Higher-order-structure analysis of proteins by native size-based separations coupled to optical and mass-spectrometric detectors

Ventouri, I.K.

Publication date
2023

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 1

General Introduction
1.1 Higher-order structures of macromolecules

Continuous developments focused on ever larger and more-complex molecules, including their supramolecular assemblies, has opened new frontiers and opportunities in many domains of chemistry and, especially, in life sciences, medicine, and biotechnology [1,2]. Innovative materials, such as functional coatings, medical implants, and biopharmaceuticals, are based on (assemblies of) macromolecules [3]. These highly complex structures possess advanced macroscopic properties and improved functions, which are directly related to their overall structure and chemical composition.

Biopharmaceuticals and biotechnological proteins are among the most sophisticated and elegant products of modern science. The field of biopharmaceuticals includes a very broad range of molecules, including recombinant proteins and monoclonal antibodies, vaccines and nucleic acids [4]. To date, most commercially available biopharmaceuticals feature recombinant proteins as active pharmaceutical ingredient [2]. Advances in recombinant DNA technology have led to the production of protein therapeutics to address previously untreatable human diseases and conditions [5,6]. In comparison with the traditional small-molecule drugs, the production process of biopharmaceuticals and biotechnological enzymes is based on living cells. Biopharmaceuticals are large molecules, the majority being 150 kDa or higher in molar mass, with high structural complexity and heterogeneity, comprised of many post-translational modifications (PTMs) [2,7]. The ability of proteins and protein therapeutics to carry out a great variety of functional roles is directly related to their well-defined three-dimensional conformation, as well as intrinsic characteristics that protect and stabilize their functional native form [8,9]. The higher-order structure (HOS) of a (bio)macromolecule refers to the complex, three-dimensional configuration of the molecules. The secondary, tertiary and quaternary structure are essentially responsible for the correct folding and shape of a protein [10]. The efficacy and safety of biotechnologically produced proteins rely strongly on the exact chemical composition, but also on the conformation of the protein. Even slight deviations from the native form and HOS of a protein may have a detrimental impact on the biological functionality and activity [11].

Degradation of the native protein form and disturbance of the HOS may follow different pathways (chemical and physical modifications), which essentially are correlated and often interdependent [11]. Chemical modifications of the protein, such as oxidation, deamidation, or glycation, may occur during storage or production and may have significant impact on the biological function of the protein. Physical modifications, involving unfolding, dissociation of protein complexes, denaturation and aggregation, may also be triggered [4,11]. Disturbances of the protein HOS and subsequent creation of intrinsically disordered proteins or mis-folded proteins may lead to aggregate formation [11]. Changes in the shape and conformation of specific proteins have been associated with significant aggregate formation, which has been implicated in the generation of “conformational diseases”, including Alzheimer’s, Parkinson’s and Huntington’s diseases [11–14]. Aggregates have been classified as reversible or irreversible, soluble or insoluble, covalent or non-covalent, and small or large assemblies. They range in size from a few nanometers to micrometers (microparticles) or even hundreds of micrometers (visible particles) [11,15]. Characterization of the HOS of (bio)macromolecules and fundamental understanding of the causes of disturbances of the native functional form are pivotal for quality evaluation, for establishing structure-function relationships, and ultimately for developing and producing pharmaceuticals and materials with improved, tailored properties.
HOS can be monitored by a multitude of physical and chemical characterization techniques [10,16]. However, the analytical challenges for the characterization of the HOS of (bio)macromolecules remain enormous. Whether macromolecules are synthetic, natural or biotechnological, experimental investigation of their architecture, conformational heterogeneity, and their association with other molecules is extremely challenging. The increasing size, complexity, and structural diversity of molecules within one sample often still exceed the limits of analytical methodologies. Additionally, preserving the structural integrity of the HOS and the functional native form during the analysis is a prerequisite to draw meaningful and reliable conclusions, especially in the characterization of labile macromolecules, such as non-covalent protein complexes and aggregates. In the case of proteins, it is essential to analyze these at conditions that resemble the environment in which they operate and in which their native functional form exists. Protein structures are dynamic and equilibria often exist between various HOS [4]. Unravelling these equilibria and understanding the factors that may disturb them are also critical. No single analytical technique is yet capable of reliably revealing the detailed structural complexity of large (bio)macromolecules and their HOS [5,10,16]. Therefore, there is a high demand for continuous advances in characterization techniques, such as combinations of high-resolution and selective separations with advanced detection methodologies.

X-ray crystallography and nuclear-magnetic-resonance (NMR) spectroscopy have been the premier methods for the determination of protein HOS, because they offer high structural resolution and allow detailed elucidation of the macromolecular structures [10,16]. However, there are certain limitations. For X-ray crystallography, protein crystals must be formed, which requires very pure proteins, and which is often a very tedious process [17]. Furthermore, X-ray crystallography provides an illustration of a protein structure in the solid state, whereas in reality complex molecular dynamics and equilibria may occur in solution, which is the native environment of the protein [18]. NMR may provide information on solutions. However, a relatively large amount of sample of high concentration is required, and large proteins give rise to decreased sensitivity and increased spectral complexity [19][20]. Also, relatively pure proteins are required to obtain accurate information from NMR.

