doi:10.1016/j.jmb.2005.11.040 J. Mol. Biol. (2006) 355, 1095–1111

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

The MS2 Coat Protein Shell is Likely Assembled Under Tension: A Novel Role for the MS2 Bacteriophage A Protein as Revealed by Small-angle Neutron Scattering

Deborah A. Kuzmanovic[1,3], Ilya Elashvili[2], Charles Wick[2] Catherine O'Connell[1] and Susan Krueger[3]*

[1]*Biotechnology Division NIST, 100 Bureau Drive, Stop 8311, Gaithersburg, MD 20899-8311, USA*

[2]*Edgewood Chemical Biological Center, 5183 Blackhawk Rd Aberdeen Proving Ground MD 21010, USA*

[3]*NIST Center for Neutron Research, NIST, 100 Bureau Drive, Stop 8562, Gaithersburg MD 20899-8562, USA*

**Corresponding author*

Recombinant forms of the bacteriophage MS2 and its RNA-free (empty) MS2 capsid were analyzed in solution to determine if RNA content and/or the A (or maturation) protein play a role in the global arrangement of the virus protein shell. Analysis of the (coat) protein shell of recombinant versions of MS2 that lack the A protein revealed dramatic differences compared to wild-type MS2 in solution. Specifically, A protein-deficient virus particles form a protein shell of between 31($\pm$1) Å and 37($\pm$1) Å. This is considerably thicker than the protein shell formed by either the wild-type MS2 or the RNA-free MS2 capsid, whose protein shells have a thickness of 21($\pm$1) Å and 25($\pm$1) Å, respectively. Since the A protein is known to separate from the intact MS2 protein shell after infection, the thin shell form of MS2 represents the pre-infection state, while the post-infection state is thick. Interestingly, these A protein-dependent differences in the virus protein shell are not seen using crystallography, as the crystallization process seems to artificially compact the wild-type MS2 virion. Furthermore, when the A protein is absent from the virus shell (post-infection), the process of crystallization exerts sufficient force to convert the protein shell from the post-infection (thick) state to the pre-infection (thin) conformation. In summary, the data are consistent with the idea that RNA content or amount does not affect the structure of the MS2 virus shell. Rather, the A protein influences the global arrangement of the virus coat dramatically, possibly by mediating the storage of energy or tension within the protein shell during virus assembly. This tension may later be used to eject the MS2 genomic RNA and A protein fragments into the host during infection.

Published by Elsevier Ltd.

Keywords: bacteriophage MS2; contrast variation; MS2 A protein; RNA ejection; small-angle neutron scattering

Introduction

MS2 is a small enteric RNA bacteriophage that belongs to the family Leviviridae, which infects F+ (male) *Escherichia coli*. Bacteriophage MS2 is an ideal model system for the study of many macromolecular processes but is especially well-suited for the study of virus structure and assembly. The infective MS2 virion is composed of a coat protein (M_r= 13,700), a maturation (or A) protein (M_r=44,000) and a single-stranded molecule of RNA. The coat protein is the major component of the MS2 protein shell. In solution, the MS2 coat protein has been shown to form a homodimer. In all, 90 coat protein homodimers, in conjunction with a single A protein

Present addresses: D. A. Kuzmanovic, Geo-Centers, Inc., Gunpowder Branch, P.O. Box 68, Aberdeen Proving Ground, MD 21010, USA; I. Elashvili, Chem/Bio Technology Transition Division, Defense Threat Reduction Agency, 8725 John. J. Kingman Rd, Ft. Belvoir, VA 22060, USA; C. O'Connell, Tetracore, Inc., 11 Firstfield Rd., Suite C, Gaithersburg, MD 20878, USA.

Abbreviations used: SANS, small-angle neutron scattering; SAXS, small-angle X-ray scattering; SLD, scattering length density; cryo-EM, cryo-electron microscopy.

E-mail address of the corresponding author: susan.krueger@nist.gov

0022-2836/$ - see front matter Published by Elsevier Ltd.

molecule, assemble to form the 275 Å MS2 virus (coat) protein shell.[1] The three-dimensional X-ray crystal structures of the coat protein dimers, as well as the complete MS2 virus, have been reported and a great deal is known about their structures.[2–5]

The crystal structure of MS2 has revealed a number of key features about the overall structure of the coat protein shell. (1) The coat protein dimer is the major subunit of virus assembly and can adopt three possible quasi-equivalent conformers (A, B, and C) in the formation of the MS2 icosahedral shell. (2) These conformers, although very similar, vary primarily at the F-G β-strand loop (F-G-loop). In the A and C subunits, which interact at the quasi 6-fold axis, the FG loop is extended, while in the B subunit it is folded back in the direction of the protein at the 5-fold axis. (3) A consequence of the interaction of the FG-loop is that a hole or channel is created at both the 5-fold axis and the quasi 6-fold axis; one of these channels may be a site of RNA extrusion during infection.[3–5]

Crystallographic analysis of the MS2 virus could not resolve the location of the A protein in the structure of the MS2 protein shell.[3–5] However, chemical labeling and antibody-binding experiments suggest that the A protein is at least partially exposed on the capsid surface, which supports the idea that the A protein comprises one of the vertices of the MS2 icosahedral shell.[6,7]

The MS2 virion contains a 3569 nt positive-sense, single-stranded RNA and its complete genomic sequence has been reported.[8] The genomic RNA contains a number of structural motifs, mainly stem–loop structures, which have been shown to act as sites for the control of virus assembly and translation.[8–11] In particular, assembly of the MS2 virus coat has been shown to be catalyzed by binding to a specific 19 nt stem–loop operator sequence of the genomic RNA *in vitro*.[1,12] In addition to this function, the coat protein binds to the MS2 operator site and acts as a transcriptional repressor of the MS2 replicase cistron.[1]

In addition to MS2 coat protein–RNA interactions, a number of lines of evidence point to tight interactions between the A protein and the MS2 genomic RNA. First, in sedimentation experiments, differentially radiolabeled A protein and genomic RNA are known to be co-sedimented selectively, forming a peak in the same part of the gradient.[13,14] Secondly, RNAse protection and competition binding experiments indicate that the MS2 RNA is bound tightly by the A protein at its 5′ and 3′ ends.[15] Thirdly, this tight association between the genomic RNA and the A protein has been shown to be important for RNA packing *in vivo*. Mutant viruses that lack the A protein contain loosely packed RNA that protrudes from the protein shell and then becomes degraded by nucleases in the medium.[16,17]

Recent cryo-EM analysis of the MS2 virus reveals a network of RNA bound to the inner surface of the protein shell, which forms an icosahedron with RNA-free holes at the 5-fold axis and the quasi 6-fold axis. This ordered RNA, which is thought to represent only a subset of the genomic RNA, is presumably attached to the coat protein dimers *via* RNA stem–loop motifs.[18]

This intimate association between the RNA and the protein shell, as well as the interactions between the RNA and the A protein, suggest the possibility of a link between RNA packing and the global packing arrangement of the final virus protein shell. This study seeks to determine if the A protein or the RNA content influences the overall packing arrangement of the MS2 protein shell, using small-angle neutron scattering (SANS) and the contrast variation technique.

