# 3 Protein Stability and Characterization

*Atanas Koulov*

## INTRODUCTION

One of the main tasks in the development of protein therapeutics is the detailed characterization of the recombinant protein—drug candidate. Gaining intimate knowledge of the molecular characteristics of the protein is required for understanding and controlling the manufacturing process, and also the stability of the molecule. The latter is critical for developing a stable and fit-for-purpose drug product, as well as defining an appropriate control strategy for monitoring the stability during long-term storage.

## ANALYTICAL TOOLBOX: GENERAL OVERVIEW

The large diversity of possible protein modifications necessitates the use of a broad array of analytical approaches. Naturally, the largest share of analytical methods comes from well-established and traditional technologies of separation science, such as chromatography and electrophoresis. Over the last couple of decades, mass spectrometry approaches have also gained vast popularity, largely owed to the very rapid development of the technology in this field. Unlike traditionally used separation technology approaches, mass spectrometry allows for elucidating the structure of protein modifications and in turn monitoring specific molecular modifications (e.g. oxidation at a specific amino acid residue), as opposed to measuring global (population) changes in the protein structure.

### ■ Chromatography

Chromatography techniques are extensively used in biotechnology, not only in protein purification procedures (see Chap. 4) but also in assessing the integrity of the product. Routine procedures are highly automated so that comparisons of similar samples can be made. An analytical chromatographic system consists of an autosampler, which will take a known amount (usually a known volume) of material for analysis and automatically places it in the solution stream (mobile phase) headed toward a separation column used to fractionate the sample. Another part of this system is a pump module, which provides a reproducible flow rate. In addition, the pumping system can provide a gradient, which changes the properties of the mobile phase such as pH, ionic strength, and polarity. A detection system (or possibly multiple detectors in series) is located at the outlet of the column. This measures the relative amount of protein exiting the column. Coupled to the detector is a data acquisition system. This takes the signal from the detector and integrates it into a value related to the amount of material (see Fig. 3.1). When the protein solution emerges from the column, the signal begins to increase, and as the protein passes through the detector, the signal subsequently decreases. The area under the peak of the signal is proportional to the amount of material that has passed through the detector. By analyzing known amounts of protein, a peak area versus amount of protein plot can be generated and this may be used to estimate the amount of this protein in the sample under other circumstances. Another benefit of this integrated chromatography system is that low levels of components which appear over time can be estimated relative to the major, desired protein being analyzed. This is a particularly useful function when the long-term stability of the product is under evaluation.

During the more than 100 years of history of chromatography, a large variety of separation modes has been developed and many of these are actively used today for characterization of proteins. Proteins and peptides can be chromatographically separated based on their polarity (reversed-phase chromatography, hydrophobic interaction chromatography, hydrophilic interaction chromatography), charge distribution (ion exchange chromatography), size (size exclusion chromatography), etc. In addition, mixed-mode chromatography (using

A. Koulov (✉)
Analytical Development and Quality Control, Drug Product Services, Lonza AG, Basel, Switzerland
e-mail: atanas.koulov@lonza.com

© Springer Nature Switzerland AG 2019
D. J. A. Crommelin et al. (eds.), *Pharmaceutical Biotechnology*, https://doi.org/10.1007/978-3-030-00710-2_3

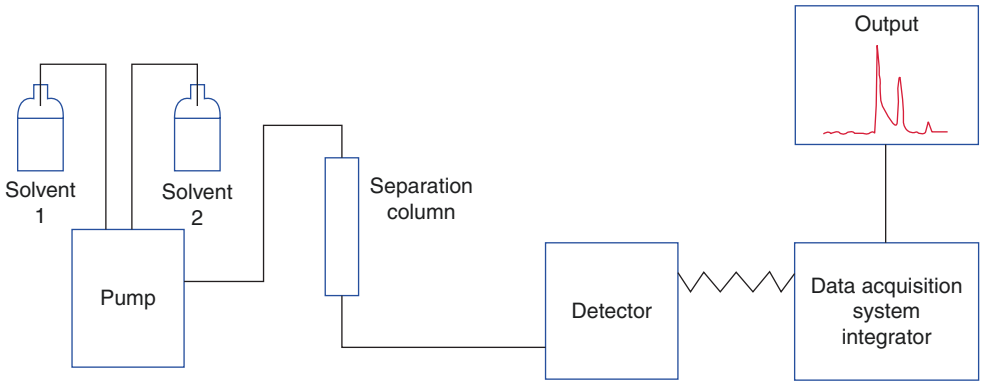


**Figure 3.1** ■ *Components* of a typical chromatography station. The pump combines solvents one and two in appropriate ratios to generate a pH, salt concentration, and/or hydrophobic gradient. Proteins that are fractioned on the column pass through a detector, which measures their occurrence. Information from the detector is used to generate chromatograms and estimate the relative amount of each component

columns with both hydrophobic and charged groups, i.e., a combination of ion-exchange and hydrophobic interaction chromatography) and two-dimensional chromatographic approaches (using a sequential combination of separation modes e.g. reversed phase and ion exchange chromatography) are regularly used when characterizing proteins and peptides.

## ■ Electrophoresis

Generally speaking, the family of electrophoretic techniques separates proteins in an electrical field, based on their charge-to-mass ratio. The charge of the protein depends on the presence of acidic and basic amino acids (cf. Chap. 2) and can be controlled by the pH of the solution in which the protein is separated. The farther away the pH of the solution is from the pI value of the protein, that is, the pH at which it has a net charge of zero, the greater is the net charge and hence the greater is its charge to mass ratio. Alternatively, additives, such as sodium dodecyl sulphate (SDS), may impart an overwhelming negative charge to the protein molecules. This phenomenon forms the basis of the SDS-PAGE (sodium dodecyl sulphate polyacrylamide gel electrophoresis) technique that is/was extensively used to determine the molecular weight of proteins (see section "Polyacrylamide Gel Electrophoresis ((SDS)-PAGE)" and section "Capillary Electrophoresis Sodium Dodecyl Sulfate (CE-SDS)").

Throughout the twentieth century, gel electrophoresis was one of the main methods of choice for characterization of proteins. Since the advent of capillary electrophoresis in the 1980s, significant improvements were achieved in the electrophoretic separation of proteins. Today, capillary electrophoresis is the main electrophoretic method used in the biopharmaceutical analytics.

## ■ Mass Spectrometry

Mass spectrometry (MS) is a technique in which ions of the various species present in the sample are generated using different ionization techniques and where their molecular masses are measured with high accuracy. This technique is one of the most impactful analytical methods in the current biopharmaceutical analytical practice. While this method was used in the past to analyze small, relatively volatile molecules, the molecular weights of highly charged proteins with masses of over 100 kDa can now be accurately determined. Together with the rapid development of informatics and MS analytical instrumentation incorporating different ionization and detection modes, a large number of different variants of MS have been developed and are currently in use.

One of the main advantages of MS is its ability to determine molecular masses with unparalleled accuracy. This attribute has enabled measuring posttranslational modifications with mass differences of only 1 Da and specific modifications that arise during stability studies. For example, an increase in mass of 16 Da suggests that an oxygen atom has been added to the protein as happens when a methionyl residue is oxidized to a methionyl sulfoxide residue. The molecular mass of peptides obtained after proteolytic digestion and separation by high performance liquid chromatography (HPLC) indicates from which region of the primary structure they are derived. Such an HPLC chromatogram is called a "peptide map." An example is shown in Fig. 3.2. This is obtained by digesting a protein with pepsin and by subsequently separating the digested peptides by reverse HPLC. This highly characteristic pattern for a protein is called a "protein fingerprint." Peaks are identified by elution times on HPLC. If peptides have molecular masses differing from those expected from the primary sequence, the nature of the modification to that peptide can be implicated. Moreover, molecular mass estimates can be made for peptides obtained from unfractionated proteolytic digests. Molecular masses that differ from expected values indicate that a part of the protein molecule has been altered, e.g. that a different glycosylation pattern occurs, that a different or degraded amino acid has been found, or that the protein under investigation still contains contaminants.

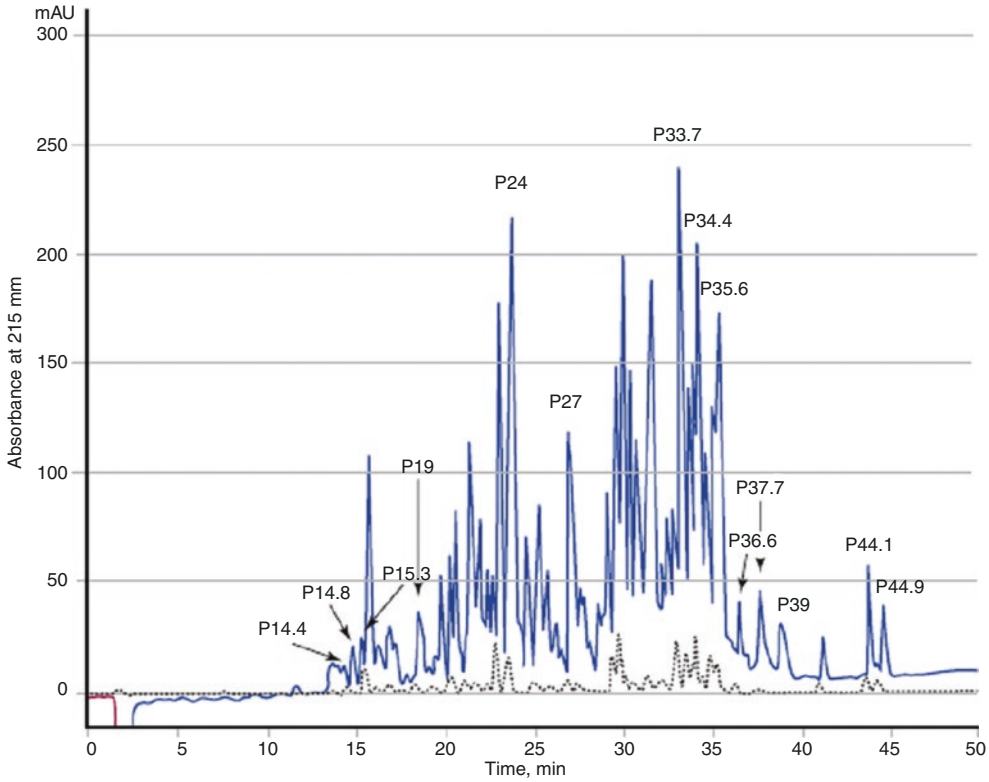


**Figure 3.2** ■ Peptide map of a pepsin digest of recombinant human β-secretase. Each peptide is labeled by elution time in HPLC

Another way of using mass spectrometry as an analytical tool is in the sequencing of peptides. A recurring structure, the peptide bond, in peptides tends to yield fragments of the mature peptide, which differ stepwise by an amino acyl residue. The difference in mass between two fragments indicates the amino acid removed from one fragment to generate the other. Except for leucine and isoleucine, each amino acid has a different mass and hence a sequence can be read from the mass spectrometer. Stepwise removal can occur from either the amino terminus or carboxy terminus. In addition, subsequent fragmentation of the individual peptides ($MS^2$) yields a highly regular fragmentation pattern, which enables sequencing of the parent peptide.

By changing three basic components of the mass spectrometer, i.e. the ion source, the analyzer, and the detector, different types of measurement may be undertaken. Typical ion sources that volatilize proteins are electrospray ionization, fast atom bombardment, and liquid secondary ionization. Common analyzers include quadrupole, magnetic sector, and time-of-flight, electrostatic sector, quadripole ion trap and ion cyclotron resonance instruments. The function of the analyzer is to separate the ionized biomolecules based on their mass-to-charge ratio. The detector measures a current whenever impinged upon by charged particles. Electrospray ionization (El) and matrix-assisted laser desorption (MALDI) are two sources that can generate high-molecular-weight volatile proteins. In the former method, droplets are generated by spraying or nebulizing the protein solution into the source of the mass spectrometer. As the solvent evaporates, the protein remains behind in the gas phase and passes through the analyzer to the detector. In MALDI, proteins are mixed with a matrix, which vaporizes when exposed to laser light, thus carrying the protein into the gas phase. An example of MALDI-mass analysis is shown in Fig. 3.3, indicating the singly charged ion (116, 118 Da) and the doubly charged ion (58,036.2) for a purified protein. Since proteins may carry multiple charges, a number of components are observed representing mass-to-charge forms, each differing from the next by one charge. By imputing various charges to the mass-to-charge values, a molecular mass of the protein can be estimated. The latter step is empirical since only the mass-to-charge ratio is detected and not the net charge for that particular particle.

## ■ Spectroscopic and Other Techniques for Studying Higher Order Structure

A variety of spectroscopic techniques have found broad use for studying protein structure. These techniques differ significantly by the information content provided and by the amount of expert knowledge required for operation and data interpretation. As a rule: the higher the information content (spatial resolution of the structure) provided by a given method, the more laborious, complex to operate and interpret the data obtained it is.