During the last decades, (ultra-)high-resolution MS (HRMS) has been established as a valuable tool for structural biology and elucidation of HOS of proteins [5,21–24]. MS in principle allows characterization of the (bio)macromolecular structures at different levels, including intact covalent structure (native MS), (post-translational) modifications, conformation (recently also with ion-mobility spectrometry, IMS), and complex formation [22]. However, the sample complexity and heterogeneity, for example in mixtures of similar polymeric compounds or mixtures of proteins in their native, unfolded and aggregated forms, result in convoluted spectra, which may hamper spectral interpretation. Therefore, (pre-) separation of macromolecules remains imperative for sample purification, reduction of the complexity, and maximizing the amount of information obtained [21].

Research has been focused on establishing separation techniques capable of analyzing samples containing a wide range of species under near-native conditions (in terms of solvents, buffers, electrolytes and temperature) under conditions that do not disturb the natural and higher-order structures of (bio-)macromolecules. A number of liquid-chromatographic (LC) techniques, such as size-exclusion (SEC), ion-exchange (IEX), and hydrophobic-interaction (HIC) chromatography, but
also capillary electrophoresis (CE) and field-flow-fractionation (FFF) can be classified as non-disturbing (“native”) separation techniques [25–28]. Analytical platforms based on the synergistic combination of such separation techniques with HRMS [21], light-scattering [29–31] and/or fluorescence detection [32–34] have been developed to yield essential and highly representative information on the HOS of both natural and industrial (bio-) macromolecules.

The challenge of dealing with samples containing species that vary greatly in size has put the size-based separation techniques, especially SEC [4, 21, 35] and FFF [36–38], in the spotlight. A key objective of our investigations has been to establish whether (bio)macromolecules, their HOS, and supramolecular aggregates can be preserved intact during size-based separations. We also set out to gain an understanding of the analytical factors of analytical platforms and methods that may affect the macromolecular structures, to ultimately ensure that the obtained information is representative of the in-solution state. Additionally, because HRMS has become one of the most powerful tools for structure elucidation, we must establish the extent to which MS measurements are reflecting the in-solution state of the molecules and complexes [39, 40]. Combining information obtained by size-based separations, non-destructive liquid-phase detectors (e.g., multi-angle light scattering), and gas-phase MS experiments may reveal new information on the stability, shape, composition, and conformation of (bio)macromolecules in complex mixtures.
1.2 Size-exclusion chromatography

Size-exclusion chromatography is one of the most-commonly used techniques for the separation, purification and molecular-weight (MW) determination of a plethora of biological and synthetic (macro-)molecules, such as proteins, peptides and polymers [41–44]. The underlying separation mechanism in SEC is based on the molecular size and hydrodynamic volume of the analytes with respect to the pore-size distribution of the packing material [4,41]. SEC is an entropy driven separation. It requires the use of inert packing materials to avoid physicochemical interactions between the (bio)-macromolecules and the surface [43]. The pore-size distribution of the column particles is the critical factor determining the resolution between molecules of different size. Relative small (low-molecular-weight) analyte molecules enter most of the pores and stay in the column longest, whereas large (high-molecular-weight) analytes do not diffuse into the pores and elute faster [41]. In SEC isocratic conditions are invariably employed. The mobile phase solely serves to transport the analytes.

SEC is nowadays considered a reference method for the qualitative and quantitative characterization of aggregates and oligomers of large biomolecules and the determination of the molar mass, molar-mass distribution (MMD) and the degree of branching of synthetic polymers [4,7,45–47]. Extension of the range of particle sizes and column dimensions available, and improvements in the inertness of the column materials have expanded the number of applications of SEC for the purification and characterization of higher-order structures, such as proteins [48,49] (between 10 and 500 kDa), protein aggregates [4,41,50], carbohydrates [51], DNA [52] and RNA [53] and lipids and liposomes [54], viral particles [55,56] and nanoparticles [57,58].

To achieve separation based strictly on the size (i.e., the hydrodynamic radius) of the analyte molecules in solution, any interactions, either ionic and hydrophobic, between the biomolecule and the packing material in the column must be eliminated or effectively minimized [41,59,60]. This can be achieved using an optimized mobile phase (salt type, pH, ionic strength, additives) [61]. The secondary interactions not only can disturb the separation - reduced recoveries, or distorted peaks - they can also disturb the structure of the native protein [62–64]. Two main types of electrostatic interactions i.e., attractive and repulsive interactions can jeopardize SEC separations. Ionic adsorption may occur when the protein and the surface of the column material hold opposite charges. This will result in low sample recoveries and distortion of the peak shape, which raise serious concerns about the deterioration of the separation between the parent protein from protein aggregates and other size variants [48,59,64]. When the analyte and column material have the same charge, “ion-exclusion” may cause the analytes to elute earlier due to repulsive forces. To eliminate secondary interactions mobile phases of relatively high ionic strength, consisting of phosphate buffers (50-100 mM) with chloride or sulphate salts (up to 250 mM) are commonly used [48,65]. The nature of the column material is also critical, as it can alleviate effects of analyte-surface interactions [64]. Silica-based column materials derivatized with, for instance, hydrophilic silanes containing diol groups, are preferred for protein analysis [47,50]. The shielding of negatively charged silanol groups significantly reduces their interactions with positively charged proteins that could lead to adverse chromatographic performance or sample adsorption [47,50,64][35,66].

