Strategies for the Characterisation of Biopharmaceuticals

by Koen Sandra[1], Maureen Joseph[2] and Pat Sandra[1]*

[1] Research Institute for Chromatography (RIC) and Metablys, President Kennedypark 26, B-8500 Kortrijk, Belgium

[2] Agilent Technologies, 2850 Centerville Rd, Wilmington DE 19808, USA

Corresponding author: koen.sandra@richrom.com

Protein biopharmaceuticals such as monoclonal antibodies and recombinant proteins have emerged as important therapeutics for the treatment of life-threatening diseases including cancer and autoimmune diseases. Close to 250 products are approved for human use in the United States and the European Union accounting for a sales value of $ 140 billion [1,2]. In addition, more than 2,000 protein biopharmaceuticals are currently in clinical development [2]. Protein therapeutics are large, heterogeneous and subject to a variety of enzymatic and chemical modifications during recombinant expression, purification and long-term storage and, hence, have a complexity far exceeding that of small molecule drugs. Unravelling this complexity represents an immense analytical challenge. The present contribution, representing a further reflection of a presentation given at the ChromSoc meeting on Advances in Bio-Separations (Medimmune, Cambridge, 12-13 May, 2015), describes common chromatographic and mass spectrometric (MS) strategies for the structural characterisation of protein biopharmaceuticals.

Important characteristics such as amino acid sequence and composition, molecular weight and structural integrity, N- and O-glycosylation, N- and C-terminal processing, S–S bridges, free cysteine residues, deamidation (asparagine, glutamine), aspartate isomerisation, oxidation (methionine, tryptophan), clipping, sequence variants, charge variants, aggregation, as well as higher order structural information can be extracted out of the generated chromatographic and mass spectrometric data. This typically requires an assessment at different levels being the protein, peptide, amino acid and sugar level [3]. For that, a wide range of chromatographic modes are available tackling different physicochemical properties of the molecules under investigation. Figure 1 gives an overview of separations performed at the different levels providing access to the above described features. The mAb Herceptin, a blockbuster used in the treatment of HER2 positive breast cancer, is used for illustrative purposes. In assessing the characteristics, important roles are played by reversed phase liquid chromatography (RPLC), ion exchange chromatography (IEX), size exclusion chromatography (SEC), hydrophobic interaction chromatography (HIC) and hydrophilic interaction chromatography (HILIC) (3).

Protein level

A variety of the above described characteristics can be revealed at the protein level. With protein level it is meant the intact protein or large fragments thereof generated following reduction or enzymatic cleavage. In the case of mAbs, large fragments represent the light (Lc) and heavy chain (Hc) generated following reduction of the disulphide bonds, antigen binding (Fab) and crystallisable fragment (Fc) generated following papain cleavage, F(ab')2 and Fc/2 generated following pepsin or immunoglobulin-degrading enzyme of *Streptococcus pyogenes* (IdeS) cleavage. The rationale behind this fragmentation strategy is that molecules often become more amenable to chromatographic and mass spectrometric measurement and particular modifications can readily be traced back to specific domains.