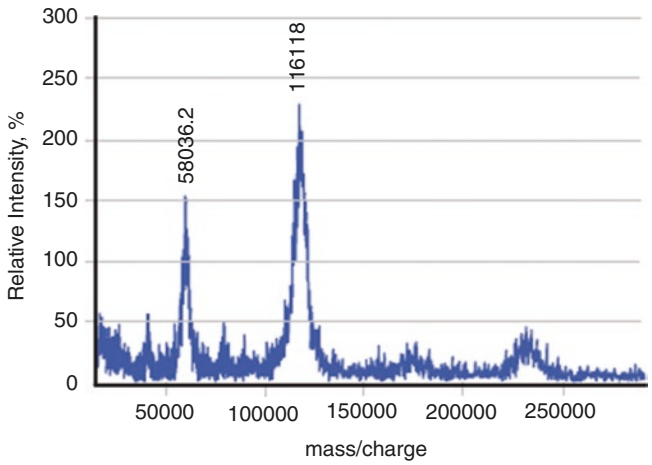


**Figure 3.3** ■ MALDI-mass analysis of a purified recombinant human β-secretase. Numbers correspond to the singly charged and doubly charged ions

Circular Dichroism (CD) is a method that utilizes the property of proteins as chiral molecules to differently absorb the right- and left-handed polarized light across the UV and visible parts of the spectrum. CD is used extensively for probing the secondary and tertiary structure of proteins.

Fourier Transformed Infrared Spectroscopy (FTIR) is used to measure the absorption of infrared (IR) light by proteins due to vibrational transitions of various functional groups. In this technique, the absorption of IR light over a broad wavelength range is measured simultaneously, using a device called interferometer. Different factors (such as hydrogen bonding, redox state, bond angles and conformation) can influence the absorption of the vibrating group, which is why FTIR spectroscopy is widely used to probe a protein structure. The repeat units in proteins give rise to several characteristic absorption bands, of which the Amide I (1700–1600 $cm^{-1}$) bands are perhaps the most useful, due to their sensitivity to the protein secondary structure.

Fluorescence is another widely used spectroscopic technique to study protein conformation, its secondary, tertiary and quaternary structure. In this technique, the fluorescence of the two main fluorophores (Tyr and Trp) present in proteins is used. Trp fluorescence is of particular interest due to the peculiar properties of this fluorophore. One of the two overlapping transitions in Trp is highly sensitive to hydrogen bonding of the indole's –NH group, which in practical terms gives rise to a high sensitivity of the Trp fluorescence spectrum to exposure to water. This property of Trp fluorescence is used to indirectly measure the protein structure as a function of Trp exposure to the aqueous environment.

Some mass spectrometry techniques have been developed with the specific goal to study the higher order protein structure. The most prominent example is Hydrogen-Deuterium exchange MS (HDX-MS). In this technique the different exchange rates of the amide hydrogens over the peptide backbone of the protein are measured using a highly specialized LC-MS system after exposing the protein to deuterium ($D_2O$) for a brief period of time, followed by quenching of the exchange and a rapid peptide mapping measurement (enzymatically digesting the protein into peptide fragments and measuring the deuteration levels of the individual peptides). Kinetic experiments allow measuring the rates of deuteration of the different peptides, which largely depend on the local structural environment—the level of exposure to the aqueous environment, as well as the conformational flexibility of the given peptide. The results from HDX-MS experiments are visualized using protein maps indicating the rates of exchange in the different protein regions. These maps are extremely useful in understanding the protein dynamics, flexibility and accessibility.

Ion mobility MS (IMS-MS) is another MS technique to study the higher order structure of a protein. Briefly, this technique measures the differential mobility of different protein species in the gas phase in an electric field. This method is mostly used to measure the aggregation state of protein mixtures.

Nuclear Magnetic Resonance (NMR) is a technique that has made major contributions to elucidate the 3D structure of proteins and is becoming more and more popular with the concomitant development of the analytical instrumentation. Very briefly, this technique measures the magnetic properties of atomic nuclei (more specifically the interaction of the magnetic moment of an atomic nucleus with an external magnetic field) which strongly depends of their local environment. Thus, when employing various experimental strategies including 2-D (two dimensional) NMR approaches (also including measurements of different nuclei) in principle allows the determination of the 3D (three dimensional) structure of a protein at atomic resolution. One advantage that NMR has over other high-resolution techniques such as X-ray diffraction, see below) is that one can directly observe protein dynamics in kinetic experiments in solution.

X-ray diffraction is still considered the ultimate technique for studying the structure of proteins. This technique uses the phenomenon of diffraction of a monochromatic X-ray beam by protein crystals. In a protein crystallography experiment the diffraction pattern of the protein crystal is captured from many different orientations of the crystal. From the diffraction patterns obtained (intensity and location of the resulting spots) the positions of the atoms in the molecule can be determined which in turn allows for a calculation of a molecular model of the protein in crystal form, often at atomic resolution.

In the next sections of this chapter, the wide-ranging arsenal of analytical methods to separate and characterize various protein structural modifications is presented in the specific context of these protein modifications. More precisely, applicable analytical techniques are discussed in the context of specific protein attribute(s) altered by a given modification. For example, analytical methods to characterize protein charge heterogeneity are discussed in the context of the modifications which introduce change in protein charge, etc.

## PROTEIN STABILITY: WHAT CAN GO WRONG AND HOW TO MEASURE IT?

All levels of structural organization of proteins (See Chap. 2) are susceptible to damage as a consequence of physical or chemical stress (Table 3.1). Different modifications of the protein structure may be manifested as changes in various attributes (properties) of the protein. This is why assessing the stability of protein therapeutics is a complex and multifaceted task. In the following sections of this chapter the most common structural modifications of proteins are presented together with the typical analytical approaches currently applied to measure these modifications.

### ■ Protein Modifications Introducing Changes in Charge Heterogeneity

*Deamidation and Isomerization*

Some of the most common and most significant modifications in terms of impact on the properties of protein biopharmaceuticals are the deamidation of asparagine (Asn) and isomerization of aspartate (Asp). The mechanism of deamidation involves the formation of a cyclic imide intermediate (succinimide), which in turn hydrolyzes spontaneously to a mixture of isoaspartic/aspartic acid at an approximate ratio of 3:1. This reaction may be accompanied by further racemization of the

| Protein modification | Typical causes and important factors | Physical property affected | Method of analysis |
|---|---|---|---|
| Oxidation<br>Cys<br>Disulfide<br>Intrachain<br>Interchain<br>Met, Trp, Tyr | Light, metal ions, peroxides | Hydrophobicity | RP-HPLC, HIC and mass spectrometry |
| Fragmentation | pH, sequence (nearest AA neighbor) | Size | Size-exclusion chromatography, SDS-PAGE |
| N to O migration<br>Ser, Thr | | Hydrophobicity | RP-HPLC inactive in Edman reaction |
| α-Carboxy to β-carboxy migration<br>Asp, Asn | | Hydrophobicity | RP-HPLC inactive in Edman reaction |
| Deamidation<br>Asn, Gln | pH, sequence (nearest AA neighbor), HOS | Charge | Ion-exchange chromatography |
| Acylation<br>α-Amino group, ε-amino group | | Charge | Ion-exchange chromatography mass spectrometry |
| Esterification/carboxylation<br>Glu, Asp, C-terminal | | Charge | Ion-exchange chromatography mass spectrometry |
| Secondary structure changes | | Hydrophobicity | RP-HPLC |
| | | Size | Size-exclusion chromatography |
| | | Sec/tert structure | CD |
| | | Sec/tert structure | FTIR |
| Aggregation | | Sec/tert structure | Fluorescence<br>Light scattering |
| | | | Analytical ultracentrifugation, AF4 |

**Table 3.1** ■ Common protein modifications and corresponding methods of analysis

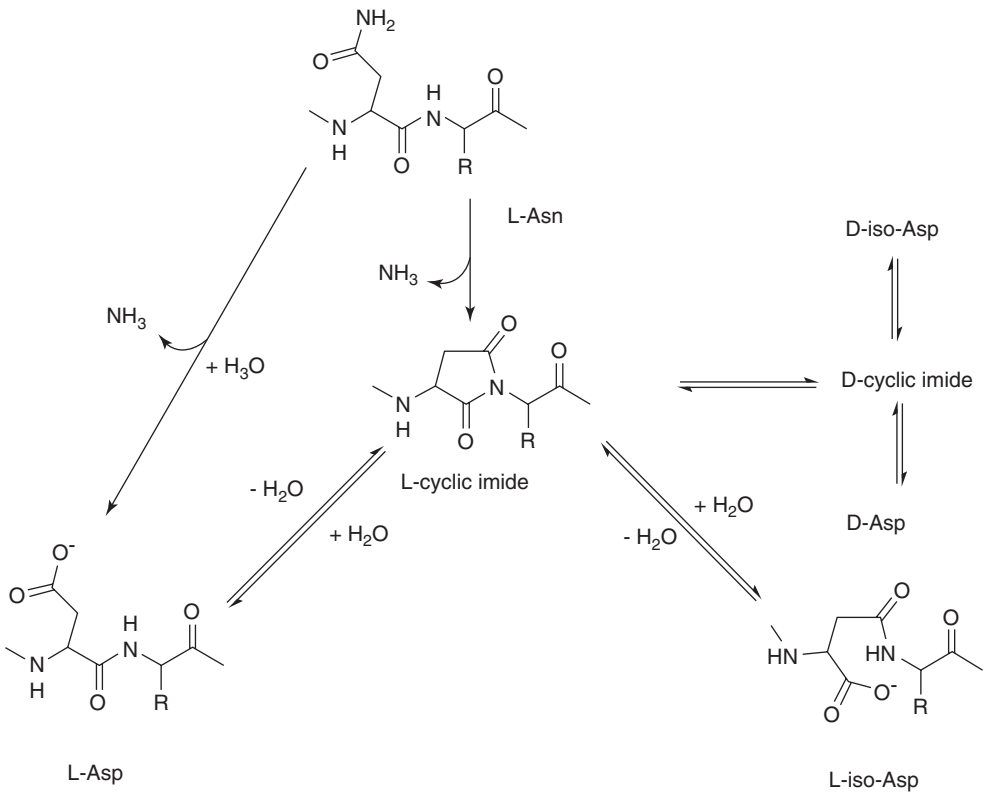


**Figure 3.4** ■ Deamidation, isomerization and succinimide formation

isoaspartyl and aspartyl residues via the succinimide intermediate (See Fig. 3.4). Typically, the succinimide intermediate is short-lived at neutral pH, but in some cases may be stabilized.

Isomerization of Asp to isoAsp and deamidation of Asn occur frequently in biotherapeutics. In some cases, these modifications may be benign, but in others they may result in severe consequences for the product, for example in cases when the complementarity-determining regions (CDR) regions of MABs are affected. The biological activity of these molecules may be altered. The most important factors that influence deamidation and isomerization rates are temperature, pH, local protein structure and flanking aminoacyl residues. All chemical reactions mentioned above either result in changes in the charge of the affected protein (deamidation and succinimide formation), or in changes of the surface charge distribution (isomerization). Whereas deamidation results in an increase of acidic species, succinimide formation contributes to an increase of basic variants of the protein.

In should be noted that deamidation can also occur in glutamine (Gln) residues. However, the rates of Gln deamidation are much slower than Asn deamidation rates.

### *Pyro-Glu Formation*

Another common modification that results in a change in the protein charge heterogeneity is the formation

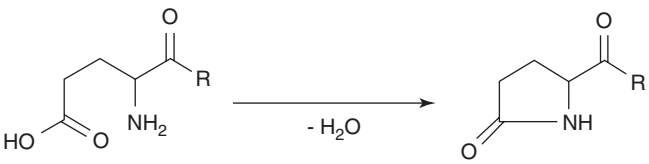


**Figure 3.5** ■ Pyro-Glu formation

of pyroglutamate (pyro-Glu). Cyclization of the N-terminal glutamate (Glu) to pyroGlu (see Fig. 3.5) may occur either enzymatically or spontaneously. PyroGlu formation typically results in protein species with a higher pI than the main isoform.

### *Glycation*

Glycation of proteins is the addition of reducing sugars (e.g. glucose or lactose) to the primary amine of lysine residues. It typically occurs during manufacturing in glucose-containing culture media. The glycation of proteins results in the increase of acidic protein variants.

