Published in final edited form as:

Rapid Commun Mass Spectrom. 2010 October 30; 24(20): 3033–3042. doi:10.1002/rcm.4732.

Determining the topology of virus assembly intermediates using ion mobility spectrometry–mass spectrometry

Tom W. Knapman[#], Victoria L. Morton[#], Nicola J. Stonehouse, Peter G. Stockley, and Alison E. Ashcroft^{*}

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

[#] These authors contributed equally to this work.

Abstract

We have combined ion mobility spectrometry–mass spectrometry with tandem mass spectrometry to characterise large, non-covalently bound macromolecular complexes in terms of mass, shape (cross-sectional area) and stability (dissociation) in a single experiment. The results indicate that the quaternary architecture of a complex influences its residual shape following removal of a single subunit by collision-induced dissociation tandem mass spectrometry. Complexes whose subunits are bound to several neighbouring subunits to create a ring-like three-dimensional (3D) architecture undergo significant collapse upon dissociation. In contrast, subunits which have only a single neighbouring subunit within a complex retain much of their original shape upon complex dissociation. Specifically, we have determined the architecture of two transient, on-pathway intermediates observed during *in vitro* viral capsid assembly. Knowledge of the mass, stoichiometry and cross-sectional area of each viral assembly intermediate allowed us to model a range of potential structures based on the known X-ray structure of the coat protein building blocks. Comparing the cross-sectional areas of these potential architectures before and after dissociation provided tangible evidence for the assignment of the topologies of the complexes, which have been found to encompass both the 3-fold and the 5-fold symmetry axes of the final icosahedral viral shell. Such insights provide unique information about virus assembly pathways that could allow the design of anti-viral therapeutics directed at the assembly step. This methodology can be readily applied to the structural characterisation of many other non-covalently bound macromolecular complexes and their assembly pathways.

Approximately half of all virus capsids are spherical in shape and consist of a defined number of copies of a single, or very few, type(s) of coat protein subunit assembled into a protective shell (capsid) based on icosahedral surface lattices.¹ Such lattices tessellate the surface of a sphere and their allowed configurations have been predicted using quasi-equivalence theory.² This permits otherwise identical proteins to adopt different symmetry-related conformations, defined by the triangulation number (T), at appropriate positions in

^{*}Correspondence to: A. E. Ashcroft, Astbury Centre for Structural Molecular Biology, Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, UK. a.e.ashcroft@leeds.ac.uk.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article.

the icosahedral shell, allowing larger capsids to be built from multiple copies of smaller coat proteins, and thus increasing the genetic economy within the viral genome.³ Single-stranded (ss) RNA viruses represent one of the largest groups of viruses and include major human, animal and plant pathogens.⁴ We have made major contributions to the study of their capsid assembly pathways *in vitro* using a model ssRNA virus, the bacteriophage MS2.^{5–7}

The MS2 phage capsid consists of a $T = 3$ icosahedral protein lattice comprising 180 copies of the viral coat protein (mass 13.7 kDa; denoted CP) in the form of 90 non-covalent dimers (CP₂) (Fig. 1). The quasi-symmetry is accommodated by the CP₂ adopting three distinct conformations, defined by the geometry of the polypeptide loop connecting the F and G β -strands in each monomer: in the B subunit the FG-loop is folded back towards the protein subunit, while it is fully extended but in slightly differing orientations in A and C conformers.^{8,9} The FG-loops define the type of symmetry axis formed in the icosahedral shell: five B-type loops surround the particle 5-fold axis and three inter-digitating A-type and C-type loops create the 3-fold axis. Understanding capsid assembly requires the molecular mechanisms of protein conformer switching and its impact on initiation and capsid growth to be determined. A 19-nucleotide RNA stem-loop from within the genomic sequence, denoted TR, is a known initiator of assembly both *in vitro* and, presumably, *in vivo*.^{10,11} Addition of TR (mass 6 kDa) to the CP₂ *in vitro* results in the rapid formation of a [CP₂:TR] complex (mass 33.4 kDa) and initiation of the assembly process. We have shown that the RNA-free CP₂ in solution is symmetrical (i.e., C/C-like) and that binding of TR promotes a conformational switch to an asymmetric (A/B-like) dimer.⁷

By monitoring the TR-induced assembly of the capsid from the protein building blocks using electrospray ionisation mass spectrometry (ESI-MS),^{5–7,12,13} we showed previously that both C/C- and A/B-like dimers are required for the efficient assembly of $T = 3$ capsids and that the building block is a CP₂ unit.⁷ More recently, we reported the mass and stoichiometry of two significantly populated non-covalently bound intermediate species detected by ESI-MS *en route* to capsid formation: one consisting of six CP₂ subunits, three with bound RNA stem-loops and three without ([3(CP₂:TR) + 3CP₂]; mass 182.9 kDa; denoted hexamer) and the other consisting of ten CP₂ subunits, five with bound RNA stem-loops and five without ([5(CP₂:TR) + 5CP₂]; mass 304.8 kDa; denoted decamer).^{6,7} Analysis of the time-dependent concentrations of these intermediates has allowed us to derive a kinetic model that provides the first accurate description of an assembly pathway that incorporates the influence of the genomic RNA being packaged.⁶ This model is consistent with the hexamer and decamer corresponding to sub-fragments of the final $T = 3$ capsid encompassing a 3-fold axis or an extended 5-fold axis. Here we validate this assumption by determining the topology of the intermediates directly.

