INSIGHTS INTO VIRUS CAPSID ASSEMBLY FROM NON-COVALENT MASS SPECTROMETRY

Victoria L. Morton, Peter G. Stockley, Nicola J. Stonehouse, and Alison E. Ashcroft*

Astbury Centre for Structural Molecular Biology,
Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, UK

Received 23 January 2008; received (revised) 7 March 2008; accepted 7 March 2008

Published online in 22 May 2008 Wiley InterScience (www.interscience.wiley.com) DOI 10.1002/mas.20176

The assembly of viral proteins into a range of macromolecular complexes of strictly defined architecture is one of Nature's wonders. Unraveling the details of these complex structures and the associated self-assembly pathways that lead to their efficient and precise construction will play an important role in the development of anti-viral therapeutics. It will also be important in bio-nanotechnology where there is a plethora of applications for such well-defined macromolecular complexes, including cell-specific drug delivery and as substrates for the formation of novel materials with unique electrical and magnetic properties. Mass spectrometry has the ability not only to measure masses accurately but also to provide vital details regarding the composition and stoichiometry of intact, non-covalently bound macromolecular complexes under near-physiological conditions. It is thus ideal for exploring the assembly and function of viruses. Over the past decade or so, significant advances have been made in this field, and these advances are summarized in this review, which covers the literature up to the end of 2007. © 2008 Wiley Periodicals, Inc., Mass Spec Rev 27:575–595, 2008

Keywords: virus assembly; mass spectrometry; non-covalent interactions

I. INTRODUCTION

Viruses are pathogens found in every kingdom of life. They are incapable of independent replication, and therefore they have common aspects within their lifecycles. These include entry into host cells, subsequent, at least partial, disassembly to allow for replication of their genetic material, followed by assembly of progeny virus particles, and finally escape of these progeny from the host cell. Despite the similarity of these overall processes, there is a vast range of distinct molecular mechanisms that underlie the specifics of each step that are virus and host-dependent. In composition, viruses are much simpler than their

hosts and exploit the cellular machinery of those hosts to complete their lifecycles. The study of viruses has, therefore, been fundamental to our understanding of modern cell biology, as well as being vital for the design of therapeutic or preventative strategies against viral diseases.

Viral particles are essentially macromolecular aggregates, comprising in the simplest cases only a protective protein shell, known as a capsid, and the enclosed genetic material. The latter can be either DNA or RNA, which in turn can be either single-stranded or double-stranded, and present as single or multiple molecules. This basic structural description, however, can be misleading because viral lifecycles can only be completed by a co-ordinated series of highly dynamic molecular events. Viruses are more accurately described as nanoscale macromolecular machines that have evolved in the highly unforgiving Darwinian world where the better machines flourish whereas those that perform less well “die out.” Thus, the viruses we see around us are the products of extensive molecular refinement, containing some of the most exquisite molecular machinery known. They are also, despite their scale and total dependence on their hosts, the most successful organisms on the planet in terms of their numbers. Bacteriophages, viruses that grow in bacteria, are the most successful of all viruses, due to the fact that their hosts include cyanobacteria and algae, which exist in coastal sea water at a level of many millions per milliliter and each host cell can be used to generate tens of thousands of phage particles (Hendrix et al., 1999).

Simple viruses, namely those that consist of just a shell of coat proteins that protect a nucleic acid molecule, have played important roles in the development of our modern understanding of biology. Indeed, it was Crick and Watson (1956) who realized that there was a basic problem to be overcome by virus particles. Each amino acid of a coat protein subunit must be encoded by a three nucleotide codon in a messenger RNA (mRNA). Since the largest amino acid, tryptophan, has a molecular weight of 204 Da whereas that of the codon specifying it is ~1,000 Da, it is immediately clear that it is impossible to construct a shell from a single giant polypeptide that can enclose the genome that encoded it. Instead, it was proposed that capsid shells were composed of multiple copies of one or very few coat proteins; that structure implied they would be highly symmetrical.

At roughly the same time it was shown that certain spherical viruses that crystallize, such as the plant virus tomato bushy stunt virus (TBSV), have X-ray diffraction patterns consistent with icosahedral symmetry (Caspar, 1956; Table 1). The roles of symmetry in viral architecture were explored systematically with

Contract grant sponsor: BBSRC; Contract grant sponsor: Leverhulme Trust; Contract grant sponsor: Wellcome Trust.

*Correspondence to: Alison E. Ashcroft, Astbury Centre for Structural Molecular Biology, Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, UK.

E-mail: a.e.ashcroft@leeds.ac.uk

TABLE 1. Summary of intact viruses and assembly intermediates studied by mass spectrometry

Virus	Structure*	<i>T</i>	Genome	Family	Citation
MS2		3	linear ssRNA	<i>Leviviridae</i>	(Golmohammadi, Valegard, Fridborg, et al., 1993, Shepherd, Borelli, Lander, et al., 2006, Stockley, Rolfsson, Thompson, et al., 2007, Tito, Tars, Valegard, et al., 2000) 2ms2
TMV(tobacco mosaic virus)		<i>n/a</i>	linear ssRNS	<i>Tobamovirus</i>	Fuerstenau, Benner, Thomas, et al., 2001)
RYMV (rice yellow mottle virus)		3	linear ssRNA	<i>Sobemovirus</i>	(Fuerstenau, Benner, Thomas, et al., 2001, Qu, Liljas, Opalka, et al., 2000, Shepherd, Borelli, Lander, et al., 2006, Thomas, Bothner, Traina, et al., 2004) 1f2n
Adenovirus		1	linear dsDNA	<i>Adenoviridae</i>	(Nie, Tzeng, Chang, et al., 2006, Thomas, Bothner, Traina, et al., 2004, Zubieta, Chroboczek and Cusack, 2005) 1x9p
CPMV (cowpea mosaic virus)		pT3	segmented linear ssRNA	<i>Comoviridae</i>	(Lin, Chen, Usha, et al., 1999, Shepherd, Borelli, Lander, et al., 2006,

(Continued)

TABLE 1. (Continued)

					Thomas, Bothner, Traina, et al., 2004, Wang, Kaltgrad, Lin, et al., 2002) 1ny7
Rhinovirus (HRV2)		pT3	linear ssRNA	<i>Picornaviridae</i>	(Bacher, Szymanski, Kaufmann, et al., 2001, Shepherd, Borelli, Lander, et al., 2006, Verdaguer, Blaas and Fita, 2000) 1fpn
T2		13	linear dsDNA	<i>Myoviridae</i>	(Hogan, Kettleson, Ramaswami, et al., 2006)
T4		13	linear dsDNA	<i>Myoviridae</i>	(Forkine, Chipman, Leiman, et al., 2004, Hogan, Kettleson, Ramaswami, et al., 2006)
BMV (brome mosaic virus)		3	segmented ssRNA	<i>Bromoviridae</i>	(Berthet-Colominas, Cuillel, Kock, et al., 1987, Lucas, Larson and McPherson, 2002, Shepherd, Borelli, Lander, et al., 2006, Wang, Lane and Smith, 2001) 1js9
CCMV (Cowpea chlorotic mottle virus)		3	segmented ssRNA	<i>Bromoviridae</i>	(Kaddis, Lomeli, Yin, et al., 2007, Shepherd, Borelli, Lander, et al., 2006, Speir, Munshi, Wang, et al., 1995) 1cwp

(Continued)

TABLE 1. (Continued)

HIV-1 (human immunodeficiency virus-1)			ssRNA-RT	<i>Retroviridae</i>	(Bardi, 2007, Ganser-Pornillos, Cheng and Yeager, 2007, Lanman, Lam, Emmett, et al., 2004, Sundquist and Hill, 2007)
Hong Kong 97 virus		7	linear dsDNA	<i>Siphoviridae</i>	(Helgstrand, Wikoff, Duda, et al., 2003, Noorden, 2007, Shepherd, Borelli, Lander, et al., 2006) lohg
HBV (Hepatitis B virus)		4	circular partially dsDNA	<i>Hepadnaviridae</i>	(Lorenzen, Uetrecht, Versluis, et al., 2007, Shepherd, Borelli, Lander, et al., 2006, Wynne, Crowther and Leslie, 1999) lqgt
Bacteriophage λ		7	linear dsDNA	<i>Siphoviridae</i>	(Dokland and Murialdo, 1993, Hogan, Kettleson, Ramaswami, et al., 2006)
FHV (Flock house virus)		3	segmented linear ssRNA	<i>Nodaviridae</i>	(Broo, Wei, Marshall, et al., 2001, Shepherd, Borelli, Lander, et al., 2006)
Foot and mouth virus		1, pseudo 3	linear ssRNA	<i>Picornaviridae</i>	(Broo, Wei, Marshall, et al., 2001, Fry, Acharya and Stuart, 1993, Shepherd, Borelli, Lander, et al., 2006) lbbt
Bacteriophage Q β		3	linear ssRNA	<i>Leviviridae</i>	(Ashcroft, 2005, Golmohammadi, Fridborg, Bundule, et al., 1996, Shepherd,

(Continued)

TABLE 1. (Continued)

					Borelli, Lander, et al., 2006) lqbe
VEE-TRD(Venezuelan equine encephalitis Trinidad Donkey virus)		4	linear ssRNA	<i>Togaviridae</i>	(Paredes, Alwell-Warda, Weaver, et al., 2003, Thomas, Falk and Fenselau, 1998)
VV(Vaccinia virus)		n/a	linear dsDNA	<i>Poxviridae</i>	(Cyrklaff, Risco, Fernandez, et al., 2005, Nie, Tzeng, Chang, et al., 2006)
GIV(Grouper iridovirus)			linear dsDNA	<i>Iridoviridae</i>	(Nie, Tzeng, Chang, et al., 2006)
P22		7	linear dsDNA	<i>Podoviridae</i>	(Kang, Hawkrige, Johnson, et al., 2006, Lander, Tang, Casjens, et al., 2006, Tuma, Coward, Kirk, et al., 2001)
P22 portal		n/a	n/a		(Lander, Tang, Casjens, et al., 2006, Poliakov, van Duijn, Lander, et al., 2007)
φ 29 portal		n/a	n/a	<i>Podoviridae</i>	(Poliakov, van Duijn, Lander, et al., 2007, Simpson, Tao, Leiman, et al., 2000)
SPP1 portal		n/a	n/a	<i>Siphoviridae</i>	(Orlova, Dube, Backmann, et al., 1999, Poliakov, van Duijn, Lander, et al., 2007)

(Continued)

TABLE 1. (Continued)

Φ8/ Φ12 viral packaging motor, helicase P4		n/a	n/a	<i>Cystoviridae</i>	(Lisal, Kainov, Lam, et al., 2006, Lisal, Lam, Kainov, et al., 2005, Mancini, Kainov, Grimes, et al., 2004) 1w4c
--	---	-----	-----	---------------------	--

* All structures are taken from Shepherd et al. (2006) except for the following: P22 phage and portal (Lander et al., 2006), reprinted with full permission from AAAS; SPPI portal reprinted with permission from the Nature Publishing Group from Orlova et al. (1999); VEE-TRD reprinted with permission from the American Society for Microbiology from Paredes et al. (2003); T4 reprinted with permission from Forkine et al. (2004), copyright (2004) National Academy of Sciences, USA; VV reprinted with permission from Cyrklaff et al. (2005), copyright (2005) National Academy of Sciences, USA; TMV Grasp image courtesy of Dr. J.Y Sgro, UW-Madison, USA (virology.wisc.edu/virusworld). [Color table can be viewed in the online issue, which is available at www.interscience.wiley.com.]

the result that the inherent efficiency of helical and icosahedral designs, which allow the largest number of small coat protein subunits and hence the least amount of nucleic acid, to enclose the largest volumes, was recognized (Caspar & Klug, 1962). This concept became known as genetic economy; namely, that viruses will respond to the selective pressure to minimize the amount of their respective genomes dedicated to production of their capsids, whilst at the same time trying to maximize that capsid's volume, to allow encapsidation of more viral genes dedicated to replication within the host. For icosahedral viruses, such arguments lead to a capsid composed of 60 coat protein subunits. In fact, most spherical viruses have more than 60 subunits, and this observation was explained via the concept of quasi-equivalence (Caspar & Klug, 1962) (Fig. 1) in which subtriangulation of the facets of an icosahedron can be used to predict "allowed" coat protein stoichiometries.