### *Additional Modifications Inducing Changes in Protein Charge Heterogeneity*

All protein charge modifications mentioned above occur as a result of chemical instability of proteins, i.e. under various types of physico-chemical stress: extreme pH, high temperature, etc. Other modifications that may result in the formation of protein

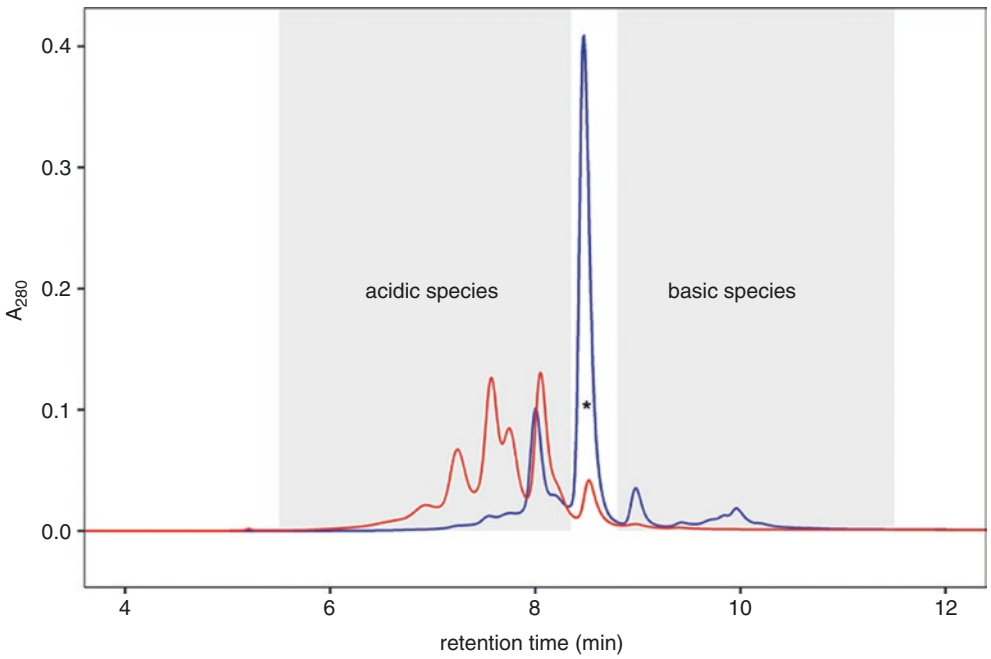


**Figure 3.6** ■ Charge heterogeneity profile of a MAB using Cationic Exchange Chromatography: control (blue line) and stressed (red line) acidic species formed during incubation at 40 °C and pH 4.5. The main peak is indicated by a star

charge variants occur as a result of enzymatic reactions—typically during the fermentation process while manufacturing biopharmaceuticals. Such modifications are, for example, the formation of C-terminal Lys variants and the sialylation of proteins.

## ■ Measuring Changes in Protein Charge Heterogeneity

A number of different protein modifications occurring either during long term storage or during fermentation (upstream processing) may result in changes of the protein charge heterogeneity profile (e.g. deamidation, isomerization, glycation, etc.—see previous section). Because these modifications commonly occur simultaneously, in practice the resulting charge heterogeneity patterns of proteins are often relatively complex. Thus, characterization of the various species underlying the complex protein charge heterogeneity may require the application of different analytical approaches.

There are two main groups of techniques commonly used to measure changes in the charge profile distribution of protein therapeutics: electrophoretic techniques (IEF/icIEF, CZE) and chromatographic techniques (IEC), see below.

### *Ion-Exchange Chromatography (IEC)*

A group of methods that has traditionally been applied for assessment of charge heterogeneity and still finds a very broad use in this context is the group of analytical ion exchange chromatographic techniques (IEC or IEX). The commercial availability of a variety of stationary phases (chromatographic columns) for separation of charge variants using HPLC provides a choice of separation modes (anionic or cationic, strong or weak) and the opportunity for a very good separation and fractionation of variants that are difficult to separate by other techniques.

This technique takes advantage of the electric charge properties of proteins. Some of the amino acyl residues are negatively charged and others are positively charged. The net charge of the protein can be modulated by the pH of its environment relative to the pI value of the protein. At a pH value lower than the pI, the protein has a net positive charge, whereas at a pH value greater than the pI, the protein has a net negative charge. IEC utilizes various resins (chromatographic stationary phases), containing functional groups with either positive or negative charges (anion- or cation-exchange chromatography, correspondingly), depending on the pI of the separated protein. Positively charged proteins bind to negatively charged matrices and negatively charged proteins bind to positively charged matrices. Proteins bound to the chromatographic column are displaced (eluted) from the resin either by increasing the salt concentration of the mobile phase (screening the protein-column charge-charge interactions), or changing the pH of the mobile phase (effectively changing the charge of the protein). Proteins or protein variants with different net charges are separated from one another during elution with the change in the gradient (salt or pH). The choice of the charged resin and elution conditions are dependent upon the protein of interest.

Figure 3.6 shows an example separation of a monoclonal antibody using cationic exchange chromatography (CEX). Acidic isoforms elute before (left side of the chromatogram) and basic isoforms after (right side of

the chromatogram) the main isoform. Upon exposure to low pH and elevated temperature stress conditions, significant changes in the charge heterogeneity of this protein can be observed. One finds a large decrease of the main isoform, accompanied by a decrease in the basic and increase in the acidic charge isoforms. Due to the complexity of the possible reactions mentioned in the previous section, it is difficult to assess what are the specific changes underlying the re-distribution of charge variants using this chromatogram alone. For this purpose, typically, it is necessary to fractionate the individual peaks (or groups of peaks) and subject them to further analyses (typically mass spectrometry) in order to establish unequivocally the specific sequence modifications (see section "Mass spectrometry").

### *Isoelectric Focusing (IEF/cIEF)*

Another family of analytical methods to separate proteins based on their electric charge properties is isoelectric focusing (IEF). Isoelectric focusing techniques rely on separating proteins based on their isoelectric point (pI). In a first run, a pH gradient is established within a polyacrylamide gel (or a capillary in cIEF) using a mixture of small-molecular-weight ampholytes with varying pI values. After introduction of the protein sample and application of an electric field, all proteins or protein species/variants migrate within the pH gradient to the pH where their corresponding net charge is zero (their apparent pI). This technique is very useful for separating protein charge variants, such as deamidated or glycated species, from the native protein.

Isoelectric focusing (IEF), or its capillary configuration: Imaged Capillary Isoelectric Focusing (icIEF), has the advantages that it can be applied to a broad variety of molecules and that it typically requires minimal method development efforts. icIEF has found particularly broad use in the biotech industry as a "platform" or "generic" charge heterogeneity assessment method due to its relative ease of use, minimal sample requirements and its broad applicability.

### *Capillary Zone Elecrophoresis (CZE)*

Another method which has gained an increased presence over the last decades is Capillary Zone Electrophoresis (CZE). Rather than separating the proteins in a matrix, as in polyacrylamide gel electrophoresis through which the proteins migrate, in CZE the proteins are free in solution in an electric field within the confines of a capillary tube with a diameter of 25–50 μm. After passing through the capillary tube, proteins encounter an UV absorbance or fluorescence detector which can be used to quantify the proteins. The movement of one protein relative to another is a function of the molecular mass and the net charge on the protein. The latter can be influenced by pH and analytes in the solution. Typically, various additives or capillary coatings (e.g. epsilon amino caproic acid (EACA), or triethylenetetramine (TETA)) are used to suppress the interaction of proteins with the capillary wall, as well as the electro-osmotic flow.

CZE offers several advantages over other analytical methods described here to assess charge heterogeneity, such as the relatively easy implementation and low development efforts required. This makes CZE very suitable as a platform method. In addition, it offers a robust and rapid separation, which makes it amenable to high throughput applications.

All analytical methods for measuring protein charge heterogeneity described in this chapter have advantages and disadvantages. The specific application of a given method depends on the given protein or protein mixture measured (e.g. the column, selection of method conditions,) and on the type of modification of the protein structure. Surface modifications for example, are easily detected by IEC, whereas modifications buried within the structure of the protein may be detected better by IEF and CZE. Additional considerations for method selection depend on the specific purpose of the analyses. In cases where the characterization goal is to simply measure the changes under given stress condition, an IEF or a CZE measurement may be sufficient. However, one disadvantage of these electrophoretic methods is the inability for direct fractionation of molecular variants and online coupling (hyphenation) to mass spectrometry for measurements of changes in the primary sequence. Thus, if the goal of the investigation is to understand the chemical nature of these changes and additional measurements (e.g. MS) may be required, IEC may be preferred. In practice, often a combined approach is used—e.g. electrophoretic methods may be used to measure charge heterogeneity routinely, while complementary IEC methods are used for fractionation of specific variants when needed.

## ■ Protein Modifications Introducing Changes in Size

Proteins can undergo a number of changes that affect their size. These changes may be either covalent modifications such as fragmentation of the polypeptide chain and intermolecular disulfide scrambling, or non-covalent modifications such as protein aggregation and particle formation. Often, such changes dramatically affect protein biopharmaceuticals' potency and safety profiles.

### *Protein Fragmentation*

Despite the fact the peptide bond as such is remarkably stable, fragmentation of the polypeptide backbone of recombinant proteins is a commonly observed modification. The reason for this apparent contradiction is that often, the adjacent amino acid side-chains

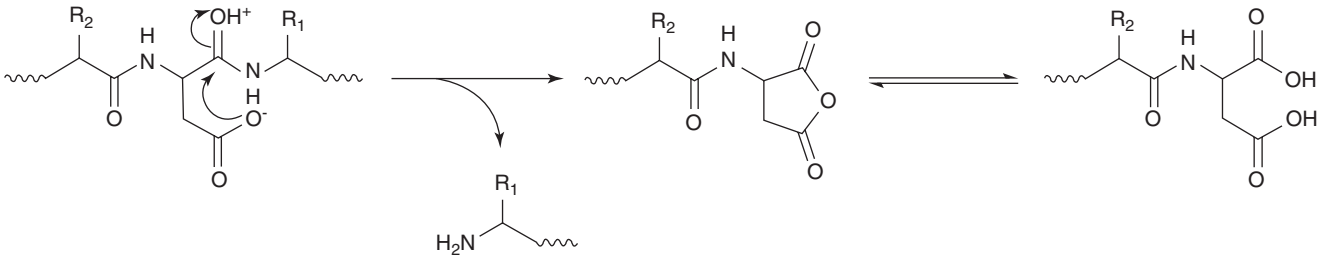


**Figure 3.7** ■ Fragmentation of a polypeptide chain at an "Asp-X" site

or local structure flexibility may contribute significantly to rendering a site susceptible to fragmentation. Neighboring amino acid side chains (most commonly Asp, Ser/Thr, Cys/Cys-Cys, Asn) may result in fragmentation which may occur via distinct mechanisms (see details below). In addition, flexible regions, such as disordered loops or the IgG hinge regions, for example, may be particularly prone to fragmentation. Various additional factors and conditions (such as pH, temperature, the presence of radicals) may contribute to increased fragmentation of the polypeptide chain of a protein as well. One of the most common examples of polypeptide chain fragmentation in recombinant proteins is fragmentation at an Asp-X site, where X is any amino acid residue (shown in Fig. 3.7). In this example, at low pH a nucleophilic attack of the ionized side-chain carboxylate of the Asp on the protonated carbonyl carbon of the peptide bond takes place, followed by release of the C-terminal peptide, cyclization and a further hydrolysis of the unstable aspartic anhydride to an Asp residue.

There are many other reported mechanisms of fragmentation of the polypeptide chain, which involve various amino acid residues. Depending on the specific mechanism, various factors affect the rate of fragmentation. In the example above, low pH and small amino acid residues at position X favor the fragmentation reaction. Typically, the structural context (three-dimensional structure of the protein) also influences fragmentation with more disordered regions being more susceptible.

### *Protein Aggregation*

The term "protein aggregation" refers to the process of agglomeration of two or more protein molecules, but it is typically distinct from functional protein-protein binding or quaternary structure formation. This term is too general for practical use, as it encompasses a vast diversity of different molecular phenomena. Protein aggregates can be reversible or irreversible, soluble or insoluble, covalent or non-covalent, etc. Furthermore, protein aggregates are typically present as a continuum of species spanning an enormous size range: from a few nm to >100 μm. Despite various attempts to categorize protein aggregation in a systematic fashion, to date there is still no sufficient clarity and agreement on nomenclature.

Protein aggregation may take place as a result of a variety of phenomena, such as local unfolding or perturbation of the protein secondary, tertiary or quaternary structure. This general mechanism is typically evoked in cases when the system (e.g. protein solution) receives an excess of energy, such as during thermal stress. The physical stability of a protein is expressed as the difference in free energy, $\Delta G_U$, between the native and denatured states. Thus, protein molecules are in equilibrium between the above two states. As long as this unfolding is reversible and $\Delta G_U$ is positive, it does not matter how small the $\Delta G_U$ is. In many cases, this reversibility does not hold. This is often seen when $\Delta G_U$ is decreased by heating. Most proteins denature (i.e., unfold) upon heating and subsequent aggregation of the denatured molecules results in irreversible denaturation. Thus, reversible unfolding is made irreversible by aggregation:

$$\text{Native state} \overset{\Delta G_U}{\Leftrightarrow} \text{Denatured state} \overset{k}{\Rightarrow} \text{Aggregated state}$$

Therefore, any stress that decreases $\Delta G_U$ and increases $k$ will cause the accumulation of irreversibly inactivated forms of the protein. Such stresses may include chemical modifications as described above. Protein aggregation may occur as a result of oxidative processes such as disulfide scrambling or the chosen physical conditions, such as pH, ionic strength, protein concentration, and temperature. Development of a suitable formulation that prolongs the shelf life of a recombinant protein is essential when it is to be used as a human therapeutic (cf. Chap. 5).

Despite the variety of molecular mechanisms via which protein aggregation may ensue, the final outcome is the generation of protein species larger than the original molecule. Thus, all techniques for measuring protein aggregation are size-based.