Ion mobility spectrometry (IMS) coupled to ESI-MS (ESI-IMS-MS) is a unique and powerful tool that can separate and analyse non-covalently bound macromolecular complexes and provide both mass and shape (via collision cross-sectional area, Ω) information on individual species within an ensemble in a single, rapid, experiment.^{14–18} Here we have gone one step further to exploit the powerful combination of IMS-MS coupled to collision-induced dissociation (CID) tandem mass spectrometry (MS/MS)^{19–21} in a single experiment (namely, ESI-MS-CID-IMS-MS) to provide direct insights into the quaternary

architecture of these two transient, co-populated viral assembly intermediates in a real-time assembly reaction.^{5,7} In order to relate the gas-phase dissociation characteristics of the non-covalently bound intermediates to their native topologies, we have compared their behaviour following collision-induced expulsion of a single subunit with that of other complexes of known structure. There are characteristic changes in the observed cross-sectional areas of some species with ring-like topologies which allow us to assign the viral intermediates.

EXPERIMENTAL

Sample preparations

Equine cytochrome c, horse heart myoglobin, bovine serum albumin (BSA), GroEL, human haemoglobin, and other proteins and reagents were obtained from Sigma-Aldrich (Gillingham, UK). HPLC-grade solvents were purchased from Fisher Scientific (Loughborough, UK). GroEL was desalted²² and buffer-exchanged into 200 mM ammonium acetate (pH 6.9) at a final concentration of 2 μ M. MHC-1 was prepared as previously described^{23,24} and dialysed into 40 mM ammonium acetate, pH 6.9, to a final concentration of 20 μ M. Human haemoglobin was dissolved as supplied in 40 mM ammonium acetate, pH 7.5, to a final concentration of 20 μ M.

Wild-type recombinant MS2 coat protein in the form of $T=3$ shells was prepared by over-expression in *E. coli* as previously described.²⁵ These capsids were purified and then treated with glacial acetic acid followed by exchange into 20 mM acetic acid to generate disassembled coat protein dimers (CP₂). Virus assembly reactions were carried out as described previously⁷ in 40 mM ammonium acetate (pH 6.9) at 4°C. Initially, CP₂ was mixed with the 19-nucleotide RNA (TR; 5'-ACA UGA GGA UUA CCC AUG U-3') to a final concentration of 8 μ M to form the 1:1 reaction complex of [CP₂:TR]; an additional aliquot of CP₂ was added to adjust the ratio of CP₂:TR from 1:1 to 2:1 in order for assembly to proceed.

Mass spectrometry

All ESI-IMS-MS and ESI-MS-CID-IMS-MS analyses were performed on a Synapt HDMS (Waters UK Ltd, Manchester, UK) orthogonal acceleration quadrupole-time-of-flight (Q-ToF) mass spectrometer with a travelling wave IMS (TWIMS) device situated between the two analysers together with MS/MS facilities both before (trap region) and after (transfer region) the IMS drift cell.^{26,27} For ESI-IMS-MS analyses the quadrupole analyser was operated in radio-frequency (Rf) mode only and the nano-ESI-generated ions were separated by the IMS device before being m/z analysed by the ToF analyser. For ESI-MS-CID-IMS-MS analyses, the nano-ESI-generated precursor ions of interest were selected by the quadrupole analyser before undergoing CID MS/MS in the trap region collision cell. The product ions were then separated by the IMS device before being m/z analysed by the ToF analyser.

The Synapt HDMS was equipped with an automated nano-ESI inlet (Triversa NanoMate, Advion Biosciences Inc., Ithaca, NY, USA) which was used with a nano-ESI capillary voltage of 1.8 kV and a nitrogen nebulising gas pressure of 0.4 psi (~0.03 bar). The sample

cone voltage was optimised at 100–175 V with nitrogen as the nebulising gas. To maintain the gas-phase ions originating from biomolecular complexes intact, the first vacuum region after the sample cone was operated at a pressure of 6 mbar.²⁸ Argon was used as the MS/MS collision gas in both the trap and the transfer ion guides and the pressure in these regions was 4.12×10^{-2} mbar. CID MS/MS was performed by raising the collision energy in the trap collision cell until product ions could be detected. The IMS cell was filled with nitrogen (5.38×10^{-1} mbar). The IMS wave height and velocity were optimised at 10 V (fixed) and 250 m s^{-1} , respectively, for the analysis of the denatured protein calibrants and the MS2, GroEL, human haemoglobin and BSA samples, and at 4–12 V (ramped) and 250 m s^{-1} for the MHC-1 sample. The duty cycle of the ion mobility cell was set at 18 ms for analysis of the protein calibrants and the MHC-1, and at 51 ms for the analysis of the MS2, GroEL, human haemoglobin and BSA samples. The differences in these settings has been shown not to cause significant differences in cross-sectional area measurements.¹⁷ The pressure in the ToF analyser was 2.53×10^{-6} mbar.

Calibration of the drift time cross-section function was achieved under both the fixed and the ramped wave height conditions by the analysis of denatured equine cytochrome c and horse heart myoglobin ($10 \text{ }\mu\text{M}$ in 50:40:10, v/v/v, acetonitrile/water/acetic acid) as described previously.^{29,30} In order to extend the calibration range to allow for the cross-sectional area measurement of larger complexes with lower mobilities, BSA oligomers (dimers, trimers and tetramers) whose collision cross-sections have been measured using conventional IMS-MS ($\Omega = 56.2, 76.3, \text{ and } 93.9 \text{ nm}^2$, respectively)³¹ were used or, alternatively, the standard calibration range was extrapolated with GroEL.¹⁷ For full details of the cross-sectional area calibration, measurements and calculations, see Supporting Information.

Computational cross-sectional area calculations

Theoretical cross-sectional areas were calculated using the recently described in-house developed Monte Carlo algorithm,^{29,30} based on an earlier MOBCAL projection approximation method,³² incorporating optimised interaction radii.²⁹ This method has been described previously as giving accurate theoretical results for large complexes from known structures¹⁷ or coarse-grained models providing that the subunit is accurately represented in the calculation.³³ Coarse-grained model structures were generated on Pymol for each intermediate in a similar fashion to previously described modelling procedures³³ using X-ray crystallographic co-ordinates of the coat protein dimers⁸ as accurate subunits, ensuring that the modelled structures possessed the correct stoichiometry and similar inter-subunit interactions to that observed in the final capsid. Structures of CID MS/MS product ions were generated by selective removal of individual subunits from the X-ray co-ordinates or from modelled structures for the complex of interest.

