

PROTEIN COMPLEXES IN THE GAS PHASE: TECHNOLOGY FOR STRUCTURAL GENOMICS AND PROTEOMICS

Justin L.P. Benesch [†], Brandon T. Ruotolo [†], Douglas A. Simmons ^{†‡}, Carol V. Robinson ^{}*

Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, U.K

[‡] Present Address: MDS SCIEX, 71 Four Valley Drive, Concord, Ontario, L4K 4V8, Canada

[†] These authors contributed equally to this work,

^{*} Corresponding Author: Tel: +44 1223 763846, Email: cvr24@cam.ac.uk

Running Title: Mass Spectrometry for Structural Proteomics

CONTENTS

Contents	2
1 Introduction	3
2 Generating Ions of Protein Complexes	5
2.1 Electrospray Ionization	6
2.1.1 Mechanism of Protein Ion Formation	7
2.1.2 NESI Ionization	9
2.1.3 Limitations of Electrospray Ionization	11
3 Transmitting and Analyzing Ions of Protein Complexes	13
3.1 Transmitting Ions	13
3.1.1 Acceleration of Gas-phase Ions in the Source Region	14
3.1.2 Transmission Efficiency of Massive Ions in RF-Ion Guides	14
3.2 Analyzing Ions	17
3.2.1 Quadrupole Analyzers	17
3.2.2 Time of Flight Analyzers	18
3.2.3 Tandem/Hybrid Instruments	19
3.3 Instrument Modifications for Non-covalent Complexes	19
4 Activation and Dissociation of Protein Complexes	21
4.1 Collision-Induced Dissociation	21
4.1.1 Activation of Macromolecular Assemblies	22
4.1.2 Dissociation of Activated Macromolecular Assemblies	25
4.1.3 The CID of a Large Oligomeric Protein: Case Study	27
4.2 Other Activation Techniques	28
4.3 Use of Gas-Phase Dissociation	30
5 Companion Technologies for MS of Protein Complexes	31
5.1 Sample Isolation Technologies	31
5.2 Sample Labeling Technologies	33
5.3 Pre-Ionization Technologies	36
5.4 Post-Ionization Technologies	38
6 The Future	42
7 Acknowledgements	44
8 References	45

1 INTRODUCTION

How does the development of mass spectrometry (MS) for the study of protein complexes, intact in the gas phase, contribute to the twin fields of genomics and proteomics? Foremost it is important to recognize the fact that the vast majority of proteins do not exist as single entities in the cell, but rather interact non-covalently with additional copies of the same protein and/or other proteins ¹. Furthermore, additional interactions can occur with nucleic acids, ligands, cofactors or metal ions such that the functional form of many proteins is rarely the simple monomeric state. As such, while traditionally genomics and proteomics has focussed on determining which proteins are encoded by the genetic material, in order to understand their function their interaction to form protein complexes must also be investigated.

Studying this highest state of protein organization therefore represents a crucial part of modern genomics and proteomics efforts. These protein complexes themselves are often composed of ‘sub-complexes’, that is multi-protein modules which have some independent stability. Other articles in this issue describe MS approaches towards studying the individual protein components which make up these non-covalently bound species. These are typically performed by the cleavage of the protein backbone into smaller constituent peptide segments. In this review, however, we concentrate on the application of MS to the upper levels of protein organisation, those in which non-covalent interactions are preserved (Fig 1).

It is now well over a decade since intact protein complexes were first successfully analysed by means of MS ²⁻⁵. These early reports stimulated much excitement, and were swiftly followed by a plethora of studies wherein protein and other non-covalent complexes, largely from commercial or recombinant sources, were maintained intact in the gas phase ⁶. It soon became apparent that the precision with which the masses of these species can be measured is unrivalled by any other approach. Furthermore, when coupled with the selective disruption of the intact protein complexes, either in solution or in the gas phase, to determine the masses of the constituents, stoichiometry can be established, even in the absence of polypeptide sequence information. Such approaches have allowed the determination of the relative populations of the oligomers within polydisperse assemblies ⁷, have revealed substrate binding to molecular machines such as the proteasome ⁸ and GroEL ⁹, and defined the stoichiometry of subcomplexes within intact ribosomes ¹⁰.

Coupled to the realization that many aspects of solution phase structure are maintained in the gas phase ^{11,12}, these studies have established MS of intact complexes as a complementary approach to many structural techniques.

Further MS developments, such as the inclusion of ion mobility spectrometry (IMS) in the analysis of protein complexes ¹³, or controlled gas-¹⁴ or solution-phase ¹⁵ dissociation, has allowed the subunit architecture, overall shape and subunit contacts, of protein complexes to be realised. Moreover, the speed of analysis has enabled the study of the dynamics of these protein complexes ¹⁶. As such a number of processes including assembly ^{17,18}, disassembly ¹⁹⁻²¹, and exchange of subunits between protein complexes ²²⁻²⁴ have been monitored by means of MS. Consequently MS can now be used to examine all the levels of protein organization (Fig 1), providing information as to the sequences of the proteins involved, the stoichiometry of their interactions, their connections, and the dynamics of the assembled complex. Applying MS to the study of protein complexes, as well as more conventional applications to individual proteins, is therefore vitally important to the burgeoning field of proteomics.

Very recently combining the MS of protein complexes with affinity purification techniques ¹⁵, has highlighted the tantalising possibility of identifying the components, connectivity and shape of such species, expressed at endogenous levels. In this review we chart the progress that has enabled the MS of intact protein complexes to approach this vision. From early studies of recombinant protein complexes, through to the latest reports that couple MS of intact cellular complexes at endogenous expression levels, we describe the underpinning technological advances. Rather than consider biological applications that have been described in detail elsewhere ^{6,25-32} this review focuses on the enabling tools and technologies. We describe the processes of generating, transmitting and analysing ions formed by protein complexes and explain the underlying theory behind the experimental conditions that are becoming standard for these investigations. From the structure and dynamics of complexes in solution through to their overall topology in the gas phase we detail additional approaches that can be used in conjunction with MS of intact complexes to broaden the scope of the information available. Overall it is our aim to stimulate research in this area and to establish MS of intact complexes as an integral part of future structural genomic and proteomic investigations.

2 GENERATING IONS OF PROTEIN COMPLEXES

The first step of any mass spectrometry experiment is the generation of ions from a sample of interest. There are a number of challenges associated with ion generation, a fact which is mirrored by the number of methods available for generating gas-phase ions for analysis by mass spectrometry. For a number of these ionization methods it is necessary that the analyte already be present in the gas phase (*e.g.* electron impact ionization), or embedded in a non-physiological matrix (*e.g.* fast atom bombardment, matrix-assisted laser desorption/ionization (MALDI)) prior to ionization. The vast majority of biological processes occur in solution so it follows that for species of biochemical interest it is often necessary to generate gas-phase ions directly from molecules in solution. Furthermore, it is beneficial to do so in a manner that avoids fragmentation or other undesirable alteration of the analyte.

The above requirements effectively limit the scope of much work in biological mass spectrometry to either electrospray ionization (ESI)³³⁻³⁵ or MALDI^{36,37}. Of the two technologies, applications of ESI to the study of the structure and stability of intact protein complexes far out-weigh applications of MALDI. This is primarily because the sample-preparation requirements for MALDI typically involve evaporation of solvent from an aqueous analyte solution that contains a 10-10,000 fold excess of UV-absorbent organic acid. Such highly-acidic conditions will undoubtedly perturb protein-protein interactions present in solution, and most likely denature the protein. In addition, MALDI mass spectra often yield intense signals for protein aggregates that are thought to be artefacts of the laser desorption/ionization process. Several instrumental parameters such as laser fluence and irradiation time^{38,39}, along with the localization of complexes within the MALDI deposition⁴⁰, have also been shown to have a profound influence on the observance of non-covalently bound protein complexes. Although some examples of using specialized matrices that allow for MALDI sample preparations near neutral pH. (*e.g.*, 6-aza-2-thiothymine) have been reported⁴¹⁻⁴⁴, it is currently unlikely that UV-MALDI will be used as a general method for the desorption and ionization of intact protein-protein complexes. In addition, although significant effort has been made to elucidate the mechanism(s) of ion formation in MALDI, it remains a hotly-debated subject. MALDI is commonly used to identify precursor proteins from the peptides produced by enzymatic digestion of complex mixtures derived from biological source, as described in other articles in this issue. A detailed knowledge of the

desorption/ionization mechanism is not necessary in the interpretation of these data (MS is used as a detector for experiments carried out in solution only), therefore we will not cover it in depth here. Instead, we direct the reader to several outstanding reviews on the subject⁴⁵⁻⁴⁹.

2.1 ELECTROSPRAY IONIZATION

The most commonly used method for biological MS is ESI³³⁻³⁵. A typical electrospray setup involves passing a sample through a metallic capillary held at high electrical potential, surrounded by a concentric tube through which a parallel flow of inert gas is passed to aid the nebulization of the emerging analyte solution. At the tip of the steel capillary, the applied potential causes charges to gather preferentially at the tip, forming a “Taylor cone”⁵⁰⁻⁵². At the tip of the cone, the stream of solvent is drawn out into highly charged droplets, generally on the order of several micrometers in diameter⁵⁰. Emerging droplets are subsequently drawn down a potential and pressure gradient towards the ion sampling interface of the mass spectrometer. Aided by both parallel and then counter-directional flows of nebulizing gas, solvent evaporation from the nascent droplets results in a reduction in droplet diameter. This reduction in droplet size continues until the Coulombic repulsion between the increasingly crowded charges becomes strong enough to overcome the surface tension holding the droplet together. At this point, termed the “Rayleigh limit”, droplet fission occurs. The limiting charge on a droplet, q_R , is governed by the droplet diameter, D , and the solvent surface tension, γ , and is given by:

$$q_R = z_R e = \pi \sqrt{(8 \epsilon_0 \gamma D^3)}$$

Eqn 1

where z_R is the number of charges, e the elementary charge, and ϵ_0 the electrical permittivity of vacuum⁵³.

Equation 1 provides an upper limit of the charge density required to bring about droplet fission. In practice, a droplet undergoing evaporation and Coulombic explosion will do so at charge densities of approximately 70 – 100% q_R ⁵². Close observation of this process has revealed that rather than undergoing symmetric division into two droplets of similar dimensions, a fine jet of offspring droplets is generated from the parent. Each sequential fission event results in an offspring droplet of approximately 2% of the mass and 15% of charge of the parent droplet^{52,54,55}. Subsequent depletion of solvent by further evaporation results in several

generations of unequal droplet fission until ultimately; gas-phase ions are produced. During this process ions are drawn into the ion sampling interface of the mass spectrometer. Additional desolvation, often required to remove residual solvent and solute particles from the protein ions, is brought about by collisions with gas molecules in the low-vacuum region between the orifice and skimmer electrodes. The energy of these collisions is dictated by the difference in the potentials applied to these two electrodes.

2.1.1 Mechanism of Protein Ion Formation

ESI usually results in the generation of intact, multiply charged ions generally in the form $[M + nH]^{n+}$. While it is possible to generate both positive (*i.e.* protonated) and negative (*i.e.* de-protonated) ions during electrospray, the majority of protein mass spectrometry is performed in the positive mode, and only this will be addressed explicitly in our following discussion of ESI. While the macroscopic aspects of electrospray are generally well-understood, the mechanics of the final generation of desolvated (or nearly desolvated) ions from a charged droplet remain incompletely resolved. Two principal models have been proposed to account for this phenomenon. The charged residue model (CRM), conceived by Dole *et al.*,³³ postulates that evaporation and Coulombic fission occurs until a droplet containing a single residual analyte ion remains. Complete evaporation of the solvent comprising this droplet eventually yields a ‘naked’ analyte ion, the charged residue. A second mechanism for gas-phase ion production, based on the work of Iribarne and Thomson⁵⁶, is termed the ion evaporation model (IEM). In this model, it is argued that, prior to complete desolvation of the droplet, the repulsion between the charged protein ion and the other charges in the droplet becomes strong enough to overcome solvation forces, and the ion is ejected from the droplet surface and into the gas phase.

The consensus in the literature is that neither the CRM nor IEM can account for all experimental observations⁵⁷. Rather, it appears that one or the other, or both in combination, can be invoked to describe ion formation depending on the type of analyte. For the case of proteins of mass > 6000 Da, and thus those species most relevant to this review, there is considerable evidence that Dole’s CRM is the dominant mechanism of ion formation during ESI^{52,55,58,59}. By examining the considerable data that had been generated for native proteins during earlier work, de la Mora found that the maximum charge acquired by globular proteins during ESI matches closely to the charge expected on spherical solvent droplets with similar radii⁵².

This is the behavior one expects from a protein ion formed through the total evaporation of an encompassing droplet which would presumably be only slightly larger than the protein itself. Furthermore, this relationship has been shown to hold up even when the number of strongly basic amino acids available for protonation exceeds the number of charges expected in the droplet, given by Equation 1^{60,61}. Additionally, ions formed by the CRM would be expected to acquire a number of adducts due to the presence of solutes (*e.g.* buffers) in the final precursor droplet. This phenomenon is not only a common empirical observation of any practitioner of ESI-MS of native proteins, but has also been shown to be quantitatively predictable based on CRM calculations⁶⁰.

The immediate implication of the dominance of the CRM mechanism in protein ion formation is that the charge acquired by a protein during ESI is dictated by the size of the precursor droplet surrounding the protein immediately before desolvation, and therefore on the size of the protein itself. This is consistent with the wealth of observations in the literature concerning the relationship between protein conformation and ESI charge state. In the early 1990s, Chait and co-workers noticed that large changes in solvent conditions were accompanied by significant changes in protein ESI mass spectra⁶². It is now well-documented that an unfolded protein in solution results in ions having higher charge states compared with the same experiment performed using solution conditions that promote folded conformations. Moreover, the width of the charge state distributions is also larger for unfolded proteins, and is thought to be related to the structural heterogeneity of the corresponding protein states in solution^{61,63-67}. Based on these empirical relationships, but with the caveat that other processes may be contributing to charge state variation⁶⁸⁻⁷⁰, ESI-MS has become a standard qualitative method for probing protein conformational changes in solution^{62-65,71-76}.

More recently, in light of the strong evidence for the CRM mechanism, protein charge states have been used as quantitative indicators of protein structure. In particular, it has been shown that the charge states acquired by native proteins can be used as a means to estimate the surface area of proteins⁶¹. From the results of two recent studies, however, the precise nature of the relationship appeared to be in some dispute³¹. Kaltashov and Mohimen compiled a data set of the average ESI charge states, z_{avg} , of native (or near-native) proteins and compared those to protein surface areas, S , as determined from crystallographic data. The resulting curve showed a power relationship well-described by $z_{\text{avg}} \propto S^{0.69}$ ⁶¹. This power function is in contrast to the linear

relationship between z_{avg} and S observed by Hautreux *et al.* in a previous study⁷⁷. This disagreement may be due to the different methods by which surface areas were calculated in the two studies: whereas Kaltashov and Mohimen derived surface areas from crystallographic data, Hautreux *et al.* extrapolated surface areas from protein mass, assuming constant density and a spherical shape³¹. While the precise details of the relationship between charge state and surface area have yet to be completely resolved, it is apparent that this aspect of ESI-MS may represent a powerful means to probe the 3D structure of proteins in solution.

2.1.2 NESI Ionization

A conventional electrospray apparatus uses a spray capillary on the order of 0.5 mm in diameter, and requires flow rates of several $\mu\text{L}/\text{min}$ to maintain the stable Taylor cone necessary for droplet formation. Consequently, it is usually necessary to have a minimum of $\sim 50 \mu\text{L}$ of sample for most ESI-MS analyses. In 1994, however, Wilm and Mann introduced an important variant of conventional ESI, termed nanoelectrospray (nESI), a name chosen to reflect the low flow rates involved⁷⁸. While this technique uses the same fundamental sequence of charged-droplet generation followed by multiple asymmetric Coulombic fission events, and finally ion formation, it is distinguished from regular ESI in a number of important ways. nESI is typically performed using glass or quartz capillaries which have been pulled to a fine tip ($\sim 1 \mu\text{m}$ inner diameter) and given a metallic (usually gold) coating to hold the electric potential in place of the metallic capillary used for conventional ESI. Approximately $1 - 3 \mu\text{L}$ of sample is injected into the glass capillary and electrosprayed at flow rates in the range of $\sim 1 \text{ nL}/\text{min}$ to several tens of nL/min ^{79,80}. Flow is driven primarily by the approximately $0.5 - 1.5 \text{ kV}$ potential applied to the capillary, although it is often necessary to provide an auxiliary backing gas pressure to the sample in order to initiate and/or maintain a steady stream of the solution through the tip.

nESI addresses a number of key issues which limit the suitability of ESI for the analysis of biological samples, and more specifically non-covalent protein assemblies. First and perhaps most obviously is the reduction in the quantity of sample consumed during analysis. Samples of protein complexes are often generated via painstaking molecular biology or isolation techniques, and are consequently often limited in sample volume. Using nESI, it is possible to produce mass spectra using only a few picomoles of native protein complex. Even attomolar quantities are detectable when the analysis is performed under ideal

electrospray conditions with organic co-solvents (*i.e.* when preserving non-covalent interactions is unimportant)⁸¹. This particular characteristic of nESI alone has made mass spectrometric analysis possible for a number of biochemical species which are difficult to produce in quantities required for conventional ESI.

Secondly, it has been shown in a number of studies that nESI is more tolerant of salt contamination of the protein solution than conventional ESI^{79,82}. This is particularly important for the study of non-covalent protein complexes since these compounds often require the presence of certain buffer salts to remain stable in solution. An increased tolerance of the presence of non-volatile salts can be explained through examination of the subtle differences in the ESI process between nanospray and conventional ESI. With both techniques, the first generation of charged droplets undergo evaporation of solvent to reach the Rayleigh limit necessary to bring about droplet fission (Equation 1). This results in an increase in the concentrations of the analyte as well as any non-volatile salts which do not undergo field evaporation. It follows that the higher the number of fission events which are necessary to form a charged residue, the greater the salt concentration in the final droplet containing a single protein complex molecule. The extent of salt enrichment during the multiple steps of evaporation/asymmetric fission is therefore affected greatly by the size of the primary droplets formed from the Taylor cone at the capillary tip. As mentioned above, the initial droplets formed during conventional ESI are of the order of several micrometers in diameter. In contrast, the diameters of primary droplets formed during nESI have been measured to be on the order of 150-200 nm^{79,82,83}. Since offspring droplets are generally one order of magnitude smaller than primary droplets during ESI, the difference in droplet size achieved by swapping from ESI to nESI removes approximately one round of Rayleigh fission, and therefore one salt concentration step, prior to releasing the desolvated gas-phase ions. This has been verified experimentally by comparing the size of salt $[(\text{NaCl})_n\text{Na}^+]$ clusters which are formed using both techniques; the greater salt enrichment experienced by droplets during ESI was found to result in larger, more highly-charged salt clusters than those formed during nESI⁸².

Thirdly, the nESI desolvation process has been shown to be a gentler, more reliable, method of introducing extremely labile protein assemblies into the gas-phase intact. It is generally accepted that the small initial droplet sizes produced by the nESI source reduces the number and the energy of the collisions required to

desolvate the macromolecules of interest. Figure 2 shows two mass spectra of the 800kDa *E. coli* chaperone GroEL acquired using the same MS instrument. The upper spectrum was acquired using nESI for ion generation. A number of well-resolved charge states are observed yielding sufficient information to measure the intact mass of the GroEL assembly. The lower spectrum was acquired using conventional ESI, and, even after re-optimizing the instrument settings to desolvate the larger droplets, a very different spectrum is obtained. Several peaks are observed which vary in shape and width. Using the information presented in the lower panel of Figure 2, it is extremely difficult, if not impossible, to obtain a reliable mass measurement of GroEL. Furthermore, rather than just the one peak series corresponding to the 800kDa complex, additional signal is observed, likely due to dissociation, and artefactual formation of more massive species (Figure 2). These differences highlight how nESI possesses several significant advantages over conventional ESI as regards the study of protein complexes by means of MS.