## ■ Measuring Changes in Protein Size

Similar to all other analytical techniques, the size-based protein characterization methods each have their advantages and disadvantages. Thus, some find more extensive use in characterizing small protein fragments and others in measuring large molecular weight species and proteinaceous particles. Because of the huge diver-

sity of protein aggregation mechanisms and products, naturally there is a large array of analytical techniques, which have been developed to study specifically protein aggregation and assess its various aspects. Some techniques are suitable for measuring and characterizing small aggregates (oligomers), whereas others are useful for measuring exclusively high-molecular weight species (protein particles). Some techniques are specifically utilized to evaluate covalent aggregates, whereas others are used for non-covalent aggregates. However, in reality nearly all of the common size-based protein characterization analytical techniques find dual use—for assessment of fragmentation and aggregation alike. This section will present only the most commonly used techniques, without delving into details and discussing the highly specialized technologies with very limited, niche use.

### *Size-Exclusion Chromatography*

As the name implies, this technique separates proteins based on their size or molecular weight or shape. The matrix consists of very fine beads containing cavities and pores accessible to molecules of a certain size or smaller, but inaccessible to larger molecules. The principle of this technique is the distribution of molecules between the volume of solution within the beads and the volume of solution surrounding the beads (cf. Fig. 4.6). Small molecules have access to a larger volume than large molecules. As the solution flows through the column, molecules can diffuse back and forth, depending upon their size, in and out of the pores of the beads. Smaller molecules can reside within the pores for a finite period whereas larger molecules, unable to enter these spaces, continue along in the fluid stream. Intermediate-sized molecules spend an intermediate amount of time within the pores. They can be fractionated from large molecules that cannot access the matrix space at all and from small molecules that have free access to this volume and spend most of the time within the beads.

Size-exclusion chromatography can be used to estimate the mass of proteins by calibrating the column with a series of globular proteins of known mass. However, the separation depends on molecular shape (conformation) as well as mass and highly elongated proteins—proteins containing flexible, disordered regions—and glycoproteins will often appear to have masses as much as two to three times the true value. Other proteins may interact weakly with the column matrix and be retarded, thereby appearing to have a smaller mass. Size-exclusion chromatography can be used to measure both protein fragmentation and protein aggregation, with the latter being more the common application. Sedimentation (ultracentrifugation), light scattering or MS methods are preferred for accurate mass measurement.

### *Polyacrylamide Gel Electrophoresis ((SDS)-PAGE)*

One of the earliest methods for analysis of proteins is polyacrylamide gel electrophoresis (PAGE). Polyacrylamide gels are used as a sieve. By adjusting the concentration of acrylamide that is used in these gels, one can control the migration rate of material within the gel. The more acrylamide, the more hindrance for the protein to migrate in an electrical field. The polyacrylamide gel provides not only a separation matrix, but also a matrix to hold the proteins in place until they can be detected with suitable reagents.

The direction and speed of migration of the protein in a gel depend on the pH of the gel. If the pH of the gel is above its pI value, then the protein is negatively charged and hence migrates toward the positive electrode. The higher the pH of the gel, the faster the migration. This type of electrophoresis is called native gel electrophoresis.

The addition of a detergent, sodium dodecyl sulfate (SDS), to the electrophoretic separation system allows for the separation to take place primarily as a function of the size of the protein. Dodecyl sulfate ions form complexes with proteins, resulting in an unfolding of the proteins, and the amount of detergent that is complexed is proportional to the mass of the protein. The larger the protein, the more detergent that is complexed. Dodecyl sulfate is a negatively charged ion. When proteins are in a solution of SDS, the net effect is that the own charge of the protein is overwhelmed by that of the dodecyl sulfate complexed with it, so that the proteins take on a net negative charge proportional to their mass. Polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate is commonly known as SDS-PAGE. All the proteins take on a net negative charge, with larger proteins binding more SDS but with the charge to mass ratio being fairly constant among the proteins. Since all proteins have essentially the same charge to mass ratio, how can separation occur? This is done by controlling the concentration of acrylamide in the path of proteins migrating in an electrical field. The greater the acrylamide concentration, the more difficult it is for large protein molecules to migrate relative to smaller protein molecules. This is sometimes thought of as a sieving effect, since the greater the acrylamide concentration, the smaller the pore size within the polyacrylamide gel. Indeed, if the acrylamide concentration is sufficiently high, some high-molecular-weight proteins may not migrate at all within the gel. Since in SDS-PAGE the proteins are denatured, their hydrodynamic size, and hence the degree of retardation by the sieving effects, is directly related to their mass. Proteins containing disulfide bonds will have a much more compact structure and higher mobility for their mass unless the disulfides are reduced prior to electrophoresis.

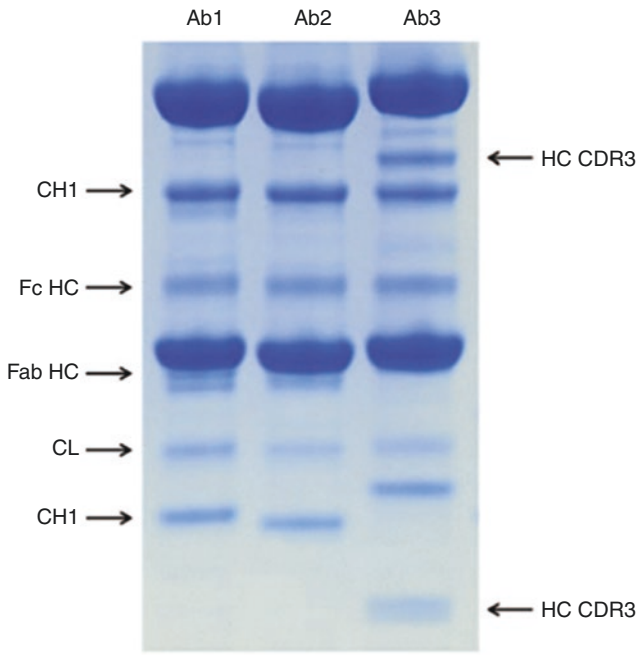


**Figure 3.8** ■ Fragmentation of three (Ab1, Ab2 and Ab3) monoclonal antibodies monitored using SDS-PAGE (*taken from Vlasak et al., MAbs (2011) 3:3, 253–263*); CH1—heavy chain constant domain 1, Fc HC—heavy chain Fc domain, Fab HC—heavy chain Fab domain, CL—light chain constant domain, HC—heavy chain, CDR3—complementarity-determining region 3

An example of SDS-PAGE is shown in Fig. 3.8. Here, SDS-PAGE is used to monitor the polypeptide chains of three different monoclonal antibodies and their various fragmentation products.

As described above, native gel electrophoresis and SDS-PAGE are quite different in terms of the mechanism of protein separation. In native gel electrophoresis, the proteins are in the native state and migrate on their own charges. Thus, this electrophoresis can be used to characterize proteins in the native state. In SDS-PAGE, proteins are unfolded and migrate based on their molecular mass.

### Detection of Proteins Within Polyacrylamide Gels

Although the polyacrylamide gels provide a flexible support for the proteins, with time the proteins will diffuse and spread within the gel. Consequently, the usual practice is to fix the proteins using fixing solutions (rendering the proteins insoluble) or trap them at the location where they migrated to.

There are many methods for staining proteins in gels, but the two most common and well-studied methods are either staining with a dye called Coomassie blue or by a method using silver. The latter method is used if a very low limit of detection needs to be achieved. The principle of developing the Coomassie blue stain is the hydrophobic interaction of a dye with the protein. Thus, the gel takes on a color wherever a protein is located. Using standard amounts of proteins, the amount of protein or contaminant may be estimated. Quantification using the silver staining method is more complex.

### Blotting Techniques

Blotting methods form an important niche in the analytical toolbox of biotech products. Typically, they are used to detect very low levels of unique molecules in a milieu of proteins, nucleic acids, and other cellular components. For example, they can be used to detect components from the host cells used for the production of recombinant proteins (cf. Chap. 4). When blotting, biomolecules are transferred to a synthetic membrane ("blotting"), and this membrane is then probed with specific reagents to detect the molecule of interest. Membranes used in protein blots are made of a variety of materials including nitrocellulose, nylon, and polyvinylidene difluoride (PVDF), all of which avidly bind protein.

Liquid samples can be analyzed by methods called dot blots or slot blots. A solution containing the biomolecule of interest is filtered through a membrane, which captures the biomolecule. Often, the sample is subjected to some type of fractionation, such as polyacrylamide gel electrophoresis, prior to the blotting step. An early technique, Southern blotting, named after the discoverer, E.M. Southern, is used to detect DNA fragments. When this procedure was adapted to RNA fragments and to proteins, other compass coordinates were chosen as labels for these procedures, i.e., northern blots for RNA and western blots for proteins. Western blots involve the use of labeled antibodies to detect specific proteins.

Following polyacrylamide gel electrophoresis, the transfer of proteins from the gel to the membrane can be accomplished in a number of ways. Originally, blotting was achieved by capillary action. The transfer of proteins to the membrane can occur under the influence of an electric field, as well. The electric field is applied perpendicularly to the original field used in separation so that the maximum distance the protein needs to migrate is only the thickness of the gel, and hence, the transfer of proteins can occur very rapidly. This latter method is called electroblotting.

Once the transfer has occurred, the next step is to identify the presence of the desired protein. In addition to various colorimetric staining methods, the blots can be probed with reagents specific for certain proteins, as for example, antibodies to a protein of interest. This technique is called immunoblotting. In the biotechnology field, immunoblotting is used as an identity test for the product of interest. An antibody that recognizes the desired protein is used in this instance. Secondly,

| |
|---|
| 1. Antibodies are labeled with radioactive markers such as $^{125}I$ |
| 2. Antibodies are linked to an enzyme such as horseradish peroxidase (HRP) or alkaline phosphatase (AP). On incubation with substrate, an insoluble colored product is formed at the location of the antibody. Alternatively, the location of the antibody can be detected using a substrate which yields a chemiluminescent product, an image of which is made on photographic film |
| 3. Antibody is labeled with biotin. Streptavidin or avidin is added to strongly bind to the biotin. Each streptavidin molecule has four binding sites. The remaining binding sites can combine with other biotin molecules which are covalently linked to HRP or to AP |

**Table 3.2** ■ Detection methods used in blotting techniques (further explained in the text below)

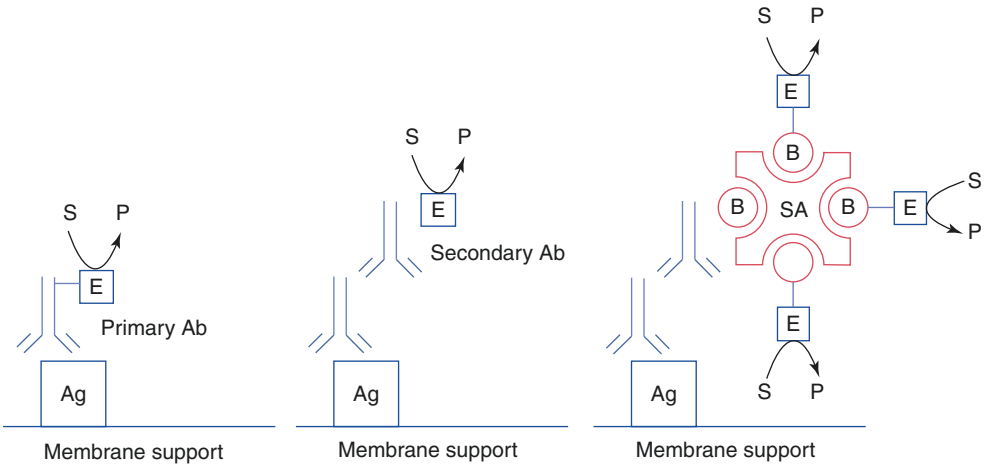


**Figure 3.9** ■ Common immunoblotting detection systems used to detect antigens (protein of interest). Ag, on membranes. Abbreviations used: *Ab* antibody, *E* enzyme, such as horseradish peroxidase or alkaline phosphatase, *S* substrate, *P* product, either colored and insoluble or chemiluminescent, *B* biotin, *SA* streptavidin

immunoblotting is sometimes used to show the absence of host proteins. In this instance, the antibodies are raised against proteins of the organism in which the recombinant protein has been expressed. This latter method can attest to the purity of the desired protein.

The antibody reacts with a specific protein on the membrane only at the location of that protein because of its specific interaction with its antigen. When immunoblotting techniques are used, methods are still needed to recognize the location of the interaction of the antibody with its specific protein. A number of procedures can be used to detect this complex (see Table 3.2).

