

How Far Can We Go with Structural Mass Spectrometry of Protein Complexes?

Michal Sharon

Department of Biological Chemistry, Weizmann Institute of Science, Rehovot, Israel

Physical interactions between proteins and the formation of stable complexes form the basis of most biological functions. Therefore, a critical step toward understanding the integrated workings of the cell is to determine the structure of protein complexes, and reveal how their structural organization dictates function. Studying the three-dimensional organization of protein assemblies, however, represents a major challenge for structural biologists, due to the large size of the complexes, their heterogeneous composition, their flexibility, and their asymmetric structure. In the last decade, mass spectrometry has proven to be a valuable tool for analyzing such noncovalent complexes. Here, I illustrate the breadth of structural information that can be obtained from this approach, and the steps taken to elucidate the stoichiometry, topology, packing, dynamics, and shape of protein complexes. In addition, I illustrate the challenges that lie ahead, and the future directions toward which the field might be heading. (J Am Soc Mass Spectrom 2010, 21, 487–500)
© 2010 American Society for Mass Spectrometry

The Human Genome Project has revealed that the biological complexity of organisms is not reflected solely by the number of proteins, but rather by the number of physiologically relevant protein interactions that organisms display [1, 2]. Current predictions estimate that about 35,000 protein–protein interactions occur in the yeast *Saccharomyces cerevisiae*, yielding approximately six unique interactions per protein [3, 4]. The human interactome is predicted to be much more complex: It is estimated that its 26,000 to 38,000 genes [2] form 650,000 protein interactions [5, 6]. These numbers demonstrate that most proteins function within a multi-component assembly, forming hundreds of different functional modules within the cellular environment. Currently, the precise mechanisms of assembly, function, and regulation of the vast majority of these protein complexes remain unclear. Revealing the structural and dynamic interactions among proteins that constitute the proteome, and correlating this information with function creates one of the major challenges in the post-genomic era.

To understand how a macromolecular structure translates into function, it is crucial to determine the characteristics of a complex, and reveal the biological role and mode of action of the protein assembly. For such analyses, it is necessary to define how the function of the complex as a whole differs from that of its individual components, and then determine the contribution of each subunit—based on its identity, stoichiometry, shape, and structure—to the overall activity

(Figure 1). In addition, the relative organization of subunits within the complex should be defined: information concerning subunit–subunit interactions, relative orientations, proximity, and assembly packing is imperative. The ultimate goal of the analysis, and its final step, is to retrieve a high-resolution structure of the protein assembly. Nevertheless, many hurdles stand in the way of defining these structural properties; for instance, the low copy numbers of protein assemblies in the cell, their heterogeneous composition, asymmetric contour, fast dynamics, and the abundance of fleeting interactions.

A variety of methods currently exist to experimentally determine the structure of protein complexes (Figure 1). These include X-ray crystallography and nuclear magnetic resonance (NMR), which yield atomic resolution structures, and lower resolution methods such as cryo-electron microscopy, and small-angle X-ray scattering and labeling techniques [7, 8]. Another emerging structural biology approach, which will be the focus of this article, is mass spectrometry (MS). While the scientific community is well aware of the contribution that MS has made to proteomic initiatives, here, I would like to focus on its critical importance to structural studies of multi-protein assemblies [9–13].

The field of structural MS came into being with the realization that noncovalent interactions can be maintained in the gas phase [14–16]. This meant that in the right set-up, entire protein complexes could be transferred, intact, into the mass spectrometer, and analyzed. This ability was then utilized for structural analysis of numerous protein complexes; for example, in large MDa complexes such as the ribosome [17], and highly ordered virus capsids [18–20], in defining substrate

Address reprint requests to Dr. M. Sharon, Department of Biological Chemistry, Weizmann Institute of Science, Rehovot 76100, Israel. E-mail: michal.sharon@weizmann.ac.il

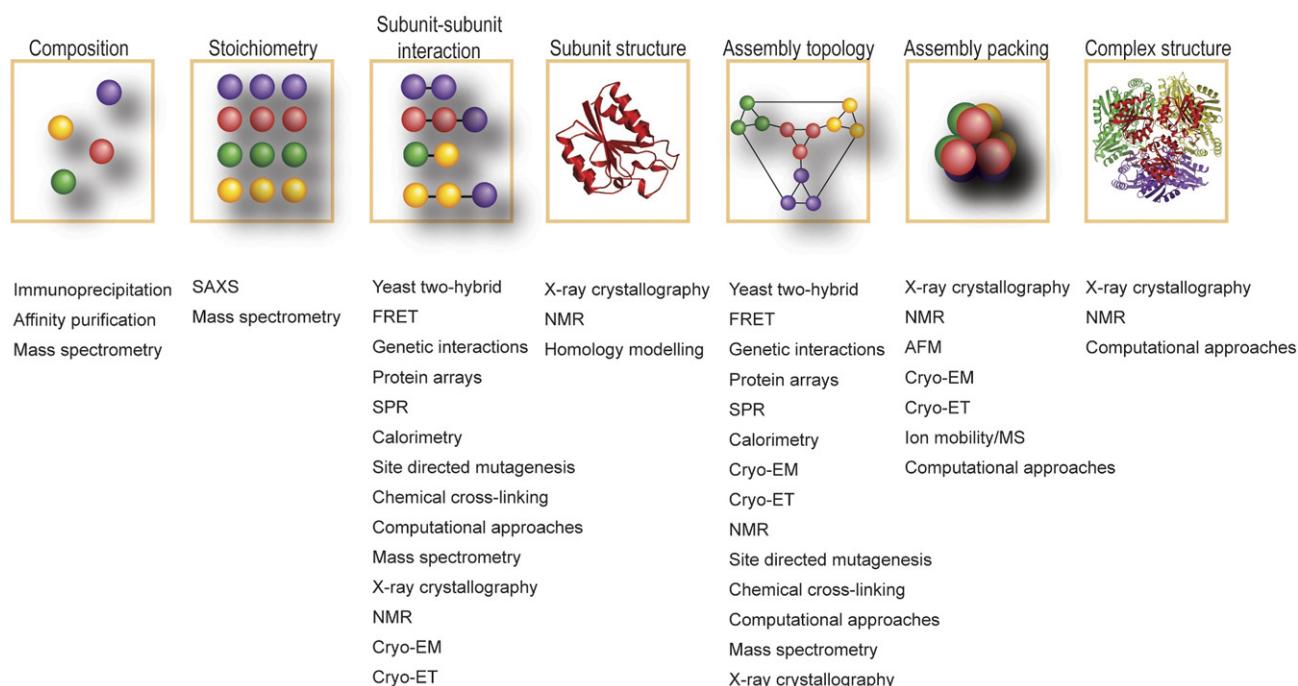


Figure 1. Different levels of information obtained along the path toward structural determination of macromolecular assemblies. The range of experimental methods that can provide the relevant structural information are listed. AFM: atomic force microscopy; Cryo-ET: cryo-electron tomography; EM: electron microscopy; FRET: fluorescence resonance energy-transfer; NMR: nuclear magnetic resonance; SAXS: small-angle X-ray scattering; SPR: surface plasmon resonance.

binding to molecular machines [21–23], and in the characterization of polydisperse assemblies [24]. The fundamental advantage of using MS lies in its ability to probe transient, asymmetric, and heterogeneous macromolecular complexes, using very low concentrations of sample (pmol). In addition, the speed of the analysis enables dynamic reactions to be monitored in real time.

Most of the experiments conducted on multi-protein complexes are performed using a nano-electrospray ionization source (nESI), and a quadrupole-time-of-flight (Q-TOF) instrument [9–13]. In this critical insight, I focus on the contributions made by nESI, although recently other ionization techniques [25–27] have been shown to provide excellent alternatives. I will describe MS analysis, step-by-step, and illustrate the type of information gained at each level. Overall, I will summarize the current features of the MS technique, highlight its capabilities, and point up current methodological gaps in obtaining structural details of protein complexes. In addition, attention will be given to the integration of the MS strategy with other structural approaches, as well as challenges that lie ahead.

Using MS as a tool for generating a subunit interaction map of protein complexes is a multi-step process (Figure 2). First, the composition and stoichiometry of the protein building blocks are defined. Next, the core and peripheral subunits are identified, and information regarding the assembly dynamics and the packing of the subunits is gathered. Subse-

quently, the set of protein–protein interactions is clarified. Ultimately, at this last stage, a protein interaction map may be created. These steps are described in detail below.

Protein Complex Composition and Stoichiometry

One of the first applications of structural mass spectrometry involved defining the stoichiometry of protein assemblies [28, 29]. When the analyzed sample is a homo-oligomer, stoichiometry determination is quite straightforward, as the obtained mass is simply divided by the mass of the monomeric subunit to obtain the subunit copy number. This copy number can then be validated by applying the tandem MS (MS/MS) approach, in which one of the charge states corresponding to that of the intact complex is selected and subjected to collision induced dissociation (CID) by colliding the ions with neutral gas atoms or molecules [30]. These collisions lead to the charge-driven unfolding of a monomer, which may then be expelled from the complex. Generally, the CID dissociation pathway of protein complexes is asymmetric in terms of charge, and yields highly charged monomers, and oligomers of relatively low charge (“stripped complex”) [31]. The asymmetric nature of the charge is an advantage when assigning the spectrum, as it increases the separation between adjacent peaks. The stripped complex mass

can then be accurately measured, and should match that of the intact complex, missing one subunit. Notably, for such analyses, conditions within the mass spectrometer must be optimized, to maintain the non-covalent interactions of large complexes. Usually, higher pressures are used in the initial vacuum stages of the instrument, to stabilize even relatively labile macromolecular protein complexes [32, 33]. In addition, the radio frequency (RF) of the quadrupole analyzer is reduced, to enable the selection and transmission of high mass to charge (m/z) ions ($>4000\ m/z$) [33].

More challenging are cases where the protein complex is composed of different subunits. In such cases an additional step of MS analysis is needed before stoichiometry determination, to define the exact mass of each subunit in the complex. One possibility is to apply the tandem MS approach and measure the mass of the expelled monomeric subunits. The advantage of this approach, which involves the removal of water and buffer molecules following collisions with the neutral gas buffer, is its accuracy. A possible disadvantage is that some core subunits may always remain in the stripped complex, making it difficult to determine their monomeric mass. For that reason soft denaturing conditions (mainly by adding small amounts of organic solvents) are used, to generate smaller subcomplexes. The latter can then be further dissociated into their composing subunits by tandem MS (as above), enabling mass measurement of additional individual subunits that may be otherwise buried within the complex (Figure 3a) [34]. Another possibility for defining the exact masses of individual subunits is MS analysis of a denatured solution of the complex [35] (Figure 3b). While this method enables accurate mass measurements of the individual components in a single step, overlapping charge states may often complicate the analysis. In such situations, the separation capabilities of a hybrid MS/ion mobility instrument (as the Synapt, Waters, Manchester, UK) [36, 37] may be utilized, to separate the different ions into a third dimension, drift time, thereby resolving the overlapping charge states. (The MS/ion mobility approach is further discussed below).