2.1.3 LIMITATIONS OF ELECTROSPRAY IONIZATION

Although electrospray has been under continuous and intense development since its introduction as an ionization source for MS over 15 years ago, and despite the fact that it is the preferred ionization method for MS experiments in structural biology, several limitations persist. For example, overall ionization efficiencies for proteins from solution is often relatively poor. While ionization efficiency is higher in nESI than in conventional ESI, less than 10% of sample molecules are detected even under favorable conditions⁸⁴. Also, the nESI response has been shown to be highly dependant upon the geometry of the nanoflow needle, and consequently some needle-to-needle irreproducibility is to be expected⁸⁴. Moreover, two related phenomena that are a consequence of the ion formation mechanisms in ESI affect the mass analysis of protein complexes. The first, artefactual oligomer formation in ESI droplets, is alluded to above and in figure 2. Figure 3 demonstrates how under conditions of high protein concentration ($\geq 50\mu\text{M}$), though the majority of the droplets remain empty, there is still a significant probability of trapping more than one protein complex within the small progeny droplets generated from the larger initial droplets emitted from the nESI tip. As such, following from the CRM it is possible for protein complexes to associate into higher non-specific oligomers that are not present in solution at low concentration ($<10\mu\text{M}$); see Figure 3. A second consequence of the nESI mechanism is that complete desolvation of macromolecular protein complex ions is

often not possible, because increasing the activation of the ions leads to disruption of non-covalent interactions before full desolvation is reached ⁸⁵. Approximately 1% of the total ion mass of a protein assembly can, in some cases, be attributed to buffer and solvent adducts ⁸⁶. The results of this incomplete desolvation are broad peaks and mass measurements that are several percent larger than those predicted by the amino acid sequence alone. Therefore, despite the mass resolution and resolving power capabilities of modern mass analyzers, the peak width achieved for large protein complexes is determined primarily by the electrospray process and subsequent ion desolvation. It has been shown, however, that the width of the mass spectrum peaks is correlated with the excess mass observed due to adducts for large ions of known mass (Figure 4) ⁸⁶. This basic approach was used to measure accurately the mass of the intact ribosome and its subcomplexes, making it possible to detect the presence of multiple components within peaks observed in the mass spectrum ⁸⁶.

3 TRANSMITTING AND ANALYZING IONS OF PROTEIN COMPLEXES

The generation of ions is only the first hurdle to be overcome in the mass spectrometry of large non-covalent protein complexes. In order for successful application of this technique, the ions must also be retained intact in the various differentially-pumped regions of the mass spectrometer, focused and guided along their intended flight path, separated according to their mass to charge ratio, and subsequently detected. All of these aspects provide their own complications when large macromolecular assemblies are studied. Various instrumental and methodological advances have been made in the last fifteen years such that non-covalent complexes may be transmitted, mass analyzed, and detected. In this section we detail the most important of these developments, and the different technological options available in order to perform successful mass spectrometry of non-covalent complexes.

3.1 TRANSMITTING IONS

The first methodological breakthrough made for the improved transmission of large non-covalent complexes occurred when it was found that increasing the pressures in the ion guides in the first vacuum stages of the mass spectrometer resulted in increased signal for high m/z ions⁸⁷⁻⁸⁹. Prior to these reports, Douglas and French had shown a clear correlation between ion mass and the pressure required for optimal ion transmission⁹⁰. This effect has often been referred to as ‘collisional cooling’, though this term is misleading as all collisions experienced by an ion result in a momentary increase in internal energy.

Early investigations of the impact of source pressure on the observation of non-covalent protein assemblies were performed in 2001^{88,89}. In one study, under standard pressure conditions (2 mbar) only trace signals were assigned to dimeric protein. However, upon increasing the pressure to 3.5 mbar a significant improvement in the intensity of peaks assigned to the protein dimer were observed. Further increases in pressure, up to 7.0 mbar, resulted in the emergence of a charge state distribution corresponding to an octameric protein assembly which had not been observed at lower pressures. This phenomenon was generally rationalized as being due to a reduction in dissociation by the somewhat vague concept of ‘gentle’ MS conditions^{2,4}. These conditions enabled efficient desolvation of macromolecular ions while maintaining

their non-covalent interactions. The general opinion at this time was that the higher pressure resulted in frequent low-energy collisions, and that this resulted in less dissociation than fewer higher energy collisions.

In recent years, however, significant progress has been made in understanding this phenomenon. These studies have focused not on how higher pressures and hence more collisions might increase the transmission of non-covalent complexes by reducing the amount of dissociation, but rather on how they might act to improve the focusing of these ions such that more pass through the mass spectrometer^{91,92}. We describe below the various steps along this collisional focusing pathway.

3.1.1 Acceleration of Gas-phase Ions in the Source Region

As described above ESI, the preferred ionization method for the study of non-covalent complexes by MS, is performed at atmospheric pressure. Carried along in the flow of bath gas, ions enter the first vacuum stage through a small aperture. The consequent free expansion of the gas stream during this traversal from atmospheric pressure to rough vacuum, as well as space-charge effects, results in most of the ions deviating from the ideal trajectory. Furthermore, the expansion of the gas stream results in ion velocities of several hundred meters per second. Even very massive ions, such as those of protein complexes, exhibit little variation from the velocity of the gas stream. The corresponding mass-dependent kinetic energies of such ions range from ~ 1 eV for ions of 1 kDa, up to ~ 1 keV for ions of 1 MDa^{91,92}. As such, the ions formed from protein complexes at this stage are not only defocused, but are traveling along these inappropriate trajectories with high kinetic energy. Typical ion guide parameters which are used to focus ions back onto the intended route are generally ineffective, as protein complexes carry fewer charges than denatured proteins of similar mass. Hence their manipulation is highly problematic in the optics of a conventional mass spectrometer which has electrostatic lens voltages operating generally at less than 100 V. Therefore ion trajectories are poorly focused with respect to the center-line of the instrument, resulting in poor transmission efficiency at the apertures that separate the initial stages of differential pumping. Various strategies for focusing massive ions have been proposed, such as aerodynamic focussing⁹³, but currently the most widely used method is collisional focusing in a radio-frequency (RF)-only ion guide.

3.1.2 Transmission Efficiency of Massive Ions in RF-Ion Guides

For some time it has been recognized that collisions with neutral gas molecules can be used to focus the trajectories of ions stored in three-dimensional Paul traps. Collisions between comparatively large ions and small neutral gas molecules serve to dissipate ion kinetic energy into the surrounding gas. This is functionally important because it allows more efficient trapping of ions in the potential well at the centre of the ion-trap. This effect has often been termed ‘collisional cooling’, which is in some ways misleading as during a collision with a neutral though the ion’s kinetic energy is decreased, its internal energy, or temperature, is increased. Therefore the ion is not cooled, but rather is heated. As such we will use the expression ‘collisional focusing’ throughout this article.

Long after this collision focussing became standard operating procedure for quadrupole ion-trap mass spectrometers, Douglas and French extended this methodology to linear multipole devices, specifically the RF quadrupole ion guide⁹⁰. In this study, it was found that the transmission efficiency of ions produced by ESI is pressure-dependent and increased pressure in the RF ion guide can result in increased transmission efficiency. While this effect was at first counter-intuitive, when stopping curves (plots of barrier voltage vs. ion transmission efficiency) were generated for ions at different pressures it was discovered that the applied potential necessary to stop ions at low pressure was much higher than that required to stop ions under high-pressure conditions. This result indicated that at higher pressure the ions possessed diminished kinetic energies, and therefore velocities.

As the axial component of the velocity (along the ion beam) is reduced by the increased collisions with bath gas at higher pressures, the same must also be true for the radial component. It was hypothesized that the increased ion transmission was therefore due to collisional damping of radial trajectories, allowing ions to be more efficiently captured in the dynamic potential well along the centre axis of the RF quadrupole⁹⁰. The focusing effect in turn allowed ions to be more efficiently transmitted through the exit aperture and into the subsequent ion optics. The trajectories of ions undergoing collisional focusing in an RF multipole were first calculated *in silico* by Krutchinsky *et al.*⁹¹, however, such an approach has only recently been applied to examine the transmission of large protein complexes⁹². In this study it was shown that while normal operating pressures (~10 μ bar) in the RF ion guide resulted in sufficient collisional focusing such that small protein ions were successfully transmitted through the aperture immediately after the RF quadrupole, the

ions of the 692 kDa 20S proteasome were poorly focused. Increasing the pressure to ~ 43 μ bar, however, resulted in sufficient collisions between the proteasome and the background gas to enable efficient focusing of the ion beam⁹². An alternative approach towards increasing the number of collisions experienced while operating at the same pressure is to increase the effective length of the ion guide. This can be achieved, without having to incorporate an ion guide which is physically longer, by trapping the ions under conditions where they undergo numerous collisions with neutrals⁹².

Figure 5 describes this collisional focusing principle, as it applies to a quadrupole-time-of-flight (Q-ToF) instrument (Figure 5A). The simulated trajectories (in a plane perpendicular to the quadrupole axis) of the ~ 147 kDa tetrameric protein ADH at three different pressures are shown (Figure 5B). As the ions enter the RF ion guide their oscillations result in large deviations from the centre of the ion axis (their ideal position for transmission through the instrument). As they experience collisions with bath gas, the velocity of the ions decreases and hence their oscillations deviate less from the central axis. At the higher simulated pressures there are sufficient collisions such that the focusing results in a very narrow ion beam. At the lower pressure (~ 9.3 μ bar, standard operating pressure) the collisional focusing is insufficient and results in a broad ion beam leading to losses of ions that fail to pass through the various small apertures downstream.

In addition to the improvement in transmission efficiency gained by collisional focusing of the ions, there are further consequences to the deceleration of protein complex ions in the initial stages of the mass spectrometer. Most of the experiments conducted on these molecules are performed with instruments incorporating an orthogonal time-of-flight mass analyzer (o-ToF). The trajectories of ions in an o-ToF analyzer are determined not only by the acceleration imparted in the orthogonal extraction source, but also by the axial velocity of the ions as they enter this region (Figure 5). Ions which enter the orthogonal extraction source with excess velocity may overshoot the detector, and are consequently not recorded⁹².

To demonstrate this effect we recorded the arrival position of tetrameric ADH ions on a four-part segmented-anode detector as a function of gas pressure in the primary RF ion-guide of a Q-ToF instrument (Figure 5C). At the lowest pressure (10 μ bar) overall signal intensity is low and the majority of the ion current is detected by anode segments furthest from the extraction source (anodes 3 and 4). As the pressure is increased to 20

μbar, total ion current increases and presents a more uniform distribution of impacts across all detection surfaces. At the highest pressures, ion current is primarily detected on anode segments nearest the extraction source (plates 1 and 2), and the total ion current had again decreased. The data in Figure 5C indicate that large macromolecular ions may not be sufficiently decelerated at low pressures. Consequentially, many ions retain excess velocity as they enter the ToF orthogonal extraction source, and overshoot the position of the detector during their passage through the ToF analyzer. The opposite problem, where ions have insufficient lateral velocity to reach the ToF detector, becomes significant at very high source pressures. Though other possibilities, such as scattering by background gas at higher pressure may also contribute⁹², these observations provide a plausible explanation as to why there appears to be an optimal pressure for ion detection, rather than an approach towards 100% transmission value with increasing pressure.^{88,89,94,95}

3.2 ANALYZING IONS

Although a wide range of mass analyzers has been developed and applied to the analysis of biomolecules, only a small subset of these has been successfully applied to the study of large protein complexes. Typically, large protein complex ions appear at high m/z , which limits the scope of applicable mass analyzers considerably. Specifically, some large multiply-charged protein complexes have been detected at > 20000 m/z ^{10,96-98}, whereas the ions typically formed from individual peptides and denatured proteins have considerably lower m/z values (< 3000). Some studies of protein complexes have been reported which employed ion-trapping instruments including 3D traps⁹⁹, Fourier-transform ion cyclotron resonance (FTICR)^{100,101}, and Orbitrap mass analyzers¹⁰². The majority of the initial experiments performed were with quadrupole analyzers, but more recently, ToF analyzers have been the preferred tool in the analysis of large protein complexes, often in the form of hybrid Q-ToF instruments. As such, we cover the elements of these instruments in detail below.

3.2.1 Quadrupole Analyzers

While quadrupoles can be used in RF-only mode as ion guides, *vide supra*, the application of a simultaneous DC voltage enables them to be used as mass analyzers. Their low cost, relatively high operating pressures,

and compatibility with the continuous ion beam generated by ESI made them the mass analyzer of choice in early MS experiments of protein complexes. The maximum m/z that can be resolved ($[m/z]_{max}$) is given by:

$$[m/z]_{max} = 7 \times 10^6 V_m / (v^2 r_0^2)$$

Eqn 2

where $V_m \cos(2\pi vt)$ is the RF voltage applied between adjacent rods (with $2V_m$ being the peak to peak amplitude, v the frequency, and t the time), and r_0 is the inner radius between the rods in meters. Lowering the operating frequency of the quadrupole therefore increases the $[m/z]_{max}$ in mass-resolving mode^{103,104}. Theoretically V_m could be increased, and/or r_0 decreased, as an alternative strategy, but the increased chance of high-voltage breakdown between the rods makes this impractical. The main disadvantage in operating quadrupole mass analyzers at such reduced frequencies, however, is a decrease in resolution. This implications of this drawback can, however, be largely negated by introducing a subsequent analysis step as is the case in the tandem instruments described below¹⁰⁵.

3.2.2 Time of Flight Analyzers

Time of flight (ToF) analysis was first coupled to an electrospray source in 1991¹⁰⁶, and relatively shortly thereafter the first spectra of non-covalent protein complexes were recorded on an ESI ToF mass spectrometer in 1994¹⁰⁷. ToF analyzers provided several immediate benefits to the study of macromolecular assemblies, the most significant being their theoretically unlimited m/z range. Due to the pulsed-nature of the ToF experiment, ESI-ToF experiments are usually conducted in an orthogonal mode, where the primary ion beam is deflected from its original direction of motion into the ToF analyzer. This arrangement typically exhibits poor duty cycle, especially when large masses (having long flight times) are observed. This limitation has been addressed by including a ion trapping step prior to orthogonal extraction of the primary ion beam^{108,109}. This step modulates the originally continuous ion beam into a source of ions pulsed at a frequency timed to match the extraction timing of the ToF analyzer. This trapping increases the overall sensitivity and limit-of-detection of the instrument, as a larger proportion of the total ion current is mass analyzed.

3.2.3 Tandem/Hybrid Instruments

One of the main strengths of MS is not only the ability to measure the mass of the intact species under investigation, but also to selectively dissociate the species in the gas phase and to mass-measure the product ions. Tandem MS, or MS/MS, has widespread applications in both ‘bottom-up’ and ‘top-down’ proteomics experiments, and also has tremendous potential in the study of protein complexes. Different types of tandem mass spectrometers have been used in the study of macromolecular complexes. Although some early experiments were conducted using triple quadrupoles instruments^{2,110,111}, the most prevalent tandem mass spectrometer in this field today is the Q-ToF instrument. The arrangement of such an instrument is shown in Figure 5A.

In Q-ToF instruments, quadrupole and ToF analyzers are arranged in series, separated by a collision cell. Ions of interest can be selected according to their mass to charge ratios in the quadrupole analyzer, then subsequently activated in the collision cell, and their products analyzed in the ToF analyzer. Activation is typically performed by ion-neutral collisions; however, ion-surface collisions¹¹² and laser-irradiation¹¹³ have also been successfully implemented on these instruments. In order to be able to select ions at high m/z ratios, quadrupoles operating at low frequency, *vida supra*, have been implemented in Q-ToF instruments by several groups^{92,105,114}. The reduction in resolution that results from operating at this lowered frequency generally does not impair the function of the mass spectrometer, as the product ion spectrum is acquired by the ToF analyzer¹⁰⁵. When a quadrupole is not in mass-resolving mode, i.e. it is not being used to select a particular ion, ions up to approximately five times the $[m/z]_{max}$ (Equation 2) can be transmitted¹⁰⁵. As such, lowering the operating RF frequency of the quadrupole not only allows the selection of high m/z ions, but also improves the transmission of ions such that their detection at over 85000 m/z has been reported¹⁰⁵.

3.3 INSTRUMENT MODIFICATIONS FOR NON-COVALENT COMPLEXES

The modification of various components and conditions within a mass spectrometer can contribute strongly towards successful mass analysis of non-covalent complexes. Higher pressures in the front end of the instruments are required to focus and decelerate high m/z ions, low frequency quadrupoles are often required for their selection and/or transmission, ToF analyzers need to be operated at low sampling frequencies, and

benefits can be gained from implementing detectors with high sensitivity at high m/z values. Higher ion guide pressures have been achieved in several ways: reducing the pumping efficiency by throttling the vacuum lines ¹⁰⁵, leaking additional gas into the source region ¹⁰⁵, or incorporating a flow-restricting sleeve between the ion guide and the pump orifice ⁹². The flow-restricting sleeve has the advantage of not increasing the load on the turbo pumps, though this benefit might be outweighed by the reduction in flexibility it affords in finding the pressure for optimal transmission of a particular species, *vide supra*.

The use of low-frequency/high m/z quadrupoles for purposes of tandem mass spectrometry ¹⁰⁵ has led to the ability to selectively dissociate very large complexes, and further modifications have since been implemented to aid in this dissociation process. Recently constructed Q-ToF instrumentation has allowed significantly higher accelerating potentials upstream of the collision cell than is possible in analogous commercial instruments ¹¹⁵. This access to higher voltages provides the ability to perform higher-energy dissociation experiments, and consequently obtain additional product ions. Early results show that both the loss of further protein subunits and covalent fragmentation of individual protein subunits can be accessed at these high collision energies (>200 V accelerating potential) ¹¹⁵.

In summary, since the first spectra of non-covalent protein complexes were obtained over fifteen years ago, numerous methodological and instrumental developments have taken place. It is now possible to control and interrogate ions that are significantly more massive and heterogeneous than ever before. As the desire to study samples of greater complexity continues to drive developments in tandem mass spectrometry, further progress will be required to effect greater control of protein complex dissociation while an increased mechanistic understanding is necessary to garner the maximum amount of information from macromolecular assemblies.

4 ACTIVATION AND DISSOCIATION OF PROTEIN COMPLEXES

While much effort over the last decade has been directed at maintaining non-covalent assemblies intact *in vacuo* for subsequent mass spectrometric detection, a lot of work is now focused on their gas-phase activation and dissociation³¹. Just as data derived from protein fragmentation has made MS a driving force in traditional proteomics¹¹⁶, the gas-phase dissociation of protein complexes is being used to deliver compositional information which may prove to be equally important.

Numerous different ion activation techniques have been developed over the years, only a relatively small subset of which has been successfully used to dissociate non-covalent protein assemblies. The activation techniques amenable to this research are necessarily limited to those that can be implemented in MS instrumentation which itself is capable of transmitting and analyzing non-covalent complexes. As such, the majority of work so far has been performed using collision-induced dissociation (CID – sometimes known as CAD, or collisionally-activated dissociation). However, black-body infrared radiative dissociation (BIRD)¹⁰⁰, and more recently electron-capture dissociation (ECD)¹¹⁷ and surface-induced dissociation (SID)¹¹² have also been used. While the dissociation of protein complexes in the gas phase is becoming ever more widespread, methodological challenges remain. In particular, further elucidations as to the mechanism of dissociation, understanding the relationship between gas-phase dissociation and solution-phase disruption, and establishing the limits of structural information that can be obtained, require further attention. Once these questions are answered, it is our belief that the use of protein complex dissociation in the gas phase will become an even more powerful tool for structural genomics and proteomics.

4.1 COLLISION-INDUCED DISSOCIATION

The majority of studies in which non-covalent complexes are disassembled in the gas phase have employed collision-induced dissociation (CID). This is performed by colliding the ions of interest with neutral gas. Activation occurs as a portion of an ion's kinetic energy is converted into internal energy during each collision event, and may lead to its subsequent dissociation if sufficient internal energy is accumulated. The use of CID in MS is well established, and several excellent reviews exist on this subject¹¹⁸⁻¹²⁰. The majority

of examples which use CID for dissociating non-covalent protein complexes employ Q-ToF mass spectrometers. It is this instrumental arrangement on which we base our discussion below.

4.1.1 Activation of Macromolecular Assemblies

During typical CID experiments ions will undergo many collisions, depending primarily on the size of the ion and the gas pressure. The number of collisions experienced by the ion during its transit through the collision cell, n , can be calculated from the following relationship, which is derived from the mean free path equation and includes a unit conversion factor:

$$n = 102430 (l\sigma p/T)$$

Eqn 2

where l is the collision cell length in m, σ the collision cross-section in \AA^2 , p is pressure in mBar, and T is temperature in K.