This idea of quasi-equivalence has proven central to all subsequent structural studies of viruses and was given molecular detail in 1977 when Harrison's group described the atomic structure of TBSV, an example of a $T=3$ capsid that contains 180 copies of a single type of coat protein subunit packaged as 90 non-covalently bound dimers (Winkler et al., 1977). Conformational changes between coat protein domains were seen that allowed these proteins to adopt different structures at defined symmetry contexts within the protein shell. Subsequent structural studies showed these conformational changes to be a common feature of these types of viruses (Suck et al., 1978; Liddington et al., 1991). These ideas saw minor revision with the realization that some animal viruses can be even more efficient in creating capsids that disobey quasi-equivalence by being built solely out of pentameric aggregates (Rayment et al., 1982; Liddington et al., 1991). Subsequent progress has been very rapid, and a number of X-ray crystal structures at atomic resolution have been reported for membrane-containing viruses (Grimes et al., 1998; Cockburn et al., 2004). Modern techniques, such as cryo-electron microscopy (cryo-EM), have allowed us to investigate the dynamical aspects of much more complex viral machines, including specialized appendages for injection of the genome into target cells (Gan et al., 2006; Lander et al., 2006).

Although work on the structures of many important viruses is now very advanced, the same can not be said of a number of critical questions with respect to the formation of these viral machines. Understanding how such machines are assembled is vital because the assembly pathways offer unique drug targets for anti-viral therapy. In recent years, established and novel emerging viruses have posed increasing threats to public health. Effective vaccines and anti-viral drugs are unavailable for the majority of these viruses, whilst those drugs that are available target viral replication, maturation, or receptor attachment but not yet assembly/disassembly or genome packaging. Many viruses have been successfully targeted by RNA interference (RNAi), a conserved sequence-specific, gene-silencing mechanism that is induced by double-stranded RNA and can trigger inhibition of virus replication, including important human pathogens such as human immunodeficiency virus type 1 (HIV-1), hepatitis C virus (HCV), hepatitis B virus (HBV), severe acute respiratory syndrome coronavirus (SARS-CoV), and influenza virus (Haasnoot, Westerhout, & Berkhout, 2007). Monoclonal antibody therapeutics can also inhibit viral infection by targeting the steps that viruses take to enter into cells (Marasco & Sui, 2007). Interfering with capsid assembly and stability is potentially a powerful target for anti-viral therapeutics. For example, the low molecular weight hydrophobic dye 1,1'-bis(4-anilino)naphthalene-5,5'-disulfonic acid (bis-ANS) has been shown to bind dimers of the HBV capsid protein with a 1:1 stoichiometry to inhibit capsid assembly (Zlotnick & Stray, 2003). Another example of this type of anti-viral activity is the use of the tripeptide Cys-Pro-Gly-NH₂, which is thought to target viral assembly or maturation in HIV-1 (Miller & Hazuda, 2001). A family of small-molecule drugs, the WIN compounds, have been identified that bind to surface "pockets" on the coat proteins of rhinovirus capsids and other members of the picornavirus family. This binding results in stabilization of the viral capsids and inhibition of viral disassembly/uncoating (Hadfield, Diana, & Rossman, 1999). However, widespread use of such anti-virals is prevented by our lack of understanding of the assembly process.

FIGURE 1. Virus architecture and quasi-equivalence (Zlotnick, 2004). **A:** A $T=1$ viral particle has 20 triangular facets, each having threefold symmetry, implying that it would consist of 60 coat protein subunits (asymmetric units). Higher triangulation numbers result from subtriangulating this basic arrangement. Higher T numbers can be calculated using the formula $T=h^2+hk+k^2$, where h and k are integers with values $h \geq 1$ and $k \geq 0$. Thus a $T=3$ structure is expected to have 180 coat protein subunits in its shell. **B:** A series of cryo-EM reconstructions of viral particles illustrating their quasi-equivalent structures. Reprinted with permission from Zlotnick (2004). Copyright (2004) National Academy of Sciences, USA.

Many virus particles and their specialized ancillary appendages can be disassembled *in vitro*, and their components fractionated and re-mixed in differing buffer conditions to lead to a rapid, precise reassembly of infectious/functional samples. The detailed molecular mechanisms that underlie these fascinating *in vitro* self-assembly reactions have been difficult to investigate because intermediates on the assembly pathways are present in limited amounts and are transient. It is, however, in these aspects of viral structure and function that non-covalent mass spectrometry (MS) has begun to make a major impact and is the subject of this review. The ability to understand the structure, assembly, and dynamics of viruses also promises to further their use in

nanotechnology: a variety of viruses has now been developed for such applications which include biomaterials, vaccines, chemical tools, imaging, and molecular electronic materials (reviewed in Singh, Gonzalez, and Manchester (2006)). Viral DNA packaging motors also have potential nanotechnological applications, including the delivery of drugs or therapeutic DNA or RNA, tissue and organ repair, the diagnosis of diseases, and the detection of pathogens (Guo & Lee, 2007).

II. THE USE OF MASS SPECTROMETRY TO PROBE VIRUS STRUCTURE

A. Why Use Mass Spectrometry?

For the past half century, the methods of choice for the determination of the 3-D structures of proteins and their macromolecular complexes at atomic resolution have been high resolution X-ray crystallography and, in some cases, nuclear magnetic resonance spectroscopy (NMR), both of which require milligram amounts of protein. For X-ray diffraction studies, the protein must be amenable to crystallization, whereas NMR is limited to soluble proteins of relatively low molecular mass. Cryo-EM has begun to solve some of these problems by allowing 3-D image reconstruction of small amounts of macromolecular complexes in vitreous ice at near-native conditions. In this way, functional dynamics as well as structures at moderate resolution can be investigated. Indeed, for viruses there is a pleasingly strong correlation between molecular features at moderate resolution in the cryo-EM and high-resolution structures from X-ray research work. Cryo-EM, however, lacks precise molecular information and here MS can provide unique insights.

The field of non-covalent MS has recently seen significant advances in detection and ionization technology that allow key questions about protein complexes to be addressed by their direct observation and characterization in the gas phase (Loo, 1997; Ashcroft, 2005a). One significant advantage of MS over other techniques is that relatively tiny volumes of material are required, often only at micromolar concentrations; another is the speed of the analysis, which can take just a few minutes. Also of importance is the fact that MS has the capability to monitor macromolecular reactions on-line, in real-time, making it ideal to follow biomolecular assembly pathways and to observe changes in protein structure. Hence, MS is proving to be an important and informative technique to monitor virus capsid assembly in real-time as well as disassembly achieved by perturbation of solution or instrument conditions. The detectable mass range in modern instruments, which now exceeds several million Daltons, allows the mass analysis of large, biomolecular assemblies, including intact viruses.

These significant recent advances have been made possible by the emergence of techniques such as electrospray ionization (ESI; Fenn et al., 1989) and the associated development of nano-ESI (Wilm & Mann, 1994), together with matrix-assisted laser desorption/ionization (MALDI; Karas & Hillenkamp, 1988; Tanaka et al., 1988), which rapidly made MS amenable for the routine mass analysis of large biomolecules such as proteins. Even from the early days, reports quickly emerged whereby the

boundaries of these techniques, in particular ESI-MS, were being extended from simply measuring the molecular mass of a protein to the characterization of proteins in their native conformations, either alone or non-covalently bound to other proteins, metals, or small ligand molecules (Chowdhury, Katta, & Chait, 1990; Ganem, Li, & Henion, 1991; Katta & Chait, 1991). In such analyses, the molecules of interest are subjected to ESI from an aqueous solution of a volatile buffer such as ammonium acetate. Careful desolvation of the protein assembly within the ion source generates multiply charged ions of the intact complex for analysis by MS. Thus, it became possible to obtain accurate information concerning not only the mass of the protein complex, but also important details about its stoichiometry, stability, and dynamics. However, such non-covalent analyses are not trivial to perform. As they are carried out under near-physiological conditions at neutral pH, they tend to suffer from low ionization yields, which result not only in weak signals (in mass spectrometric terms), but also in high mass-to-charge (m/z) ratio ions that require instrumentation with higher m/z ranges. The presence of varying quantities of salt and buffer ions specifically and non-specifically bound within the 3-D architecture of such complexes gives rise to broad, heterogeneous signals for which mass measurements can be challenging to deduce accurately, and this broadening is particularly apparent in the case of protein–polynucleotide interactions, where non-specific cation binding to nucleic acids is to be expected (Hanson & Robinson, 2004).

Thus, developments in MS analyzer design have also contributed significantly to the successful analysis of large, non-covalently bound macromolecules. These enhancements have included interfacing larger m/z range instruments to ESI (Verentchikov, Ens, & Standing, 1994), and experimental techniques such as “collisional cooling,” whereby nitrogen gas has been introduced into the intermediate pressure region of an ESI-MS instrument with a time-of-flight (ToF) analyzer to reduce the internal energy of the ions that originate from non-covalently bound macromolecular complexes. This “cooling” ensures that large, fragile ions remain sufficiently stable to traverse the mass spectrometer and reach the detector intact, thus avoiding dissociation into the individual biomolecular components (van Berkel et al., 2000; Sobott et al., 2002).

The application of these technological advances has paved the way not only for the analysis of intact virus capsids but also to provide valuable insights into understanding virus architecture and functionality. This review summarizes these approaches and covers the literature to the end of 2007.