One of the chromatographic modes used most often in characterisation workflows is RPLC. RPLC separates proteins based on hydrophobicity and two types of interactions contribute to the chromatographic process, solvophobic and electrostatic interactions. The latter is governed by an ion-pairing additive. When making use of widepore fully porous (sub 2 µm) or superficially porous particles in combination with elevated column temperatures (80°C) and trifluoroacetic acid (TFA) as ion-pairing reagent, highly efficient separations can be obtained on intact mAbs (±150 kDa), F(ab')2 (±100 kDa), Hc, Fab and Fc (±50 kDa) and Lc, Fc/2 and Fd' (±25 kDa) (Figure 1) [3,4]. Moreover, the conditions used are compatible with mass spectrometry allowing identification of the observed peaks. This is illustrated in Figure 2 which shows the LC-MS analysis of a reduced and non-reduced IdeZ digest of Herceptin. IdeZ or Immunoglobulin-degrading enzyme from *Streptococcus equi* ssp *zooepidemicus* is a highly specific protease similar to IdeS that cleaves mAbs at a single site below the hinge region, yielding F(ab')2 and Fc/2 fragments [5]. Following reduction using tris(2-carboxyethyl)phosphine (TCEP), the F(ab')2 fragment is converted into the Lc and Fd'. On the top one observes the UV chromatograms and at the bottom the deconvoluted quadrupole time-of-flight (Q-TOF) MS spectra associated with the different annotated peaks simultaneously acquired with the UV data. Peak a corresponds to the Fc/2 part of the mAb and the existence of several glycoforms with characteristic 146 Da and 162 Da spacings indicative of fucose and galactose and consequently N-glycosylation is highlighted. Four main glycoforms can be revealed containing the typical mammalian complex type N-glycans termed G0, G0F, G1F and G2F. Measuring glycosylation is of utmost importance for mAbs since this represents a critical quality attribute that can impart on product safety and efficacy. Note that mass deviations are below 10 ppm for all four glycoforms. In the reduced sample, a 4 Da mass increase is noticed on the Fc/2 glycoforms (peak c). This can be explained by the reduction of the 2 internal S-S bridges adding up to four protons. Peak d with a MW of 23,443.5 can be identified as the fully reduced light chain of Herceptin. The partially resolved post-peak e with a MW of 23,444.8 is only one Da higher and represents a deamidated variant of the Lc. It is remarkable that such subtle mass differences can be revealed.