The antibody itself can be labeled with a radioactive marker such as $^{125}I$ and placed in direct contact with X-ray film. After exposure of the membrane to the film for a suitable period, the film is developed and a photographic negative is made of the location of radioactivity on the membrane. Alternatively, the antibody can be linked to an enzyme that, upon the addition of appropriate reagents, catalyzes a color or light reaction at the site of the antibody. These procedures entail purification of the antibody and specifically label it. More often, "secondary" antibodies are used. The primary antibody is the one that recognizes the protein of interest. The secondary antibody is then an antibody that specifically recognizes the primary antibody. Quite commonly, the primary antibody is raised in rabbits. The secondary antibody may then be an antibody raised in another animal, such as a goat, which recognizes rabbit antibodies. Since this secondary antibody recognizes rabbit antibodies in general, it can be used as a generic reagent to detect rabbit antibodies in a number of different proteins of interest that have been raised in rabbits. Thus, the primary antibody specifically recognizes and complexes a unique protein, and the secondary antibody, suitably labeled, is used for detection (see also section "ELISA" and Fig. 3.9).

The secondary antibody can be labeled with a radioactive or enzymatic marker group and used to detect several different primary antibodies. Thus, rather than purifying a number of different primary antibodies, only one secondary antibody needs to be purified and labeled for recognition of all primary antibodies. Because of their wide use, many common secondary antibodies are commercially available in kits containing the detection system and follow routine, straightforward procedures.

In addition to antibodies raised against the amino acyl constituents of proteins, specific antibodies can be used which recognize unique posttranslational components in proteins, such as phosphotyrosyl residues, which are important during signal transduction, and carbohydrate moieties of glycoproteins.

Figure 3.9 illustrates the above mentioned detection methods that can be used on immunoblots. The primary antibody, or if convenient, the secondary antibody, can have an appropriate label for detection. They may be labeled with a radioactive tag as mentioned previously. Secondly, these antibodies can be coupled with an enzyme such as horseradish peroxidase (HRP) or alkaline phosphatase (AP). Substrate is added and is converted to an insoluble, colored product at the site of the protein-primary antibody-secondary antibody-HRP product. An alternative substrate can be used which yields a chemiluminescent product. A chemical reaction

leads to the production of light that can detected by photographic or X-ray film. The chromogenic and chemiluminescent detection systems have comparable sensitivities to radioactive methods. The former detection methods are displacing the latter method, since problems associated with handling radioactive material and radioactive waste solutions are eliminated.

As illustrated in Fig. 3.9, streptavidin, or alternatively avidin, and biotin can play an important role in detecting proteins on immunoblots. This is because biotin forms very tight complexes with streptavidin and avidin. Secondly, these proteins are multimeric and contain four binding sites for biotin. When biotin is covalently linked to proteins such as antibodies and enzymes, streptavidin binds to the covalently bound biotin, thus recognizing the site on the membrane where the protein of interest is located.

### *Capillary Electrophoresis Sodium Dodecyl Sulfate (CE-SDS)*

Today, the traditional slab gel SDS-PAGE is increasingly being replaced by capillary electrophoresis sodium dodecyl sulfate (CE-SDS) due to the improved convenience and possibilities for automation, superior separation and reproducibility of this newer technique. In CE-SDS the separation is carried out in a capillary in the presence of a sieving matrix. Whereas the basic electrophoretic separation principle of CE-SDS is the same as the one of SDS-PAGE, there are also some significant differences. Unlike SDS-PAGE, where only cross-linked polyacrylamide is used as a sieving matrix, in CE-SDS a variety of linear or slightly branched polymers may be used for the same purpose (e.g. linear polyacrylamide, polyethylene oxide, polyethylene glycol, dextran, pullulan). This contributes to the method's flexibility. Another aspect where CE-SDS differs is the detection mode. In this technique, the laborious step of post-separation staining of the SDS-PAGE gels is eliminated and replaced by online UV or highly sensitive fluorescence detection. The elimination of the staining/destaining step as well as the need for scanning of the gels in CE-SDS (online detection generates quantitative electropherograms), together with the CE instrument design contributes to faster and more reproducible analysis, as well as amenability to automation. Altogether, CE-SDS is considered a superior method, demonstrating better accuracy, linearity and precision than SDS-PAGE, which is why the latter has been effectively replaced by CE-SDS in the current pharmaceutical analytical practice.

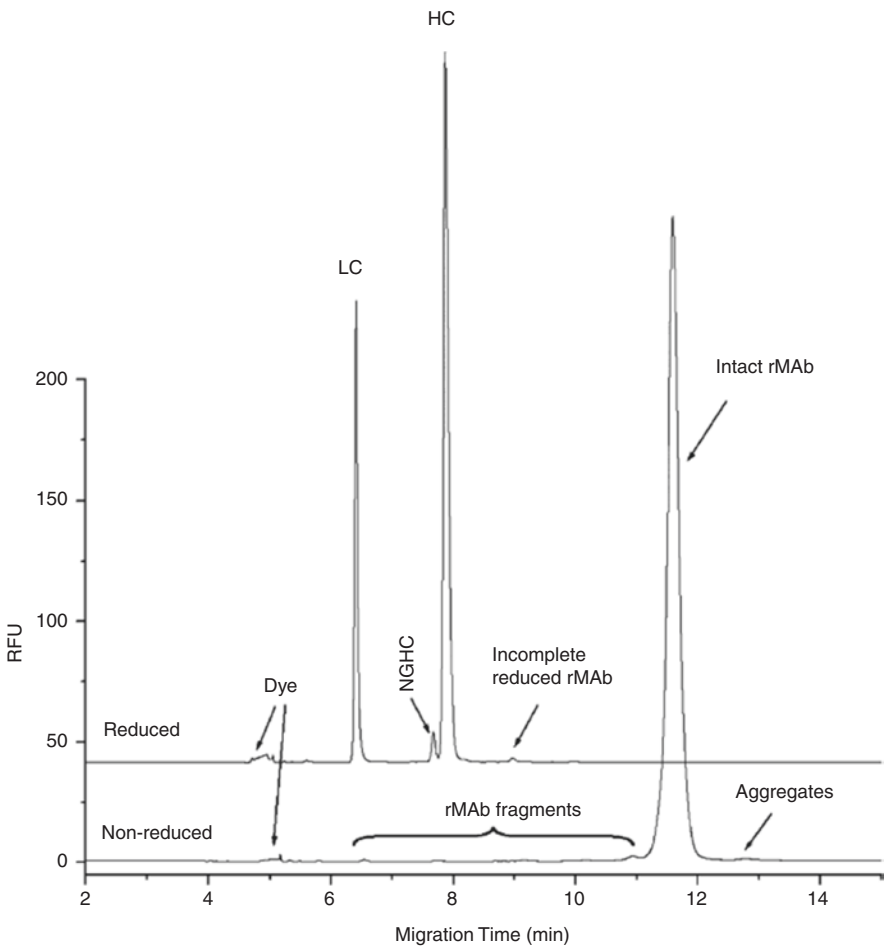


**Figure 3.10** ■ CE-SDS reduced and non-reduced separations of a monoclonal antibody, showing the intact MAB, the heavy (HC) and light chains (LC) as well as various fragments, non-glycosylated form and incompletely reduced recombinant rMAB. Taken from Salas-Solano and Felten (2008)

Figure 3.10 shows an example of reduced and non-reduced CE-SDS separations of a monoclonal antibody.

An interesting new trend in the CE instrument development that has emerged recently—the development of new systems based on a microchip technology—has promised to revolutionize these analyses even further. These new chip- or cartridge-based configurations of CE-SDS (also IEF) enable the fully automatic separation of a larger number of samples and improve the ease of use even further. They achieve even faster and higher throughput separations, while maintaining the advantages of CE-SDS over SDS-PAGE.

### *Asymmetric Field Flow Field Fractionation*

Protein aggregation is a process that can produce species spanning a vast size-range, stretching from a few nanometers (oligomeric species) to hundreds of micrometers (visible particles). Measuring the various species across this continuum is impossible using one single technique. Whereas the methods described above (section "Electrophoresis") are limited by the corresponding separation matrices (columns, gels) and can measure protein aggregates up to a certain range (depending on the protein size typically large oligomeric species), for quantifying larger protein species the application of other techniques is required. One of the best techniques for measuring high order protein aggregates is Asymmetric Field Flow Fractionation (AF4). Although this technique was discovered in the 1960s and its application for separation of proteins developed in the 1980s, it has only gained popularity in the last decade.

AF4 is based on the migration of analytes in a mobile phase flowing through a channel with a semi-permeable bottom wall. During the separation, as the analytes advance through the channel they are subjected to an asymmetric field, generated by the application of a flow perpendicular to the sample flow. This leads to the differential migration of the analytes—smaller species eluting faster due to their faster lateral diffusion and larger species eluting slower. Thus, a separation of aggregates of various sizes is achieved without using a stationary phase. The lack of a stationary phase is an advantage as it eliminates the filter effect of columns, column frits and gels, which often leads to exclusion of the large aggregate/particle species from separation altogether. A second advantage is its very wide size range of separation—it can separate aggregates ranging from several nanometers to hundreds of nanometers and even micrometers.

### *Techniques for Measuring Sub-Visible Particles*

Protein aggregate species of tens of nanometers and larger are commonly termed "sub-viosible particles" (SvP). Because the sub-visible particles are a critical quality attribute of protein therapeutics (cf. Chap. 7), their accurate and precise measurements are of high importance for the development of biotherapeutics. Due to the broad size-range span of these species, the simultaneous application of several techniques is required in order to measure all applicable species (See Table 3.3).

The traditional and "gold standard" method for measuring particles in the micrometer size-range is Light Obscuration (LO). This method uses a flow cell through which the sample is led. A Laser illuminates the flow cell. A particle passing with the liquid flow casts a shadow over the photodiode detector, which is registered and quantified via the resulting current drop.

A newer technique is Flow Imaging Microscopy (FIM). In this technique, instead of photodiode detector, a high-speed camera is used to capture the images of all individual particles imaged via a microscope. This invention allows for studying the morphology of the particles detected and potentially provides the option to draw conclusions about their composition and origin.

Other methods for particle characterization include some newly emerged techniques, such as Resonance Mass Measurement (RMM) or Nano Tracking Analysis (NTA). Both of these techniques allow measuring the concentrations of sub-micrometer particles, which is their major application. Nanotracking Analysis uses single particle tracking to calculate the diffusion coefficient of each individual particle and in turn—its size. The concentrations of particles in solution are then inferred from the small subset measured. Due to the unique capability of this technique to measure particles >30 nm it finds extensive use in vaccine development and recombinant virus characterization. RMM also offers some unique features, namely the ability to distinguish between particles with different densities. The latter is very useful in discriminating proteinaceous particles from silicone oil droplets (often present in biopharmaceutical drug products in pre-filled syringes or cartridges), for example.

The availability of the various SvP methods allows for coverage of the entire particle size-range. However, as with all analytical methodologies, an important consideration when using different SvP characterization methods in parallel is to recognize the specific advantages and shortcomings that apply to each of them.

| | Method principle and data analysis | Size (μm) | | | | | | | | | | | | Optimal sample concentration (part./ mL) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 0.03 | 0.05 | 0.20 | 0.30 | 0.50 | 0.60 | 0.80 | 1.00 | 2.00 | 5.00 | 10.00 | 25.00 | |
| **NTA** | ***Tracking of Brownian motion of individual particles:*** Hypothetical hard spheres that diffuse at the same speed of the tracked particles are assumed. The hydrodynamic diameter is obtained according to the 2D-modified Stokes-Einstein equation. For count determinations the averaged particle abundance (average number of particles per frame) is divided by the estimated volume of the sample chamber. | | | | | | | | | | | | | $3x10^8$-$1x10^9$<br><br>~20-70 centers per frame |
| **RMM** | ***Changes in frequency due to added mass:*** Shifts in frequency with respect to sensor baseline resonance are convert ed into buoyant mass using the sensor-specific sensitivity. Sensitivity is obtained using size standards as calibrators. Knowing the fluid's and particle's density, buoyant mass is converted into dry mass. Assuming a sphere shape, particle diameter is calculated. Concentration is obtained relating the number of events (particles) registered with the volume of sample dispensed | | | | | | | | | | | | | < $8x10^6$ |
| **EZS** | ***Changes in resistance due to volume displacement:*** The impedance pulses generated as particles are pumped through an orifice in a glass tube are individually analyzed by the instrument electronic components. As the electrical current is constrained in the aperture orifice each pulse is directly proportional to the volume that the particle displaced and its size. Concentration is obtained relating the number of events (particles) registered with the volume of sample dispensed | | | | | | | | | | | | | ~ $2x10^5$<br><br>Coincidence <5% |
| **FIM** | ***Image analysis of single particles:*** Digital images of the particles in the sample are captured and analyzed by the instrument software. Following background comparison, intensity values are assigned to each activated pixel. Adjoining pixels below 96% of the maximum brightness is grouped as particles. Internal algorithms are used to generate morphological descriptors per each particle. Concentration is obtained relating the number of events (particles) registered with the volume of sample dispensed. | | | | | | | | | | | | | < $9x10^4$ |
| **LO** | ***Drop in current due to light obscuration:*** A calibration curve size vs. voltage is defined using calibration size standards. Particle size is obtained by direct interpolation in the calibration curve of the voltage recorded when a particle blocks the sensor. Concentration is obtained relating the number of events (particles) registered with the volume of sample dispensed. | | | | | | | | | | | | | < $1.8x10^4$ |

Reproduced with modifications from Ríos Quiroz et al. (2016)
*NTA* nano tracking analysis, *RMM* resonance mass measurement, *EZS* electrical zone sensing, *FIM* flow imaging microscopy, *LO* light obscuration

**Table 3.3** ■ Analytical techniques commonly used for measuring and characterizing sub-visible particles

## Protein Modifications Introducing Changes in Hydrophobicity

### *Protein Oxidation*

Oxidation is a common degradative pathway for proteins. It often has profound effects on their physicochemical properties. Such major property changes may in turn result in alteration of the biological functions of the affected protein, such as loss of binding, reduction of enzymatic activity, unexpectedly rapid clearance. Thus, monitoring protein oxidation is very critical for the successful development of biopharmaceuticals.

