0.09	0.11	0.10	0.11	0.10	0.13					0.11	0.01
	C88L	0.16	0.19	0.15	0.20	0.21	0.22					0.19	0.02
C134L-C194L	C134L	0.24	0.24	0.22	0.28	0.24	0.34					0.26	0.02
	C194L	0.07	0.04	0.05	0.08	0.08	0.10					0.07	0.02
Hinge	C223H/226H	0.00	0.00	0.00	0.00	0.00	0.00					0.00	0.00
C131H-C214L	C131H	0.18	0.18	0.19	0.20	0.22	0.21					0.20	0.01
	C214L	0.20	0.18	0.24	0.11	0.12	0.17					0.17	0.05

While in the foregoing specification this invention has been described in relation to certain embodiments thereof, and many details have been put forth for the purpose of illustration, it will be apparent to those skilled in the art that the invention is susceptible to additional embodiments and that certain of the details described herein can be varied considerably without
5 departing from the basic principles of the invention.

All references cited herein are incorporated by reference in their entirety. The present invention may be embodied in other specific forms without departing from the spirit or essential attributes thereof and, accordingly, reference should be made to the appended claims, rather than
10 to the foregoing specification, as indicating the scope of the invention.