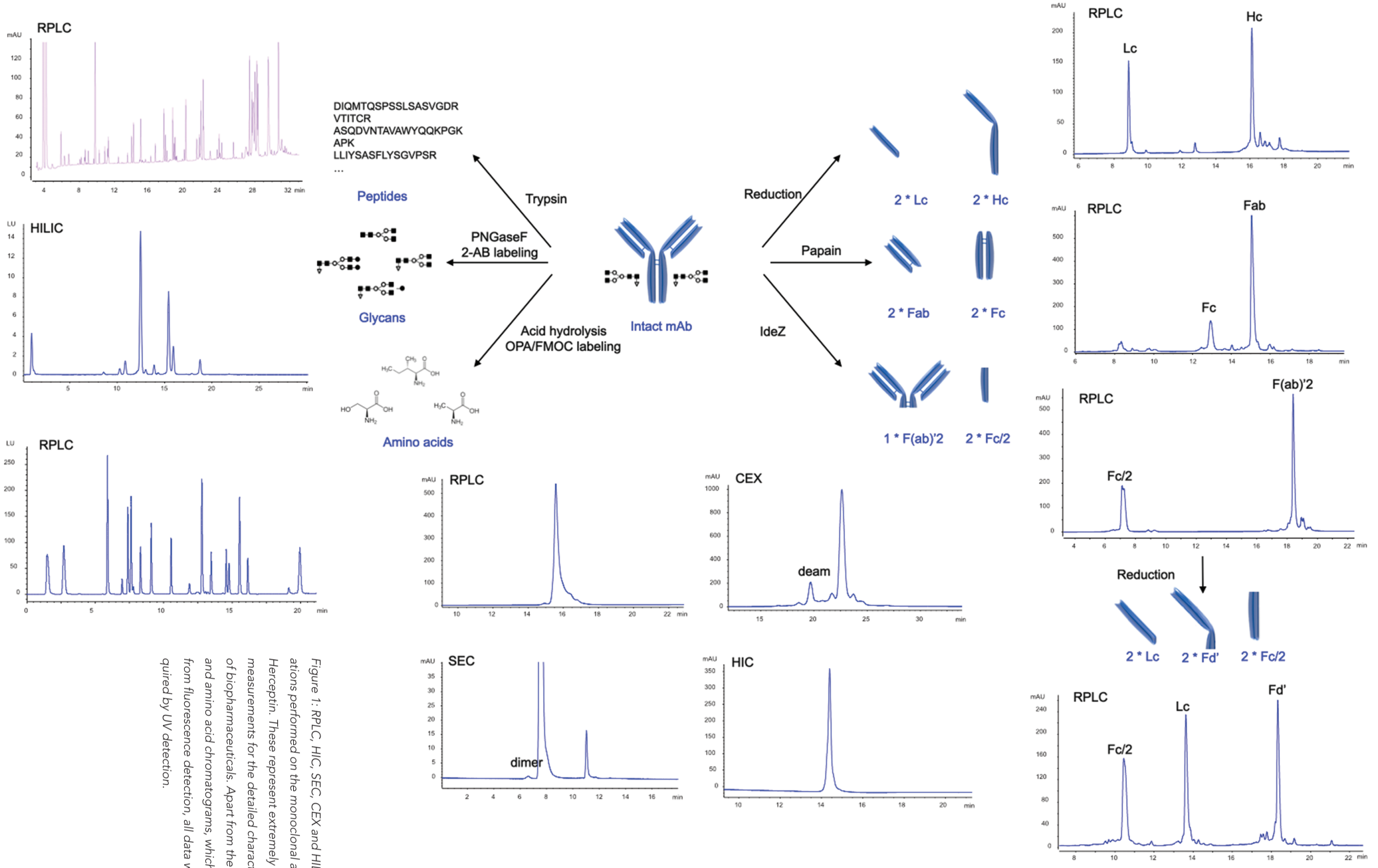

Figure 1: RPLC, HIC, SEC, CEX and HILIC separations performed on the monoclonal antibody Herceptin. These represent extremely powerful measurements for the detailed characterisation of biopharmaceuticals. Apart from the glycan and amino acid chromatograms, which result from fluorescence detection, all data were acquired by UV detection.