RESULTS

Determining the topology of virus assembly intermediates

Previously, we have shown that the reaction profile of the MS2 $T=3$ capsid assembly *in vitro*, triggered by adding the 19-nucleotide RNA stem-loop TR to the CP₂, can be monitored in real-time using ESI-MS.^{5–7} Two non-covalently bound, transient capsid

intermediates of unknown topology, the hexamer and decamer, that are only available in small quantities within an ensemble of species, were detected. Here we have employed ESI-IMS-MS and ESI-MS-CID-IMS-MS technologies to gain key structural insights into these intermediates, both of which have been shown to be *en route* to the final capsid (i.e., true intermediates rather than dead-end products or non-specifically bound subunits).⁷ The ESI-IMS-MS spectrum (Fig. 2) recorded during an *in vitro* capsid assembly reaction shows the [CP₂:TR] assembly initiation complex, which forms immediately on mixing the CP₂ with the RNA stem-loop, together with the hexamer and decamer intermediates of 182.9 and 304.8 kDa, respectively. ESI-IMS-MS analysis also allowed the measurement of the collision cross-sectional areas (Ω values) of these co-populated species. To obtain these values, a calibration to align IMS drift times with charge and cross-sectional area was created²⁹ (see Supporting Information). In this case the calibration was extended to cover the cross-sectional areas of the two intermediates, which was achieved either by extrapolation with the chaperonin complex GroEL³⁴ or by calibration with BSA oligomers.³¹ The IMS-measured Ω values were within experimental error regardless of the calibration method used. In addition, a broad peak corresponding to the intact 2.5 MDa capsid appears at ca. m/z 25 000 in the spectrum but the low resolution of this signal precludes any charge state assignments and Ω value measurements (data not shown).

To relate the measured Ω values of these components to the structural arrangement of their subunits, theoretical models of the intermediate species were generated using the CP₂ and [CP₂:TR] crystallographic co-ordinates as accurate subunits. This is similar to previous studies of other macromolecular complexes, where quaternary structures have been modelled using spherical subunits;¹⁵ however, the use of atomistic subunits here provides a more detailed and accurate description of each potential structure for these assembly intermediates. The use of these subunits, in conjunction with an in-house projection approximation algorithm^{29,30} based on an earlier MOBCAL projection approximation method³² with improved speed and ease of file importation (see Supporting Information), was justified by the close agreement (within 1% difference) observed between the experimentally measured and theoretically calculated Ω values for both the CP₂ before RNA stem-loop addition (measured 19.2 nm², calculated 19.4 nm²), and the [CP₂:TR] initiation complex (measured 20.1 nm², calculated 20.0 nm²) (Fig. 3). In addition, although the molecular mass and cross-sectional area of the intact virus capsid could not be measured in the IMS-MS experiments due to its unresolved charge state distribution, which is indicative of sample heterogeneity arising from small molecules and ions forming adducts with the non-covalently bound complex, the calculated Ω value using the in-house algorithm was found to be 606 nm². This compares well with the cross-sectional area of 616 nm² inferred from the 28 nm capsid diameter reported for the crystal structure⁸ and provides further confidence in the modelling methods.

The Ω values for the hexamer and decamer were measured as 80.4 and 111.6 nm², respectively. These values were each compared with an extended, linear arrangement of their proposed stoichiometries and observed to be significantly smaller (~17 and ~40%, respectively), suggesting that both intermediates adopt a compact quaternary structure. Potential intermediate structures were generated from different arrangements of the CP₂ and [CP₂:TR] subunits, adhering to the measured stoichiometry in all cases and also ensuring

that the geometry of the interactions between the subunits matched that observed in the overall capsid structure (PDB accession code 2MS2).⁸ In the case of the hexamer, three potential models, denoted 3-fold ring, crescent and horseshoe, were identified (Fig. 3 and Table 1). All were found to have calculated Ω values within 7% of the measured Ω value of 80.4 nm², so no one species alone could be assigned as the definitive structure on this basis. In the case of the decamer, however, the only potential model structure based on the above criteria was that corresponding to an extended 5-fold ring model, comprising a pentamer of [CP₂:TR] subunits decorated around the periphery with a further five CP₂ subunits, whose calculated Ω value (114.6 nm²) is in good agreement with the measured Ω value (111.6 nm²) (Fig. 3 and Table 1).

Correlating topology to gas-phase dissociation characteristics

To clarify the structural arrangement of the subunits within the hexamer intermediate, we investigated the stability of the virus capsid intermediates using ESI-MS-CID-IMS-MS. The accepted CID MS/MS mechanism of non-covalently bound protein complexes involves the unfolding of a single subunit and its subsequent expulsion from the complex, with concurrent charge transfer from the intact complex to the ejected subunit, leaving a species less one subunit with significantly fewer charges than the parent complex.^{35–41} Using a Synapt HDMS hybrid tandem mass spectrometer, precursor ions were selected by the first, quadrupole, analyser and then fragmented in the collision cell situated in the trap region immediately before the IMS drift cell. IMS separation of the product ions was followed by m/z analysis with the second, ToF, analyser.^{26,27} Thus, the measured Ω values of the intact complexes pre-dissociation (ESI-IMS-MS) could be compared with the measured Ω values of the expelled subunits and residual complexes post-dissociation (ESI-MS-CID-IMS-MS) to investigate any relationship between gas-phase dissociation patterns to original native topology.