In SANS experiments, samples are exposed to neutrons in solution and the resulting scattering pattern is used to determine the average size, shape, molecular mass and orientation of the sample.[19,20] For complex macromolecular structures like viruses, SANS can be combined with the contrast variation technique to gain structural information about the protein, nucleic acid and lipid components separately.[20,21] Since the contrast variation technique simply involves varying the ratio of solvent to deuterated water, biological samples are analyzed under physiological conditions and the data obtained can reveal *in situ* structural information.[19,21,22]

Analysis of SANS data typically occurs in three steps. (1) The SANS scattering pattern of the protein or complex is determined experimentally and the data are compared or "fit" to a number of different idealized shapes until a shape is found that best describes the experimental structure seen in solution.[22,23] (2) If a crystal structure of the protein or complex is known, it is used as a starting point for the model solution structure. This is useful, because crystal structures provide structural information of higher resolution than SANS. The crystal structure coordinates are used to generate model SANS data.[24] (3) Finally, both the measured experimental SANS data and the model SANS data generated from the crystal structure are compared to the best real-space physical model (from step (1)) that describes the shape and other dimensions of the structure.[22] This sequence of analysis of SANS data is often especially powerful in allowing *in situ* comparisons of very similar structures.[25–27]

Examination of the structure of the MS2 virus is made particularly amenable to SANS analysis because the wild-type and a variety of mutant versions of MS2 viruses have been described genetically,[28,29] and high-resolution crystal structures of mutant and wild-type MS2 viruses are available.[3,4,30–32] Additionally, it is now possible to generate recombinant forms of MS2 that lack the A protein and contain various amounts of RNA.[33–35] It is possible to create RNA-free versions of MS2, called MS2 capsids.[29,36,37] Furthermore, a recent experimental analysis of the wild-type MS2 virus using SANS and small-angle X-ray scattering (SAXS) has been reported, and a model of its physical structure in solution has been described.[38,39]

To better understand the influence of RNA and the A protein on protein shell arrangement *in vivo*, we present the use of SANS and the contrast variation technique to solve the structures in solution of the (coat) protein shell component of two types of MS2 viruses: (a) wild-type MS2 (WT MS2) containing coat protein, A protein and RNA; and (b) two recombinant forms of MS2 virus lacking the A protein but containing various amounts of RNA, (MS2-HCV and MS2-λ). In addition, empty capsids (MS2-capsid), containing coat protein and A protein, but lacking RNA, were measured for comparison to the protein shell structures of the above-mentioned samples.

Results

RNA content does not affect the packing arrangement of the MS2 protein shell

The effect of RNA content on the overall packing arrangement of the MS2 (coat) protein shell was analyzed by comparing the structures of the wild-type MS2 protein shell with that of its RNA-free capsid. The quality of the WT MS2 and MS2-capsid samples used for these studies were analyzed by SDS/PAGE. The results are shown in Figure 1 and a complete strain list is given in Table 1. Notice that all samples are free of contaminants, and that the A protein is visible in the empty MS2-capsid sample as well as in the wild-type MS2 virion. The presence of the A protein in the MS2 empty capsids was surprising, given the fact that the MS2 genomic RNA and A protein have been shown to be bound tightly at the 5′ and 3′ ends,[15] and the MS2 capsids were purified under conditions that selectively precipitated the RNA (and presumably the A protein) from the MS2 virus coat protein prior to permitting the coat proteins to reassemble into capsids. Since these MS2 capsids were generated *in vitro*, we first sought to determine if the structure of the MS2-capsid coat protein shell was similar to that of WT MS2.

Figure 2 shows the measured SANS data for the MS2-capsid sample, as well as for, selectively, the protein component of the WT MS2 protein shell. The scattering from the WT MS2 protein shell was separated from that of its RNA component using the contrast variation technique, as described.[38] MS2-capsid was measured in 100% ^2H_2O TSM buffer (see Materials and Methods). Note that the locations of the peaks in the scattered intensity are the same, at approximately $Q=0.03$ Å^{-1}, $Q=0.06$ Å^{-1} and $Q=0.09$ Å^{-1} in both cases. This indicates that the scattering from WT MS2 and the MS2-capsid protein shells essentially arise from structures that are very similar. On the other hand, the peak intensities differ somewhat, with those of the MS2-capsid sample being less intense. This is likely due to more polydispersity in the empty MS2-capsid sample as compared to the WT MS2 protein shell.

The existence of polydispersity can be determined by calculating the molecular mass (M) of the MS2-capsid sample from the forward scattered intensity, $I(0)$, using equation (2) (see Materials and Methods). Both the radius of gyration, R_g, and $I(0)$ can be obtained from the data by two methods. Guinier fits can be made to the data, using only the lowest Q data points. On the other hand, R_g and $I(0)$ are also deduced from the calculation of the distance distribution function, $P(r)$, which makes use of all of the data points, as explained in Materials and Methods. As such, the R_g and $I(0)$ values obtained from $P(r)$ are less influenced by aggregation effects, but more influenced by smaller particles that may be present in the sample. Thus, by comparing the R_g and $I(0)$ values obtained by both methods, the existence of aggregation or dissociation in the sample can sometimes be discerned.

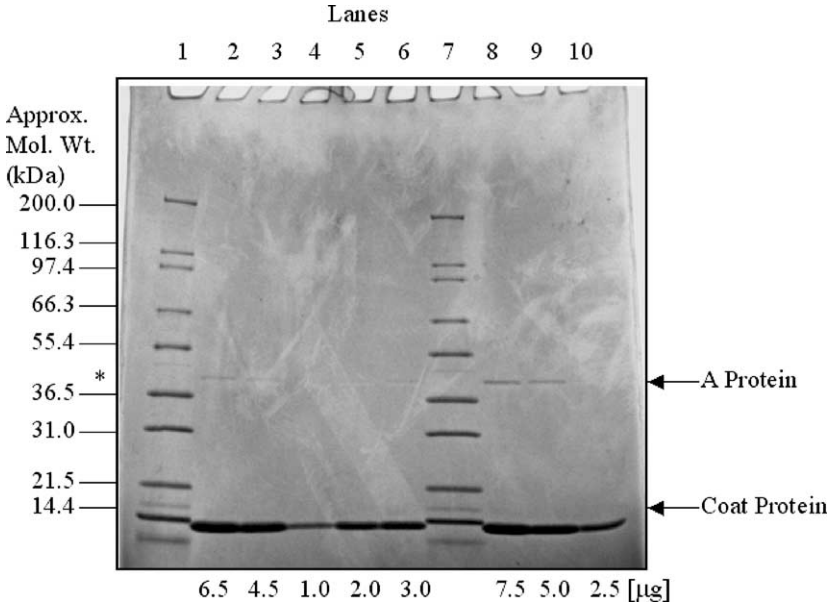

Figure 1. Results from denaturing SDS/PAGE. Lanes 1 and 7 contain molecular mass markers. Lanes 2 and 3 contain empty MS2-capsid. Lanes 4–6 and 8–10 contain dilutions of the WT MS2 virus.