The methods described above are generally used to define the stoichiometry of complexes composed of 6–8 different subunits. This is due to the relatively low ionization efficiency and overlapping charge states of some subunits, making it difficult to define the exact masses of multi-component complexes. Moreover, when studying endogenous protein complexes isolated directly from cells, protein subunits are frequently modified to such an extent that it is impossible to identify their sequences by simply measuring their mass and searching protein databases. To this end, it was necessary to establish new methods of analysis for larger and more heterogeneous complexes.

In one such approach, the complex is decomposed to its component building blocks by using denaturing conditions. The individual subunits are then separated

using a liquid chromatography (LC) column, such as C₄ or monolithic (Figure 3c). The eluted flow is initially sprayed directly into the mass spectrometer for accurate mass determination of subunits. In addition, the same LC separation is repeated, and the eluted peaks are spotted directly onto a MALDI plate, followed by tryptic digestion, MS/MS analysis, and a database search for protein sequence identification. The UV chromatogram of the two LC separations enables a correlation between mass and identity for each of the eluted subunits [6, 38]. This step provides basic information on the building elements (i.e., subunit mass and identity) of the complex; subsequently, the noncovalent interactions between subunits can be investigated (as detailed in the subunit–subunit interaction section below). While the maximum number of subunits that can be separated using this approach still needs clarification, one of its major strengths is that it makes it easier to detect and identify isoforms or post-translational modifications of subunits.

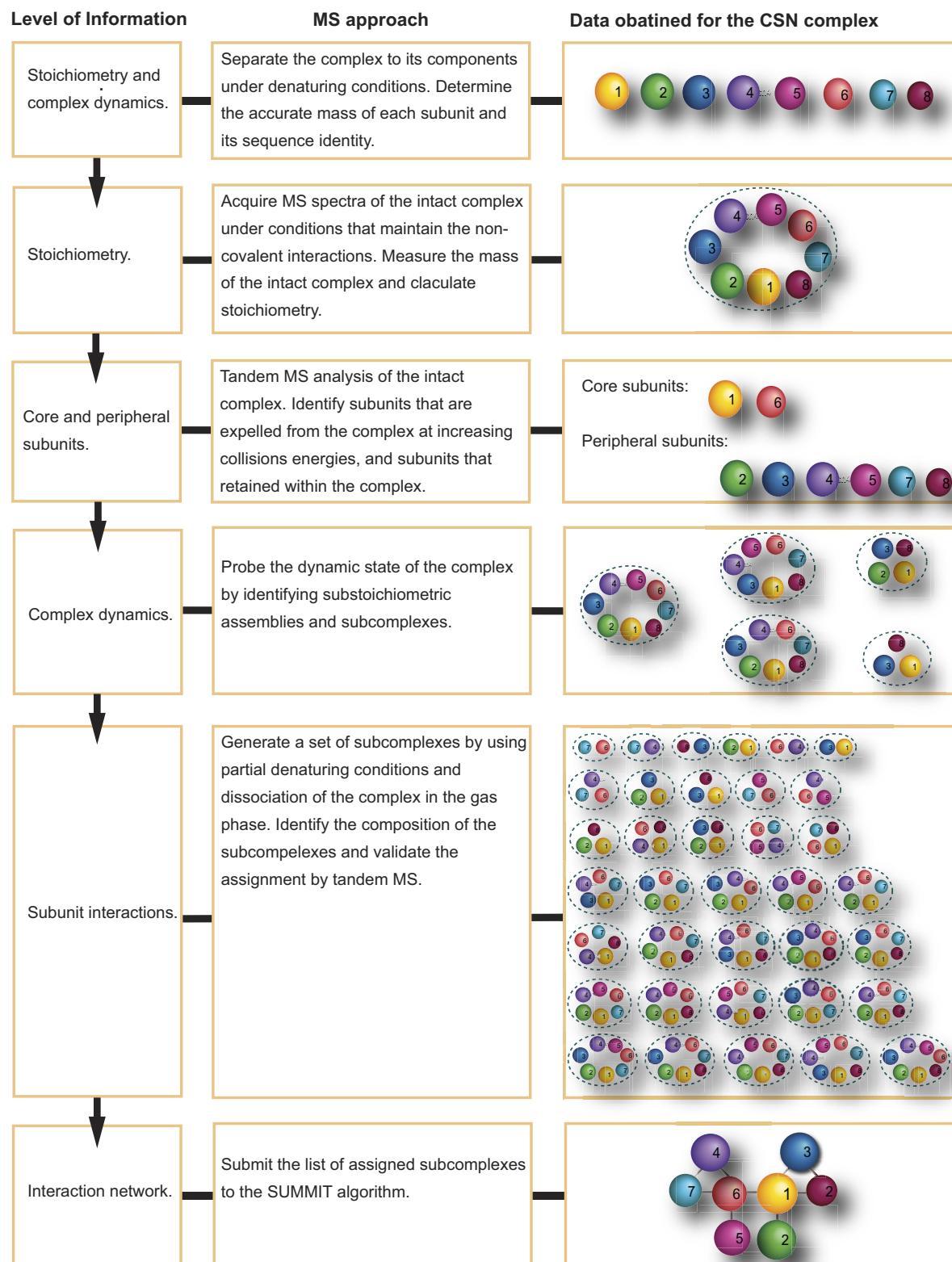
Once the composition of the complex has been identified, the next step is to define the mass of the intact complex, and the relative abundance of each subunit. Often, this can be a challenging task. To enable a more efficient and comprehensive analysis of complex mass spectra, computational algorithms were developed. One way to analyze MS data is to deconvolute the spectra by means of the software program MaxEnt [39, 40]. This software uses maximum entropy to provide protein profiles such that the actual mass is determined rather than the observed m/z results. The MaxEnt maximum entropy approach can produce very good results; however, it is not always possible to deconvolute very complex mass spectra in which the measured ions from different species often overlap. Moreover, when measuring a wide mass range spectrum, data processing can become time-consuming. In general, the deconvolution approach can greatly simplify data analysis, and help to define the intact mass of the complex, though it does not provide information on the stoichiometry of its protein components.

The determination of subunit copy number may be achieved by means of the summing masses for interaction topology (SUMMIT) [41] algorithm which, given the molecular masses of subunits, calculates all of their possible compositions. This algorithm performs an exhaustive search for masses that together, add up to the target mass over all allowable boundaries within a given error range. The output of this routine is a list of possible combinations of subunits that sum up to a similar mass. An elegant way to assist and validate such analyses, is to use the solving complex macromolecular mass spectra (SOMMS) program, which can simulate and subsequently resolve complex mass spectra [42]. Developed by the Heck group, the SOMMS program uses a multinomial distribution with Gaussian curve-fitting to simulate putative mass spectra of protein complexes and subcomplexes. The SOMMS program requires prior knowledge of factors such as potential macromolecular complexes in solution, molecu-

lar masses of subunits, and predicted numbers of charges, to correctly simulate charge-state distribution for the candidate complexes. Overall, unlike the field of proteomics, in which the availability of various software

programs enables automated data analysis, structural MS now is far from such a goal.

Even with the aid of computational algorithms, when the masses of the protein subunits are so similar that their



mass difference divided by the total mass of the protein complex is less than the peak width at half-height, defining their stoichiometry is a very difficult task. This difficulty is more pronounced in large protein complexes, in which the peak width is broadened due to incomplete desolvation, and adherence of buffer and solvent adducts [43]. In such cases, genetic manipulations may be introduced, to increase or decrease the mass of one of the subunits, by either introducing N- or C-terminal truncations or adding a protein tag [44, 45]. However, such genetically engineered constructs may potentially interfere with the biological activity of the assembly. Therefore, it is desirable to validate the activity of such modified complexes, before MS analysis.

Core and Peripheral Subunits

When the composition, stoichiometry, and exact mass of each of the subunits are known, identifying core and peripheral subunits is the next step in the structural analysis of the complex. Subunits that are located at the periphery are more exposed; as a consequence, they have fewer interaction partners, and their binding interface is smaller [31, 46]. Therefore, these subunits are often the first to undergo unfolding and dissociation when protein complexes are activated during the tandem MS process [31, 47–49]. A stepwise increase of the accelerating voltage during the MS/MS analysis, and subsequent analysis of the dissociation products, will pinpoint peripheral constituents. While core subunits will be retained within the complex (Figure 3a). This method, although crude, is quite effective, as has been proven by the growing number of MS studies on macromolecular complexes with known high-resolution structures [47, 48, 50]. However, it is currently limited to less than 10 dissociation steps (depending on the complex), and so cannot uncover the subunit architecture of larger, multi-component assemblies.

The key questions that arise are whether new mass spectrometers that enable the user to achieve higher activation energies would increase the sequential dissociation steps, or whether there are other limits which