Protein oxidation may occur during all stages of protein manufacturing, processing and storage, whenever the proteins may be exposed to oxidative agents. The latter may include peroxides, transition metal ions, exposure to light, etc.

Whereas theoretically all amino acids can be oxidized, in practice the most commonly oxidized amino acid residues are Trp, Met, Tyr, His, Phe and Cys.

Tryptophan residues are particularly susceptible to oxidation due to the relatively high reactivity of the aromatic indole with reactive oxygen species. Tryptophan oxidation typically requires some level of exposure of the Trp residues, which a commonly buried in the three-dimensional structure of proteins. However, when oxidized, tryptophan residues may convert to a large variety of products (see Figs. 3.11 and 3.12), all of which with properties very different from the original Trp. The most common pathway for Trp oxidation includes the formation of N-formylkinurenine.

Another commonly oxidized amino acid is methionine. The sulfur atom in the Met residue can accept either one or two oxygen atoms leading to the formation of sulfoxide or sulfone, correspondingly (see Fig. 3.12). Due to the typically high surface exposure of Met, this modification is relatively common.

## Measuring Changes in Protein Hydrophobicity

Most oxidative modifications of proteins result in some changes of the polarity of the affected residues. In the Met and Trp oxidation examples shown here, the resulting products differ from the original residues by their relative hydrophobicity. Thus, protein oxidation is commonly detected and quantified with analytical methods utilizing polarity-based separation. The most common technique using this separation principle is reversed-phase chromatography.

### *Reversed-Phase High-Performance Liquid Chromatography*

Reversed-phase high-performance liquid chromatography (RP-HPLC) takes advantage of the hydrophobic properties of proteins. The functional groups on the column matrix may contain from one to up to 18 carbon atoms in a hydrocarbon chain. The longer this chain, the more hydrophobic is the matrix. The hydrophobic patches of proteins interact with the hydrophobic chromatographic matrix. Proteins are then eluted from the matrix by increasing the hydrophobic nature of the solvent passing through the column. Acetonitrile is a solvent commonly used, although other organic solvents such as ethanol also may be employed. The solvent is made acidic by the addition of trifluoroacetic acid, since proteins have increased solubility at pH values further removed from their pI. A gradient with increasing concentration of hydrophobic solvent is passed through the column. Different proteins have different hydrophobicities and are eluted from the column depending on the "hydrophobic potential" of the solvent.

This technique can be very powerful. It may detect the addition of a single oxygen atom to the protein, as is the case when a methionyl residue is oxidized or when the hydrolysis of an amide moiety on a glutamyl or asparaginyl residue occurs. Disulfide bond formation or shuffling also changes the hydrophobic characteristic of the protein. Hence, RP-HPLC can be used not only to assess the homogeneity of the protein but also to follow degradation pathways occurring during long-term storage.

RP-HPLC does not always provide sufficient resolution for separation of oxidized species of an intact protein, particularly when larger and more complex proteins are analyzed (such as monoclonal antibodies). In such cases various methods can be applied to solve this problem. For example, the intact protein can be digested into subdomains, such as the Fab and Fc fragments in the case of a MAB, or even smaller fragments (cf. Chap. 8). This latter approach typically employs more frequently-cutting enzymes, such as trypsin, or Lys-C, in order to generate a large number of small peptide fragments, which can be better separated on a RP-HPLC from their oxidized isoforms (see Fig. 3.13).

Such RP-HPLC separations of proteolytic digests of recombinant proteins typically yield complex and unique separation patterns ("peptide maps"), which are often used as a method to identify a protein. Several different proteases, such as trypsin, chymotrypsin, and other endoproteinases, are used for these identity tests (see below under section "Mass Spectrometry").

### *Hydrophobic Interaction Chromatography*

A companion to RP-HPLC is hydrophobic interaction chromatography (HIC). In principle, this latter method is normal-phase chromatography, i.e., here an aqueous solvent system rather than an organic one is used to fractionate proteins. The hydrophobic characteristics of the solution are modulated by inorganic salt concentrations. Ammonium sulfate and sodium chloride are often used,

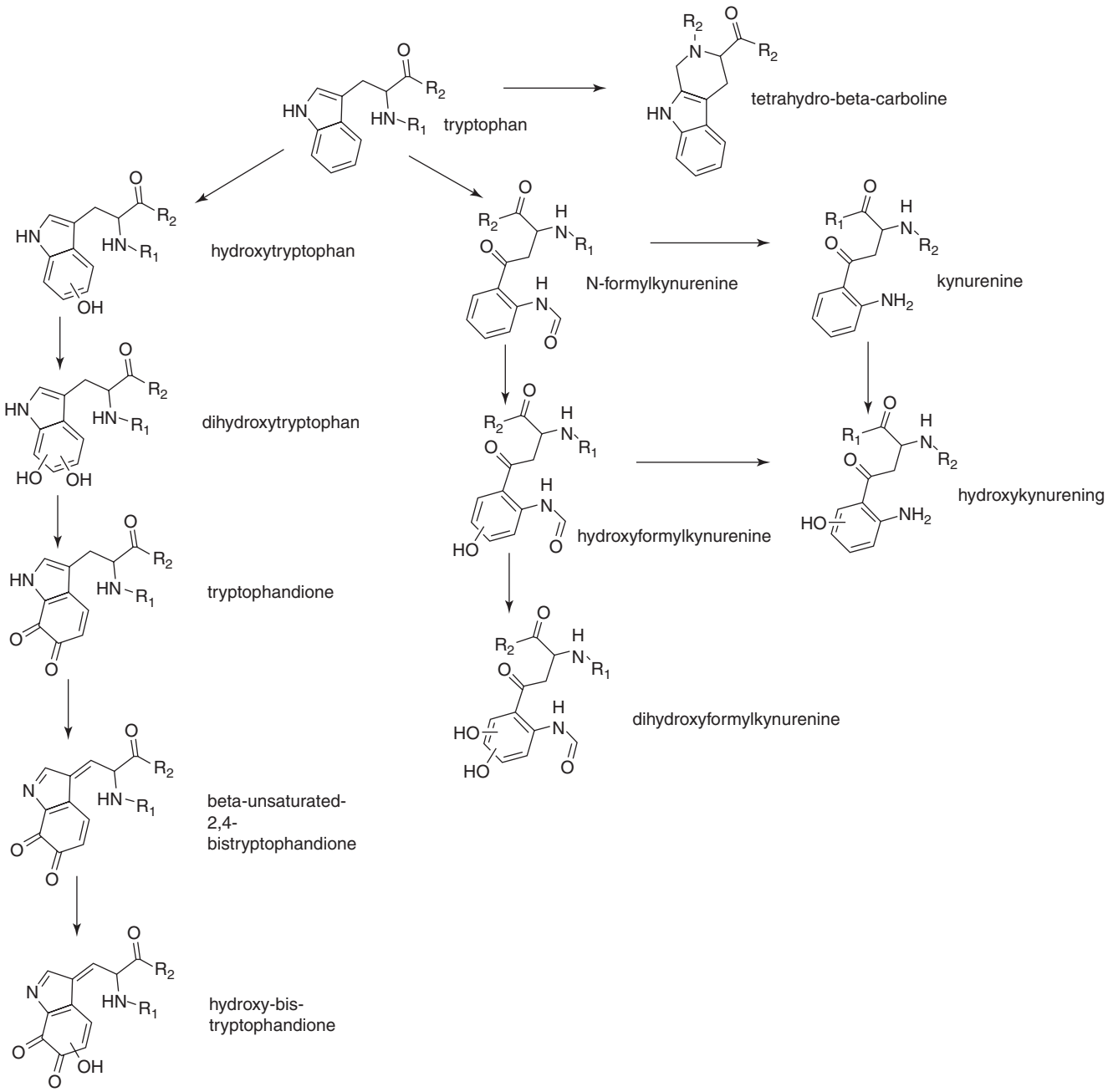


**Figure 3.11** ■ Possible oxidation products of tryptophan

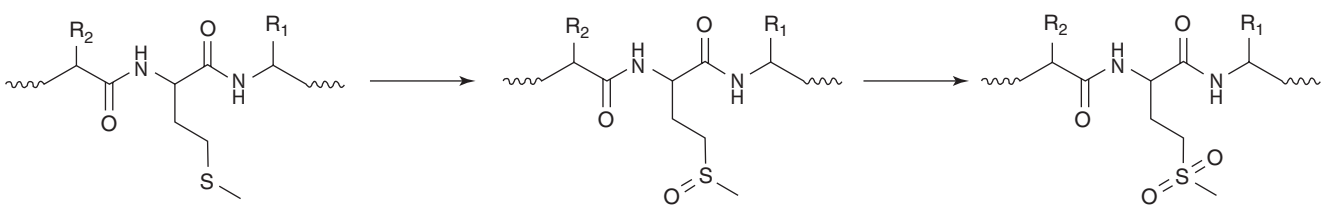


**Figure 3.12** ■ Oxidation of a methionine-containing peptide

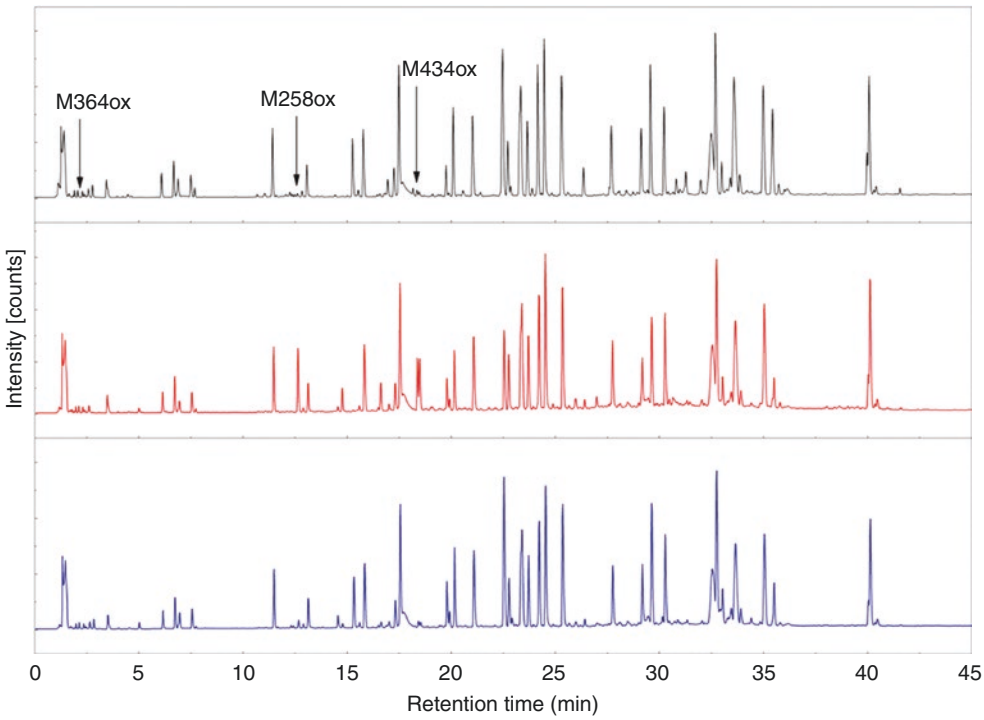


**Figure 3.13** ■ A RP-HPLC peptide map of a MAB showing the oxidation of individual methionine residues (indicated by arrows). Top trace shows the peptide map of a non-oxidized protein sample, middle trace—sample oxidized using $H_2O_2$ and bottom—sample to which $H_2O_2$ and an antioxidant have been added

since these compounds are highly soluble in water. In the presence of high salt concentrations (up to several molar), proteins are attracted to hydrophobic surfaces on the matrix of resins used in this technique. As the salt concentration decreases, proteins have less affinity for the matrix and eventually elute from the column. This method lacks the resolving power of RP-HPLC, but is gentler, since low pH values or organic solvents as used in RP-HPLC can be detrimental to some proteins.

## ■ Two-Dimensional (Hyphenated) Techniques

Some analytical techniques can be combined (hyphenated) to achieve additional functionality. Two prominent examples are discussed below.

### *2-Dimensional Gel Electrophoresis and Differential Gel Electrophoresis*

Isoelectric focusing and SDS-PAGE can be combined into a procedure called 2-D gel electrophoresis. Briefly, proteins are first fractionated by isoelectric focusing based upon their pI values. They are then subjected to SDS-PAGE run perpendicular to the first dimension and fractionated based on the molecular weights of the proteins. These separations produce a gel on which each protein appears as a separate spot, corresponding to a specific molecular weight and pI value combination. This setup allows for separating very complex protein mixtures (e.g. extracted from cells or tissues) and is commonly used in the proteomic field (cf. Chap. 9). Another situation where 2D Gel Electrophoresis is regularly applied concerns profiling of host cell proteins.

