RECOMBINANT PROTEIN CHARACTERIZATION

A Primer

The Measure of Confidence

Agilent Technologies

CONTENTS

1 INTRODUCTION TO RECOMBINANT PROTEINS

Historically, most drugs have been based on chemically synthesized small molecules. However, biological entities now represent a significant share of pharmaceutical sales and future growth potential, due to their targeted mode of action, increasing patent expirations in small molecule drugs, and competition caused by generic versions of small-molecule drugs. While biologics can be small RNAs (e.g., microRNA), oligonucleotides, peptides, or vaccines, they are most commonly recombinant proteins. Their specificity has been the secret to their success, due to their cell-specific modality and relative lack of compound-specific toxicity. The discovery and development of protein therapeutics is growing rapidly and soon will equal and even surpass the efforts in chemical entity drug development. For example, more than 30 monoclonal antibodies have been approved for therapeutic use, and every major pharmaceutical company has a significant protein therapeutic development program.

Producing recombinant protein therapeutics

While small-molecule drugs are typically produced by chemical synthesis and have well-defined chemical structures, therapeutic proteins pose a much more difficult production and characterization challenge. These proteins can be isolated from a variety of sources, including human, animal, or microorganism, but they are most often produced by biotechnological means, using recombinant DNA technologies. Typically, the recombinant protein of interest is produced in a cell culture (eukaryotic or bacterial) that can be grown to high density and thus produce large amounts of the putative biopharmaceutical. The product must then be purified from the cellular components, which is a more complex process than that used for small-molecule drugs. Finally, like small molecule drugs, the protein therapeutic must be formulated to produce a stable drug.

Taking biopharmaceuticals from discovery to development

Recombinant protein drug development starts with the identification of disease pathways that can be used to identify lead candidates. Recombinant proteins are derived from recombinant DNA (rDNA), which consists of at least two DNA sequences not normally found together in nature. These rDNA libraries and the proteins created from them are then screened for use as inhibitors of those disease pathways, as replacement therapies, or as other anti-disease agents.

One technique used for screening recombinant proteins is phage display. Phages packed with rDNA are introduced to well plates with isolated or immobilized disease pathway protein targets. The phages produce and secrete the desired protein, and protein-protein interactions with the targets can then be measured. Once candidate proteins are selected, the rDNA for each candidate is inserted into a plasmid which is then cloned in either prokaryotic or eukaryotic cells. Small-scale fermentation of the cloned cells can then be carried out, and purification schemes can be developed to test the proteins for potential safety parameters and efficacy against *in vitro* and *in vivo* disease models. A variety of methodologies and unique approaches can be used to identify, isolate, and characterize these recombinant protein leads. After a protein therapeutic has been identified, validated, and selected for drug development, the manufacturing processes become regulated.

Characterizing the protein therapeutic

The focus in the development stage of recombinant protein drugs is the production and purification of the biological entity. Cost-effective scale-up of the cell culture or fermentation process, and purification of the biologic must be developed. In recombinant protein manufacturing, the final structure of the product is highly dependent on post-translational modifications that can occur during fermentation, as well as other modifications caused by the manufacturing process. These must be characterized and controlled. Impurities, particularly those that may be pyrogenic or induce an immunologic reaction, must be identified, characterized, and minimized in the final product. Proper formulation of the protein drug is also important, in order to minimize degradation and loss of potency over time, which can be caused by chemical changes such as oxidation and hydrolysis. Accurate and reproducible characterization methods are therefore an absolute requirement to support and guide decisions made in developing the manufacturing process and product formulation of protein therapeutics.

Putting critical QA/QC procedures in place

Drug manufacturers must ensure product consistency, quality, and purity by rigorously monitoring both the manufacturing process over time, as well as the finished product. Robust, rigorous, and accurate methods for quality assurance and quality control are accordingly required to confirm conformance to pre-determined specifications and assure final product safety and efficacy. Biologic products must be tested for identity, impurities, quantity (e.g., protein content), and stability. Thorough quality control is required by all GMP regulations in order to obtain FDA and EMA approval, and is frequently subject to regulatory inspections.

Meeting the recombinant protein therapeutics characterization challenge

A biologic must be "well-characterized" to assure its safety and efficacy. The FDA Center for Biologics Evaluation and Research (CBER) established the use of the term in 1996, and the concept has been further refined as the technology for characterization of biological molecular diversity has evolved. For protein pharmaceuticals, well-characterized means that the natural molecular heterogeneity, impurity profile, and potency can be defined with a high degree of confidence. The complexity of the proteins being developed ensures that there is no one analytical platform or application that can meet all of these needs.

Biologics are often heterogeneous mixtures of closely related molecular weights and charged isoforms. They are derived from living cells and typically include a complex pattern of product- and process-related impurities. In addition, recombinant proteins undergo complex post-translational modifications, have a highly specific three-dimensional structure that depends in part on disulfide bridges, and have the potential for aggregation, adsorption, and truncation. A comprehensive chemical, physical, and conformational characterization is essential to understand the heterogeneity, impurity profile, and potency of a recombinant protein destined for pharmaceutical use. The precise amino acid sequence, molecular weight, charge variances, glycosylation, aggregation level, and oxidation level are all key components of thorough characterization of a protein drug.

The following pages describe several categories of analyses that are required to fully characterize recombinant proteins intended for therapeutic use, and the current state of the methodologies and instrument platforms used to perform those analyses.

2 BIOLOGIC PROCESS MONITORING

Appropriate analytical techniques are essential to effective process monitoring during the development and manufacturing of a protein pharmaceutical. The demand for rapid and effective monitoring techniques is being driven by the need to better understand the biological production processes, as well as a need for improved control of feeding of the organisms producing the recombinant protein and other process parameters. Production processes must also be robust, standardized, transferable, and operator-independent, further impelling the need for effective monitoring. The acceleration of process development to reduce the time to market and ensure optimal exploitation of the biological production process is also a strong driver. Finally, accurate and reliable monitoring techniques are required to comply with the FDA and EMA Process Analytical Technology (PAT) guidances. Attaining the proper efficiency of the fermentation, extraction, and purification processes is critical to the success of the product.

Protein A chromatography is routinely used to affinity-purify immunoglobulin G (IgG) antibodies for monoclonal antibody production. Analytical Protein A formats are also used to prepare samples for analysis and determine titer, which is the concentration of the target protein in a bioreactor. The Agilent Bravo for Protein Purification system equipped with AssayMAP Protein A cartridges is an excellent automation tool to rapidly determine antibody titers and to prepare small amounts of antibodies for other analytical characterization techniques such as glycosylation, charge isoform analysis, and sizing assays. The complete capture workflow takes place in a reusable, microscale chromatographic cartridge (Figure 1). The Protein A tips are available in 96-well format, making high-throughput analysis possible. The AssayMap cartridges can also be packed with other resins useful for rapid screening of antibodies and other recombinant proteins like Protein G, Protein L, ion exchange, reverse phase, and streptavidin. AssayMap plates can also be packed with process media for screening purposes. For example, an ion exchange plate could be screened for binding conditions at various pH and salt concentrations.

Figure 1. The complete capture workflow for the Agilent Bravo for Protein Purification system equipped with AssayMAP Protein A cartridges, as it would be used for protein titer analysis. Protein A captures the mAb at neutral pH; the mAb is then eluted at low pH using an acidic buffer. A calibration curve can be created to accurately measure the concentration of a mAb produced in a bioreactor.

The speed of the Agilent 2100 Bioanalyzer for microfluidic on-chip analysis of proteins makes it a useful process monitoring tool. In combination with an immunoprecipitation method, it can deliver sensitive, selective, and quantitative detection of proteins from crude *E. coli* lysates. For example, this method has been used to quantify β-galactosidase in a crude cell lysate, and provided highly reproducible results over multiple orders of magnitude. The 2100 Bioanalyzer can also be used to confirm proper protein purification and removal of unwanted impurities.

The production of immunoglobulin M (IgM) monoclonal antibodies requires an accurate, rapid, and simple analytical method to measure levels in cell culture supernatants, and to document the distribution of IgM and protein contaminants in fractions from various process chromatography steps. The Agilent Bio Monolith QA column (a strong anion exchange column) can be used to rapidly monitor the IgM purification process and quantify IgM concentration from cell culture supernatants (Figure 2).

One parameter that can require monitoring during the production of recombinant protein therapeutics is the degree of PEGylation (the attachment of polyethylene glycol, or PEG) of the protein. PEGylation of proteins can help improve the safety and efficiency of many therapeutics by enhancing drug solubility, reducing dosage frequency and toxicity, and increasing drug stability. The degree of PEGylation can be monitored by determining the increase in molecular weight using size-exclusion chromatography (SEC) HPLC. Since the surface charge of the protein often varies depending on the position of the PEG moiety, ion exchange chromatography (IEC) HPLC can also be used to resolve PEGylated isoforms. The degree of PEGylation can also be determined by electrophoretic analysis, and the Agilent 2100 Bioanalyzer offers an automated alternative to SDS-PAGE analysis that simplifies sample preparation and shortens analysis time, providing digital data within 30 minutes (Figure 3).

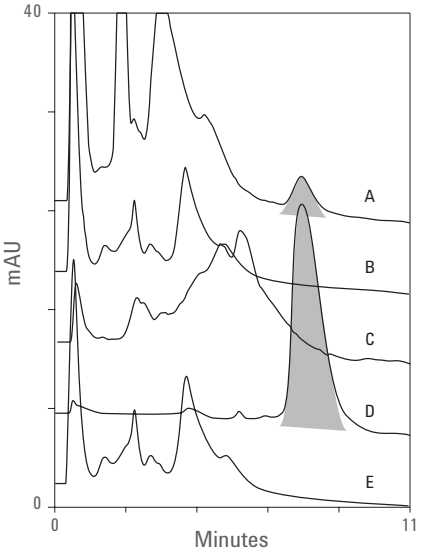


Figure 2. A series of hydroxyapatite chromatography fractions (A through E) that were collected and analyzed using the Agilent Bio-Monolith QA analytical column to confirm proper purification and column regeneration. A. Cell culture supernatant; B. Flow-through from sample load; C. Pre-elution wash; D. IgM elution peak; E. Post-elution cleaning.

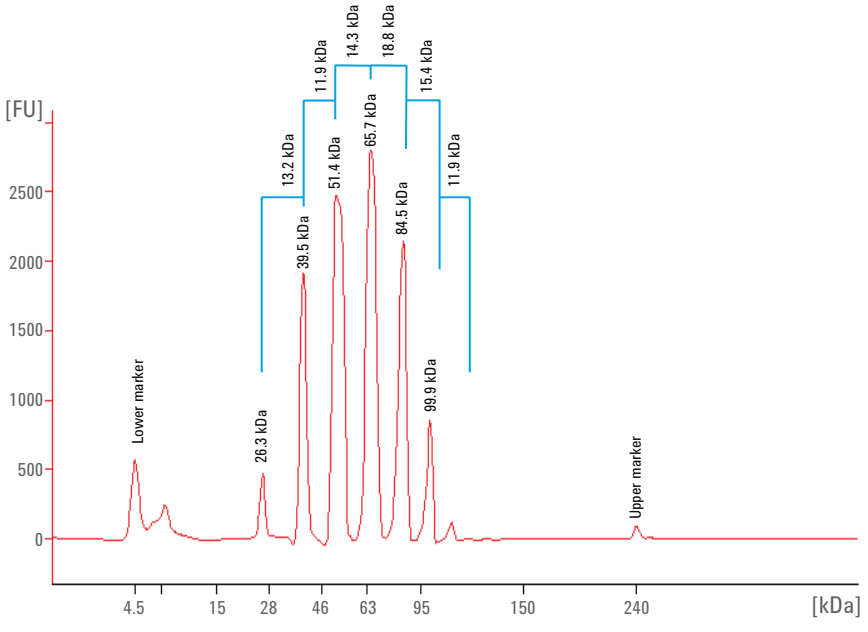


Figure 3. Analysis of the PEGylated isoforms of a recombinant therapeutic protein using the Agilent 2100 Bioanalyzer and the Protein 230 Kit. The free protein containing no PEG migrates at 26.3 kDA. Since each PEG moiety adds about 5 kDa of molecular weight, the PEGylated forms containing three (39.5 kDa), five (51.4 kDa), and more PEG moieties are clearly resolved. Further analysis by mass spectrometry analysis is required to definitively assign the number of PEG molecules present in each isoform.

