

Published in final edited form as:

*J Mol Biol.* 2007 June 01; 369(2): 541–52. doi:10.1016/j.jmb.2007.03.020.

## A Simple, RNA-Mediated Allosteric Switch Controls the Pathway to Formation of a $T = 3$ Viral Capsid

Peter G. Stockley\*, Ottar Rolfsson, Gary S. Thompson, Gabriella Basnak, Simona Francese, Nicola J. Stonehouse, Steven W. Homans, Alison E. Ashcroft

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

### Abstract

Using mass spectrometry we have detected both assembly intermediates and the final product, the  $T = 3$  viral capsid, during reassembly of the RNA bacteriophage MS2. Assembly is only efficient when both types of quasiequivalent coat protein dimer seen in the final capsid are present in solution. NMR experiments confirm that interconversion of these conformers is allosterically regulated by sequence-specific binding of a short RNA stem-loop. Isotope pulse-chase experiments confirm that all intermediates observed are competent for further coat protein addition, i.e., they are all on the pathway to capsid formation, and that the unit of capsid growth is a coat protein dimer. The major intermediate species are dominated by stoichiometries derived from formation of the particle threefold axis, implying that there is a defined pathway toward the  $T = 3$  shell. These results provide the first experimental evidence for a detailed mechanistic explanation of the regulation of quasiequivalent capsid assembly. They suggest a direct role for the encapsidated RNA in assembly *in vivo*, which is consistent with the structure of the genomic RNA within wild-type phage.

### Keywords

virus assembly; RNA-protein interaction; mass spectrometry

### Introduction

Simple spherical viruses consisting of a nucleic acid genome enclosed in a protein shell composed of multiple copies of one, or very few, coat protein (CP) subunits generally have isometric capsids. The positioning of the CP subunits is predictable from the allowed tessellations of a sphere.<sup>1,2</sup> We studied the RNA bacteriophage MS2, which conforms to a  $T = 3$  shell in the Caspar-Klug nomenclature. The triangulation number ( $T = 3$ ) implies that the phage CP subunit can exist in three quasiequivalent conformations (A, B, and C), allowing 180 copies of the subunit to create an enclosed capsid. This is indeed the case (Figure 1); the CP subunits form interdigitated noncovalent dimers (CP<sub>2</sub>) that act as the basic capsid building blocks composed of A/B or C/C conformers.<sup>3,4</sup> Within each monomer, these conformers are defined by the orientation of the loop of the polypeptide connecting the F and G  $\beta$ -strands, with A and C subunits having extended loops while B subunit loops

\*Corresponding author: stockley@bmb.leeds.ac.uk .

fold back toward the globular core of the protein (Figure 1(c)). In the complete capsid, A/B dimers surround the particle fivefold axes with a ring of B-type loops, while A- and C-type loops alternate around the threefold axes.

A major unresolved problem in structural virology is understanding the detailed molecular mechanism(s) that gives rise to the assembly of capsids of the correct size and symmetry. Various proposals for the formation of  $T=3$  shells based on the initial formation of three<sup>5</sup> or fivefold<sup>6</sup> assembly initiation complexes have been made. However, capsid assembly is spontaneous and usually very rapid, with the result that it has been technically difficult to isolate and characterize intermediates beyond these “initiation complexes” on the pathway to the final products.<sup>7,8</sup> The MS2 system is an ideal model for investigating such phenomena due to the extensive biochemical and structural information available.<sup>3, 9–13</sup> Capsid reassembly can be triggered *in vitro* by a sequence-specific RNA–protein interaction between a CP<sub>2</sub> and an RNA stem–loop (TR) of just 19 nt<sup>14,15</sup> that encompasses the start codon of the viral replicase on the RNA genome (Figure 1). RNA–protein binding thus achieves two functions *via* a single molecular recognition event: translational repression of replicase and creation of an assembly competent complex on viral RNA.

We used electrospray ionization-mass spectrometry (ESI-MS) to investigate the mechanism of capsid reassembly and here show for the first time the time-dependent formation of both intermediates and the final product in a single assay. ESI-MS has proven to be a powerful technique for the analysis of noncovalently bound protein complexes because the mild ionization conditions are able to preserve often fragile macromolecular structures, thus permitting mass analysis of the intact complex<sup>16,17</sup> Our analysis of the yield and kinetics of the assembly of both intermediates and the final capsid leads to one of the first mechanistic descriptions of the control of quasiequivalent interactions during virus assembly, in this case *via* an allosteric switch controlled by interaction with phage RNA.

## Results and Discussion

### An RNA-mediated conformational change

The reassembly experiments were analyzed using a customized nanoelectrospray ionization–time-of-flight instrument with an extended  $m/z$  range (up to  $m/z$  60,000), collisional cooling capabilities,<sup>18–21</sup> and a temperature-controlled, automated injection and ionization system (see Materials and Methods). Capsid reassembly *in vitro* was achieved by raising the pH of a solution of acid-disassembled coat protein in the presence of TR RNAs.<sup>22,23</sup> Working in solutions of final pH 5.2–5.7 and in buffers compatible with mass spectrometry slows the reassembly process, allowing us to observe early and intermediate stages.

In our ESI-MS experiments the possible molecular components in the reassembly reaction consist of the following: the CP; the CP dimer (CP<sub>2</sub>); the TR RNA (TR), the RNA–protein dimer complex (CP<sub>2</sub>:TR), and various higher-order species. Systematic analysis of reassembly conditions by gel filtration-light-scattering prior to the MS assay allowed us to choose conditions for the reaction that slow the production of the final product, the  $T=3$  capsid. These conditions favor very high overall yields of the desired end product

and essentially eliminate the production of mis-structured aggregates that could otherwise complicate the analysis.

At pH  $\sim 3.0$  the CP obtained by acid disassembly of recombinant  $T=3$  capsids<sup>24</sup> shows a spectrum dominated by a mixture of CP (13.7 kDa) and CP<sub>2</sub> (Figure 2(a); components labeled A and B, respectively), even at concentrations well above the dissociation constant for dimer formation determined at neutral pH, suggesting that the CP<sub>2</sub> is not completely stable at this pH, at least in the gas phase of the mass spectrometer. These major species are accompanied by minor species with masses consistent with the CP trimer, tetramer, and pentamer (Figure 2(a); unlabeled, minor peaks (< 1%) in the range  $m/z$  4000–7000). The addition of TR to a solution of coat protein at a final pH  $\sim 5.2$ – $5.7$  and at a stoichiometry of 1:1 (CP<sub>2</sub> to TR) changes the spectrum radically, resulting in the complete disappearance of CP, CP<sub>2</sub>, and the other minor species and the emergence of the complex CP<sub>2</sub>:TR of mass 33.5 kDa (Figure 2(b); component C). The presence of this complex as the major species implies that all of the CP<sub>2</sub> binds TR RNA and is thus stabilized. The narrow ESI charge state distribution (9+–10+ ions) of the complex reflects its well-defined structure. This species is accompanied by masses consistent with trace amounts of higher-order oligomers in the region  $m/z$  4000–10,000, which do not appear to alter significantly in concentration over time (Figure 2(c)) until after some 5 h, when certain higher-order species increase slightly in intensity (e.g., component D in Figure 2(d); see below for an explanation).

The spectra are significantly different, however, when reactions are performed at a final stoichiometry of 2:1 CP<sub>2</sub> to TR (Figure 2(f)–(h)), by allowing pre-equilibration of the complex (1:1, CP<sub>2</sub>:TR) for 10 min (Figure 2(e)) followed by the addition of an equal aliquot of CP<sub>2</sub> lacking RNA (Figure 2(f)). The proportion of the higher-order species increases rapidly and dramatically and is not consistent with a simple mass action effect (Figure 2, cf. (c) and (g)). This effect was not seen when equal aliquots of CP<sub>2</sub> and TR were added to 1:1 reactions. Several higher-order species are clearly visible in the 2:1 reaction, some of which we are able to assign to unambiguous stoichiometries, as well as weaker species whose signals are too broad to measure accurately. For the mass assignment, the measured masses, based on a minimum of three adjacent charge states, were compared with the calculated masses of all possible multiples of CP and TR. A significant higher mass species of 182.7 kDa is consistent with the complex  $[3(\text{CP}_2:\text{TR}) + 3\text{CP}_2]$  (Figure 2(g) and (h); component D). In addition to CP, CP<sub>2</sub>, CP<sub>2</sub>:TR, and  $[3(\text{CP}_2:\text{TR}) + 3\text{CP}_2]$  (components A–D, respectively), other, minor, higher-order species consistent with  $[\text{CP}_2:\text{TR} + \text{CP}]$  (mass 47 kDa),  $[2(\text{CP}_2:\text{TR})]$  (mass 67 kDa), and  $[3(\text{CP}_2:\text{TR}) + \text{CP}_2 + \text{CP}]$  (mass 142 kDa) were unambiguously identified (Figure 2; not labeled). Other species of masses > 200 kDa (e.g., component(s) E, mass  $\sim 288$ – $300$  kDa) were more difficult to identify due to low-intensity broad peaks, which made unambiguous charge state assignment difficult and hence gave rise to the possibility of several theoretically compatible stoichiometries. The presence of complexes with CP monomer components may arise because of the presence of the monomer under the starting conditions. Over time the peaks corresponding to all of these species decreased in intensity. In contrast to the time-course for the 1:1 CP<sub>2</sub> to TR reaction (Figure 2(b)–(d)), the higher-order aggregates ( $m/z$  4000–10,000) are accompanied by broad unresolved higher mass to charge signals ( $> m/z$  14000) visible in the spectra for the 2:1 reaction from the earliest time-point sampled ( $\sim 1$  min; Figure 2(f)–(h)). Such broad peaks

have been mass assigned to the MS2  $T=3$  capsid by others<sup>25</sup> and are also seen when we analyze the intact recombinant capsid.<sup>24</sup>