B. Intact Virus Capsid Analysis

1. Intact Virus Analysis by ESI-MS

A pertinent starting point for this review is the key result published by Siuzdak in 1996 (Siuzdak, 1996). This study showed, for the first time, that it was possible to preserve the native conformation and activity of non-covalently bound viral particles throughout an entire mass spectrometric analysis. This finding was a major milestone for virologists and mass spectrometrists alike, and paved the way for subsequent mass

spectrometric research on a variety of viruses. Although early ESI-MS studies had shown that non-covalent interactions could be maintained during the ionization and m/z analysis processes of proteins (Chowdhury, Katta, & Chait, 1990), the investigation of supramolecular complexes was severely limited by the m/z ranges available on the instruments of that time. However, using an ESI-triple quadrupole mass spectrometer with an upper m/z limit of only 2300, Siuzdak (1996) successfully ionized two viruses, rice yellow mottle virus (RYMV) and tobacco mosaic virus (TMV), which had been chosen for their contrasting viral structures (Table 1). RYMV is a single-stranded RNA virus with an icosahedral capsid, ~29 nm in diameter, comprised of 180 copies of a coat protein monomer with a total mass of 6.2 MDa; TMV is also a single-stranded RNA virus, consisting of ~2,140 identical protein subunits that form a rod-shaped helical capsid of diameter 17.5 nm and length 300 nm with a mass of 40.5 MDa (Bharyavhatla, Watowich, & Caspar, 1998). For this experiment, a glycerol-coated brass plate was fitted within the instrument to act as an ion collector between the second and third quadrupoles. With all of the quadrupoles operating in non-analyzing Rf mode only, the virus-related ions were transmitted through the first and second quadrupoles, and collected on the inserted brass plate. This plate was subsequently removed from the instrument and when the two collected viruses were examined by transmission electron microscopy (TEM), both were found to have retained their distinctive morphologies. Furthermore, the viability of the TMV species was established by showing retention of infectivity. These data provided the first direct observation that the biological activity of a supramolecular complex could be preserved throughout the ESI-MS process, and specifically that viral structure and activity could be retained after subjection to ionization, high vacuum conditions, and finally impact on a collector plate. Although the m/z ions generated were much too high for a mass analysis to be performed on this particular instrument, the vital proof that such complexes could survive this type of experiment had been obtained, and from this result the implication that MS was indeed a viable and important technique for the investigation of virus architecture.

Marriage of the continuous stream of ESI-generated ions to the pulsed ToF analyzer (Verentchikov, Ens, & Standing, 1994) paved the way to a range of commercially available ESI-ToF and ESI-QToF mass spectrometers with much larger (theoretically infinite) m/z ranges, and hence made feasible, although still demanding, the mass analysis of intact supramolecular complexes. Thus, in 2000 the mass of the bacteriophage MS2 capsid, a $T = 3$ icosahedral virus consisting of 180 copies of a single coat protein (13.7 kDa) and of total mass 2.5 MDa, which forms a shell around a single-stranded RNA molecule, was measured with an error of only $\pm 1\%$ (Tito et al., 2000; Fig. 2). The fact that the observed mass was 0.55% higher than the actual mass was accounted for by the ambiguity in the assignment of the barely separated, broad peaks associated with the observed charge states ($n = 101\text{--}124$, m/z 20,000–24,000, where $m/z = [M + nH^+]/n$) and also to the propensity of small molecules and counterions to remain bound to protein assemblies to increase their observed mass. However, this measurement marked an outstanding achievement and another milestone in virus analysis.

Recent work from the Heck group (Lorenzen et al., 2007) on HBV, an important human pathogen, that causes chronic liver

disease in an estimated 250 million individuals worldwide, confirmed the co-existence of icosahedral capsids of $T=3$ (180 coat protein monomers) and $T=4$ (240 coat protein monomers) geometry by mass spectrometric measurements from their respective charge-state distributions ($T=3$: 2.98 MDa, $n=128-147$, m/z 20,000–23,500; $T=4$: 3.99 MDa, $n=144-170$, m/z 23,500–28,000), using an ESI-QToF instrument with improved performance (van den Heuvel et al., 2006) and an ESI-QToF with integrated ion-mobility spectrometry (IMS; Giles et al., 2004; Pringle et al., 2007). ESI-IMS-MS of the intact capsid shells permitted ionization, separation, and m/z analysis of these co-populated, individual species, and led to an estimation of the cross-sectional areas of 600 and 800 nm² for the $T=3$ and $T=4$ icosahedra, respectively (Lorenzen et al., 2007). These estimations were slightly lower ($\sim 10\%$) than would be expected from the diameters of these capsids measured by X-ray crystallographic studies (Zlotnick et al., 1999). Thus, mass and physical size were both estimated in one experiment.

2. Intact Virus Analysis by Charge-Detection MS

An alternative approach to conventional m/z analysis for the mass analysis of intact viral capsids is the use of charge-detection mass spectrometry, a technique that involves simultaneous measurements of charge and m/z for individual ions, and which is

especially useful in the case of samples that generate high charge states, which are difficult to assign unambiguously (Smith, Cheng, & Bruce, 1994; Benner, 1997; Schultz, Hack, & Benner, 1998, 1999; Fuerstenau et al., 2001). Employing ESI as the first stage in the analysis, this approach has been used with ToF (Benner, 1997) and Fourier transform ion cyclotron resonance (FT-ICR; Smith, Cheng, & Bruce, 1994) m/z analyzers. After the ESI process, the virus ions pass through a metal flight tube attached to a charge-sensitive pre-amplifier that captures their image current, the magnitude of which is proportional to their charge. The m/z ratio is measured by the analyzer of the mass spectrometer, and from this combination of charge and m/z , the mass of each ion can be deduced. The measurements of RYMV and TMV have been reported by Suizdak et al., using this methodology coupled to a ToF mass analyzer (Fuerstenau et al., 2001), following on from their early pioneering studies with the same two viruses. In their experiment, thousands of individual ions were measured and charge-state distributions were obtained. The charge-state distribution of large particles depends on particle shape, together with the pH and solvent used for the ESI process. The RYMV and TMV ions each possessed a charge distribution of between 300 and 1,000 positive unit charges, with higher charge states also being observed for TMV. From these distributions, masses of 6–7 MDa and 39–42 MDa for RYMV and TMV, respectively, were calculated (Fig. 3). These observed

FIGURE 2. Mass spectrum of the intact MS2 virus capsid showing the histogram produced for the raw data (a) and the onset of the charge state distribution after data processing (b): (i) a single 10 sec acquisition step in which single ions are visible in each bin above the level of background noise; (ii and iii) summation of 10 and 50 acquisitions in which a maximum of 13 ions are detected in each bin. The total ion current measured for the charge state distribution was 1.67×10^3 . Spectra were recorded on a LCT mass spectrometer (Micromass UK Ltd, Manchester, UK) modified to allow increased pressure in the first hexapole. The following voltages were found to be optimal: needle 1.3 kV, cone 150 V, RF lens 480 V, extractor 0 V, and MCP 2.7 kV (Tito et al., 2000). Reprinted with permission from Tito et al. (2000). Copyright 2000 American Chemical Society.

FIGURE 3. Mass spectra of rice yellow mottle virus (RYMV, top) and tobacco mosaic virus (TMV, bottom) particles analyzed with an ESI charge-detection ToF-MS. Insets: EMs of the icosahedral RYMV (diameter 28.8 nm; top) and the helical TMV (~300 nm long and 17 nm diameter; bottom). The known molecular weights of RYMV and TMV are 6.5 and 40.5 MDa, respectively. From these distributions, masses of 6–7 and 39–42 MDa respectively were calculated (Fuerstenau et al., 2001). Reprinted from Fuerstenau et al. (2001) with permission from Wiley-VCH Verlag GmbH & Co.

masses are within $\pm 15\%$ of the calculated masses (6.5 and 40.5 MDa, respectively), and instrumental improvements were suggested by the authors to better these values.

3. Intact Virus Analysis by Gas-Phase Electrophoretic Mobility Molecular Analysis (GEMMA)

Charge reduction of multiply charged ions is employed in gas-phase electrophoretic mobility molecular analysis (GEMMA; Bacher et al., 2001). As the ESI-MS spectra of large protein complexes can be difficult to interpret due to the presence of many, often partially resolved, multiply charged ions, charge reduction of these species to generate singly charged molecular ions provides a means to simplify such spectra. GEMMA, which provides molecular size information rather than a direct mass measurement, works on the principle whereby multiply charged ions generated by ESI are transformed by a bipolar neutralizer into singly charged molecular ions that are subsequently separated according to their electrophoretic mobility as they pass through a differential mobility analyzer. The electrophoretic mobility diameter is defined as the diameter of a singly charged sphere with the same electrophoretic mobility as the particle. Depending on the applied air flow rate and voltage in this region, ions with a certain mobility leave the differential mobility

analyzer and are detected by a condensation particle counter. Finally, the number of detected ions is plotted against the inverse electrophoretic mobility, and this plot is transformed into a GEMMA spectrum, which shows the concentration of charged particles as a function of their calculated electrophoretic mobility diameters.

Human rhinovirus (HRV) is the major cause of common cold infections, and natural viral populations contain many different antigenically distinct serotypes that hamper the development of effective vaccines. Serotypes 2 and 14 (HRV2 and HRV14) have been analyzed with GEMMA. Analysis of an intact preparation of HRV2 gave rise to a spectrum that showed a broad peak with an electrophoretic mobility diameter of 29.8 ± 0.3 nm (Bacher et al., 2001), consistent with the value of ~ 30 nm obtained from cryo-EM images (Hewat & Blaas, 1996). The molecular mass of HRV2 was estimated from the measured electrophoretic mobility diameter to be ~ 5.20 MDa, rather lower than the calculated mass of ~ 8.08 MDa based on the known building blocks. Two explanations were proposed for this discrepancy: one was that the density of single-stranded RNA and of the space between RNA and the capsid protein shell are different from the density of the protein alone (on which the calibration was based), and secondly that the nearest calibration point to convert electrophoretic mobility data to molecular mass was that of a protein of mass < 2 MDa. Experiments with HRV14 gave similar results. The broadness and hence heterogeneity of the peaks, a familiar phenomenon of macromolecular complexes, was attributed again to the binding of small molecules such as water and buffer.

Four icosahedral viruses, namely RYMV, cowpea mosaic virus (CPMV), the bacteriophage MS2, and the human adenovirus, a double-stranded DNA virus that is a major cause of human respiratory ailments, have also been examined by the GEMMA technique (Thomas et al., 2004). The calculated diameters of RYMV, MS2, and CPMV were within 15% of the values obtained from X-ray crystallography, but in each case were consistently smaller. For example, RYMV gave a spectrum that corresponded to a diameter of 28.5 nm, slightly smaller than the diameter calculated from the crystal structure (29–32.2 nm), and the observed value for CPMV was 25.9 nm diameter compared with the 28–32 nm measured from X-ray crystallography data. These discrepancies were suggested to be due to the removal during the ESI-GEMMA analysis of most of the water molecules that are associated with the virus ensembles, which would result in an overall compaction of the macromolecular structure due to the loss of these water molecules from the protein–protein and protein–nucleic acid interfaces. However, the GEMMA technique was successful in resolving a 1:1 mixture of MS2 and RYMV, both composed of 180 copies of a single protein with $T = 3$ symmetry, whose structures differ in diameter by only 4 nm (15%). The data from the analysis of the adenovirus capsid had lower resolution and signal-to-noise than the data obtained from the other capsids.

Two of the large tailed bacteriophages, T2 and T4, together again with MS2, have also been investigated with mixed success with ESI-GEMMA (Hogan et al., 2006). T2 and T4 are double-stranded DNA viruses with masses > 1 GDa. The two are almost identical, with the exception that the capsid head of T4 contains additional protein subunits (Kostyuchenko et al., 2005). In this study, to determine the viability of viruses after the ESI process,

the electrosprayed particles were collected in a liquid medium, and virus activity was subsequently determined by a microbiological plaque assay. The mean diameter of the bacteriophage MS2 was measured to be 24.1 nm, slightly lower than the particle diameter of 27.5 nm measured by X-ray crystallography; however, the virus remained viable after the ESI process. The lengths of T2 and T4 are >210 nm, and thus as they are larger than most of the droplets produced by the ESI used in these experiments, the authors gave this reason to explain the fact that intact capsids could not be detected in either of these cases, but rather that virus fragmentation was observed (Hogan et al., 2006). The conclusion from this study was that the production of sub-micrometer electrospray droplets can constrain large macromolecules and cause fragmentation at the non-covalently bound interfaces.