Bioprocess analytics techniques summary

Techniques	Technique Description	Benefits	Limitations
MAb Titer, Capture, Sample Prep: Protein A-LC/UV	HPLC separation based on Protein A affinity for IgG (monoclonal antibodies), separation based on a step gradient from low to high pH	On-line separation and detection	Off-line, commonly hands-on calibration curve preparation
		Ability to connect LC to bioreactor for real-time analysis	Expensive LC columns, non-disposable, not single-use capable
			Lower throughput, single sample per LC system run
MAb Titer, Capture, Sample Prep: Microchromatography Cartridge Protein A, Spectral Reader for Titer	BRAVO liquid handling system, AssayMap microchromatography head, AssayMap microchromatography Protein A cartridge, separation of monoclonal antibodies from culture or other matrices, separation based on a strp gradient from high to low pH, titer detection by spectral reader, other analysis by LC, LE/MS, CE and other techniques	Calculate titers while simultaneously preparing samples for analytical characterization	Currently no online connection for real-time reactor analysis
		96-well plate format, high throughput	For titers plates need to read by separate spectral reader
		Multi-use tips, can be single use disposable at much lower cost than Protein A LC columns	
		Hands-off calibration curve preparation and separations	
Monitor Process Chromatography: Analytical Ion Chromatography	HPLC system with ion exchange columns (strong and weak cation, or anion exchange packing material), interaction with the column particles based on total net charge of the protein, elution occurs across either an increasing salt or pH gradient or a combination of the two	Provides chromatographic separation of protein charge variants, various cation and anion exchange resins available offering many selectivity choices	Commonly long gradients with high buffer consumption
		Similar chromatographic phases/mechanisms as the process media used, can easily be used to determine separation efficiencies and recoveries	Lower resolution separations than cIEF
		Ability to fractionate and collect separated charge isoforms for further analysis (MS analysis possible after desalting, buffer exchange)	Not mass-spectrometry compatible
Monitor Process Chromatography: Analytical Size-Exclusion Chromatography	HPLC system, size-exclusion HPLC column, separation is based on protein molecular radius and passage speed through packed porous particles, UV detection	High-resolution separations of aggregates and impurities based on their size	Limitation of separation window, not easy to resolve high molecular weight aggregates and low molecular weight impurities in the same separation
		Short analysis time compared to SDS-PAGE	Does not provide mass information, difficult to connect to mass spectrometer
		Simple method, commonly used in process monitoring to monitor aggregate purification and removal of impurities	
		Ability to collect fractions for further analysis	
Monitor PEGylation process: Chip-based (Bioanalyzer) Protein Electrophoresis Assays	2100 Bioanalyzer system, High Sensitivity Protein 250 or 230 Kit (includes chip, fluorescence labeling buffer, and all required reagents), electrophoretic separation combined with fluorescence detection	Ability to distinguish between single and multiple PEGylation forms of the same proteins based on their size	Currently no online connection for real-time reactor analysis
		High-resolution separations	No sample collection for further analysis
		Quantitative results	
		Short analysis time compare to SDS-PAGE	

Agilent applications literature

Publication #	Title
5990-7203EN	High throughput purification of human IgG using the Agilent Bravo for Protein Purification and AssayMAP1 protein A cartridges
5989-9733EN	Rapid Human Polyclonal IgG Quantification Using the Agilent Bio-Monolith Protein A HPLC Column
5990-6153EN	Monitoring protein fate during purification with the Agilent 2100 Bioanalyzer
5989-9674EN	Rapid IgM Quantification in Cell Culture Production and Purification Process Monitoring Using the Agilent Bio-Monolith QA Column

INTACT PROTEIN IDENTITY, PURITY, AND IMPURITIES

3

Although protein biologics are for the most part relatively stable molecules, a number of chemical modifications and degradation reactions can occur during manufacturing, formulation, and storage. Many proteins can be proteolytically cleaved during the purification process, which may have major ramifications due to the fact that the cleaved contaminants may cause immune reactions in patients. The existence of modification and degradation reactions necessitates reliable and sensitive methods to assess protein purity and structural integrity during process development, as well as during manufacturing.

Accurate-mass measurements of intact proteins, whole subunits, or domains are useful for the rapid verification of sequence composition and identification of post-translational modifications, degradation, and sample handling artifacts. Electrophoresis, chromatography, and mass spectrometry (MS) techniques are most often used to determine the molecular weight of intact proteins and their degradation products. One of the most common analyses is size-exclusion chromatography (SEC), which is used to determine if a protein has a monomeric structure and maintains that structure throughout manufacturing and formulation.

SDS electrophoresis can provide an accurate estimate of protein molecular weights and purity, and automated chip-based systems are available to provide monitoring capability throughout the development and production process. The Agilent 2100 Bioanalyzer is designed to automatically size and quantitate a wide range of proteins. Heavily glycosylated proteins can be problematic for any electrophoretic method, because of the large carbohydrate attachments. However, removal of the glycan components enables proper migration and more accurate molecular weight estimation. In addition, many of the different species, including glycosylated, deglycosylated, and various isoforms, can be clearly visualized using the Agilent 2100 Bioanalyzer.

One of the methods of choice for molecular weight analysis of protein biologics is capillary electrophoresis (CE). Methods range from CE with UV or laser-induced fluorescence (LIF) detection, to combined solutions with both LIF and MS detectors (CE/LIF/MS). Due to the fact that it separates primarily on the basis of charge, mass, and shape of the molecule, CE can be a very effective tool for assessing the primary structure and glycosylation state of antibodies. Simultaneous coupling to both an LIF detector and a time-of-flight (TOF) MS provides previously unattainable sensitivity in the detection and analysis of minor components in mAb samples (Figure 1).

Liquid chromatography coupled to TOF or quadrupole time-of-flight (QTOF) MS can also provide high-mass-accuracy information for therapeutic proteins, with much higher accuracy than electrophoretic methods. A typical analysis can be completed in less than an hour and provide vital information about the protein and any degradation or modification impurities. For example, electrospray ionization (ESI) LC/QTOF MS analysis of a monoclonal antibody can be performed in only 9 minutes, with mass accuracy better than 25 ppm, and the system can resolve the various antibody subpopulations present in the sample.

Many proteins of therapeutic interest, including mAbs, are protein complexes held together by disulfide bridges. These complexes can be dissociated into their subunits by reduction, and HPLC followed by mass spectrometry can then be used to analyze the subunits. While intact mAb analysis provides a lot of insight regarding purity and glycosylation level, analysis of the

light and heavy chain subunits released by reduction of the disulfide bridge can generate even more information. The level and types of glycosylation of the heavy chain can be detected as well as any unwanted modifications such as non-enzymatic, non-specific glycosylation (Figure 2).

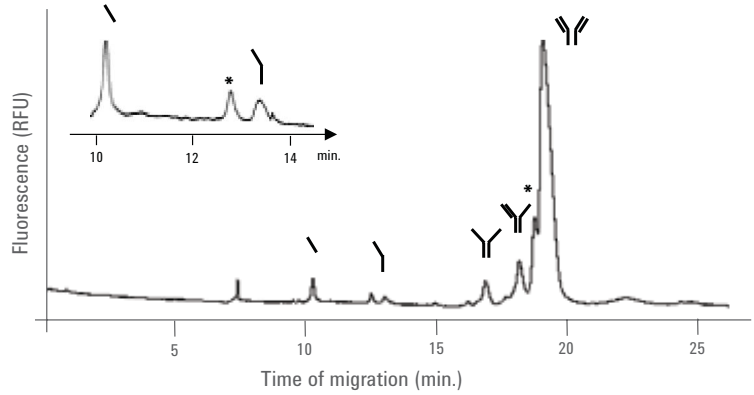


Figure 1: Separation of partially reduced IgG using CE/LIF with native fluorescence (266 nm excitation; >290 nm emission) at 1 mg/mL. The native glycosylated antibody complex is resolved from the complex lacking glycosylation on the heavy chain (*). The inset shows an expansion of the region between 10 and 14 minutes, with the light chain, the heavy chain without N glycosylation (*) and the heavy chain with glycosylation.

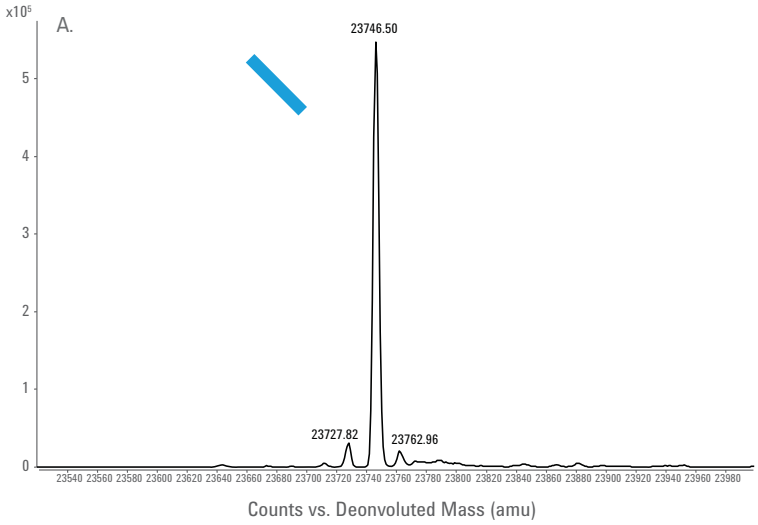


Figure 2: Analysis of a reduced monoclonal antibody using the Agilent 1260 Infinity HPLC-Chip/MS system with ZORBAX 300SB-C8 (300) enrichment and analytical columns, coupled to an Accurate-Mass Q-TOF MS.
A) Deconvoluted spectrum of light chain of mAb (shown as blue bar).
B) Deconvoluted spectrum of heavy chain of mAb (shown as red bar with different sugar attachments).