To confirm that the high mass peaks corresponded to the formation of  $T=3$  shells, identical reassembly reactions were analyzed by gel filtration-light-scattering (Figure 3). The 1:1 reaction forms capsids only very slowly, whereas the 2:1 equivalent has capsids present from the earliest time-points. These data imply that the small amounts of reassembly intermediates seen in Figure 2(b)-(d) are caused by a limiting concentration of an RNA-free CP component in the assembly pathway. Increasing the concentration of CP<sub>2</sub> triggers a dramatic increase in the efficiency of assembly of the higher-order species observed in Figure 2(f)-(h). Further increasing the concentration of RNA-free CP subunits in the MS assay (to a 4:1 CP<sub>2</sub> to TR ratio) results in more rapid loss of these higher-order species, including the CP<sub>2</sub>:TR complex, from the  $m/z$  spectrum, suggesting that they do participate in capsid formation rather than being simply MS artifacts. The spectrum is then dominated by peaks originating from CP monomer and dimer, together with a trace of trimer (data not shown). A reasonable inference from these effects is that the RNA-free coat protein added to the preformed CP<sub>2</sub>:TR complex is in a different quasiequivalent conformation from the one in the complex. Because both A/B and C/C dimers are required to build the  $T=3$  capsid, mixing these conformers at high concentration allows the assembly reaction to proceed efficiently. The 1:1 CP<sub>2</sub>:TR mixture also makes  $T=3$  shells, although under these conditions the RNA-free sample does not. Because the life-time of the CP<sub>2</sub>:TR is 42 s,<sup>26,27</sup> both types of dimer are available in the 1:1 reaction but in these reactions the concentration of the RNA-free form would be limiting.

To probe whether TR binding leads to conformational change within the CP subunit, using NMR we carried out measurements with a mutant CP subunit (W82R) that does not assemble beyond dimer but does bind TR RNA with the same affinity as the wild-type (unpublished results). Residue 82 lies outside both the FG-loop and the TR binding site. The backbone resonances from the <sup>1</sup>H, <sup>13</sup>C, <sup>15</sup>N triple resonance NMR spectrum of the *apo*-CP<sub>2</sub>(W82R) were analyzed with 87% of the backbone NH resonances being assigned, including virtually all the nonproline resonances in the FG loop region. The full details of these experiments will be published elsewhere. The degeneracy of all pairs of backbone resonances present in *apo*-CP<sub>2</sub>(W82R) resulting from pairs of symmetry-related atoms in the two subunits of the dimer, encompassing the FG loop, show that in the absence of RNA the protein subunits are equivalent on the NMR chemical shift timescale (Figure 4). Furthermore, <sup>1</sup>H-<sup>15</sup>N heteronuclear nuclear Overhauser effect spectra (data not shown) show considerable motion within the FG loop on picosecond to nanosecond timescales, indicating that the chemical shift equivalence seen in these regions results from an average of a large number of conformers.

The addition of TR to the labeled sample leads to a change in the spectrum of CP<sub>2</sub>(W82R), with resonances for the *apo*-CP<sub>2</sub> state disappearing and pairs of peaks appearing at perturbed positions, which are associated with the [CP<sub>2</sub>(W82R):TR] complex. The data show that the complex is in slow exchange with the *apo* state, as expected for a RNA-protein complex with a nanomolar binding affinity. The sequential assignment for the backbone HN, N, and C<sup>α</sup> resonances of residues 71–76 were made from HNCA, HncoCA, and <sup>1</sup>H-<sup>15</sup>N

heteronuclear single-quantum coherence spectra and the position of the fragment within the CP<sub>2</sub>(W82R) primary sequence assured by the presence of distinctive C<sup>α</sup> resonances for residues 73 and 74, which are a GG pair, one of only two such pairs within the primary sequence of the CP. Pairs of resonances can be seen for residues 71, 72, 75, and 76, which are clearly delineated by shift differences in the <sup>1</sup>H+<sup>15</sup>N dimensions (Figure 4). Comparison of the observed differences in shifts for the FG loop <sup>1</sup>H amide resonances were made with those predicted by the Heigh Mallion ring current model using the program MOLMOL and ring current intensities suitable for RNA.<sup>28,29</sup> These calculations show that the observed changes in shifts in the presence and absence of TR are approximately an order of magnitude greater than that expected from theory (data not shown), confirming that the observed shifts are not caused by RNA binding per se but arise either from RNA-induced changes in the conformations of the loops or from fixing of a defined conformational state from within an equilibrium of states present in the *apo* protein. In summary, binding of TR results in conformational changes involving the FG loops and creates asymmetry in the complex. We propose that the conformations of the loops in the RNA-bound dimer are likely to be closer to those in an A/B coat protein dimer, compared to C/C-like or ill-defined conformers in the *apo* protein.

Why then does the W82R mutant fail to assemble beyond the TR:CP<sub>2</sub> complex? To address this question we carried out an ESI-MS chasing experiment with combinations of wild-type and W82R CP. The W82R mutant dimer forms a CP<sub>2</sub>(W82R):TR complex when mixed 1:1 with TR with very little evidence of any higher-order species. Interestingly, the charge state distribution of the CP<sub>2</sub>(W82R):TR complex was slightly but noticeably wider ( $n = 7^+ - 10^+$ ) than that of the wild-type complex ( $n = 9^+ - 10^+$ ) analyzed under the same buffer and instrumental conditions, indicating a more dynamic structure (data not shown). Upon the addition of wild-type CP<sub>2</sub> to the preformed CP<sub>2</sub>(W82R):TR complex, both the CP<sub>2</sub>:TR and the CP<sub>2</sub>(W82R):TR complexes were detected with no higher-order species visible. The reverse experiment, i.e., performing the wild-type CP<sub>2</sub>:TR complex (with its associated dimer, trimer, and traces of other higher-order species) followed by the addition of CP<sub>2</sub>(W82R), resulted again in the observation of only the CP<sub>2</sub>:TR and CP<sub>2</sub>(W82R):TR complexes, in this case with the concomitant disappearance of the higher-order species (data not shown). There were no signals corresponding to a mixed CP-CP(W82R):TR complex. These results are consistent with our expectation that W82R can bind TR normally but is unable to support further interactions with other coat protein dimers, thus acting as an inhibitor of reassembly.

It is a reasonable assumption that the RNA-induced conformational change in the FG loops observed with W82R protein also occurs with the wild-type coat protein. This RNA-mediated allosteric effect at the FG loops explains the dependency of the kinetics and yield of assembly products on the presence of RNA-free and RNA-bound forms of the coat protein dimer.

### The pathway to $T = 3$ shell formation

To confirm that the intermediates seen in the MS spectra were on-pathway to capsid formation and to determine the mass step size of capsid growth, we repeated the reassembly

assay using isotopically labeled wild-type CP. The CP<sub>2</sub>:TR complex was performed as above with a 1:1 ratio of reactants but using <sup>15</sup>N-labeled CP<sub>2</sub> (Figure 5(a), component F). The intermediate complex [3(CP<sub>2</sub>:TR) + 3CP<sub>2</sub>] discussed above was also visible in low yield in the region m/z 6700–7200, as expected. The measured molecular mass of this component indicated a totally <sup>15</sup>N-labeled [3(CP<sub>2</sub>:TR) + 3CP<sub>2</sub>] complex, i.e., with all 12 CP subunits <sup>15</sup>N-labeled, and hence is labeled 12:0 (<sup>15</sup>N to <sup>14</sup>N). RNA-free <sup>14</sup>N-CP<sub>2</sub> was then added to make the 2:1 reaction (Figure 5(b)). There were then two peaks of approximately equivalent intensity corresponding to <sup>14</sup>N and <sup>15</sup>N-labeled CP<sub>2</sub>:TR complexes (components C and F), but there is no evidence for a heterodimer complex, indicating that while the TR molecule can dissociate from the CP<sub>2</sub>:TR complex and reassociate with a different CP<sub>2</sub>, the dimer does not dissociate to monomer under these conditions. Inspection of the region m/z 6700–7200 containing the multiply charged peaks for the complex [3(CP<sub>2</sub>:TR) + 3CP<sub>2</sub>] in the 2:1 reaction reveals that the <sup>15</sup>N-labeled complex peaks are still present (Figure 5(b), inset, labeled 12:0) with additional peaks of the same charge state but with lower m/z values, i.e., of lower molecular mass. Further addition of <sup>14</sup>N-CP<sub>2</sub> (<sup>15</sup>N to <sup>14</sup>N ratio now 1:3) causes these peaks to intensify (Figure 5(c)) and finally (<sup>15</sup>N to <sup>14</sup>N ratio now 1:4) disappear (Figure 5(d)), as expected from the previous experiments. The masses of these components indicated the presence of complexes consisting of <sup>15</sup>N to <sup>14</sup>N CP in the ratios of 10:2, 8:4, 6:6, 4:8, 2:10, and 0:12 subunits, respectively, i.e., equivalent to all stoichiometries differing by dimer addition. The mass differences between these species consolidate the assignment of this complex and confirm the presence of 12 CP molecules. Again, no evidence was found for complexes of this stoichiometry with mixed <sup>15</sup>N:<sup>14</sup>N CP<sub>2</sub> units. The predominance of the <sup>15</sup>N to <sup>14</sup>N 4:8 ratio species in Figure 5(c) is consistent with the ratio of <sup>15</sup>N to <sup>14</sup>N CP present in this particular reassembly reaction (i.e., 1:3).