We first examined the gas-phase dissociation of a simple two-subunit complex which shows minimal structural change upon dissociation: the major histocompatibility complex class 1 (MHC-1).²⁴ Human MHC-1 is a heterogeneous complex consisting of a 31.9 kDa heavy chain and a 11.8 kDa light chain, with a non-specific peptide subunit (~1 kDa) bound to the heavy chain.⁴² The ESI-IMS-MS spectrum of intact recombinant MHC-1, accompanied by minor signals consistent with the *apo* form of the complex (i.e., without the peptide), the unbound peptide, and the unbound light chain, is shown in Fig. 4(a). The IMS-MS-measured Ω value of the intact MHC-1 complex was 29.0 nm², which differs by only ~4% from the calculated Ω value of 27.8 nm² from its PDB co-ordinates (Table 1). The 12+ charge state ions of intact MHC-1 were subjected to ESI-MS-CID-IMS-MS and the Ω values of the dissociation products determined (Fig. 4(b)). The measured Ω values of the dissociated heavy and light chain subunits were found to be within 7% of calculated Ω values which were modelled on segments of the intact complex assuming no rearrangement of the residual subunits. This indicates that these dissociation products had experienced minimal unfolding, or any other change in shape, upon gas-phase dissociation (Fig. 4(c)). Furthermore, the dissociated light chain has a measured Ω value of 12.6 nm² which correlates well (within 5%) with our previously reported measured Ω value of 13.2 nm² for this protein in its native conformation.³⁰ Relating these observations to the topology of the complex, the data suggest

that as the heavy and light chain subunits are arranged in a sequential fashion with a limited number of inter-subunit contacts, dissociation from each other can then occur without significant destabilisation of either entity. This inference does not preclude any other factors that could contribute to the mode of dissociation, including the number and strength of intra-subunit interactions.

In contrast, the ESI-MS-CID-IMS-MS analysis of the chaperonin GroEL produced quite different results. GroEL is a non-covalently bound complex of 14 subunits, each 57.2 kDa in mass, arranged in two heptameric rings stacked back-to-back.⁴³ The characteristic topological feature of GroEL is a large hydrophobic cavity approximately 47 Å in diameter in which a polypeptide binds during protein folding events *in vivo*.⁴³ Figure 5(a) shows the ESI-IMS-MS spectrum of GroEL dominated by the intact 14-mer accompanied by a trace of unbound monomer. In this experiment, the collision cross-sectional area of the 14-mer was measured and gave a Ω value of 225 nm², in good agreement with data reported elsewhere.^{17,34,44} A Ω value of 221 nm² was calculated for intact GroEL from its crystal structure, which is similar to the reported MOBCAL projection approximation of 220 nm².¹⁷ The ESI-MS-CID-IMS-MS product ions arising from the intact complex showed a characteristic dissociation pattern, i.e., a monomer subunit carrying a disproportionately high number of charges, together with a 13-mer with approximately half the number of charges carried by the intact complex^{45,46} (Fig. 5(b)). In this experiment, the measured Ω value of the residual 13-mer was found to be 120 nm², a significant reduction of 45% in cross-sectional area compared with the intact complex (Table 1). A notional structure of the 13-mer has been modelled, assuming that loss of a single subunit is not accompanied by any structural changes, i.e., that the residual bulk of the complex is unchanged by the dissociation event (Fig. 5(c)). The calculated Ω value for this proposed 13-mer is 217 nm², and thus the estimated difference in calculated collision cross-section between the intact 14-mer and the 13-mer in this scenario would be (221–217 =) 4 nm², a decrease of only ~2%. This is clearly not the case here; indeed, the measured Ω value is less than the reported value of 145–163 nm² for the single ring, 7-mer GroEL complex,³⁴ suggesting a major structural collapse of the complex upon dissociation. However, on consideration of the quaternary structure of GroEL in which each subunit is non-covalently bound to two adjacent subunits within the same seven-membered ring as well as having contacts with subunits in the second seven-membered ring, removal of a single subunit from the intact 14-mer may therefore be expected to cause serious perturbations to the overall structure.

Haemoglobin is another example of a complex whose subunits, in this case two haem-bound α -globin (15.9 kDa) and two haem-bound β -globin (15.1 kDa) proteins, are each in contact with more than one neighbouring subunit. The ESI-MS-CID-IMS-MS-measured Ω values for the two dissociated trimer product ions resulting from loss of a single globin subunit reveal topological behaviour reminiscent of GroEL. Both $\alpha_2\beta$ (23.4 nm²) and $\alpha\beta_2$ (23.2 nm²) show some 18–24% reduction in cross-section compared with the calculated Ω values modelled on the elimination of a monomer from the complex but assuming no other structural rearrangement (29.8 and 27.8 nm², respectively), as well as being significantly (~44%) less than the measured Ω value of the intact complex (41.5 nm²) (see Supporting Information for details).

The observations from these examples suggest that although the phenomenon of gas-phase dissociation is common to many non-covalently bound complexes, the structure of the product ions produced may well be correlated to the original native topology of the complex by inspection of the interactions that undergo perturbation during the dissociation event.