Loo and co-workers have used ESI-GEMMA to calculate the mass of cowpea chlorotic mottle virus (CCMV), a member of the *Bromoviridae* family of plant viruses (Kaddis et al., 2007). The capsid has $T=3$ icosahedral symmetry and is composed of 180 identical protein monomers of 19.8 kDa each. The intact shell has a diameter of 28.6 nm, determined by X-ray crystallography (Speir et al., 1995). ESI-GEMMA analysis led to an observed diameter of ~10% less than the measurement from the X-ray crystal structure, as others had found previously (Hogan et al., 2006; Thomas et al., 2004), (Fig. 4). From these studies, the overall conclusion is that there is a general correlation between molecular mass and electrophoretic mobility for most protein complexes observed from the ESI-GEMMA technique. Protein complexes held together by weak non-covalent forces survive the ESI and subsequent desolvation processes and can be measured; however, the supposition is that on desolvation, protein complexes collapse somewhat although key features of their overall structures are preserved and biological function can be restored on return to the solubilized state.

4. Intact Virus Analysis by MALDI-MS

The field of mass spectrometric analysis of virus capsids is dominated by ESI, a solution-based ionization technique that lends itself to sampling biomolecules in a native-like environment and that generates ions with multiple charges, and hence lower m/z ratios, than MALDI, a technique that produces predominantly singly charged, molecule-related ions. An exception to this domination is the report that a MALDI-Tof instrument with cryodetection has successfully measured the mass of the capsid head of the lamboid double-stranded DNA phage, Hong Kong 97 virus (13 MDa), to within 1% accuracy (Bier, 2007). The ultra-sensitive cryodetector utilizes a superconducting tunnel junction detector that allows high sensitivity for the detection of singly charged, high molecular mass ions (Wenzel et al., 2005).

C. Virus Capsid Assembly

1. Assembly Pathways Monitored Directly by ESI-MS

The confidence gained from these impressive high-mass measurements on intact viral capsids, and the evidence that

FIGURE 4. The ESI-GEMMA spectrum of the 4.6 MDa cowpea chlorotic mottle virus (CCMV) complex (a) and its corresponding structure (b). The structure figure was created using WebLab Viewer with the coordinates from the Protein Data Bank (1cwp) (Kaddis et al., 2007). ESI-GEMMA led to an observed diameter ~10% less than that of 28.6 nm determined by X-ray crystallography for the intact shell (Speir et al., 1995). Reprinted from Kaddis et al. (2007) with permission from Elsevier.

such supramolecular complexes can be preserved in their active form throughout the mass spectrometric analysis, has resulted in the wider application of MS to investigate viral structure and function. The architectures of these highly ordered structures have been probed by following their assembly pathways in real-time and by disassembling intact structures into individual building blocks for detailed investigation. As capsid assembly/disassembly under appropriate, defined conditions is generally spontaneous and rapid, the isolation and characterization of the intermediates *en route* to the final products is technically challenging.

Electrospray ionization-mass spectrometry (ESI-MS) has been used successfully to examine the stoichiometry of several bacteriophage portal complexes, key parts of the molecular motors that drive the DNA packaging in this highly successful class of viruses (Poliakov et al., 2007). Double-stranded DNA viruses rely on an active packaging mechanism for encapsidation

of their genomes, in which a molecular motor forces genomic DNA into a pre-formed icosahedral procapsid. The packaging motors consist of at least two parts, the terminase (or ATPase) complex and the connector or portal complex. The portal motor complex replaces the normal capsid coat proteins at one of the fivefold vertices of the procapsid. In $\Phi 29$, which has one of the strongest molecular motors known to date, virally encoded RNA molecules are also involved in this process. The action of the packaging motor results in the compaction of double-stranded DNA to almost crystalline density (Smith et al., 2001). The origin of force generation in these systems has long remained a mystery. Almost 30 years ago, Hendrix proposed that symmetry mismatching between a portal complex with sixfold symmetry and the procapsid with local fivefold symmetry would allow a Brownian ratchet type of mechanism to occur (Hendrix, 1978). The implication of this model is that rotation of the portal with respect to the procapsid must occur, although to date no one has demonstrated such movement. A second inference is that the protein stoichiometries of the portal complexes must be very tightly regulated to ensure the symmetry mismatching. More recently, conformational change in motor components has been suggested as an alternative model of motor function (Hugel et al., 2007; Robinson et al., 2006). Understanding the assembly of the portal complex and its role(s) in DNA packaging has therefore been a major goal of structural virology.

The bulk of the structural information on this group of molecules has been obtained by TEM. Although the stoichiometry of smaller intermediate assemblies *en route* to the capsid can be analyzed by size-exclusion chromatography or analytical ultracentrifugation measurements, larger intermediate assemblies that consist of higher order oligomerization products present a challenge to these analytical techniques. Heck and co-workers have recently used ESI-MS to determine and semi-quantify the oligomeric states of *in vitro* reconstituted phage portals. Their findings showed conclusively that the isolated portals of the bacteriophages $\Phi 29$, SPP1, and P22 consist of 12-mers, 13-mers, and a mixture of 11- and 12-mers, respectively. Confirmation that the 11-mer assemblies of mass ~ 1 MDa were the predominant species in the case of reconstituted P22 was obtained from negatively stained EM images (Poliakov et al., 2007). The complexity of the m/z spectra obtained and the need for sophisticated data interpretation are illustrated with the P22 portal data shown in Figure 5. Such molecular details are essentially inaccessible by any other technique except cryo-EM coupled with single particle image reconstruction, because the

mixed stoichiometries make it unlikely that these species would crystallize, and hence eliminate the possibility of X-ray crystallography.

The capsid structure and *in vitro* reassembly pathways of the $T = 3$ RNA bacteriophages Q β and MS2 have been investigated with ESI-MS by the Leeds Group (Ashcroft et al., 2005b; Stockley et al., 2005, 2007). The X-ray structure of Q β at 3.5 Å resolution suggested that two cysteine residues within loops of neighboring coat protein subunits were sufficiently close to form inter-molecular disulfide bridges at the capsid five- and threefold axes (Golmohammadi et al., 1996). These interactions were investigated by ESI-MS (Ashcroft et al., 2005b). Analysis of Q β capsids under neutral solvent conditions after disulfide bridge reduction showed the existence only of coat protein dimers, the expected capsomer (principal building block) in these systems; subsequent denaturation led to the detection of protein monomers (14,123 kDa). However, disassembly of the intact capsids led to an ESI-MS spectrum that indicated the presence of protein pentamers and, to a much lesser extent, hexamers; disulfide bridge reduction of this mixture again led to monomer detection. Thus, MS suggests that a contributing factor to the stability of Q β capsids is the formation of inter-molecular, disulfide linkages at the fivefold axes that lead to five covalently coupled monomers. The low yield of similarly linked hexamers suggests that the orientation of the cysteines at the particle threefold axes is non-ideal for disulfide bond formation. These data have obvious implications for capsid assembly and genomic RNA release in such systems.

Re-assembly of the complete $T = 3$ MS2 capsid has also been monitored directly by ESI-MS (Stockley et al., 2007). *In vitro* re-assembly was triggered by a sequence-specific RNA–protein interaction between a coat protein dimer (CP₂) and an RNA stem-loop (TR) of 19 nucleotides encompassing the well-characterized packaging signal on the RNA genome (Fig. 6A). The re-assembly was monitored by nano-ESI-Tof by a temperature-controlled, automated injection and ionization system (Stockley et al., 2007). The addition of TR RNA to a solution of coat protein at pH 5.2–5.7 with a 1:1 molar stoichiometry (CP₂ to TR) resulted in the observation of a CP₂:TR complex of mass 33.5 kDa. Minor, higher-order oligomers in the region m/z 4000–10,000 were also observed, but their concentration did not change significantly over time. However, when the stoichiometry of CP₂:TR was increased to 2:1, the proportion of higher-order species increased dramatically, and at least one significant higher mass species of 182.7 kDa was observed. This species was consistent with the

FIGURE 5. ESI mass spectra of the reconstituted P22 portal sprayed from 50 mM aq. ammonium acetate. **A:** Zoomed in part of the spectrum focusing on the undecameric and dodecameric ion signals. The inset in (A) shows the zero-charge convolution of this spectrum, which indicates that the ratio between undecamers and dodecamers of the portal protein P22 is $\sim 2:1$. **B:** Whole range mass spectrum. Oligomeric assemblies with stoichiometries varying from 1 to 8 and from 11 to 12 are present. **C:** Analysis and Solving complex Macromolecular Mass Spectra (SOMMS) simulation (van Breukelen et al., 2006) of the whole ESI-MS spectrum. Peak position and intensity coincidence confirms correctness of assignment of different peaks. The inset shows the SOMMS simulation-derived relative abundance of ions of different oligomeric states. The simulation confirms that the ratio between undecamers and dodecamers is $\sim 2:1$ and furthermore reveals that P22 is partly unstable under the solution conditions used for ESI, and dissociates primarily into species ranging from trimers to octamers. Interestingly, the nonamers and decamers are nearly absent in the ESI spectrum (Poliakov et al., 2007). Reprinted from van Breukelen et al. (2006) with permission from Elsevier.

FIGURE 5.

complex $[3(\text{CP}_2:\text{TR}) + 3\text{CP}_2]$ which corresponds to the species expected at a threefold axis of symmetry in the final capsid (Fig. 6B). Over time, these higher-order species decreased in intensity, and were accompanied by broader, unresolved higher m/z signals ($m/z > 14,000$), consistent with signals observed for the intact $T=3$ capsid. The complete regeneration of intact capsids under these conditions was confirmed by gel-filtration/light-scattering techniques and TEM. The results imply that entire viral assembly pathways can be analyzed via MS techniques to reveal insights into the details of the assembly pathways and intermediates.

As an example of the utility of this approach, it was possible to confirm that the complexes observed in the ESI-MS spectra were on-pathway to capsid formation and to determine the mass step size of capsid growth. The reassembly reaction was repeated, using isotopically labeled coat protein. The $\text{CP}_2:\text{TR}$ complex was pre-formed as above with a 1:1 molar ratio of reactants, but using ^{15}N -labeled CP_2 . The intermediate complex $[3(\text{CP}_2:\text{TR}) + 3\text{CP}_2]$ observed in this sample was fully ^{15}N -labeled and hence all 12 coat protein monomers were labeled. RNA-free ^{14}N - CP_2 was added to initiate the 2:1 reaction. Immediately, the ESI-MS spectrum indicated two peaks of approximately equivalent intensity that corresponded to either all ^{14}N or all ^{15}N -labeled $\text{CP}_2:\text{TR}$, but with no evidence for a heterodimer complex; those data indicated that, whereas the TR RNA molecule can dissociate from the $\text{CP}_2:\text{TR}$ complex and re-associate with a different CP_2 , the protein dimer does not readily dissociate to monomer under these conditions. Titration of increasing amounts of ^{14}N - CP_2 resulted in the detection of $[3(\text{CP}_2:\text{TR}) + 3\text{CP}_2]$ complexes that contained 12, 10, 8, 6, 4, 2, and 0 ^{15}N -labeled coat protein molecules; that is, corresponding to all possible stoichiometries that differ by dimer addition. Thus, it appears that, as the initially formed $[3(\text{CP}_2:\text{TR}) + 3\text{CP}_2]$ ^{15}N -labeled complexes continue on the assembly pathway to form capsids, new $[3(\text{CP}_2:\text{TR}) + 3\text{CP}_2]$ are formed from a mixture of the remaining ^{15}N -labeled CP_2 species and the newly-added ^{14}N material. Eventually, the percentage of all ^{15}N -labeled $[3(\text{CP}_2:\text{TR}) + 3\text{CP}_2]$ decreases with increasing CP_2 concentration, whereas the mixed $^{14}\text{N}:$ ^{15}N and eventually all ^{14}N complexes predominate. Finally, this complex disappears as capsids are formed.