Thus, it appears that the TR in the reassembly reaction rapidly equilibrates between the <sup>14</sup>N and <sup>15</sup>N-labeled coat protein dimer subunits, presumably because of the short half-life of the CP<sub>2</sub>:TR complex.<sup>26,27</sup> As the initially formed [3(CP<sub>2</sub>:TR) + 3CP<sub>2</sub>] <sup>15</sup>N-labeled complexes continue on the assembly pathway to form capsids, new [3(CP<sub>2</sub>:TR) + 3CP<sub>2</sub>] complexes are formed from a mixture of the remaining <sup>15</sup>N-labeled CP<sub>2</sub> species and the added <sup>14</sup>N material. Gradually, the percentage of all <sup>15</sup>N-labeled [3(CP<sub>2</sub>:TR) + 3CP<sub>2</sub>] decreases with increasing CP<sub>2</sub> concentration, while the mixed <sup>14</sup>N:<sup>15</sup>N (Figure 5(b) and (c)), and eventually all <sup>14</sup>N complexes, become predominant. Eventually, the addition of further CP<sub>2</sub> dimers results in the disappearance of this complex, as well as other higher-order species, as capsid formation progresses (Figure 5(d)). One prediction of this interpretation is that performing the experiment in the opposite direction, i.e., performing the complex with <sup>14</sup>N CP<sub>2</sub> and chasing with <sup>15</sup>N-labeled material, should produce an equivalent, but mirror image, spectrum. This is indeed the case (see Supplementary Information).

The data confirm that the intermediate species detected in the MS spectra are competent to bind additional coat proteins and are therefore likely to be on the capsid assembly pathway. The unit of capsid growth is clearly a coat protein dimer under these conditions. Although some intermediate species are visible for extended periods during the assembly process, there does not appear to be a significant build-up of partially formed structures that later coalesce. This last observation is consistent with cryo-EM reconstructions of species formed in reassembly reactions that show particles of increasing completeness over time.<sup>23</sup>

The results lead directly to a molecular mechanism for  $T=3$  capsid formation (Figure 6). RNA-free CP<sub>2</sub> in solution has a preponderant, quasiequivalent conformation distinct from that created when TR binds, both of which are needed to allow the capsid to be built. This is consistent with the previous observation that some CP<sub>2</sub> subunits will not bind TR until they have undergone a conformational change.<sup>12</sup> The TR binding site lies entirely within a CP<sub>2</sub> and it makes distinct contacts with each monomer, suggesting that it might bind preferentially to an asymmetric dimer, e.g., an A/B unit. However, the NMR data from the *apo*-CP<sub>2</sub> suggest that the dynamic motions of the FG loop regions also propagate along the attached F and G β-strands and overlap the RNA binding site, providing a potentially direct mechanism to couple RNA binding and FG loop conformational preference. A 1:1 reassembly reaction would thus be dominated by A/B-like species, whereas the 2:1 reactions described above would, in addition, contain more symmetrical dimers but with very dynamic FG loop regions. These could be C/C-like. Formation of A/B dimers would in principle permit formation of a capsid fivefold axis via interacting B-type loops. Such dimers could also contribute to the formation of a threefold axis by the A conformers interacting with C/C dimers. The stimulatory effect of adding RNA-free CP<sub>2</sub> to the preformed CP<sub>2</sub>:TR suggests that formation of particle threefold axes is the dominant pathway toward  $T=3$  shell formation under these conditions. This is consistent with the fact that none of the assigned mass peaks in the spectra corresponds to a fivefold intermediate, 5(CP<sub>2</sub>:TR); rather, they are essentially all consistent with the threefold stoichiometry, [3(CP<sub>2</sub>:TR) + 3CP<sub>2</sub>], implying that there is a unique assembly pathway. Formation of the threefold complex determines that assembly yields a  $T=3$  rather than a  $T=1$  capsid; the latter would require only complexes with fivefold stoichiometries.

### Biological implications

The mechanism described above immediately begs an obvious question: what triggers the C/C-like dimer in solution to adopt the A/B conformer at appropriate positions in the growing capsid? In vivo there is only one copy of the TR stem-loop per genome and there are no obvious candidates for closely related sequences/structures in the remainder of the phage RNA. However, we have shown that very many RNA stem-loop variants can bind at the TR binding site. Indeed, we have characterized over 30 such species by X-ray crystallography.<sup>13,30</sup> In such experiments preformed capsids are soaked with RNAs that penetrate into the interior of the particles via the pores at the symmetry axes, from where they can bind to all possible binding sites within the interior, including both A/B and C/C conformers. The orientation of the RNA at C/C dimers is degenerate but it is unique at A/B, allowing the details of the RNA-protein complex to be determined. These results suggest that the difference in RNA affinity between A/B and C/C dimers in a capsid is probably modest, although for many of the RNA variants we have studied the measured in vitro affinity for isolated dimers is very low. Note that the recombinant  $T=3$  shells used to generate the CP<sub>2</sub> starting material here were described previously as “empty,” because they appear to lack any MS2 RNA or other long RNA sequences.<sup>24</sup> They are not, however, free of cellular RNA fragments, which may serve as the triggers of quasiequivalent switching when the recombinant protein is expressed at high levels.

It appears, therefore, that if there were suitably oriented genomic sequences/structures outside the TR region, coat protein dimers adding to a growing shell could bind them, thus fixing the appropriate quasiequivalent conformation. Cryo-EM reconstruction of the intact phage reveals the existence of such ordered RNA segments located underneath the protein shell,<sup>31</sup> supporting this idea. We have also shown that RNAs that encompass the TR sequence but have 5' or 3' genomic extensions appear to alter the efficiency of capsid formation.<sup>23</sup> This mechanism may well play a direct role in the life-cycle of the phage. The TR sequence is a transient signal that is only generated when ribosomes read through the coat protein gene, disrupting a long-range base-pairing interaction, the Min Jou sequence<sup>32</sup>. Binding of a CP<sub>2</sub> to the TR stem-loop would lead to translational repression of the replicase cistron and simultaneously mark the MS2 genome for packaging into capsids. Many of the genomes marked in this way, however, might only be partially replicated and it might be inappropriate to begin assembly immediately after the repression complex forms. If protein-free genomic RNAs present higher affinity binding sites for CP<sub>2</sub> at low concentrations, this would allow all progeny phage genomes to complete replication and carry an assembly initiation species. As the concentration of CP<sub>2</sub> rises to levels in excess of the genomic RNA copy number, these C/C-like species can begin the assembly process by interacting with the A/B dimer bound at TR. Presumably, the additional A/B dimers required to complete assembly would then be triggered to form at appropriate points by interactions with genomic sequences outside the TR region, where both RNA-protein and protein-protein binding energy can contribute to the affinity of the interaction.

The pathway to capsid formation described above corresponds to the mechanism of autostery first described by Caspar, in which free subunits are held in an assembly-incompetent state until they make contact with the growing viral shell or another viral ligand.<sup>33</sup> In many other viral systems quasiequivalent conformer formation is dominated by alteration of the orientation of distinct protein domains or extended "arm" regions.<sup>5</sup> Many such proteins share structurally conserved protein domains and these have been co-opted by viruses from diverse sources from bacteria to humans.<sup>34</sup> The RNA phage coat proteins, in contrast, contain only a single domain and lack such extended arms. They form a unique protein fold that has to date no other viral homologues. This isolation may be because in these cases assembly is so dependent on interactions with the viral genome that it places such strict evolutionary constraints on the genomic sequence that the fold has not been used elsewhere. A related issue is the conformation of the phage genome during the assembly process. Does CP<sub>2</sub> binding to the growing shell provide the driving force for genome folding and condensation or does the protein bind to a prefolded RNA? AFM measurements on a number of other single-stranded RNA viruses suggest that the genomic RNAs can fold stably into a compact structure with dimensions similar to that of the interior of the final protein shell, suggesting that the latter proposal is more likely.<sup>35</sup>

The data presented here demonstrate the power of mass spectrometry to dissect the molecular details of a viral assembly pathway and gain insight into MS2 phage capsid construction by the identification of on-pathway intermediates. The ability to monitor mass and intensity changes in co-existing, noncovalently bound, high mass oligomeric species in real time is a vital asset for these studies. Such information is potentially very important for designing novel inhibitors of viral life-cycles, e.g., in the RNA phage, reagents



that block the conformational changes required to create the A/B dimer would be potent inhibitors of assembly. Additionally, there are growing applications of virus-like particles in therapeutic and bionanotechnologic applications in which fundamental information on assembly mechanisms would be extremely useful.<sup>36,37</sup> ESI-MS can clearly contribute significantly to interdisciplinary studies of such systems.