Table 1. List of compared structures for various MS2 and related virus strains

	Genotype				
	A protein	Coat protein	RNA	Description	Reference
A. *Measured data*					
MS2	+	+	+	WT phage MS2	ATCC strain 15597-B1
MS2-capsid	+	+	−	MS2/empty capsid	This work
MS2-HCV	−	+	+*	MS2/RNA fragment	33
MS2-λ	−	+	+*	MS2/RNA fragment	Ambion, personal communication
B. *Model data PDB ID*					
2MS2 (I)	+	+	+	WT phage MS2	3
1BMS	−	+*	−	MS2/coat protein mutation P78N	30
1MST	−	+*	−	MS2/coat protein mutation E76D	30
1ZDI	−	+	+*	MS2/RNA fragment	31
1ZDH	−	+	+*	MS2/RNA fragment	31
1FRS (I)	−	+	−	Phage fr	32
1FR5	−	+*	−	Phage fr/4 aa coat protein deletion	44
1GAV (II)	+	+	+	WT phage GA	40
1QBE (III)	+	+	+	WT phage QB	41

+*, Present but mutant or recombinant. The virus family Leviviridae is divided into four groups I, II, III, and IV. MS2 (and fr), Qβ and GA are members of groups I–III, respectively. No group IV family member has been crystallized to date.

Guinier fits to the MS2-capsid data resulted in $R_g = 130(\pm 3)$ Å and $I(0) = 1.3(\pm 0.1)\ cm^{-1}$. The same parameters were determined from the $P(r)$ calculation, using D_{max} values between 280 Å and 320 Å, resulting in $R_g = 123(\pm 2)$ Å and $I(0) = 1.28(\pm 0.06)\ cm^{-1}$. Using a value of $I(0) \sim 1.3\ cm^{-1}$, the M value for the MS2-capsid sample was calculated to be 2.7 $(\pm 0.4) \times 10^6$ g/mol, as shown in Table 2. This is in good agreement with the expected value of 2.5×10^6 g/mol, indicating that the MS2-capsids are essentially intact, containing 180 copies of the coat protein. This suggests that the polydispersity in the sample is not due to incomplete capsid formation. Rather, it might be caused by the existence of free coat protein dimers in solution. The low value of M for these dimers, compared to that of the capsid particles, means that their existence would have little effect on $I(0)$. The existence of free dimers would also have little effect on the R_g value of the MS2-capsid sample if only the low Q data points are used, such as in the Guinier analysis. However, the effect of these free dimers is seen by a decrease in R_g when all of the data points are used in the calculation of $P(r)$. Table 2 shows the calculated M for the protein shell component of WT MS2, determined from our earlier SANS studies,[38] to be 2.5 $(\pm 0.3) \times 10^6$ g/mol. Note that the uncertainties in the M determinations for both the MS2-capsid and WT MS2 protein shell are too large to ascertain

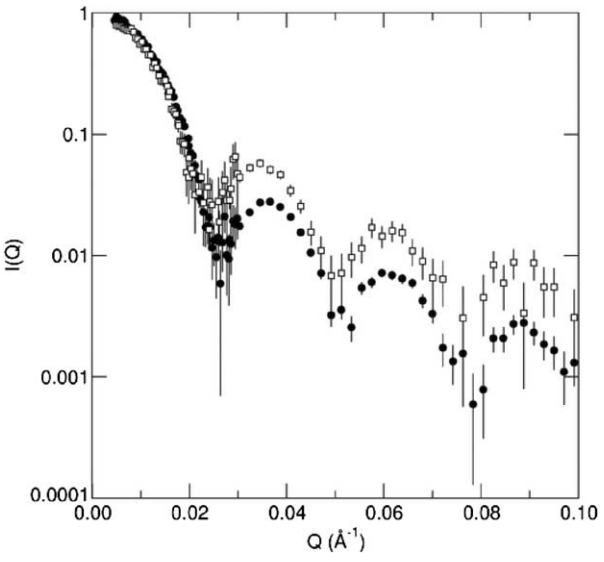

Figure 2. Measured SANS data for the WT MS2 protein shell (open squares) and MS2-capsid (filled circles).

Table 2. Molecular mass determinations

	Concentration (mg/ml)	Mass density (g/cm^3)	M ($\times 10^6$ g/mol)
WT MS2 A-plus protein shell	2.0±0.1	1.38±0.01	2.5±0.3
MS2-capsid A-plus protein	0.55±0.05	1.32±0.02	2.7±0.4
MS2-HCV A-minus protein shell	0.55±0.03	1.36±0.02	2.1±0.2
MS2-λ A-minus protein shell	0.65±0.03	1.37±0.02	–

Errors on the concentration and mass density are standard deviations of the mean. The M value for the MS2-capsid sample was obtained from equation (2), using $\Delta\rho = 3.0 \times 10^{10}\ cm^{-2}$. The M values for WT MS2 and MS2-HCV are M_{PROT} values from equation (5), using independent number densities measured by the Integrated Virus Detection System.[38] Since independent number density measurements are not available for MS2-λ, M_{PROT} cannot be determined in this case.

whether the A protein is present in the sample. That determination was made from SDS-PAGE results, as explained above.

The shape of the structure that MS2 particles form in solution has been described using SANS.[38] The shape of an MS2 particle can be approximated very well by a spherical shell, with inner radius, R_1, outer radius, R_2, and shell thickness, t, where t is equal to $R_2 - R_1$. This shape is generally known as a core-shell model.[23] To compare the shape of the structures formed by the MS2-capsid protein shell with that of the wild-type MS2 in solution, the MS2-capsid SANS data were also fit to the core-shell model. The MS2-capsid SANS data are shown again in Figure 3, along with the best fit to a core-shell model structure. Because these are experimental data, the resolution of the SANS instrument was taken into account in the fitting process. Figure 3 indicates that the shape of the structures formed by the MS2-capsid protein shell also fits the core-shell model. Therefore, wild-type MS2 and RNA-free MS2-capsid protein shell both form structures in solution that are best described by the core-shell model.