Intact protein identity, purity, and impurities techniques summary

Techniques	Technique Description	Benefits	Limitations
SDS electrophoresis	Electrophoresis chamber, acrylamide gel, buffers and reagents, Coomassie Blue or silver stain (proteins will separate based on their charge and size, smaller proteins or monomers will migrate further through the gel than dimers and aggregates)	Ability to visually identify proteins, aggregates, and impurities in the migrating bands	Low resolution for intact proteins and impurities with similar molecular size, but different modification and structure
		Samples can be collected for further analysis by cutting out gel bands	Non-quantitative
Analytical Size-Exclusion Chromatography	HPLC system, size-exclusion HPLC column, separation is based on protein molecular radius and passage speed through packed porous particles, UV detection	High-resolution separations of aggregates and impurities based on their size	Limitation of separation window, not easy to resolve high molecular weight aggregates and low molecular weight impurities in the same separation
		Short analysis time compared to SDS-PAGE	Does not provide mass information, difficult to connect to mass spectrometer
		Simple method, commonly used in process monitoring to monitor aggregate purification and removal of impurities	
		Ability to collect fractions for further analysis	
Chip-based (Bioanalyzer) Protein Electrophoresis Assays	2100 Bioanalyzer system, High Sensitivity Protein 250 Kit (includes chip, fluorescence labelling buffer, and all required reagents), electrophoretic separation combined with fluorescence detection	Quantitative results	Currently no online connection for real-time reactor analysis
		High-resolution separations	No sample collection for further analysis
		Ability to resolve low molecular weight impurities along with intact protein and possible aggregates	
		Short analysis time compare to SDS-PAGE	
CE/UV	Capillary electrophoresis system (7100 CE) used in CZE mode, capillaries and buffers, separation based on molecular radius, separation of intact proteins from impurities, UV detection	Orthogonal separation technique to reversed phase HPLC	Does not provide mass information, requires further verification of molecular structure to determine if impurities come from protein of interest
		In some cases, improved resolution of proteins compared to HPLC separations	No sample collection for further analysis
CE/LIF	Capillary electrophoresis system (7100 CE) used in CZE mode, capillaries, APTS label and buffers, separation based on molecular radius, separation of intact proteins from impurities, laser-induced fluorescence detection of APTS labelled proteins and impurities, identification based on ladder standard	Orthogonal separation technique to reversed-phase HPLC	Does not provide mass information, requires further verification of molecular structure to determine if impurities come from protein 1
		In some cases, improved resolution of proteins compared to HPLC separations	No sample collection for further analysis
		Increased specificity with fluorescent label and LIF detection	Potential issues with incomplete labelling of intact proteins and impurities

Techniques	Technique Description	Benefits	Limitations
CE/LIF/MS	Capillary electrophoresis system (7100 CE) used in CZE mode, capillaries, APTS label and buffers, separation based on molecular radius, separation of intact proteins from impurities, laser-induced fluorescence detection of APTS labelled proteins and impurities, MS detection, identification of intact proteins, isoforms and impuritites based on ladder standard and molecular weight	Orthogonal separation technique to reversed-phase HPLC	Does not provide mass information, requires further verification of molecular structure to determine if impurities come from protein of interest
		In some cases, improved resolution of proteins compared to HPLC separations	No sample collection for further analysis
		Further increased specificity with fluorescent label and MS detection in tandem	Potential issues with incomplete labelling of intact proteins and impurities
			Potential MS signal supression by free APTS label
RP-LC/UV	HPLC system, separation based on hydrophobicity across an increasing organic gradient commonly by C8 or smaller carbon chain reversed-phase HPLC column, UV detection (commonly used for monitoring disulfide patterns/shifts and protein degradation)	Wide variety of columns available to separate various variants and impurities (e.g., C18, C8, C4, C3, diphenyl, Hillic)	No mass information obtained, further analysis requires fraction collection
		Fractions can be collected and are mass-spectrometry compatible	Fractions can be collected and are mass-spectrometry compatible
		Provides robust methods that are commonly transferred to QA/QC	
RP-LC/QTOF MS (ESI)	HPLC system, separation based on hydrophobicity across an increasing organic gradient by C8 or smaller carbon chain reversed-phase HPLC column, MS detection	Wide variety of columns avaiable to separate various variants and impurities (e.g., C18, C8, C4, C3, diphenyl, Hillic)	Not commonly used in QA/QC, requires expert user for LC/MS use and data analysis
		Fractions can be collected and are mass-spectrometry compatible	
		Provides mass information for intact protein, possible variants, impurities and non-target proteins derived impurities (these can be collected, digested, and mapped)	

Agilent applications literature

Publication #	Title
5990-3445EN	Primary Characterization of a Monoclonal Antibody Using Agilent HPLC-Chip Accurate-Mass LC/MS Technology
5989-7406EN	Accurate Mass LC/TOF MS for Molecular Weight Confirmation of Intact Proteins
5989-0332EN	Glycoprotein sizing on the Agilent 2100 Bioanalyzer
5989-8940EN	Performance characteristics of the High Sensitivity Protein 250 Assay for the Agilent 2100 Bioanalyzer
5989-6840EN	Comparison of ZORBAX StableBond 300 Å LC Columns to Optimize Selectivity for Antibody Separations Using HPLC and LC/MS

GLYCAN ANALYSIS AND PROFILING

4

Glycoproteins are involved in immune defense, cell growth, and cell-cell adhesion, and the glycans that help mediate these functions can take on a myriad of complex structures, requiring analytical techniques that can elucidate them. More than 90% of the protein drugs in existence are glycoproteins. For instance, many different recombinant forms of immunoglobulins (e.g., monoclonal antibodies, mAbs) are produced as therapeutic glycoprotein drugs for treating life-threatening conditions such as metastatic breast cancer and non-Hodgkin's lymphoma. These glycoprotein pharmaceuticals contain complex oligosaccharide moieties whose presence, absence, and profile can have significant impact on therapeutic efficacy, pharmacokinetics, immunogenicity, folding, and stability of the biologic drug. For example, certain glycan structures are known to cause aggregation and decrease drug efficacy. In turn, the degree and types of glycosylation are strong functions of the expression system and the cell culture conditions used in the production of the antibody. As a result, throughout fermentation, purification, and formulation the types and relative amounts of N-linked glycosylation structures contained in a protein biologic are monitored to ensure that the drug product is consistent and stable.

A variety of approaches can be used to characterize glycoproteins and their glycan moieties. For example, HPLC followed by quadrupole time-of-flight mass spectrometry (LC/QTOF MS) can distinguish the number of glycan units attached to an intact protein, thus providing the ability to distinguish an active form from an inactive one (Figure 1). In addition, glycosylation sites can be elucidated by digesting the protein with trypsin and using LC/QTOF MS to separate and identify the resulting glycopeptides. Comparison of the masses of these peptides to those generated by a theoretical

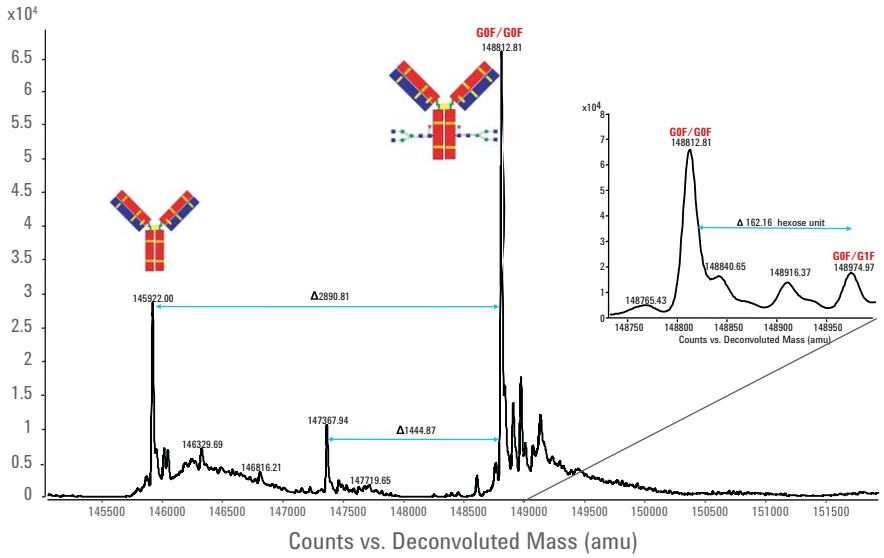


Figure 1: Deconvoluted Q-TOF MS spectrum of intact antibody with inset showing the expanded view of a small amount of the G1F form (addition of hexose unit to G0F) of the mAb. The data was generated on an Agilent 1260 Infinity HPLC-Chip/MS system coupled to a 6520 Accurate-Mass Q-TOF LC/MS platform, utilizing a 75 µm X 43 mm ZORBAX 300SB-C8 analytical column.

digestion of the desired glycan form of the protein can determine if the protein is properly glycosylated. Capillary electrophoresis (CE) coupled to QTOF MS can also be used for this purpose, while delivering excellent separation efficiency, short run times, and minimal sample/solvent consumption.

Analysis of the glycan moieties attached to a protein is most commonly done by enzymatic deglycosylation and hydrolysis (Figure 2). N-Glycosidase F (PNGase F), an amidase, is used to cleave asparagine-linked (N-linked) oligosaccharides from glycoproteins. Most commonly, the removed glycans are derivatized, labeled, and analyzed by fluorescence detection. Capillary electrophoresis with laser-induced fluorescence (LIF) or MS detection can then be used to identify and quantify glycans labeled with amino pyrene tri-sulfonic acid (APTS). HPLC with fluorescence detection (FLD) can also be used to separate and identify glycans labeled with 2-aminobenzamide (2-AB). The HPLC-FLD method often uses a hydrophilic interaction chromatography (HILIC) column or other separation media to separate the glycans. These techniques do not provide mass information, do not have the ability to separate all isomers, lack the sensitivity required to detect low-abundance glycoforms, and can take as long as two days to complete due to long enzymatic reactions and labeling steps.

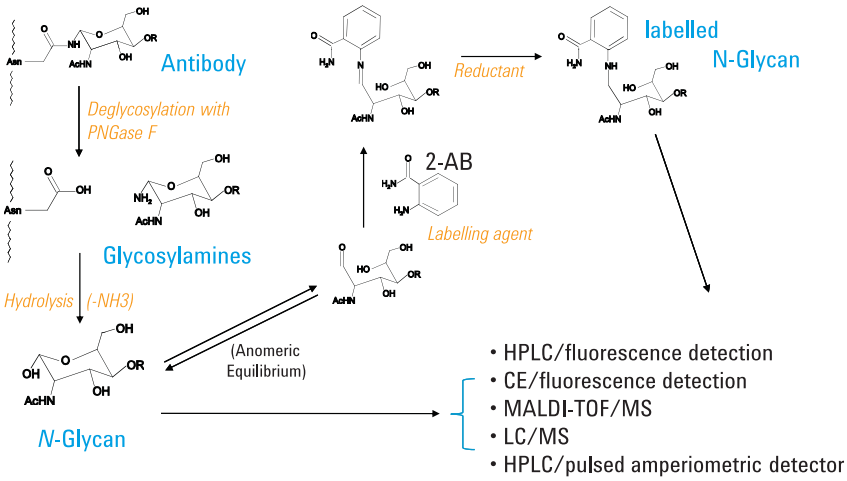


Figure 2: Glycan characterization workflows

Unlabeled glycans can also be analyzed by various techniques. Traditionally, high-performance anion exchange chromatography with pulsed ampiometric detection (HPAEC-PAD) is used to separate unlabeled glycans based on hydroxyl group interaction with the stationary phase. HPAEC-PAD is also used to determine sialic acid content, which must be measured to ensure the product is safe and for batch-to-batch reproducibility. While this method removes the potential issues with labeling efficiency, it still provides fingerprint-like information and requires standards for glycan identification. Gas chromatography (GC) is also frequently used for monosaccharide compositional analysis because it is robust and has high resolution.

Analysis of unlabeled glycans can also be performed by reversed-phase (RP) LC/MS and matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry (MALDI-TOF/MS) in order to provide mass information about the glycans. However, these techniques still do not provide complete separation of potential isomers.

Agilent has developed a unique, chip-based LC/MS solution for the characterization of N-linked glycans that delivers results in 10 to 30 minutes, rather than two days, and is run in automated mode. The mAb-Glyco Chip kit performs deglycosylation using an on-line PNGase F reactor, concentrates glycans on a graphitized carbon enrichment column, and separates glycans on a graphitized carbon analytical column. Porous graphitized carbon has the ability to separate all potential glycan structures, including isomers, which are often present (Figure 3). All of these steps are performed on a single nano-LC chip that interfaces with an Agilent mass spectrometer. The analyst simply dilutes and centrifuges the mAb sample and then places it in the autosampler. The rest of the analysis takes place on the chip. MassHunter software and a provided glycan database are then used to identify and determine the glycan structures, quantities, and ratios. Agilent capillary and nano pumps, the Chip-Cube interface and an Agilent MS are required. The chip reproducibly separates and quantitates all of the common N-linked glycans, making it a useful Process Analytical Technology (PAT) method for assuring batch-to-batch reproducibility of the glycan composition of a therapeutic protein (Figure 4). Agilent also provides a porous graphitized carbon chip without the enzyme reactor to allow for analysis of samples that are deglycosylated off-line.