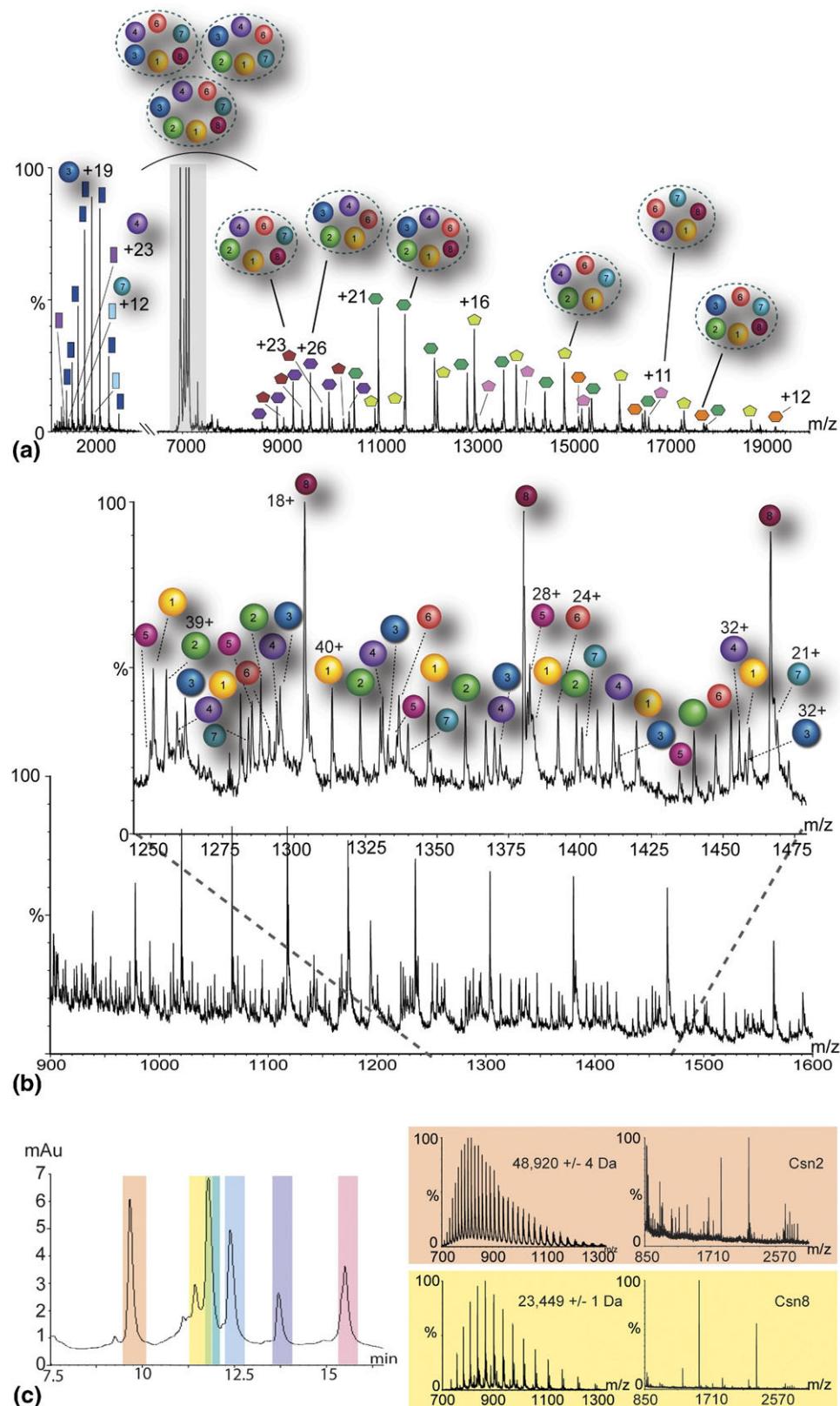
are not yet understood. In a recent study, a QTOF instrument capable of accessing high activation energies was used to examine the dissociation pathway of HSP16.5, a 24-mer complex, and up to four dissociation steps of monomeric subunits were observed [51]. Even when high collision energies were employed, a fifth dissociation step was not induced; rather, the excess energy caused the monomeric subunits to fragment into peptides. This study suggests that the extent of noncovalent dissociation by CID is limited by the amount of charge available, and not merely by the activation energy. These results raise the question of whether the number of subunits, or the size of the stripped subunits, is the limiting factor in defining the core of the complex. Given that charge partitioning during the CID process is proportional to the surface area of the unfolded subunits [9], will large subunits siphon off more charge and, as a result, limit the number of possible dissociation states? In general, a possible route for overcoming the limited number of dissociation steps might be to increase the amount of charge allotted to the complete complex, by using a small amount of chemical adducts such as *m*-nitrobenzyl alcohol [52], or alternatively by changing the pH [53].

Another option would be to exploit alternative dissociation methods such as surface-induced dissociation (SID) [54–57]. In the SID activation strategy, a large amount of energy is deposited within the protein complexes in a very short time, resulting in more extensive dissociation than that caused by CID. Recent findings [56] indicated that unlike CID, protein assemblies exposed to SID tend to dissociate in such a way that symmetrical charge partitioning of the product ions is obtained. This is attributed to the high-energy of collisions, which leads to dissociation of subunits before large-scale structural rearrangements. It would be interesting to examine whether such symmetric dissociation would also occur in large protein complexes (greater than ~500 kDa), in which the deposited energy might be absorbed internally. Nevertheless, a comparison between SID and CID dissociation pathways might reveal information regarding the asymmetry of a com-

Figure 2. Generating a subunit interaction map of a protein complex is a multi-step process. The diagram describes the different steps taken, the structural information gained at each step, and the relevant MS approach. The process is illustrated in a study of the human COP9 signalosome complex (CSN) [6]. Under denaturing conditions, the eight subunits of the complex were separated chromatographically; the masses of the individual proteins, together with peptide sequencing, enabled identification of each subunit by means of its unique mass. Using a volatile buffer at physiologic pH and MS conditions optimized for preserving noncovalent interactions, the MS spectrum of the intact complex was acquired. The measured mass of the intact complex confirmed that all subunits were present at unit stoichiometry. The MS spectrum also revealed the integrity and stability of the complex. Substoichiometric complexes in which Csn2 and Csn5 were absent, and smaller subcomplexes corresponding to Csn1/2/3/8 and Csn1/3/8, were obtained. A series of MS/MS spectra revealed that unlike the other subunits, Csn1 and Csn6 are not dissociated from either complex, suggesting that these subunits are located within the core. A set of overlapping subcomplexes was then generated by stepwise addition of methanol, concomitantly with harsh MS conditions. MS and MS/MS analyses enabled the identification of 35 subcomplexes. This list of subcomplexes was then submitted to SUMMIT for determination of an interaction network. The data yielded a single interaction network, leading to a high confidence interaction map.

plex, or the existence of hinge regions with weaker protein-protein interactions. In general, SID complements information obtained by CID experiments [54–57]. A future challenge would be to combine these

dissociation methods within a single mass spectrometer, to facilitate comprehensive data analysis. Moreover, I anticipate that the ability of lasers to induce assembly dissociation will also be explored.



Subunit–Subunit Interactions

Many large macromolecular complexes are composed of subcomplexes that come together to perform a function. Identifying the interactions between the subunits of these complexes and generating their topological arrangements is similar to creating a mosaic: the pieces of information linking the small subcomplexes are assembled, to provide a full description of the intact complex organization (Figure 2). In practice, the complex is initially fragmented into a set of subcomplexes. The composition of each subcomplex is defined, and the overlapping components are used to generate a detailed subunit interaction map. Recently, the subunit architecture of several complexes was defined by means of this approach [6, 34, 35, 38]. In most cases, the complexes are disrupted by partial denaturation in solution, through changing the pH, increasing the ionic strength, or adding organic solvents to the solution [58]. Active dissociation of the complex may also be induced in the gas-phase by collisional activation methods. MS and MS/MS analyses are then applied to identify the composition of the subcomplexes generated. Subsequently, the list of assigned subcomplexes is used to build the interaction map.

Given the large number of possible spatial conformations, defining the unique subunit organization map of a complex is not trivial. For example, an eight-subunit complex can be arranged in 2^{28} adjacency structures. Even if topologies that are not physically feasible are removed it is still a large number of structural possibilities. Therefore, utilizing computational algorithms would greatly assist in the generation of the protein interaction network. Such an algorithm

should enable sampling of all the possible interactions between subunits compatible with the observed subcomplexes. SUMMIT, which uses a genetic algorithm to generate a protein interaction network [41], is an example of such a software program. SUMMIT takes a simplified approach, ignoring the 3D aspect of protein complexes, and deconstructs complexes into their smallest possible protein building blocks. The software then generates all viable interactions between these building blocks. A list of subcomplexes is submitted to SUMMIT to determine the interaction network in which the area of each subunit is scaled according to its mass; the connections are weighted according to the number of times they occur within the various networks. When performing such analyses, care must be taken not to overreach the goal by inferring incorrect connectivities. Zhou et al. used a similar approach, based on a different analytical algorithm, termed ‘minimal connectivity’ [38]. According to this algorithm, the full complex contains the smallest set of connections between subunits required to explain those present in multiple subcomplexes. This assignment strategy was performed manually; however, there is no doubt that the community would benefit from further development of these computational packages.

When using the overlapping sets of subcomplexes to generate a connectivity map, the following points should be considered: if a large set of subcomplexes has been generated, it is recommended that only a fraction of them be used for generalizing the connectivity map, and all those remaining used for structure validation [6]. Moreover, within the pool of generated subcomplexes, the distribution of dimers, trimers, tetramers,

Figure 3. Methods for defining the exact mass of individual subunits. Three different methods for mass measurements of subunits are illustrated on the COP9 signalosome (CSN), an eight subunit (Csn1-8) complex. (a) MS/MS analysis of subcomplexes that comprise the CSN complex. The activation of the complex supports removal of associated water or buffer molecules, and consequently increases the accuracy of mass measurement of expelled subunits. Data were obtained by isolating a mixed population of ions at $7000\text{ }m/z$ ($\Delta\text{Csn}5$, $\Delta\text{Csn}2\Delta\text{Csn}5$, and $\Delta\text{Csn}5\Delta\text{Csn}8$), giving rise to the dissociation of individual subunits Csn3, Csn4, Csn7, and Csn8 in the low- m/z region (Csn8 is not labeled in the Figure). In the high- m/z region of the spectrum, charge states of both hexamers and pentamers were identified, all containing Csn1 and Csn6, suggesting that these subunits are located within the core of the complex [6]. The MS analysis was performed in the presence of 1 M ammonium acetate (pH 7.5) and 12.5% methanol (vol/vol). (b) A nanoflow ESI spectrum was measured from a denatured solution of the COP9 complex. The complex was eluted from a C₄ ZipTip column (50% acetonitrile and 0.1% formic acid). All eight subunits of the COP9 complex could be identified by this method. However, applying this approach to the analysis of multi-component complexes is not straightforward, for two reasons: the complexity of the spectrum, which complicates the assignment; and the fact that the method catalogues all proteins within the sample, and the identification is based on their theoretical masses. However, if a protein subunit has been modified only to some extent, it might be difficult to define its identity. (c) Capillary-LC separation of COP9 subunits (in the absence of Csn5) coupled with MS analysis. Subunit separation was carried out using a monolithic column (initial solution conditions: 10% acetonitrile, 90% H₂O, and 0.05% trifluoroacetic acid (TFA)). Peaks representing the different isolated subunits are color-coded (left panel). This separation was repeated under identical conditions, and each peak was analyzed by ESI-QTOF, to determine the molecular weight of the intact protein subunit, and by MALDI-TOF/TOF, following tryptic digestion of the fractionated proteins, for protein identification. The chromatographic separation was highly reproducible, enabling the correlation of each LC peak with its mass and identity [6]. An example of two of the resulting ESI-QTOF (left) and MALDI TOF/TOF (right) spectra are shown in the right panel. The identified subunits and the masses measured are labeled.

etc. is critical for unambiguous clarification of the complex topology. For example, generating 12 tetrameric subcomplexes would only yield 75% of the subunit interactions for a 10-mer complex, while 13 dimers would enable correct predictions of all interactions [41]. Thus, the minimum number of subcomplexes required to define explicit interaction determination depends on both the size of the subcomplexes, and the total number of subunits comprising the intact complex. Another point to consider is that structural enantiomers cannot be dictated by the constraints determined by MS; for example, a ring structure can run either clockwise or counterclockwise. In such cases, other experimental or computational methods could be brought into play, to choose between possibilities. Overall, defining subcomplex connectivity from MS data is still in its early days; further development is required to make this approach generic and applicable to a wide range of complexes.