For comparison, a similar fit was performed to a modeled SANS intensity calculated from the published crystal structure of wild-type MS2 (2MS2).[3] Recall that crystal structure data can be used to derive model SANS solution data. That is, the MS2 crystal structure can essentially be "hydrated" so that its structure in solution can be inferred. The modeled WT MS2 SANS curve and corresponding best fit to the core-shell model are shown in Figure 4. Instrumental resolution effects are not taken into account in this case, since the modeled SANS data are calculated assuming perfect resolution under idealized conditions and, thus, the peaks in the model data are sharper than in the experimental data. It is clear from Figure 4 that the structures formed by the crystal structure-derived modeled wild-type MS2 data (2MS2) are also well described by the core-shell model.

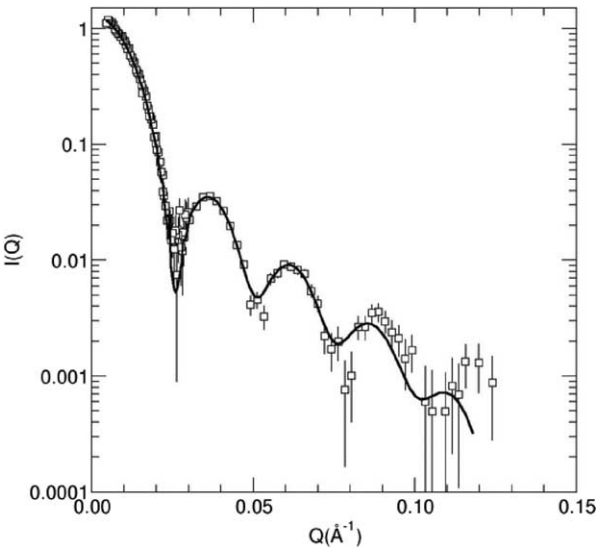

Figure 3. MS2-Capsid in 100% ^2H_2O buffer (open squares). The continuous line is the best fit core-shell model structure, taking into account instrument resolution effects.

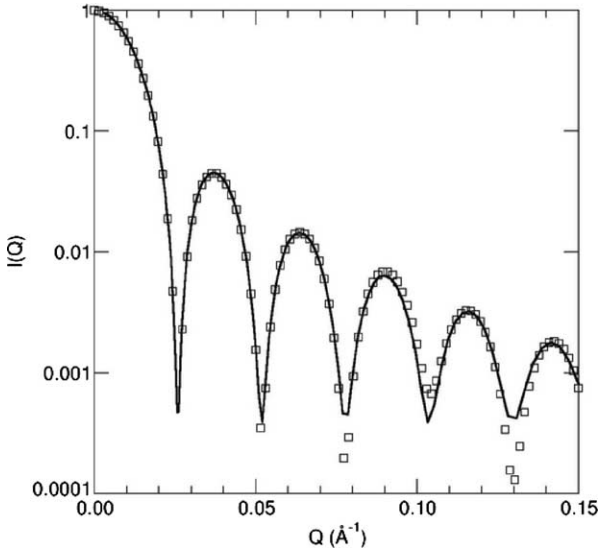

Figure 4. Model SANS data (open squares) for the MS2 protein shell obtained from X-ray crystallography coordinates (2MS2).[3] The continuous line represents the best fit core-shell model structure.

The resulting fit parameters, which describe the dimensions of the core-shell model structures, are shown in Table 3. The experimental MS2-capsid and WT MS2 protein shell results are listed as MS2-capsid A-plus protein shell and WT MS2 A-plus protein shell under Measured data. The crystal

Table 3. Core-shell model fit parameters

	R_1 (Å)	R_2 (Å)	t (Å)
A. *Measured data*			
WT MS2 A-plus protein shell	115±1	136±1	21±1
MS2-capsid A-plus protein shell	115±1	140±2	25±1
MS2-HCV A-minus protein shell	105±2	135±1	31±1
MS2-λ A-minus protein shell	101±1	136±1	37±1
B. *Model data*			
2MS2	111±1	130±1	19±1
1BMS	111±1	130±1	19±1
1MST	111±1	131±1	20±1
1FR5	111±1	131±1	20±1
1FRS	111±1	130±1	19±1
1ZDI	111±1	130±1	19±1
1ZDH	111±1	131±1	20±1
1QBE	117±1	136±1	19±1

The measured data values are determined by SANS empirically in solution. The model data values are determined from model SANS data derived from crystal structures. Both sets of data are fit to the same core-shell model. Instrument resolution effects are taken into account for the measured data values. Data for the MS2-capsid were obtained in 100% ^2H_2O buffer. Errors are standard deviations of the mean for several equally good fits to the data. Data for the WT MS2, MS2-HCV and MS2-λ protein shell were obtained using the contrast variation technique.[38] Errors are standard deviations of the mean for fits to the data at several different contrasts.

structure derived results are listed as 2MS2 under Model data. The parameters obtained from the model crystal structure derived data (2MS2) indicate that the protein shell has a core (or inner) radius (R_1) of 111($\pm$1) Å, an overall (outer) radius (R_2) of 130($\pm$1) Å and a thickness (t) of 19($\pm$1) Å. Note that the virus core was assumed to have 100% water for the fit to the model crystal structure-derived wild-type MS2 (2MS2) data. The experimental WT MS2 protein shell has R_1=115($\pm$1) Å, R_2=136($\pm$1) Å, and t=21($\pm$1) Å in solution. The RNA-free MS2-capsid protein shell has R_1=115($\pm$1) Å, R_2=140($\pm$2) Å and t=25($\pm$1) Å. The 4 Å difference in the shell thickness between the WT MS2 protein shell and MS2-capsid represents a difference in shell thickness of 10–20%. Compared to the 4 Å difference in total radius of these two samples (2.9–3%), the difference in shell thickness is more significant. Since the shell thickness parameter is determined more by the shape of the peaks in Figure 2, rather than their Q-value, this difference in thickness may be an artifact of the polydispersity in the MS2-capsid sample, as described above.

Although, the shape of the crystal structure-derived (model) wild-type MS2 SANS data also fit a core shell model, the predicted dimensions of the protein shell vary somewhat from the wild-type MS2 crystal structure-derived SANS data vary somewhat from that of the wild-type MS2 and MS2-capsid protein shell measured in solution. From the crystal structure-derived wild-type MS2 data, the protein shell of MS2 in solution would be predicted to have an inner radius (R_1) of 111($\pm$1) Å and outer radius (R_2) of 130($\pm$1) Å. These values are somewhat smaller than an R_1 value of 115($\pm$1) Å and an R_2 value between 136 Å and 140 Å measured for the protein shell of WT MS2 and MS2-capsid, respectively, in solution. These data indicate that the model (crystal structure-derived) MS2 structures are reduced in size compared to the experimentally measured WT MS2 and MS2-capsid structures in solution, although this difference is only at the 3% level. However, the thickness (t=19($\pm$1) Å) of the protein shell, based on crystal structure-derived SANS data for MS2, is in good agreement with the thickness of the WT MS2 and MS2-capsid protein shells measured in solution.