There is a continuous search for non-denaturing aqueous buffers suitable to minimize the unwanted interactions between the analyte and the column surface material while also preserving the
structural integrity of the proteins and their assemblies during analysis [63,64]. The mild, near-native elution conditions associated with such mobile phases enable characterization of biologically active proteins and their aggregates with minimal disturbance of their conformational structure [49,63,64].

Analytical platforms featuring SEC coupled with several detectors have emerged for the comprehensive study of protein aggregation and extensive characterization of the resulting structures [61]. The multiple detectors must typically include at least one concentration-sensitive detector, such as ultraviolet absorbance (UV), refractive-index (RI), or fluorescence (FL) detection. Multiangle light scattering (MALS) and viscometry (Visc) provide detailed information on the size and conformation of the protein structures in solution [67,68]. MALS is especially useful in combination with SEC, because it enables estimation of the molecular weight of the various species without a need for column calibration [45,68]. However, there are certain limitations associated with SEC-MALS for protein characterization [4]. Separation of the various species only based on size implies that molecules with the same size in solution cannot be resolved and properly characterized. Additionally, MALS detectors are much more sensitive for high-MW species. The presence of large aggregates, even in low amounts, can cause a shift in the light-scattering signal of the native protein, introducing significant errors in the calculated molar mass. In the analysis of oligomeric aggregates in dynamic equilibrium, SEC-MALS may provide accurate information only for well-resolved species [30,63].

More recently, hyphenation of SEC with native MS (nMS) has emerged and this technique has found considerable resonance for the characterization of therapeutic proteins, protein complexes and oligomers, and protein fragments [21,44,66]. The additional separation dimension provided by MS allows distinction of variants that cannot be separated by SEC [64,69–71]. This is essential, because the different protein variants may show different activity and immunogenicity [72]. From an MS perspective the coupling with SEC is highly advantageous, as it allows on-line reduction of the sample complexity, resolving the sample components according to size, as well as desalting the sample prior to its introduction in the MS [44,70,72]. Moreover, separation of the (non-covalent) protein aggregates or oligomeric species may reveal whether all species were already present in the sample or were formed during the ionization process.

1.3 Field-flow fractionation

The technique of field-flow fractionation (FFF) was introduced by Giddings in 1966 [73,74]. A whole family of analytical fractionating techniques evolved, because of the potential to separate analytes ranging from a few nanometers up to micrometers in size [75,76]. The members of the FFF family are diverse. While all FFF techniques are based on the same principle, each one may be suitable for specific applications. Analytes are fractionated thanks to an unequal distribution across the velocity profile of a continuous laminar flow under the influence of a perpendicular external field (e.g., flow, centrifugal, thermal, gravitational) [77] (Figure 1). In contrast with liquid-chromatographic techniques, FFF fractionations take place in an open channel, without a packing material or a stationary phase. In most cases the FFF channel is formed by a thin, ribbon-like spacer (thickness (w) from 100 up to 500 μm, width typical breadth between 10 and 30 mm) clamped between a top and a bottom plate. Because of the high aspect ratio (breadth/thickness) and the low flow velocities (typically less than 1 mm/s) a nearly perfect parabolic flow velocity profile is generated in the channel (Figure 1) [78]. The flow velocity is highest in the centre and lowest close to the channel
walls. The perpendicularly applied force drives the sample toward the accumulation wall (typically the bottom plate of the channel). This force is counteracted by diffusion of the molecules away from the wall, creating a concentration distribution. Depending on the physicochemical properties of the sample components, they are positioned at a certain average distance from the accumulation wall. Because the flow velocity also depends on the distance from the wall, due to the parabolic flow velocity, different analytes will elute at different times.

Figure 1. Schematic representation of separation in a field-flow fractionation channel under normal mode.

Depending on the type of the applied external field analytes of various physicochemical properties can be separated (Table 1). The most widely used mode is flow-FFF (FIFFF) comprising the asymmetrical FFF (AF4) and hollow-fiber FFF (HF5) variants. FFF is further classified as electrical FFF (ElFFF), thermal FFF (ThFFF or TF3), sedimentation or centrifugal FFF (SdFFF or CF3), magnetic FFF (MgFFF). AF4 is the most successful variant of the FFF techniques showing highly useful for the characterization of nanoparticles, natural and synthetic (bio)-macromolecules and supramolecular assemblies [38,76,79–81]. Besides AF4, TF3, ElFFF and SdFFF/CF3 have also been commercialized.