Macromolecule Packing and Shape

While the subunit interaction map of a protein assembly provides valuable clues regarding the three-dimensional spatial arrangement of the complex, it cannot supply a full description of its shape. This information, however, may be obtained by applying ion mobility separation coupled to mass spectrometry (Synapt, Waters) [36, 37]. Ion mobility (IM) measures the time it takes for an ion to traverse an environment of gas molecules of neutral charge, under the influence of a weak electric field [59]. Large ions will experience more collisions with the background gas and, consequently, travel for longer (drift time or t_D). The measured drift time is proportional to the collision cross-section (Ω) of the ion and, in turn, can be related to its quaternary structure. Acquiring IM-MS data yields a three-dimensional spectrum containing information regarding the mass-to-charge ratio, abundance of ions, and drift time. Consequently, the ion mobility capability not only provides structural information, but also enables a decrease in the spectral overlap by distributing the data into a third dimension, thereby enabling analysis of heterogeneous complexes with very similar compositions. This feature is nicely demonstrated in a recent IM/MS study, which highlighted the co-existence of two distinct conformational states of hepatitis B virus capsids [19].

The correlation between the drift time of an ion and its Ω in an IM/MS instrument is based on a calibration curve generated using a set of proteins with defined cross-sections [60]. Thus, once the drift time of the complex of interest is measured, its Ω can be calculated from the calibration curve fit [61, 62]. Computational methods and modeling techniques are then utilized, to relate the measured cross-section to the quaternary structure of the protein complex.

The field of structure elucidation by ion mobility is still in its infancy, and methods for data interpretation

and prediction of structural models must still be developed. A weak point of the current method is that Ω is defined by a calibration approach, rather than by absolute measurements. Currently, three proteins are routinely used for calibration: ubiquitin (8.5 kDa), cytochrome *c* (12.2 kDa), and myoglobin (16.9 kDa) [60]. The ability of these relatively low mass calibrant ions to produce high-precision Ω measurements for very large protein complexes is an issue that should be addressed. However, likely larger ions with measured cross sections (larger than myoglobin, $\sim 3500 \text{ \AA}^2$, would enable more precise Ω calculations.

Another key question in this field involves the resolution of the IM/MS method. How similar are the topologies that can be distinguished from one another? Is it possible to differentiate between the shape of a protein complex with or without an encapsulated substrate? Can small structural rearrangements induced by ligand binding be detected? In general, the resolution of ion mobility is defined as the centroid of the drift time distribution divided by the width of the distribution at half height ($t/\Delta t$). The current resolution of ion mobility measurements for large protein complexes is about 10–15. The higher the separation power, the smaller the Ω differences that can be distinguished by means of IM separation. One important parameter for resolving subunit packings is the number of subunits in an assembly. As the number of subunits increases, structural families diverge, making ion mobility separation more facile at a fixed resolution value. For instance, with a resolution about 10, it is possible to distinguish between ring and collapsed topologies only for complexes containing nine or more subunits [60]. Another factor that influences the ability to distinguish between different topologies is the mass of individual subunits: more massive subunits will, on average, increase the overall dimensions of the final quaternary structure adopted by the complex [60]. In summary, there is no doubt that the emerging field of IM/MS holds great promise for the study of large assemblies, and would greatly benefit from the development of more advanced technologies that will increase the ability to distinguish between different topological states.

Assembly Dynamics

Many protein assemblies display dynamic characteristics. While most structural biology methods provide static snapshots, a major advantage of the MS approach is that it is capable of monitoring such processes in real time. This capacity is driven by the fast time scale of the analysis (ms), and the ability to simultaneously detect multiple populations of protein assemblies. MS may therefore be used to elucidate the biogenesis pathway of a protein complex, since intermediate states along the assembly pathway may be captured [63, 64]. In such analyses, the decrease in the intensity of the monomers, and the increase in the intermediate and fully assembled states, is monitored. This data enables extraction of

information on the timescale and kinetics of the assembly processes, and the presence or absence of rate-limiting steps. Insights into the assembly formation pathway may also be gained by disassembling intact structures into their building blocks followed by detailed analyses of the generated subcomplexes [65]. Recently, to track the assembly pathway of two dodecameric chaperone proteins, Painter et al. introduced an automated approach for rapid and repeated monitoring of the reaction in real time [66]. The time resolution of this method is 32 s, far better than that achieved using offline methods, in which the dead time (i.e., the time between mixing of components and detection) is ~1 min.

Since MS is able to capture different species within heterogeneous ensembles, a glimpse into the dynamic properties of intact protein complexes may be obtained. For instance, in the study of the proteasome 19S lid [35] and the COP9 signalosome complex [6], a dynamic equilibrium between the intact form of the complexes and the stoichiometric subcomplexes was revealed. This observation adds another layer of information to that provided by static interaction maps, and may shed light on the molecular mechanism underlying the function of these complexes.

Another dynamic property of some protein complexes is their ability to freely exchange subunits. This process possibly helps to regulate their cellular function, and enables assemblies to change size and adopt different compositions [67]. Subunit exchange reactions are often spontaneous and rapid, and the identification and characterization of the intermediates along the route are technically challenging. Several experimental methods have been used to study such processes, including fluorescence resonance energy-transfer [68], native gel electrophoresis [69], and affinity chromatography [70]. These methods, however, require the use of a potentially invasive tag; moreover, they provide an “average” of the whole reaction mixture, or result in a poor resolution of separation, or a combination thereof [66]. MS is ideally suited for monitoring such reactions, as the mass-resolving detection system enables label-free monitoring of reactants, products, and intermediate species in the reaction, without the need to separate these components before analysis [24, 66, 71].

The use of ion mobility coupled with mass spectrometry for the analysis of macromolecular assembly pathways is an area waiting to be explored. By applying such an approach, an increase in size of the suboligomeric states could be directly correlated with shape. Integration of these results with computational modeling would provide a comprehensive view of the assembly process. Such MS tools could also be used to characterize the dynamic reorganization of protein complexes that takes place in response to different signaling events.

Hybrid Methods

In some cases, the data generated by the steps described above are not sufficient for defining the subunit architecture of a protein complex. This can happen in highly complex samples, due to limited amounts of sample and low quality of spectra. In such scenarios, it is suggested that other experimental approaches be integrated with MS. For example, methods that provide pairwise protein interactions such as co-immunoprecipitation, the yeast two-hybrid system, fluorescence resonance energy-transfer (FRET), or protein arrays, can offer valuable input. These spatial constraints can be then integrated with the MS results acquired using the SUMMIT algorithm [41].

Chemical cross-linking is another approach that can readily be coupled with MS. In this method, a bi-functional reagent containing two reactive groups that can react with the side chains of amino acids such as amine or carboxylic acid, covalently binds two interacting partners [72, 73]. This technique may be utilized to stabilize protein complexes before MS analysis [74], and enable the use of MALDI MS for the analysis of noncovalent complexes [75–77]. Moreover, cross-linking can provide information on the identity of neighboring subunits, by subjecting the partially cross-linked sample to SDS separation, tryptic digestion, and proteomic analysis (Figure 4). Identifying the cross-linked sites themselves provides additional information. As the location of the created cross-links imposes a distance constraint on the location of the respective side chains, it is possible to draw conclusions as to the distance geometries of individual proteins located within a protein complex [78–81]. The cross-linker arm can also act as a ruler to map spatial proximities of amino acids in proteins [80]. Consequently, distance restraints derived from cross-linking experiments may then be used, much like the distance restraints obtained by the nuclear Overhauser effect (NOE) in NMR studies. De novo modeling based on distance restraints from cross-linking, coupled with MS analysis, may presently seem far-fetched. However, the contributions of this technology may effectively advance structural research, especially when cross-links are used as restraint data, in conjunction with structural information from other sources, as well as from computer modeling.

The ultimate goal in structural analysis of protein assemblies is to obtain models of intact complexes at atomic resolution. Although MS analysis can tolerate large, heterogeneous, asymmetric, and dynamic protein complexes, which are often very challenging for NMR and X-ray crystallography studies, it cannot provide such structures at high resolution. Therefore, a major challenge is to integrate the spatial restraints generated from MS analysis with other structural biology approaches, to obtain atomic models. Such integration of data is generally achieved by means of computational approaches. Along these lines, an atomic model of the 10-subunit yeast exosome was constructed in an elegant

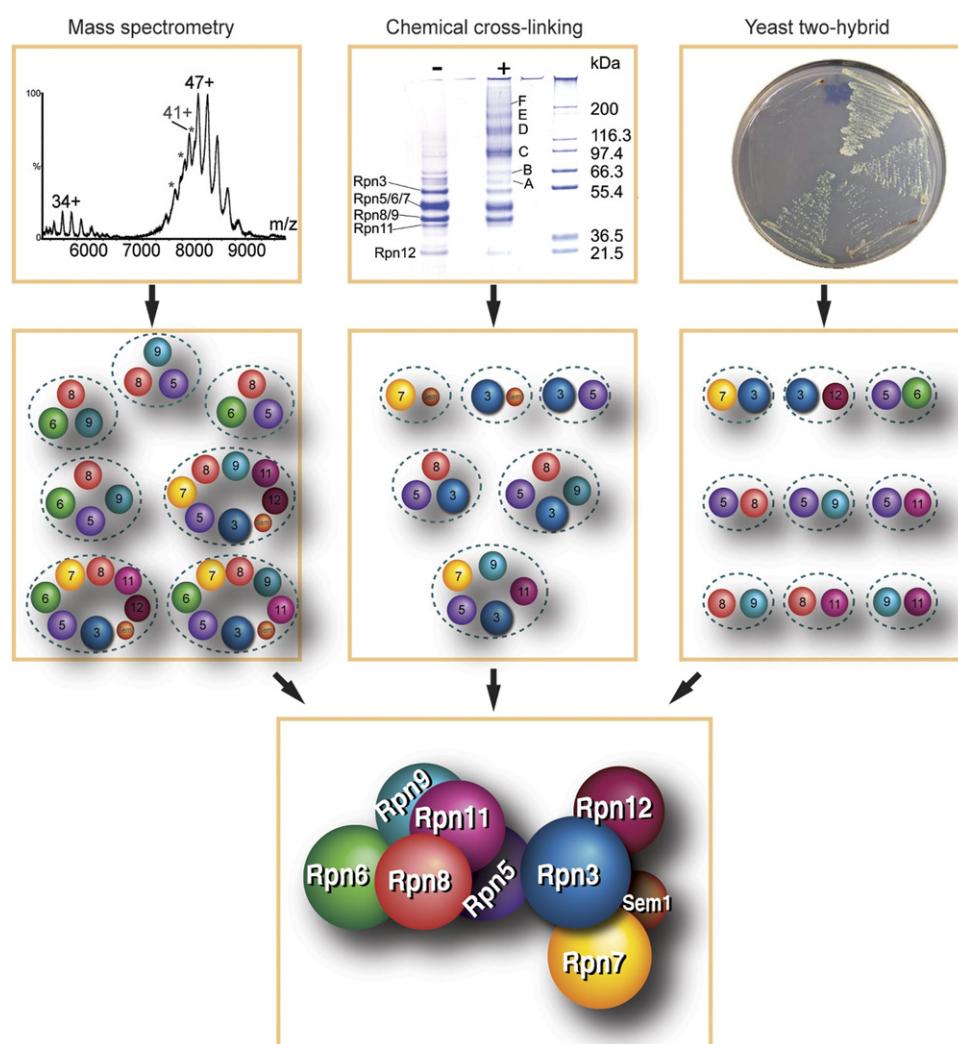


Figure 4. A hybrid method approach was taken for structural elucidation of the 19S proteasome lid complex [33]. Information generated from structural MS and chemical cross-linking was combined with yeast two-hybrid data (top boxes). The protein interactions identified in each step (middle boxes) were integrated to generate a comprehensive interaction map (bottom box).

study that combined MS analysis with homology modeling [34, 41]: the subunit network of the complex, generated from MS data, was used as a template for homology calculations based on the known structure of the archaeal exosome. Currently, the fundamental prerequisites for such hybrid approaches are accessible homologous proteins with known atomic structures. Although these requirements restrict its application to a limited number of complexes, likely soon, as more structures are solved, this will no longer be the case.