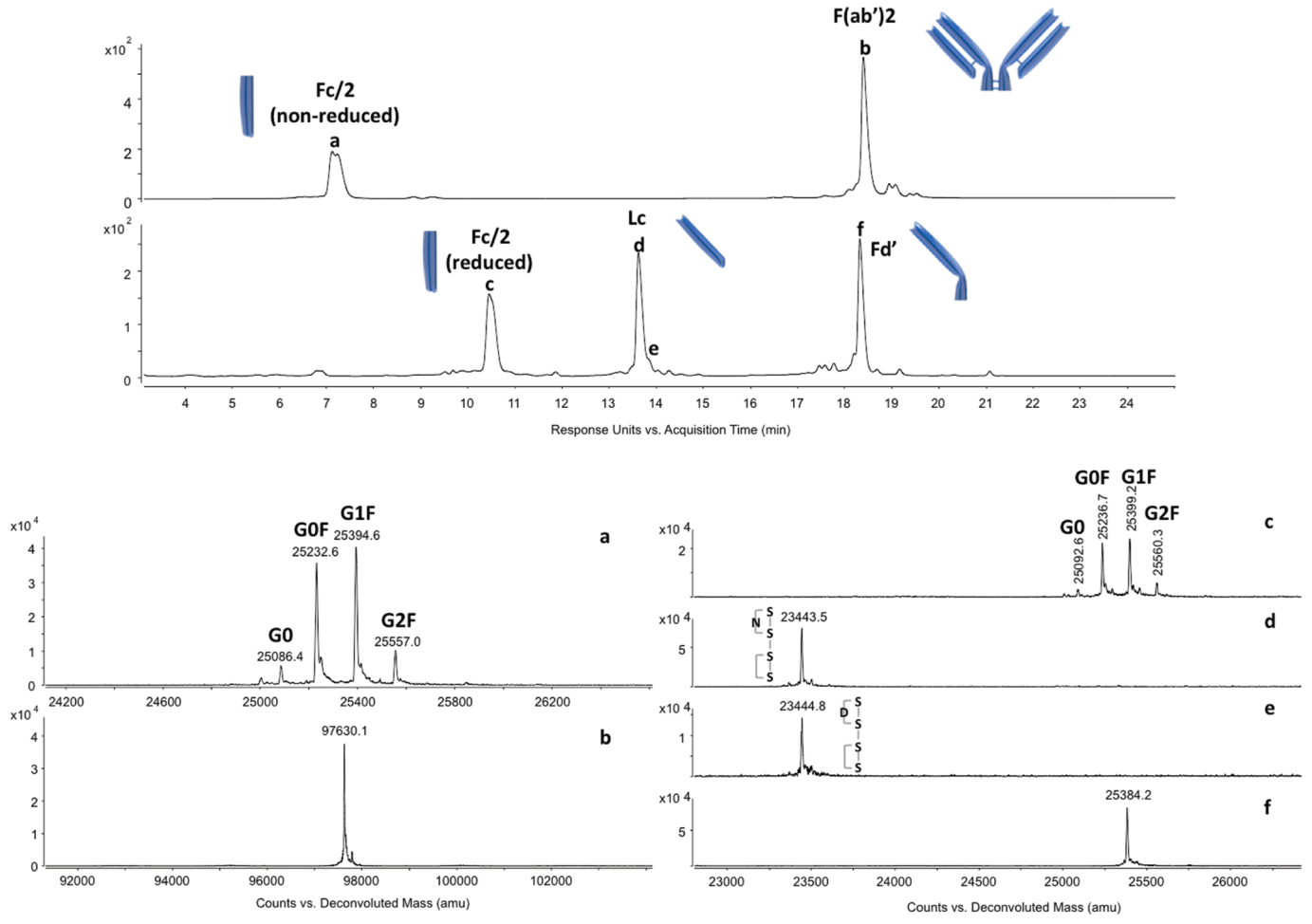

Figure 2: RPLC-UV-MS analysis of non-reduced (upper chromatogram) and TCEP reduced (lower chromatogram) IdeZ cleaved Herceptin and deconvoluted MS spectra associated with the annotated peaks.

In contrast to RPLC; SEC, IEX and HIC are non-denaturing techniques providing complementary information to the former chromatographic mode. Due to the use of non-volatile salts in the mobile phases, these modes are typically not directly compatible with MS. The identification of peaks requires their collection and subsequent desalting or dilution prior to MS measurement. Desalting of the collected fractions can be performed in an automated manner using, for example, a small RP cartridge hyphenated to mass spectrometry.

SEC is the chromatographic mode with the lowest efficiency or resolution but is extremely powerful to determine aggregation and fragmentation. It is recognised that aggregates may stimulate immune responses and it is therefore very important to measure this critical quality attribute. SEC typically makes use of a high ionic strength phosphate buffer at neutral pH which is delivered isocratically to a column packed with particles with a fixed pore size which govern the separation by size. Figure 1 shows the detection of dimers in a commercial Herceptin batch at 0.4% illustrating the sensitivity of the method.

IEX is an excellent tool to highlight charged variants that might arise from modifications such as deamidation, lysine truncation, N-terminal cyclisation, etc. The pre-peak observed in the cation exchange (CEX) chromatogram of Herceptin (Figure 1) corresponds to an asparagine deamidation in the light chain. Deamidation renders a protein more acidic explaining its earlier elution on CEX. This separation makes use of a phosphate buffer at pH 7.6, conditions where mAbs are positively charged and interact with the negatively charged functionalities of the cation exchanger. Elution is achieved by applying a NaCl gradient. Figure 3 shows the CEX analysis of papain digested Herceptin [6]. Due to the lower isoelectric point of the Fab and Fc fragments compared to the intact antibody, separations are performed at lower pH values, i.e. pH 5.6 as opposed to pH 7.6 for intact mAbs. In order to unravel the identity of the different peaks, fractions were collected and analysed by MS following on-line desalting. Deconvoluted spectra are shown in Figure 3 providing great structural detail. Deamidation of the Fab fragment as well as Fc glycosylation is demonstrated.