## Materials and Methods

### Preparation of coat proteins and TR RNA

**Proteins**—Wild-type recombinant coat protein in the form of  $T=3$  shells lacking specifically bound RNA were prepared by overexpression in *Escherichia coli* as described previously.<sup>24</sup> These capsids were purified and then disassembled coat protein dimers isolated by treatment with glacial acetic acid, followed by exchange into 20 mM acetic acid.<sup>12,22</sup>

NMR experiments were carried out on samples of the MS2 mutant coat protein, W82R, that does not assemble beyond the dimer.<sup>38</sup> W82R protein was purified by successive chromatography over diethylaminoethanol and sulphopropyl exchange media following the methodology described previously<sup>38</sup> with slight modifications. Following elution from the sulphopropyl column in 20 mM phosphate, pH 5.8, all samples were treated with 10  $\mu$ L/ml of 2 M Tris, pH 8.5, to prevent crystallization and were then dialyzed into 50 mM Tris, pH 8.5, 100 mM KCl, 10 mM DTT, and 20 mg/l complete protease inhibitor for long-term storage at 4 °C. The purity and identity of the protein samples were assessed by SDS–PAGE, ESI-MS, and N-terminal sequencing of the first eight residues in the protein.

Isotopically labeled reagents were purchased from (Spectra Stable Isotopes Inc., Columbia, MD, USA), Goss Scientific Instruments (Great Baddow, Essex, UK), or Cambridge Isotope Laboratories (Andover, MA, USA). W82R samples were prepared by growing cells in M9 medium prepared with 50% D<sub>2</sub>O/50% H<sub>2</sub>O (v/v) and supplemented with 40 mg/l FeCl<sub>3</sub> and 1% (v/v) 100 $\times$  BME vitamin solution. Isotopes were introduced into the growth medium by the addition of 4 g/l of uniformly <sup>13</sup>C-labeled 50% <sup>2</sup>H-random-labeled glucose and 1 g/l of <sup>15</sup>NH<sub>4</sub>Cl. Plasmid selection during expression was maintained using 100  $\mu$ g/l carbenicillin. Cells were grown to an A<sub>600</sub> of 0.8 starting from a labeled starter culture grown overnight and were induced with 0.1 mM isopropyl- $\beta$ -D-thiogalactopyranoside and harvested after a further 8 h of growth.

<sup>15</sup>N-labeled wild-type MS2 coat protein was prepared identically with wild-type protein, with the exception that the cells were grown in <sup>15</sup>NH<sub>4</sub>Cl-containing minimal medium.

**RNAs**—Synthetic oligonucleotides encompassing the TR sequence (Figure 1) or the –5C variant were prepared as described previously.<sup>39</sup> The –5C variant binds to the coat protein dimer significantly more tightly than the wild-type sequence, with half-lives of > 300 and 42 s, respectively.<sup>27</sup> The crystal structure of the –5C variant bound to the coat protein in a preformed capsid<sup>10</sup> shows that this effect is likely the result of an additional intramolecular hydrogen bond within the RNA, from the exocyclic amino group of C to the phosphodiester bridging the – 6 and – 7 positions. This presumably stabilizes the conformation of the RNA

in solution closer to that resulting from protein binding reducing the free energy cost of distortion. There are no differences in contacts to the protein. Therefore, the – 5C variant was used in the NMR experiments.

## Reassembly reactions

**Reaction conditions**—Assembly reactions were carried out in 40 mM ammonium acetate, pH 6.8, at 8 °C. Initially, CP<sub>2</sub> was mixed with TR to a final concentration of 8 μM, and further aliquots of CP<sub>2</sub> were added as indicated. The concentration of the CP<sub>2</sub> solution was calculated from UV absorbance, and the sample was passed through a 0.2-μm syringe filter prior to measurement.

**Electrospray ionization-mass spectrometry**—Samples were analyzed by positive ionization nanoelectrospray using an LCT Premier Mass Spectrometer (Waters Corp., Manchester, UK) with collisional cooling capabilities equipped with a NanoMate (Advion, Inc., Ithaca, NY, USA) temperature-controlled automated sample handling and ionization interface, which was thermostated at 8 °C. A capillary voltage of 1.9 kV was set with a nitrogen gas flow of 0.5 psi for sample introduction and ionization. The sampling cone voltage was optimized at 80 V, ion guide 1 at 130 V, and aperture 1 at 60 V. The analyzer pressure was  $7.2 \times 10^{-7}$  mBar and the pressure in the transfer optics region was  $4.9 \times 10^{-3}$  mBar. Data were acquired over the range *m/z* 500–30,000 and data processing was performed using the MassLynx software supplied with the mass spectrometer. An external calibration using CsI clusters was applied to the data.

**Gel filtration-light-scattering assays**—For assays of reassembly by gel filtration, reactions at differing stoichiometries were set up with slight variations on the protocol for mass spectrometry. For convenience, the concentration of CP<sub>2</sub> in the 1:1 reaction was 10 and 20 μM for the 1:2 sample, with the final TR concentration being 10 μM for both. The final 100-μl reaction volume was made up with assembly buffer (40 mM ammonium acetate, pH 6.8), and the samples were incubated on ice for various times until injection onto a Tricon high-performance column of Sepharose 6 10/300 GL (GE Healthcare Bio-Sciences AB, Sweden) thermostated at 12 °C. The eluant was 50 mM Tris acetate, pH 7.4 (chosen to avoid degassing problems with ammonium acetate), at a flow rate 0.36 ml/min.

The outflow of the column was monitored for UV absorption and for light-scattering using a static light-scattering detector PD 2000 (Precision Detectors, Inc., Bellingham, Mass., USA).

**NMR data collection**—Conditions providing optimum solubility of CP<sub>2</sub> were determined from a set of hanging drop trials<sup>40</sup> in a range of buffers at a pH between 6.0 and 7.0 at 25 °C. Suitable solution conditions were thus determined to be 1% v/v glycerol, 100 mM Bis-Tris, pH 7.0, 100 mM KCl. In addition, the buffer also contained 10% D<sub>2</sub>O, 2 mg/ml complete protease inhibitor, 10 mM NaN<sub>3</sub>, 2 mM EDTA, 1 mM DTT, and 0.05 mM sodium 2,2-dimethyl-2-silapentane-5-sulfonate. Samples were introduced into the required solution conditions and concentrated using Vivaspin centrifugal concentrators (Sartorius, Ltd., Epsom, UK). All experiments were carried out in Shigemi microcells (Shigemi, Inc. Allison Park, PA, USA) containing 250–300 μl of solution with concentrations of

CP<sub>2</sub>(W82R) between 0.5 and 1.0 mM. NMR experiments were measured on a Unity Inova 750-MHz NMR spectrometer (Varian, Inc., Palo Alto, CA USA) equipped with four radiofrequency channels and a triple-resonance (<sup>1</sup>H–<sup>15</sup>N–<sup>13</sup>C) room temperature probe with Z axis magnetic field gradients. The set of triple-resonance NMR experiments applied to the CP<sub>2</sub> complex consisted of <sup>1</sup>H–<sup>15</sup>N heteronuclear single-quantum coherence, HNCA, HNcoCA, HNCO, HNcaCO, HNcaCB, and HncocaCB spectra, which were acquired using Varian Biopack pulse sequences. In the case of the HNCA, HNcoCA, HncaCB, and HncocaCB experiments the pulse sequences were modified<sup>41</sup> so as to maintain optimal transverse relaxation optimized spectroscopy effects in partially deuterated proteins during the C<sup>α</sup>/β chemical shift evolution period. The HNcoCA sequence was modified to have a C<sup>α</sup> selective inversion pulse. <sup>1</sup>H–<sup>15</sup>N heteronuclear nuclear Overhauser effect spectra were acquired on the CP<sub>2</sub> dimer using a Varian Unity Inova 600-MHz NMR spectrometer equipped with four radiofrequency channels and a triple-resonance (<sup>1</sup>H–<sup>15</sup>N–<sup>13</sup>C) probe with Z axis magnetic field gradients.

In contrast to the experiments measured on CP<sub>2</sub> the, HNCA and HNcoCA spectra of the CP<sub>2</sub>:TR complex were measured using a Varian cryogenic probe with receiver coils and preamplifiers cooled to – 253 °C. The limited assignment of the CP<sub>2</sub>:TR complex used spectra from <sup>1</sup>H–<sup>15</sup>N heteronuclear single-quantum coherence, HNCA, and HNcoCA experiments. The buffer for the CP<sub>2</sub>:TR samples used for NMR spectroscopy was supplemented with 800 U RNasin (Promega, Southampton, UK), which was brought to the final solution conditions using a Vivaspin centrifugal concentrator. Solutions of the CP<sub>2</sub>:TR complex were prepared by addition of a slight excess of TR to CP<sub>2</sub> as determined by the loss of signals from the *apo* CP<sub>2</sub> dimer.

All experiments were measured with the use of transverse relaxation optimized spectroscopy to reduce T2 relaxation effects and improve dispersion. Data processing and analysis were carried out using NMRPipe<sup>42</sup> and Nmrview<sup>43</sup> on computers running under Linux. Segments of assigned residues from CP<sub>2</sub> were mapped to the primary sequence of the protein using the program Mapper2.<sup>44</sup> <sup>1</sup>H, <sup>13</sup>C, and <sup>15</sup>N chemical shifts were referenced directly (<sup>1</sup>H) and indirectly (<sup>13</sup>C, <sup>15</sup>N) to sodium 2,2-dimethyl-2-silapentane-5-sulfonate.