In cases where the collision cross-section is not known, estimates can be used. For example, the collision cross section can be estimated by approximating the shape of the molecule to a sphere, at which point the collision cross-section (\AA^2) is given by:

$$\sigma = \pi (R_i + R_g)^2$$

Eqn 3

Where R_i and R_g are the radii of the ion and gas molecule in \AA , respectively. When structural information sufficient to perform the above approximation is not available, one can assume a spherical shape and a typical density ρ (in $\text{Da}/\text{\AA}^3$) of proteins and their complexes. For collision cross-section estimates reported in this review, we use a value of $0.84 \text{ Da } \text{\AA}^{-3}$ for ρ , though different values for protein density can be found throughout the literature^{52,77,121}. From these assumptions, a spherical volume may be calculated from the mass M_i of the ion, and subsequently the radius. Therefore Equation 3 becomes:

$$\sigma = \pi (\sqrt[3]{3M_i/4\pi\rho} + R_g)^2$$

Eqn 4

In Figure 6 we give an overview of the number of collisions experienced, and length of time spent in the collision cell, for different proteinaceous species in typical experimental conditions for the CID of large non-

covalent complexes on a modified Q-ToF instrument^{92,105,114}. The top panel shows the number of collisions experienced by the different species (based on Equations 2 and 4) and the mass-dependency of the trend. It can be seen that large protein complexes typically experience from thousands to tens of thousands of collisions as they pass through the cell, and that the number of these collision can be modulated by adjusting the pressure (inset).

The bottom panel shows how long the ions spend in the collision cell, *i.e.* the amount of time during which they experience the collisions. This residence time depends on the initial kinetic energy of the ions, and hence on charge state. Therefore, to model a mass-dependency we used the relationship between ion mass and average charge state (Z_{av})⁵²: $Z_{av} \approx 0.0778\sqrt{m}$. By modeling all the collisions experienced by an ion, and the resulting reductions in velocity, one can see that the amount of time spent in the collision cell is dependent on mass, and is typically on the order of about 200-400 μ s (Figure 6). Only at the low end of the mass scale, does this residence time deviate considerably, reaching, in the case of cytochrome *c* over 1700 μ s (under the conditions used to generate Figure 6). This dramatic increase in residency time in the collision cell is due to the considerably higher transition efficiency of kinetic to internal energy for the relatively small cytochrome *c*. This allows the ion to reach a steady-state velocity (determined largely by the kinetic energy of the neutrals and space-charge effects) early in the transit of the ion through the collision cell. It is important to note that at no time are ions assumed to be held stationary in the collision cell under the conditions described in Figure 6. The residence time will also depend on both the accelerating voltage and the pressure in the cell. Inset into the lower panel is a surface plot of these dependencies. In all of the conditions displayed, the ions experience collisions over a time period on the order of hundreds of microseconds.

Each of the thousands of collisions a protein complex experiences in the collision cell causes its internal energy to increase. The maximal increase in internal energy accumulated by an ion, ΔE_{Int} , (assuming no energy dissipation in between collisions), after it has experienced n collisions can be given by the following relationship, where the initial translational energy is assumed to be determined only by the acceleration into the collision cell:

$$\Delta E_{Int} = zV_a (1 - [(M_i^2 + M_g^2)/(M_i + M_g)^2]^n)$$

where M_g and M_i are the mass of the gas and ion respectively, z is the charge state of the ion, and V_a the voltage accelerating the ions into the collision cell (adapted from ^{122,123}).

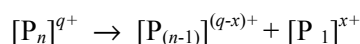
To demonstrate the transfer of translational energy into internal modes of ions as they traverse the collision cell we have performed a series of explicit individual-collision simulations based on Equation 5. The top panel of Figure 7 demonstrates the conversion of energy as a function of distance travelled along the collision cell for the same five species shown in figure 6, ranging from the 12 kDa cytochrome *c*, to the 2.3 MDa 70S ribosome from *Thermus thermophilus*. It can be seen how, for the relatively small and light cytochrome *c*, essentially all the translational energy is converted into internal modes by half-way along the collision cell, after less than 1000 collisions. At this distance into collision cell, however, the much larger 70S ribosome does not achieve such complete conversion, despite having experienced over 47000 collisions. This demonstrates that under these conditions, very large species will retain significant velocities as they exit the collision cell, which will affect their transmission efficiency to the MCP detector ⁹². To overcome this problem, from Equation 5 we can see there are several possibilities: either increasing the number of collisions n , which in turn can be effected by increasing the gas pressure or length of cell (Equation 5), or increasing the mass of the gas M_g . The middle and lower panels of Figure 7 show how the effective conversion efficiency between kinetic and internal energy is altered by changes in pressure or neutral mass respectively. Clearly, higher pressures and heavier target gases are beneficial for more rapid energy conversion. However, the insets of Figure 7 show how altering these variables meets with diminishing returns, and as such the practical benefit is somewhat limited.

Ion activation can also be achieved in the skimmer-cone region of the instrument through careful management of the pressure and accelerating potential experienced by the ions in the first regions of differential-pumping within the instrument. The dissociation products formed here from protein complex ions are similar to those generated using the collision cell ¹⁴. Dissociation of protein complexes is sometimes also observed if the pressure in the first region of differential pumping is too low, which is attributed to the ions undergoing a small number of high-energy collisions rather than a large number of very low-energy collisions ⁸⁸. Upon first inspection, this result may seem counterintuitive when viewed in the context of

Figure 6. This is because vibrational/rotational relaxation which occurs in between collisions is not taken into account in these simulations. Therefore, while the calculations shown in Figure 6 offer a useful description of the number collisions experienced by an ion in the collision cell and the ion kinetic energy losses experienced, the same approach is limited in describing the internal energy of the ions. In general, the activation of ions in this region of the instrument is effected most easily under conditions of high accelerating potentials, and lower pressures.

4.1.2 Dissociation of Activated Macromolecular Assemblies

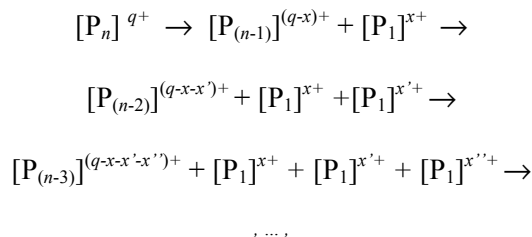
In the previous section we have dealt briefly with the way in which macromolecular ions are activated by gas-phase collisions. In this section, we turn our attention to describing the unimolecular decay process for non-covalent protein complexes. Collisional activation has been used to dissociate gas-phase protein-protein complexes since the very earliest studies that demonstrated their detection^{2,3}. In these studies the activation of the non-covalent complexes was shown to break the non-covalent bonds, rather than causing covalent-bond cleavage. While initial studies were conducted on dimers, when complexes comprised of more than two subunits were dissociated by CID, some unexpected characteristics of the dissociation process were observed. The dissociation products of tetramers in the gas-phase were found not to be dimers, as expected, but rather complementary monomers and trimers^{4,5,124}. Furthermore, it was noted that the charge partitioning between the dissociation products also appeared to be surprisingly asymmetric. For one of the proteins studied, avidin, the monomers and trimers both were centred on the 8+ charge state despite there being a three-fold difference in mass⁵. This dissociation pathway of protein complexes into highly charged monomers and relatively lowly charged oligomers missing one subunit ('stripped oligomers') has been demonstrated for a large variety of protein complexes ranging from dimers to species with more than 100 subunits¹⁴. As such the general scheme of the reaction can be described by



Scheme 1

where n is the number of subunits in the oligomer of species P , z is the number of charges on the oligomer and x is the average charge carried by the monomer P_1 , and the charge density by mass is much higher in P_1 than $P_{(n-1)}$.

Furthermore it has also been shown that the asymmetric dissociation process can re-occur, such that several monomeric units can be removed ¹⁴. This has been shown to proceed in a sequential manner as described by



Scheme 2

where n and q are the number of subunits and charges in the original oligomer, and x , x' and x'' are the average charges carried by the monomers P_1 at different dissociation steps ¹⁴. The number of dissociation steps which are observed is governed by the initial kinetic energy of the ions. Under the highest energy CID conditions currently accessible on a modified Q-ToF instrument, at 350V acceleration, up to four dissociation steps have been observed¹¹⁵.

The many observations made of the dissociation of protein complexes, including the typically high charge states for the monomeric product ions (indicative of unfolded gas-phase conformations)⁶², led one early report to suggest that “dissociation of the [oligomer] may occur by a Coulombically-driven process in which a monomer species becomes ‘unravelling’ and ejected ... with a disproportionately large share of the charge”⁵. Compelling evidence that an unfolding event occurs during the asymmetric dissociation pathway observed during CID of a multimeric assembly comes from studies performed on non-specific cytochrome *c* and α -lactalbumin dimers with and without the presence of conformational restraints. When the individual protein chains were covalently cross-linked, thereby restricting their unfolding, dissociation was shifted from an asymmetric pathway to a symmetric one^{125,126}. Furthermore, it has been shown that the estimated increase in surface area induced by monomer unfolding correlates well with the charge partitioning observed, indicating that the observed charge re-distribution is symmetric with respect to surface area while asymmetric with respect to mass ¹⁴. Recently this CID process has been visualized by means of ion-mobility mass spectrometry¹²⁷, wherein the activated state of the transthyretin tetramer was shown to populate multiple states much larger than the inactivated form. The sizes measured for these ions are consistent with a structure comprised of a single unfolded and three folded protein subunits.

Current thinking on the mechanism of dissociation of a multimeric protein complex stands as follows^{14,100,125}. An oligomeric complex is produced by ESI with the charges distributed, on average, evenly over its surface. The ion is accelerated into a gas-filled collision cell, with an increase in kinetic energy dependent on its charge state and the accelerating voltage. During each of *circa* 10^3 – 10^4 collisions with neutral gas molecules, some of this translational energy is converted into internal energy, where it is distributed among the vibrational and rotational modes of the proteins that comprise the complex. The increase in internal energy enables local unfolding/disordering events to occur. Local unfolding increases the surface area, and mobile charge carriers re-distribute to the freshly-exposed area to minimize Coulombic repulsion. Regions close to the locally-unfolded section are destabilized by this process, and further unfolding follows, with consequent further charge migration. This continues until a subunit is essentially fully unfolded, and its interactions with the oligomer are broken. An intermediate state is reached when the intra-oligomer forces holding the monomer to the rest of the oligomer are equal to the Coulombic forces pushing them apart. With the activation barrier having been overcome, the two charged species then separate, with the total charge partitioned according to their surface areas. The magnitude of the activation energy is therefore dependent on the energy required to overcome both intra-monomer and intra-oligomer interactions.

4.1.3 The CID of a Large Oligomeric Protein: Case Study

In this section we demonstrate the CID of a large oligomeric protein complex from both an experimental standpoint, and the theoretical background described above. Figure 8A shows the CID of the 200 kDa, 12 subunit oligomer formed by *TaHSP16.9*^{128,129}. The 32+ charge state was selected by using the quadrupole analyzer of a Q-ToF instrument and accelerated into a collision cell containing 30 μ bar of argon. At the lowest acceleration voltages, only signal corresponding to the intact oligomeric species is observed. Therefore the ions do not accumulate sufficient internal energy through transfer from translational energy to dissociate on the timescale of their passage through the mass spectrometer. When the accelerating voltage is increased to 80 V, with the ions therefore entering the collision cell with 2560 eV, the 12mer signal is reduced and peaks corresponding to monomers and 11mers are observed. These first dissociation products become of increasingly high intensity as the voltage is increased, until at ~120 V two additional series are observed: one corresponding to 10mers, and a second monomer distribution at lower charge states (higher *m/z*) than the first. As the voltage is increased still further, up to the maximum possible on this instrument

(200 V), no further dissociation steps occur, but the intensity of the 10mer peaks increases while those of the 11mers decrease.

Figure 8B shows how the relative intensities of the parent 12mer and the different stripped oligomers vary as a function of acceleration voltage. This plot clearly demonstrates that protein complex dissociation is a sequential reaction, as described in Scheme 2¹⁴. By setting a definition of ‘onset for dissociation’ as being the accelerating voltage at which the dissociation product amounts to 10% of the total ion current, we can use this graph to read off the threshold voltage at which dissociation occurs. In this case, the 11mer forms at 80 V and the 10mer at 145 V. As the selected ion was the 32+ charge state, these acceleration voltages correspond to initial translational energies of 2560 eV and 4640 eV respectively. As the process is sequential, the 10mer being formed from the 11mer, a crude estimate of the energy required for the 11mer to 10mer process is the difference of the two, namely 2080 eV.

Having established the dissociation threshold for the 12mer to 11mer, we examine the effect of using different accelerating voltages on energy deposition. Figure 8C shows simulations of the distance along the cell (left panel) and time taken (right) for the 12mer³²⁺ ion to reach the dissociation threshold. We can see that with the voltages accessible in this instrument, activation of these ions can be accomplished on the order of 11-341 μ s. From the inset it is clear that, even if higher energies were attainable, as on other instruments platforms described in the literature¹¹⁵, the activation process would not occur on a significantly faster timescale.

4.2 OTHER ACTIVATION TECHNIQUES

Though the vast majority of studies in which protein complexes have been dissociated have employed CID, there have been a few reports of the successful application of other activation techniques, notably SID¹¹², BIRD^{100,112,130} and ECD¹¹⁷. SID can be viewed almost as an extension of CID, but with a solid surface rather than gas molecules being the colliding neutral. The immediate benefits of SID in effectively increasing the mass of the collision partner are however difficult to realize, and it is only very recently that the first examples of successful SID of multi-protein complexes has been demonstrated. Wysocki and coworkers

have implemented an SID system into a Q-ToF, to provide an alternative technique to CID¹³¹. An initial study using non-specific cytochrome *c* dimers compared the symmetry of charge partitioning of the product ions created using SID or CID as the activation technique¹¹². In the case of the 11+ ion, asymmetric partitioning was observed by CID whereas symmetric dissociation was observed by SID. The authors suggest that this could be explained by the much faster time scale of activation of SID (on the order of picoseconds) relative to CID (tens to hundreds of microseconds). SID of larger, specific, protein complexes has also revealed that symmetric pathways can be accessed: SID of tetramers of TTR and concanavalin A revealed a significant amount of dimeric dissociation products¹³². The dodecameric sHSPs TaHSP16.9 (dissociation data shown in Figure 8) and PaHSP18.1 dissociated with the same asymmetry in mass as observed using CID, though differences in the charge partitioning were observed¹³¹.

ECD has seemed an attractive approach to the dissociation of protein complexes as several reports have suggested that ECD can initiate covalent fragmentation while maintaining non-covalent interactions¹³³⁻¹³⁵. This has been successfully applied to locate a ligand binding site on a protein implicated in Parkinson's disease, α -synuclein¹³⁶. ECD of larger protein complexes has also recently been implemented on an FT-ICR system¹¹⁷. Relatively intense signal for charge-reduced products and some limited fragmentation was observed, with analytically-useful ECD product ions only being produced from a few charge states. In those cases where fragment ions were observed, they were identified as resulting from both dimer and monomer ions ejected from the heptameric precursor ion.

BIRD, an activation technique where ions are slowly heated by absorption of black-body photons, is an alternative approach to ion activation typically implemented on ion trap-type mass analyzers^{137,138}. Due to the intrinsically temperature-resolved nature of the BIRD approach, temperature-dependant rate constants can be easily extracted from the dissociation data acquired. The first study to employ BIRD for gas-phase dissociation of a multi-protein complex was reported on the pentameric Shiga-like Toxin I¹⁰⁰. Asymmetric dissociation into monomers and tetramers was observed, even with this activation happening over seconds, and Arrhenius parameters for the dissociation were determined. Of particular note are the remarkably large pre-exponential factors determined for the dissociation which indicate that the loss of the monomeric subunit results in a substantial favorable entropy gain. This is consistent with a dissociation mechanism that proceeds

via the unfolding of a monomeric unit, as originally suggested by Smith et al.⁵. Further insights into the mechanism of dissociation of multimeric proteins^{130,139,140} and stability of the shiga toxins^{101,141} have been obtained by further BIRD experiments.

4.3 USE OF GAS-PHASE DISSOCIATION

Strikingly, despite the different mechanics and timescales involved in the various methods described above, most product ions generated by protein complex dissociation remain very similar. The dominant dissociation pathway is highly-asymmetric with respect to mass (Scheme 1) and very different to that which is observed in solution¹⁹. In the last year, however, gas-phase dissociation pathways have been observed in which the loss of a highly-charged monomer ion is not the chief decay pathway^{114,132}. In these two cases, noncovalently bound tetrameric proteins were found to undergo symmetric dissociation into dimers. Such exceptions provide us with the opportunity to increase our understanding, and possibly manipulate the mechanism of protein complex dissociation in the gas phase. Based on an increased understanding, a future challenge is to use gas-phase activation and dissociation to establish the location of subunits within protein interaction networks^{8,15}.

As more information on the mechanism of dissociation comes to light it is becoming clear that considerable structural information can be obtained for protein complexes via such dissociative approaches^{31,32}. Information can be generated regarding the nature of intra- and inter-subunit interactions, on both global and local levels, as well as details of the overall organization of subunits within an oligomer (Figure 9)¹⁴. Coupled with the thermodynamic information attainable from BIRD¹⁰⁰, and the differing dissociation pathways now being accessed^{112,114,132}, the extra dimension afforded by dissociative approaches is as an essential part of future MS investigations in structural proteomics.

5 COMPANION TECHNOLOGIES FOR MS OF PROTEIN COMPLEXES

In addition to insight from the mass-measurement of intact protein complexes and their fragments in the gas-phase, technologies can be employed in conjunction with MS to provide further insight into their composition, structure, dynamics and topology. In this section, we limit our scope to those technologies that have been used in conjunction with MS to investigate protein-protein interactions. We have organized these technologies according to the order in which they occur relative to the typical MS experiment (Figure 10). Some of the technologies, for example hydrogen/deuterium exchange, are long established, while others have only recently been coupled to MS however in all cases, the technology described provides further information on some aspect of the protein complex. For instance, affinity-based purification not only provides an effective means of obtaining protein complexes for direct analysis by MS, but also provides information on binding partners within uncharacterised protein complexes expressed at endogenous levels. While many of the technologies described in this section do not involve the detection of intact complexes themselves, though in many cases this extension is theoretically possible, the information garnered from their application can provide complementary insight into protein complex structure and function. In some cases, these technologies overlap considerably in terms of their capabilities, however in general figure 10 demonstrates the complementarity of the enabling technologies covered here. While the progress of MS as described in the previous sections of this review will play an important role in proteomics, it is when coupled with these companion technologies that the full scope and potential of MS in this field can be realized.

5.1 SAMPLE ISOLATION TECHNOLOGIES

The majority of MS studies of protein complexes have been performed using recombinantly-expressed proteins, purified using standardized molecular biology procedures. However, an exciting frontier in MS is the isolation and analysis of previously unknown protein-protein complexes directly from cells at endogenous levels of expression. Tandem affinity purification (TAP)^{142,143} is a widely-used method for MS analysis of such protein assemblies¹⁴⁴. In this experiment, a protein is expressed containing a tag consisting of two sections, at least one of which is chemically cleavable. Both sections exhibit a high-degree of affinity towards resin-bound substrates¹⁴⁵. Various additions and refinements have been made to the original TAP-

tag protocol ¹⁴⁶⁻¹⁴⁹. For example, several popular variants of the strategy incorporate poly-histidine tags (for Ni²⁺ binding affinity) as either the primary or secondary section of the tag ^{146,147}.

Since the initial landmark TAP-tag – MS experiments were reported ^{150,151}, several applications of the technology have been reported. These experiments have provided valuable data on the global organization of proteins for several organisms including examples from yeast^{147,152-154} and human ¹⁵⁵ cells. Typically, these experiments follow protocols similar to those used in bottom-up proteomics experiments. Proteins are isolated after TAP-tag purification, separated by gel-electrophoresis, enzymatically digested, and then identified by a combination of accurate peptide masses and sequence tags generated by tandem MS. This isolation strategy has been used to generate the most comprehensive map of the yeast interactome, totalling 547 protein-protein interactions, averaging 4.9 proteins per complex ¹⁵⁶.