Combining data obtained from mass spectrometry, NMR, light-scattering, and TEM experiments led to the formulation of the most-detailed molecular mechanism for $T=3$ capsid formation to date. The RNA-free MS2 CP_2 in solution has a preponderantly symmetrical (C/C), quasi-equivalent conformation distinct from the asymmetric dimer created when TR binds, but both forms of dimer are required to allow the capsid to be assembled efficiently (Fig. 6A). Thus, the MS experiment leads directly to a model in which the mechanism of coat-protein quasi-equivalent selection during capsid assembly, an unsolved issue despite the many structures of viral capsids that are available, is seen to be due to an allosteric effect of RNA binding.

An elegant investigation into the processes of genome dimerization and packaging in HIV-1, which are mediated by specific interactions between the nucleocapsid protein and structural elements formed by the genome packaging signal, ψ -RNA, has also been carried out (Fabris et al., 2007). HIV-1 selectively packages two copies of its genomic RNA which are non-covalently bound by an intermolecular base-paired region. The formation of this structure involves the initial annealing of a self-complementary sequence located in the stem-loop domain of the genome-packaging signal, ψ -RNA, a domain that consists of a highly conserved stretch of RNA folded into four discrete stem loops. Subsequent dimer stabilization involves structural rearrangements that are mediated by the chaperone activity of the nucleocapsid domain of the viral Gag polyprotein. Fabris et al. (2007) have used nano-ESI-MS coupled to an FT-ICR mass spectrometer to determine the stoichiometry and binding constants of the assemblies formed by the nucleocapsid with the four stem loops, and to investigate the possible inhibitory effects of several small ligands on the nucleocapsid-stem loop complexes. Competitive-binding experiments that involve the isolated RNA elements were performed to evaluate the ability of these species to sustain specific protein interactions. Using RNA mutants, two distinct binding sites on one stem loop were identified, and the preferred site for protein binding was identified on a second stem loop. These results provide the first evidence that interactions between contiguous structures modify the binding modes of the different sites in the context of full-length ψ -RNA.

FIGURE 6. A: The structures and solution components associated with MS2 bacteriophage assembly. (a) shows a space-filling representation of the X-ray structure of the $T=3$ capsid of wild-type MS2 (PDB 2MS2). (b) shows the sequence and secondary structure of the TR RNA. Panel (c) shows the distribution of the five- and threefold symmetry axes (black and white symbols respectively) within the icosahedral surface lattice of the MS2 capsid and the arrangement of the quasi-equivalent dimers, AB (blue/green) and CC (red) within it. Alongside are ribbon models for the structures of each, with the A/B dimer bound to TR RNA, and an enlarged cartoon of their relationship within the capsid (Stockley et al., 2007). **B:** Bacteriophage MS2 capsid re-assembly monitored by real-time nano-ESI-MS. The spectra were acquired over the range m/z 500–30000 for samples in ammonium acetate (40 mM) at pH 5.2–5.7. The components are labeled as follows: A = CP; B = CP_2 ; C = $\text{CP}_2:\text{TR}$; D = 182.7 kDa, assigned as $[3(\text{CP}_2:\text{TR}) + 3\text{CP}_2]$, and E = ~288–300 kDa, unassigned. The number immediately following each letter is the charge state of those particular ions. The spectra are as follows: (a) CP alone (8 μM); pH 3.2; (b–d) are for re-assembly at pH 5.2–5.7 at a stoichiometry of CP_2 to TR 1:1 (8 μM :8 μM) at times 1, 90, and 300 min, respectively. Spectrum (e) is the 1:1 pre-equilibrated starting point for re-assembly at a stoichiometry of CP_2 to TR 2:1 (16 μM :8 μM), spectra (f–h) at times 1, 120, 180 min, respectively. Note, spectra (b) and (e) are essentially equivalent. The bars below the spectra indicate a magnification factor of 5 for all ions above m/z 6,000 to enhance clarity. Reprinted from Stockley et al. (2007) (Elsevier).

2. Capsid Architecture Explored by Hydrogen-Deuterium Exchange MS

Viral capsid assembly involves dynamic protein–protein and protein–nucleic acid interactions. Hydrogen-deuterium exchange (HDX), followed by rapid proteolytic cleavage and mass spectrometric analysis to determine the amount of label incorporated into specific regions of the protein, can be used to follow the differential exposure, or protection, of sites during the process. HDX does not involve non-covalent MS; rather, the exchange experiments are performed on the intact supra-molecular complexes which are then denatured and proteolyzed before MS analysis. However, because the data produced provide important insights into the non-covalently bound architecture of these complexes, the topic is included in this review.

Hydrogen-deuterium exchange-mass spectrometry (HDX-MS) experiments are extremely demanding in terms of precision of timing and strict pH and temperature control (Engen & Smith, 2000). After the protein or protein complex has undergone (H → D) HDX, the pH and temperature of the reaction mixture must be lowered to minimize back-exchange (D → H) of the amide backbone residues of the protein during subsequent analysis. Hence, the rapid proteolysis step is usually accomplished with pepsin, an enzyme that works optimally at low pH, and proteolysis often takes place within the injection loop or on a specially packed, on-line column for fast throughput. Following digestion, if using ESI-MS, an HPLC step is generally required to separate the complex peptide mixtures generated by proteolysis; typically, the injector and column must be held at low temperature to minimize back-exchange during elution *en route* to the mass spectrometer. Any back-exchange during the analysis leads to greater uncertainty in the structural information generated, and therefore control experiments to calculate the expected back-exchange of amide residues during the proteolysis and HPLC-MS steps are required (Zhang & Smith, 1993).

The double-stranded DNA bacteriophage P22 assembles in two distinct stages. Firstly, the coat protein (46,700 Da) assembles with the assistance of ~300 scaffolding protein monomers into a $T=7$ procapsid that contains 420 subunits (19.6 MDa) with the scaffolding proteins encapsidated and a portal protein complex located at one of the icosahedral vertices. Secondly, the scaffolding protein is released in a reaction triggered by DNA packaging, and the capsid lattice expands to become the mature, stable $T=7$ capsid. *In vitro*, capsid expansion can be initiated by elevated temperature (Galisto & King, 1993). The use of MS techniques with HDX followed by limited pepsin proteolysis has had some success in identifying the specific residues that participate in this transformation (Tuma et al., 2001). MALDI-Tof was used to identify the peptide fragments generated and to determine the amount of deuterium uptake (i.e., the accessibility of each region) that occurred. Approximately 30% of the monomeric P22 protein sequence was monitored by this approach, and the rate of HDX varied from one region of the protein to another. Some regions were exposed to exchange before procapsid expansion, whereas on formation of the mature capsid, about half of the amino acid residues in these regions became fully protected.

A further investigation reported the expression of individual domains of the P22 coat protein to enable individual HDX studies

to be carried out with ESI-Tof and ESI-FT-ICR mass spectrometry (Kang & Prevelige, 2005). In this instance, ~85% of the coat protein was successfully mapped from the peptic peptide fragments: a significant improvement over the earlier work. Again, significant differences between the corresponding protein regions were observed when comparing the procapsid shells with the mature capsids. In general, HDX occurred rapidly in the empty procapsid shells, but the exchangeable hydrogens were well protected in the mature capsids; in fact, there was only one region that exchanged more rapidly (i.e., was more exposed) in the mature capsids compared with the procapsids. From this experiment, it was concluded that the mature capsids have more extensive inter-subunit interactions than the procapsids, but that the series of rearrangements in which new contacts are created can cause a slight local exposure of the other sites after maturation.

Insights into the mode of action of the bacteriophage $\Phi 29$ have also been made with HDX ESI-MS methodology (Fu & Prevelige, 2006). The $\Phi 29$ scaffolding protein consists of three α -helices with a disordered C-terminal; it is this protein that interacts with the capsid protein monomers to promote efficient assembly into icosahedral shells. The interactions between the scaffolding and capsid proteins are transient and are disrupted during ATP-driven DNA packaging into the procapsids. Following DNA packaging, the scaffolding protein is lost either by proteolysis or via a conformational switch of the capsid that allows dissociation. To investigate these reactions, the conformational differences between free and procapsid-bound scaffolding protein have been explored in an elegant study that uses HDX-MS and revealed highly dynamic and co-operative opening motions in the N-terminal region. At low temperatures (4 or 10°C) when these motions were dampened, or in the case of a mutant scaffolding protein in which the N-terminal helix-loop-helix motif was tethered via a disulfide bond, the co-operative opening motions of this region were restricted; those data confirm the interpretation of the HDX-MS results. The opening rate was increased in the procapsid-bound compared to the free form to suggest that the N-terminal helix-loop-helix motif might be the region that interacts with the capsid protein.

Similar HDX-MS techniques have been used to explore the hexameric ATPase, helicase P4, the viral genome packaging motor of double-stranded RNA bacteriophage $\Phi 8$ (Lisal et al., 2005). The HDX kinetics of the highly respectable ~85% protein sequence of P4 detected by MS showed no large changes in the protection of subunit interfaces during ATP binding and hydrolysis. However, nucleotide- and RNA-binding did result in localized changes of HDX kinetics, and different nucleotide-bound states revealed a co-operative unit that spanned two neighboring nucleotide-binding sites that links ATP-binding sites with the central RNA-binding channel and thus identifies the mode of coupling of ATP hydrolysis to RNA translocation. A later report on this system expanded the study to monitor the interaction(s) of P4 with the procapsid, and showed that virus-bound P4 is more stable than free P4 in solution (Lisal et al., 2006). Further work on the P4-procapsid complex, employing HDX-ESI-MS and tandem mass spectrometry (MS/MS) with both FT-ICR and ion trap instruments, has revealed interactions of the packaging motor with the viral core and provided novel insights into the roles of different regions in assembly, genome packaging, and transcription regulation (Lisal et al., 2006).

Hydrogen-deuterium exchange (HDX) followed by rapid proteolysis and HPLC-ESI-MS with FT-ICR analysis was also employed to map the protein–protein interfaces for *in vitro* assembled HIV-1 capsid protein (Lam et al., 2002; Lanman et al., 2003) and mapped an impressive ~95% of the primary sequence. At low salt concentration, the capsid protein (25,601 Da) exists in a monomer-dimer equilibrium, whereas at high ionic strength it polymerizes into long, tube-like structures that consist of hexameric subunits. HDX of the monomeric protein was compared with that of the assembled capsid, and a difference in exchange rates was noted when comparing peptides from both complexes; that difference suggested that the dimer interface is important in the formation of the tube-like structure. The HDX patterns of the immature and mature capsids of HIV-1 have also been compared, using ESI-MS—both on the intact protein and on peptide fragments that result from pepsin proteolysis. The immature capsid is spherical in morphology, with the Gag polypeptide arranged radially. Upon maturation, the capsid protein collapses to form a conical cone and this encases the complex of the nucleocapsid protein and the RNA. The changes in protection upon maturation have been interpreted in terms of an interaction between the *N*- and *C*-terminal domains of the capsid protein. Only half of the capsid protein assembles into the conical core; that is, only a sub-population of the total capsid protein molecules in the virion contribute to the mature viral core (Lanman et al., 2004).