## Acknowledgements

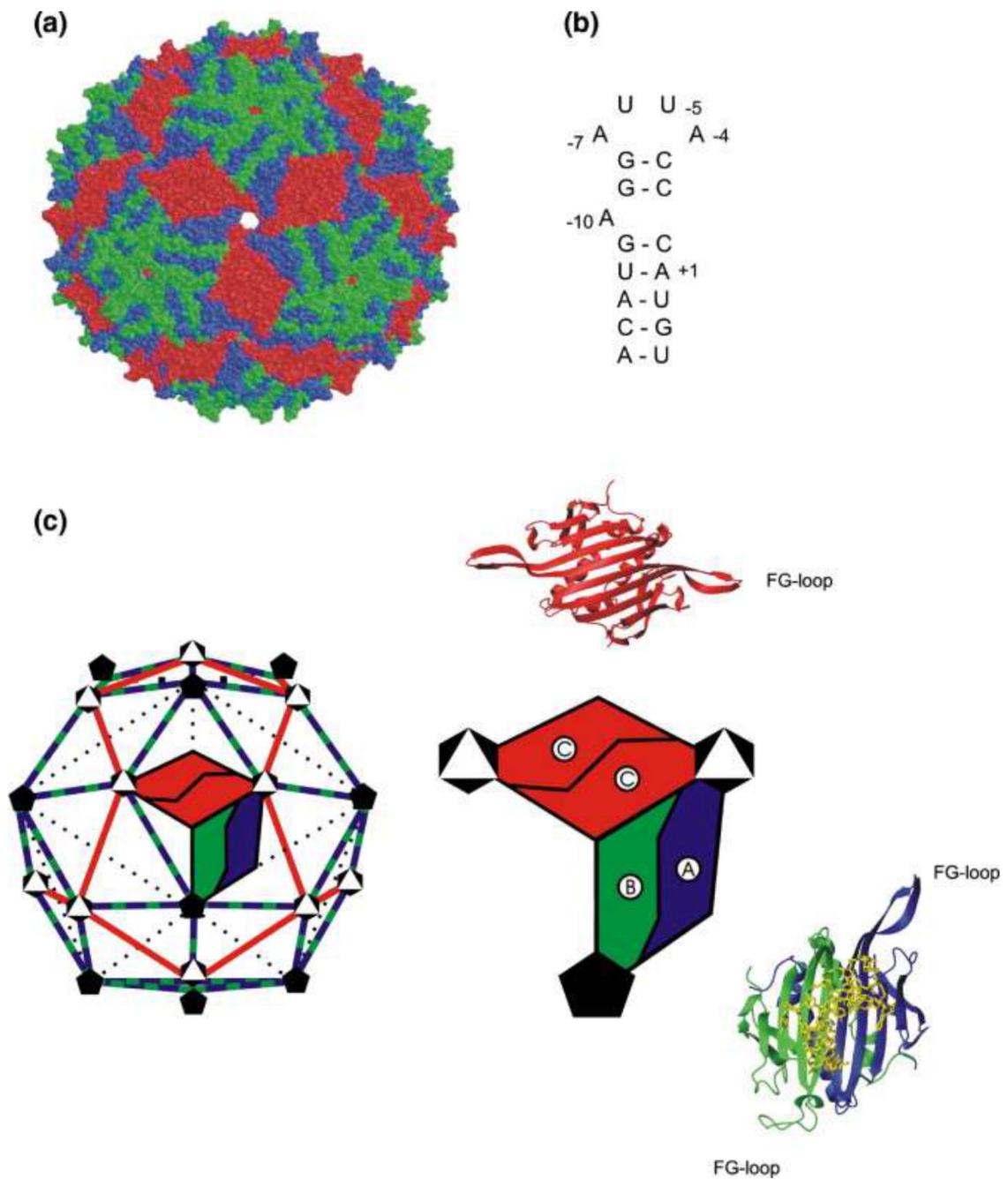
We thank Drs Hugo Lago and Andrew Smith for preliminary mass spectrometry and help with data collection, respectively, and Dr Abdul Rashid for production of some of the materials used in these experiments. We also thank Professors Sheena Radford (Leeds), Bentley Fane (Arizona), and Jon King (Cambridge, Mass.) and Dr Reidun Twarock (York) for their helpful discussions of the results and the manuscript. This work was supported in part by the Leverhulme Trust, the Wellcome Trust, and the UK BBSRC. OR is a Wellcome Trust-funded PhD student.

## References

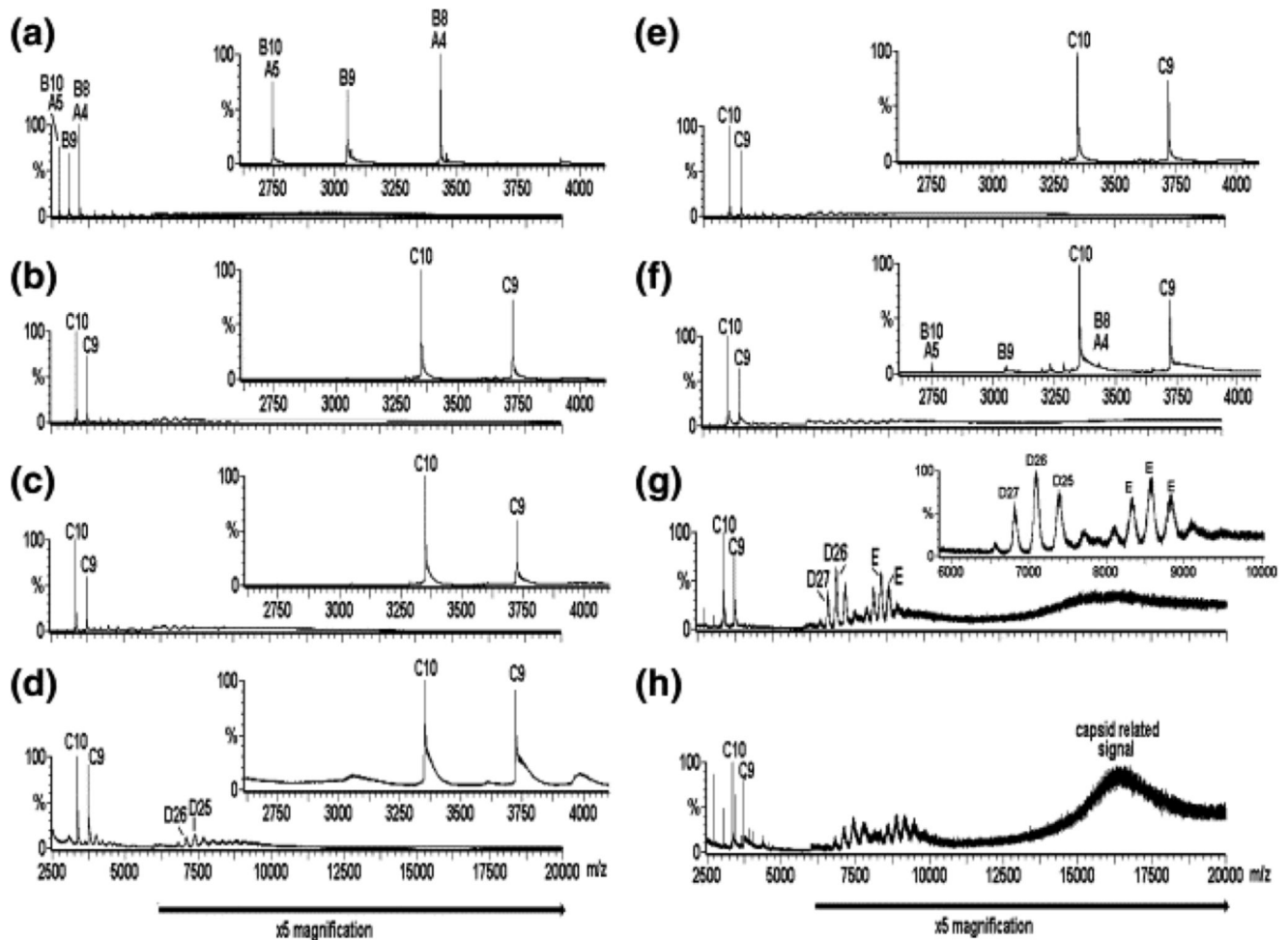
1. Caspar D, Klug A. Physical principles in the construction of regular viruses. Cold Spring Harb Symp Quant Biol. 1962; 27: 1–24. [PubMed: 14019094]
2. Twarock R. A tiling approach to virus capsid assembly explaining a structural puzzle in virology. J Theoret Biol. 2004; 226: 477–482. [PubMed: 14759653]
3. Valegård K, Liljas L, Fridborg K, Unge T. The three-dimensional structure of the bacterial virus MS2. Nature. 1990; 345: 36–41. [PubMed: 2330049]