In addition to defining the global architecture of interacting proteins within a given organism, several studies have taken a more focused approach: using affinity purification in conjunction with MS to study a single signaling pathway, regulation system, or even a single protein complex. As before, the composition of the protein complex or complexes involved is frequently the principal target of the analysis. Using such an approach, the components of the yeast nucleopore complex were identified using MALDI-ToF-ToF analysis following TAP-tag isolation ¹⁵⁷. Isolation of intact protein complexes using the TAP-tag strategy and MS analysis of the non-covalent complex is also having important implications for structural genomics. By studying MS of intact protein assemblies, along with sub-complexes generated under mildly-denaturing conditions, a high-confidence model for the subunit architecture of the yeast exosome was recently proposed, derived entirely from MS experiments ¹⁵.

The ability to examine protein complexes expressed at endogenous levels has enormous implications for MS in structural proteomics. The high-sensitivity and consequently low sample requirement relative to other structural biology techniques means that MS is well placed to provide initial low-resolution structural data on novel protein complexes. Moreover, the fact that sub-stoichiometric binding of protein subunits is readily apparent in spectra enables this technique to be applied to heterogeneous complexes that are often not amenable to other structural biology approaches.

5.2 SAMPLE LABELING TECHNOLOGIES

Nuclear magnetic resonance spectroscopy (NMR) or X-ray diffraction analysis require relatively high concentrations of homogenous protein complexes for successful structure determinations. In addition, NMR and X-ray diffraction are difficult to automate and operate in a high-throughput mode. Consequently, in situations where high-resolution structures are difficult to obtain due to the scarcity or low purity of the available protein complex, MS can contribute significantly to structural data. A general MS approach toward obtaining low-resolution structural data involves the use of chemical labeling, the most common forms being chemical cross-linking, hydrogen/deuterium exchange, and oxidative footprinting.

Often used in conjunction with TAP-Tag purification of protein complexes, chemical cross-linking methods have been developed in order to determine the topology of protein-protein complexes¹⁵⁸⁻¹⁶⁰. In these experiments, a chemical agent is added to the sample that induces a reaction to covalently link proteins together for subsequent analysis, often *via* enzymatic digestion and MS/MS. The identification of cross-linked peptides, together with knowledge of the dimensions of the cross-linking agent can be used to give distance constraints for determining the overall topology of the assembly under investigation¹⁵⁸.

A variety of cross-linking agents, that differ in terms of tag distance, amino acid specificity, and architecture have been developed^{152,153,161-178}. Generally speaking, chemical cross-linking agents fall into five categories: zero-length, homobifunctional, heterobifunctional, trifunctional, and cleavable chemical agents¹⁵⁸. Zero-length reagents act to create covalent bonds directly between two protein sites, consequently, the length constraints created by the linking event are minimal¹⁶⁴. Bifunctional cross-linking agents, using either two different or two identical reactive groups, are spaced by an inert 'linker', which imposes a longer distance constraint for subsequent structural analysis. Trifunctional and cleavable reagents impart additional functionality to the bifunctional cross-linker concept. Often, the third functional group incorporated into the cross-linking agent acts as a tag for subsequent purification of cross-linked peptides. A useful strategy is the incorporation of isotopic labeling into the cross-linking agents, as this allows low-intensity cross-linked peptides to be readily identified by inspection of their isotope pattern in mass spectra. As an additional technical enhancement to the cross-linking strategy, one study incorporated cross-linking agents that are

readily cleavable through activation in the gas-phase. Using bisuccinimidyl-succinamyl-aspartyl-prolylglycine (SuDPG) or bisuccinimidyl-succinamyl-aspartyl-proline (SuDP) as the linking agent, CID can be used to break the cross-linked peptides as a tandem MS check of their cross-linked status¹⁶¹.

Limited examples have been reported where mass spectrometry of intact protein-protein complexes has been used in conjunction with cross-linking agents to ascertain the structure of protein assemblies. In one study, intact mass spectra of the 19S lid assembly of the yeast proteasome complex were used in combination with chemical cross-linking (using Bis(Sulfosuccinimidyl)suberate (BS3) as the chemical linker) to ascertain the topology of the protein complex in solution. It was found that, while tandem MS of the intact protein complex provided a substantial amount of information on the organization of the lid complex, additional pair-wise associations were needed to choose between structures compatible with the MS data and complete the interaction map of the complex. Chemical cross-linking was used to supplement the MS/MS data and resulted in the identification of four additional protein-protein contacts that were not apparent from CID data alone⁸.

Mass spectrometry in combination with hydrogen/deuterium (H/D) exchange has developed into a robust and versatile tool for the analysis of protein-protein complexes¹⁷⁹⁻¹⁸¹. A generalized H/D exchange MS method includes the exposure of either the intact complex or the various interaction partners to deuterated solvent to induce H/D exchange, digestion of the exchanged protein material (often with pepsin, which is active at the low pH range where exchange is slowest), and identification of the exchange levels on the individual peptides observed¹⁸². Alternatively, mass spectra of the intact protein can be recorded in order to define the extent of global hydrogen exchange. Ideally, base-line levels of exchange are recorded for each protein within the complex in isolation and these values are compared to the site-specific exchange observed for the intact complex in order to determine sites of protein-protein contact. If this comparison is carried out after proteolytic digestion, region-specific information about sites of protein-protein contact can be obtained. As an additional check of exchange results, it is also possible to carry out the exchange reaction in reverse *i.e.*, in a protic solvent. In this case proteins are exposed to protonated solvent after complete deuteration of the complex has been accomplished. Both ESI and MALDI ionization have been employed to characterize the level of hydrogen exchange within peptides generated from protein complexes. ESI allows for facile online

monitoring of the exchange reaction and coupling with separation technology, often necessary to observe high sequence coverage of the complex mixture of the exchanged peptides. MALDI offers a higher tolerance to salt and detergents that often suppress signal in ESI instrumentation but requires an off-line approach for efficient coupling with chromatographic separation. Mass spectrometry serves as an ideal detector for H/D exchange studies as it is capable of distinguishing co-existing populations which differ in their exchange profile. This is in contrast to NMR detection which, while yielding excellent spatial resolution, provides only sample-wide average exchange values.

Many aspects of protein-protein complexes have been probed using hydrogen exchange mass spectrometry¹⁸³⁻¹⁹¹. In one set of experiments employing H/D exchange and ESI-MS, the location of protected residues in the multi-protein ubiquitin ligase SCF^{skp2} complex was monitored upon binding with Cks1¹⁸⁶. Hydrogen exchange data indicated that the binding event causes global conformational changes, some remote from the binding site of the protein. In another set of hydrogen exchange experiments, the flexibility of the 55 protein components of the 70S bacterial ribosome particle were investigated¹⁸⁷. These results are consistent with previous experiments, as they identify the ‘stalk’ region of the complex as being particularly flexible. The data also enable prediction of the assembly of the intact ribosome, as regions of enhanced flexibility are thought to be among the final proteins to be incorporated into the 70S particle. H/D exchange data have been used to detect the dynamic properties of amyloid fibrils derived from aggregated SH3 domains¹⁸⁸. The kinetics of the H/D exchange data revealed that dissociation and re-association of SH3 domains is common in the amyloid fibrils studied. On-line H/D exchange was used to probe the conformations of the S100A11 protein dimer during its assembly¹⁸³. The same group has utilized hydrogen exchange to distinguish artefactual protein complexes, produced during the ESI process, from *bona fide* protein-protein complexes that exist in solution¹⁹⁰. Hydrogen exchange, despite having a relatively long track-record in the analysis of protein complexes, continues to be a valuable tool in their analysis.

A method closely-related to hydrogen exchange, oxidative footprinting, uses hydroxyl radicals to oxidize the solution-accessible areas of a protein complex¹⁹²⁻¹⁹⁵. An obvious advantage of this approach over that of hydrogen exchange is that the sites of oxidation remain stable through the CID process, whereas deuterium labels are scrambled during ion activation¹⁹⁶. This enables the application of tandem MS to identify the

site(s) of solvent accessibility at the residue level, rather than the peptide level, which is generally the limit of hydrogen exchange studies by MS. Also, the irreversible nature of the modification means that back-exchange of the identified sites will not occur, a process that complicates interpretation of hydrogen exchange studies. Several methods have been suggested for producing the hydroxyl radicals necessary for protein oxidation, including Fenton chemistry (transition metal-mediated production of hydroxyl radicals) and the radiolysis of water ¹⁹⁴. In all other aspects the experimental procedure followed is similar to that used in hydrogen exchange experiments monitored using MS, *vide supra*. Proof-of-principle data have been presented for small protein-protein complexes (*e.g.*, the TfR dimer) ¹⁹⁷, and the technology is likely to be extended to larger protein assemblies in the near future.

5.3 PRE-IONIZATION TECHNOLOGIES

Many methods have been coupled with MS for the analysis of protein-protein complexes, and a significant number of those are techniques that manipulate the complex in solution prior to its ionization. For example, time-dependent data is imperative for understanding the kinetics involved in the assembly, disassembly, and subunit exchange pathways available to protein-protein complexes ¹⁹⁸. MS is ideally suited for monitoring such reactions, as the mass-resolving detection system allows for the label-free monitoring of reactants, products, and intermediate species in the reaction, without the need to separate these components prior to analysis. This approach relies on there being a sufficient difference in mass between the species of interest. If this is not the case, non-invasive labels such as isotopic labeling with ¹³C and/or ¹⁵N can be used. Reactions can also be monitored in the presence of other, potentially interfering, species (*e.g.*, protein impurities). There are many modes of operation for a time-resolved mass spectrometry experiment (Figure 11). A simple approach to recording the progress of a biological reaction with MS is to employ an off-line sampling methodology (Figure 11A). Off-line measurements have been made on systems like the subunit exchange of the transthyretin tetramer ²⁴, and the assembly of the 30S ribosome ¹⁹⁹. For higher time resolution data, alternative approaches are often used, typically the continuous sampling of the reaction mixture (Figure 11B). This approach has been employed to measure the subunit exchange kinetics of different small heat-shock proteins ^{22,23} as well as the assembly of hemoglobin¹⁸ and the MtGimC complex

An alternative methodology uses robotic sampling of a reacting mixture, electrosprayed via a chip-based array of nESI emitters²⁰⁰. While this technology has been used for screening protein-ligand interactions^{201,202}, protein-metal uptake²⁰³, protein aggregation²⁰⁴, and the binding of ligands to tetrameric transthyretin²⁰⁵. These studies highlight run-to-run reproducibility as a major advantage of the technology over manual sampling methods. Recently this robotic sampling approach also be used to achieve many modes of automated time-resolved data acquisition²⁰⁶. This automated approach can acquire data continuously as sample is injected periodically (Figure 11C), record different reaction mixtures into the same data file (Figure 11D), or record each isolated reaction mixture into its own data file (Figure 11E). Aside from the reproducibility benefits, the time resolution that the technology can provide is much greater than that achieved using off-line methods and negates the possibility of electro-chemical effects associated with electrospraying of the same sample for a prolonged period of time²⁰⁶⁻²⁰⁸.

Several flow injection-based technologies have been reported for monitoring the kinetics of biological reactions on the millisecond time scale. This instrumentation typically includes the ability to use both variable length capillary reaction tubes and variable flow-rate (including stopped-flow) to record reaction products at different time points²⁰⁹. Recently, this technology has been used to measure the disassembly of hemoglobin²¹⁰, as well as the unfolding of nitric oxide synthase²¹¹.

The reactions of protein assemblies are highly temperature dependent. Consequently, recording temperature-dependant rate constants is of primary importance, and MS-based methods provide several advantages for this application. Among these are the ability to measure the kinetics of several species in parallel, binding competition experiments, and low-sample consumption. Online temperature control of the nESI capillary containing the protein complex allows reversible thermal effects to be investigated¹⁹. In these experiments, a Peltier-type heating device was used to raise the temperature uniformly within the nESI capillary and observe the thermally-induced dissociation of protein complexes in solution^{19,212}. A device for cooling an ESI emitter to as low as 4°C also been reported and applied to protein-ligand complexes²¹³. Such a device also has obvious potential for protein-protein complexes as well.

While a few successes have been reported, protein-protein binding constants are often difficult to measure using MS-only technologies²¹⁴. The ion intensities observed in mass spectra depend not only on the amounts of different protein-protein complexes present in solution, but also on their relative ionization efficiency and transmission through the instrument. Therefore, ion intensity is often a poor quantitative indicator of the amount of bound or unbound protein complex in solution, as the ionization efficiencies of the two species may differ substantially from each other²¹⁵. Moreover, for instruments that employ micro-channel plate (MCP)-type detectors, the velocity-dependent response of the ion detector may have to be taken into account if the binding partners differ significantly in m/z ^{216,217}. To overcome these limitations, other methodologies have been developed for use in conjunction with MS. One such method, surface plasmon resonance (SPR), uses the interaction of proteins with a surface-specific evanescent wave propagating on a conductive surface irradiated at a specific wavelength and angle to detect the presence and strength of protein-protein binding events. This technology has been successfully coupled to mass spectrometry for the high-throughput identification of protein-protein complexes in terms of both their composition and binding strength²¹⁸⁻²²¹. In typical experiments, solutions containing various potential binding partners are washed over the SPR sensor surface (containing either a single molecule or an array of molecules). Bound protein or ligand is then either recovered for MS analysis, or desorbed/ionized directly from the chip surface. SPR-MS technology has been applied to a variety of protein complexes²²²⁻²²⁶. For example, binding events between tetrameric transthyretin (TTR) and retinol binding protein (RBP) were detected in the analysis of human plasma²²⁵. Similarly, complexes between insulin-like growth factor proteins were investigated and a novel truncated form of IGF-2 was identified through SPR-MS analysis²²⁶. Such experiments will undoubtedly prove valuable for protein complexes, adding a quantitative measure of interaction strength to the connectivity information that can be obtained for protein-protein interaction networks.

5.4 POST-IONIZATION TECHNOLOGIES

Ion mobility (IM) separation has been a tool for physicists²²⁷ and analytical chemists²²⁸ for decades, but has been applied to proteins and protein-protein complexes only recently. Several outstanding reviews are available on both the evolution of the ion mobility-mass spectrometry (IM-MS) method and its application to biomolecules^{11,229-232}. The basic concept of IM separation involves the injection of a temporally-defined

packet of ions into an environment of inert neutral gas (Figure 12A). Ions are directed through this environment by an electric field. Those ions that are large undergo a greater number of collisions with the inert neutrals, *i.e.*, their collision cross-section is large, and thus travel more slowly through the IM device than those ions that present a smaller cross-section. It is important to note that there are other methods of determining the size of protein-protein complexes in the gas-phase, although IM is currently the only method that has been used to analyze protein assemblies. One particularly important method, ion energy-loss measurements^{123,233}, uses a stopping potential at the terminus of a gas-filled region of the instrument, and can be implemented on widely available commercial instrumentation.

IM-MS instrumentation has evolved dramatically since it was first applied to the analysis of biomolecules²³⁴, and falls into two basic types: drift-tube devices (where the flight time of the ion of interest is recorded) and differential devices (where the voltage required to observe the ion of interest is recorded). Figure 12B summarizes some of the most prominent methods of separating ions according to their ion mobility. Of the two most widely-used types of IM instrumentation, drift tube-type devices (Figure 12BI) have been applied more extensively to biomolecules. Initial experiments were performed with instruments constructed using an electrode ring-stack that defined a strictly linear potential gradient across the drift region²³⁴⁻²³⁹. These devices were designed to operate at both low pressure (for ease of interfacing to mass analyzers that operate at high-vacuum) and high pressure (for higher-resolution ion mobility separations). While very effective at making high-resolution ion mobility measurements, these devices exhibited poor transmission efficiency (<0.01%, in some cases, due to radial diffusion of the ions during IM separation). This rendered the application of this technology to problems in biology and biochemistry extremely difficult. Instrument development has included a number of devices designed to increase transmission efficiency. These include the use of alternative DC field gradients²⁴⁰⁻²⁴², adding RF focusing fields throughout the drift tube^{243,244}, strong RF/DC focusing at the end of the drift tube²⁴⁵ and traveling DC pulses in combination with RF^{109,246}, all designed to either confine ions to the center-axis of the device or to recover ions at the end of the drift tube after radial diffusion has taken place. These advancements, some of which purport to have increased IM transmission efficiency to 100%, have made IM-MS a growing technology in the area of protein analysis.

Differential-type IM separators, while a primary technology for defense and security applications, have been employed less widely than drift tube analyzers for the analysis of large biomolecules in conjunction with MS (Figure 12BIII). Stand-alone differential mobility analyzers (DMA) have been used by several groups for the analysis of large proteins and protein complexes^{121,247-249}. However, most reports that use a DMA-MS combination are still focused on non-protein or non-biological analytes²⁵⁰. One possible exception to this statement is high-field asymmetric waveform ion mobility spectrometry (FAIMS)²⁵¹, a variant form of ion mobility that uses a combination of strong and weak field strengths to separate ions (Figure 12BIV). Not only has this technology been extensively applied to the analysis of proteins, but the last few years have seen a rapid development of the technology towards higher resolution and more informative analysis of proteins²⁵²⁻²⁵⁸.

The development of new IM-MS instrumentation is continuing at a rapid pace. Several hybrid-type ion mobility devices have been reported recently, including tandem arrangements of drift tubes (IM-IM) connected via high-transmission efficiency ion funnel devices^{259,260} and FAIMS in combination with drift tube separation^{256,261}. Alternative methods of performing high-resolution IM separations have also recently been reported^{262,263}. For example, diffusion-limited resolution in excess of 600 ($t/\Delta t$) has been predicted for an ion mobility separator where ions are driven by an electric field against a counter-current of gas²⁶² (Figure 12BII). This general method creates a scanning-type IM separator capable of both high resolution and high-transmission efficiency separation.

IM-MS has been applied in a limited number of cases to the analysis of protein-protein complexes. The aim of these studies is primarily at the determination of the quaternary structure of the complex in the gas phase. Studies reported on smaller non-covalent assemblies include studies of various forms of the A β peptide indicating that up to dodecamers of the peptide are formed and exhibit a barrel-like topology²⁶⁴⁻²⁶⁶. In addition, the structure of the α -synuclein dimer has been investigated by IM-MS²⁶⁷. The structural information obtained from IM can often increase as the number of proteins that comprises an assembly increases. This is primarily due to the fact that having a larger number of protein subunits allows a larger number of quaternary arrangements, with larger differences in collision cross-section. Stand-alone IM studies, presented alongside MS results performed on a separate instrument, of the 20S proteasome complex

have been reported. The results indicate that charge-reduced proteasome ions, carrying a single ionic charge, retain their overall shape in the gas-phase²⁶⁸. Using an IM-MS approach, it was possible to examine the gas phase structure of the undecameric *trp* RNA-binding attenuation protein (TRAP)¹³. The data presented show that a population of desolvated TRAP ions retain a ring-type topology similar to that indicated by X-ray diffraction analysis. It was also possible to perform activation experiments on this protein complex and observe deformation of quaternary structure as a function of internal energy. IM-MS is a particularly promising technology for the analysis of protein complexes as it promises to add a certain amount of topological information to the complement the compositional available from the MS step of the experiments.

6 THE FUTURE

The material covered in this review is focused on the synergy achieved between technology and application of that technology to problems in both structural genomics and proteomics. Specifically, the growth in understanding of the conditions required to maintain and detect protein complexes has driven the continued and rapid growth of MS for intact protein complexes. As such, MS is now established in structural genomics, with numerous studies where key information about the stoichiometry of non-covalent complexes has facilitated deduction of high resolution structures³². Moreover with tandem mass spectrometers capable of transmitting and dissociating such high mass species becoming available in a number of laboratories worldwide, and the progress in understanding the mechanism of collisional activation and dissociation of protein complexes, we anticipate that this methodology will also become increasingly important for determining the architecture of protein complexes³¹. As the MS field moves from abundant protein complexes into low copy number cellular extracts, new challenges have to be overcome to establish the study of non-covalent complexes firmly in the area of proteomics. Specifically, these challenges require an increase in the ionization efficiency of the electrospray process and further optimisation of the instrumentation for their detection.