The separation principle and the open channel format of FFF result in certain advantages that have led to a number of significant application fields for FFF [36–38,80,82]. The shear stress induced by the velocity profile is significantly lower than the stress exerted on molecules during their migration through a column packed with very small (micrometer-sized) particles, as in LC techniques, including SEC. Thanks to the low stress exerted, FFF is particularly suited for fragile analytes, such as biomacromolecules, molecular aggregates, and higher-order structures [83,84]; Chapters 2, 3 and 4 of this thesis]. FFF can be used for fractionating high-molecular-weight species, and nano- or microparticles that typically cannot be analyzed by column-based separation methods [80,82,85]. Moreover, most FFF techniques are compatible with numerous types of detectors, such as differential-refractive-index (dRI), UV-absorbance, fluorescence, MALS, and dynamic-light-scattering (DLS) detectors and inductively coupled plasma (ICP)-MS [37,81,86–90]. The online coupling of FFF, with several such detectors in one system, provides valuable information on the sample [Chapter 2 of this thesis].
Table 1. Commercially available FFF techniques, relevant analyte properties and the applied field.

<table>
<thead>
<tr>
<th>FFF technique</th>
<th>Analyte properties</th>
<th>Applied field</th>
<th>Refs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flow FFF (F4, AF4, HF5)</td>
<td>Diffusion coefficient / hydrodynamic diameter</td>
<td>Perpendicular flow through top and bottom plates (conventional or symmetrical flow FFF, F4) Perpendicular through bottom plate only (asymmetrical flow FFF, AF4) Radial flow through a tubular membrane (hollow-fiber flow FFF, HF5)</td>
<td>[77,91]</td>
</tr>
<tr>
<td>Thermal (ThFFF or TF3)</td>
<td>Diffusion coefficient / thermal diffusion</td>
<td>Thermal field between a hot and a cold plate</td>
<td>[85]</td>
</tr>
<tr>
<td>Sedimentation FFF (SdFFF) or Centrifugal FFF (CF3)</td>
<td>Diffusion coefficient / hydrodynamic diameter, effective mass (density)</td>
<td>Gravitational field or Centrifugal field</td>
<td>[92,93]</td>
</tr>
<tr>
<td>Electrical FFF (EF3)</td>
<td>Diffusion coefficient / Electrophoretic mobility</td>
<td>Electrical field between top and bottom plates</td>
<td>[94,95]</td>
</tr>
</tbody>
</table>

1.4 Asymmetrical flow field-flow fractionation

The first AF4 system was presented by Giddings and Wahlund in 1987 [96]. The asymmetrical channel replaced quickly the previous symmetrical version. In the original, symmetrical version of field-flow fractionation, (SyFFF or SyF4) both walls were permeable [77]. In AF4 the top wall of the channel is impermeable, whereas the bottom wall, the so-called accumulation wall, is semi-permeable. The latter consists of an ultrafiltration membrane of an inert material and molecular-weight cut-off (MWCO), placed on top of a ceramic or metal support plate (“frit”). The MWCO of the membrane determines the smallest analytes that will remain inside the channel until they are eluted and characterized at the detector. In comparison with the SyF4 configuration, the AF4 configuration offers a higher resolution and greater separation speed [97].

In AF4, the sample components are fractionated in a trapezoidal-shaped open channel (i.e., no packing material present) based on two exerted flows, viz. the laminar channel flow ($F_{out}$) and the perpendicular cross-flow ($F_c$) through the bottom plate. Upon sample injection, the first step is a focusing and relaxation period. There is some variation in the exact procedures, including stop-flow (used in SyFFF), a focusing flow opposite to the channel flow (the most-common approach in AF4), and frit-inlet sample introduction based on hydrodynamic relaxation [88]. During the focusing step in AF4 the analytes are concentrated in a narrow zone based on three flows, i.e., the two opposite flows (laminar inlet flow and focusing flow) and the cross-flow. To achieve adequate focusing, a certain focusing time and an optimal ratio of inlet and focus flow have to be applied. After focusing a steady-state distribution of the particles or macromolecules is achieved, which is essential to achieve an efficient fractionation. Too little focusing can lead to peak broadening, loss of resolution and elution of (a fraction of the sample) at the void volume. However, focusing for too long can lead to very high local concentrations of analytes and can induce interactions between molecules or particles mutually or between analytes and the ultrafiltration membrane. Such a crowding effect
may result in sample aggregation or degradation and a loss in recovery \([56,83,84]\). Consequently, careful optimization of the focusing step is essential. After focusing (i.e., stopping the focusing flow) the actual fractionation commences. Programming a decaying cross-flow (linear or exponential decay) is often necessary during this step, especially for complex samples containing analytes of vastly different size or molecular weight.

AF4 has a wide range of applications involving a wide variety of analytes of vastly different physicochemical properties. Naturally, each AF4 method has to be optimized and tuned, depending on the properties of the analytes and the sample matrix. The optimization of an AF4 method involves the proper selection of the carrier-liquid composition and the membrane (chemistry and MWCO), and establishing suitable channel dimensions (thickness, width and length), flow rates during the relaxation-and-focusing step and the elution step, focusing time, and injected mass. Although there are many parameters involved in optimizing AF4 methods and although many of these parameters appear to be interdependent, AF4 theory and various published studies offer assistance to the user \([76,79,98]\) [99].