4. Golmohammadi R, Valegård K, Fridborg K, Liljas L. Allosteric Conformer Switching in Virus Assembly. The refined structure of bacteriophage MS2 at 2.8 Å resolution. *J Mol Biol.* 1993; 234: 620–639. [PubMed: 8254664]
5. Sorger PK, Stockley P, Harrison S. Structure and assembly of turnip crinkle virus. 2. Mechanism of Reassembly in vitro. *J Mol Biol.* 1986; 191: 639–658. [PubMed: 3806677]
6. Rossmann MG, Abad-Zapatero C, Erickson JW, Savithri HS. RNA-protein interactions in some small plant viruses. *J Biomol Struct Dyn.* 1983; 1: 565–579. [PubMed: 6401119]
7. Rossmann MG, Johnson JE. Icosahedral RNA virus structure. *Annu Rev Biochem.* 1989; 58: 533–573. [PubMed: 2673017]
8. Zlotnick A, Aldrich R, Johnson JM, Ceres P, Young MJ. Mechanism of capsid assembly for an icosahedral plant virus. *Virology.* 2000; 277: 450–456. [PubMed: 11080492]
9. Witherell GW, Gott JM, Uhlenbeck OC. Specific interaction between RNA phage coat proteins and RNA. *Prog Nucl Acid Res Mol Biol.* 1991; 40: 185–220.
10. Valegård K, Murray JB, Stockley PG, Stonehouse NJ, Liljas L. Crystal structure of a bacteriophage RNA coat protein operator system. *Nature.* 1994; 371: 623–626. [PubMed: 7523953]
11. Stockley PG, Stonehouse NJ, Valegård K. Molecular mechanism of RNA phage morphogenesis. *Int J Biochem.* 1994; 26: 1249–1260. [PubMed: 7851629]
12. Lago H, Parrott AM, Moss T, Stonehouse NJ, Stockley PG. Probing the kinetics of formation of the bacteriophage MS2 translational operator complex: identification of a protein conformer unable to bind RNA. *J Mol Biol.* 2001; 305: 1131–1144. [PubMed: 11162119]
13. Horn WT, Tars K, Grahn E, Helgstrand C, Baron AJ, Lago, et al. Structural basis of RNA binding discrimination between bacteriophages Q $\beta$  and MS2. *Structure.* 2006; 14: 487–495. [PubMed: 16531233]
14. Beckett D, Uhlenbeck OC. Ribonucleoprotein complexes of R17 coat protein and a translational operator analog. *J Mol Biol.* 1988; 204: 927–938. [PubMed: 3221400]
15. Beckett D, Wu H-N, Uhlenbeck OC. Roles of operator and non-operator RNA sequences in bacteriophage R17 capsid assembly. *J Mol Biol.* 1988; 204: 939–947. [PubMed: 3221401]
16. Loo JA. Studying noncovalent protein complexes by electrospray ionization mass spectrometry. *Mass Spectrom Rev.* 1997; 16: 1–23. [PubMed: 9414489]
17. Ashcroft AE. Recent developments in electrospray ionisation mass spectrometry: noncovalently bound protein complexes. *Nat Prod Rep.* 2005; 22: 452–464. [PubMed: 16047045]
18. Krutchinsky AN, Chernushevich IV, Spicer VL, Ens W, Standing KG. Collisional damping interface for an electrospray ionization time-of-flight mass spectrometer. *J Am Soc Mass Spectrom.* 1998; 9: 569–579.
19. Chernushevich IV, Thomson BA. Collisional cooling of large ions in electrospray mass spectrometry. *Anal Chem.* 2004; 76: 1754–1760. [PubMed: 15018579]
20. Van Berkel WJH, Van den Heuvel RHH, Versluis C, Heck AJR. Detection of intact megaDalton protein assemblies of vanillyl alcohol oxidase by mass spectrometry. *Protein Sci.* 2000; 9: 435–439. [PubMed: 10752605]
21. Sobott F, Hernandez H, McCammon MG, Tito MA, Robinson CV. A tandem mass spectrometer for improved transmission and analysis of large macromolecular assemblies. *Anal Chem.* 2002; 74: 1402–1407. [PubMed: 11922310]
22. Sugiyama T, Nakada D. Control of translation of MS2 RNA cistrons by MS2 coat protein. *Biochemistry.* 1967; 57: 1750–1774.
23. Stockley PG, Ashcroft AE, Francese S, Thompson GS, Ranson NA, Smith AM, et al. Dissecting the fine details of assembly of a T=3 phage capsid. *J Theoret Med.* 2005; 6: 119–125.
24. Mastico RA, Talbot SJ, Stockley PG. Multiple presentation of foreign peptides on the surface of an RNA-free spherical bacteriophage capsid. *J Gen Virol.* 1993; 74: 541–548. [PubMed: 7682249]
25. Tito MA, Tars K, Valegard J, Hajdu J, Robinson CV. Electrospray time-of-flight mass spectrometry of the intact MS2 virus capsid. *J Am Chem Soc.* 2000; 122: 3550–3551.
26. Carey J, Uhlenbeck OC. Kinetic and thermodynamic characterization of the R17 coat protein-ribonucleic acid interaction. *Biochemistry.* 1983; 22: 2610–2615. [PubMed: 6347248]

27. Talbot SJ, Goodman S, Bates SRE, Fishwick CWG, Stockley PG. Use of synthetic oligoribonucleotides to probe RNA-protein interactions in the MS2 translational operator complex. *Nucl Acids Res.* 1990; 18: 3521–3528. [PubMed: 1694577]
28. Koradi R, Billeter M, Wüthrich K. MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graphics.* 1996; 14: 51–55.
29. Case DA. Calibration of ring-current effects in proteins and nucleic acids. *J Biomol NMR.* 1995; 6: 341–346. [PubMed: 8563464]
30. Valegård K, Murray JB, Stonehouse NJ, van den Worm S, Stockley PG, Liljas L. The three dimensional structures of two complexes between recombinant MS2 capsids and RNA operator fragments reveal sequence specific protein-RNA interactions. *J Mol Biol.* 1997; 270: 724–738. [PubMed: 9245600]
31. Koning R, van den Worm S, Plaisier JR, van Duin J, Pieter Abrahams J, Koerten H. Visualization by cryo-electron microscopy of genomic RNA that binds to the protein capsid inside bacteriophage MS2. *J Mol Biol.* 2003; 332: 415–422. [PubMed: 12948491]
32. Berkhout B, van Duin J. Mechanism of translational coupling between coat protein and replicase genes of RNA bacteriophage MS2. *Nucl Acids Res.* 1985; 13: 6955–6967. [PubMed: 3840590]
33. Caspar D. Movement and self-control in protein assemblies. Quasi-equivalence revisited. *Biophys J.* 1980; 32: 103–138. [PubMed: 6894706]
34. Bamford DH, Grimes JM, Stuart DI. What does structure tell us about virus evolution? *Curr Opin Struct Biol.* 2005; 15: 655–663. [PubMed: 16271469]
35. Kuznetsov YG, Daijogo S, Zhou J, Semler BL, McPherson A. Atomic force microscopy analysis of icosahedral virus RNA. *J Mol Biol.* 2005; 347: 41–52. [PubMed: 15733916]
36. Wu M, Brown WL, Stockley PG. Cellspecific delivery of bacteriophage-encapsidated ricin- A chain. *Bioconjug Chem.* 1995; 6: 587–595. [PubMed: 8974458]
37. Wang Q, Lin TW, Tang L, Johnson JE, Finn MG. Icosahedral virus particles as addressable nanoscale building blocks. *Angew Chem Int Ed Engl.* 2002; 41: 459–462. [PubMed: 12491378]
38. Ni CZ, Kodandapani R, Wickersham J, Peabody DS, Ely KR, Syed R. Crystal structure of the MS2 coat protein dimer: implications for RNA binding and virus assembly. *Structure.* 1995; 3: 255–263. [PubMed: 7788292]
39. Murray JB, Collier AK, Arnold JRP. A general purification procedure for chemically synthesised oligoribonucleotides. *Anal Biochem.* 1994; 218: 177–184. [PubMed: 7519835]
40. Lepre CA, Moore JM. Microdrop screening: a rapid method to optimize solvent conditions for NMR spectroscopy of proteins. *J Biomol NMR.* 1998; 12: 493–499. [PubMed: 9862127]
41. Eletsky A, Kienhofer A, Pervushin K. TROSY NMR with partially deuterated proteins. *J Biomol NMR.* 2001; 2: 177–180.
42. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe—a multidimensional spectral processing system based on unix pipes. *J Biomol NMR.* 1995; 6: 277–293. [PubMed: 8520220]
43. Johnson BA, Blevins RA. NMR View—a computer-program for the visualization and analysis of NMR data. *J Biomol NMR.* 1994; 4: 603–614. [PubMed: 22911360]
44. Güntert P, Salzmann M, Braun D, Wüthrich K. Sequence-specific NMR assignment of proteins by global fragment mapping with the program MAPPER. *J Biomol NMR.* 2000; 18: 129–137. [PubMed: 11101217]