The core-shell model fit to the measured MS2-capsid SANS data was also used to estimate the amount of RNA in the capsid core, using equation (6) as described in Materials and Methods. The fitted scattering length density (SLD), which is defined as the sum of the neutron-scattering lengths of all the atoms in the molecule divided by the total volume of the molecule, of the virus core is 6.1×10^{-6} Å^{-2} $\pm$ 0.1×10^{-6} Å^{-2}. When compared to the SLD for 100% ^2H_2O buffer, 6.4×10^{-6} Å^{-2}, this suggests that the MS2-capsid core contains 96($\pm$1) % water. Thus, at the resolution of SANS, the measured MS2-capsid sample is essentially free of RNA, as expected. This result is in agreement with that obtained from SAXS measurements of MS2-capsid.[39]

In summary, the core-shell model describes the shape of the structures formed in solution by the wild-type MS2 virus and its empty MS2 capsid. Using the core-shell model, the dimensions of the wild-type MS2 and empty MS2-capsid protein shells (in solution) are found to be virtually indistinguishable at the resolution of SANS. The overall shape of the crystal structure-derived model MS2 SANS data also can be fit to the core shell model, but with dimensions slightly different from those of the WT MS2 virus and MS2-capsid found in solution. Finally, it was confirmed that the MS2-capsid sample forms a coat protein shell like that of the wild-type MS2, except that it is essentially RNA-free with a virus core composed of 96($\pm$1)% water. The fact that the structure of the MS2-capsid and WT MS2 protein shell do not differ significantly indicates that RNA content has no detectable effect on the dimensions or packing arrangement of the protein shell of the MS2 virus in solution.

The A protein affects the MS2 protein shell arrangement

As the RNA content does not affect the protein shell packing arrangement of MS2, we next analyzed the influence of the A protein on MS2 protein shell structure. Two recombinant forms of MS2 that lack the A protein, MS2-HCV and MS2-λ, were examined using SANS. Figure 5 shows the measured SANS data for wild-type MS2 protein shell, as well as that from the protein shell component of the MS2-HCV recombinant virus. Again, the scattering from the protein shell was separated from that of the MS2-HCV RNA core using the contrast variation technique, as described.[38] Note that, unlike the case of the MS2-capsid, the locations of the peaks in the scattered

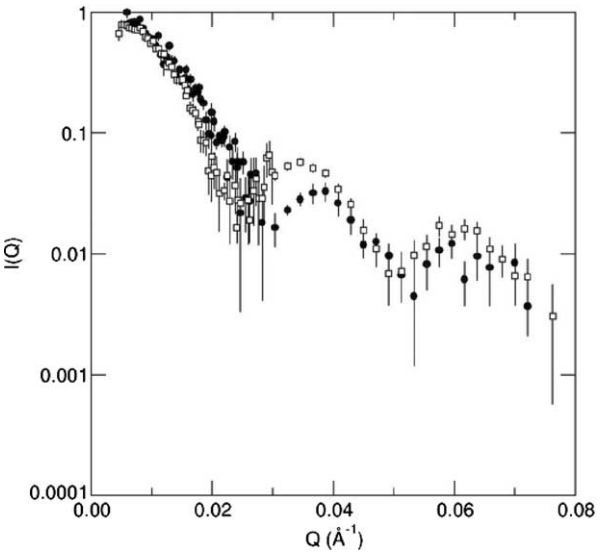

Figure 5. Measured SANS data for the WT MS2 protein shell (open squares) and MS2-HCV protein shell (filled circles).

intensity are not the same for the wild-type MS2 and the MS2-HCV protein shell components. This indicates that the scattering arises from structures that are different. A similar result was found for the protein shell component of the MS2-λ recombinant virus (data not shown).

The parameters obtained from the fits to the core-shell model, which are measures of the dimensions of the MS2-HCV and MS2-λ protein components, are given in Table 3 as MS2-HCV A-minus protein shell" and "MS2-λ A-minus protein shell", respectively, under Measured data. The results show a core radius of $R_1 = 105(\pm 2)$ Å and a protein shell thickness of $t = 31(\pm 1)$ Å for the MS2-HCV sample. For MS2-λ, the results are $R_1 = 101(\pm 1)$ Å and $t = 37(\pm 1)$ Å. Table 2 shows the calculated M for the protein shell component of MS2-HCV, determined from equation (5) to be $2.1(\pm 0.2) \times 10^6$ g/mol. This result is in good agreement with the M values of both the WT MS2 protein shell component and the MS2-capsid sample. Again, the uncertainty in the M determination for both the MS2-HCV protein shell is too large to ascertain whether the A protein is present in the sample. That determination was made from the SDS-PAGE results.

When compared to the values obtained for the MS2-capsid and WT MS2 protein shell, it is clear that the protein shell is thicker for the MS2-HCV and MS2-λ viruses, while the overall radius, R_2, of all the viruses are the same in solution. Another way to illustrate this difference in protein shell thickness (Figure 6) is by plotting the distance distribution functions, $P(r)$, for the wild-type MS2 protein shell and the MS2-HCV and MS2-λ protein shells. $P(r)$ gives a real space representation of the SANS data as a function of radial distance, r, in the molecule. The $P(r)$ function is essentially a histogram of the distances between all possible pairs of points in the molecule, and it indicates the probability that a certain distance between two pairs of points in the molecule will occur. For instance, the $P(r)$ function for a solid sphere would be symmetric with a peak at the distance, r, that represents its radius. On the other hand, a hollow sphere with a thin shell has no mass in the core. All of the mass lies in the shell and there is a very low probability that there are distances represented by two points in the interior of the particle. Therefore, $P(r)$ would be asymmetric with a peak at $r = R_1 + t/2$, Thus, the $P(r)$ function provides a more intuitive representation of the shape of the scattering particles than does the scattered intensity, $I(Q)$.[22]

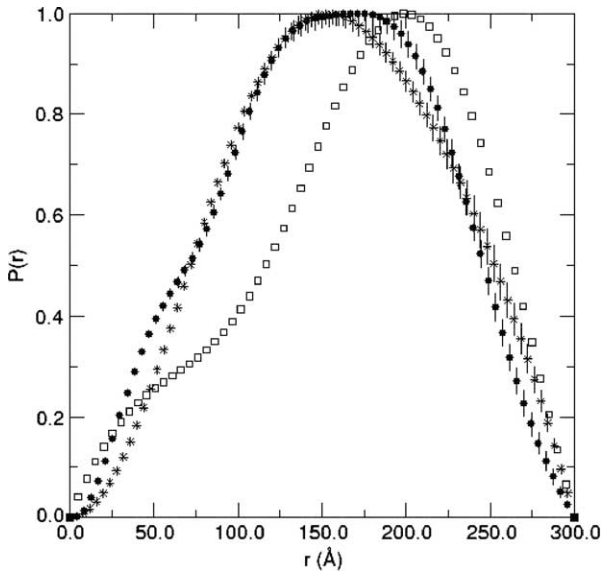

Figure 6. Distance distribution functions, $P(r)$, for the WT MS2 protein shell (open squares), MS2-HCV protein shell (filled circles) and MS2-λ protein shell (stars).