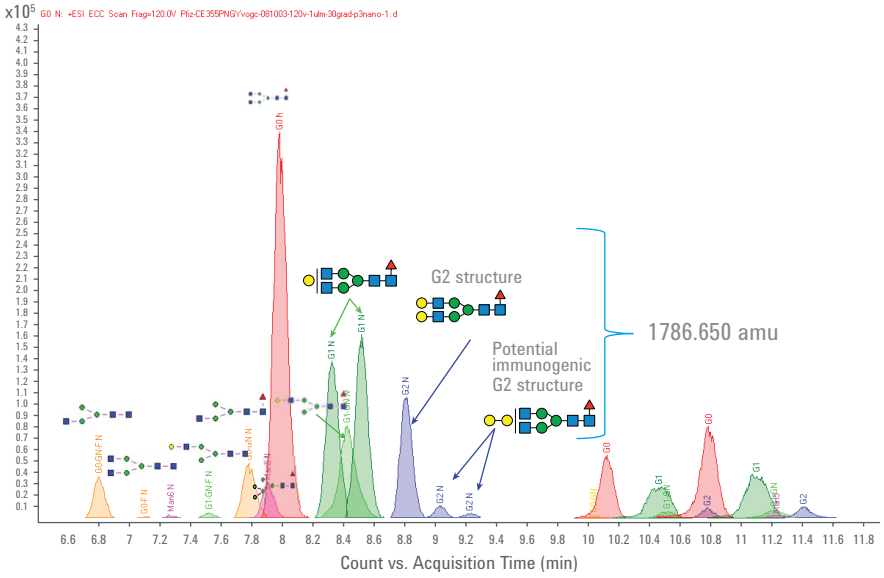
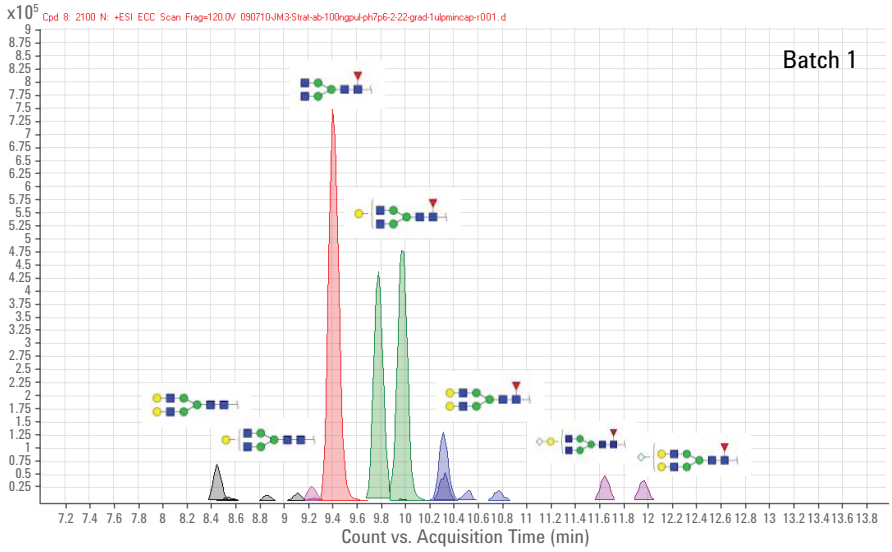


Figure 3. Separation of monoclonal antibody glycans including isomers using the mAb-Glyco Chip. Three G2 peaks are shown in blue, between 8.7 and 9.3 minutes. The first G2 peak represents the common G2 structure where each arm contains 1 galactose sugar, and the smaller G2 peaks in blue are alpha-linked galactose structures which are known to be potentially immunogenic.

When a three-dimensional glycan structure is required to fully understand the properties of the protein drug, nuclear magnetic resonance (NMR) has been used to reveal essential features of the glycoprotein and glycan structures. The two-dimensional (2D) NMR made possible by the two naturally active nuclei (^{13}C and ^{1}H) can provide unambiguous structural assignment and specific detection of glycan spatial arrangements. NMR is often indispensable in the determination of unusual or previously undescribed sugars, such as those present in bacterial glycoconjugates, and ^{1}H NMR has been used as a powerful tool for structural characterization of free glycans.



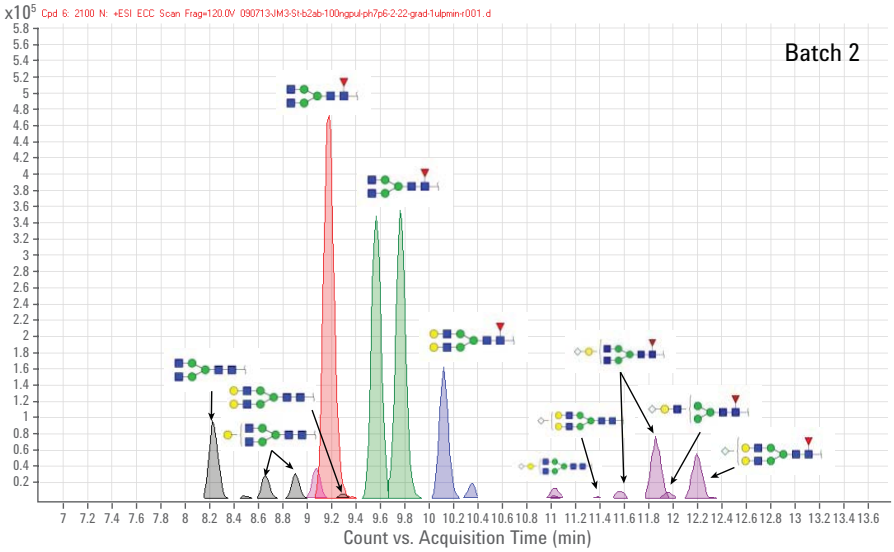


Figure 4. Comparison of the glycan profiles of two monoclonal antibody production batches, analyzed using the mAb-Glyco Chip. Batch 2 has a higher percentage and greater variety of sialylated glycan forms.

Glycan profiling techniques summary

Solution	Technique Description	Benefits	Limitations
CE(CZE)/LIF	PNGase F removal of N-linked glycans, fluorescence labelling with APTS, Ionic state separation (CZE), laser-induced fluorescence excitation and detection, identification based on position across ladder standard	High-resolution separation methods capable of differentiating most positional isomers	Time consuming enzymatic and labelling steps
			Requires use of glycan ladder standard for identification
			Potential labelling efficiency issues
			No mass identification, reliance on ladder standard
CE(CZE)/MS	PNGase F removal of N-linked glycans, fluorescence labelling with APTS, ionic state separation (CZE), mass spectrometry detection, identification based on position across ladder standard and molecular weight database	High-resolution separation methods capable of differentiating most positional isomers	Time-consuming enzymatic and labelling steps
		Provides accurate-mass information of glycans	Potential labelling efficiency issues
HILIC-LC/FLD	PNGase removal of N-linked glycans, fluorescence labelling with 2-AB, separation based on hydrophobicity across a declining organic concentration (HILIC), fluorescence label detection, identification based on retention-time database	High resolution separation methods capable of differentiating most positional isomers	Time-consuming enzymatic and labelling steps
		Fractions can be collected for further analysis	Potential labelling efficiency issues
			No mass identification, reliance on retention times
HPAEC-PAD	PNGase F removal of N-linked glycans, fluorescence labelling, separation based on charge state interaction with anion exchange media across increasing salt concentrations, pulse amperometric detection (chemical conversion and electrode interaction), identification based on retention-time database	Unlabeled analysis, no labelling efficiency issues	Time-consuming enzymatic removal step
		More rapid analysis than labelled technique, no labelling step	Co-elution of glycans is possible
		Commonly used as a "fingerprint" for batch analysis, not used for quantitation or identification	Lacks resolution and sensitivity of CE and LC/MS techniques
GC or GC/MS	Free monosaccharide compositional analysis (not bound to protein), separation from proteins, UV detection or MS detection, identification based on either retention time and/or molecular weight	Separation removes proteins from monossacharides prior to detection, less inteference	Limited applicability for oligosaccharides
		Commonly used to monitor and identify monosaccharides in cell culture and ensure removal during purification	Lacks resolution and sensitivity of CE and LC/MS techniques
MALDI-TOF/MS	PNGase F removal of N-glycans, spotting free glycans on MALDI plate with ionization matrix, matrix-assisted laser-desorption/ionization, TOF/QTOF MS detection, identification based on molecular weight database	Does not require time-consuming labelling step	Time-consuming enzymatic removal of glycans
		Provides accurate-mass information of glycan structures, making exact mass identification possible	Possible matrix-related suppression causing a decrease in sensitivity, making low-level glycans less likely to be detected

Solution	Technique Description	Benefits	Limitations
Porous Graphitized Carbon (PGC) Chip-LC TOF/QTOF MS	PNGase F removal of N-glycans, enrichment of glycans on porous graphitized carbon (PGC) enrichment column, separation of glycans on PGC nanocolumn across an organic gradient, ESI-TOF/QTOF MS detection, identification based on retention-time and molecular weight database	Does not require time-consuming labelling step	Time-consuming enzymatic removal of glycans
		Provides accurate-mass information of glycans	Deglycosylated protein is immobilized on PGC enrichment column, cannot be analyzed
		High resolving power and unique selectivity for oligosaccharide moieties provides excellent separation of glycan forms, including more isomers, than other techniques	
MAb Glyco Chip-LC TOF/QTOF MS	Online PNGase F removal of N-glycans, enrichment of glycans on porous graphitized carbon (PGC) enrichment column, separation of glycans on PGC nanocolumn across an organic gradient, ESI-TOF/QTOF MS detection, identification based on retention-time and molecular weight database	Enables on-line enzymatic glycan removal, decreases reaction time to minutes rather than hours or overnight	Deglycosylated protein is immobilized on PGC enrichment column, cannot be analyzed
		Does not require time-consuming labelling step	
		Provides accurate-mass information of glycan structures, making exact mass identification possible	
		High resolving power and unique selectivity for oligosaccharide moieties provides excellent separation of glycan forms, including more isomers, than any other technique	
Two-dimensional (2D) NMR	Provides three-dimensional glycan structure for basic protein research or structural confirmation, requires probe and high-field magnet for analysis of large proteins	Provides unambiguous structural assignment and specific detection of glycan spatial arrangements	Usually requires expert in NMR to perform analysis
		Determination of unusual or previously undescribed sugars	Difficult for whole protein structural analysis, commonly used for specific protein region or site analysis (e.g., Fab or Fc region of a mAb)

Agilent applications literature

Publication #	Title
5990-6924EN	The Agilent mAb-Glyco Chip Kit for rapid and fully automated characterization of N-linked glycans from monoclonal antibodies
5990-5190EN	Glycopeptide and glycan analysis of monoclonal antibodies using a microfluidic-based HPLC-Chip coupled to an Agilent Accurate-Mass Q-TOF LC/MS
5990-7138EN	Glycopeptide Analysis of Antibodies by Capillary Electrophoresis and Q-TOF Mass Spectrometry
5990-5155EN	Success Story at Boston University School of Medicine – Custom HPLC-Chip enables new research in glycan expression

PEPTIDE MAPPING

5

Peptide mapping is the most widely used identity test for proteins, particularly those produced by recombinant means. It most commonly involves enzymatic digestion (usually using trypsin) of a protein to produce peptide fragments, followed by separation and identification of the fragments in a reproducible manner. Peptide mapping is a very powerful method that has become an invaluable tool to the biotechnology sector, allowing the detection and monitoring of single amino acid changes, oxidation, deamidation, and other degradation products. It also enables the direct detection of common monoclonal antibody variants such as N-terminal cyclization, C-terminal lysine processing, and N-glycosylation, as well as unexpected variations such as a translated intron.

A peptide map is a fingerprint of a protein and the end product of several processes that provide a comprehensive understanding of the protein being analyzed. It involves four major steps: isolation and purification of the protein; selective cleavage of the peptide bonds; chromatographic separation of the peptides; and analysis and identification of the peptides. A test sample is digested and assayed in parallel with a reference standard or a reference material. A map should contain enough peptides to be meaningful; it should not only provide a positive identification of the protein, but should also maximize coverage of the complete peptide sequence.