The theoretical basis of AF4 has been well described in the literature \([77,97]\)[99]. In this thesis a brief description will be given. The approximations that have been used to simplify calculations will be discussed along with their limitations. A detailed mathematical treatise, can be found in the above-cited literature or in the “Field-Flow Fractionation Handbook” [77].

AF4 is most commonly conducted in the so-called “normal” or “Brownian” mode, where the fractionation relies on a balance between the force field (cross-flow) and the diffusion of the analyte molecules or particles [97]. This normal mode applies to dissolved macromolecules and to submicron (<0.5 μm diameter) particles [97][99] [100]. The cross-flow is directing the particles closer to the accumulation wall. The counteracting force due to the diffusivity of the analytes adheres to Fick’s law. The channel flow rate in AF4 is sufficiently low to reach quasi-equilibrium (“steady-state”) conditions. A concentration gradient is formed, with the analyte concentration decreasing exponentially from the accumulation wall, i.e.,

\[
c(x) = c_0 e^{-\frac{x}{\ell}}
\]  

(1)

In this equation, \(c_0\) is the concentration at the accumulation wall, \(\ell\) is the equilibrium layer thickness, and \(x\) is the distance from the accumulation wall.

Independently of the size or diffusivity of the analytes, the highest concentration is near the accumulation wall [Equation 1]. Smaller analyte particles with higher diffusion coefficients exhibit concentration profiles that stretch further into the channel than larger particles with lower diffusivity. As a result, the smaller analytes reach a higher average distance from the accumulation wall. In combination with the parabolic velocity profile this results in these small molecules eluting earlier form the channel.

For larger particles, with diameters above 1 μm, steric or hyperlayer modes are encountered [101]. The diffusion of these large analytes away from the accumulation wall is negligible and the average distance from the wall is essentially determined by the physical radius (size) of the particles. In these cases, the elution proceeds from larger to smaller analyte particles [101]. In this thesis, all relevant samples are well below 1 μm in diameter, thus following the normal “Brownian” behavior in AF4.
The characteristic distance, $\ell$, for each analyte is determined by the ratio between the analyte diffusion coefficient $D$ and the absolute field-induced velocity $|U|$ (Equation 2). Therefore, the separation in AF4 is driven by differences in the diffusivity of the analytes ($D$).

$$\ell = \frac{D}{|U|} \tag{2}$$

For AF4 $|U|$ is the transverse (cross-flow) velocity.

The retention parameter $\lambda$ is related to the equilibrium layer thickness $\ell$ and the channel thickness $w$ (Equation 3).

$$\lambda = \frac{\ell}{w} \tag{3}$$

The retention ratio or retention level ($R$) is the ratio of the void time ($t_0$) and the observed retention time ($t_R$) is expressed by Equation 4.

$$R = \frac{t_0}{t_R} = 6\lambda \left[ \coth \left( \frac{1}{2\lambda} \right) - 2\lambda \right] \tag{4}$$

For $\lambda < 0.02$, Equation (4) can be approximated within an error of 5% as

$$R = \frac{t_0}{t_R} = 6\lambda \tag{5}$$

The retention time $t_R$ is dependent on the diffusivity of the analytes ($D$), the flow rate at the channel out ($F_{out}$), and the cross-flow rate ($F_c$). The following equation can be derived [78].

$$t_R = \frac{w^2}{6D} \ln \left( 1 + \frac{F_c}{F_{out}} \right) \tag{6}$$

It should be noted that this equation is valid only if a constant cross-flow rate is applied, and if self-interactions and interactions between the analyte and the membrane are absent. Using the Stokes-Einstein relationship for the diffusion coefficient of analytes (Equation 7),

$$D = \frac{k_B T}{6\pi \eta r_H} \tag{7}$$

where $k_B$ is the Boltzmann constant, $T$ the absolute temperature and $\eta$ the viscosity of the eluent, $t_R$ can be related to the hydrodynamic radius $r_H$ (Equation 8),

$$t_R = \frac{\pi \eta w^2 F_c t_0 r_H}{k_B T V_0} \tag{8}$$

To calculate $r_H$, the void time ($t_0$), void volume ($V_0$) and the channel thickness ($w$) have to be estimated. The calculation of these parameters is dependent on the application of a cross-flow rate that is constant over time or cross-flow programming, involving a decaying cross-flow or a series of program steps. Solutions for these calculations have been described by Giddings and co-workers [77,97,100,102]. The calculation of the diffusivity of analytes based on the retention time requires complete absence of non-ideal effects, so that the elution behavior of the analytes follows the FFF theory. The use of calibration standards to relate $t_R$ to the molecular weight (MW) or to $r_H$ is also possible. However, standards with similar properties should be used and these are not always available. Therefore, AF4 is commonly coupled to size-specific detectors, such as MALS or DLS. This allows obtaining information on the “absolute” MW and $r_H$. 