In recent years, HIC has been revisited mainly from the perspective of antibody-drug-conjugates (ADCs) governing a separation based on the number of conjugated drugs to determine the drug-to-antibody ratio (DAR). In the separation of naked mAbs it is useful to highlight heterogeneities originating from oxidation, aspartate isomerisation, deamidation, succinimide formation, C-terminal lysine and clipping. The mobile phase in HIC consists of a salting-out agent, which at high concentration, retains the protein by increasing hydrophobic interaction between the solute and the stationary phase. Bound proteins are eluted by decreasing the salt concentration. Figure 4 shows the analysis of papain digested non-stressed and high pH-stressed Herceptin on a HIC column. A deamidation event in the Fab fragment is clearly highlighted which is in agreement with the CEX profile shown in Figure 3. High pH values are known to induce deamidations.

Peptide level

Protein measurement is extremely powerful but does not provide the complete picture. While it is indicative for identity and highlights dominant modifications, it

typically does not provide the actual amino acid sequence nor does it adequately pinpoint the location of the modifications. Further structural detail is provided at the peptide level following proteolytic digestion with, for example, trypsin that cleaves peptides C-terminally of an arginine or lysine residue. The highly efficient RPLC separation of a Herceptin digest is shown in Figure 1. Similar to protein analysis, this separation benefits from the use of superficially porous C18 particles, elevated temperatures (60°C) and the use of TFA as ion-pairing reagent. When combining this separation with mass spectrometry, in our case Q-TOF MS, over 98% of the Herceptin sequence is covered thereby confirming identity. Next to identity peptides, the peptide map is also rich in modified peptides. Some at levels down to 0.1%, others at more pronounced levels. Figure 5 shows the extracted ion chromatograms of selected modified peptides present in the peptide map. Of particular interest is peptide ASQDVNTAVAWYQQKPGK located in the light chain. This peptide contains four potential deamidation sites (3 Q and 1 N). Based on MS measurement one cannot discriminate between the 4 sites. If one wants to determine the actual deamidation site, it is necessary to perform MS/MS measurements on the peptides. Based on the fragment ions generated, the deamidation can be traced back to the N. This deamidation in fact corresponds to the deamidation observed at the intact mAb, Fab and Fd' level using RPLC, CEX and HIC. Using these techniques, the deamidation cannot be traced back to a specific residue. Working at the peptide level together with performing MS/MS links the deamidation to a specific residue.

Glycan level

Glycosylation can be revealed at both the protein and peptide level. To get more detail in the glycans themselves it is advisable to remove them from the protein and to measure them as such. Glycan analysis starts with liberating the sugars from the protein backbone using PNGase F and is followed by labelling with 2-aminobenzamide (2-AB) to improve the chromatographic as well as the detection properties. Labelled sugars are subsequently analysed by HILIC with fluorescence and MS detection on-line. Fluorescence detection is mainly used for quantitative purposes while MS is applied for identification purposes. Figure 1 shows the fluorescence chromatograms of the 2-AB-labelled Herceptin N-glycans separated on a HILIC column. The chromatographic principle