Peptide mapping utilizing reversed-phase HPLC separation of proteolytic peptides combined with mass spectrometry has become the method of choice for establishing the identity of a recombinant protein, a crucial requirement for the release of the product. The stability of a protein biologic is also an important aspect of characterization, requiring long-term monitoring at normal storage conditions throughout the shelf-life of a product for modifications such as oxidation, reduction, glycosylation, and truncation.

Many peptide separations are performed on electrospray ionization (ESI) LC/MS instruments due to the convenience of LC coupling and better quality of tandem mass spectra for confident protein identification. A QTOF instrument often gives more structural information, especially for larger peptides, due to its high resolving power and mass accuracy. Analytical- and capillary-scale HPLC systems and columns are commonly used for peptide mapping. In sample-limited situations and for mAb subunits or smaller recombinant proteins, the Agilent microfluidic chip-based LC/QTOF MS system provides highly accurate peptide maps very quickly, while consuming very little sample. This is especially valuable when mapping minor impurities or components.

Proteins and their potential modifications are identified from their LC/MS peptide patterns (maps) by first using molecular feature extraction software to assign a true mass to each peptide and thus confirm its structure. Software is also used to generate a theoretical digestion list of peptide masses, based on the known sequence of the protein. Matching the experimental data to the theoretical peptide pattern generates a confirmed identity of the protein. A peptide profile may consist of over 60 peaks representing individual peptides and their derivatives, requiring a very powerful separation method. Superficially-porous reversed-phase particle technology, with an

optimal pore size, greatly improves the efficiency and speed of these separations, providing 100% sequence coverage of both the light and heavy chains of a monoclonal antibody, while reducing the separation time more than two-fold relative to conventional packing materials (Figure 1). Totally porous reversed-phase particles are also used to generate high-coverage peptide maps, often using longer columns with smaller particles requiring higher pressure UHPLC systems like the Agilent 1290 Infinity LC system.

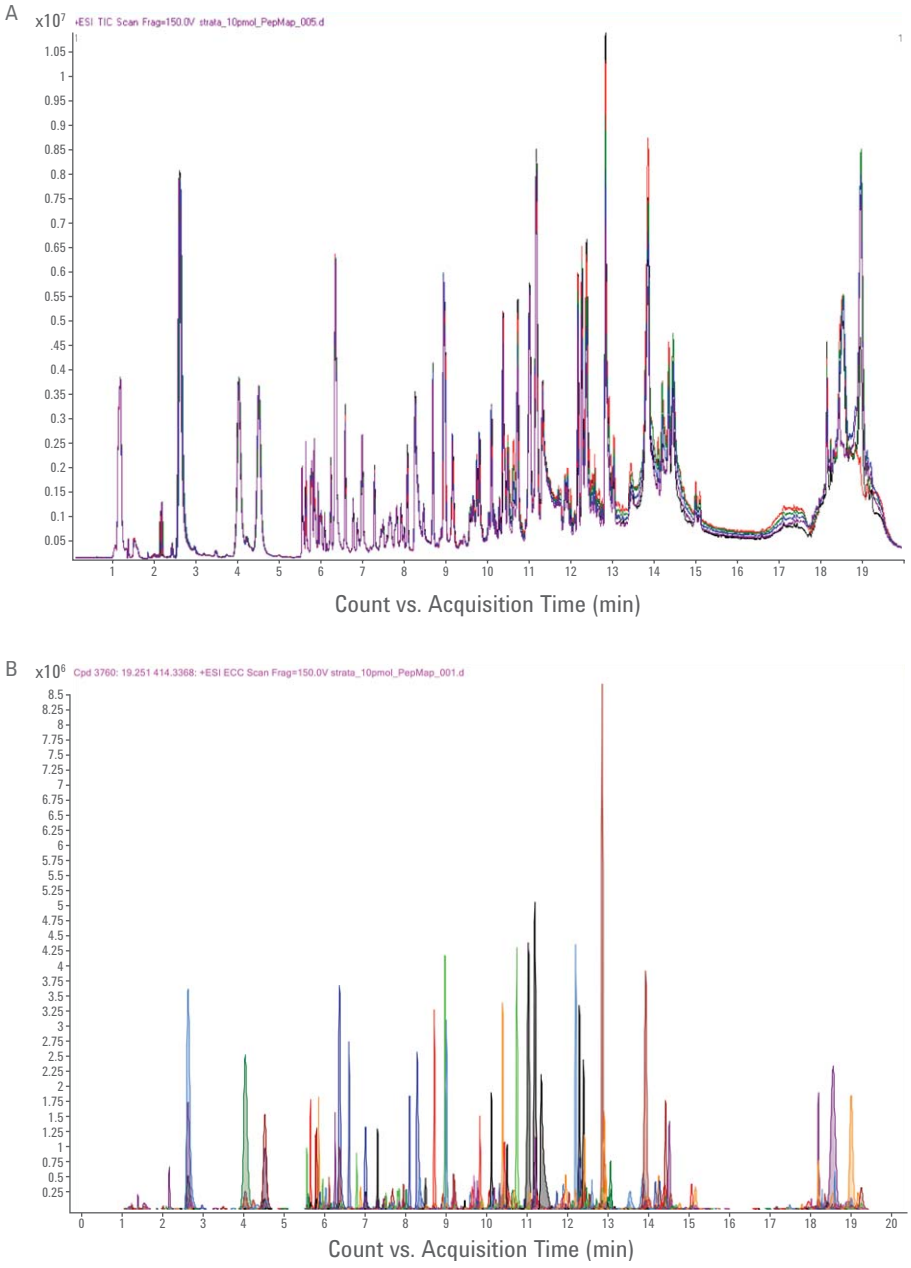


Figure 1: (A) Total Ion Current Chromatogram (TIC) of a tryptic peptide digest of a monoclonal antibody preparation separated using a Poroshell 120 reverse-phase column. Data were collected on an Agilent 6530 Accurate-Mass Q-TOF LC/MS equipped with an ESI source. (B) Extracted ion chromatogram (EIC) of the resulting peptides from the mAb with 100% sequence coverage.

Once the identities of the peptides in the map have been confirmed, the HPLC separation pattern itself becomes a fingerprint used to identify the proper form of the protein, using UV as the detector (Figure 2). This rapid method is then useful for process monitoring as well as QA/QC. However, for UV-based methods it is of utmost importance to obtain as much separation efficiency as possible in order to resolve all peptide peaks present in the mixture.

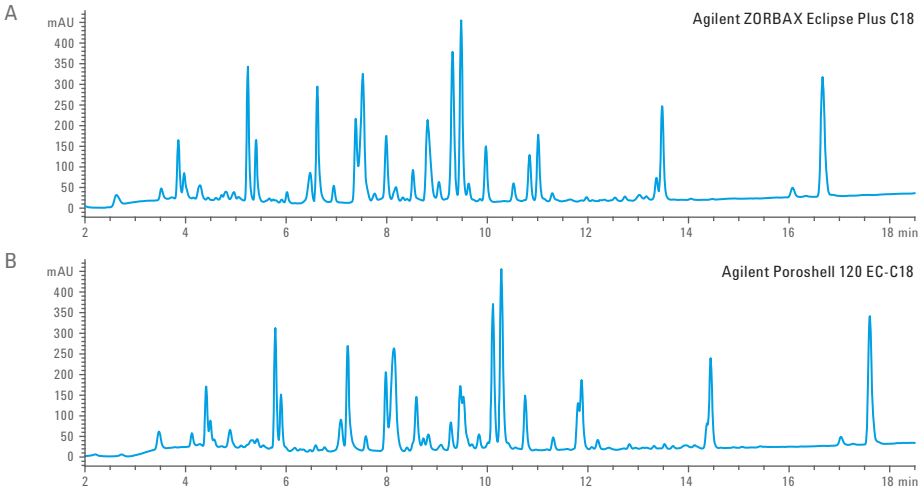


Figure 2: HPLC/UV peptide map of a 26-kDa recombinant protein digest on the Agilent 1260 Bio-inert LC system and columns. (A) Agilent ZORBAX Eclipse Plus C18 3.0 mm × 100 mm, 1.8 µm. (B) Agilent Poroshell 120 EC C18 3.0 mm × 150 mm, 2.7 µm at 0.6 mL/min flow rate with the corresponding gradient.

Recently, CE has been applied to peptide analysis because it is an excellent complementary tool to HPLC, due to its different separation mechanisms. Capillary zone electrophoresis (CZE) is the CE mode that is most commonly used for peptide analysis. The benefits that make CE an attractive technique include the speed of analysis, minimal sample consumption, lack of organic waste, and its general versatility. The versatility results from the ease of changing the separation mode and affecting selectivity simply by altering buffer composition. CE peptide mapping is not considered a replacement for LC/UV and LC/MS peptide mapping, but can be used to obtain complementary or confirmatory data.

Peptide mapping techniques summary

Techniques	Technique Description	Benefits	Limitations
RP-LC/UV	HPLC system, separation based on hydrophobicity across an increasing organic gradient by C18 reversed-phase HPLC column, UV detection (for larger proteins with large resulting peptide fragments, larger pore sizes are used, 300 Å for example)	Data output commonly called a protein "fingerprint" that can be monitored throughout QA/QC	Does not provide mass information for peptides, not possible to determine sequence information or coverage
		Robust and reliable methods, lower cost equipment	Separation of majority of peptide requires long HPLC columns and very long gradients, time consuming
RP-LC/MS	HPLC system, separation based on hydrophobicity across an increasing organic gradient by C18 reversed-phase HPLC column, MS detection (for larger proteins with large resulting peptide fragments, larger pore sizes are used, 300 Å for example)	Provides peptide mass information that can be correlated back to a protein sequence	Not commonly used in QA/QC, requires expert user for LC/MS use and data analysis
		Enables verification of protein sequence and structure, allows for site identification of post-translational modifications such as glycosylation and deamidation	
		MS detection allows for use of shorter columns and gradients, less separation required	
CE (CZE)/DAD or CE(CZE)/MS	Capillary electrophoresis system used in CZE mode (label free), capillaries, buffers, separation based on peptide charges and size, detection by diode array detection (DAD) or MS	Orthogonal technique to LC/UV or LC/MS that may provide different selectivity and resolution	Possible reproducibility issues with separations
		CE/MS will also provide mass information of peptide for sequence confirmation and coverage	CE interface with mass spectrometry possible but does not provide the same sensitivity as LC/MS

Agilent applications literature

Publication #	Title
5990-4712EN	Rapid Peptide Mapping Method with High Resolution Using a Sub 2 µm Column
5990-4587EN	Peptide Mapping of a Monoclonal Antibody using a Microfluidic-based HPLC-Chip coupled to an Agilent Accurate-Mass Q-TOF LC/MS
5990-5096EN	Faster, More Accurate Characterization of Proteins and Peptides with Agilent MassHunter BioConfirm Software
5990-6313EN	Increased peak capacity for peptide analysis with the Agilent 1290 Infinity LC System
5990-6192EN	Physicochemical characterization of a therapeutic protein by peptide mapping, SEC and IEX using the Agilent 1260 Infinity Bio-inert Quaternary LC system
5990-4031EN	Tryptic digest analysis using the Agilent 1290 Infinity LC System
5990-7631EN	An orthogonal view of peptide mapping – analysis of bovine serum albumin digest using CE and quadrupole time-of-flight mass spectrometry
5990-8244EN	Analyze MAb and BSA digests by UHPLC with UV detection and Agilent ZORBAX RRHD 300SB-C18

CHARGE ISOFORMS

6

During production and purification processes, proteins can exhibit changes in charge heterogeneity. These changes may not only impact stability but also activity, and they can cause immunologically adverse reactions. Hence, the analysis of charge isoforms in therapeutic protein preparations is key during the development and manufacturing processes. Charge variant characterization is commonly performed using isoelectric focusing or ion exchange chromatography.