Envisioning the Future

The field of structural mass spectrometry has come a long way since the earliest protein complexes were shown to be maintained within the mass spectrometer [82]. Today, the method has its own niche within the field of structural biology: over the last few years, a considerable body of evidence has emerged regarding

the contribution of MS to the structural elucidation of protein complexes [9–13]. I would therefore like to end this critical insight by offering a perspective on the possible array of future directions the field may take.

Studying Membrane Protein Complexes

While it is accepted that membrane-associated protein complexes perform a wide range of essential cellular functions ranging from signaling to nutrient uptake, trafficking, lipid biosynthesis, and ion homeostasis, the study of their structure still remains one of the most challenging aspects of structural biology. Though MS may contribute extensively to this field, the challenges still remain. First, the transition into the gas phase usually leads to the disruption of interactions, particularly between cytoplasmic and membrane subunits. Secondly, the mass spectrum tends to be dominated by large aggregates of detergent molecules. Moreover, the

poor solubility of membrane proteins in aqueous buffers often causes blockage in the electrospray needle, whereas the large quantities of detergent suppress the protein signal.

Nevertheless, exciting progress was recently made when a new strategy was utilized to study membrane proteins [83–86]. This approach involves the application of MS to a micellar solution: the introduction of energy facilitates the desolvation process, and induces dissociation of detergent protein clusters. This method enables determination of the oligomeric state of membrane complexes. In the coming years, I anticipate that the application of MS to the study of membrane proteins will expand, and new strategies will be developed to enable it to not only reveal the assembly state of membrane assemblies, but also their comprehensive interaction maps and overall shape.

Elucidating the Diversity of Protein Complexes

Traditionally, protein complexes are seen as largely uniform in both composition and functional capacity. However, accumulating evidence suggests that the cell's population of a specific protein complex actually represents a diverse group of functionally distinct entities. The capacity of protein complexes to convert between oligomeric forms, as well as to generate novel compositions may be highly developed and subject to sophisticated modes of regulation [87, 88]. Therefore, there is a real need to develop novel experimental approaches that would make it possible to view protein complexes as dynamic entities which alter their composition from one tissue to another, between normal versus diseased cells, or even at different locations and time points within a single cell. I anticipate that in the coming years, together with technological developments, structural MS will play a key role in the analysis of coexisting structural states of protein complexes.

Revealing the Interaction Network Between Protein Complexes

Every major process in the cell is carried out by large assemblies of protein molecules [89]. Recently, increasing evidence indicates that these functional machines also interact with each other. For example, the translation and degradation machineries were shown to be linked [90], as well as the degradation and transcription machineries [91]. The entire cellular entity may be viewed as a factory containing an elaborate network of interactions between macromolecular complexes. Revealing the physical interactions between these protein assemblies, identifying the common subunits, and elucidating the mechanisms of communication and cooperation between them, is crucial for understanding their integrated and coordinated function. The challenge will be to recognize the specificity of these numerous inter-

actions, and to eliminate contaminant proteins or non-specific associations.

Structural mass spectrometry is capable of resolving specific interactions between protein complexes; however, the methodology depends on the strength and nature of these interactions. Interactions with micromolar dissociation constants are clearly detected by MS; however, even weakly bound complexes with millimolar dissociation constants can often survive the ESI process [92, 93]. Another point to consider is that MS is biased toward detecting electrostatic interactions, as hydrophobic interfaces are weakened in the gas phase [94]. Therefore, if hydrophobic interactions dominate the interface between complexes, it may not be possible to maintain/detect the interactions. Nevertheless, the analysis of multi-protein complexes relies heavily on purification techniques. Progress in the biochemical isolation of protein complexes is necessary, so as to maintain the weak interactions between assemblies under conditions compatible with MS.

Characterizing Protein Complexes from Cell Lines

Currently, sample consumption for MS studies of non-covalent complexes is in the picomole range; as a result, the method is mainly implemented in recombinant proteins or complexes endogenous to bacteria or yeast. Scaling up the production of cells in higher eukaryotes to isolate their protein complexes would be the natural way to enable their analysis. However, the facilities required for such massive production are limited in the academic environment; moreover, the process itself is very expensive. In the future, such limitations may be overcome by increasing the sensitivity of mass spectrometers. These technological developments will enable detection and analysis of cellular complexes present in the femtomole range, and will pave the way toward analysis of less abundant protein assemblies and complexes isolated directly from cells of higher eukaryotes, greatly enhancing the biological relevance of the results. Overall, a development that will diminish the bottleneck of sample availability may resemble the revolution that occurred in the field of X-ray crystallography, with the introduction of robots that reduced by a factor of 10 the amount of protein needed for crystallization [95, 96] and, as a consequence, extensively expanded the application of the method.

Overall, many challenging frontiers lie in the way of understanding the integrated and regulated function of cellular machineries, though clearly, structural MS will play a critical role in such studies. It is hard to envision what scientific revolutions will shape the future capabilities of this method, when only two decades ago, such a field could not even have been imagined. I anticipate that in the coming years, innovative technological developments, side-by-side with novel methodological approaches and computational methods, will

certainly broaden the scope of application of structural MS, and make new research directions feasible.

Acknowledgments

The author thanks Maya Schuldiner, Ron Milo, Amos Tanay, and Gideon Schreiber (Weizmann Institute of Science) for their critical review, and the Sharon group members for their contribution to the manuscript. She is grateful for the support of the Morasha Program, Israel Science Foundation (grant no. 1823/07), the Josef Cohn Minerva Center for Biomembrane Research, the Chais Family Fellows Program for New Scientists, the Abraham and Sonia Rochlin Foundation, the Wolfson Family Charitable Trust, the Helen and Milton A. Kimmelman Center for Biomolecular Structure and Assembly, the estate of Shlomo and Sabine Beirzinsky; Meil de Botton Aynsley and Karen Siem, UK.