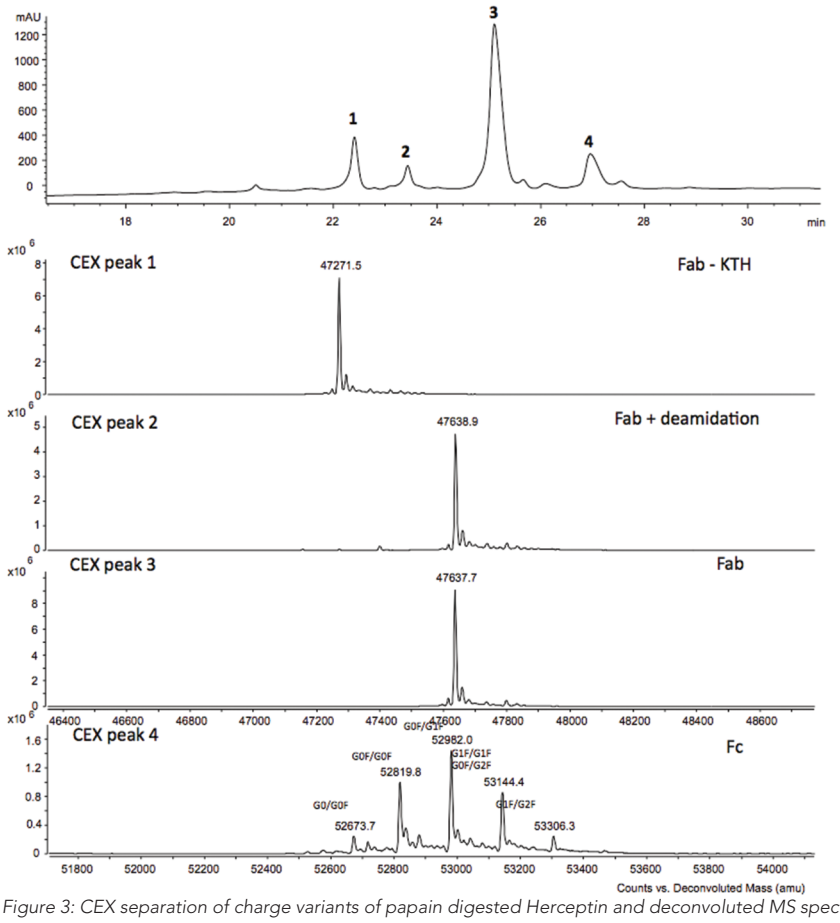

Figure 3: CEX separation of charge variants of papain digested Herceptin and deconvoluted MS spectra associated with the annotated peaks.

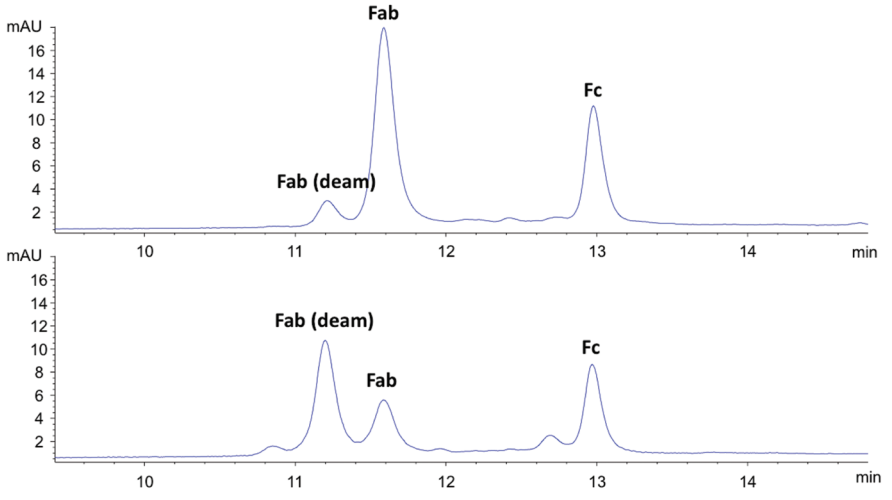

Figure 4: HIC separation of papain digested non-stressed (top) and high pH-stressed (bottom) Herceptin clearly highlighting Fab deamidation.