The most probable pair correlation distance for the WT MS2 protein shell occurs at a value of approximately 200 Å (Figure 6). The most probable distances for the protein shells of the recombinant viruses occur at much lower values, at approximately 150 Å and 170 Å for the MS2-λ and MS2-HCV cases, respectively. The distribution of the protein shell is also broader for the recombinant viruses. While there are also some differences in $P(r)$ for r values below 50 Å, these correspond to the higher Q region of the data ($Q \geq 0.12$ Å^{-1}), where subtle differences in background subtraction might change $P(r)$. Thus, the observed differences in $P(r)$ between MS2-HCV and MS2-λ in this region are likely not real.

These results indicate that, in general, the MS2-HCV and MS2-λ particles form a shape in solution that can be described by the core-shell model. The core-shell model structure formed is not exactly like that of the (A protein-containing) WT MS2 and MS2-capsid particles. MS2-HCV and MS2-λ form particles of approximately the same size or outer radius (R_2 between $135(\pm 1)$ Å and $136(\pm 1)$ Å) as compared to the size of the WT MS and MS2-capsid particles, which are $136(\pm 1)$ Å and $140(\pm 2)$ Å, respectively. However, the inner radius (R_1) of the MS2-HCV and MS2-λ is significantly smaller than the WT MS2 or MS2-capsid particles. Since the thickness (t) of the protein shell is defined as the size or outer radius (R_2) minus the inner radius (R_1), the maintanence of overall size with a decreased inner radius results in particles with a protein shell of greater thickness. In this case, the thickness of the protein shell of MS2-HCV and MS2-λ is $31(\pm 1)$ Å and $37(\pm 1)$ Å, respectively. This is considerably thicker than the protein shell formed by either the WT MS2 or the MS2-capsid, whose protein shells have thickness of $21(\pm 1)$ Å and $25(\pm 1)$ Å, respectively. The recombinant viruses (MS2-HCV and MS2-λ) differ from the wild-type MS2 and MS2-capsid samples in one significant respect: the recombinant viruses lack the A protein, while the WT MS2 and MS2-capsid viruses contain the A protein. Therefore, these results are consistent with the idea that the A protein plays a role in the structural arrangement of the MS2 protein shell.

Crystal structure data do not reveal the A protein-dependent protein shell rearrangement seen by SANS

To determine if the A protein-dependent protein shell rearrangements observed by SANS in solution can be seen using other methods, we examined the known X-ray crystal structures of a number of different MS2-related viruses and mutants, and then converted the crystal structure data into model (or simulated) SANS data as described (see Introduction).[38] The results from the core-shell model fitting of the model data for these viruses are presented in Table 3, under Model data. The crystal structure-derived model SANS data for wild-type MS2, GA and Qβ viruses and the capsids of fr virus are labeled according to Protein Data Bank identification (PDB ID) designations as 2MS2, 1GAV, 1QBE, 1FRS and 1FR5, respectively.[3,32,40,41] GA, Qβ, and fR, like bacteriophage MS2, are members of the Leviviridae family of viruses.[1]

1BMS and 1MST are versions of MS2 that contain point mutations in the coat protein subunits in the region thought to be involved in quasi-equivalent interactions necessary for virus assembly.[42,43] In particular, 1BMS has been shown to contain a mutation in a conserved proline residue that is thought to interact with the A protein.[42] 1ZDI and 1ZDH are versions of MS2 that lack the A protein,[31] and are most like the recombinant MS2 particles (MS2-HCV and MS2-λ) when RNA content is ignored. Both 1FRS and 1FR5 are A protein-deficient, RNA-free capsid versions of the FR virus.[32,44] However, 1FR5 contains a four amino acid residue deletion in the coat protein in the FG β-loop region thought to generate the quasi-equivalent interactions necessary for virus assembly, while the 1FRS coat protein sequence is wild-type.[32,44] The crystal structure of the wild-type FR virus has not been reported.

Notice that all of the crystal structure-derived model SANS results are essentially in agreement with each other, except for 1QBE, which is a larger virus. However, note that the protein shell thickness (t) of 1QBE matches that of the other crystal structures as well. Table 1 gives the genotype information for the viruses that are compared in Table 3. It can be seen from Table 1 that, with the exception of the wild-type viruses (2MS2, 1GAV and 1QBE), all of the other structures were formed in the absence of the A protein. Yet, Table 3 shows no significant difference in the radii (R_1 and R_2) or thicknesses (t) for all of the viruses, with the exception of 1QBE discussed above. Thus, it appears that, at least as far as the crystal structures are concerned, the A protein does not affect the thickness of the coat protein shell. In fact, all of the protein shell thicknesses are between 19 Å and 20 Å, even if the virus itself is larger than MS2 overall, as is the case with 1QBE, or if the virus contains point mutations (1BMS and 1MST) compared to wild-type versions of the viruses (2MS2). In particular, the empty MS2 capsids (1ZDI and 1ZDH) and empty FR capsids (1FRS and 1FR5) that were formed in the absence of A protein also produce the same result.

It is important to note that both the core (or internal) radius (R_1) and overall size (R_2) differ significantly between the wild-type MS2 particles samples (WT MS2) measured in solution, when compared to the modeled SANS data derived from the crystal structure, 2MS2. For instance the inner radius, R_1 is 115($\pm$1) Å for the wild-type MS2 sample in solution, whereas it is 111($\pm$1) Å for the 2MS2 crystal structure. Similarly, the overall radius (or size), R_2, of the MS2 virus is 6 Å larger for the WT MS2 sample in solution than that predicted from the 2MS2 crystal structure. The crystallization process itself appears to affect the virus structure, having the net effect of compacting the virus, which reduces its physical dimensions. Whether this is due to crystal contacts or different solvent conditions between the SANS solution studies and those used for crystallization is not clear.