Isoelectric focusing (IEF) separates proteins by their isoelectric point (pI) and is routinely used to profile the charge isoforms of recombinant proteins. Compared to conventional slab gel IEF, IEF performed on a capillary electrophoresis system (cIEF) offers higher resolution, speed, quantitation, and automation capabilities (Figure 1).

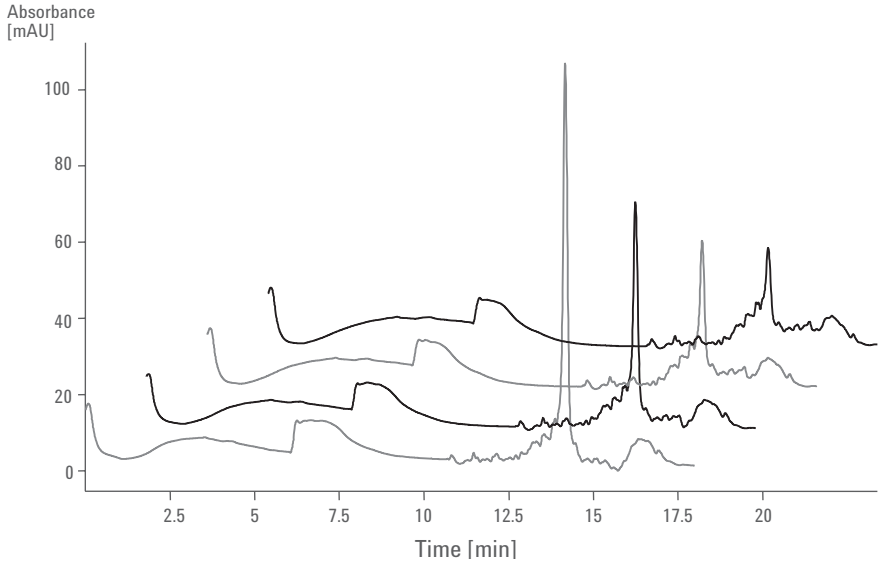


Figure 1: cIEF separation using the Agilent 7100 capillary electrophoresis system of carbonic anhydrase IIa at 500, 250, 125 and 62.5 mg/L, showing the linearity of the peak area over a broad range of concentrations.

Off-gel, in-solution isoelectric focusing methods can also be used to fractionate protein mixtures based on immobilized pH gradients. In contrast to gel-based techniques, samples can be easily recovered in solution for downstream applications such as size analysis or LC/MS. The OFFGEL system from Agilent allows isoelectric point (pI)-based fractionation of peptides and denatured as well as native proteins down to a resolution of 0.1 pH units. This technique yields 80% of the fractions in the liquid phase and 20% remains in the gel strip, making it an exploratory technique, not a quantitative one.

Ion exchange chromatography is a very useful tool for the characterization of protein charge variants, as it provides good separation of proteins with similar isoelectric points, and it has the advantage over isoelectric focusing of being able to detect differences in

the surface charge distribution of protein molecules. The separations are usually done using aqueous buffer systems and a salt gradient, but pH gradients are also used, and are sometimes combined with salt gradients in a technique called chromatofocusing.

Isoform characterization is particularly important in the production of mAbs, and analytical weak cation exchange chromatography is often used to determine the acidic and basic charge isoforms that are present after manufacturing and purification. Typically, charge isoform characterization involves calculation of the percent area of acidic forms, which lie to the left of the main peak, and the percent of basic forms, which lie to the right of the main peak (Figure 2). These isoforms are commonly deamidation and glycosylation products. Each peak from such a separation is often collected and analyzed on a mass spectrometer for verification. Once an elution profile is put into a standard operating procedure (SOP), the total number of peaks, their retention time and the percentage of acidic and basic isoforms are monitored throughout the production process. Monoclonal antibodies and other recombinant proteins can be susceptible to interactions with metallic surfaces. Agilent has developed the 1260 Infinity Bio-inert LC system with a completely metal-free flow path, eliminating the potential for these interactions and the need for passivation procedures. The system uses metal-cladded PEEK (inside) capillaries, allowing it to operate at 600 bar and at high salt and pH concentrations, while still being able to run reversed-phase methods in organic solvents.

Columns:	Agilent Bio MAb, NP10 4.6x250 mm
Mobile phase:	A, 10 mM phosphate, pH 7.5 B,A + 0.1M NaCl
Gradient:	A) 15 to 75%B in 30 min B) 15 to 65%B in 30 min C) 15 to 55%B in 30 min D) 15 to 47.5%B in 30 min E) 15 to 40%B in 30 min
Flow rate:	0.8 mL/min
Sample:	Monoclonal Antibody
Injection:	10 µL (1.5 mg/mL)
Temperature:	25°C
Detection:	UV 214 nm

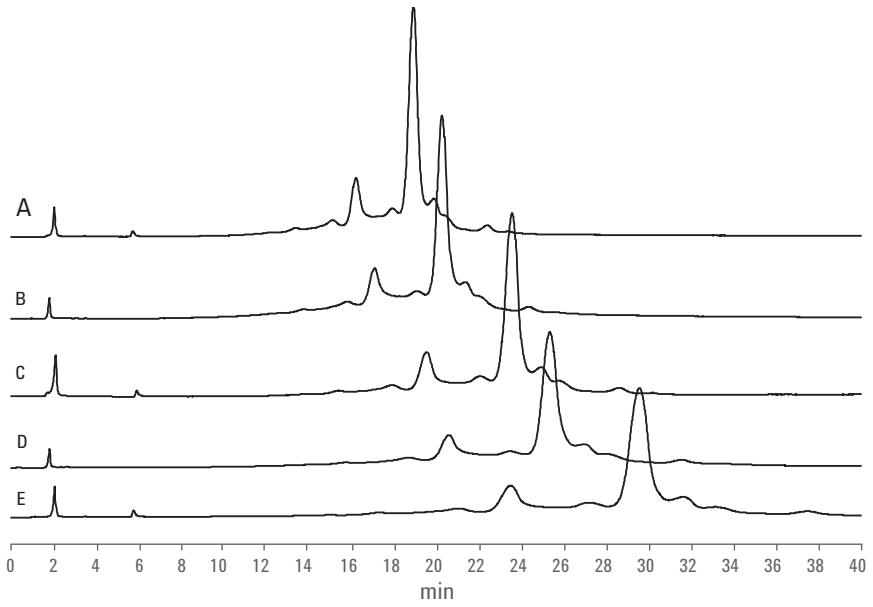


Figure 2: Weak ion exchange chromatographic characterization of the charge isoforms of a monoclonal antibody preparation using the Agilent Bio mAb NP10 column.

Charge isoform analysis techniques summary

Techniques	Technique Description	Benefits	Limitations
Gel Isoelectric Focusing (IEF)	Acrylamide gel matrix co-polymerized with a pH gradient, reagents, proteins migrate through the pH gradient in the gel based on their charge state	Ability to visually separate and identify proteins based on their charge states and pI	Low resolution for separating charge isoforms of a purified protein, better results with protein mixtures
		Samples can be collected for further analysis by cutting out gel bands	Non-quantitative
Capillary Isoeletric Focusing (cIEF)	Capillary electrophoresis system used in IEF mode, protein sample is mixed with ampholytes, mixture introduced into a capillary which is then subject to electrophoretic separation (charge is applied proteins or protein isoforms migrate to their respective pI and are then detected by either UV or LIF	Provides high-resolution separation of charge isoforms of recombinant proteins	Limited ability to fractionate and collect separated proteins for further analysis
		Faster separation than ion exchange chromatography	Not mass-spectrometry compatible
			Lacks selectivity options that ion exchange phases provide
Whole-Column Capillary Isoelectric Focusing (iCE)	Isoelectric focusing system where protein sample is mixed with ampholytes, mixture introduced into a capillary which is then subject to electrophoretic separation (charge is applied, proteins migrate to their respective pI), UV light is applied over the whole column and a digital camera records the protein migration, no migration step into a UV detector	Comparable resolution to traditional gel IEF but incorporates the advantages of a column-based separation technology	Not compatible with fluorescence detector (LIF), required for separating and analyzing fluorescently labeled proteins
		Automated sample introduction and reduced analysis time with no migration step to a UV detector required	Lack of compatibility with mass spectrometry
		Quantitative results	
OFFGEL electrophoresis	Electrophoresis system for pI-based fractionation of proteins or peptides in-gel or in OFFGEL mode, which means that 80% of migrated proteins are in liquid and 20% remains in the IEF gel	Can be used for screening for pH buffer conditions, supports cIEF and ion exchange method development	Non-quantitative fractionation of proteins and mixtures
		Commonly used to fractionate complex protein mixtures (e.g., plasma, cell lysates) prior to mass spectrometric analysis	No detection available, only used as an offline tool for fractionation
IEX-LC/UV	HPLC system with ion exchange columns (strong and weak cation or anion exchange packing material), interaction with the column particles based on total net charge of the protein, elution occurs across either an increasing salt or pH gradient or a combination of the two	Provides chromatographic separation of protein charge variants, various cation and anion exchange resins available offering many selectivity choices	Commonly long gradients with high buffer consumption
		Ability to fractionate and collect separated charge isoforms for further analysis (MS analysis possible after desalting, buffer exchange)	Lower resolution separations than cIEF
			Not mass-spectrometry compatible

Agilent applications literature

Publication #	Title
5990-6192EN	Physicochemical characterization of a therapeutic protein by peptide mapping, SEC and IEX using the Agilent 1260 Infinity Bio-inert Quaternary LC system
5989-9852EN	Capillary isoelectric focusing on the Agilent Capillary Electrophoresis system
5990-6521EN	Monitoring antibody charge variants using a combination of Agilent 3100 OFFGEL Fractionation by isoelectric point and high sensitivity protein detection with the Agilent 2100 Bioanalyzer

7 AGGREGATION

The amount, type, and size of the aggregates in protein biopharmaceuticals can have important consequences for their safety and efficacy. Several mechanisms are involved in the formation of protein aggregates, including non-covalent interactions between hydrophobic domains and disulfide bond formation. The presence of aggregates of any type in protein therapeutics is undesirable because of the concern that the aggregates may lead to an immunogenic reaction (small aggregates) or may cause adverse events on administration (particulates). Irreversible aggregation is a major concern with therapeutic proteins, particularly during long-term storage and shipping. Some of the methods used for aggregate analysis are ultracentrifugation, size-exclusion chromatography with light scattering, and native gel electrophoresis.

Native gel electrophoresis carried out near neutral pH can be used to study conformation and the self-association or aggregation of proteins. SDS-PAGE can be used to separate aggregates, but the technique is considered very time consuming and does not provide high-resolution separation of aggregates or potential impurities close in molecular weight. The Agilent 2100 Bioanalyzer and the High Sensitivity Protein 250 Assay make it possible to rapidly resolve the intact protein, aggregates, and lower molecular-weight impurities, such as non-glcosylated intact mAb and free light and heavy chains. This assay provides highly reproducible (%CV <6%) results and can easily be adopted for quality control throughout development and manufacturing (Figure 1).