Despite the tremendous progress that has occurred over last decade, further opportunities are emerging within the context of structural genomics and proteomics initiatives. Determining the interaction networks between proteins through MS methods is an exciting avenue of research, though new approaches and supporting software are needed to determine the most likely arrangement of protein subunits within any given assembly²⁶⁹. IM-MS, *vide supra*, of protein complexes is an area of research that is only in its infancy. While it has been shown that solution-phase structure can be maintained intact in the gas phase¹³, the challenge is now to use this technology to determine topologies of protein complexes for which no other structural data exists. Particularly tantalising is the prospect of visualising intact protein complexes. Recently recognised is the potential for MS to be used in a preparative manner, by isolating a defined m/z range from a population of protein ions. By decelerating ions and collecting them on a surface *in vacuo* (termed ‘soft-landing’), it has been possible to recover gaseous protein downstream in a mass spectrometer²⁷⁰⁻²⁷². In preliminary experiments, we have shown the 800kDa GroEL 14-mer was transmitted intact through the mass

spectrometer and soft landed onto an EM grid placed in the ion beam of a quadrupole ToF mass spectrometer. Subsequent EM imaging revealed that the intact GroEL 14-mer was maintained intact ²⁷³. While further optimization of the technique is required to generate 3D reconstructions, we believe these developments will provide an exciting possibility to enable structure determination of hitherto intractable complexes – an enticing prospect for mass spectrometry and structural biology more generally.

At the start of this article we reviewed progress in this field of research over the last decade. Having reached the end, it is now prudent to speculate about developments that the next decade will bring. Undoubtedly further strengthening of the links between MS of non-covalent complexes and structural genomics will ensue. These links will not only aid in the definition of protein interactions, but also provide the all important detail of stoichiometry, connectivity, binding strengths, and shape. Furthermore, direct observation of the dynamic reactions of these complexes, such as assembly, disassembly, and subunit exchange reactions will combine with the structural information to provide a detailed view of their function. As we write, however, MS has already entered a new era for characterising intact cellular complexes at endogenous expression levels. The challenge that remains therefore is to examine dynamic reactions for protein complexes within the confines of the cellular environment. In theory comparison of intact complexes isolated from cells exposed to different stimuli could reveal the re-organization that takes place in response to signalling events. While quantitative proteomics techniques can address temporal changes in individual proteins within a complex, visualizing the intact complex and adding dynamics to static protein interaction maps is now the tantalising goal.

7 ACKNOWLEDGEMENTS

The authors thank Alan Sandercock (University of Cambridge) for help collecting data (Fig 5) and critical appraisal of the manuscript, and Tara Pukala (University of Cambridge) for additional comments; and Giorgio Favrin (University of Cambridge) and Bruce Thomson (MDS Sciex) for performing the Monte-Carlo (Fig 3) and ion trajectory (Fig 5) simulations respectively. JLPB and BTR are supported by the Engineering and Physical Sciences Research Council Basic Technology Program, DAS was supported by the Natural Sciences and Engineering Research Council of Canada and CVR is a Professor of the Royal Society.

8 REFERENCES

- (1) Sali, A.; Glaeser, R.; Earnest, T.; Baumeister, W. *Nature* **2003**, 422, 216.
- (2) Baca, M.; Kent, S. B. H. *J. Am. Chem. Soc.* **1992**, 114, 3992.
- (3) Huang, E. C.; Pramanik, B. N.; Tsarbopoulos, A.; Reichert, P.; Ganguly, A. K.; Trotta, P. P.; Nagabhushan, T. L.; Covey, T. R. *J. Am. Soc. Mass Spectrom.* **1993**, 4, 624.
- (4) Light-Wahl, K. J.; Winger, B. E.; Smith, R. D. *J. Am. Chem. Soc.* **1993**, 115, 5869.
- (5) Light-Wahl, K. J.; Schwartz, B. L.; Smith, R. D. *J. Am. Chem. Soc.* **1994**, 116, 5271.
- (6) Loo, J. A. *Mass. Spectrom. Rev.* **1997**, 16, 1.
- (7) Aquilina, J. A.; Benesch, J. L. P.; Bateman, O. A.; Slingsby, C.; Robinson, C. V. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, 100, 10611.
- (8) Sharon, M.; Taverner, T.; Ambroggio, X. I.; Deshaies, R. J.; Robinson, C. V. *PLoS Biol.* **2006**, 4, e267.
- (9) van Duijn, E.; Simmons, D. A.; van den Heuvel, R. H.; Bakkes, P. J.; van Heerikhuizen, H.; Heeren, R. M.; Robinson, C. V.; van der Vies, S. M.; Heck, A. J. *J. Am. Chem. Soc.* **2006**, 128, 4694.
- (10) Ilag, L. L.; Videler, H.; McKay, A. R.; Sobott, F.; Fucini, P.; Nierhaus, K. H.; Robinson, C. V. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, 102, 8192.
- (11) Ruotolo, B. T.; Robinson, C. V. *Curr. Opin. Chem. Biol.* **2006**, 10, 402.
- (12) Wyttenbach, T.; Bowers, M. T. *Annu. Rev. Phys. Chem.* **2007**, 58, 511.
- (13) Ruotolo, B. T.; Giles, K.; Campuzano, I.; Sandercock, A. M.; Bateman, R. H.; Robinson, C. V. *Science* **2005**.
- (14) Benesch, J. L. P.; Aquilina, J. A.; Ruotolo, B. T.; Sobott, F.; Robinson, C. V. *Chem. Biol.* **2006**, 13, 597.
- (15) Hernández, H.; Dziembowski, A.; Taverner, T.; Seraphin, B.; Robinson, C. V. *EMBO Rep.* **2006**, 7, 605.
- (16) Hernández, H.; Robinson, C. V. *J. Biol. Chem.* **2001**, 276, 46685.
- (17) Fändrich, M.; Tito, M. A.; Leroux, M. R.; Rostom, A. A.; Hartl, F. U.; Dobson, C. M.; Robinson, C. V. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, 97, 14151.
- (18) Boys, B. L.; Konermann, L. *J. Am. Soc. Mass Spectrom.* **2007**, 18, 8.
- (19) Benesch, J. L. P.; Sobott, F.; Robinson, C. V. *Anal. Chem.* **2003**, 75, 2208.
- (20) Griffith, W. P.; Kaltashov, I. A. *Biochemistry* **2003**, 42, 10024.
- (21) Simmons, D. A.; Wilson, D. J.; Lajoie, G. A.; Doherty-Kirby, A.; Konermann, L. *Biochemistry* **2004**, 43, 14792.
- (22) Sobott, F.; Benesch, J. L. P.; Vierling, E.; Robinson, C. V. *J. Biol. Chem.* **2002**, 277, 38921.
- (23) Aquilina, J. A.; Benesch, J. L. P.; Ding, L. L.; Yaron, O.; Horwitz, J.; Robinson, C. V. *J. Biol. Chem.* **2005**, 280, 14485.
- (24) Keetch, C. A.; Bromley, E. H.; McCammon, M. G.; Wang, N.; Christodoulou, J.; Robinson, C. V. *J. Biol. Chem.* **2005**, 280, 41667.
- (25) Loo, J. A. *Int. J. Mass Spectrom.* **2000**, 200, 175.
- (26) Miranker, A. D. *Curr. Opin. Struct. Biol.* **2000**, 10, 601.
- (27) Sobott, F.; Robinson, C. V. *Curr. Opin. Struct. Biol.* **2002**, 12, 729.
- (28) Heck, A. J. R.; van den Heuvel, R. H. H. *Mass Spectrom. Rev.* **2004**, 23, 368.
- (29) van den Heuvel, R. H.; Heck, A. J. R. *Curr. Opin. Chem. Biol.* **2004**, 8, 519.
- (30) McCammon, M. G.; Robinson, C. V. *Curr. Opin. Chem. Biol.* **2004**, 8, 60.
- (31) Benesch, J. L. P.; Robinson, C. V. *Curr. Opin. Struct. Biol.* **2006**, 16, 245.
- (32) Sharon, M.; Robinson, C. V. *Annu. Rev. Biochem.* **2007**, In Press; Corrected Proof.
- (33) Dole, M.; Mack, L. L.; Hines, R. L. *J. Chem. Phys.* **1968**, 49, 2240.
- (34) Yamashita, M.; Fenn, J. B. *J. Phys. Chem.* **1984**, 88, 4451.
- (35) Aleksandrov, M. L.; Gall, L. N.; Krasnov, V. N.; Nikolaev, V. I.; Pavlenko, V. A.; Shkurov, V. A. *Dokl. Akad. Nauk SSSR* **1984**, 277, 379.
- (36) Karas, M.; Hillenkamp, F. *Anal. Chem.* **1988**, 60, 2299.

- (37) Tanaka K. ; Waki H. ; Ido Y.; Akita S. ; Yoshida Y.; T., Y. *Rapid Comm. Mass Spectrom.* **1988**, 2, 151.
- (38) Cohen, L. R. H.; Strupat, K.; Hillenkamp, F. *J. Am. Soc. Mass Spectrom.* **1997**, 8, 1046.
- (39) Rosinke B; Strupat K; Hillenkamp F; Rosenbusch J; Dencher N; Krüger U; H-J, G. *J. Mass Spectrom.* **1995**, 30, 1462.
- (40) Horneffer, V.; Strupat, K.; Hillenkamp, F. *J. Am. Soc. Mass Spectrom.* **2006**, 17, 1599.
- (41) Lecchi, P.; Pannell, L. K. *J. Am. Soc. Mass Spectrom.* **1995**, 6, 972.
- (42) Woods, A. S.; Buchsbaum, J. C.; Worrall, T. A.; Berg, J. M.; Cotter, R. J. *Anal. Chem.* **1995**, 67, 4462.
- (43) Woods, A. S.; Huestis, M. A. *J. Am. Soc. Mass Spectrom.* **2001**, 12, 88.
- (44) Juhasz, P.; Biemann, K. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, 91, 4333.
- (45) Zhigilei, L. V.; Leveugle, E.; Garrison, B. J.; Yingling, Y. G.; Zeifman, M. I. *Chem. Rev.* **2003**, 103, 321.
- (46) Karas, M.; Kruger, R. *Chem. Rev.* **2003**, 103, 427.
- (47) Knochenmuss, R.; Zenobi, R. *Chem. Rev.* **2003**, 103, 441.
- (48) Dreisewerd, K. *Chem. Rev.* **2003**, 103, 395.
- (49) Georgiou, S.; Hillenkamp, F. *Chem. Rev.* **2003**, 103, 317.
- (50) Kebarle, P.; Tang, L. *Anal. Chem.* **1993**, 65, 972.
- (51) de la Mora, J. F.; Loscertales, I. G. *J. Fluid Mech.* **1994**, 260, 155.
- (52) de la Mora, J. F. *Anal. Chim. Acta* **2000**, 406, 93.
- (53) Rayleigh, L. *Philos. Mag.* **1882**, 14, 184.
- (54) Gomez, A.; Tang, K. Q. *Phys. Fluids* **1994**, 6, 404.
- (55) Kebarle, P.; Peschke, M. *Anal. Chim. Acta* **2000**, 406, 11.
- (56) Iribarne, J. V.; Thomson, B. A. *J. Chem. Phys.* **1976**, 64, 2287.
- (57) Gamero-Castaño, M.; de la Mora, J. F. **2000**, 406, 67.
- (58) Kebarle, P. *Journal of Mass Spectrometry* **2000**, 35, 804.
- (59) Verkerk, U. H.; Kebarle, P. *J. Am. Soc. Mass Spectrom.* **2005**, 16, 1325.
- (60) Felitsyn, N.; Peschke, M.; Kebarle, P. *Int. J. Mass Spectrom.* **2002**, 219, 39.
- (61) Kaltashov, I. A.; Mohimen, A. *Anal. Chem.* **2005**, 77, 5370.
- (62) Chowdhury, S. K.; Katta, V.; Chait, B. T. *J. Am. Chem. Soc.* **1990**, 112, 9012.
- (63) Dobo, A.; Kaltashov, I. A. *Anal. Chem.* **2001**, 73, 4763.
- (64) Grandori, R. *Protein Sci.* **2002**, 11, 453.
- (65) Gumerov, D. R.; Dobo, A.; Kaltashov, I. A. *European J. Mass Spectrom.* **2002**, 8, 123.
- (66) Vis, H.; Heinemann, U.; Dobson, C. M.; Robinson, C. V. *Protein Sci.* **1998**, 120, 6427.
- (67) Sogbein, O. O.; Simmons, D. A.; Konermann, L. *J. Am. Soc. Mass Spectrom.* **2000**, 11, 312.
- (68) Samalikova, M.; Grandori, R. *J. Am. Chem. Soc.* **2003**, 125, 13352.
- (69) Samalikova, M.; Matecko, I.; Muller, N.; Grandori, R. *Anal. Bioanal. Chem.* **2004**, 378, 1112.
- (70) Samalikova, M.; Grandori, R. *J. Mass. Spectrom.* **2005**, 40, 503.
- (71) Kaltashov, I. A.; Eyles, S. J. *Mass Spectrom. Rev.* **2002**, 21, 37.
- (72) Konermann, L.; Douglas, D. J. *Rapid Commun. Mass Spectrom.* **1998**, 12, 435.
- (73) Konermann, L.; Collings, B. A.; Douglas, D. J. *Biochemistry* **1997**, 36, 5554.
- (74) Mirza, U. A.; Cohen, S. L.; Chait, B. T. *Anal. Chem.* **1993**, 65, 1.
- (75) Loo, J. A.; Edmonds, C. G.; Udseth, H. R.; Smith, R. D. *Anal. Chem.* **1990**, 62, 693.
- (76) Wagner, D. S.; Anderegg, R. J. *Anal. Chem.* **1994**, 66, 706.
- (77) Hautreux, M.; Hue, N.; de Kerdaniel, A. D.; Zahir, A.; Malec, V.; Laprévotte, O. *Int. J. Mass Spectrom.* **2004**, 231, 131.
- (78) Wilm, M. S.; Mann, M. *Int. J. Mass Spectrom. Ion Proc.* **1994**, 136, 167.
- (79) Wilm, M.; Mann, M. *Anal. Chem.* **1996**, 68, 1.
- (80) El-Faramawy, A.; Siu, K. W.; Thomson, B. A. *J Am Soc Mass Spectrom* **2005**, 16, 1702.
- (81) Smith, R. D.; Shen, Y.; Tang, K. *Acc. Chem. Res.* **2004**, 37, 269.
- (82) Juraschek, R.; Dulcks, T.; Karas, M. *J. Am. Soc. Mass Spectrom.* **1999**, 10, 300.

- (83) Karas, M.; Bahr, U.; Dulcks, T. *Fresenius J. Anal. Chem.* **2000**, 366, 669.
- (84) El-Faramawy, A.; Siu, K. W. M.; Thomson, B. A. *J. Am. Soc. Mass Spectrom.* **2005**, 16, 1702.
- (85) Sobott, F.; Robinson, C. V. *Int. J. Mass Spectrom.* **2004**, 236, 25.
- (86) McKay, A. R.; Ruotolo, B. T.; Ilag, L. L.; Robinson, C. V. *J. Am. Chem. Soc.* **2006**, 128, 11433.
- (87) Rostom, A. A.; Robinson, C. V. *J. Am. Chem. Soc.* **1999**, 121, 4718.
- (88) Schmidt, A.; Bahr, U.; Karas, M. *Anal. Chem.* **2001**, 73, 6040.
- (89) Tahallah, N.; Pinkse, M.; Maier, C. S.; Heck, A. J. R. *Rapid Commun. Mass Spectrom.* **2001**, 15, 596.
- (90) Douglas, D. J.; French, J. B. *J. Am. Soc. Mass Spectrom.* **1992**, 3, 398.
- (91) Krutchinsky, A. N.; Chernushevich, I. V.; Spicer, V. L.; Ens, W.; Standing, K. G. *J. Am. Soc. Mass Spectrom.* **1998**, 9, 569.
- (92) Chernushevich, I. V.; Thomson, B. A. *Anal. Chem.* **2004**, 76, 1754.
- (93) Liu, P.; Ziemann, P. J.; Kittelson, D. B.; McMurtry, P. H. *Aerosol Sci. Tech.* **1995**, 22, 293.
- (94) Nettleton, E. J.; Sunde, M.; Lai, Z.; Kelly, J. W.; Dobson, C. M.; Robinson, C. V. *J. Mol. Biol.* **1998**, 281, 553.
- (95) Chernushevich, I.; Loboda, A. V.; Thomson, B. A. *J. Mass Spectrom.* **2001**, 36, 849.
- (96) Tito, M. A.; Tars, K.; Valegard, K.; Hajdu, J.; Robinson, C. V. *J. Am. Chem. Soc.* **2000**, 122, 3550.
- (97) Sanglier, S.; Leize, E.; Van Dorsselaer, A.; Zal, F. *J. Am. Soc. Mass Spectrom.* **2003**, 14, 419.
- (98) Pinkse, M. W.; Maier, C. S.; Kim, J. I.; Oh, B. H.; Heck, A. J. *J. Mass Spectrom.* **2003**, 38, 315.
- (99) Wang, Y.; Schubert, M.; Ingendoh, A.; Franzen, J. *Rapid Commun. Mass Spectrom.* **2000**, 14, 12.
- (100) Felitsyn, N.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2001**, 73, 4647.
- (101) Kitova, E. N.; Kitov, P. I.; Bundle, D. R.; Klassen, J. S. *Glycobiology* **2001**, 11, 605.
- (102) Hu, Q. Z.; Noll, R. J.; Li, H. Y.; Makarov, A.; Hardman, M.; Cooks, R. G. *J. Mass Spectrom.* **2005**, 40, 430.
- (103) Labastie, P.; Doy, M. *Int. J. Mass Spectrom. Ion Proc.* **1989**, 91, 105.
- (104) Winger, B. E.; Light-Wahl, K. J.; Ogorzalek-Loo, R. R.; Udseth, H. R.; Smith, R. D. *J. Am. Soc. Mass Spectrom.* **1993**, 4, 536.
- (105) Sobott, F.; Hernández, H.; McCammon, M. G.; Tito, M. A.; Robinson, C. V. *Anal. Chem.* **2002**, 74, 1402.
- (106) Boyle, J. G.; Whitehouse, C. M.; Fenn, J. B. *Rapid Commun. Mass Spectrom.* **1991**, 5, 400.
- (107) Tang, X. J.; Brewer, C. F.; Saha, S.; Chernushevich, I.; Ens, W.; Standing, K. G. *Rapid Commun. Mass Spectrom.* **1994**, 8, 750.
- (108) Loboda, A.; Krutchinsky, A.; Loboda, O.; McNabb, J.; Spicer, V.; Ens, W.; Standing, K. *European J. Mass Spectrom.* **2000**, 6, 531.
- (109) Giles, K.; Pringle, S. D.; Worthington, K. R.; Little, D.; Wildgoose, J. L.; Bateman, R. H. *Rapid Commun. Mass Spectrom.* **2004**, 18, 2401.
- (110) Ganem, B.; Li, Y. T.; Henion, J. D. *J. Am. Chem. Soc.* **1991**, 113, 6294.
- (111) Ganem, B.; Li, Y. T.; Henion, J. D. *J. Am. Chem. Soc.* **1991**, 113, 7818.
- (112) Jones, C. M.; Beardsley, R. L.; Galhena, A. S.; Dagan, S.; Cheng, G.; Wysocki, V. H. *J. Am. Chem. Soc.* **2006**, 128, 15044.
- (113) Raspopov, S. A.; El-Faramawy, A.; Thomson, B. A.; Siu, K. W. *Anal. Chem.* **2006**, 78, 4572.
- (114) van den Heuvel, R. H.; van Duijn, E.; Mazon, H.; Synowsky, S. A.; Lorenzen, K.; Versluis, C.; Brouns, S. J.; Langridge, D.; van der Oost, J.; Hoyes, J.; Heck, A. J. *Anal. Chem.* **2006**, 78, 7473.