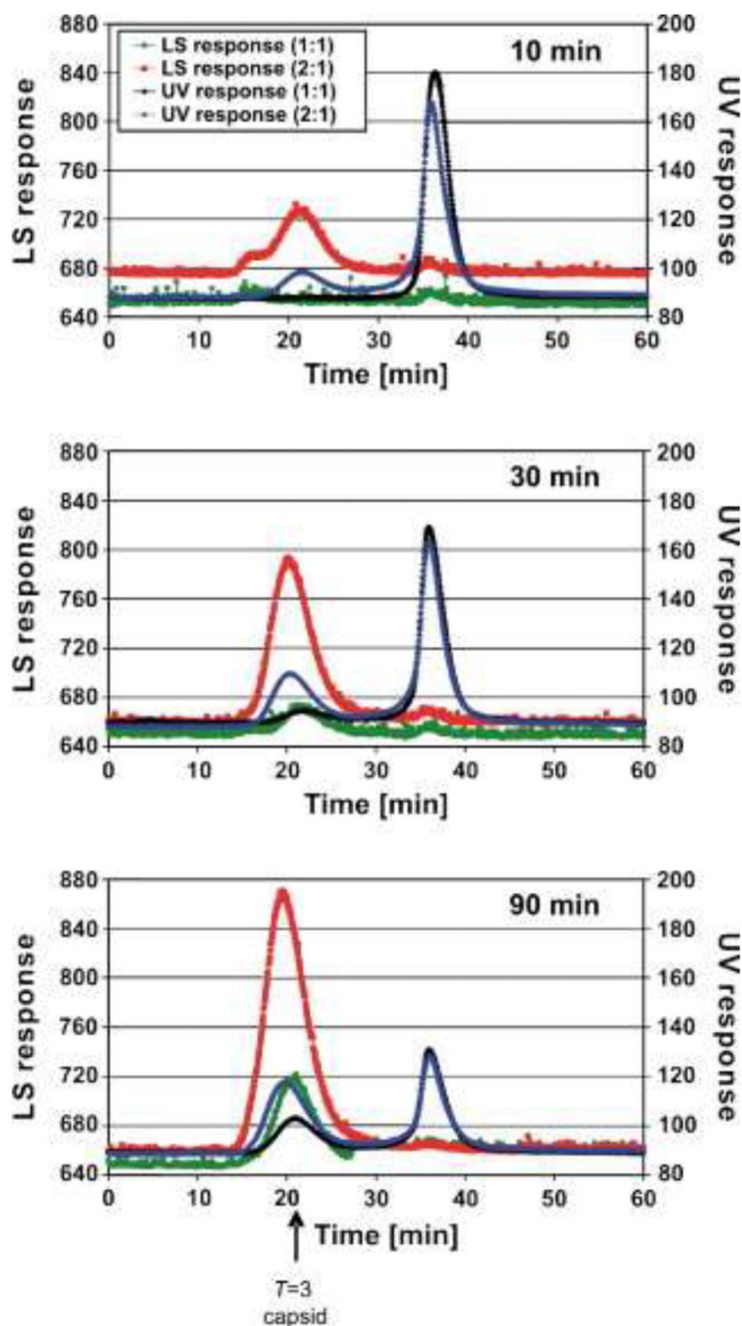


**Figure 1. The structures and solution components associated with MS2 bacteriophage assembly.** (a) A space-filling representation of the X-ray structure of the  $T=3$  capsid of wild-type MS23 (Protein Data Bank 2MS2). (b) The sequence and secondary structure of the TR RNA. (c) The distribution of symmetry axes (black and white symbols) within the icosahedral surface lattice of the MS2 capsid and the arrangement of the quasiequivalent 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.



**Figure 2.**

Virus capsid reassembly monitored over time by real-time nanoESI-MS. The spectra were acquired over the range  $m/z$  500–30,000 for samples in ammonium acetate (40 mM) at pH 5.2–5.7. The components are labeled as follows: A = CP; B = CP<sub>2</sub>; C = CP<sub>2</sub>:TR; D = 182.7 kDa, assigned as [3(CP<sub>2</sub>:TR)+3CP<sub>2</sub>], and E = ~ 288–300 kDa, unassigned. The number immediately following each letter is the charge state of those particular ions, i.e.,  $[M+nH]^{n+}$ . The spectra are as follows: (a) CP alone (8  $\mu$ M); pH 3.2; (b)–(d) are for reassembly at pH 5.2–5.7 at a stoichiometry of CP<sub>2</sub> to TR 1:1 (8  $\mu$ M:8  $\mu$ M), at 1, 90, 300 min, respectively. Spectrum (e) is the 1:1 pre-equilibrated starting point for reassembly at a stoichiometry of CP<sub>2</sub> to TR 2:1 (16  $\mu$ M:8  $\mu$ M), spectra (f)–(h), at 1, 120, and 180 min, respectively. Note that spectra (b) and (e) are essentially equivalent. The bars below the spectra indicate a magnification factor of 5 for all ions above  $m/z$  6000 to enhance clarity. The insets in (a)–(f) highlight the range  $m/z$  2600–4100 to enhance clarity in the region containing ions arising from components A, B, and C. The inset in (g) highlights the range  $m/z$  6000–10,000 to enhance clarity of the ions arising from components D and E and is normalized to the D26 ions, which are ~10% of the intensity of the full spectrum.

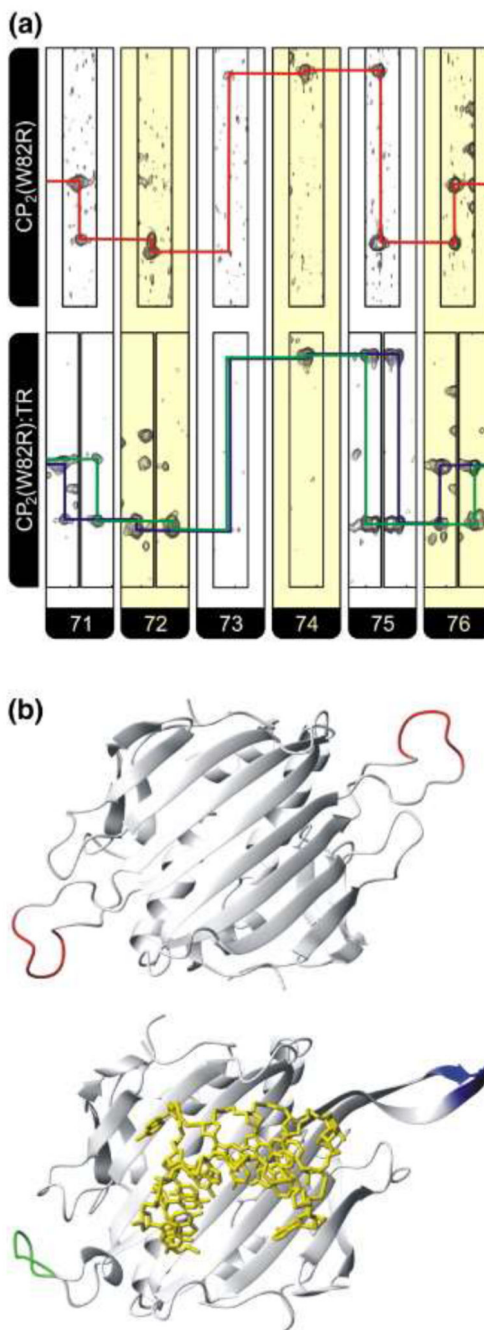


**Figure 3.**

Gel filtration-light-scattering assays of capsid reassembly. MS2 CP<sub>2</sub> and TR were mixed in 40 mM ammonium acetate, pH 5.2–5.7, to form a 1:1 reaction, incubated at 4 °C for 10, 30, or 90 min, and then loaded onto a Sepharose 6 gel filtration column equilibrated in 50 mM Tris–acetate, pH 7.4, and eluted at 0.36 ml/min. The outflow from the column was analyzed simultaneously via UV absorbance (black trace) and light-scattering (green trace). Similar reactions were then set up, preincubated at a molar ratio of 1:1 for 10 min, and then an additional aliquot of protein was added to create the 2:1 reaction. The traces are shown as



UV absorbance (blue trace) and light-scattering (red trace). The position at which authentic  $T=3$  capsids elute is marked with an arrow; the peak at  $\sim 35$  min corresponds to the starting materials and smaller complexes.