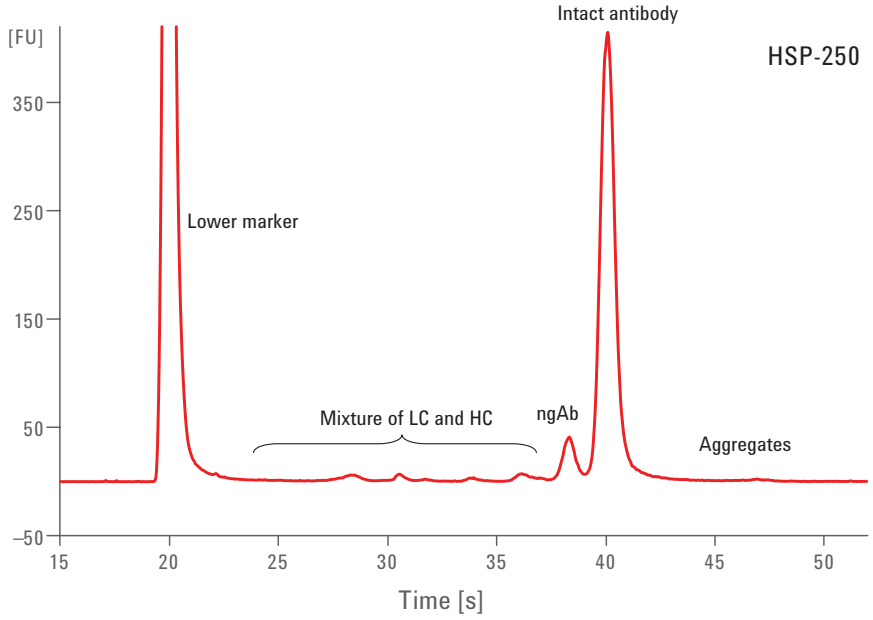
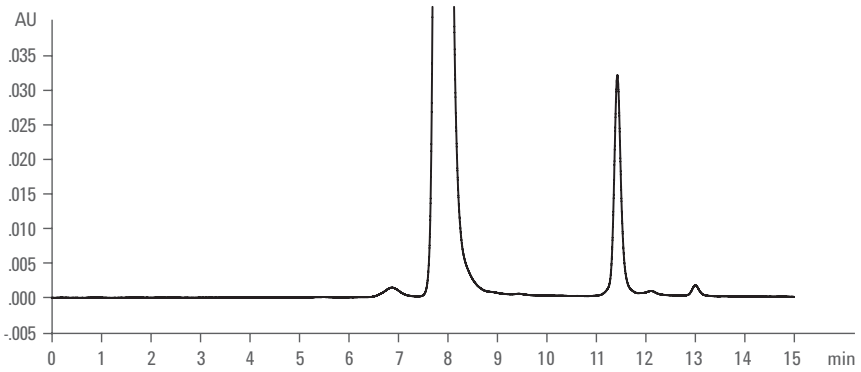


Figure 1: Analysis of an IgG2 preparation under non-reducing conditions; electropherogram generated with the Agilent 2100 Bioanalyzer and the High Sensitivity Protein 250 Assay. Abbreviations: LC: Light Chain; HC: Heavy Chain; ngAb: non Glycosolyated antibody.

Asymmetrical Field Flow Fractionation (AF4) is a matrix-free technique that provides broad dynamic range as an analytical method for the separation of mAb aggregates. Monomer and dimer peaks are well resolved, and separations are very rapid. Crude cell culture samples can be injected directly onto the FFF membrane, providing an analytical method for aggregate measurement that takes only 10 minutes. The advantage of AF4 is the ability to separate both soluble and colloidal components over a wide size range.

Sedimentation velocity can be particularly valuable for determining aggregates and homogeneity, changes in protein conformation, and comparing different, engineered variants of a protein. The signal from the light-scattering detector is directly proportional to the molecular mass of the protein multiplied by the concentration, so it is most commonly combined with size-exclusion chromatography to measure the molecular mass of each peak coming off the column.

High-performance liquid chromatography (HPLC) can be used to accurately characterize the sizes of the components of the biologic product, while offering the advantages of ease-of-use and a high level of automation. Size-exclusion chromatography (SEC) with either UV or light scattering detection has been a workhorse for detecting and quantifying protein aggregation and is a common QC/QA method. In SEC separations, large aggregates elute first, followed by the protein dimers, the monomers, and in some cases low molecular weight forms called clips (smaller degradation products). The linear response of these separations can span almost three orders of magnitude. Resolution is paramount, to assure detection of aggregrates and degradation products. These separations are commonly run on size-exclusion columns with 5-micron (µm) particles. However, smaller particle sizes can provide higher resolution and faster separations. Using a column with 3-µm particles, the resolution of the dimers from the active monomer present in a monoclonal antibody preparation is significantly better, allowing an accurate estimation of the percent contamination (Figure 2). Using smaller particles allows for separation at higher flow rates (1.0 mL/min) than traditional separations on 5-µm particle columns (0.5 mL/min), without a significant loss of resolution and efficiency.



Columns:	Bio SEC-3 300 Å, 7.8 x 300 mm
Mobile phase:	150 mM Phosphate, pH7
Flow rate:	1.0 mL/min
Temperature:	Ambient
Sample:	Monoclonal antibody (10 µL 5 mg/mL)

Figure 2: Size exclusion chromatographic separation of a monoclonal antibody preparation illustrating the resolution of the monomer (first peak), dimer (second), and buffer-related peaks (third peak), using the Agilent Bio SEC-3 300-Å column, which uses 3-µm particles.

Aggregation analysis techniques summary

Techniques	Technique Description	Benefits	Limitations
SDS-PAGE	Electrophoresis chamber, acrylamide gel, buffers and reagents, Coomassie Blue or silver stain (proteins will separate based on their size, smaller proteins or monomers will migrate further through the gel than dimers and aggregates)	Visual results, can see protein migration in gel	Qualitative results
		Protein bands can be extracted for further analysis	Low resolution, impurities similar in size migrate together
			Low throughput, very long separation times
Chip-based (Bioanalyzer) Protein Electrophoresis Assays	2100 Bioanalyzer system, High Sensitivity Protein 250 Kit (includes chip, fluorescence labelling buffer and all required reagents), electrophoretic separation combined with fluorescence detection	Quantitative results	Currently no online connection for real-time reactor analysis
		High-resolution separations	No sample collection for further analysis
		Ability to resolve low molecular weight impurities along with monomer, dimer, and larger aggregates	
		Short analysis time compared to SDS-PAGE	
SEC-LC/UV	HPLC system, size-exclusion HPLC column, separation is based on protein molecular radius and passage speed through packed porous particles, UV detection	Quantitative results	Limitation of separation window, not easy to resolve high molecular weight aggregates and low molecular weight impurities in the same separation
		High-resolution separations of dimers and aggregates	Does not provide mass information
		Short analysis time compared to SDS-PAGE	
		Simple method, commonly used in QA/QC to monitor aggregate ratios	
		Ability to collect fractions for further analysis	
SEC-LC/Light Scattering	HPLC system, size-exclusion HPLC column, separation is based on protein molecular radius and passage speed through packed porous particles, UV and light scattering detection in tandem, light scattering provides relative mass information	Quantitative results	Limitation of separation window, not easy to resolve high molecular weight aggregates and low molecular weight impurities in the same separation
		High-resolution separations of dimers and aggregates	
		Short analysis time compared to SDS-PAGE	
		Simple method, commonly used in QA/QC to monitor aggregate ratios	
		Ability to collect fractions for further analysis	
		Addition of light scattering provides relative mass information and identifies column particle shedding	

Agilent applications literature

Publication #	Title
5990-5283EN	Protein analysis with the Agilent 2100 Bioanalyzer – An overview of the protein kit portfolio
5990-6416EN	Characterization of monoclonal antibodies on the Agilent 1260 Infinity Bio-inert Quaternary LC by Size Exclusion Chromatography using the Agilent BioSEC columns
SI-02395	Static Light Scattering Analysis of Globular Proteins with ProSEC 300S Columns

OXIDATION

8

Oxidative modification of enzymes has been shown to inhibit enzymatic activity. Similarly, oxidation of recombinant proteins can cause a loss of function. Oxidation is therefore a serious concern during the development and manufacturing of a protein therapeutic, as it can alter a protein drug's biological activity, half-life, and immunogenicity. Amino acids within a therapeutic protein are susceptible to oxidation during processing and storage, and these oxidations must be monitored continuously. Some of the methionine residues in a protein can be oxidized to methionine sulfoxide or even methionine sulfone. Methionine oxidation can lead to inactivity, aggregation, and increased immunogenicity. Deamidation of asparagine residue to form aspartic acid and iso-aspartic acid is another cause of protein degradation, particularly during long-term storage. While glutamine residues can also be deamidated, the rate of this reaction is one hundred times slower than that for asparagine, and it is rarely detected in recombinant proteins. Oxidation of tryptophan and cysteine can also occur.

Methionine oxidation can be detected by mass spectrometry or UV detection of peptide fragments. Intact proteins containing oxidized methionines can also be separated from their non-oxidized counterparts using ion-exchange or hydrophobic interaction chromatography (HIC), which uses a reverse-salt gradient, from high to low salt concentrations. Asparagine residues prone to deamidation can be identified through peptide mapping of the recombinant protein.

One method for determining the susceptibility of a protein drug target to oxidation is the use of forced oxidation studies. In these studies, the protein is exposed to a mild oxidant, often hydrogen peroxide or t-butyl hydroperoxide (t-BHP), in order to investigate the possible sites that are susceptible to oxidation and the effect on the activity of the protein biologic due to these oxidations. In one such study, a monoclonal antibody oxidized with hydrogen peroxide was analyzed using an Agilent HPLC-Chip in reverse-phase mode coupled with an Accurate-Mass Q-TOF LC/MS; BioConfirm software was used to find changes in the mAb's mass (Figure 1). Subsequent analysis of the peptide fragments generated by tryptic digestion in MS and MS/MS mode identified the sites of oxidative modification (Figures 2 and 3). More importantly, the degree and rate of oxidation susceptibility needs to be measured, and in this respect LC/MS is a valuable tool. Excellent chromatographic resolution and highly accurate peptide mass determination provided by the Agilent 6520 Accurate-Mass Q-TOF LC/MS enables accurate assignment of oxidized peptide peaks to the mAb sequence under study. The MS/MS analyses help increase confidence in the peptide sequence assignments by mapping the exact location of the oxidative modification in peptides.

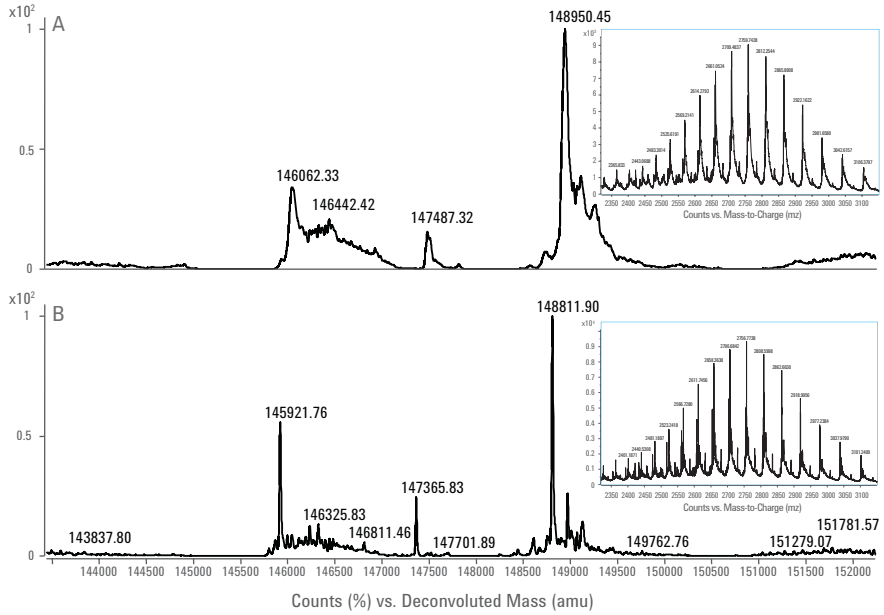


Figure 1: Deconvoluted spectrum of whole intact antibody (with inset showing the mass spectrum of intact antibody). (A) Oxidatively modified mAb and (B) unmodified mAb.