Following these initial trials, the effect of ESI-MS-CID-IMS-MS on the two virus capsid intermediates was assessed (Fig. 6). Each one dissociated by loss of a single, relatively highly charged CP monomer. The measured Ω values for the residual complexes after dissociative loss of a monomer, together with the calculated Ω values for a series of modelled, potential structures, are summarised in Table 1. In the case of the hexamer intermediate (Fig. 6(a)), a Ω value of 61.6 nm² was measured for the major [3(CP₂:TR) + 2CP₂ + CP] product ions resulting from CP monomer dissociation, a reduction of ~25% from the intact species. This significant reduction, as observed for GroEL after monomer expulsion, suggests collapse of a structure whose stability depends on a cohesive arrangement of its subunits. The calculated Ω values for the three modelled structures, assuming loss of a single CP monomer but with their residual structures otherwise unperturbed, are as follows: 3-fold ring 81.4 nm², crescent 74.9 nm², and horseshoe 80.6 nm², all of which are ~18–25% greater than the measured Ω value (Table 1). However, on inspection of these structures the only one which might be expected to undergo collapse on expulsion of a subunit is the proposed 3-fold ring, in which each CP monomer is bound to two adjacent species, i.e., its dimeric partner and an adjacent CP₂ or [CP₂:TR] subunit, in an arrangement which is vital to sustain the ring-like structure (Fig. 6(b)). For the crescent and horseshoe models, one could envisage that the simplest way to remove a CP monomer would be from one of the distal CP₂ subunits which would involve perturbing only one set of intramolecular bonds and would involve no significant opportunity for a complete structural collapse. Thus, the ESI-MS-CID-IMS-MS results support a structure of 3-fold ring topology for the hexamer (Figs. 3 and 6(b)).

Interestingly, the decamer intermediate shows somewhat different ESI-MS-CID-IMS-MS dissociation characteristics. Upon expulsion of a highly charged CP monomer, the IMS-MS measured Ω value of the residual [5(CP₂:TR) + 4CP₂ + CP] complex is 107.2 nm², compared with that of 111.6 nm² measured for the intact complex, indicating that minimal change in shape or conformational form has taken place (Fig. 6(c)). Modelling and subsequent Ω value calculation of the dissociated complex using the 5-fold ring structure proposed (Fig. 6(d) and Table 1), while assuming no further structural change other than removal of a monomer, gave a Ω value of 113.0 nm², again very close to the measured Ω value. These observations lend further weight to the proposed 5-fold ring structural arrangement, as the CP monomer expelled probably originates from one of the peripheral CP₂ subunits that decorate the central core of five [CP₂:TR] subunits. Therefore, in contrast to the 3-fold hexamer structure, monomer expulsion does not interrupt the core architecture and therefore does not bring about a significant structural collapse, consistent with the experimental result.

It is interesting to note the manner in which the charges associated with each complex are proportioned between the products of CID. In the case of the hexamer, upon dissociation of the 24+ precursor ions the expelled CP monomer takes ~25% of the charges on average (i.e., 6+), despite accounting for only ~8% of the total mass of the complex. The monomer

expelled from the 34+ decamer precursor ions has the same number of charges, 6+, which in this case accounts for ~18% of the total number of charges but only ~5% of the total mass. The measured Ω values for the expelled CP monomer are the same for both hexamer and decamer dissociation, and indicate a structure ~15% expanded compared with the calculated Ω value for a CP monomer modelled assuming no perturbation from its defined structure within the intact complex. It is not possible to analyse the CP monomer alone in its folded state as it exists only as a dimer under neutral solution conditions. In comparison, the monomer expelled from dissociation of the GroEL complex takes almost 50% of the available charges and has a measured Ω value ~21% larger than the calculated Ω value. In general terms, therefore, it appears from these data that the greater the proportion of charges taken by the expelled monomer, the (albeit only slightly) more extended, or unfolded, the monomer. However, at this stage we do not believe that it is possible to correlate the quaternary structure of a complex with the distribution of charges between the products of dissociation with any degree of certainty. For example, on dissociation of MHC-1 the heavy and the light chain subunits share the available charges in a ~60%:40% ratio, respectively, and both have measured Ω values consistent with a native-like structure. This can be compared with the dissociation of GroEL, where the charges are shared between the 13-mer and the monomer in a ~53%:47% ratio, respectively, but the Ω value of the 13-mer is ~45% less than expected while that of the expelled monomer is ~21% greater.

CONCLUSIONS

ESI-IMS-MS and ESI-MS-CID-IMS-MS methodologies have been applied to elucidate the quaternary architectures of two viral capsid assembly intermediates of known stoichiometry but unknown topology, with the experimental data being complemented by modelling methods. In the case of the decamer, from the ESI-IMS-MS-measured cross-sectional area only one plausible model structure was apparent: a ring-like arrangement of five [CP₂:TR] subunits surrounded by five CP₂ subunits. The calculated Ω value for the model was in good agreement with the measured Ω value. In further support of this interpretation, ESI-MS-CID-IMS-MS dissociation of the decamer resulted in loss of a monomer leaving a residual species which showed very little change in cross-sectional area from the intact species, which accords well with the model of this intermediate from which a monomer could dissociate from one of the CP₂ subunits decorating the complex without disturbing the core structure (Fig. 6(d)). In the case of the hexamer intermediate, three plausible structures of the correct mass and stoichiometry were modelled. However, their calculated Ω values were all very similar to the ESI-IMS-MS-measured Ω value of this species, and assignment of a unique structure was not possible at this stage. ESI-MS-CID-IMS-MS dissociation of the hexamer led to loss of a coat protein monomer accompanied by a significant decrease in the cross-sectional area of the residual complex. This observation is consistent with an interconnected structural arrangement of subunits in which each one exhibits non-covalent interactions with several others in order to maintain a scaffold-based quaternary structure which could be expected to suffer significant destabilisation and collapse on dissociation. Therefore, the proposed 3-fold model, in which three CP₂ subunits and three [CP₂:TR] subunits are positioned in a ring-like arrangement, is the most likely structure for the hexamer. Removal of a coat protein monomer would involve disruption of the non-covalent

interactions with its dimeric partner as well as with another CP₂ subunit, which would result in cleavage of the ring (Fig. 6(b)). Conversely, dissociation of a monomer from the other structural possibilities, the crescent and horseshoe models, would not be expected to result in a collapse of their residual structures. Thus, the structures of the hexamer and the decamer have been assigned and both have been found to contain symmetry elements (3-fold and 5-fold, respectively) characteristic of the $T=3$ capsid structure.