14
1.5 Multi-angle light scattering

Multi-angle light scattering (MALS) is a well-established technique used to determine the average molar mass (M) and the molecular size (root-mean-square radius, approximately equal to the radius of gyration, \( r_g \)) of various macromolecules and particles [103]. It is most often used in conjunction with size separation techniques, such as SEC or FFF. In particular, AF4 coupled to a multi-detector system (MALS in combination with UV absorbance and/or dRI) represents a very useful tool to obtain independent size information [104–106]. AF4-UV-MALS-dRI was used throughout this thesis for characterizing higher-order protein structures, including oligomers and larger aggregates. A MALS instrument features between three and twenty-one photodiodes, positioned at different fixed angles in a single plane. In the present study a twenty-one detector configuration was used. The principle of static light scattering at multiple angles (MALS) is described by the equations for light scattering at a single angle. Polarized light is used, and this is focussed on the sample. The scattered light is detected simultaneously at all angles by the photodiodes, which are oriented perpendicularly to the polarization of the laser beam. The intensity of the scattered light is proportional to the molar mass of the analyte molecules (\( M \)), the concentration (\( c \)) of the analytes and the refractive index increment (\( \text{dn}/\text{dc} \)) [103,107]. To explain how \( M \) and \( r_g \) are calculated from MALS experiments, some fundamental relationships will be briefly presented. The intensity of the scattered light is related to the molar mass by the following equations.

\[
\frac{K_{LS}c}{R(\theta)} = \frac{1}{M} \frac{P(\theta)}{R(\theta)} + 2A_2c + \cdots \tag{9}
\]

\( R(\theta) \), the so-called Rayleigh ratio, is the excess intensity of scattered light at a given angle \( \theta \)

\[
R(\theta) = \frac{I(\theta) - I_s}{I_0} \cdot \frac{d^2}{V} \tag{10}
\]

\( I(\theta) \) is the intensity at angle \( \theta \), \( I_s \) the scattering intensity of the solvent, \( I_0 \) the intensity of the incident beam, \( V \) the scattering volume and \( d \) the distance from the scattering centre. The optical constant \( K_{LS} \) is equal to

\[
K_{LS} = \frac{\left( \frac{2m_0\eta_0}{\pi} \right)^2}{\lambda_0^4N_A} \tag{11}
\]

where \( c \) is the sample concentration of the scattering species (information obtained by UV or dRI detectors), \( M \) the weight-average molar mass and \( N_A \) is Avogadro’s number. In the case that a separation or fractionation technique is used to separate the various analytes, each fraction is assumed to be monodisperse and \( M \) can be determined for each slice and from these values the entire molar-mass distribution can be obtained. \( A_2 \) in Eq. (9) is the second virial coefficient, which is usually neglected, because of the high dilution during the fractionation process (low \( c \)) [108]. \( P(\theta) \) is the particle scattering function, which describes the angular dependence of the scattered light. It can be related to \( r_g \) as follows

\[
P(\theta) \approx 1 - \left( \frac{16\pi^2n_0^2}{3\lambda_0^2} \right) \sin^2 \left( \frac{\theta}{2} \right) \left( \frac{r_g^2}{2} \right) + \cdots \tag{12}
\]

In this approximate expression, \( n_0 \) is the refractive index of the solvent, \( \lambda_0 \) is the vacuum wavelength of the laser. From the above equations the relation between the size and angular dependence of
the scattered light is known. For larger particles \(r_g > \lambda_0/20\) it is necessary to include higher moments in the expansion of \(P(\theta)\) \[109,110\].

The resulting equation is the following

\[
\frac{Kc}{R(\theta)} = \frac{1}{M} + \frac{16\pi^2}{3\lambda^2 M} (r_g)^2 \sin^2 \left(\frac{\theta}{2}\right)
\]

(13)

In theory the simplest way to obtain the \(M\) from static-light-scattering experiments is by measuring the scattered intensity at a 0° angle, where \(\sin^2 \left(\frac{\theta}{2}\right)\) equals 0. However, in practice such an experiment is impossible. With MALS the measurement of the light scattering at various angles allows extrapolation of the intensity to 0°. \(r_g\) can be obtained from a plot of the intensity of the scattered light intensity versus the scattering angle (i.e., \(R(\theta)/Kc\) against \(\sin^2 \left(\frac{\theta}{2}\right)\)). \(M\) is obtained by extrapolation to 0° after empirical curve fitting. Various mathematical equations have been proposed to describe the relationship between the intensity of the scattered light and the scattering angle, including the Debye, Zimm, and Berry equations \[108\].

The equation used for the extrapolation may affect the accuracy of the calculated values for the molar mass and radius of gyration. For smaller macromolecules the Zimm-equation is most commonly applied \[103,108,109\]. For larger and more-complex structures, a Debye or Berry plot should be used, and the extrapolation has to be carefully investigated \[103\]. The Zimm-method was used for the characterization of proteins and protein complexes in this work. More information on the use of light-scattering in conjunction with FFF, as well as on some of the many possible application can be found in Chapter 2 of this thesis.