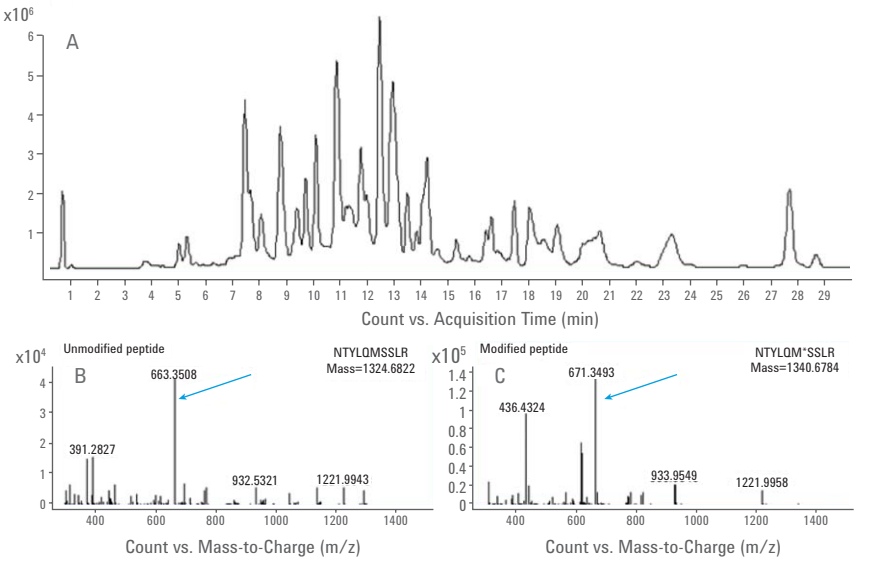


Figure 2: (A) Total ion chromatogram (TIC) of trypsin-digested mAb after oxidative modification with H_2O_2 on a C18 HPLC-Chip. Arrows within the representative mass spectra show the (B) unmodified and (C) modified peptide.

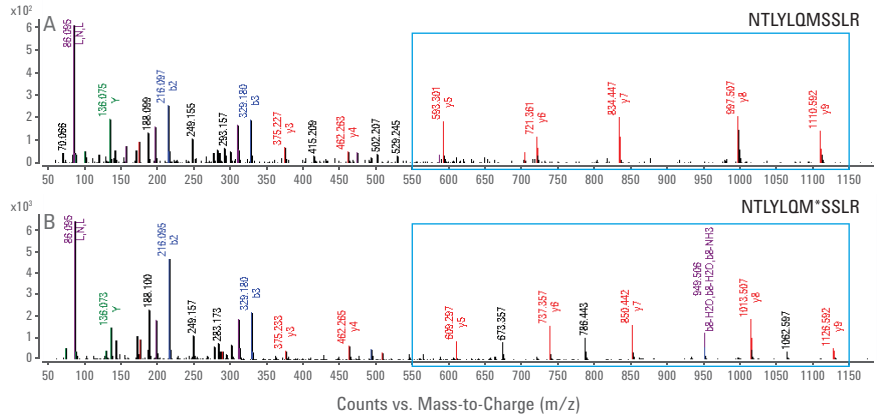


Figure 3: (A) Unmodified and (B) modified representative MS/MS spectra of peptides; these assignments were done using a feature of the BioConfirm software package. The blue box shows the difference of ~16Da for y5 to y9 ions between the unmodified and modified peptide containing a methionine that has been oxidized.

Oxidation analysis techniques summary

Techniques	Technique Description	Benefits	Limitations
HIC-LC/UV	HPLC system with Hydorphobic Interaction Chromatography (HIC) column, separation of oxidized from unoxidized proteins based on their hydrophobicity across a declining salt gradient, UV detection	Provides ratio information for oxidized and unoxidized protein product	Not able to provide mass verification or oxidation site information
		Rapid and simple QA/QC method to monitor for oxidation	Less applicable for QA/QC than HIC-LC/UV due to method complexity
Intact LC/MS	HPLC system with reversed-phase (RP) column, separation of oxidized from unoxidized proteins based on their hydrophobicity and increasing organic gradient, MS detection	Accurate mass identification of oxidized protein states	Not able to provide site of oxidation, to do so requires enzymatic digestion
		Commonly used to confirm and verify oxidation issues identified in QA/QC	Less applicable for QA/QC than HIC-LC/UV due to method complexity
Enzymatic digestion + LC/MS	Enzymatic digestion, most commonly using trypsin, of protein into peptide fragments, HPLC system with reversed-phase (RP) column, separation of peptides based on their hydrophobicity and increasing organic gradient, MS detection of oxidized peptides	Combined with intact LC/MS analysis, enables identification of oxidation forms and after digestion the site of oxidation, MS/MS confirmation possible	Not commonly used in QA/QC due to method complexity and long analysis time
		Commonly used to confirm and verify oxidation issues identified in QA/QC	

Agilent applications literature

Publication #	Title
5990-8768EN	Identification of Oxidation Sites on a Monoclonal Antibody Using an Agilent 1260 Infinity HPLC-Chip/MS System Coupled to an Accurate-Mass 6520 Q-TOF LC/MS
5990-8769EN	Quantitation of Oxidation Sites on a Monoclonal Antibody Using an Agilent 1260 Infinity HPLC-Chip/MS System Coupled to an Accurate-Mass 6520 Q-TOF LC/MS

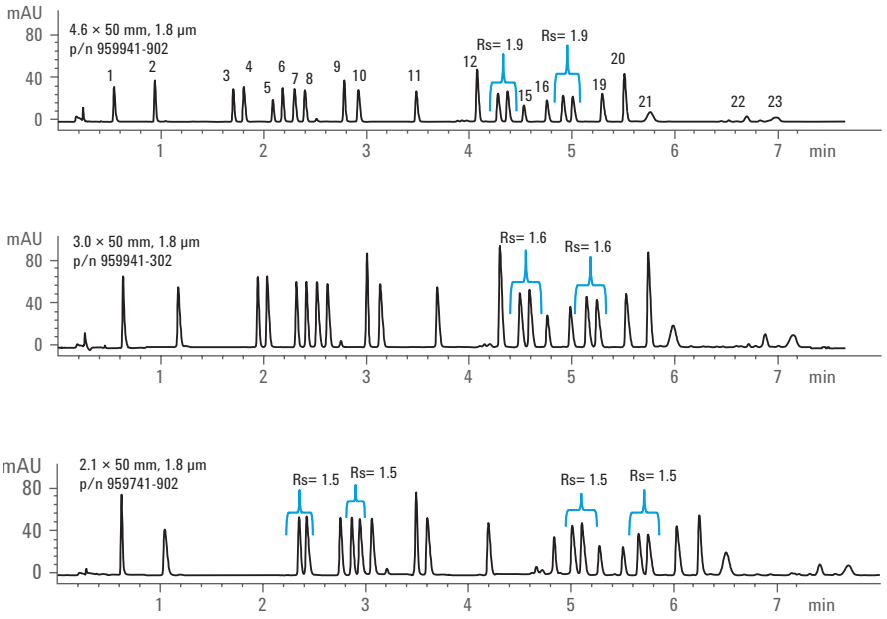
9 AMINO ACID ANALYSIS

Each protein or peptide has a unique amino acid sequence and, therefore, amino acid composition. The procedure of determining the amino acid composition of a protein or peptide can consequently be used as an identification test. Amino acid analysis is used from drug discovery through manufacturing to demonstrate comparability and consistency between batches. It is also often used as a decision support for selecting proteases for protein fragmentation. Amino acid analysis is also a suitable tool for precise determination of protein quantities.

Amino acid analysis involves four basic steps, starting with acid hydrolysis of the protein to individual constituent amino acids. The amino acids are then labeled with a detectable UV-absorbing or fluorescent marker, and the derivatized amino acids are separated by chromatography. Lastly, determination of the relative amounts of each amino acid type is done based on the intensity of the detectable marker.

An ideal, quantitative amino acid analysis combines speed and sensitivity with reliability of both the derivatization reaction and the analytical technique. These goals are achieved with automated, online derivatization using o-phthalaldehyde (OPA) for primary amino acids and 9-fluorenylmethyl chloroformate (FMOC) for secondary amino acids; the automated derivatization is then integrated with rugged HPLC analysis. The complete procedure is rapid, accurate, sensitive, and reproducible using the Agilent 1260 Infinity or 1290 Infinity LC system and Agilent ZORBAX Eclipse Plus C18 stationary phase columns.

Combining OPA and FMOC chemistries enables fast pre-column derivatization of amino acids (AA) for chromatographic analysis. The primary amino acids are reacted first with OPA using 3-mercaptopropionic acid (3-MPA). The secondary amino acids do not react with the OPA, but are then derivatized using FMOC. The derivatization process is fast and is easily automated using an Agilent autosampler. The automated procedure provides a high degree of reproducibility. Total analysis from injection to injection can be achieved in as little as 14 minutes (10-minute analysis time) on the 50-mm Agilent ZORBAX Eclipse Plus C18, 1.8-µm columns (Figure 1). This method provides good retention of the first two eluting amino acids, aspartic and glutamic acids, and high resolution of several closely eluting amino acid pairs, depending on the column configuration used.



1. Aspartic acid
2. Glutamic acid
3. Asparagine
4. Serine
5. Glutamine
6. Histidine
7. Glycine
8. Threonine
9. Arginine
10. Alanine
11. Tyrosine
12. Cysteine
13. Valine
14. Methionine
15. Norvaline
16. Tryptophan
17. Phenylalanine
18. Isoleucine
19. Leucine
20. Lysine
21. Hydroxyproline
22. Sarcosine
23. Proline

Figure 1. Amino acid analysis using 50-mm Agilent ZORBAX Eclipse Plus C18, 1.8-µm columns.

Amino acid analysis techniques summary

Techniques	Technique Description	Benefits	Limitations
OPA derivitization + RP-LC/UV	HPLC system with reversed-phase columns, OPA/tiol reaction for derivitization and analysis of primary amino acids, separation across an increasing organic gradient	Enables primary amino acid separation and quantification	No amino acid mass information, identification and quantification based on standard retention time and concentration
		Ability to connect LC to bioreactor for real-time analysis	
FMOC derivitization + RP-LC/UV	HPLC system with reversed-phase columns, FMOC-Cl reaction for derivitization and analysis of secondary amino acids, separation across an increasing organic gradient	Enables secondary amino acid separation and quantification	No amino acid mass information, identification and quantification based on standard retention time and concentration
		Ability to connect LC to bioreactor for real-time analysis	
Derivitization + RP-LC/MS	HPLC system with reversed-phase columns, FMOC or OPA derivitization and analysis of primary or secondary amino acids, separation across an increasing organic gradient, MS detection requiring change in buffer conditions for ionization	Ability to connect LC to bioreactor for real-time analysis	Possible changes in LC/UV chromatography when changing to buffer modifications for mass-spectrometry compatibility
		Added benefit of MS detection, verification of primary and secondary amino acid based on mass	MS not commonly used in QA/QC for amino acids

Agilent applications literature

Publication #	Title
5990-4547EN	Improved Amino Acid Methods using Agilent ZORBAX Eclipse Plus C18 Columns for a Variety of Agilent LC Instrumentation and Separation Goals
5990-5977EN	Separation of two sulfurated amino acids with other seventeen amino acids by HPLC with pre-column derivatization
5990-3283EN	Rapid and Precise Determination of Cellular Amino Acid Flux Rates Using HPLC with Automated Derivatization with Absorbance Detection
5989-6297EN	High-Speed Amino Acid Analysis (AAA) on 1.8 µm Reversed-Phase (RP) Columns
5980-1193EN	Rapid, Accurate, Sensitive and Reproducible Analysis of Amino Acids

For more information

Learn more:
www.agilent.com/chem/biopharma

Buy online:
www.agilent.com/chem/store

Find an Agilent customer center in your country:
www.agilent.com/chem/contactus

U.S. and Canada
1-800-227-9770
agilent_inquiries@agilent.com

Europe
info_agilent@agilent.com

Asia Pacific
inquiry_lsca@agilent.com

Printed in USA August 08, 2011
5990-8561EN

Agilent Technologies