Another 2-dimensional gel technique is Differential Gel Electrophoresis (DIGE). This technique is essentially 2D Gel Electrophoresis where 2 or 3 different samples are separated simultaneously. The proteins in the samples are labeled with differently colored fluorescent dyes (typically Cy2, Cy3 and Cy5, which are charge- and mass- matched). The proteins co-migrate on the gel and are typically detected simultaneously using a multi-channel scanner or a camera. The overlay of the different channels allows for identifying/visualization of individual proteins being over- or underrepresented in the different samples, which is otherwise very difficult to find out with complex protein mixtures.

### *2-Dimensional Chromatography*

One important strategy for improving the selectivity (specificity) of chromatographic separations is the coupling of two or more columns—for example, an ion exchange column, directly followed by a reversed phase column. This strategy allows for the separation of highly complex analyte mixtures, such as the mixtures of peptides generated during shotgun proteomic experiments. In these experiments multiprotein (up to thousands of proteins) mixtures are digested using a protease (typically trypsin) to produce an even richer mixture of peptides. These peptide mixtures are separated on a 2D chromatographic system in order to reduce the complexity

and are commonly analyzed using an MS as a detector. MS/MS fragmentation of the individual peptides allows for the sequencing and identification of each peptide and correspondingly—protein, thus permitting the semi-quantitative analysis of the original protein mixtures. Such approaches are very commonly used in host cell protein (HCP) profiling, and biomarker research.

## MODIFICATIONS TO THE HIGHER ORDER STRUCTURE OF PROTEINS

All levels of protein structural organization can be altered as a consequence of physical or chemical stress. These alterations can be manifested in a large diversity of protein modifications, each changing the physicochemical and biological properties of the protein. See Table 3.1.

In addition to covalent modifications (modifications to the primary structure/amino-acid sequence) described in earlier sections, the higher order structure of proteins can undergo changes as well. Such changes can be relatively minor, such as alterations of the quaternary structure (subunit configuration) of a protein complex due to an incorrect disulfide bridge, or more substantial like perturbation of the tertiary structure of a given domain. Some of these possible modifications are described in the following section.

### ■ Measuring Changes in Higher Order Structure of Proteins

A large number of analytical techniques can be used to measure the structural organization of proteins. However, all these methods differ significantly from each other by the level of information content they provide and the ease of use. Typically, the most accessible and easy to use methods provide a relatively low information content, whereas higher resolution methods (providing specific information about the structure of separate domains and even functional groups and atoms) require highly specialized and expensive equipment and dedicated, highly trained specialized personnel.

#### *Lower Resolution Techniques: CD, FTIR, Fluorescence*

Most of the spectroscopic techniques used for characterization of the higher order protein structure (See section "Spectroscopic and Other Techniques for Studying Higher Order Structure") are relatively easy to use, although interpretation of the experimental results typically requires expert knowledge. Very often in biotech development spectroscopic measurements such as CD, FTIR or fluorescence spectroscopy are applied in order to compare protein therapeutic products from different manufacturing batches (typically after changes introduced into the manufacturing process), asking the specific question whether significant differences (alterations of the secondary or tertiary structure) are present between the different batches. In these, so called "comparability studies" (cf. Chap. 12) the first goal is to identify if such changes are present at all. To answer this question, it is sufficient to overlay the CD, FTIR or fluorescence spectra from the different batches and look for any differences. However, one common downside of CD, FTIR and fluorescence spectroscopy is the fact that if differences are seen, it is difficult to judge to which region of the molecule these differences are related. The reason for this is that all of the above-mentioned techniques provide a summary/population information for all of the spectroscopically active functional groups in the molecule (chromophores in the case of CD, fluorophores in the case of fluorescence and amide absorption bands in case of FTIR) and do not provide specific spacial information for individual groups. This means that using the results from such experiments one cannot pinpoint where exactly in the structure of the molecule the detected changes are located. To answer that question, additional analytical work employing higher resolution techniques (see below) is required.

One additional complication resulting from the fact that the methods mentioned above measure the overall molecular population present in the test solution is the fact that the limit of detection (LOD) of specific structural changes is relatively high. More specifically, if a given structural modification has occurred only within a small portion of the overall population (for example, let us say in only 5% of the molecules), this change is unlikely to give sufficiently strong spectral signals to modify the overall (summary) spectrum collected in the experiment. Thus, the techniques described above are typically useful for the detection of gross modifications of the secondary, tertiary (CD, FTIR and fluorescence) and quaternary (fluorescence) structure.

#### *Higher Resolution Techniques*

In contrast to the spectroscopic methods described in the previous section, some techniques are capable of providing specific spatial information for specific domains, functional groups or even individual atoms in proteins. The degree of structural detail available varies from method to method and typically, the higher the information content (structural details) —the more complex and specialized the method.

An increasingly popular technique for higher order structures (HOS) determination is HDX-MS (see section "Spectroscopic and Other Techniques for Studying Higher Order Structure"). This method typically provides much higher spatial resolution than the spectroscopic techniques mentioned above, although not as high as X-ray crystallography or NMR

(see below). What is typically achieved using HDX-MS is at peptide–level information, allowing to map individual sub-domains according to their mobility and solvent accessibility. These maps (using 3D molecular models) can be extremely useful in understanding structural alterations limited to small regions of the protein. Compared to the higher resolution methods described below, HDX-MS is more accessible (both in terms of instrumentation and also expertise) and it is not limited by the protein size, crystallographic properties or required sample amounts, which are some of the downsides for NMR and X-ray crystallography.

X-ray crystallography is the ultimate spatial resolution method. Using this technique, it is quite common to determine protein structures at an atomic resolution. This technique is indispensable in research focused on enzymes or specific protein binding sites. A number of X-ray crystallographic structure analyses of Ab-Ag complexes, for example, or drug-target molecule complexes have been very illuminating and were critical with respect to advancing drug discovery. In drug development this technique typically does not find as broad use, due to the huge efforts required. Despite advances in this field, including the use of robotics and machine learning approaches, it is not uncommon to take year(s) for solving a given structure, not to mention the fact that solving some structures is currently impossible due to the fact that some proteins are exceedingly difficult to crystalize.

Another very high resolution technique is NMR (See section "Spectroscopic and Other Techniques for Studying Higher Order Structure"). One very significant advantage that this method offers is the fact that experiments can be carried out in solution, meaning that often important aspects of the protein dynamics can be interrogated, a feature not available with other high resolution techniques. Until recently, major limitations of this technique came from the need to label the proteins prior to analyses using stable isotopes and also the size limit to the proteins analyzed. More recent advances in the protein NMR field led to the possibility to obtain 2D $^{13}C$ NMR methyl fingerprint data for structural mapping of an intact MAB at natural isotopic abundance (Arbogast et al. 2015) which significantly ameliorated the shortcomings mentioned above.

One technique which has undergone a development boom in the twenty-first century is Cryo Electron Microscopy (CryoEM). This technique also allows looking at native structures and complexes in some cases to near-atomic resolution. The huge impact of this technique on biological research and studies of protein complexes was recognized in 2017, when the co-discoverers of the method received the Nobel Prize for Chemistry (Henderson 2018).

## BIOLOGICAL ACTIVITY (POTENCY) ASSAYS/ BIOASSAYS

### Binding Assays (ELISA), Surface Plasmon Resonance (SPR)

*Immunoassays*

Enzyme-linked immunosorbent assay (ELISA) provides a means to quantitatively measure extremely small amounts of proteins. This procedure utilizes the fact that plastic surfaces are able to adsorb low but detectable amounts of proteins. Typically, antibodies against a certain protein of interest are allowed to adsorb to the surface of microtitration plates. Each plate may contain up to 96 or 384 wells so that multiple samples can be assayed. After incubating the antibodies in the wells of the plate for a specific period, excess antibody is removed and residual protein binding sites on the plastic are blocked by incubation with an inert protein. Several microtitration plates can be prepared at one time since the antibodies coating the plates retain their binding capacity for an extended period. During the ELISA, the sample solution containing the protein of interest is incubated in the wells and the protein (Ag) is captured by the antibodies coating the well surface. Excess sample is removed and other antibodies which now have an enzyme (E) linked to them are added to react with the bound antigen.

The format described above is called a sandwich assay since the antigen (protein of interest) is located between the antibody on the titer well surface and the antibody containing the linked enzyme. Figure 3.14 illustrates a number of formats that can be used in an ELISA. A suitable substrate is added and the enzyme linked to the antibody-antigen-antibody well complex converts this compound to a colored product. The amount of product obtained is proportional to the enzyme adsorbed in the well of the plate. A standard curve can be prepared if known concentrations of antigen (protein of interest) are tested in this system and the amount of antigen in unknown samples can be estimated from this standard curve. A number of enzymes can be used in ELISAs. However, the most common ones are horseradish peroxidase and alkaline phosphatase. A variety of substrates for each enzyme is available; they yield colored products catalyzed by the antibody-linked enzyme. Absorbance of the colored product solutions is measured on plate readers, instruments that rapidly measure the absorbance in all 96 wells of the microtitration plate. Data processing can be automated for rapid throughput of information. Note that detection approaches partly parallel those discussed in the section on "Blotting." The above ELISA format (sandwich assay) is only one of many different ELISA set ups. For example, the microtitration wells may be coated directly with the antigen

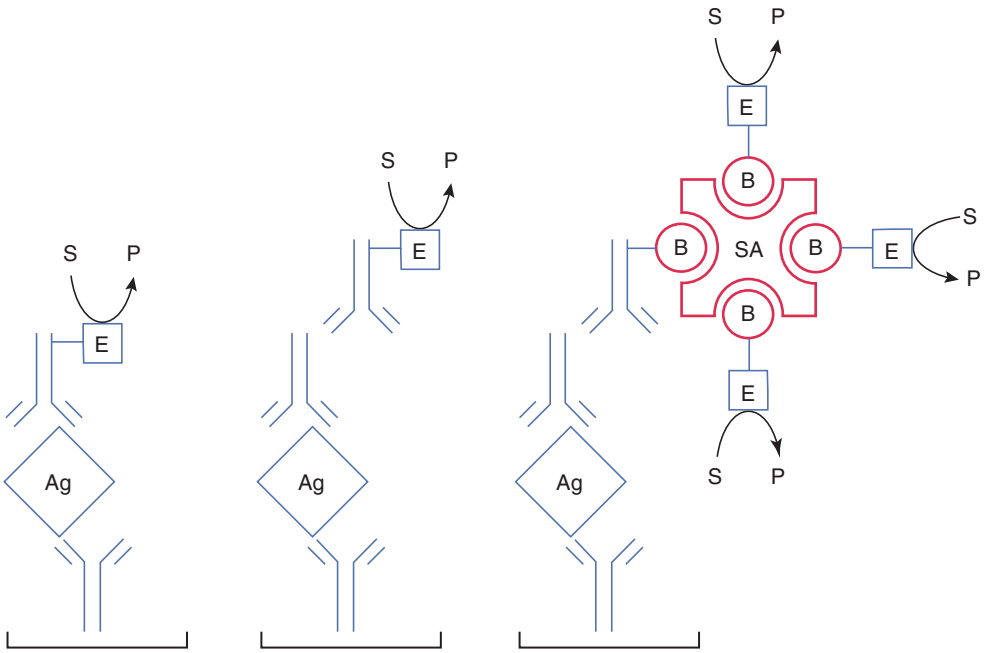


**Figure 3.14** ■ Examples of several formats for ELISA in which the specific antibody is adsorbed to the surface of a microtitration plate. See Fig. 3.9 for abbreviations used. The antibody is represented by the Y-type structure. The product *P* is colored and the amount generated is measured with a spectrometer or plate reader

rather than having a specific antibody attached to the surface. The concentration of antigen is established by comparison with known quantities of antigen (protein of interest) used to coat individual wells.

Another approach is to use—this time subsequent to the binding of the antigen (protein of interest) either directly to the surface or to an antibody on the surface—an antibody specific to the antibody binding the protein antigen, that is, a secondary antibody (see section "Blotting Techniques"; Fig. 3.9). This latter, secondary, antibody contains the linked enzyme used for detection. As already discussed in the section on "Blotting," the advantage of this approach is that such antibodies can be obtained in high purity and with the desired enzyme linked to them from commercial sources. Thus, a single source of enzyme-linked antibody can be used in assays for different protein antigens. Should a sandwich assay be used, then antibodies from different species need to be used for each side of the sandwich. A possible scenario is that rabbit antibodies are used to coat the microtitration wells; mouse antibodies, possibly a monoclonal antibody, are used to complex with the antigen; and then, a goat anti-mouse immunoglobulin containing linked HRP or AP is used for detection purposes.

As with immunoblots discussed above, streptavidin or avidin can be used in these assays if biotin is covalently linked to the antibodies and enzymes.