The fits to the core-shell model parameters in Table 3 show that there is a difference between the solution structures of the viruses that contain the A protein and those that do not. In particular, the viruses that do not contain the A protein (MS2-HCV and MS2-λ) have thicker protein shells, t, while still maintaining approximately the same overall radius, R_2, as those viruses containing the A protein. These differences are clear also from Figures 5 and 6

To illustrate this difference further, Figure 7(a) shows cross-sectional views of the 2MS2 structure obtained from X-ray crystallography (red lines), along with the low-resolution core-shell model spherical structure for the protein shell obtained from SANS for the WT MS2 sample (blue circles). It is clear from Figure 7(a) that the shape of the protein shell of WT MS2 in solution is approximated well by the core-shell model. However, the outer radius of the WT MS2 core-shell (SANS) model structure is somewhat larger than that of the 2MS2 crystal structure, except for a few loops of the 2MS2 structure that protrude beyond the boundaries of the WT MS2 spherical shell. Also, in general, the inner radius of the WT MS2 core-shell model structure is larger than that of 2MS2, with a large number of loops extending inward beyond the boundaries of the inner radius of the WT MS2 spherical shell. These results are in agreement with the smaller R_1 and R_2 values found in Table 3 for the 2MS2 crystal-derived model structure, as compared with the WT MS2 solution structure.

Figure 7(b) shows cross-sectional views of the core-shell model spherical structures for the protein shells obtained from SANS for both the WT MS2 (blue circles) and the MS2-HCV (green circles) samples. The MS2-HCV spherical shell is clearly thicker than the wild-type MS2 protein shell. While the outer shell boundaries of both models are the same, the MS2-HCV spherical shell extends inward beyond the inner boundary of the wild-type MS2 spherical shell. Thus, while the protein shell of MS2-HCV is thicker than that of the WT MS2 protein

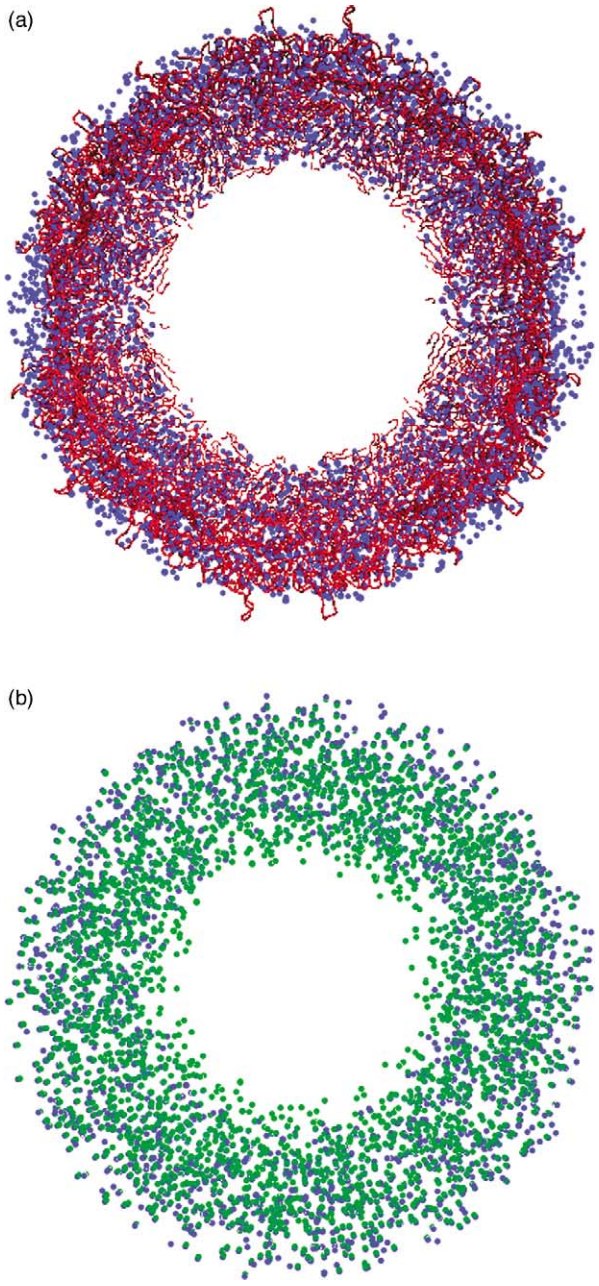

Figure 7. Cross-section view of the core-shell models using the best fit model parameters in Table 3. (a) WT MS2 protein shell (blue) derived from SANS compared to the crystal structure 2MS2 (red). (b) WT MS2 protein shell (blue) compared to MS2-HCV protein shell (green), both derived from SANS.

shell, the overall size of the viruses is approximately the same.

This difference in protein shell arrangement is not seen in the crystal structures. This might explain why the change in thickness or organization of the protein shell in the absence of the A protein observed by SANS in solution is not seen in the crystal-derived structures (1ZDI and 1ZDH), which are analogous to the MS2-HCV and MS2-λ viruses used in this study.

Discussion

An ideal or perfect icosahedron would contain no more than 60 identical subunits to achieve perfect symmetry. The quasi-equivalence theory explains how viruses can produce icosahedral protein shells of various sizes by simply producing more copies of a single protein subunit without the need for additional genetic material.[45] According to this model, only certain arrangements of subunits will lead to closed low-energy icosahedral structures. These arrangements are defined by the triangulation number T, where $T=1,3,4,7$, etc. produce virus protein shells made of 60, 180, 240, etc. subunits.[45]

MS2 virus, like a number of small RNA viruses, forms an icosahedral protein shell from 180 identical subunits plus a single A protein. X-ray crystallographic analysis of the MS2 virus suggests that FG β-loops of the coat protein subunits are capable of the quasi-equivalent interactions necessary for the formation of the MS2 protein shell.[3] The role of the A protein in the structure of the MS2 protein shell is not entirely clear. The A protein is, for practical purposes, an extra structural component for which there are few theories to explain either its structural positioning or biological function in the structural assembly of the virus coat.[1,45]

In general, what is known about the role of the MS2 A protein has been determined primarily by genetic analysis and labeling experiments.[46] In the intact MS2 virus, the A protein has been shown to be required for host recognition of the F-+ pilus.[13,46] Subsequent interaction between the A protein and its *Escherichia coli* host leads to cleavage of the A protein into two fragments during the absorption step of infection. These A protein fragments (which are tethered to the virus genomic RNA) dissociate from the virus protein shell and both the A protein fragments and RNA are injected into the host cell.[46] After infection, the MS2 protein shell is left intact in solution but lacks RNA and A protein.[13,47] Sedimentation gradient analysis of versions of MS2 containing mutations in the A protein indicate that the A protein is also necessary for RNA packing (or compaction within the virion).[16] MS2 viruses with mutations in the A protein, in contrast to the wild-type MS2 virus, contain RNA that dangles outside of the protein shell and is sensitive to the action of nucleases.[16,17] Therefore, with respect to the life-cycle of the MS2 virus coat protein shell, in its pre-infection state, the protein shell contains the A protein, but post-infection, the A protein is absent.