- (115) Benesch, J. L. P.; Ruotolo, B. T.; Sobott, F.; Wildgoose, J.; Gilbert, A.; Bateman, R.; Robinson, C. V. *In Preparation*.
- (116) Aebersold, R.; Mann, M. *Nature* **2003**, 422, 198.
- (117) Geels, R. B.; van der Vies, S. M.; Heck, A. J.; Heeren, R. M. *Anal. Chem.* **2006**, 78, 7191.
- (118) Senko, M. W.; McLafferty, F. W. *Annu. Rev. Biophys. Biomol. Struct.* **1994**, 23, 763.
- (119) Bowers, M. T.; Marshall, A. G.; McLafferty, F. W. *J. Phys. Chem.* **1996**, 100, 12897.
- (120) Brodbelt, J. S. *Mass Spectrom. Rev.* **1997**, 16, 91.
- (121) Kaufman, S. L.; Kuchumov, A. R.; Kazakevich, M.; Vinogradov, S. N. *Anal. Biochem.* **1998**, 259, 195.
- (122) Douglas, D. J. *J. Phys. Chem.* **1982**, 86, 185.
- (123) Gill, A. C.; Jennings, K. R.; Wyttenbach, T.; Bowers, M. T. *Int. J. Mass Spectrom.* **2000**, 196, 685.
- (124) Schwartz, B. L.; Bruce, J. E.; Anderson, G. A.; Hofstadler, G. A.; Rockwood, A. L.; Smith, R. D.; Chilkoti, A.; Stayton, P. S. *J. Am. Soc. Mass Spectrom.* **1995**, 6, 459.
- (125) Jurchen, J. C.; Williams, E. R. *J. Am. Chem. Soc.* **2003**, 125, 2817.
- (126) Jurchen, J. C.; Garcia, D. E.; Williams, E. R. *J. Am. Soc. Mass Spectrom.* **2004**, 15, 1408.
- (127) Ruotolo, B. T.; Hyung, S.-J.; Robinson, P. M.; Giles, K.; Bateman, R.; Robinson, C. V. *Angew Chem Int Ed Engl* **2007**, Submitted.
- (128) van Montfort, R. L.; Basha, E.; Friedrich, K. L.; Slingsby, C.; Vierling, E. *Nat. Struct. Biol.* **2001**, 8, 1025.
- (129) van Montfort, R.; Slingsby, C.; Vierling, E. *Adv. Protein Chem.* **2002**, 59, 105.
- (130) Felitsyn, N.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2002**, 13, 1432.
- (131) Galhena, A. S.; Dagan, S.; Jones, C. M.; Beardsley, R. L.; Wysocki, V. H. *In Preparation*.
- (132) Beardsley, R. L.; Jones, C. M.; Galhena, A. S.; Wysocki, V. H. *In Preparation*.
- (133) Haselmann, K. F.; Jorgensen, T. J.; Budnik, B. A.; Jensen, F.; Zubarev, R. A. *Rapid Commun. Mass Spectrom.* **2002**, 16, 2260.
- (134) Zubarev, R. A. *Curr. Opin. Biotechnol.* **2004**, 15, 12.
- (135) Yang, J.; Mo, J.; Adamson, J. T.; Hakansson, K. *Anal. Chem.* **2005**, 77, 1876.
- (136) Xie, Y.; Zhang, J.; Yin, S.; Loo, J. A. *J. Am. Chem. Soc.* **2006**, 128, 14432.
- (137) Price, W. D.; Schnier, P. D.; Jockusch, R. A.; Strittmatter, E. F.; Williams, E. R. *J. Am. Chem. Soc.* **1996**, 118, 10640.
- (138) Dunbar, R. C.; McMahon, T. B. *Science* **1998**, 279, 194.
- (139) Sinelnikov, I.; Kitova, E. N.; Klassen, J. S.; Armstrong, G. D. *J. Am. Soc. Mass Spectrom.* **2007**, 18, 688.
- (140) Sinelnikov, I.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2007**, 18, 617.
- (141) Kitova, E. N.; Daneshfar, R.; Marcato, P.; Mulvey, G. L.; Armstrong, G.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2005**, 16, 1957.
- (142) Puig, O.; Caspary, F.; Rigaut, G.; Rutz, B.; Bouveret, E.; Bragado-Nilsson, E.; Wilm, M.; Seraphin, B. *Methods* **2001**, 24, 218.
- (143) Rigaut, G.; Shevchenko, A.; Rutz, B.; Wilm, M.; Mann, M.; Seraphin, B. *Nat. Biotechnol.* **1999**, 17, 1030.
- (144) Zhao, Y.; Muir, T. W.; Kent, S. B.; Tischer, E.; Scardina, J. M.; Chait, B. T. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, 93, 4020.
- (145) Vasilescu, J.; Figeys, D. *Curr. Opin. Biotechnol.* **2006**, 17, 394.
- (146) Markillie, L. M.; Lin, C. T.; Adkins, J. N.; Audberry, D. L.; Hill, E. A.; Hooker, B. S.; Moore, P. A.; Moore, R. J.; Shi, L.; Wiley, H. S.; Kery, V. *Proteome Res.* **2005**, 4, 268.
- (147) Tagwerker, C.; Flick, K.; Cui, M.; Guerrero, C.; Dou, Y.; Auer, B.; Baldi, P.; Huang, L.; Kaiser, P. *Mol. Cell. Proteomics* **2006**, 5, 737.
- (148) Junttila, M. R.; Saarinen, S.; Schmidt, T.; Kast, J.; J., W. *Proteomics* **2005**, 5, 1199.
- (149) Drakas, R.; Prisco, M.; Baserga, R. *Proteomics* **2005**, 5, 132.
- (150) Gavin, A. C.; Bosche, M.; Krause, R.; Grandi, P.; Marzioch, M.; Bauer, A.; Schultz, J.; Rick, J. M.; Michon, A. M.; Cruciat, C. M.; Remor, M.; Hofert, C.; Schelder, M.;

- Brajenovic, M.; Ruffner, H.; Merino, A.; Klein, K.; Hudak, M.; Dickson, D.; Rudi, T.; Gnau, V.; Bauch, A.; Bastuck, S.; Huhse, B.; Leutwein, C.; Heurtier, M. A.; Copley, R. R.; Edelmann, A.; Querfurth, E.; Rybin, V.; Drewes, G.; Raida, M.; Bouwmeester, T.; Bork, P.; Seraphin, B.; Kuster, B.; Neubauer, G.; Superti-Furga, G. *Nature* **2002**, *415*, 141.
- (151) Ho, Y.; Gruhler, A.; Heilbut, A.; Bader, G. D.; Moore, L.; Adams, S. L.; Millar, A.; Taylor, P.; Bennett, K.; Boutilier, K.; Yang, L.; Wolting, C.; Donaldson, I.; Schandorff, S.; Shewnarane, J.; Vo, M.; Taggart, J.; Goudreau, M.; Muskat, B.; Alfarano, C.; Dewar, D.; Lin, Z.; Michalickova, K.; Willems, A. R.; Sassi, H.; Nielsen, P. A.; Rasmussen, K. J.; Andersen, J. R.; Johansen, L. E.; Hansen, L. H.; Jespersen, H.; Podtelejnikov, A.; Nielsen, E.; Crawford, J.; Poulsen, V.; Sorensen, B. D.; Matthiesen, J.; Hendrickson, R. C.; Gleeson, F.; Pawson, T.; Moran, M. F.; Durocher, D.; Mann, M.; Hogue, C. W.; Figeys, D.; Tyers, M. *Nature* **2002**, *415*, 180.
- (152) Kruppa, G. H.; Schoeniger, J.; Young, M. M. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 155.
- (153) Guerrero, C.; Tagwerker, C.; Kaiser, P.; Huang, L. *Mol. Cell. Proteomics* **2006**, *5*, 366.
- (154) Gavin, A. C.; Aloy, P.; Grandi, P.; Krause, R.; Boesche, M.; Marzioch, M.; Rau, C.; Jensen, L. J.; Bastuck, S.; Dumpelfeld, B.; Edelmann, A.; Heurtier, M. A.; Hoffman, V.; Hoefert, C.; Klein, K.; Hudak, M.; Michon, A. M.; Schelder, M.; Schirle, M.; Remor, M.; Rudi, T.; Hooper, S.; Bauer, A.; Bouwmeester, T.; Casari, G.; Drewes, G.; Neubauer, G.; Rick, J. M.; Kuster, B.; Bork, P.; Russell, R. B.; Superti-Furga, G. *Nature* **2006**, *440*, 631.
- (155) Bouwmeester, T.; Bauch, A.; Ruffner, H.; Angrand, P. O.; Bergamini, G.; Croughton, K.; Cruciat, C.; Eberhard, D.; Gagneur, J.; Ghidelli, S.; Hopf, C.; Huhse, B.; Mangano, R.; Michon, A. M.; Schirle, M.; Schlegl, J.; Schwab, M.; Stein, M. A.; Bauer, A.; Casari, G.; Drewes, G.; Gavin, A. C.; Jackson, D. B.; Joberty, G.; Neubauer, G.; Rick, J.; Kuster, B.; Superti-Furga, G. *Nat. Cell Biol.* **2004**, *6*, 97.
- (156) Krogan, N. J.; Cagney, G.; Yu, H.; Zhong, G.; Guo, X.; Ignatchenko, A.; Li, J.; Pu, S.; Datta, N.; Tikuisis, A. P.; Punna, T.; Peregrin-Alvarez, J. M.; Shales, M.; Zhang, X.; Davey, M.; Robinson, M. D.; Paccanaro, A.; Bray, J. E.; Sheung, A.; Beattie, B.; Richards, D. P.; Canadien, V.; Lalev, A.; Mena, F.; Wong, P.; Starostine, A.; Canete, M. M.; Vlasblom, J.; Wu, S.; Orsi, C.; Collins, S. R.; Chandran, S.; Haw, R.; Rilstone, J. J.; Gandi, K.; Thompson, N. J.; Musso, G.; St Onge, P.; Ghanny, S.; Lam, M. H.; Butland, G.; Altaf-Ul, A. M.; Kanaya, S.; Shilatifard, A.; O'Shea, E.; Weissman, J. S.; Ingles, C. J.; Hughes, T. R.; Parkinson, J.; Gerstein, M.; Wodak, S. J.; Emili, A.; Greenblatt, J. F. *Nature* **2006**, *440*, 637.
- (157) Huang, L.; Baldwin, M. A.; Maltby, D. A.; Medzihradszky, K. F.; Baker, P. R.; Allen, N.; Rexach, M.; Edmondson, R. D.; Campbell, J.; Juhasz, P.; Martin, S. A.; Vestal, M. L.; Burlingame, A. L. *Mol. Cell. Proteomics* **2002**, *1*, 434.
- (158) Sinz, A. *Mass Spectrom. Rev.* **2006**, *25*, 663.
- (159) Robinette, D.; Neamati, N.; Tomer, K. B.; Borchers, C. H. *Expert Rev. Proteomics* **2006**, *3*, 399.
- (160) Ethier, M.; Lambert, J. P.; Vasilescu, J.; Figeys, D. *Anal. Chim. Acta* **2006**, *564*, 10.
- (161) Tang, X.; Munske, G. R.; Siems, W. F.; Bruce, J. E. *Anal. Chem.* **2005**, *77*, 311.
- (162) Chu, F.; Mahrus, S.; Craik, C. S.; Burlingame, A. L. *J. Am. Chem. Soc.* **2006**, *128*, 10362.
- (163) Seebacher, J.; Mallick, P.; Zhang, N.; Eddes, J. S.; Aebersold, R.; Gelb, M. H. *J. Proteome Res.* **2006**, *5*, 2270.
- (164) El-Shafey, A.; Tolic, N.; Young, M. M.; Sale, K.; Smith, R. D.; Kery, V. *Protein Sci.* **2006**, *15*, 429.
- (165) Nazabal, A.; Wenzel, R. J.; Zenobi, R. *Anal. Chem.* **2006**, *78*, 3562.
- (166) Sinz, A.; Kalkhof, S.; Ihling, C. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1921.
- (167) Petrotchenko, E. V.; Olkhovik, V. K.; Borchers, C. H. *Mol. Cell. Proteomics* **2005**, *4*, 1167.
- (168) Vasilescu, J.; Guo, X.; Kast, J. *Proteomics* **2004**, *4*, 3845.
- (169) Chang, Z.; Kuchar, J.; Hausinger, R. P. *J. Biol. Chem.* **2004**, *279*, 15305.
- (170) Schulz, D. M.; Ihling, C.; Clore, G. M.; Sinz, A. *Biochemistry* **2004**, *43*, 4703.

- (171) Davidson, W.; McGibbon, G. A.; White, P. W.; Yoakim, C.; Hopkins, J. L.; Guse, I.; Hambly, D. M.; Frego, L.; Ogilvie, W. W.; Lavallee, P.; Archambault, J. *Anal. Chem.* **2004**, *76*, 2095.
- (172) Fujii, N.; Jacobsen, R. B.; Wood, N. L.; Schoeniger, J. S.; Guy, R. K. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 427.
- (173) Bernhard, O. K.; Sheil, M. M.; Cunningham, A. L. *Biochemistry* **2004**, *43*, 256.
- (174) Trester-Zedlitz, M.; Kamada, K.; Burley, S. K.; Fenyo, D.; Chait, B. T.; Muir, T. W. *J. Am. Chem. Soc.* **2003**, *125*, 2416.
- (175) Papaioannou, M.; Ruppert, T.; Dotzlaw, H.; Dressel, U.; Baniahmad, A. *Protein Expr. Purif.* **2002**, *26*, 462.
- (176) Taverner, T.; Hall, N. E.; O'Hair, R. A.; Simpson, R. J. *J. Biol. Chem.* **2002**, *277*, 46487.
- (177) Brown, K. C.; Yu, Z.; Burlingame, A. L.; Craik, C. S. *Biochemistry* **1998**, *37*, 4397.
- (178) Winters, M. S.; Day, R. A. *Anal. Biochem.* **2002**, *309*, 48.
- (179) Engen, J. R. *Analyst* **2003**, *240*, 285.
- (180) Komives, E. A. *Int. J. Mass Spectrom.* **2005**, *240*, 285.
- (181) Yao, Z. P.; Tito, P.; Robinson, C. V. *Methods Enzymol.* **2005**, *402*, 389.
- (182) Zhang, Z.; Smith, D. L. *Protein Sci.* **1993**, *2*, 522.
- (183) Pan, J.; Rintala-Dempsey, A. C.; Li, Y.; Shaw, G. S.; Konermann, L. *Biochemistry* **2006**, *45*, 3005.
- (184) Yamada, N.; Suzuki, E.; Hirayama, K. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 293.
- (185) Wintrode, P. L.; Friedrich, K. L.; Vierling, E.; Smith, J. B.; Smith, D. L. *Biochemistry* **2003**, *42*, 10667.
- (186) Yao, Z. P.; Zhou, M.; Kelly, S. E.; Seeliger, M. A.; Robinson, C. V.; Itzhaki, L. S. *J. Mol. Biol.* **2006**, *363*, 673.
- (187) Yamamoto, T.; Izumi, S.; Gekko, K. *FEBS Lett.* **2006**, *580*, 3638.
- (188) Carulla, N.; Caddy, G. L.; Hall, D. R.; Zurdo, J.; Gairi, M.; Feliz, M.; Giralt, E.; Robinson, C. V.; Dobson, C. M. *Nature* **2005**, *436*, 554.
- (189) Horn, J. R.; Kraybill, B.; Petro, E. J.; Coales, S. J.; Morrow, J. A.; Hamuro, Y.; Kossiakoff, A. A. *Biochemistry* **2006**, *45*, 8488.
- (190) Hossain, B. M.; Konermann, L. *Anal. Chem.* **2006**, *78*, 1613.
- (191) Law, D.; Hotchk, M.; Ten Eyck, L. *Proteins* **2005**, *60*, 302.
- (192) Xu, G.; Chance, M. R. *Anal. Chem.* **2004**, *76*, 1213.
- (193) Xu, G.; Takamoto, K.; Chance, M. R. *Anal. Chem.* **2003**, *75*, 6995.
- (194) Takamoto, K.; Chance, M. R. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 251.
- (195) Guan, J. Q.; Chance, M. R. *Trends Biochem. Sci.* **2005**, *30*, 583.
- (196) Demmers, J. A.; Rijkers, D. T.; Haverkamp, J.; Killian, J. A.; Heck, A. J. *J. Am. Chem. Soc.* **2002**, *124*, 11191.
- (197) Liu, R.; Guan, J.-Q.; Zak, O.; Aisen, P.; Chance, M. R. *Biochemistry* **2003**, *42*, 12447.
- (198) Griffith, W. P.; Kaltashov, I. A. *Curr. Org. Chem.* **2006**, *10*, 535.
- (199) Talkington, M. W. T.; Siuzdak, G.; Williamson, J. R. *Nature* **2005**, *438*, 628.
- (200) Schultz, G. A.; Corso, T. N.; Prosser, S. J.; Zhang, S. *Anal. Chem.* **2000**, *72*, 4058.
- (201) Zhang, S.; Van Pelt, C. K.; Wilson, D. B. *Anal. Chem.* **2003**, *75*, 3010.
- (202) Benkestock, K.; Van Pelt, C. K.; Akerud, T.; Sterling, A.; Edlund, P. O.; Roeraade, J. *J. Biomol. Screen.* **2003**, *8*, 247.
- (203) de Vriendt, K.; Sandra, K.; Desmet, T.; Nerinckx, W.; Van Beeumen, J.; Devreese, B. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 3061.
- (204) Smith, A. M.; Jahn, T. R.; Ashcroft, A. E.; Radford, S. E. *J. Mol. Biol.* **2006**, *364*, 9.
- (205) Keetch, C. A.; Hernandez, H.; Sterling, A.; Baumert, M.; Allen, M. H.; Robinson, C. V. *Anal. Chem.* **2003**, *75*, 4937.
- (206) Painter, A. J.; Jaya, N.; Basha, E.; Vierling, E.; Robinson, C. V.; Benesch, J. L. P. **2007**, *In Preparation*.
- (207) van Berkel, G. J.; Zhou, F.; Aronson, J. T. *Int. J. Mass Spectrom. Ion Proc.* **1997**, *162*, 55.

- (208) van Berkel, W. J.; van den Heuvel, R. H.; Versluis, C.; Heck, A. J. *Protein Sci.* **2000**, *9*, 435.
- (209) Wilson, D. J.; Konermann, L. *Anal. Chem.* **2003**, *75*, 6408.
- (210) Simmons, D. A.; Wilson, D. J.; Lajoie, G. A.; Doherty-Kirby, A.; Konermann, L. *Biochemistry* **2004**, *43*, 14792.
- (211) Wilson, D. J.; Rafferty, S. P.; Konermann, L. *Biochemistry* **2005**, *44*, 2276.
- (212) Lentze, N.; Aquilina, J. A.; Lindbauer, M.; Robinson, C. V.; Narberhaus, F. *Eur. J. Biochem.* **2004**, *271*, 2494.
- (213) Nishimura, S.; Nagahori, N.; Takaya, K.; Tachibana, Y.; Miura, N.; Monde, K. *Angew Chem. Int. Ed. Engl.* **2005**, *44*, 571.
- (214) Daniel, J. M.; Friess, S. D.; Rajagopalan, S.; Wendt, S.; Zenobi, R. *Int. J. Mass Spectrom.* **2002**, *216*, 1.
- (215) Liang, H. R.; Foltz, R. L.; Meng, M.; Bennett, P. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 2815.
- (216) Ayed, A.; Krutchinsky, A. N.; Ens, W.; Standing, K. G.; Duckworth, H. W. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 339.
- (217) Fraser, G. W. *Int. J. Mass Spectrom.* **2002**, *215*, 13.
- (218) Gilligan, J. J.; Schuck, P.; Yergey, A. L. *Anal. Chem.* **2002**, *74*, 2041.
- (219) Nelson, R. W.; Nedelkov, D.; Tubbs, K. A. *Electrophoresis* **2000**, *21*, 1155.
- (220) Nedelkov, D.; Nelson, R. W. *Trends Biotechnol.* **2003**, *21*, 301.
- (221) Lopez, F.; Pichereaux, C.; Burlet-Schiltz, O.; Pradayrol, L.; Monsarrat, B.; Esteve, J. P. *Proteomics* **2003**, *3*, 402.
- (222) Boerner, R. J.; Kassel, D. B.; Barker, S. C.; Ellis, B.; DeLacy, P.; Knight, W. B. *Biochemistry* **1996**, *35*, 9519.
- (223) Taoka, M.; Ichimura, T.; Wakamiya-Tsuruta, A.; Kubota, Y.; Araki, T.; Obinata, T.; Isobe, T. *J. Biol. Chem.* **2003**, *278*, 5864.
- (224) Swietnicki, W.; O'Brien, S.; Holman, K.; Cherry, S.; Brueggemann, E.; Tropea, J. E.; Hines, H. B.; Waugh, D. S.; Ulrich, R. G. *J. Biol. Chem.* **2004**, *279*, 38693.
- (225) Nedelkov, D.; Nelson, R. W. *Proteomics* **2001**, *1*, 1441.
- (226) Nedelkov, D.; Nelson, R. W.; Kiernan, U. A.; Niederkofler, E. E.; Tubbs, K. A. *FEBS Lett.* **2003**, *536*, 130.
- (227) Mason, E. R.; McDaniel, E. W. *Transport Properties of Ions in Gases*; J. Wiley and Sons: New York, 1988.
- (228) Carr, T. W. *Plasma Chromatography*; Plenum Press: New York, 1984.
- (229) McLean, J. A.; Ruotolo, B. T.; Gillig, K. J.; Russell, D. H. *Int. J. Mass Spectrom.* **2005**, *240*, 301.
- (230) Wyttenbach, T.; Bowers, M. T. In *Modern Mass Spectrometry*, 2003; Vol. 225.
- (231) Jarrold, M. F. *Annu. Rev. Phys. Chem.* **2000**, *51*, 179.
- (232) Hoaglund-Hyzer, C. S.; Counterman, A. E.; Clemmer, D. E. *Chem. Rev.* **1999**, *99*, 3037.
- (233) Covey, T.; Douglas, D. J. *J. Am. Soc. Mass Spectrom.* **1993**, *4*, 616.
- (234) von Helden, G.; Wyttenbach, T.; Bowers, M. T. *Science* **1995**, *267*, 1483.
- (235) Henderson, S. C.; Valentine, S. J.; Counterman, A. E.; Clemmer, D. E. *Anal. Chem.* **1999**, *71*, 291.
- (236) Dugourd, P.; Hudgins, R. R.; Clemmer, D. E.; Jarrold, M. F. *Rev. Sci. Instrum.* **1997**, *68*, 1122.
- (237) Wu, C.; Siems, W. F.; Asbury, G. R.; Hill, H. H. *Anal. Chem.* **1998**, *70*, 4929.
- (238) Hoaglund, C. S.; Valentine, S. J.; Clemmer, D. E. *Anal. Chem.* **1997**, *69*, 4156.
- (239) Wyttenbach, T.; Kemper, P. R.; Bowers, M. T. *Int. J. Mass Spectrom.* **2001**, *212*, 13.
- (240) Gillig, K. J.; Ruotolo, B.; Stone, E. G.; Russell, D. H.; Fuhrer, K.; Gonin, M.; Schultz, A. J. *Anal. Chem.* **2000**, *72*, 3965.
- (241) Gillig, K. J.; Ruotolo, B. T.; Stone, E. G.; Russell, D. H. *Int. J. Mass Spectrom.* **2004**, *239*, 43.
- (242) Valentine, S. J.; Koeniger, S. L.; Clemmer, D. E. *Anal. Chem.* **2003**, *75*, 6202.