References

- Lander, E. S.; Linton, L. M.; Birren, B.; Nusbaum, C.; Zody, M. C.; Baldwin, J.; Devon, K.; Dewar, K.; Doyle, M.; FitzHugh, W.; Funke, R.; Gage, D.; Harris, K.; Heaford, A.; Howland, J.; Kann, L.; Lehoczky, J.; LeVine, R.; McEwan, P.; McKernan, K.; Meldrim, J.; Mesirov, J. P.; Miranda, C.; Morris, W.; Naylor, J.; Raymond, C.; Rosetti, M.; Santos, R.; Sheridan, A.; Sougnez, C.; Stange-Thomann, N.; Stojanovic, N.; Subramanian, A.; Wyman, D.; Rogers, J.; Sulston, J.; Ainscough, R.; Beck, S.; Bentley, D.; Burton, J.; Cleo, C.; Carter, N.; Coulson, A.; Deadman, R.; Deloukas, P.; Dunham, A.; Dunham, I.; Durbin, R.; French, L.; Graham, D.; Gregory, S.; Hubbard, T.; Humphray, S.; Hunt, A.; Jones, M.; Lloyd, C.; McMurray, A.; Matthews, L.; Mercer, S.; Milne, S.; Mullikin, J. C.; Mungall, A.; Plumbe, R.; Ross, M.; Shownkeen, R.; Sims, S.; Waterston, R. H.; Wilson, R. K.; Hillier, L. W.; McPherson, J. D.; Marra, M. A.; Mardis, E. R.; Fulton, L. A.; Chinwalla, A. T.; Pepin, K. H.; Gish, W. R.; Chissoe, S. L.; Wendt, M. C.; Delehaunty, K. D.; Miner, T. L.; Delehaunty, A.; Kramer, J. B.; Cook, L. L.; Fulton, R. S.; Johnson, D. L.; Minx, P. J.; Clifton, S. W.; Hawkins, T.; Branscomb, E.; Predki, P.; Richardson, P.; Wenning, S.; Slezak, T.; Doggett, N.; Cheng, J. F.; Olsen, A.; Lucas, S.; Elkin, C.; Überbacher, E.; Fraizer, M.; Gibbs, R. A.; Muzny, D. M.; Scherer, S. E.; Bouck, J. B.; Sodergren, E. J.; Worley, K. C.; Rives, C. M.; Gorrell, J. H.; Metzker, M. L.; Naylor, S. L.; Kucherlapati, R. S.; Nelson, D. L.; Weinstock, G. M.; Sakaki, Y.; Fujiyama, A.; Hattori, M.; Yada, T.; Toyoda, A.; Itoh, T.; Kawagoe, C.; Watanabe, H.; Totoki, Y.; Taylor, T.; Weissenbach, J.; Heilig, R.; Saurin, W.; Artiguenave, F.; Brottier, P.; Bruls, T.; Pelletier, E.; Robert, C.; Wincker, P.; Smith, D. R.; Doucette-Stamm, L.; Rubenfeld, M.; Weinstock, K.; Lee, H. M.; Dubois, J.; Rosenthal, A.; Platzer, M.; Nyakatura, G.; Taudien, S.; Rump, A.; Yang, H.; Yu, J.; Wang, J.; Huang, G.; Gu, J.; Hood, L.; Rowen, L.; Madan, A.; Qin, S.; Davis, R. W.; Federspiel, N. A.; Abola, A. P.; Proctor, M. J.; Myers, R. M.; Schmutz, J.; Dickson, M.; Grimwood, J.; Cox, D. R.; Olson, M. V.; Kaul, R.; Shimizu, N.; Kawasaki, K.; Minoshima, S.; Evans, G. A.; Athanasiou, M.; Schultz, R.; Roe, B. A.; Chen, F.; Pan, H.; Ramser, J.; Lehrach, H.; Reinhardt, R.; McCombie, W. R.; de la Bastide, M.; Dedhia, N.; Blocker, H.; Hornischer, K.; Nordsiek, G.; Agarwala, R.; Aravind, L.; Bailey, J. A.; Bateman, A.; Batzoglou, S.; Birney, E.; Bork, P.; Brown, D. G.; Burge, C. B.; Cerutti, L.; Chen, H. C.; Church, D.; Clamp, M.; Copley, R. R.; Doerks, T.; Eddy, S. R.; Eichler, E. E.; Furey, T. S.; Galagan, J.; Gilbert, J. G.; Harmon, C.; Hayashizaki, Y.; Haussler, D.; Hermjakob, H.; Hokamp, K.; Jang, W.; Johnson, L. S.; Jones, T. A.; Kasif, S.; Kasprzyk, A.; Kennedy, S.; Kent, W. J.; Kitts, P.; Koonin, E. V.; Korf, I.; Kulp, D.; Lancet, D.; Lowe, T. M.; McLysaght, A.; Mikkelsen, T.; Moran, J. V.; Mulder, N.; Pollara, V. J.; Ponting, C. P.; Schuler, G.; Schultz, J.; Slater, G.; Smit, A. F.; Stupka, E.; Szustakowski, J.; Thierry-Mieg, D.; Thierry-Mieg, J.; Wagner, L.; Wallis, J.; Wheeler, R.; Williams, A.; Wolf, Y. I.; Wolfe, K. H.; Yang, S. P.; Yeh, R. F.; Collins, F.; Guyer, M. S.; Peterson, J.; Felsenfeld, A.; Wetterstrand, K. A.; Patrinos, A.; Morgan, M. J.; de Jong, P.; Catanese, J. J.; Osoegawa, K.; Shizuya, H.; Choi, S.; Chen, Y. J. Initial Sequencing and Analysis of the Human Genome. *Nature* **2001**, *409*, 860–921.
- Venter, J. C.; Adams, M. D.; Myers, E. W.; Li, P. W.; Mural, R. J.; Sutton, G. G.; Smith, H. O.; Yandell, M.; Evans, C. A.; Holt, R. A.; Gocayne, J. D.; Amanatides, P.; Ballew, R. M.; Huson, D. H.; Wortman, J. R.; Zhang, Q.; Kodira, C. D.; Zheng, X. H.; Chen, L.; Skupski, M.; Subramanian, G.; Thomas, P. D.; Zhang, J.; Gabor Miklos, G. L.; Nelson, C.; Broder, S.; Clark, A. G.; Nadeau, J.; McKusick, V. A.; Zinder, N.; Levine, A. J.; Roberts, R. J.; Simon, M.; Slayman, C.; Hunkapiller, M.; Bolanos, R.; Delcher, A.; Dew, I.; Fasulo, D.; Flanigan, M.; Florea, L.; Halpern, A.; Hannenhalli, S.; Kravitz, S.; Levy, S.; Mobarry, C.; Reinert, K.; Remington, K.; Abu-Threideh, J.; Beasley, E.; Biddick, K.; Bonazzi, V.; Brandon, R.; Cargill, M.; Chandramouliswaran, I.; Charlab, R.; Chaturvedi, K.; Deng, Z.; Di Francesco, V.; Dunn, P.; Eilbeck, K.; Evangelista, C.; Gabrielian, A. E.; Gan, W.; Ge, W.; Gong, F.; Gu, Z.; Guan, P.; Heiman, T. J.; Higgins, M. E.; Ji, R. R.; Ke, Z.; Ketchum, K. A.; Lai, Z.; Lei, Y.; Li, Z.; Li, J.; Liang, Y.; Lin, X.; Lu, F.; Merkulov, G. V.; Milshina, N.; Moore, H. M.; Naik, A. K.; Narayan, V. A.; Neelam, B.; Nusskern, D.; Rusch, D. B.; Salzberg, S.; Shao, W.; Shue, B.; Sun, J.; Wang, Z.; Wang, A.; Wang, X.; Wang, J.; Wei, M.; Wides, R.; Xiao, C.; Yan, C.; Yao, A.; Ye, J.; Zhan, M.; Zhang, W.; Zhang, H.; Zhao, Q.; Zheng, L.; Zhong, F.; Zhong, W.; Zhu, S.; Zhao, S.; Gilbert, D.; Baumhueter, S.; Spier, G.; Carter, C.; Cravchik, A.; Woodage, T.; Ali, F.; An, H.; Awe, A.; Baldwin, D.; Baden, H.; Barnstead, M.; Barlow, I.; Beeson, K.; Busam, D.; Carver, A.; Center, A.; Cheng, M. L.; Curry, L.; Danaher, S.; Davenport, L.; Desilets, R.; Dietz, S.; Dodson, K.; Douc, L.; Ferriera, S.; Garg, N.; Gluecksmann, A.; Hart, B.; Haynes, J.; Haynes, C.; Heiner, C.; Hladun, S.; Hostin, D.; Houck, J.; Howland, T.; Ibegwam, C.; Johnson, J.; Kalush, F.; Kline, B.; Koduru, S.; Love, A.; Mann, F.; May, D.; McCawley, S.; McIntosh, T.; McMullen, I.; Moy, M.; Moy, L.; Murphy, B.; Nelson, K.; Pfannkoch, C.; Pratts, E.; Puri, V.; Qureshi, H.; Reardon, M.; Rodriguez, R.; Rogers, Y. H.; Romblad, D.; Ruhfel, B.; Scott, R.; Sitter, C.; Smallwood, M.; Stewart, E.; Strong, R.; Suh, E.; Thomas, R.; Tint, N. N.; Tse, S.; Vech, C.; Wang, G.; Wetter, J.; Williams, S.; Williams, M.; Windsor, S.; Winn-Deen, E.; Wolfe, K.; Zaveri, J.; Zaveri, K.; Abril, J. F.; Guigo, R.; Campbell, M. J.; Sjolander, K. V.; Karlak, B.; Kejariwal, A.; Mi, H.; Lazareva, B.; Hatton, T.; Narechania, A.; Diemer, K.; Muruganujan, A.; Guo, N.; Sato, S.; Bafna, V.; Istrail, S.; Lippert, R.; Schwartz, R.; Walenz, B.; Yoosoph, S.; Allen, D.; Basu, A.; Baxendale, J.; Blick, L.; Caminha, M.; Carnes-Stine, J.; Caulk, P.; Chiang, Y. H.; Coyne, M.; Dahlke, C.; Mays, A.; Dombroski, M.; Donnelly, M.; Ely, D.; Esperham, S.; Fosler, C.; Gire, H.; Golanowski, S.; Glasser, K.; Glodek, A.; Gorokhov, M.; Graham, K.; Gropman, B.; Harris, M.; Heil, J.; Henderson, S.; Hoover, J.; Jennings, D.; Jordan, C.; Jordan, J.; Kash, J.; Kagan, L.; Kraft, C.; Levitsky, A.; Lewis, M.; Liu, X.; Lopez, J.; Ma, D.; Majoros, W.; McDowell, M.; Pan, S.; Peck, J.; Peterson, M.; Rowe, W.; Sanders, R.; Scott, J.; Simpson, M.; Smith, T.; Sprague, A.; Stockwell, T.; Turner, R.; Venter, E.; Wang, M.; Wen, M.; Wu, D.; Wu, M.; Xia, A.; Zandieh, A.; Zhu, X. The Sequence of the Human Genome. *Science* **2001**, *291*, 1304–1351.
- Hart, G. T.; Ramani, A. K.; Marcotte, E. M. How Complete are Current Yeast and Human Protein-Interaction Networks? *Genome Biol.* **2006**, *7*, 120.
- Reguly, T.; Breitkreutz, A.; Boucher, L.; Breitkreutz, B. J.; Hon, G. C.; Myers, C. L.; Parsons, A.; Friesen, H.; Oughtred, R.; Tong, A.; Stark, C.; Ho, Y.; Botstein, D.; Andrews, B.; Boone, C.; Troyanskaya, O. G.; Ideker, T.; Dolinski, K.; Batada, N. N.; Tyers, M. Comprehensive Curation and Analysis of Global Interaction Networks in *Saccharomyces cerevisiae*. *J. Biol.* **2006**, *5*, 11.
- Stumpf, M. P.; Thorne, T.; de Silva, E.; Stewart, R.; An, H. J.; Lappe, M.; Wiuf, C. Estimating the Size of the Human Interactome. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 6959–6964.
- Sharon, M.; Mao, H.; Boeri Erba, E.; Stephens, E.; Zheng, N.; Robinson, C. V. Symmetrical Modularity of the COP9 Signalosome Complex Suggests Its Multifunctionality. *Structure* **2009**, *17*, 31–40.
- Robinson, C. V.; Sali, A.; Baumeister, W. The Molecular Sociology of the Cell. *Nature* **2007**, *450*, 973–982.
- Sali, A.; Glaeser, R.; Earnest, T.; Baumeister, W. From Words to Literature in Structural Proteomics. *Nature* **2003**, *422*, 216–225.
- Benesch, J. L.; Ruotolo, B. T.; Simmons, D. A.; Robinson, C. V. Protein Complexes in the Gas Phase: Technology for Structural Genomics and Proteomics. *Chem. Rev.* **2007**, *107*, 3544–3567.
- Heck, A. J. Native Mass Spectrometry: A Bridge Between Interactomics and Structural biology. *Nat. Methods* **2008**, *5*, 927–933.
- Loo, J. A. Studying Noncovalent Protein Complexes by Electrospray Ionization Mass Spectrometry. *Mass Spectrom. Rev.* **1997**, *16*, 1–23.
- Sharon, M.; Robinson, C. V. The Role of Mass Spectrometry in Structure Elucidation of Dynamic Protein Complexes. *Annu. Rev. Biochem.* **2007**, *76*, 167–193.
- van den Heuvel, R. H.; Heck, A. J. Native Protein Mass Spectrometry: From Intact Oligomers to Functional Machineries. *Curr. Opin. Chem. Biol.* **2004**, *8*, 519–526.
- Ganem, B.; Li, Y. T.; Henion, J. D. Detection of Noncovalent Receptor-Ligand Complexes by Mass Spectrometry. *J. Am. Chem. Soc.* **1991**, *113*, 6294–6296.
- Ganem, B.; Li, Y. T.; Henion, J. D. Observation of Noncovalent Enzyme-Substrate and Enzyme-Product Complexes by Ion-Spray Mass Spectrometry. *J. Am. Chem. Soc.* **1991**, *113*, 7818–7819.
- Karas, M.; Bahra, U.; Ingendoh, A.; Nordhoff, E.; Stahl, B.; Strupata, K.; Hillenkamp, F. Principles and Applications of Matrix-Assisted UV-Laser Desorption/Ionization Mass Spectrometry. *Anal. Chim. Acta* **1990**, *241*, 175–185.
- Videler, H.; Ilag, L. L.; McKay, A. R.; Hanson, C. L.; Robinson, C. V. Mass Spectrometry of Intact Ribosomes. *FEBS Lett.* **2005**, *579*, 943–947.
- Uetrecht, C.; Versluis, C.; Watts, N. R.; Roos, W. H.; Wuite, G. J.; Wingfield, P. T.; Steven, A. C.; Heck, A. J. High-Resolution Mass Spectrometry of Viral Assemblies: Molecular Composition and Stability of Dimorphic Hepatitis B Virus Capsids. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 9216–9220.
- Uetrecht, C.; Versluis, C.; Watts, N. R.; Wingfield, P. T.; Steven, A. C.; Heck, A. J. Stability and Shape of Hepatitis B Virus Capsids In Vacuo. *Angew. Chem. Int. Ed. Engl.* **2008**, *47*, 6247–6251.
- Morton, V. L.; Stockley, P. G.; Stonehouse, N. J.; Ashcroft, A. E. Insights into Virus Capsid Assembly from Noncovalent Mass Spectrometry. *Mass Spectrom. Rev.* **2008**, *27*, 575–595.
- van Duijn, E.; Simmons, D. A.; van den Heuvel, R. H.; Bakkes, P. J.; van Heerikhuijsen, H.; Heeren, R. M.; Robinson, C. V.; van der Vies, S. M.;