Hydrogen-deuterium exchange (HDX) coupled to on-line capillary HPLC and nano-ESI-MS has also been used to map the in-solution structure and dynamics of the HRV14 capsid (Wang & Smith, 2005). Comprehensive local exchange profiles of each of the four capsid proteins of HRV14 were generated, and these exchange profiles could be correlated with the known crystal structures whereby the extended flexible terminal and surface-loop regions were unprotected whereas the folded helical and sheet regions were protected. The beta-strands of the VP3 *N*-terminal domain also showed a very slow exchange that suggested stable pentamer contacts. It was noted, however, that the interface around the fivefold axis had a fast and intermediate exchange that indicated relatively more flexibility. One potential application of the availability of such detailed information would be to study structural changes in the HRV capsid brought about, for example, by the binding of anti-viral agents.

3. Capsid Architecture Explored by Chemical Cross-Linking and Hydrogen-Deuterium Exchange MS

Although HDX can identify protected regions that might be assumed to be involved in protein–protein interactions, identifying the partner regions paired in such interactions is more difficult. Chemical cross-linking, using bi-functional cross-linking reagents, which involves the covalent joining of two molecules, or two regions within a single molecule, has been identified as a method to overcome this difficulty.

Prevelige and co-workers have measured the amide hydrogen-exchange protection factors of the soluble HIV-1 capsid protein (see also Section II.C.2) and used these results, together with chemical cross-linking experiments, to map the subunit–subunit interfaces in the assembled hexameric ring-based tube

structures (Lanman et al., 2003). The homo-bi-functional lysine-reactive cross-linking molecule disuccinimidyl tartarate, a short cross-linker that covers a distance of 6.4 Å (0.64 nm), was used to link sub-units in tube-like structures. The authors dissociated the tubes and isolated the covalently linked dimers to find ~50% carried a single cross-linker molecule, ~35% carried two cross-linker molecules, and the remainder carried three cross-linker molecules. Proteolysis followed by HPLC-ESI-MS identified the primary cross-link between Lys70 in the *N*-terminal domain of one protein subunit and Lys182 in the *C*-terminal domain of the other subunit; those data demonstrated the existence of a previously unrecognized interaction between helices III and IV and the *C*-terminal domain of another capsid protein. These two regions also become protected upon assembly.

Subunit–subunit interactions in bacteriophage P22 procapsids have also been examined by chemical cross-linking, and specific inter- and intra-subunit interactions have been identified (Kang et al., 2006). All of the inter-subunit cross-linking was found to occur in a particular loop region to thus confirm previous HDX-MS results and to suggest a protein–protein interaction in this region. Three amine-specific cross-linking molecules were used, all of which generate an amide bond with a lysine residue: disuccinimidyl tartarate, dithiobis(succinimidyl propionate), and bis(sulfosuccinimidyl) suberate; the latter two have similar lengths of 12 Å (1.2 nm) and 11.4 Å (1.14 nm), respectively. Disuccinimidyl tartarate, the shortest cross-linker, was found to link Lys175 in one subunit to Lys183 in another, whereas the longer reagents linked Lys183 in one subunit to the same residue in a neighboring subunit. The difference in cross-linking results reflects the different lengths of the cross-linking reagents, and provides distance constraints that can be used to build a structural model of the interactions.

4. Capsid Protein Characterization by Protein Mass Mapping

Detailed studies have been published on the detection, identification, and characterization of viral coat proteins with protein mass mapping. Although not a non-covalent MS technique, protein mass mapping can also be used to probe the quaternary structure of multi-component assemblies. For example, limited proteolysis and MALDI-MS has been carried out on HRV14 and flock house virus (FHV; Bothner et al., 1998, 1999; Thomas, Bakhtiar, & Suizdak, 2000). Based on the crystal structure, the four proteins that constitute the capsid shell of HRV are VP1, VP2, and VP3, which comprise the viral surface, and VP4, which lies in the interior of the shell at the capsid/RNA interface. FHV also has a protein capsid that consists of two proteins with regions that are distinctly internal and external to the virus surface. Using time-resolved proteolysis, the expected cleavages on the surface-accessible regions of the coat proteins were observed, together with unexpected cleavages in regions that lie in the inside of the viral capsids in the crystal structures. These results suggest that portions of these internal proteins are transiently exposed on the viral surface, and indeed that virus capsids are highly dynamic. The effect(s) of capsid-binding anti-viral agents on the dynamics of the capsids was also studied. In the case of HRV, proteolysis was inhibited dramatically in the presence of

such anti-viral agents, presumably because the drugs cover the hydrophobic binding pockets and these are also the site of cell surface receptor attachment.

Characterization of the monomeric protein building blocks of viral capsids has been achieved by MALDI-MS mass measurements for TMV, MS2, and Venezuelan equine encephalitis (VEE; monomeric masses 13–31 kDa) in the low femtomole range (Thomas, Falk, & Fenselau, 1998). Also, direct detection of MS2 in *E. coli* lysates has been performed, using ESI-tandem ion trap MS analyses on the multiply charged coat protein monomer ions (Cargile, McLucky, & Stephenson, 2001). Proteomics-style mass mapping has been used to identify naturally occurring mutants of HRV and TMV, and a modification on the HRV capsid protein (Thomas, Bakhtiar, & Suizdak, 2000) whereas a non-enzymatic proteomics strategy that employed microwave digestion under acidic conditions has been demonstrated successfully with the MS2 coat protein (Swatkoski et al., 2007). Enzyme-accessibility experiments, using nano-liquid chromatography ESI-MS, have been used to probe the dynamic nature of the HRV capsid in the presence of potential anti-viral compounds (Reisdorph et al., 2003); whereas some anti-virals stabilized the capsid against proteolysis, others had the opposite effect.

III. SUMMARY OF THE STUDY OF VIRUSES BY MASS SPECTROMETRY

Understanding how individual biomolecular components interact and assemble into functional complexes at the molecular level is of fundamental importance to our knowledge of biological processes. With this goal in mind, the simple definition of a virus as a protein shell that surrounds the virus genome does not do justice to the true complexity of the molecular processes involved in virion assembly. Advances in virology and the design of anti-viral therapeutics rely strongly on an appreciation of the viral replication cycle and particularly of capsid architecture and assembly mechanisms. This review has set out to show that the application of MS to probe the assembly, structure, and function of viruses is proving fruitful, and does indeed represent a reliable, highly sensitive, and viable alternative to EM and X-ray crystallography with the significant advantages of being able to monitor assembly reactions on-line, in real-time, and permit the observation of individual components in co-populated ensembles. The viruses studied to date with MS techniques range from the mass measurement of intact virus capsids, through the characterization of assembly intermediates, to the use of HDX and chemical cross-linking to pin-point localized interactions that are important in capsid assembly, are summarized in Table 1. With on-going improvements in MS technologies and the advent of coupled separation techniques such as IMS-MS (Giles et al., 2004; Pringle et al., 2007), the input into this field is certain to flourish.

IV. ABBREVIATIONS

ATP	adenosine triphosphate
Bis-ANS	1,1'-bis(4-anilino)naphthalene-5,5'-disulfonic acid

CID	collision-induced dissociation
DNA	deoxyribonucleic acid
EM	electron microscopy
ESI	electrospray ionization
FT-ICR	Fourier transform ion cyclotron resonance
GEMMA	gas-phase electrophoretic mobility mass analyzer
HDX	hydrogen-deuterium exchange
HPLC	high-pressure liquid chromatography
IMS	ion-mobility spectrometry
MALDI	matrix-assisted laser desorption/ionization
mRNA	messenger ribonucleic acid
MS	mass spectrometry
MS/MS	tandem mass spectrometry
<i>m/z</i>	mass-to-charge ratio
NMR	nuclear magnetic resonance
Rf	radio-frequency
RNA	ribonucleic acid
RNAi	ribonucleic acid interference
<i>T</i>	triangulation number
TEM	transmission electron microscopy
Tof	time-of-flight analyzer

ACKNOWLEDGMENTS

We acknowledge with gratitude the BBSRC, the Leverhulme Trust, and the Wellcome Trust for funding. We also thank colleagues past and present for the many constructive discussions and collaborations we have had over the years, and in particular Prof. Mark Harris and Dr. Roman Tuma for reading this manuscript.

REFERENCES

- Ashcroft AE. 2005a. Recent developments in electrospray ionisation mass spectrometry: Noncovalently bound protein complexes. *Nat Prod Rep* 22:452–464.
- Ashcroft AE, Lago H, Macedo JMB, Horn WT, Stonehouse NJ, Stockley PG. 2005b. Engineering thermal stability in RNA phage capsids via disulphide bonds. *J Nanosci Nanotechnol* 5:2034–2041.
- Bacher G, Szymanski WW, Kaufmann SL, Zollner P, Blaas D, Allmaier G. 2001. Charge-reduced nano electrospray ionization combined with differential mobility analysis of peptides, proteins, glycoproteins, noncovalent protein complexes and viruses. *J Mass Spectrom* 36:1038–1052.
- Bardi JS. 2007. Structure of HIV capsid protein reveals potential weakness at inner core of virus. *News & Views*, October 15, Vol 7, Issue 30, Online weekly published by the Scripps Institute.
- Benner WH. 1997. A gated electrostatic ion trap to repetitiously measure the charge and *m/z* of large electrospray ions. *Anal Chem* 69:4162–4168.
- Berthet-Colominas C, Cuillel M, Kock MHJ, Vachette P, Jacrot B. 1987. Kinetic study of the self assembly of brome mosaic virus capsid. *Eur Biophys J* 15:159–168.
- Bhyravbhatla B, Watowich SJ, Caspar DL. 1998. Refined atomic model of the four-layer aggregate of the tobacco mosaic virus coat protein at 2.4 Å resolution. *Biophys J* 74:604–615.
- Bier M. 2007. Mass spectrometry tackles big problem: ACS Fall meeting, Boston. Cambridge, UK: Chemistry World, Royal Society of Chemistry. p 12.