- (243) Javahery, G.; Thomson, B. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 697.
- (244) Guo, Y. Z.; Wang, J. X.; Javahery, G.; Thomson, B. A.; Siu, K. W. M. *Anal. Chem.* **2005**, *77*, 266.
- (245) Tang, K.; Shvartsburg, A. A.; Lee, H. N.; Prior, D. C.; Buschbach, M. A.; Li, F. M.; Tolmachev, A. V.; Anderson, G. A.; Smith, R. D. *Anal. Chem.* **2005**, *77*, 3330.
- (246) Pringle, S. D.; Giles, K.; Wildgoose, J. L.; Williams, J. P.; Slade, S. E.; Thalassinou, K.; Bateman, R. H.; Bowers, M. T.; Scrivens, J. H. *Int. J. Mass Spectrom.* **2007**, *261*, 1.
- (247) Bacher, G.; Szymanski, W. W.; Kaufman, S. L.; Zollner, P.; Blaas, D.; Allmaier, G. *J. Mass Spectrom.* **2001**, *36*, 1038.
- (248) Kaufman, S. L.; Skogen, J. W.; Dorman, F. D.; Zarrin, F.; Lewis, K. C. *Anal. Chem.* **1996**, *68*, 3703.
- (249) Kaufman, S. L.; Skogen, J. W.; Dorman, F. D.; Zarrin, F.; Lewis, K. C. *Anal. Chem.* **1996**, *68*, 1895.
- (250) Ude, S.; de la Mora, J. F.; Thomson, B. A. *J. Am. Chem. Soc.* **2004**, *126*, 12184.
- (251) Guevremont, R. *J. Chrom. A* **2004**, *1058*, 3.
- (252) Shvartsburg, A. A.; Tang, K. Q.; Smith, R. D. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1447.
- (253) Robinson, E. W.; Williams, E. R. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1427.
- (254) Robinson, E. W.; Leib, R. D.; Williams, E. R. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1469.
- (255) Shvartsburg, A. A.; Li, F. M.; Tang, K. Q.; Smith, R. D. *Anal. Chem.* **2006**, *78*, 3706.
- (256) Shvartsburg, A. A.; Li, F. M.; Tang, K. Q.; Smith, R. D. *Anal. Chem.* **2006**, *78*, 3304.
- (257) Shvartsburg, A. A.; Mashkevich, S. V.; Smith, R. D. *J. Phys. Chem. A* **2006**, *110*, 2663.
- (258) Shvartsburg, A. A.; Li, F.; Tang, K.; Smith, R. D. *Anal. Chem.* **2007**, *79*, 1523.
- (259) Koeniger, S. L.; Merenbloom, S. I.; Valentine, S. J.; Jarrold, M. F.; Udseth, H. R.; Smith, R. D.; Clemmer, D. E. *Anal. Chem.* **2006**, *78*, 4161.
- (260) Merenbloom, S. I.; Koeniger, S. L.; Valentine, S. J.; Plasencia, M. D.; Clemmer, D. E. *Anal. Chem.* **2006**, *78*, 2802.
- (261) Tang, K. Q.; Li, F. M.; Shvartsburg, A. A.; Strittmatter, E. F.; Smith, R. D. *Anal. Chem.* **2005**, *77*, 6381.
- (262) Laiko, V. V. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 500.
- (263) Loboda, A. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 691.
- (264) Teplow, D. B.; Lazo, N. D.; Bitan, G.; Bernstein, S.; Wyttenbach, T.; Bowers, M. T.; Baumketner, A.; Shea, J. E.; Urbanc, B.; Cruz, L.; Borreguero, J.; Stanley, H. E. *Accounts Chem. Res.* **2006**, *39*, 635.
- (265) Bernstein, S. L.; Wyttenbach, T.; Baumketner, A.; Shea, J. E.; Bitan, G.; Teplow, D. B.; Bowers, M. T. *J. Am. Chem. Soc.* **2005**, *127*, 2075.
- (266) Bernstein, S.; Dupuis, N. F.; Lazo, N.; Wyttenbach, T.; Condron, M.; Bitan, G.; Teplow, D. B.; Shea, J. E.; Ruotolo, B. T.; Robinson, C. V.; Bowers, M. T. *In Preparation*.
- (267) Bernstein, S. L.; Liu, D. F.; Wyttenbach, T.; Bowers, M. T.; Lee, J. C.; Gray, H. B.; Winkler, J. R. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 1435.
- (268) Loo, J. A.; Berhane, B.; Kaddis, C. S.; Wooding, K. M.; Xie, Y.; Kaufman, S. L.; Chernushevich, I. V. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 998.
- (269) Taverner, T.; Hernandez, H.; Sharon, M.; Ruotolo, B. T.; Matak-Vinkovic, D.; D., D.; Russell, R. B.; Robinson, C. V. *In Preparation*.
- (270) Siuzdak, G.; Bothner, B.; Yeager, M.; Brugidou, C.; Fauquet, C. M.; Hoey, K.; Chang, C. M. *Chem. Biol.* **1996**, *3*, 45.
- (271) Bothner, B.; Siuzdak, G. *Chembiochem* **2004**, *5*, 258.
- (272) Ouyang, Z.; Takats, Z.; Blake, T. A.; Gologan, B.; Guymon, A. J.; Wiseman, J. M.; Oliver, J. C.; Davisson, V. J.; Cooks, R. G. *Science* **2003**, *301*, 1351.
- (273) Simmons, D. A.; Ruotolo, B. T.; Benesch, J. L.; Robinson, C. V. 54th meeting of the American Society for Mass Spectrometry, Seattle, WA, USA, 2006.

FIGURE 1

Pyramid of protein organization states

There are various levels of protein organization which can be probed in proteomic analyses by means of MS. A protein complex may be comprised of several sub-complexes, which themselves are composed of individual protein chains. A further level, that of peptide segments which are generated by proteolytic digestion, is that which most proteomics approaches interrogate. Combining investigations into all four levels of organization however can allow for a far more extensive and detailed characterization of proteins and their functional states. In this article we discuss methods and strategies for studying the upper two states of this pyramid, those wherein non-covalent interactions are maintained.

FIGURE 2

Conventional and nano electrospray MS of a protein-protein complex.

MS of the GroEL complex ionized by means of conventional (lower) and nano (upper) ESI. The nESI spectrum displays a series of peaks between around 11500 m/z which correspond to the 800kDa tetradecamer. Conventional ESI of the same solution results in poorly resolved ‘humps’ centred on 12500, 16000 and 18500 m/z. These are assigned to the tetradecamer, a dimer of tetradecamers, and a trimer of tetradecamers respectively. There is also signal at low m/z which corresponds to GroEL monomer. Though signal is observed corresponding to intact tetradecamer, the signal differs to that observed by nESI in three main ways: the peaks are less well resolved, the charge state distribution is broader and bimodal, and the formation of non-specific oligomers is increased. These results indicate some of the benefits of using nESI. A smaller initial droplet size leads to less nonspecific aggregation (both protein-protein and protein-salt), and the gentler interface conditions possible while still allowing adequate desolvation lead to less dissociation and disruption of oligomeric structure. Solution conditions were 200 mM ammonium acetate, pH 6.9, and a protein concentration of 2 μ M tetradecamer. Spectra were obtained on a modified Q-ToF 2 (Waters/Micromass)¹⁰⁵.

FIGURE 3

Nonspecific aggregation during ion formation by nESI

As the initial droplets formed during nESI undergo asymmetric fission, offspring droplets containing none, one, or several of the molecules of interest are formed (A). These droplets are on the order of 18nm⁵⁰. A large proportion of the droplets formed are vacant, but at higher concentrations more occupied droplets are formed (B). Moreover, the relative proportions of the droplets which contained none, one, two and three molecules vary according to the concentration. Those droplets containing multiple copies of proteins, will, via the CRM, subsequently lead to the formation of non-specific aggregation. These simulations were performed by using Monte-Carlo methods. A certain number of particles (variable, depending on the concentration set) were put randomly into a volume, and the number counted in a volume corresponding to an 18 nm diameter droplet. This procedure was repeated 100000 times, and averages were taken. For this simple simulation, excluded volume and solvent evaporation effects were not taken into account.

FIGURE 4

Adduction of solvent molecules and buffer ions to proteins

(A) The theoretical deconvoluted spectrum of a pure protein complex would appear at a mass value governed solely by primary sequence, with a peak width defined by the isotopic distribution and instrumental resolution ('sequence', pink). The masses of protein complexes invariably are higher than those due to the sequence alone, which is attributed to the retention of solvent molecules and buffer ions. The higher the accelerating voltages ('low', blue; 'medium', indigo; 'high', violet) and consequently the more energetic the collisions, the smaller this shift becomes, and there is a concomitant narrowing of the distribution. (B) Examination of the 68+ charge state of GroEL (see Fig 2) at different activation voltages demonstrates that the amount of additional mass and the width of the distribution can be approximated by a linear relationship⁸⁶. The amount of accelerating voltage (low to high) required to achieve the indicated peak width is indicated by the color of the points in Panel B.

FIGURE 5

A quadrupole time-of-flight type instrument customized for the transmission and analysis of protein complexes

(A) Ions are introduced into the spectrometer by nESI. The ion beam is focused by RF-only ion guides, before entering a quadrupole analyzer. This analyzer is modified such that it operates at a reduced RF frequency, allowing the selection and transmission of high m/z ions. Ions then traverse a collision cell, into which gas can be leaked in order to allow collision-induced dissociation. The dissociation of protein complex ions requires higher collision cell gas pressures and accelerating voltages relative to ‘normal’ parameters. Ions exiting the collision cell enter a pusher region, whereupon they are pulsed orthogonally in packets into a ToF analyser. Ions are detected by means of an MCP detector, which in some instruments is split into four segments along the ion axis.

The two primary areas for protein-complex ion loss, once they have entered the vacuum stages, are in the RF-only ion guides and the apertures between them, and by over- or under-shooting the ToF detector. Improved transmission is achieved by adjusting the velocity of the ions, via altering the pressures in the instrument. Simulation of the ion trajectories of a 147 kDa protein complex (ADH tetramer) at different pressures (B) shows how, at the standard pressure of 9.3 μbar , the ions are poorly focused. Increasing the pressure to 40 μbar results in a much narrower ion beam and consequently improved transmission. This amelioration is reflected if one monitors the ion intensity across the MCP detector (C). As the pressure is increased from 9.3 μbar the total signal intensity (width of the bars) increases. Once the pressure exceeds ~ 40 μbar however, the signal intensity starts to decrease again. Examining how the ion signal divides between the four segments of the MCP plate shows how at low pressures the ions mainly impact on the anodes furthest from the pusher (3 and 4), whereas at the highest pressures they primarily impact anodes 1 and 2. This suggests that at pressures in the source region below and above the optimum, ions over- and undershoot the detector, respectively. Simulations in (A) were performed as described previously⁹², assuming an initial kinetic energy of 122 eV, a parallel gas flow of 50 m s^{-1} and a collision cross section of 7000 \AA^2 for ADH (based on²⁴⁷). Data in (B) were recorded for the 25+ charge state of ADH, at 200 mM ammonium acetate, pH 6.9 and 3 μM tetramer, on an MDS Q-Star XL instrument with a 4-channel MCP detector.

FIGURE 6

Number of collisions experienced and time spent in collision cell

The upper panel shows the calculated number of collisions experienced, plotted relative to mass, by cytochrome *c* monomer (violet), transthyretin tetramer (blue), *Mj*HSP16.5 24mer (green), GroEL tetradecamer (orange), and the 70S ribosome from *Thermus thermophilus* (red) (see Equation 2). The inset shows the linear dependence of the number of collisions on pressure for *Mj*HSP16.5. The time spent in the collision cell undergoing these collisions is plotted in the lower panel (using Equations 2 and 4). The inset shows the accelerating voltage and pressure dependency of this time (for *Mj*HSP16.5, 47⁺ charge state). These calculations are based on a collision cell length of 18.5 cm, a gas pressure of 30 μbar argon, and an accelerating voltage of 200 V.

FIGURE 7

Energy conversion during collisional activation

Simulations showing the maximum percentage conversion of energy from kinetic to internal modes as ions pass through the gas filled collision cell, and the variations due to mass (upper), gas pressure (middle), and collision gas (lower). The upper panel shows the energy conversion profiles for cytochrome *c* monomer (violet), transthyretin tetramer (blue), *Mj*HSP16.5 24mer (green), GroEL tetradecamer (orange), and the 70S ribosome from *Thermus thermophilus* (red), at 30 μbar argon collision cell pressure. In order to determine a true dependency on mass, Equation 5 was used. The distance at which half of the energy is converted (dashed line) for the different species allows the extraction of the dependency of energy conversion on mass (inset). The middle panel shows the effect of varying pressure on this energy conversion process for a single species, *Mj*HSP16.5. Different pressures (labeled from 50 μbar, violet, to 10 μbar, red) of argon are simulated. Similarly, the lower panel shows (using the same ion as above) the effect of the mass of the gas (the noble gases from radon, violet, to neon, red). Inset into these panels are the dependencies for 50% conversion. These simulations show that in order for sufficient conversion of kinetic energy into internal modes to occur, in this length of collision cell, the use of higher pressures and/or heavier target gas is preferable.

FIGURE 8

Dissociation pathway of a multi-protein complex

(A) CID of the 32+ charge state of the *TaHSP16.9* dodecamer (violet), results in the formation of complementary monomers at low m/z and 11mers at high m/z (both blue). At higher accelerating voltages (>100 V) a second distribution of monomers as well as decamers are observed (both cyan). This is indicative of a sequential dissociation reaction (white arrows). Plotting the relative abundance of the different oligomeric species as a function of accelerating voltage, (B), further emphasizes this reaction sequence and, moreover, allows the determination of dissociation thresholds (set at relative intensity value of 10%). These values are then used in simulations as in Fig 6, demonstrating that at the lowest voltage that still incurs CID (80 V) the entire length of the collision cell is used and that activation takes approximately 342 μ s. At the highest accelerating voltage used in this experiment, 200 V, the time taken to reach the dissociation threshold is 11 μ s. This demonstrates that the activation time for such a molecule is on the microsecond timescale, but can vary by more than an order of magnitude. Experiments were performed on a modified Q-ToF 2¹⁰⁵. The solution of *TaHSP16.9* was infused by nESI at a concentration of 1.4 μ M in 200 mM ammonium acetate, pH6.9¹⁴.

FIGURE 9

Applications of collisional activation to the study of protein complexes

When a protein complex is collisionally activated, both the reaction products and pathway allow conclusions to be drawn about its structure. Information as to oligomeric composition (green) can be obtained from the identity of dissociation products (purple), as well as from beneficial effects regarding the parent ion (pink). A detailed examination of the reaction pathway can reveal certain parameters (blue) which can allow the determination of interaction strengths (red) and information as to oligomeric organization and size (orange).

FIGURE 10

Technologies associated with MS for the study of protein complexes

Chart demonstrating the benefits of some of the numerous technologies which have been coupled to MS for in the field of protein complexes. Here they are classified in their order relative to the MS experiment, namely under sample isolation, sample labeling, and online pre- and post-ionization categories.

FIGURE 11

Reactions monitored in real time by means of MS

Reactions can be monitored in a time-resolved manner in several ways by MS. A reaction can be sampled at various time points (A), or sampled continuously (B). Without recourse to automation, the former approach is only possible if the reaction occurs over a sufficiently long-time course. The latter has the disadvantage of occurring over an extended period of time in the ESI capillary, and hence electrochemical changes in the solution can affect the reaction. An automated ESI approach however allows much more rapid spray-to-spray turnover such that a fast reaction may be monitored through repeated sampling (C), as well as the ability to perform the experiment in parallel (D), and to create individual data files for each sample delivery (E).

FIGURE 12

Different approaches for achieving ion mobility separation

(A) The principle of ion mobility spectrometry lies in separating ions according to their ability to traverse a region of relatively high gas pressure. Ions which experience more collisions with the background gas, those with a higher collisional cross-section, reach the detector later than ions of the same mass with a lower collisional cross-section. This effectively allows the separation of ions of the same mass but with different sizes. There are four major approaches towards effecting ion mobility separation (B). Drift-tube separation (I) works by accelerating the ions in a background of neutral gas molecules. The resistance provided by this gas effects a separation proportional to the cross-section of the ion. A similar approach to this, the counter-flow method (II) separates ions by propelling the ions against a stopping potential by an opposing flow of gas molecules. Two other commonly used methods employ a gas flow and electric potential orthogonal to each other (III and IV). The differential mobility analyzer (III) uses a single-function field, whereas FAIMS uses a multi-element field gradient (IV).