Significantly, although the gas-phase ESI-MS-CID-IMS-MS dissociation characteristics of the complexes reported follow the generally accepted mechanism of monomer expulsion, leaving the residual complex less a single subunit, the measured Ω values of the post-dissociation products have revealed diagnostic features that can be correlated to the native topology of the intact biomolecular complexes. Single-stranded, positive-sense RNA viruses are one of the largest classes of viral pathogens infecting human, animal and plant hosts. The structures proposed here are of great significance, suggesting that the symmetry of the final capsid is established at an early stage in the assembly process and that the observed intermediates drive a templated assembly. Structural interpretation of these intermediates represents a break-through for understanding the capsid assembly mechanism and in determining whether they correspond to defined sub-fractions of the capsid or not. Importantly, it may be the key to providing potential targets for antiviral therapies. Furthermore, the applicability of ESI-MS-CID-IMS-MS to the elucidation of macromolecular complex topology of transient species in heterogeneous assembly reactions can be used to provide a wealth of information on other biomolecular processes, especially when supported by molecular modelling.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

TWK is funded by a White Rose DTC EPSRC studentship and VLM is a BBSRC post-doctoral research fellow (BB/E008070/1). The authors thank Dr B. M. Baker (Department of Chemistry and Biochemistry, University of Notre Dame, IN, USA) for providing the MHC-I preparation protocol, clones and peptide, and Dr J. P. Hodgkinson and Professor S. E. Radford (University of Leeds, UK) for supplying the prepared MHC-I sample. We acknowledge with gratitude the BBSRC for funding the purchase of the Synapt HDMS mass spectrometer (BB/E012558/1) and The Wellcome Trust and The Leverhulme Trust for supporting our research.

REFERENCES

1. Harrison SC, Olson AJ, Schutt CE, Winkler FK, Bricogne G. Nature. 1978; 276:368. [PubMed: 19711552]
2. Caspar DLD, Klug A. Cold Spring Harb. Sym. Quant. Biol. XXVII. 1962:1.
3. Crick FHC, Watson JD. Nature. 1956; 177:473. [PubMed: 13309339]
4. Schneemann A. Annu. Rev. Microbiol. 2006; 60:51. [PubMed: 16704342]
5. Basnak G, Morton VL, Rolfsson O, Stonehouse NJ, Ashcroft AE, Stockley PG. J. Mol. Biol. 2010; 395:924. [PubMed: 19913556]
6. Morton VL, Dykeman EC, Stonehouse NJ, Ashcroft AE, Twarock R, Stockley PG. J. Mol. Biol. 2010; 401:298. [PubMed: 20621589]
7. Stockley PG, Rolfsson O, Thompson GS, Basnak G, Francese S, Stonehouse NJ, Homans SW, Ashcroft AE. J. Mol. Biol. 2007; 369:541. [PubMed: 17434527]

8. Golmohammadi R, Valegard K, Fridborg K, Liljas L. J. Mol. Biol. 1993; 234:620. [PubMed: 8254664]
9. Valegard K, Liljas L, Fridborg K, Unge T. Nature. 1990; 345:36. [PubMed: 2330049]
10. Lago H, Parrott AM, Moss T, Stonehouse NJ, Stockley PG. J. Mol. Biol. 2001; 305:1131. [PubMed: 11162119]
11. Valegard K, Murray JB, Stonehouse NJ, van den Worm S, Stockley PG, Liljas L. J. Mol. Biol. 1997; 270:714.
12. Ashcroft AE, Lago H, Macedo JMB, Horn WT, Stonehouse NJ, Stockley PG. J. Nanosci. Nanotechnol. 2005; 5:2034. [PubMed: 16430137]
13. Stockley PG, Ashcroft AE, Francese S, Thompson GS, Ranson N, Smith AM, Homans SW, Stonehouse NJ. J. Theoret. Med. 2005; 6:119.
14. Bernstein SL, Dupuis NF, Laz ND, Wyttenbach T, Condron MM, Bitan G, Teplow BD, Shea J, Ruotolo BT, Robinson CV, Bowers MT. Nat. Chem. 2009; 1:326. [PubMed: 20703363]
15. Ruotolo BT, Giles K, Campuzano I, Sandercock AM, Bateman RH, Robinson CV. Science. 2005; 310:1658. [PubMed: 16293722]
16. Ruotolo BT, Hyung SJ, Robinson PM, Giles K, Bateman RH, Robinson CV. Angew. Chem. Int. Ed. 2007; 46:8001.
17. van Duijn E, Barendregt A, Synowsky S, Versluis C, Heck AJ. J. Am. Chem. Soc. 2009; 131:1452. [PubMed: 19138114]
18. Smith DP, Radford SE, Ashcroft E. Proc. Natl. Acad. Sci. USA. 2010; 107:6794. [PubMed: 20351246]
19. Heck AJ. Nat. Methods. 2008; 11:927. [PubMed: 18974734]
20. Pukala TL, Ruotolo BT, Zhou M, Politis A, Stefanescu R, Leary JA, Robinson CV. Structure. 2009; 17:1235. [PubMed: 19748344]
21. Taverner T, Hernandez H, Sharon M, Ruotolo BT, Matak-Vinkovic D, Devos D, Russell RB, Robinson CV. Acc. Chem. Res. 2008; 41:617. [PubMed: 18314965]
22. Hernandez H, Robinson CV. Nat. Protocols. 2007; 2:715. [PubMed: 17406634]
23. Garboczi DN, Hung DT, Wiley DC. Proc. Natl. Acad. Sci. USA. 1992; 89:3429. [PubMed: 1565634]
24. Hodkinson JP, Jahn TR, Radford SE, Ashcroft AE. J. Am. Soc. Mass Spectrom. 2009; 20:278. [PubMed: 18996721]
25. Mastico RA, Talbot SJ, Stockley PG. J. Gen. Virol. 1993; 74:541. [PubMed: 7682249]
26. Giles K, Pringle SD, Worthington KR, Little D, Wildgoose JL, Bateman RH. Rapid Commun. Mass Spectrom. 2004; 18:2401. [PubMed: 15386629]
27. Pringle SD, Giles K, Wildgoose JL, Williams JP, Slade SE, Thalassinis K, Bateman RH, Bowers MT, Scrivens JH. Int. J. Mass Spectrom. 2007; 261:1.
28. Tahallah N, Pinkse M, Maier CS, Heck AJ. Rapid Commun. Mass Spectrom. 2001; 15:596. [PubMed: 11312509]
29. Knapman TW, Berryman JT, Campuzano I, Harris SA, Ashcroft AE. Int. J. Mass Spectrom. 2009; in press. doi: 10.1016/j.ijms.2009.09.011
30. Smith DP, Knapman TW, Campuzano I, Malham RW, Berryman JT, Radford SE, Ashcroft AE. Eur. J. Mass Spectrom. 2009; 15:113.
31. Atlasevich, N.; Maze, J.; Bohrer, B.; Jarrold, M.; Clemmer, DE. Proc. 57th ASMS Conf. Mass Spectrometry and Allied Topics; Philadelphia, PA, USA. May 31–June 4, 2009; Poster: Ion Mobility II
32. Mesleh MF, Hunter JM, Shvartsburg AA, Schatz GC, Jarrold MF. J. Phys. Chem. 1996; 100:16082.
33. Ruotolo BT, Benesch JL, Sandercock AM, Hyung SJ, Robinson CV. Nat. Protocols. 2008; 3:1139. [PubMed: 18600219]
34. Uetrecht C, Versluis C, Watts NR, Wingfield PT, Steven AC, Heck AJR. Angew. Chem. Int. Ed. 2008; 47:6247.
35. Uetrecht C, Rose RJ, van Duijn E, Lorenzen K, Heck AJR. Chem. Soc. Rev. 2010; 5:1633. [PubMed: 20419213]