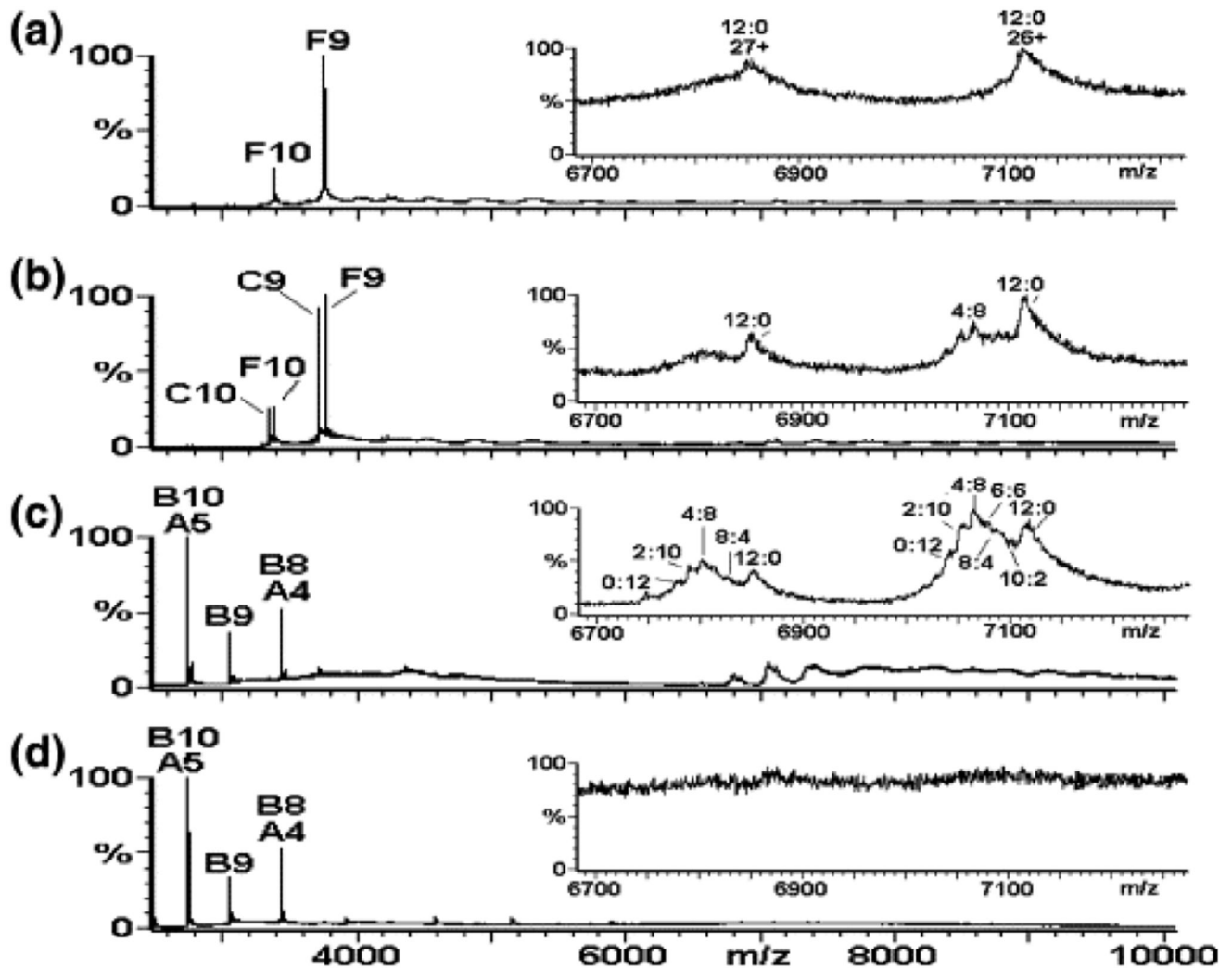


**Figure 4. Observed changes in the chemical shifts for the FG loop residues 71–76 in both *apo* and RNA-bound complexes.**

(a) Strips, alternately colored white and yellow for each residue, from HNCA spectra of CP<sub>2</sub>(W82R) (top) and CP<sub>2</sub>(W82R):TR (bottom). The assignment pathway for CP<sub>2</sub>(W82R) is shown in red while the split pathways for the complex are in blue and green. Note that no assignment of resonances to specific FG loop conformations is made or assumed.

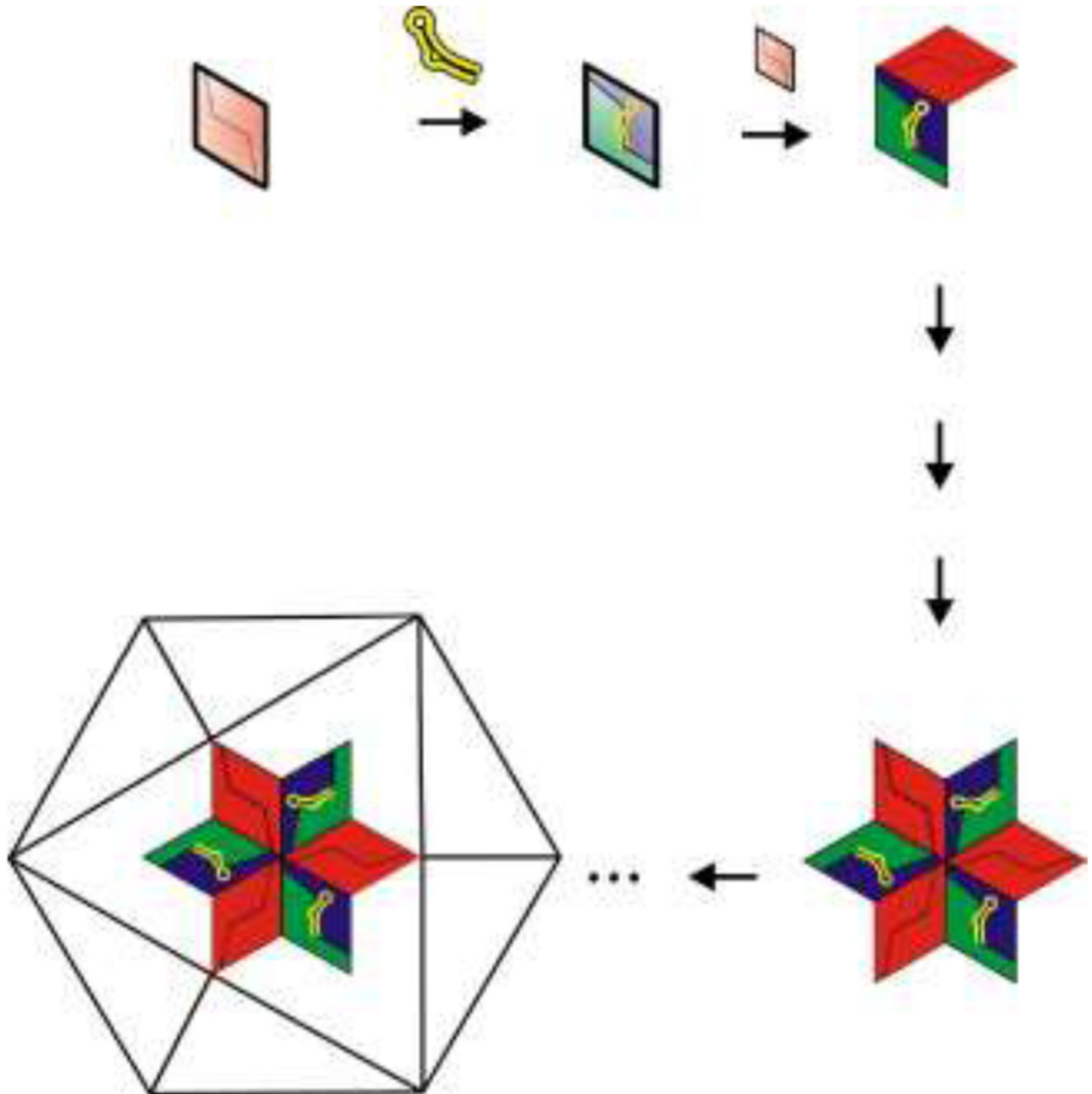
(b) Cartoon representation of (top) the backbone of CP<sub>2</sub>(W82R) showing the position of residues 71–76 in red for the two CP subunits derived from the crystal structure of this mutant<sup>38</sup> (Protein Data Bank 1MSC) and (bottom) the equivalent view for a wild-type A/B

dimer (green/blue) bound to TR RNA (yellow sticks), which is the conformation we assume is closest to that present in solution.



**Figure 5.**

Mapping the capsid reassembly pathway using isotopic chase experiments. Virus capsid reassembly from  $^{15}\text{N}$ -labeled  $\text{CP}_2$  chased by adding increasing quantities of  $^{14}\text{N}$ - $\text{CP}_2$ , with each reaction monitored by real-time nanoESI-MS at  $t = 1$  min. The spectra were acquired over the range  $m/z$  500–10,000 for samples in 40 mM ammonium acetate, pH 5.2–5.7. The larger panels show the range  $m/z$  2500–10,000 and the components are labeled as follows: A = CP; B =  $\text{CP}_2$ ; C =  $\text{CP}_2$ :TR; F =  $^{15}\text{N}$ - $\text{CP}_2$ :TR. The number immediately following each letter is the charge state of those particular ions. Spectra are (a)  $^{15}\text{N}$ - $\text{CP}_2$  to TR 1:1 (8  $\mu\text{M}$ :8  $\mu\text{M}$ ),  $t = 1$  min; (b)  $^{15}\text{N}$ - $\text{CP}_2$  to TR 1:1 (8  $\mu\text{M}$ :8  $\mu\text{M}$ ) with  $^{14}\text{N}$ - $\text{CP}_2$  (8  $\mu\text{M}$ ) added,  $t = 1$  min; (c)  $^{15}\text{N}$ - $\text{CP}_2$  to TR 1:1 (8  $\mu\text{M}$ :8  $\mu\text{M}$ ) with  $^{14}\text{N}$ - $\text{CP}_2$  (24  $\mu\text{M}$ ) added,  $t = 1$  min; (d)  $^{15}\text{N}$ - $\text{CP}_2$  to TR 1:1 (8  $\mu\text{M}$ :8  $\mu\text{M}$ ) with  $^{14}\text{N}$ - $\text{CP}_2$  (32  $\mu\text{M}$ ) added,  $t = 1$  min. Inset spectra highlight the range  $m/z$  6700–7200, showing the 26+ and 27+ charge states of the major reassembly intermediate  $[3(\text{CP}_2:\text{TR})+3\text{CP}_2]$ . The peaks are labeled with the ratio of  $^{15}\text{N}$ : $^{14}\text{N}$  CP contained in the species, i.e., 12:0, 10:2, 8:4, 6:6, 4:8, 2:10, and 0:12.



**Figure 6.**

Model for the initiation of  $T=3$  capsid assembly. A pair of chemically equivalent (C/C-like) subunits in solution (top left, shaded pink) binds a TR RNA, resulting in a conformational switch into an A/B-like asymmetric dimer (shaded green and blue). This complex binds an additional C/C-like dimer, creating the first higher-order intermediate on the pathway to capsid formation. A variety of subpathways involving the addition of coat protein dimers (arrows) can then generate the major assigned intermediate encompassing the first threefold axis of the virus particle. We have mass assignments for all the cartoon species shown in the

diagram, as well as a number of components that may be in the subassembly pathway, such as  $2[\text{CP}_2:\text{TR}] + 2\text{CP}_2$  &  $3[\text{CP}_2:\text{TR}] + 2\text{CP}_2 + \text{CP}$ . These latter species are not always present in the reactions and when present are in trace amounts.