FIGURE 1

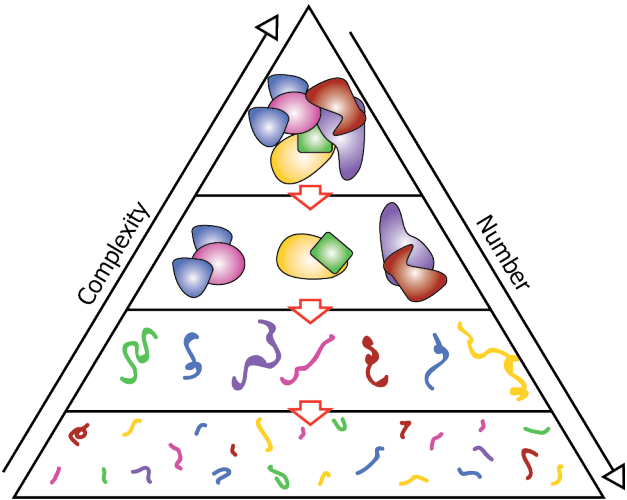


FIGURE 2

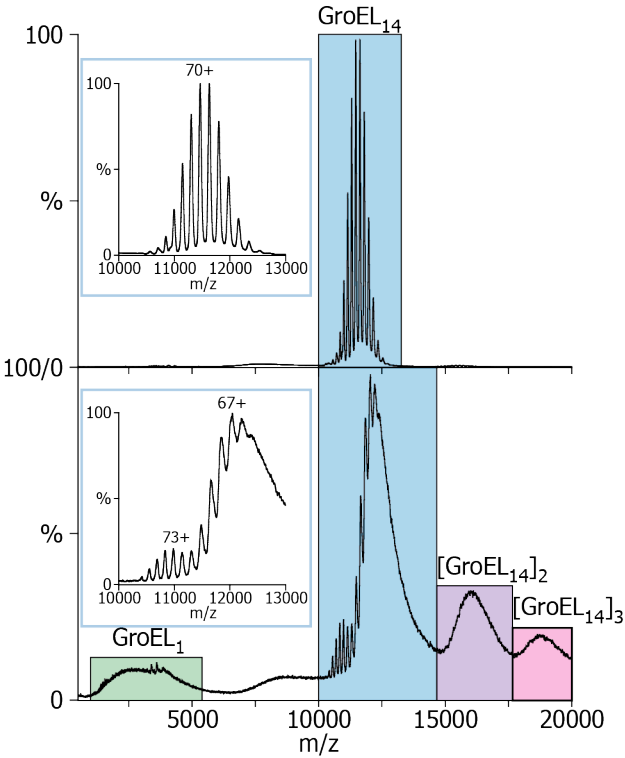


FIGURE 3

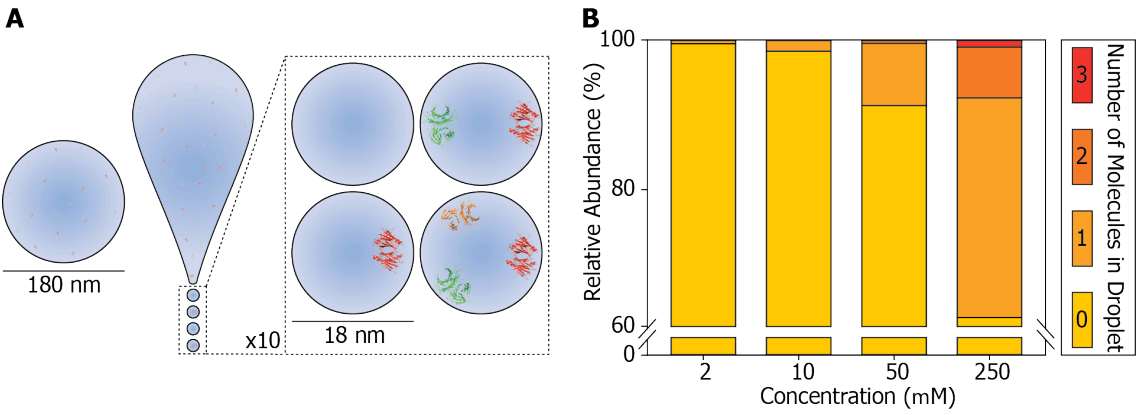


FIGURE 4

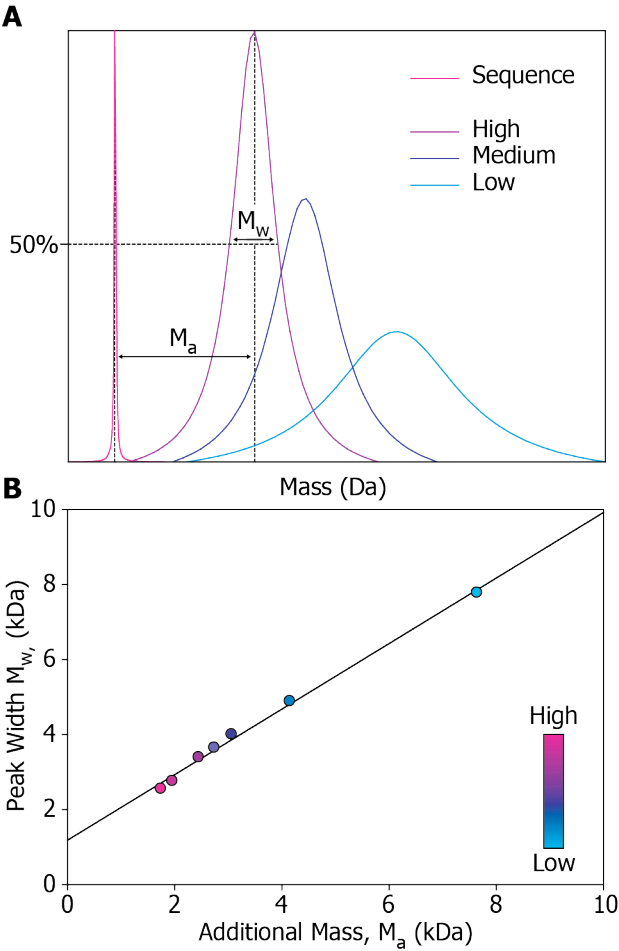


FIGURE 5

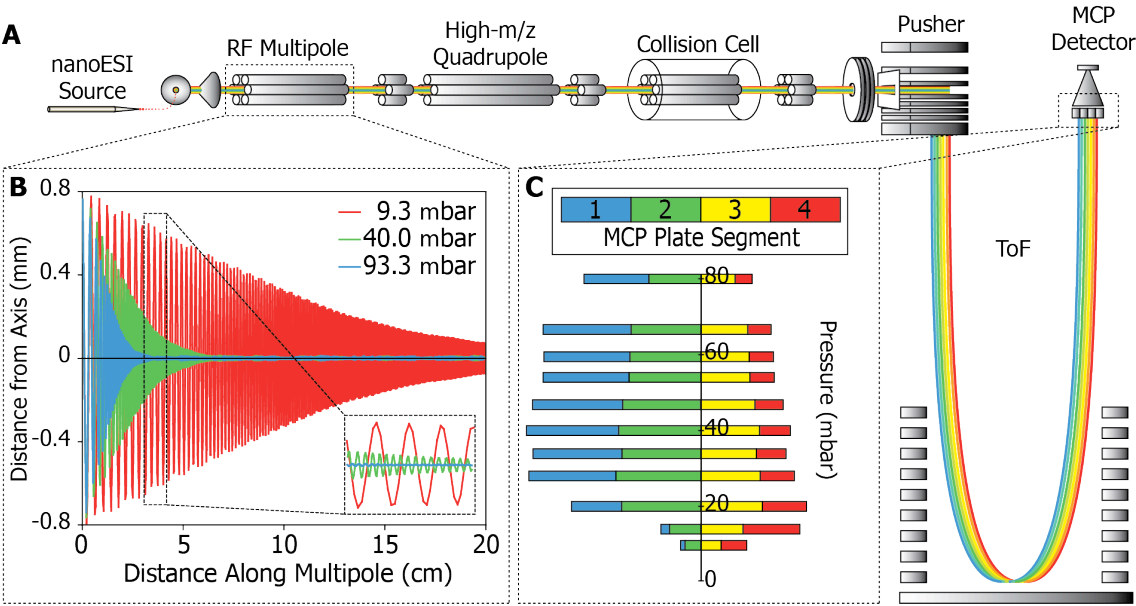


FIGURE 6

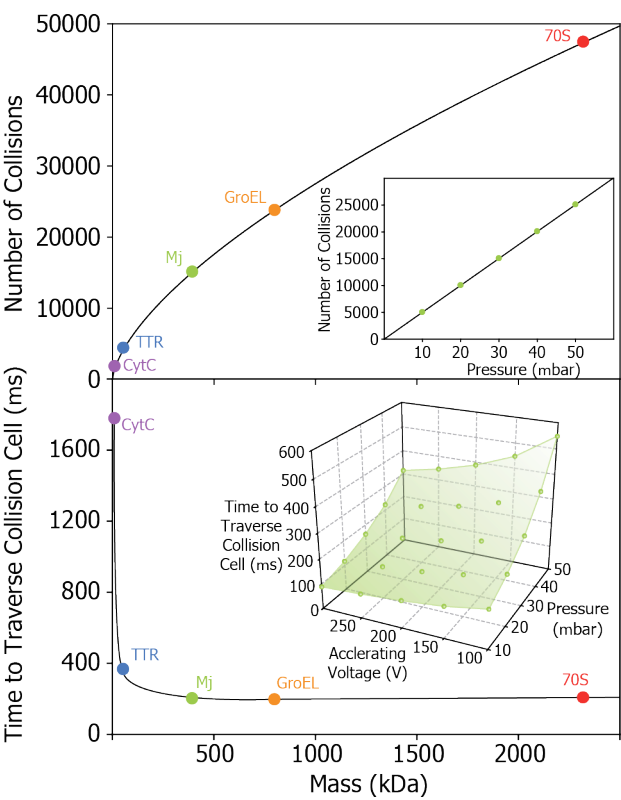


FIGURE 7

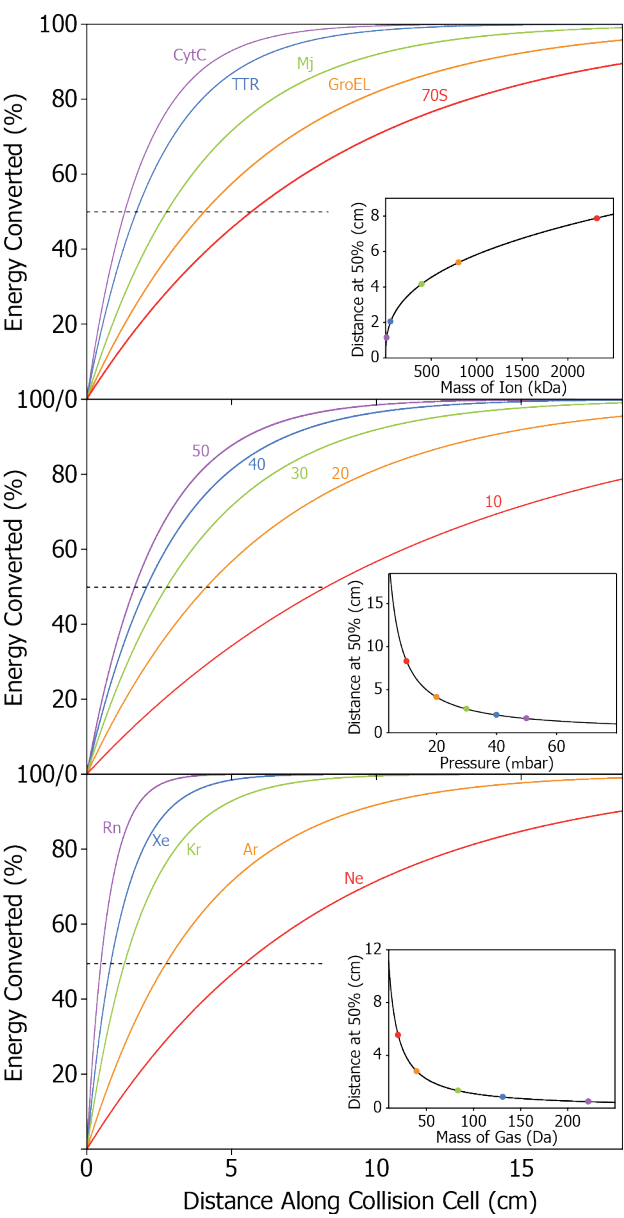


FIGURE 8

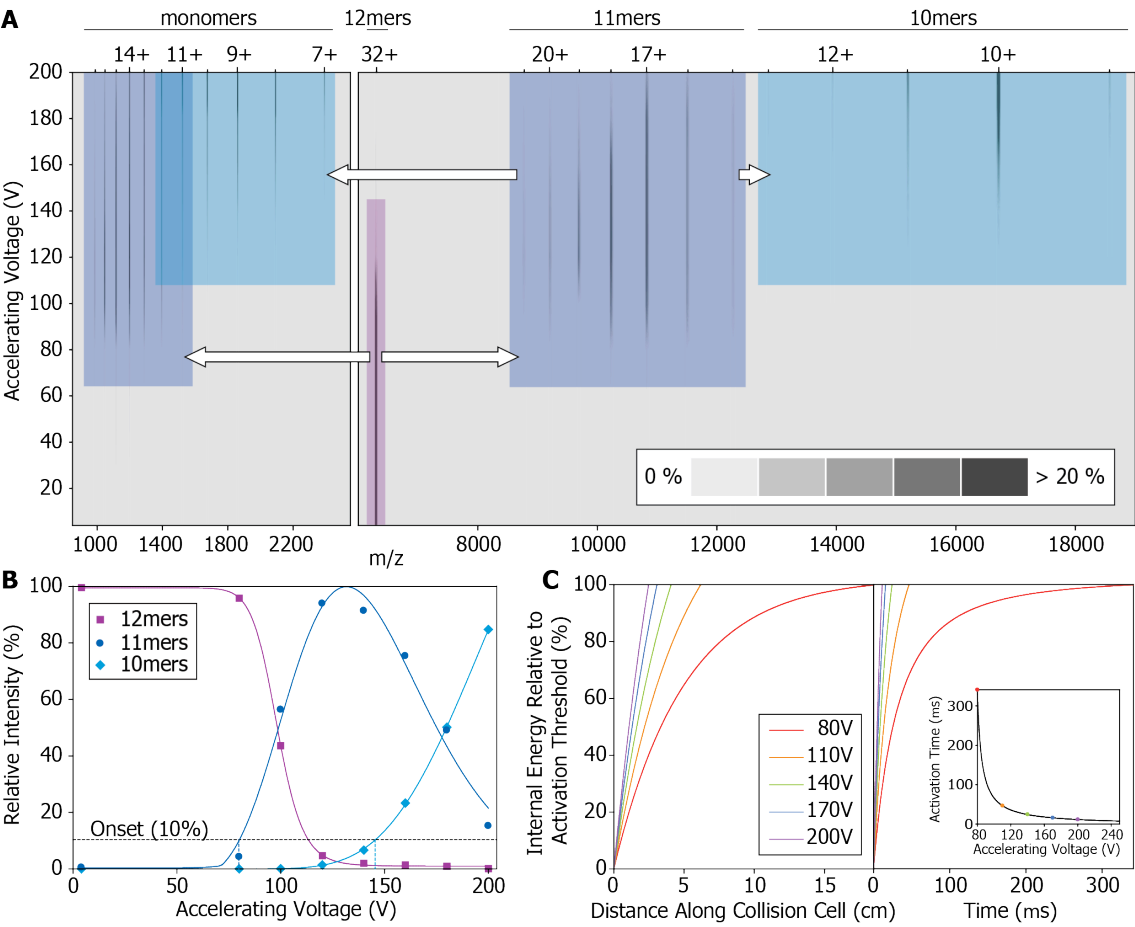


FIGURE 9

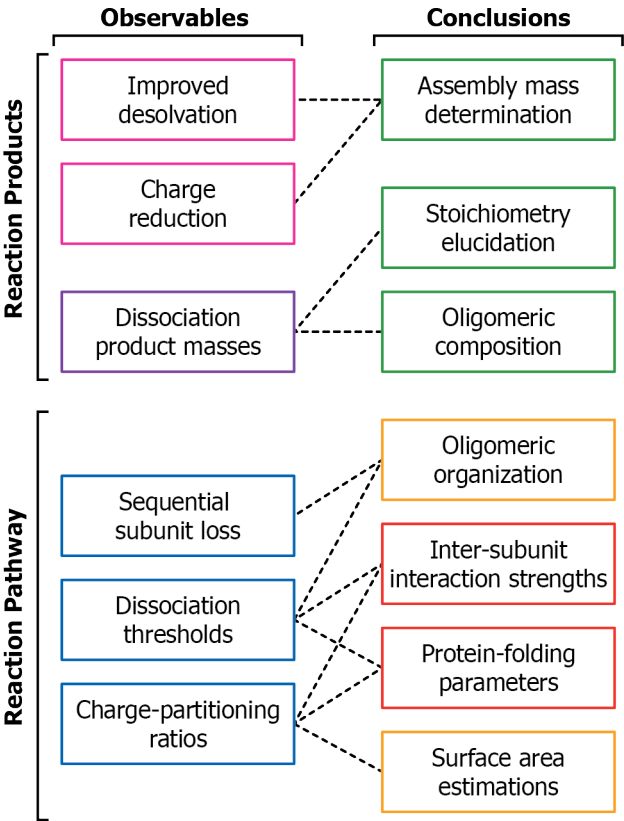


FIGURE 10

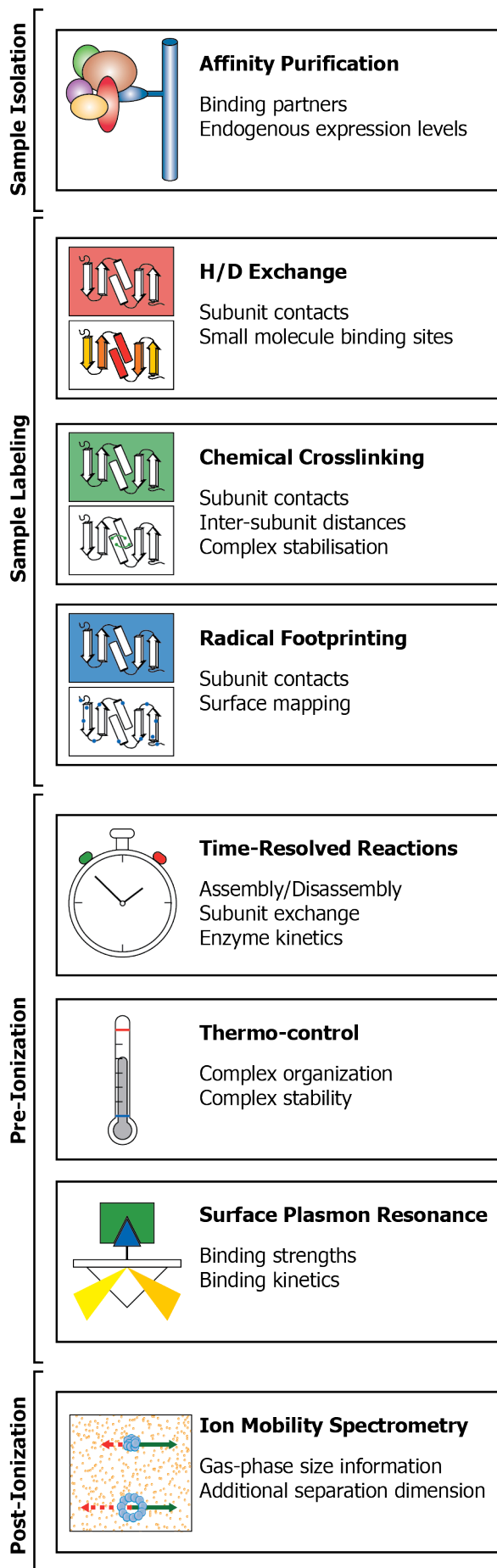


FIGURE 11

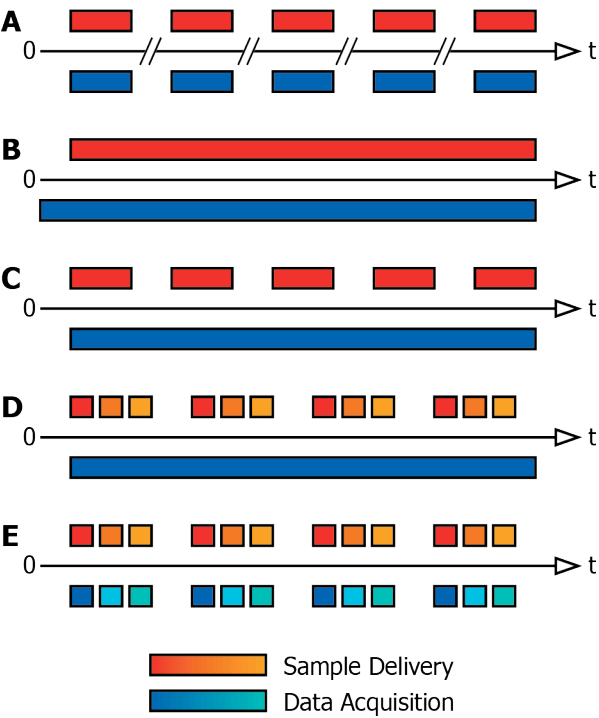


FIGURE 12