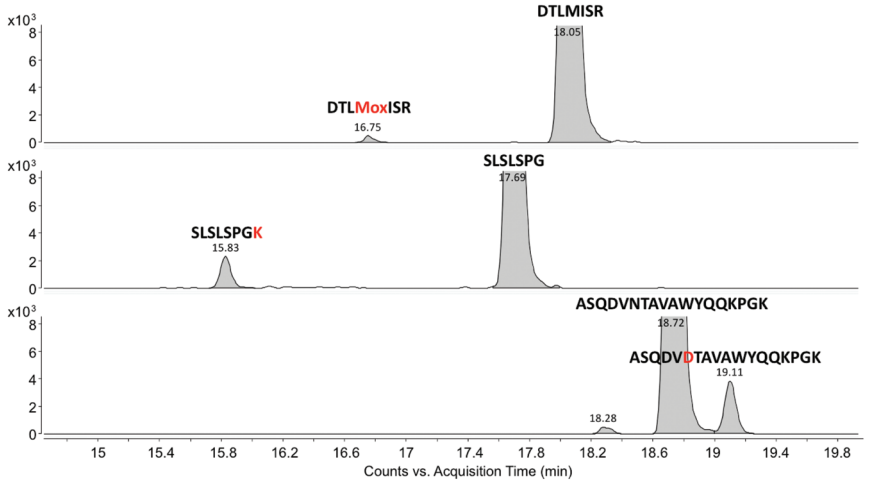

Figure 5: Extracted ion chromatograms of selected native and modified peptides. Shown are a methionine oxidation, the presence of a lysine residue at the C-terminus of the heavy chain and asparagine deamidation.

here is based on the hydrophilic partitioning between an aqueous layer formed around the stationary phase and the mobile phase. If water is absent in the mobile phase the principle of HILIC does not apply. Elution is achieved by increasing the water content. Highly efficient separations can be obtained and isomeric species can be resolved. Since HILIC uses MS friendly mobile phases, the identity of every peak can be confirmed with MS and MS/MS measurement.

Amino acid level

Amino acid analysis is still extensively used to accurately quantify a protein and to determine the amino acid composition. In a first step amino acids are liberated through acid hydrolysis (110°C, 24 h, 6 N HCl). Liberated amino acids are subsequently subjected to automated pre-column derivatisation using o-phtaldialdehyde (OPA) for primary amino acids and 9-fluorenylmethyl chloroformate (FMOC) for secondary amino acids (proline). Derivatised amino acids are then separated by RPLC and detected by fluorescence. Figure 1 shows the analysis of OPA/FMOC derivatised amino acids in a Herceptin acid hydrolysate.

Conclusion

The present contribution reports on common strategies for the characterisation of protein biopharmaceuticals. Characterising these heterogeneous molecular giants requires a wide range of methods as well as a multi-level approach. While all examples shown are related to mAbs, similar workflows are undertaken for other recombinant proteins. Chromatography and mass spectrometry are furthermore used beyond the examples discussed in the monograph, i.e. in assessing pharmacokinetic (PK) properties (7), in clone selection (using Protein A) (8), in determining higher order structures (hydrogen/deuterium exchange)(9), in identifying and quantifying host cell proteins (HPC)(10), etc.

Acknowledgement

The authors acknowledge Lindsay Mesure (Promega, Leiden, The Netherlands).

References

(1) G. Walsh, Nat. Biotechnol. 32 (2015) 992-1000.

(2) G. Walsh, ChromSoc meeting at Medimmune on Advances in Bio-separations (May12-13, 2015).

(3) K. Sandra, I. Vandenheede, P. Sandra, J. Chromatogr. B 1335 (2014) 81-103.

(4) S. Fekete, J.L. Veuthey, D. Guillarme, J. Pharm. Biomed. Anal., 69 (2012) 9-27.

(5) C. Hosfield, P. Compton, L. Fornelli, P. Thomas, N.L. Kelleher, M. Rosenblatt, M. Urh, Promega Poster Part#PS260 (2015).

(6) I. Vandenheede, E. Dumont, P. Sandra, K. Sandra, M. Joseph, Agilent Technologies Publication Part Number 5991-5274EN (2015).

(7) K. Sandra, K. Mortier, L. Jorge, L.C. Perez, P. Sandra, S. Prime, S. Poelmans, M.P. Bouche, Bioanalysis 6 (2014) 1201-1213.

(8) E. Dumont, I. Vandenheede, P. Sandra, K. Sandra, J. Martosella, P. Duong, M. Joseph, Agilent Technologies Publication Part Number 5991-5124EN (2014).

(9) H. Wei, J. Mo, L. Tao, R.J. Russel, A.A. Tymiak, G. Chen, R.E. Iacob, J.R. Engen, Drug Discov. Today 19 (2014) 95-102.

(10) C.E. Doneanu, W. Chen, Methods Mol. Biol. 1129 (2014) 341-350.