1.6 Native mass spectrometry for proteins and higher-order structures

In native mass spectrometry (nMS) large biomolecules, such as intact proteins and protein complexes, are transferred from the liquid phase to the gas phase under conditions which are so mild that their non-covalent interactions and native conformation are largely preserved \[24,40\]. Characterization of large biomolecules by MS was made possible through the introduction of soft ionisation techniques such as electrospray ionisation (ESI) \[111\] and matrix-assisted laser desorption/ionisation (MALDI) \[112\]. Essential for nMS is that native structural features of the molecules or complexes are retained during their ionization, desolvation and transfer in and through the mass spectrometer. nMS is an extremely useful approach to gain insights in the properties and functionality of proteins under physiological conditions and establish structure-function relationships \[5,21,23,113\]. Analysis of biomacromolecules under native (or, in other words, non-denaturing) conditions, may be used to confirm their molar mass, but may also provide invaluable information on homogeneity \[114\], conformation (folded vs. unfolded) \[115–117\], stoichiometry of complexes formed \[118\], and oligomerization (aggregation) state \[119,120\]. Such information is crucial for developing effective and high-quality biotherapeutics.

The main aim of nMS is to bring the biomolecules into the mass analyzer, while maintaining their native, folded state, the (intra-) molecular interactions within the molecule and the (extra-) molecular interactions within HOS. This may be a great challenge. During the ionization process solvent molecules are stripped off, before MS analysis takes place under (ultra-)high-vacuum conditions. Strictly speaking, a native state cannot be preserved during this transition, there is
evidence of a significant structural memory in proteins when transferred into the gas phase under mild-MS conditions [39]. A primary example is the analysis of intact non-covalent complexes, the native structure of which is largely preserved during their transfer from droplets of a solution to the gas phase, while they are being desolvated in the process [121–123].

In recent times nMS has been applied to study numerous proteins, including large membrane proteins [124] and intrinsically disordered proteins [125,126], protein assemblies [115,127,128], interactions between proteins and ligands [129,130], DNA and RNA molecules [131,132], and very large (MDa) systems [133,134], such as intact virus capsids [135]. For all these applications, the MS instruments had to be able to transmit and detect biomolecules with high charge states and high mass-to-charge ratios, i.e., they required mass analyzers with extended m/z ranges. Advances in MS instrumentation have pushed the technology forward [24,136]. Until recently, nMS for studying HOS of biomacromolecules has been largely dominated by the use of quadrupole time-of-flight (Q-ToF) MS analyzers, because ToF analyzers theoretically allow an unlimited m/z range [137,138]. Additionally, improvements and modifications of instruments based on ion traps, including Orbitrap and Fourier-transform ion-cyclotron resonance (FTICR) instruments, have brought them to the fore for nMS applications. Such instruments show a high resolving power [24,136]. The vast majority of modern MS systems are hybrid instruments (i.e., they feature several different mass analyzers). Typical examples are quadrupole-TOF and quadrupole-Orbitrap systems. Such hybrid instruments allow rapid, accurate, and precise mass determinations. Also, they can transmit large protein complexes without a need for high collisional-activation energies, that could damage the HOS. The ever-increasing size and complexity of analytes of interest come with an ever-increasing demand for the performance of nMS analyzers in terms of key performance metrics (mass accuracy, detection sensitivity, mass range, MS/MS capabilities, mass resolution, and data-analysis options).

Although nMS has become widely used over the past decades, some analytical challenges remain to be overcome. Not only is there a need for more-powerful instrumentation, but one of the most critical factors for nMS concerns the sample preparation. Making the analytes amenable to mild-ESI requires tailored procedures. Universal guidelines have not been found to suffice. Analysis of proteins with nMS requires highly purified samples. The presence of various protein species in diverse concentrations often leads to ionization suppression and highly convoluted mass spectra, making data interpretation extremely challenging [21]. The sample matrix often prohibits the direct introduction of specimens into the MS instrument. Therefore, laborious efforts are often needed to purify and desalt the proteins of interest. Such procedures can cause significant sample loss and severe denaturation of the protein. As an alternative to laborious sample-preparation processes, comprehensive characterization of biotherapeutic proteins and biological samples typically involves efficient analytical separations prior to analysis and detection by nMS.

To characterize the structure of proteins in complex mixtures, hyphenated techniques that combine separations with MS are the leading technology. SEC, IEX, HIC, and CZE are suitable for the separation and characterization of proteins and protein complexes, using near-physiological conditions that leave their structure, conformation and functionality intact. The online coupling of non-denaturing separation techniques with MS provides significant benefits. However, this approach is also accompanied by practical challenges [21][44]. Essentially, the separation techniques that are to be combined with ESI-MS must allow direct ionization of the biomacromolecules from the separation effluent. One of the most critical concerns is that both the
separation conditions and the MS conditions must preserve the protein structure, while optimal compatibility between the two techniques is required [64][23]. In this thesis, we describe our quest for size-based separation conditions that ensure preservation of HOS, including labile protein complexes. We specifically focused on finding eluent conditions that allowed near-native size-based separations (SEC, AF4), yet were compatible with on-line coupling to MS. We also explored the potential advantages and limitations of the use of microbore SEC columns (≤ 1 mm I.D.) operated at micro-flow conditions (15 μL/min) on the efficiency of the ionization in MS and sample consumption, especially for larger biomolecules and protein complexes.