- Bothner B, Dong XF, Bibbs L, Johnson JE, Siuzdak G. 1998. Evidence of viral capsid dynamics using limited proteolysis and mass spectrometry. *J Biol Chem* 273:673–676.
- Bothner B, Schneemann A, Marshall D, Redy V, Johnson JE, Siuzdak G. 1999. Crystallographically identical virus capsids display different properties in solution. *Nat Struct Biol* 6:114–116.
- Broo K, Wei J, Marshall D, Brown F, Smith TJ, Johnson JE, Schneemann A, Siuzdak G. 2001. Viral capsid mobility: A dynamic conduit for inactivation. *Proc Natl Acad Sci USA* 98:2274–2277.
- Cargile BJ, McLuckey SA, Stephenson JL. 2001. Identification of bacteriophage MS2 coat protein from *E. coli* lysates via ion trap collisional activation of intact protein ions. *Anal Chem* 73:1277–1285.
- Caspar DL. 1956. Structure of bushy stunt virus. *Nature* 177:475–476.
- Caspar DL, Klug A. 1962. Cold Spring Harbor Symposium Quantitative Biology. New York: Cold Spring Harbor Press. pp 1–24.
- Chowdhury SK, Katta V, Chait BT. 1990. Probing conformational changes in proteins by mass spectrometry. *J Am Chem Soc* 112:9012–9013.
- Cockburn JJ, Abrescia NG, Grimes JM, Sutton GC, Diprose JM, Benevides JM, Thomas GJ, Bamford DH, Stuart DI. 2004. Membrane structure and interactions with protein and DNA in bacteriophage P RD1. *Nature* 432:122–125.
- Crick FHC, Watson JD. 1956. Structure of small viruses. *Nature* 177:473–475.
- Cyrklaff M, Risco C, Fernandez JJ, Jimenez MV, Esteban M, Baumeister W, Carrascosa JL. 2005. Cryo-electron tomography of vaccinia virus. *Proc Natl Acad Sci USA* 102:2772–2777.
- Dokland T, Murialdo H. 1993. Structural transitions during maturation of bacteriophage lambda capsids. *J Mol Biol* 233:682–694.
- Engen JR, Smith DL. 2000. Investigating the higher order structure of proteins: Hydrogen exchange, proteolytic fragmentation & mass spectrometry. *Meth Mol Biol* 146:95–112.
- Fabris D, Chaudhari P, Hagan N, Turner K. 2007. Functional investigations of retroviral protein ribonucleic acid complexes by nanospray Fourier transform ion cyclotron resonance mass spectrometry. *Eur J Mass Spectrom* 13:29–33.
- Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM. 1989. Electrospray ionization for mass spectrometry of large biomolecules. *Science* 246:64–71.
- Forkine A, Chipman PR, Leiman PG, Mesyanzhinov VV, Rao VB, Rossmann MG. 2004. Molecular architecture of the prolate head of bacteriophage T4. *Proc Natl Acad Sci USA* 101:6003–6008.
- Fry E, Acharya R, Stuart D. 1993. Methods used in the structure determination of foot-and-mouth disease virus. *Acta Cryst* A49:45–55.
- Fu C-y, Prevelige PE. 2006. Dynamic motions of free and bound Phi29 scaffolding protein determined by hydrogen deuterium exchange mass spectrometry. *Prot Sci* 15:731–743.
- Fuerstenau SD, Benner WH, Thomas JJ, Brugidou C, Bothner B, Siuzdak G. 2001. Mass spectrometry of an intact virus. *Angew Chem Int Ed* 40:541–544.
- Galisto ML, King J. 1993. Conformational transformation in the protein lattice of phage P22 procapsids. *Biophys J* 65:227–235.
- Gan L, Speir JA, Conway JF, Lander G, Cheng N, Firek BA, Hendrix RW, Duda RL, Liljas L, Johnson JE. 2006. Capsid conformational sampling in HK97 maturation visualized by X-ray crystallography and cryo-EM. *Structure* 14:1655–1665.
- Ganem B, Li YT, Henion JD. 1991. Detecting of noncovalent receptor-ligand complexes by mass spectrometry. *J Am Chem Soc* 113:6294–6296.
- Ganser-Pornillos BK, Cheng A, Yeager M. 2007. Structure of full length HIV-1 CA; a model for the mature capsid lattice. *Cell* 131:70–79.
- Giles K, Pringle SD, Worthington KR, Little D, Wildgoose JL, Bateman RH. 2004. Applications of a travelling wave-based radio-frequency-only stacked ring ion guide. *Rapid Commun Mass Spectrom* 18:2401–2414.
- Golmohammadi R, Fridborg K, Bundule M, Valegard K, Liljas L. 1996. The crystal structure of bacteriophage Q beta at 3.5 Å resolution. *Structure* 4:543–554.
- Golmohammadi R, Valegard K, Fridborg K, Liljas L. 1993. The refined structure of bacteriophage MS2 at 2.8 Å resolution. *J Mol Biol* 234:620–639.
- Grimes JM, Burroughs JN, Gouet P, Diprose JM, Malby R, Zientara S, Mertens PP, Stuart DI. 1998. The atomic structure of the bluetongue virus core. *Nature* 395:470–478.
- Guo P, Lee TJ. 2007. Viral nanomotors for packaging of dsDNA and dsRNA. *Mol Microbiol* 64:886–903.
- Haasnoot J, Westerhout EM, Berkhout B. 2007. RNA interference against viruses: Strike and counterstrike. *Nat Biotechnol* 25:1435–1443.
- Hadfield AT, Diana GD, Rossman MG. 1999. Analysis of three structurally related antiviral compounds in complex with human rhinovirus 16. *Proc Natl Acad Sci USA* 96:14730–14735.
- Hanson CL, Robinson CV. 2004. Protein-nucleic acid interactions and the expanding role of mass spectrometry. *J Biol Chem* 279:24907–24910.
- Helgstrand C, Wikoff WR, Duda RL, Hendrix RW, Johnson JE, Liljas L. 2003. The refined structure of a protein catenane: The HK97 bacteriophage capsid at 3.44 Å resolution. *J Mol Biol* 334:885–899.
- Hendrix RW. 1978. Symmetry mismatch and DNA packaging in large bacteriophages. *Proc Natl Acad Sci USA* 75:4779–47783.
- Hendrix RW, Smith MCM, Burns RN, Ford ME, Hatfull GF. 1999. Evolutionary relationships among diverse bacteriophages and prophages: All the world's a phage. *Proc Natl Acad Sci USA* 96:2192–2197.
- Hewat EA, Blaas D. 1996. Structure of a neutralizing antibody bound bivalently to human rhinovirus 2. *EMBO J* 15:1515–1523.
- Hogan CJ, Kettleson EM, Ramaswami B, Chen D-R, Biswas P. 2006. Charge reduced electrospray size spectrometry of mega- and gigadalton complexes: Whole viruses and virus fragments. *Anal Chem* 78:844–852.
- Hugel T, Michaelis J, Hetherington CL, Jardine PJ, Grimes S, Walter JM, Falk W, Anderson DL, Bustamante C. 2007. Experimental test of connector rotation during DNA packaging into bacteriophage phi29 capsids. *PLoS Biol* 5:e59.
- Kaddis CS, Lomeli SH, Yin S, Berhane B, Apostol MI, Kickhoefer VA, Rome LH, Loo JA. 2007. Sizing large proteins and protein complexes by electrospray ionization mass spectrometry and ion mobility. *J Am Soc Mass Spectrom* 18:1206–1216.
- Kang S, Hawkrigde AM, Johnson KL, Muddiman DC, Prevelige PE. 2006. Identification of subunit-subunit interactions in bacteriophage P22 procapsids by chemical cross-linking and mass spectrometry. *J Prot Res* 5:370–377.
- Kang S, Prevelige PE. 2005. Domain study of bacteriophage P22 coat protein and characterisation of the capsid lattice transformation by hydrogen/deuterium exchange. *J Mol Biol* 347:935–948.
- Karas M, Hillenkamp F. 1988. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem* 60:2299–2301.
- Katta V, Chait BT. 1991. Observation of the heme-globin complex in native myoglobin by mass spectrometry. *J Am Chem Soc* 113:8534–8535.
- Kostyuchenko VA, Chipman PR, Leiman PG, Arisaka F, Mesyanzhinov VV, Rossmann MG. 2005. The tail structure of bacteriophage T4 and its mechanism of contraction. *Nat Struct Mol Biol* 12:810–813.
- Lam TT, Lanman JK, Emmett MR, Hendrickson CL, Marshall AG, Prevelige PE. 2002. Mapping of protein:protein contact surfaces by hydrogen/deuterium exchange, followed by on-line high-performance liquid chromatography-electrospray ionisation Fourier transform ion cyclotron resonance mass analysis. *J Chrom A* 982:85–95.