36. Benesch JL, Aquilina JA, Ruotolo BT, Sobott F, Robinson CV. *Chem. Biol.* 2006; 13:597. [PubMed: 16793517]
37. Benesch JL, Sobott F, Robinson CV. *Anal. Chem.* 2003; 75:2208. [PubMed: 12918957]
38. Heck AJ, van den Heuvel RH. *Mass Spectrom. Rev.* 2004; 23:368. [PubMed: 15264235]
39. Hyung SJ, Robinson CV, Ruotolo BT. *Chem. Biol.* 2009; 16:382. [PubMed: 19389624]
40. Jurchen JC, Garcia DE, Williams ER. *J. Am. Soc. Mass Spectrom.* 2004; 15:1408. [PubMed: 15465353]
41. Versluis C, Heck AJ. *Int. J. Mass Spectrom.* 2001; 637:210.
42. Bjorkman PJ, Saper MA, Samraoui B, Bennet WS, Strominger JL, Wiley DC. *J. Immunol.* 2005; 174:6. [PubMed: 15622565]
43. Braig K, Otwinowski Z, Hegde R, Boisvert DC, Joachimiak A, Horwich AL, Sigler PB. *Nature.* 1994; 371:578. [PubMed: 7935790]
44. Lorenzen K, Olia AS, Uetrecht C, Cingolani G, Heck AJ. *J. Mol. Biol.* 2008; 379:385. [PubMed: 18448123]
45. Van Duijn E, Simmons DA, Van den Heuvel RH, Bakkes PJ, van Heerikhuizen H, Heeren RM, Robinson CV, van der Vies SM, Heck AJ. *J. Am. Chem. Soc.* 2006; 128:4694. [PubMed: 16594706]
46. van Duijn E, Heck AJ, van der Vies SM. *Protein Sci.* 2007; 16:956. [PubMed: 17456746]

Figure 1.

(a) The two types of MS2 coat protein dimers (CP₂): the asymmetric A/B and the symmetric C/C, which form the basic building blocks of the $T=3$ capsid, with A (blue), B (green) and C (magenta). The main site of variation between these conformers occurs between the F and G β -strands (FG-loop) (taken from PDB accession code 2MS2⁸). (b) Conformational switching between the two types of dimer occurs upon binding of a 19-nucleotide RNA stem loop (TR).^{5,7} The sequence of the stem loop is shown and is also depicted as a yellow ribbon in complex with an A/B CP₂ dimer [CP₂:TR]. (c) The MS2 capsid consists of 180 coat protein (CP) monomers co-populating three quasi-equivalent conformers denoted A (blue), B (green), and C (magenta) in the form of 90 non-covalent dimers arranged within an icosahedral surface lattice.

Figure 2.

ESI-IMS-MS Driftscope plot (drift time (ms) vs. m/z) acquired during an MS2 viral capsid assembly reaction at $t = 3$ hours highlighting (A) the [CP₂:TR] initiation complex (33.5 kDa; 9+/10+ charge state ions; $\sim m/z$ 3500; measured Ω value 20.1 nm²) and the two capsid assembly intermediates: (B) the hexamer (182.9 kDa; 24+ to 27+ charge state ions; $\sim m/z$ 7500; measured Ω value 80.4 nm²) and (C) the decamer (304.8 kDa; 33+ to 37+ charge state ions; $\sim m/z$ 9000; measured Ω value 111.6 nm²).

Figure 3.