### *SPR-Based Binding Assays*

Surface plasmon resonance techniques are based on the excitation of free electrons (called surface plasmons when excited) by polarized light from a metal film at an interface with a medium having a different refractive index. Binding of molecules to this interfacial layer results in shifts in their reflection curves. Since the refractive index changes are linearly proportional to the number of molecules bound, this technique can be used to calculate a number of binding parameters such as the equilibrium association constant ($K_A$), equilibrium dissociation constant ($K_D$), as well as the concentration of a protein in solution. In practice, these measurements are carried out using commercially available SPR chips, which are typically covalently derivatized with antibodies or antigens to which the protein of interest can bind. The solution with the protein of interest flows over the chip at a defined rate and from the SPR signals the characteristics mentioned above can be calculated.

Today, both ELISA- and SPR- based binding assays are extensively used in the development of biopharmaceuticals. Although neither method has undergone fundamental changes over the last decade or so, one recent improvement which has been broadly implemented is the automation of these assays. Since these techniques typically require a significant hands-on time for analysis, recent advances as the introduction of robotic systems have increased the throughput significantly.

## ■ Cell-Based Activity Assays

Paramount to the development of a protein therapeutic is to have an assay that identifies its biological function. Chromatographic and electrophoretic methodologies can address the purity of a biotherapeutic and be useful in investigating stability parameters. However,

it is also essential to ascertain whether the therapeutic protein has adequate bioactivity. Typically, bioassays (potency assays) are carried out in vitro by monitoring the biological response in cells when the therapeutic protein is added to the system. These biological responces need to reflect the mode of action (MoA) of the therapeutic. There is a wide and ever-increasing variety of cell-based assays.

Common cell-based bioassays are:

(a) Cell proliferation/anti-proliferation assays in which the proliferation (or reduction of proliferation) of the cells in culture is measured as a response to the drug,
(b) Cytotoxicity assays in which cell death occurring as a response to the drug is measured,
(c) Adhesion assays in which the influence of the drug on the adhesion properties of the cells are measured,
(d) Kinase receptor activation assays in which the phosphorylation of a tyrosine kinase as a response to the drug as a ligand is measured by capture-ELISA after cell lysis,
(e) Cellular response to the biopharmaceutical is often monitored via the activation of a specific cellular signaling cascade,
(f) Antibody-Dependent Cell-mediated Cytotoxicity (ADCC) assays are assays in which cell lysis by immune system effector cells upon their activation by a therapeutic antibody bound to a receptor of the target cell is measured,
(g) Complement-Dependent Cytotoxicity (CDC) assays are assays in which the lysis of the target cell by components of the complement system are measured after complement activation by the therapeutic antibody bound to a cellular receptor,
(h) Cytokine release assays measure the release of cytokines by the target cell as a response to the protein therapeutic.

An interesting new approach to cell-based potency assays is the use of reporter genes. In the reporter gene assays the activation of a gene regulatory element (as a response to a signaling cascade activation) is monitored using a common reporter gene, for example luciferase. The assay readout is the expression of the reporter gene as response to the drug. The advantage of this approach is the potential for using it as a generic (platform) approach to a number of different bioassays, thus simplifying the bioassay development, cell banking, etc.

## CONCLUDING REMARKS

With the ever-increasing application and variety of human recombinant proteins as therapeutics, the need for characterization of their structure, function, and purity has also increased. Naturally, the analytical technology has also undergone a rapid evolution. Today, a large array of techniques is used to characterize the primary, secondary, tertiary and quaternary structure of proteins and to determine the quality, purity, and stability of the recombinant drug product.

The information provided in this chapter offers only a general guidance on the process and tools for analytical characterization of various protein modifications. In reality, this process is rarely straightforward and is often convoluted by a number of factors. For example, a lot of different protein modifications occur simultaneously, thus inducing simultaneous changes in a number of molecular attributes. Often the results of these simultaneous changes may be masked. For example, it is possible that no changes in the IEX charge heterogeneity profile of a protein molecule are detected when simultaneous succinimide formation (inducing the formation of more basic species) and deamidation (inducing the formation of acidic species) occur, in spite of a significant redistribution of various charged species.

Many protein modifications result in changes in more than one molecular attribute. Thus, more than one analytical technique should be used to characterize these modifications. For example, pyro-Glu formation may be measured using either charge based techniques (e.g. IEX) or techniques measuring changes in polarity (e.g. RP-HPLC). Moreover, the extent to which the molecular properties of a protein are modified as a result of a given modification, depends on the structural context in which this modification occurs (overall molecular charge, hydrophobicity, size, etc.). For example, clipping of a small fragment of a very large protein may not be easy to detect using size-exclusion chromatography, due to the resolution limits of this technique.

Because of the complexity of the interplay between various potential protein modifications, the definition of the characterization and quality control strategies is an important intellectual challenge for the analytical experts in the therapeutic protein development teams, and critical for the success of every development program.

## SELF-ASSESSMENT QUESTIONS

### ■ Questions

1. What are the most common chemical modifications resulting in changes in protein charge heterogeneity?
2. What are the three most common techniques for measuring charge heterogeneity of proteins?
3. What is the transfer of proteins to a membrane such as nitrocellulose or PDVF called?
4. What are the two most commonly oxidized amino acids in protein biotherapeutics? How can one

detect these oxidized amino acids in molecular structure?
5. In a 2-dimensional electrophoresis analysis of a protein mixture the first method of separation is an IEF run, followed by a SDS-PAGE run in a perpendicular direction to the first run? If one runs the SDS-PAGE analysis first, followed by the IEF run, would one get a similar result?
6. List three techniques for separating proteins based on molecular size.
7. Which technique provides the ultimate (atomic) resolution of the structure of a protein?
8. What are limitations of an ELISA when analyzing a protein product?

## ■ Answers

1. Deamidation and isomerization.
2. Ion exchange chromatography, Isoelectric focusing, capillary zone electrophoresis.
3. This method is called blotting. If an electric current is used, then the method is called electroblotting.
4. Methionine and tryptophan. MS, RP-HPLC and HIC are preferred analytical techniques to detect oxidation of proteins.
5. No. In the SDS-PAGE analysis the protein is denatured (unfolds) and SDS interacts with the protein giving it a uniform negative charge that masks the amino acids charges in the protein. The pI of the protein(s) cannot be determined in a subsequent IEF run.
6. Size-exclusion chromatography, SDS-PAGE, Asymmetric Field Flow Field Fractionation (AFFF).
7. X-ray crystallography.
8. The ELISA (sandwich assay set up) may measure degradation products and/or product-related variants with similar affinity to intact molecules.

**Acknowledgements** This chapter is a re-work of part (including a number of figures) of the chapter "*Biophysical and Biochemical Analysis of Recombinant Proteins*", by Tsutomu Arakawa and John Philo, which appeared in the previous editions of this book—cf. fourth edition, 2013. Despite the significant changes made to the original chapter, it still contains fragments of the excellent original text by Arakawa and Philo.

The author thanks Abbas Razvi and Marigone Lenjani for providing chromatograms and electropherograms shown in this chapter.

## FURTHER READING

Arbogast LW, Brinson RG, Marino JP (2015) Mapping monoclonal antibody structure by 2D 13C NMR at natural abundance. Anal Chem 87(7):3556–3561

Berkowitz SA, Engen JR, Mazzeo JR, Jones GB (2012) Analytical tools for characterizing biopharmaceuticals and the implications for biosimilars. Nat Rev Drug Discov 11(7):527–540

Butler JE (ed) (1991) Immunochemistry of solid-phase immunoassay. CRC Press, Boca Raton

Cavanagh J, Fairbrother WJ, Skelton NJ (2007) Protein NMR spectroscopy: principles and practice book, 2nd edn. Academic Press, San Diego

Chirino AJ, Mire-Sluis A (2004) Characterizing biological products and assessing comparability following manufacturing changes. Nat Biotechnol 22(11):1383–1391

Coligan J, Dunn B, Ploegh H, Speicher D, Wingfield P (eds) (1995) Current protocols in protein science. Wiley, New York

Crabb JW (ed) (1995) Techniques in protein chemistry VI. Academic, San Diego

Domon B, Aebersold R (2006) Mass spectrometry and protein analysis. Science 312(5771):212–217

Du Y, Walsh A, Ehrick R, Xu W, May K, Liu H (2012) Chromatographic analysis of the acidic and basic species of recombinant monoclonal antibodies. MAbs 4(5):578–585

Dunbar BS (1994) Protein blotting: a practical approach. Oxford University Press, New York

Fekete S, Guillarme D, Sandra P, Sandra K (2016) Chromatographic, electrophoretic, and mass spectrometric methods for the analytical characterization of protein biopharmaceuticals. Anal Chem 88(1):480–507

Folzer E, Diepold C, Bomans K, Finkler C, Schmidt R, Bulau P, Huwyler J, Mahler H-C, Koulov AV (2015) Selective oxidation of methionine and tryptophan residues in a therapeutic IgG1 molecule. J Pharm Sci 104:2824–2831

Gahoual R, Beck A, Leize-Wagner E, François YN (2016) Cutting-edge capillary electrophoresis characterization of monoclonal antibodies and related products. J Chromatogr B Analyt Technol Biomed Life Sci 1032:61–78

Hames BD, Rickwood D (eds) (1990) Gel electrophoresis of proteins: a practical approach, 2nd edn. IRL Press, New York

Henderson R (2018) From electron crystallography to single particle Cryo EM (Nobel lecture). Angew Chem Int Ed Engl 57(34):10804–10825

Jiskoot W, Crommelin DJA (eds) (2005) Methods for structural analysis of protein pharmaceuticals. AAPS Press, Arlington

Kahle J, Wätzig H (2018) Determination of protein charge variants with (imaged) capillary isoelectric focusing and capillary zone electrophoresis. Electrophoresis 39(20):2492–2511

Kaltashov IA, Bobst CE, Abzalimov RR, Wang G, Baykal B, Wang S (2012) Advances and challenges in analytical characterization of biotechnology products: mass spectrometry-based approaches to study properties and behavior of protein therapeutics. Biotechnol Adv 30(1):210–222

Liu H, Gaza-Bulseco G, Faldu D, Chumsae C, Sun J (2008) Heterogeneity of monoclonal antibodies. J Pharm Sci 97(7):2426–2447

McEwen CN, Larsen BS (eds) (1998) Mass spectrometry of biological materials, 2nd edn. Dekker, New York

Moritz B, Schnaible V, Kiessig S, Heyne A, Wild M, Finkler C, Christians S, Mueller K, Zhang L, Furuya K, Hassel M, Hamm M, Rustandi R, He Y, Solano OS, Whitmore C, Park SA, Hansen D, Santos M, Lies M (2015) Evaluation of capillary zone electrophoresis for charge heterogeneity testing of monoclonal antibodies. J Chromatogr B Analyt Technol Biomed Life Sci 983-984:101–110

Müllertz A, Perrie Y, Rades T (eds) (2016) Analytical techniques in the pharmaceutical sciences. Springer, New York

Pace CN, Grimsley GR, Scholtz JM, Shaw KL (2014) Protein stability. In: eLS. Wiley, Chichester

Parr MK, Montacir O, Montacir H (2016) Physicochemical characterization of biopharmaceuticals. J Pharm Biomed Anal 130:366–389

Ponniah G, Nowak C, Neill A, Liu H (2017) Characterization of charge variants of a monoclonal antibody using weak anion exchange chromatography at subunit levels. Anal Biochem 520:49–57

Rathore D, Faustino A, Schiel J, Pang E, Boyne M, Rogstad S (2018) The role of mass spectrometry in the characterization of biologic protein products. Expert Rev Proteomics 15(5):431–449

Ríos Quiroz A, Lamerz J, Da Cunha T, Boillon A, Adler M, Finkler C, Huwyler J, Schmidt R, Mahler H-M, Koulov AV (2016) Factors governing the precision of subvisible particle measurement methods—a case study with a low-concentration therapeutic protein product in a prefilled syringe. Pharm Res 33:450–461

Salas-Solano O, Felten O (2008) Capillary electrophoresis methods for pharmaceutical analysis. Sep Sci Technol 9:401

Shirley BA (ed) (1995) Protein stability and folding. Humana Press, Totowa

Strege MA, Lagu AL (eds) (2004) Capillary electrophoresis of proteins and peptides. Humana Press, Totowa

Thorpe R, Wadhwa M, Mire-Sluis A (1997) The use of bioassays for the characterisation and control of biological therapeutic products produced by biotechnology. Dev Biol Stand 91:79–88

Vlasak J, Ionescu R (2011) Fragmentation of monoclonal antibodies. MAbs 63:253–263

Vlasak J, Ionescu R (2008) Heterogeneity of monoclonal antibodies revealed by charge-sensitive methods. Curr Pharm Biotechnol 9:468–481

Wei H, Mo J, Tao L, Russell RJ, Tymiak AA, Chen G, Iacob RE, Engen JR (2014) Hydrogen/deuterium exchange mass spectrometry for probing higher order structure of protein therapeutics: methodology and applications. Drug Discov Today 19(1):95–102

Wild D (ed) (2013) The immunoassay handbook theory and applications of ligand binding, ELISA and related techniques, 4th edn. Elsevier, New York