- Lander GC, Tang L, Casjens SR, Gilcrease EB, Prevelige PE, Poliakov A, Potter CS, Carragher B, Johnson JE. 2006. The structure of an infectious P22 virion shows the signal for headful DNA packaging. *Science* 312:1791–1795.
- Lanman JK, Lam TT, Barnes S, Sakalian M, Emmett MR, Marshall AG, Prevelige PE. 2003. Identification of novel interactions in HIV-1 capsid protein assembly by high resolution mass spectrometry. *J Mol Biol* 325:759–772.
- Lanman JK, Lam TT, Emmett MR, Marshall AG, Sakalian M, Prevelige PE. 2004. Key interactions in HIV-1 maturation identified by hydrogen-deuterium exchange. *Nat Struct Mol Biol* 11:676–677.
- Liddington RC, Yan Y, Moulai J, Sahli R, Benjamin TL, Harrison SC. 1991. Structure of simian virus 40 at 3.8-Å resolution. *Nature* 354:278–284.
- Lin T, Chen Z, Usha R, Stauffacher CV, Dai J-B, Schmidt T, Johnson JE. 1999. The refined crystal structure of cowpea mosaic virus at 2.8 Å resolution. *Virology* 265:20–34.
- Lisal J, Kainov DE, Lam TT, Emmett MR, Wei H, Gottlieb P, Marshall AG, Tuma R. 2006. Interaction of packaging motor with the polymerase complex of dsRNA bacteriophage. *Virology* 351:73–79.
- Lisal J, Lam TT, Kainov DE, Emmett MR, Marshall AG, Tuma R. 2005. Functional visualisation of viral molecular motor by hydrogen-deuterium exchange reveals transient states. *Nat Struct Mol Biol* 12:460–466.
- Loo JA. 1997. Studying noncovalent protein complexes by electrospray ionization mass spectrometry. *Mass Spectrom Rev* 16:1–23.
- Lorenzen K, Uetrecht C, Versluis C, Heck AJR. 2007. Virus assembly studied by macromolecular (tandem) mass spectrometry and ion mobility. (TOD), 55th ASMS Conference on Mass Spectrometry & Allied Topics, June 3rd–7th 2007, Indianapolis, In, USA.
- Lucas RW, Larson SB, McPherson A. 2002. The crystallographic structure of brome mosaic virus. *J Mol Biol* 317:95–108.
- Mancini EJ, Kainov DE, Grimes JM, Tuma R, Bamford DH, Stuart DI. 2004. Atomic snapshots of an RNA packaging motor reveal conformational changes linking ATP hydrolysis to RNA translocation. *Cell* 118:743–755.
- Marasco WA, Sui J. 2007. The growth and potential of human antiviral monoclonal antibody therapeutics. *Nat Biotechnol* 25:1421–1434.
- Miller MD, Hazuda DJ. 2001. New antiretroviral agents: Looking beyond protease and reverse transcriptase. *Curr Opin Microbiol* 4:535–539.
- Nie Z, Tzeng Y-K, Chang C, Chiu CC, Chang C-Y, Chang C-M, Tao M-H. 2006. Microscopy-based mass measurement of a single whole virus in a cylindrical ion trap. *Angew Chem Int Ed* 45:8131–8134.
- Noorden RV. 2007. Mass spectrometry tackles big problem. *Chemistry World*, Vol. 4, Issue 10, p.12 (published by the Royal Society of Chemistry, Cambridge, UK).
- Orlova EV, Dube P, Backmann E, Zemlin F, Lurz R, Trautner TA, Tavares P, van Heel M. 1999. Structure of the 13-fold symmetric portal protein of bacteriophage SPP1. *Nat Struct Biol* 6:842–846.
- Paredes A, Alwell-Warda K, Weaver SC, Chiu W, Watowich SJ. 2003. Structure of isolated nucleocapsids from Venezuelan equine encephalitis virus and implications for assembly and disassembly of enveloped virus. *J Virol* 77:659–664.
- Poliakov A, van Duijn E, Lander G, Fu C-y, Johnson JE, Prevelige PE, Heck AJR. 2007. Macromolecular mass spectrometry and electron microscopy as complementary tools for investigation of the heterogeneity of bacteriophage portal assemblies. *J Struct Biol* 157:371–383.
- Pringle SD, Giles K, Wildgoose JL, Slade SE, Thalassinos K, Bateman RH, Bowers MT, Scrivens JH. 2007. An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument. *Int J Mass Spectrom* 261:1–12.
- Qu C, Liljas L, Opalka N, Brugidou C, Yeager M, Beachy RN, Fauquet CM, Johnson JE. 2000. 3D domain swapping modulates the stability of members of an icosahedral virus group. *Structure* 8:1095–1103.
- Rayment I, Baker TS, Caspar DL, Murakami WT. 1982. Polyoma virus capsid structure at 22.5 Å resolution. *Nature* 295:110–115.
- Reisdorph N, Thomas JJ, Katpally U, Chase E, Harris K, Siuzdak G, Smith TJ. 2003. Human rhinovirus capsid dynamics is controlled by canyon flexibility. *Virology* 314:34–44.
- Robinson MA, Wood JP, Capaldi SA, Baron AJ, Gell C, Smith DA, Stonehouse NJ. 2006. Affinity of molecular interactions in the bacteriophage phi29 DNA packaging motor. *Nucleic Acids Res* 34:2698–2709.
- Schultz JC, Hack CA, Benner WH. 1998. Mass determination of megadalton-DNA electrospray ions using charge detection mass spectrometry. *J Am Soc Mass Spectrom* 9:305–313.
- Schultz JC, Hack CA, Benner WH. 1999. Polymerase chain reaction products analyzed by charge detection mass spectrometry. *Rapid Commun Mass Spectrom* 13:15–20.
- Shepherd CM, Borelli IA, Lander G, Natarajan P, Siddavanahalli V, Bajaj C, Johnson JE, Brooks CL, Reddy VS. 2006. A relational database for structural virology. *Nucleic Acids Res* 34:D386–D389.
- Simpson AA, Tao Y, Leiman PG, Badasso MO, He Y, Jardine PJ, Olson NH, Morais MC, Grimes S, Anderson DL, Baker TS, Rossmann MG. 2000. Structure of the bacteriophage phi29 DNA packaging motor. *Nature* 408:745–750.
- Singh P, Gonzalez MJ, Manchester M. 2006. Viruses and their uses in nanotechnology. *Drug Dev Res* 67:23–41.
- Siuzdak G. 1996. Mass spectrometry and viral analysis. *Chem Biol* 9:707–715.
- Smith DE, Tans SJ, Smith SB, Grimes S, Anderson DL, Bustamante C. 2001. The bacteriophage phi29 portal motor can package DNA against a large internal force. *Nature* 413:748–752.
- Smith RD, Cheng X, Bruce JE. 1994. Trapping, detection and reaction of very large single molecular ions by mass spectrometry. *Nature* 369:137–139.
- Sobott F, Hernandez H, McCammon MG, Tito MA, Robinson CV. 2002. A tandem mass spectrometer for improved transmission and analysis of large macromolecular assemblies. *Anal Chem* 74:1402–1407.
- Speir JA, Munshi S, Wang G, Baker TS, Johnson JE. 1995. Structures of the native and swollen forms of cowpea chlorotic mottle virus determined by X-Ray crystallography and cryo-electron microscopy. *Structure* 3:63–78.
- Stockley PG, Ashcroft AE, Francese S, Thompson GS, Ransom N, Smith AM, Homans SW, Stonehouse NJ. 2005. Dissecting the fine details of assembly of a $T=3$ capsid. *J Theor Med* 6:119–125.
- Stockley PG, Rolfsson O, Thompson GS, Basnak G, Francese S, Stonehouse NJ, Homans SW, Ashcroft AE. 2007. A simple, RNA-mediated allosteric switch controls the pathway to formation of a $T=3$ capsid. *J Mol Biol* 369:541–552.
- Suck D, Rayment I, Johnson JE, Rossmann MG. 1978. The structure of southern bean mosaic virus at 5 Å resolution. *Virology* 85:187–197.
- Sundquist WI, Hill CP. 2007. How to assemble a capsid. *Cell* 131:70–79.
- Swatkoski S, Russell S, Edwards N, Fenselau C. 2007. Analysis of a model virus using residue-specific chemical cleavage and MALDI-ToF mass spectrometry. *Anal Chem* 79:654–658.
- Tanaka K, Waki H, Ido Y, Akita S, Yoshida T. 1988. Protein and polymer analyses up to m/z 100 000 by laser ionization time-of flight mass spectrometry. *Rapid Commun Mass Spectrom* 2:151–153.
- Thomas JJ, Bakhtiar R, Siuzdak G. 2000. Mass spectrometry in viral proteomics. *Acc Chem Res* 33:179–187.
- Thomas JJ, Bothner B, Traina J, Benner WH, Siuzdak G. 2004. Protein structure characterisation with mass spectrometry. *Spectroscopy* 18: 31–36.
- Thomas JJ, Falk B, Fenselau C. 1998. Viral characterisation by direct analysis of capsid proteins. *Anal Chem* 70:3863–3867.
- Tito MA, Tars K, Vølgard K, Hajdu J, Robinson CV. 2000. Electrospray time-of-flight mass spectrometry of the intact MS2 virus capsid. *J Am Chem Soc* 122:3550–3551.

- Tuma R, Coward LU, Kirk MC, Barnes S, Prevelige PE. 2001. Hydrogen-deuterium exchange as a probe of folding and assembly in viral capsids. *J Mol Biol* 306:389–396.
- van Berkel WJH, van den Heuvel RHH, Versluis C, Heck AJR. 2000. Detection of intact megaDalton protein assemblies of vanillyl-alcohol oxidase by mass spectrometry. *Prot Sci* 9:435–439.
- van Breukelen B, Barendregt A, Heck AJR, van den Heuvel RHH. 2006. Resolving stoichiometries and oligomeric states of glutamate synthase protein complexes with curve fitting and simulation of electrospray mass spectra. *Rapid Commun Mass Spectrom* 20:2490–2496.
- van den Heuvel RHH, van Duijn E, Mazon H, Synowsky SA, Lorenzen K, Versluis C, Brouns SJ, Langridge D, van der Oost J, Hoyes J, Heck AJR. 2006. Improving the performance of a quadrupole time-of-flight instrument for macromolecular mass spectrometry. *Anal Chem* 78: 7473–7483.
- Verdaguer N, Blaas D, Fita I. 2000. Structure of human rhinovirus serotype 2 (HRV2). *J Mol Biol* 300:1179–1194.
- Verentchikov AN, Ens W, Standing KG. 1994. Reflecting time-of-flight mass spectrometer with an electrospray ion source and orthogonal extraction. *Anal Chem* 66:126–133.
- Wang L, Lane LC, Smith DL. 2001. Detecting structural changes in viral capsids by hydrogen exchange and mass spectrometry. *Prot Sci* 10: 1234–1243.
- Wang L, Smith DL. 2005. Capsid structure and dynamics of a human rhinovirus probed by hydrogen exchange mass spectrometry. *Prot Sci* 14:1661–1672.
- Wang Q, Kaltgrad E, Lin TW, Johnson JE, Finn MG. 2002. Natural supramolecular building blocks: Wild-type cowpea mosaic virus. *Chem Biol* 9:805–811.
- Wenzel RJ, Matter U, Schultheis L, Zenobi R. 2005. Analysis of megadalton ions using cryodetection MALDI time of flight mass spectrometry. *Anal Chem* 77:4329–4337.
- Wilm MS, Mann M. 1994. Electrospray and Taylor-Cone theory, Dole's beam of macromolecules at last? *Int J Mass Spectrom Ion Proc* 136:167–180.
- Winkler FK, Schutt CE, Harrison SC, Bricogne G. 1977. Tomato bushy stunt virus at 5.5-Å resolution. *Nature* 265:509–513.
- Wynne SA, Crowther RA, Leslie AGW. 1999. The crystal structure of the human hepatitis b virus capsid. *Mol Cell* 3:771–780.
- Zhang Z, Smith RD. 1993. Determination of amide hydrogen exchange by mass spectrometry: A new tool for protein structure elucidation. *Prot Sci* 2:522–531.
- Zlotnick A. 2004. Viruses and the physics of soft condensed matter. *Proc Natl Acad Sci USA* 101:15549–15550.
- Zlotnick A, Palmer I, Kaufman JD, Stahl SJ, Steven AC, Wingfield PT. 1999. Separation and crystallisation of $T=3$ and $T=4$ icosahedral complexes of the hepatitis B virus core protein. *Acta Cryst D55*:717–720.
- Zlotnick A, Stray SJ. 2003. How does your virus grow? Understanding and interfering with virus assembly. *Trends Biotechnol* 21:536–542.
- Zubieta CS, Schoehn G, Chroboczek J, Cusack S. 2005. The structure of human adenovirus 2 penton. *Mol Cell* 17:121–135.

Victoria L. Morton is a post-doctoral research fellow at the University of Leeds working on the characterisation of virus assembly using mass spectrometry. She obtained her B.Sc. in Microbiology in 2001 from the University of Edinburgh and Ph.D. in Biochemistry, studying the kinetics of protein folding, from the University of Leeds in 2006.

Peter G. Stockley is Professor of Biological Chemistry, and Director of both the Astbury Centre for Structural Molecular Biology and the Inter-disciplinary Institute for Bionanosciences at the University of Leeds. He obtained his B.Sc. in Chemistry from Imperial College, London and his Ph.D. in Biochemistry from the University of Cambridge. Stockley joined the academic staff of the University of Leeds in 1986 after postdoctoral training at Harvard University. He has a long history of work with protein-nucleic acid complexes and on the molecular mechanisms of self-assembly of simple RNA viruses.

Nicola J. Stonehouse is a Senior Lecturer at the University of Leeds. Following a 4-year MRC Career Development Fellowship, Stonehouse was appointed as a lecturer and promoted to Senior Lecturer in August 2006. Her laboratory employs a reductionist approach to study the details of molecular interactions in several viral systems. Current projects include the use of RNA aptamers to study protein-protein interactions (focussed on the human papilloma virus oncogenic proteins and the genome replication complex from foot-and-mouth disease virus), and bacteriophage assembly and genome packaging.

Alison E. Ashcroft is a Reader in Biomolecular Mass Spectrometry at the University of Leeds. She obtained her B.Sc. in Chemistry from the University of Liverpool and Ph.D. in Organic Chemistry from the University of Manchester. Ashcroft joined the University of Leeds in 1997 after post-doctoral training in Switzerland and then 14 years as an industrial mass spectrometrist. Her research focuses on protein folding and conformer characterisation, the assembly of macromolecular complexes, and protein structural elucidation. She is currently developing ion mobility spectrometry coupled to mass spectrometry to separate co-populated protein species.