- Heck, A. J. Tandem Mass Spectrometry of Intact GroEL-Substrate Complexes Reveals Substrate-Specific Conformational Changes in the trans Ring. *J Am. Chem. Soc.* **2006**, *128*, 4694–4702.
22. van Duijn, E.; Heck, A. J.; van der Vies, S. M. Inter-Ring Communication Allows the GroEL Chaperonin Complex to Distinguish Between Different Substrates. *Protein Sci.* **2007**, *16*, 956–965.
 23. Sharon, M.; Witt, S.; Felderer, K.; Rockel, B.; Baumeister, W.; Robinson, C. V. 20S Proteasomes Have the Potential to Keep Substrates in Store for Continual Degradation. *J. Biol. Chem.* **2006**, *281*, 9569–9575.
 24. Aquilina, J. A.; Benesch, J. L.; Bateman, O. A.; Slingsby, C.; Robinson, C. V. Polydispersity of a Mammalian Chaperone: Mass Spectrometry Reveals the Population of Oligomers in α B-crystallin. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 10611–10616.
 25. Bich, C.; Zenobi, R. Mass Spectrometry of Large Complexes. *Curr. Opin. Struct. Biol.* **2009**, *19*, 632–639.
 26. Charvat, A.; Abel, B. How to Make Big Molecules Fly Out of Liquid Water: Applications, Features, and Physics of Laser Assisted Liquid Phase Dispersion Mass Spectrometry. *Phys. Chem., Chem. Phys.* **2007**, *9*, 3335–3360.
 27. Bolbach, G. Matrix-Assisted Laser Desorption/Ionization Analysis of Noncovalent Complexes: Fundamentals and Applications. *Curr. Pharm. Des.* **2005**, *11*, 2535–2557.
 28. Smith, D. L.; Zhang, Z. Probing Noncovalent Structural Interactions in Soybean Agglutinin by Electrofeatures of Proteins by Mass Spectrometry. *Mass Spectrom. Rev.* **1994**, *13*, 411–429.
 29. Smith, V. F.; Schwartz, B. L.; Randall, L. L.; Smith, R. D. Electrospray Mass Spectrometric Investigation of the Chaperone SecB. *Protein Sci.* **1996**, *5*, 488–494.
 30. McLuckey, S. A. Principles of Collisional Activation in Analytical Mass Spectrometry. *J Am. Soc. Mass Spectrom.* **1992**, *3*, 599–614.
 31. Benesch, J. L.; Aquilina, J. A.; Ruotolo, B. T.; Sobott, F.; Robinson, C. V. Tandem Mass Spectrometry Reveals the Quaternary Organization of Macromolecular Assemblies. *Chem. Biol.* **2006**, *13*, 597–605.
 32. Chernushevich, I. V.; Thomson, B. A. Collisional Cooling of Large Ions in Electrospray Mass Spectrometry. *Anal. Chem.* **2004**, *76*, 1754–1760.
 33. Sobott, F.; Hernandez, H.; McCammon, M. G.; Tito, M. A.; Robinson, C. V. A Tandem Mass Spectrometer for Improved Transmission and Analysis of Large Macromolecular Assemblies. *Anal. Chem.* **2002**, *74*, 1402–1407.
 34. Hernandez, H.; Dziembowski, A.; Taverner, T.; Seraphin, B.; Robinson, C. V. Subunit Architecture of Multimeric Complexes Isolated Directly from Cells. *EMBO Rep.* **2006**, *7*, 605–610.
 35. Sharon, M.; Taverner, T.; Ambroggio, X. I.; Deshaies, R. J.; Robinson, C. V. Structural Organization of the 19S Orotateosome Lid: Insights from MS of Intact Complexes. *PLoS Biol.* **2006**, *4*, e267.
 36. Giles, K.; Pringle, S. D.; Worthington, K. R.; Little, D.; Wildgoose, J. L.; Bateman, R. H. Applications of a Traveling Wave-Based Radio-Frequency Only Stacked Ring Ion Guide. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 2401–2414.
 37. Thalassinos, K.; Slade, S. E.; Jennings, K. R.; Scrivens, J. H.; Giles, K.; Wildgoose, J.; Hoyes, J.; Bateman, R. H.; Bowers, M. T. Ion Mobility Mass Spectrometry of Proteins in a Modified Commercial Mass Spectrometer. *Int. J. Mass Spectrom.* **2004**, *236*, 55–63.
 38. Zhou, M.; Sandercock, A. M.; Fraser, C. S.; Ridlova, G.; Stephens, E.; Schenauer, M. R.; Yokoi-Fong, T.; Barsky, D.; Leary, J. A.; Hershey, J. W.; Doudna, J. A.; Robinson, C. V. Mass Spectrometry Reveals Modularity and a Complete Subunit Interaction Map of the Eukaryotic Translation Factor eIF3. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 18139–18144.
 39. Ferrige, A. G.; Seddon, M. J.; Green, B. N.; Jarvis, S. A.; Skilling, J.; Staunton, J. Disentangling Electrospray Spectra with Maximum Entropy. *Rapid Commun. Mass Spectrom.* **1992**, *6*, 707–711.
 40. Ferrige, A. G.; Seddon, M. J.; Skilling, J.; Ordsmit, N. The Application of MaxEnt to High Resolution Mass Spectrometry. *Mass Spectrom.* **1992**, *6*, 765–770.
 41. Taverner, T.; Hernandez, H.; Sharon, M.; Ruotolo, B. T.; Matajk-Vinkovic, D.; Devos, D.; Russell, R. B.; Robinson, C. V. Subunit Architecture of Intact Protein Complexes from Mass Spectrometry and Homology Modeling. *Acc. Chem. Res.* **2008**, *41*, 617–627.
 42. van Breukelen, B.; Barendregt, A.; Heck, A. J.; van den Heuvel, R. H. Resolving Stoichiometries and Oligomeric States of Glutamate Synthase Protein Complexes with Curve Fitting and Simulation of Electrospray Mass Spectra. *Rapid Commun. Mass Spectrom.* **2006**, *20*, 2490–2496.
 43. McKay, A. R.; Ruotolo, B. T.; Ilag, L. L.; Robinson, C. V. Mass Measurements of Increased Accuracy Resolve Heterogeneous Populations of Intact Ribosomes. *J. Am. Chem. Soc.* **2006**, *128*, 11433–11442.
 44. Waugh, D. S. Making the Most of Affinity Tags. *Trends Biotechnol.* **2005**, *23*, 316–320.
 45. Brizzard, B. Epitope Yagging. *Biotechniques* **2008**, *44*, 693–695.
 46. Benesch, J. L. P.; Aquilina, J. A.; Ruotolo, B. T.; Sobott, F.; Robinson, C. V. Tandem Mass Spectrometry Reveals the Quaternary Organization of Macromolecular Assemblies. *Chem. Biol.* **2006**, *13*, 597–605.
 47. Fandrich, M.; Tito, M. A.; Leroux, M. R.; Rostom, A. A.; Hartl, F. U.; Dobson, C. M.; Robinson, C. V. Observation of the Noncovalent Assembly and Disassembly Pathways of the Chaperone Complex MtGimC by Mass Spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 14151–14155.
 48. Ilag, L. L.; Videeler, H.; McKay, A. R.; Sobott, F.; Fucini, P.; Nierhaus, K. H.; Robinson, C. V. Heptameric (L12)6/L10 Rather Than Canonical Pentameric Complexes are Found by Tandem MS of Intact Ribosomes from Thermophilic Bacteria. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 8192–8197.
 49. Zhang, Z.; Zheng, Y.; Mazon, H.; Milgrom, E.; Kitagawa, N.; Kish-Trier, E.; Heck, A. J.; Kane, P. M.; Wilkens, S. Structure of the Yeast Vacuolar ATPase. *J. Biol. Chem.* **2008**, *283*, 35983–35995.
 50. Loo, J. A.; Berhane, B.; Kaddis, C. S.; Wooding, K. M.; Xie, Y.; Kaufman, S. L.; Chernushevich, I. V. Electrospray Ionization Mass Spectrometry and Ion Mobility Analysis of the 20S Proteasome Complex. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 998–1008.
 51. Benesch, J. L.; Ruotolo, B. T.; Sobott, F.; Wildgoose, J.; Gilbert, A.; Bateman, R.; Robinson, C. V. Quadrupole-Time-of-Flight Mass Spectrometer Modified for Higher-Energy Dissociation Reduces Protein Assemblies to Peptide Fragments. *Anal. Chem.* **2009**, *81*, 1270–1274.
 52. Lomeli, S. H.; Yin, S.; Ogorzalek Loo, R. R.; Loo, J. A. Increasing Charge While Preserving Noncovalent Protein Complexes for ESI-MS. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 593–596.
 53. Kaltashov, I. A.; Abzalimov, R. R. Do Ionic Charges in ESI-MS Provide Useful Information on Macromolecular Structure? *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1239–1246.
 54. Beardsley, R. L.; Jones, C. M.; Galhena, A. S.; Wysocki, V. H. Noncovalent Protein Tetramers and Pentamers with “n” Charges Yield Monomers with n/4 and n/5 Charges. *Anal. Chem.* **2009**, *81*, 1347–1356.
 55. Galhena, A. S.; Dagan, S.; Jones, C. M.; Beardsley, R. L.; Wysocki, V. H. Surface-Induced Dissociation of Peptides and Protein Complexes in a Quadrupole/Time-of-Flight Mass Spectrometer. *Anal. Chem.* **2008**, *80*, 1425–1436.
 56. Wysocki, V. H.; Jones, C. M.; Galhena, A. S.; Blackwell, A. E. Surface-Induced Dissociation Shows Potential to be More Informative Than Collision-Induced Dissociation for Structural Studies of Large Systems. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 903–913.
 57. Wysocki, V. H.; Joyce, K. E.; Jones, C. M.; Beardsley, R. L. Surface-Induced Dissociation of Small Molecules, Peptides, and Noncovalent Protein Complexes. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 190–208.
 58. Hernandez, H.; Robinson, C. V. Determining the Stoichiometry and Interactions of Macromolecular Assemblies from Mass Spectrometry. *Nat. Protoc.* **2007**, *2*, 715–726.
 59. Mason, E. R.; McDaniel, E. W. Transport Properties of Ions in Gases; John Wiley and Sons: New York, 1988.
 60. Ruotolo, B. T.; Benesch, J. L.; Sandercock, A. M.; Hyung, S. J.; Robinson, C. V. Ion Mobility-Mass Spectrometry Analysis of Large Protein Complexes. *Nat. Protoc.* **2008**, *3*, 1139–1152.
 61. Ruotolo, B. T.; Giles, K.; Campuzano, I.; Sandercock, A. M.; Bateman, R. H.; Robinson, C. V. Evidence for Macromolecular Protein Rings in the Absence of Bulk Water. *Science* **2005**, *310*, 1658–1661.
 62. van Duijn, E.; Barendregt, A.; Synowsky, S.; Versluis, C.; Heck, A. J. Chaperonin Complexes Monitored by Ion Mobility Mass Spectrometry. *J. Am. Chem. Soc.* **2009**, *131*, 1452–1459.
 63. Remaut, H.; Rose, R. J.; Hannan, T. J.; Hultgren, S. J.; Radford, S. E.; Ashcroft, A. E.; Waksman, G. Donor-Strand Exchange in Chaperone-Assisted Pilus Assembly Proceeds Through a Concerted β Strand Displacement Mechanism. *Mol. Cell.* **2006**, *22*, 831–842.
 64. Sharon, M.; Witt, S.; Glasmacher, E.; Baumeister, W.; Robinson, C. V. Mass Spectrometry Reveals the Missing Links in the Assembly Pathway of the Bacterial 20 S Proteasome. *J. Biol. Chem.* **2007**, *282*, 18448–18457.
 65. Levy, E. D.; Boeri Erba, E.; Robinson, C. V.; Teichmann, S. A. Assembly Reflects Evolution of Protein Complexes. *Nature* **2008**, *453*, 1262–1265.
 66. Painter, A. J.; Jaya, N.; Basha, E.; Vierling, E.; Robinson, C. V.; Benesch, J. L. Real-Time Monitoring of Protein Complexes Reveals Their Quaternary Organization and Dynamics. *Chem. Biol.* **2008**, *15*, 246–253.
 67. Sun, Y.; MacRae, T. H. The Small Heat Shock Proteins and Their Role in Human Disease. *FEBS J.* **2005**, *272*, 2613–2627.
 68. Bova, M. P.; Ding, L. L.; Horwitz, J.; Fung, B. K. Subunit Exchange of α A-Crystallin. *J. Biol. Chem.* **1997**, *272*, 29511–29517.
 69. van den Oetelaar, P. J.; van Someren, P. F.; Thomson, J. A.; Siezen, R. J.; Hoenders, H. J. A Dynamic Quaternary Structure of Bovine α -Crystallin as Indicated from Intermolecular Exchange of Subunits. *Biochemistry* **1990**, *29*, 3488–3493.
 70. Schneider, F.; Hammarstrom, P.; Kelly, J. W. Transthyretin Slowly Exchanges Subunits Under Physiologic Conditions: A Convenient Chromatographic Method to Study Subunit Exchange in Oligomeric Proteins. *Protein Sci.* **2001**, *10*, 1606–1613.
 71. Keetch, C. A.; Bromley, E. H.; McCammon, M. G.; Wang, N.; Christodoulou, J.; Robinson, C. V. L. 55P Transthyretin Accelerates Subunit Exchange and Leads to Rapid Formation of Hybrid Tetramers. *J. Biol. Chem.* **2005**, *280*, 41667–41674.
 72. Back, J. W.; de Jong, L.; Muijsers, A. O.; de Koster, C. G. Chemical Cross-Linking and Mass Spectrometry for Protein Structural Modeling. *J. Mol. Biol.* **2003**, *331*, 303–313.
 73. Sinz, A. Chemical Cross-Linking and Mass Spectrometry to Map Three-Dimensional Protein Structures and Protein-Protein Interactions. *Mass Spectrom. Rev.* **2006**, *25*, 663–682.
 74. Kolstoe, S. E.; Ridha, B. H.; Bellotti, V.; Wang, N.; Robinson, C. V.; Crutch, S. J.; Keir, G.; Kukastenvehas, R.; Gallimore, J. R.; Hutchinson, W. L.; Hawkins, P. N.; Wood, S. P.; Rossor, M. N.; Pepys, M. B. Molecular Dissection of Alzheimer's Disease Neuropathology by Depletion of Serum Amyloid P Component. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 7619–7623.
 75. Jensen, O. N.; Barofsky, D. F.; Young, M. C.; von Hippel, P. H.; Swenson, S.; Seifried, S. E. Direct Observation of UV-Crosslinked Protein-Nucleic