1.7 Scope of the thesis

The objective of the work presented in this thesis was to establish novel analytical platforms by combining advanced separation techniques with state-of-the-art mass spectrometry. We aimed to provide detailed information on the size, chemical composition and conformation of labile (bio)macromolecules and their supramolecular structures in a single analysis. To this end, we have performed extensive investigations into the suitability of the size-based separation techniques SEC, and AF4 for the non-destructive characterization of proteins under (near-) native conditions. Emphasis has been on optimization of the separation and interfacing (“hyphenation”) with several detectors that provide complementary information. The latter include liquid-phase characterization by MALS, UV absorbance, Visc and RI detection, and gas-phase detection, by high-resolution mass spectrometry. Ultimately, the combination of the various methods should provide a reliable picture of the actual (native) state of proteins and labile protein complexes in solution.

In Chapter 2, an extensive overview is presented, which shows the developments in AF4 to study interactions between various molecular entities, such as protein-protein, polymer-polymer, nanoparticle-drug, and nanoparticle-protein interactions in diverse matrices. Based on the literature research a summary is provided of the information that is obtained by FFF techniques (e.g., size, density, molar mass, shape, metal composition), hyphenated with a great variety of detectors. The main attractiveness of FFF for interaction studies is found to lie in the absence of packing in the separation channel. FFF is shown to be suitable for monitoring structural, morphological and chemical alterations that result from inter- and intramolecular interactions and therefore, relevant interaction parameters, such as the dissociation constants, aggregation kinetics and binding stoichiometry, can be obtained. The prospects, but also the practical limitations of the technique for interaction studies are outlined.

The study described in Chapter 3 involves the investigation of the dynamic association equilibria between the various species of the biotechnological enzyme β-D-galactosidase using AF4 and frit-inlet FFF coupled to a triple-detection system comprised of UV-absorbance, multi-angle light-scattering, and differential-refractive-index detectors (UV-MALS-dRI). The main aspect of this work was to understand and evaluate possible structural changes in the conformation or changes in the association equilibria (between monomers and various the oligomeric forms) resulting from the analytical measurements and conditions. Various critical parameters of AF4, including the focusing step, cross-flow rate, and the injected amount, were studied and optimized in order not to affect oligomerization or dissociation. Batch-mode dynamic light scattering (DLS) was used to provide complementary information on protein oligomerization under the examined conditions. AF4 and
SEC were compared to evaluate, potential disturbance of the dynamic association equilibria of proteins, due to physical stress exerted on the molecules.

In Chapter 4, the coupling of AF4 to native mass spectrometry (nMS) is presented for the detailed characterization of heterogenic macromolecular structures and fragile non-covalent protein complexes. Coupling AF4 to both UV-MALS-dRI and nMS simultaneously would increase the level of structural information obtained and allows studying analytes across a broader size range (from small analyte molecules up to higher-order protein structures) in a single run. In (forced) degradation and real-time stability studies, dynamic processes of aggregation, oligomerization or dissociation of protein species may result in a mixture of species, which are difficult to discern with MALS. Dissociation may occur during the electrospray-ionization (ESI) process, if non-covalent protein assemblies are transferred from a liquid solution to the gas phase. By combining the structural information obtained from the liquid phase (AF4 and UV-MALS-dRI) and the gas phase (nMS) misinterpretations can be avoided and more-reliable conclusions can be drawn about the stoichiometry, the dynamic equilibria between the various oligomer assemblies, and stability of these structures.

In Chapter 5, the impact of SEC conditions on the structural integrity of proteins was investigated using native ESI-MS as a selective tool to reveal conformational alterations. The combination of SEC with ESI-nMS has proven to be a useful tool for structural biology, employing aqueous mobile phases with volatile salts at neutral pH. However, it is essential to ensure that the separation and ionization conditions do not change the protein conformation and structure. On-line coupling of SEC with ESI-nMS allowed monitoring of structural changes during the separation process depending on the mobile-phase composition (ionic strength, type of volatile salt, pH). ESI-nMS can be used to reveal the influence of nonspecific interactions between protein and stationary phase on the protein structure. The ultimate goal was to establish, conditions at which target proteins (myoglobin, carbonic anhydrase, and cytochrome-c) could be analyzed by SEC-ESI-MS without compromising their native state.

In Chapter 6, the use of SEC-ESI-nMS for studying labile, non-covalent protein complexes and HOS. Narrow SEC columns (1 mm internal diameter, operated at a set flow rate 15-μL/min) directly coupled to ESI-nMS was evaluated for the characterization of proteins. The goal was to achieve a significant increase of the ionization efficiency and, thus MS sensitivity, facilitating detection of low-abundant impurities and higher-order species. Desolvation and ionization of the proteins mild conditions were investigated in order to transfer the protein molecules to the gas phase while (largely) preserving their native structure. To avoid band-broadening effects due to large injection volumes on narrow SEC columns, a trap-and-elute set-up was developed and tested, with an ion-exchange (IEX) trap column installed prior to the SEC-nMS analysis. The aim was to preconcentrate highly diluted samples in a small volume with a concurrent rapid on-line buffer exchange. The trap-SEC-nMS system was evaluated for the detection of labile proteins and protein complexes allowing picogram detection limits.
References


Chapter 1


[70] G. Van der Rest, F. Halgand, Size exclusion chromatography-ion mobility-mass spectrometry coupling: a step