ESI-IMS-MS-measured Ω values (Ω_{TWIMS}) compared with computationally calculated Ω values^{29,30} (Ω_{Theory}) for the CP₂ (magenta) and [CP₂:TR] (green, blue = CP monomers; gold = TR) building blocks, and potential hexamer and decamer capsid assembly intermediates. The structures were modelled on the final capsid topology (PDB accession code 2MS2) taking into account the mass and stoichiometry discerned previously for these species,^{5,7} in addition to the interaction geometry observed in the intact capsid. Artificially extended, linear subunit arrangements were modelled for both the hexamer and decamer intermediates for comparison and were found, as expected, to yield calculated Ω values significantly larger than those measured by ESI-IMS-MS, while the 3-fold ring, crescent, horseshoe and 5-fold ring modelled structures gave calculated Ω values in good agreement with their respective intermediates.

Figure 4.

(a) ESI-IMS-MS spectrum of the major histocompatibility complex class 1 (MHC-1), showing the intact complex with an average of 12+ charge state ions (44.7 kDa; ~ m/z 3800) together with two unbound subunits, the light chain (average 6+ charge state ions; 11.8 kDa; ~ m/z 2000) and the peptide (MH^+ ; ~ m/z 1100), and the MHC-1 complex less the peptide (average 11+ charge state ions; 43.6 kDa; ~ m/z 4000). (b) ESI-MS-CID-IMS-MS spectra with increasing collision voltage from the precursor 12+ charge state ions of intact MHC-1 showing heavy chain (average 8+ charge state ions; 31.9 kDa; ~ m/z 5000), light chain (average 5+ charge state ions; 11.8 kDa; ~ m/z 2400) and peptide (MH^+ ; ~ m/z 1100) product ions. (c) Schematic model of MHC-1 dissociation assuming that the residual heavy and light chain subunits do not undergo any structural rearrangement after expulsion from the intact complex. The structures were adapted from PDB accession code 3HLA.

Figure 5.

(a) ESI-IMS-MS spectrum of the GroEL chaperonin complex dominated by the intact 14-mer with an average of 66+ charge state ions (801 kDa; $\sim m/z$ 12,000), accompanied by a trace of unbound monomer (13+/14+ charge state ions; 57.2 kDa; $\sim m/z$ 4400). (b) ESI-MS-CID-IMS-MS spectra with increasing collision voltage from the precursor 66+ charge state ions of intact GroEL, showing the residual 13-mer (28+ to 38+ charges; 744 kDa; $\sim m/z$ 22,000) and the unfolded, expelled monomer carrying a disproportionately high number of charges (average 31+ charge state ions; $\sim m/z$ 1800). (c) Schematic model of GroEL dissociation assuming that the residual 13-mer and expelled monomer subunit do not undergo any structural rearrangement after dissociation, generated by removal of atoms constituting one subunit (from PDB accession code 1GRL).

Figure 6.

(a) ESI-MS-CID-IMS-MS spectra with increasing collision voltage from the precursor 24+ charge state ions of the hexamer assembly intermediate showing the dissociated $[3(\text{CP}_2:\text{TR}) + 2\text{CP}_2 + \text{CP}]$ product (i.e., loss of a CP monomer) (average 17+ charge state ions; 169.2 kDa; $\sim m/z$ 10,000) and the expelled, unfolded CP monomer (average 6+ charge state ions; 13.7 kDa; $\sim m/z$ 2100). (b) Schematic model of hexamer assembly intermediate dissociation, assuming that the residual complex does not undergo any structural rearrangement after monomer expulsion. Expulsion of a monomer reveals a cavity within the proposed ring-like structure. The fragmented representation was generated by removal of atoms constituting one CP monomer from the previously modelled 3-fold ring hexamer structure (Fig. 3). (c) ESI-MS-CID-IMS-MS spectra with increasing collision voltage from the precursor 34+ charge state ions of the decamer assembly intermediate showing the dissociated $[5(\text{CP}_2:\text{TR}) + 4\text{CP}_2 + \text{CP}]$ product (i.e., loss of a CP monomer) (average 27+ charge state ions; 291.1 kDa; $\sim m/z$ 11 000) and the expelled, unfolded CP monomer (average 6+ charge state ions; 137. kDa; $\sim m/z$ 2100). (d) Schematic model of the decamer assembly intermediate dissociation, assuming that the residual complex does not undergo any structural rearrangement after monomer expulsion. Expulsion of a monomer reveals minimal difference in the structural size and shape. The fragmented representation was generated by removal of atoms constituting one CP monomer from the previously modelled 5-fold ring decamer structure (Fig. 3).

Table 1

Summary of the ESI-IMS-MS and ESI-MS-CID-IMS-MS cross-sectional areas measured from the lowest charge state ions detected (Ω_{IMS}) and the theoretically calculated cross-sectional areas using the Leeds algorithm (Ω_{Calc}). Data for two macromolecular complexes, GroEL and MHC-1, and their residual structures after MS/MS dissociation assuming no further change to the original structure other than loss of a single monomer subunit are shown, together with the two virus capsid assembly intermediates, the hexamer and decamer, and their potential, modelled topologies (see also Supporting Information)

Complex	Topology	Intact complex (MS/MS precursor)			MS/MS dissociation product (i.e., complex less one subunit)		
		Ω_{IMS} (nm ²)	Ω_{Calc} (nm ²)	Ω (%)	Ω_{IMS} (nm ²)	Ω_{Calc} (nm ²)	Ω (%)
GroEL	14-mer	224.6	221.3	+1.5			
	13-mer				120.2	217.1	−45
MHC-1	intact complex	29.0	27.8	+4.4			
	heavy chain (monomer)				24.8	24.1	−2.9
	light chain (monomer)				12.6	11.8	−6.6
MS2 hexamer intermediate [3(CP ₂ :TR)+3CP ₂]	3-fold ring		83.9	−4		81.4	−24
	crescent	80.4	78.4	+3	61.6	74.9	−15
	horseshoe		85.4	−6		80.6	−25
MS2 decamer intermediate [5(CP ₂ :TR)+5CP ₂]	5-fold ring	111.6	114.6	+1	107.2	113.0	+1