- Acid Complexes by Matrix-Assisted Laser Desorption Ionization Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **1993**, *7*, 496–501.
76. Nazabal, A.; Wenzel, R. J.; Zenobi, R. Immunoassays with Direct Mass Spectrometric Detection. *Anal. Chem.* **2006**, *78*, 3562–3570.
77. Farmer, T. B.; Caprioli, R. M. Determination of Protein–Protein Interactions by Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry. *J. Mass Spectrom.* **1998**, *33*, 697–704.
78. Back, J. W.; Hartog, A. F.; Dekker, H. L.; Muijsers, A. O.; de Koning, L. J.; de Jong, L. A New Crosslinker for Mass Spectrometric Analysis of the Quaternary Structure of Protein Complexes. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 222–227.
79. Sinz, A. Chemical Cross-Linking and Mass Spectrometry for Mapping Three-Dimensional Structures of Proteins and Protein Complexes. *J. Mass Spectrom.* **2003**, *38*, 1225–1237.
80. Young, M. M.; Tang, N.; Hempel, J. C.; Oshiro, C. M.; Taylor, E. W.; Kuntz, I. D.; Gibson, B. W.; Dollinger, G. High Throughput Protein Fold Identification by Using Experimental Constraints Derived from Intramolecular Cross-Links and Mass Spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5802–5806.
81. Krauth, F.; Ihling, C. H.; Ruttinger, H. H.; Sinz, A. Heterobifunctional Isotope-Labeled Amine-Reactive Photo-Cross-Linker for Structural Investigation of Proteins by Matrix-Assisted Laser Desorption/Ionization Tandem Time-of-Flight and Electrospray Ionization LTQ-Orbitrap Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2009**, *23*, 2811–2818.
82. Baca, M.; Kent, S. B. H. Direct Observation of a Ternary Complex Between the Dimeric Enzyme HIV-1 Protease and a Substrate-Based Inhibitor. *J. Am. Chem. Soc.* **1992**, *114*, 3992.
83. Barrera, N. P.; Di Bartolo, N.; Booth, P. J.; Robinson, C. V. Micelles Protect Membrane Complexes from Solution to Vacuum. *Science* **2008**, *321*, 243–246.
84. Barrera, N. P.; Isaacson, S. C.; Zhou, M.; Bavro, V. N.; Welch, A.; Schaedler, T. A.; Seeger, M. A.; Miguel, R. N.; Korkhov, V. M.; van Veen, H. W.; Venter, H.; Walmsley, A. R.; Tate, C. G.; Robinson, C. V. Mass Spectrometry of Membrane Transporters Reveals Subunit Stoichiometry and Interactions. *Nat. Methods* **2009**, *6*, 585–587.
85. Lin, H. T.; Bavro, V. N.; Barrera, N. P.; Frankish, H. M.; Velamakanni, S.; van Veen, H. W.; Robinson, C. V.; Borges-Walmsley, M. I.; Walmsley, A. R. MacB ABC Transporter is a Dimer Whose ATPase Activity and Macrolide-Binding Capacity are Regulated by the Membrane Fusion Protein MacA. *J. Biol. Chem.* **2009**, *284*, 1145–1154.
86. Morgner, N.; Kleinschroth, T.; Barth, H. D.; Ludwig, B.; Brutschy, B. A Novel Approach to Analyze Membrane Proteins by Laser Mass Spectrometry: From Protein Subunits to the Integral Complex. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 1429–1438.
87. Hanna, J.; Finley, D. A Proteasome for All Occasions. *FEBS Lett.* **2007**, *581*, 2854–2861.
88. Hartwell, L. H.; Hopfield, J. J.; Leibler, S.; Murray, A. W. From Molecular to Modular Cell Biology. *Nature* **1999**, *402*, C47–52.
89. Alberts, B. The Cell as a Collection of Protein Machines: Preparing the Next Generation of Molecular Biologists. *Cell* **1998**, *92*, 291–294.
90. Glickman, M. H.; Raveh, D. Proteasome Plasticity. *FEBS Lett.* **2005**, *579*, 3214–3223.
91. Wei, N.; Serino, G.; Deng, X. W. The COP9 Signalosome: More Than a Protease. *Trends Biochem. Sci.* **2008**, *33*, 592–600.
92. Gonzalez, C. M.; Griffey, S. M.; Naydan, D. K.; Flores, E.; Cepeda, R.; Cattaneo, G.; Madewell, B. R. Canine Transmissible Venereal Tumor: A Morphological and Immunohistochemical Study of 11 Tumors in Growth Phase and During Regression After Chemotherapy. *J. Comp. Pathol.* **2000**, *122*, 241–248.
93. Kitova, E. N.; Kitov, P. I.; Bundle, D. R.; Klassen, J. S. The Observation of Multivalent Complexes of Shiga-Like Toxin with Globotriaoside and the Determination of Their Stoichiometry by Nanoelectrospray Fourier-Transform Ion Cyclotron Resonance Mass Spectrometry. *Glycobiology* **2001**, *11*, 605–611.
94. Robinson, C. V.; Chung, E. W.; Kragelund, B. B.; Knudsen, J.; Aplin, R. T.; Poulsen, F. M.; Dobson, C. M. Probing the Nature of Noncovalent Interactions by Mass Spectrometry. A Study of Protein-CoA Ligand Binding and Assembly. *J. Am. Chem. Soc.* **1996**, *118*, 8646–8653.
95. Rupp, B. High-Throughput Crystallography at an Affordable Cost: The TB Structural Genomics Consortium Crystallization Facility. *Acc. Chem. Res.* **2003**, *36*, 173–181.
96. Service, R. F. Structural Biology. Robots Enter the Race to Analyze Proteins. *Science* **2001**, *292*, 187–188.