This study is the first to examine the contribution of the A protein and RNA content on the global packing arrangement of the MS2 virus coat in solution. Specifically, this work examines the structure of the protein shell of MS2 virus particles that assembled spontaneously in the presence of A protein (WT-MS2 and RNA-free MS2-capsid) and in the absence of A protein (MS2-HCV and MS2-λ) in solution by SANS. First, we determined that RNA

content does not effect the structure of the MS2 protein shell, as the solution structure of the A protein containing RNA-free MS2-capsid particles (MS2-capsid) particles are virtually indistinguishable from the WT MS2 protein shell. Secondly, we analyzed the structure of MS2 virus particles assembled in the presence of A protein (like the WT MS2 pre-infection state) and the absence of A protein (like the MS2 post-infection state). We found a dramatic A protein-dependent difference in the conformation of the MS2 virus protein shell arrangement. Specifically, A protein-containing MS2 protein shells have a thickness in solution of between 21 Å and 25 Å, while versions of MS2 that lack the A protein have considerably thicker shells, with thickness of between 31 Å and 37 Å. Although, the increase in thickness of A protein-deficient protein shells was not associated with a change in the overall size or diameter (R_2) of the virus, the internal radius (R_1) of the virus was decreased significantly.

These results, taken together, are the first to show that the A protein contributes to the global organization of the native structure of the MS2 protein shell. Our results are consistent with the idea that this contribution is made first during assembly. Since there is a large body of evidence that MS2 virus assembly can occur in the absence of the A protein, there are clearly many low-energy packing arrangements that generate closed MS2 icosahedral viruses that are acceptable.[29,36,37,44] However, the A protein may help the protein subunits to selectively build the most biologically relevant packing arrangement during the virus assembly process. Despite the fact that the complete genomic sequence of the A protein is known,[10] and the gene has been cloned successfully, a number of factors have hampered efforts to solve the crystal structure of the A protein. First, the level of A protein is regulated tightly at the level of translation, so that very low levels of the protein are made from its native transcript.[48] Secondly, although several purification protocols for the isolation of the A protein have been reported, it is still very difficult to isolate large quantities of the protein, due to the fact that the protein is very insoluble and adheres to a wide variety of surfaces.[49,50] The fact that the A protein has many glue-like properties might provide insight into its function.

During the assembly of the protein shell, the A protein might act as a sticky scaffold or nucleator that helps to confine the coat protein subunits to choose a particular conformation that holds the shell under tension. This tension may be stored in the protein shell for use during the ejection step of infection. At some time during or after A protein cleavage, the virus protein shell may shift from the thin (pre-infection) conformation to the thicker (A minus post-infection) conformation. The result of this conformational shift would be a decrease in the internal radius of the virus protein shell perhaps due to the repositioning of coat protein loops that would help to force the RNA and attached A protein out into the host. The data described here are consistent with this model and the idea that viruses are assembled under strain is not new.[51]

A detailed examination of the structure of the MS2 protein shell using single-particle cryo-electron microscopy (cryo-EM) and three-dimensional icosahedral image reconstruction has been reported.[18] In these studies, the structures of the wild-type MS2, RNA-free capsid and RNA were analyzed and compared to a number of different MS2 crystal structures. The RNa-free capsids were generated by heating the purified wild-type MS2 in the presence of 20 mM Hepes (pH 8.4) and 50 mM NaCl in the presence of RNase (a nuclease that degrades RNA). This type of heat treatment has been shown to cause the MS2 virus to expand, and the RNA to unfold and dangle outside the virus coat and become susceptible to nucleases.[52] However, the A protein has specifically been shown to remain attached to the virus under these conditions.[52,53] Just as in this study, the authors of the cryo-EM work reported no obvious difference between the structure of the coat protein shell of the wild-type MS2 and the RNA-free MS2 capsid, except those related to RNA content. Therefore, the RNA-free capsids analyzed in the cryo-EM studies are similar to the RNA-free MS2-capsid samples used in this study. Both the WT MS2 and the A protein-containing, RNA-free MS2-capsid samples should therefore be in the (thin) pre-infection conformation.

The cryo-EM study noted the presence of a thin icosahedral layer of RNA just below the surface of the coat protein shell in the WT MS2 sample, which leaves the 5-fold and quasi 6-fold holes RNA-free.[18] A comparison of the cryo-EM images to the crystal structure of a version of MS2 bound to an RNA fragment containing the coat protein operator site (PDB entry ZDI) confirmed that 10–20% of the genomic RNA is bound to the coat protein dimers in the same site as the coat protein operator RNA, presumably *via* RNA stem–loop motifs.[18] These results suggest that cleavage of the A protein mediates the conformational change in the coat protein shell that is necessary for the release of the icosahedral network of RNA that is attached to the inner surface of the virus protein shell during early stages of RNA ejection. The associated change in internal radius of the virus due to the conformation change of the coat protein loops may help to catapult or squeeze the RNA and attached A protein out of the virus. Although the amount of energy needed for RNA injection for MS2 (or any RNA virus to our knowledge) is not clear, it is known how much energy is required to unfold 75% of the genomic RNA and to cause an expansion of the protein coat. The energy of activation for the protein shell expansion is 85 kcal/mol.[52] Therefore, this value must reflect the order of magnitude of the lower limit on the amount of energy needed for some of the early steps of the RNA ejection process.

The idea that the energy necessary for nucleic acid ejection is stored in the protein shell for use

during viral infection is not commonplace. What is known about nucleic acid ejection is based primarily on studies using complex double-stranded DNA viruses. Historically, inspired by the experiments done by Hershey & Chase using the bacteriophage T2, viruses were thought to inject their DNA into their host in a syringe-like mechanism.[54] Implicit to this idea is that the DNA is packed under pressure or tension, which is used for the rapid expulsion of the nucleic acid. There seems to be substantial evidence to support the idea that DNA is packed tightly and, in some cases under pressure, within the virus protein shell.[55–57] This is perhaps only one of many mechanisms that can effect nucleic acid ejection. A growing body of evidence suggests that the T7 virus DNA ejection step is mediated *via* a multi-step process that includes the use of an enzyme-like mechanism as well as transcription to drive DNA translocation into its host.[58–61] However, it is not clear if simple RNA viruses, like MS2, employ any or all of these methods of nucleic acid translocation. Nevertheless, the results reported here suggest that perhaps energy stored during the assembly of the virus protein shell might provide some of the mode of force necessary for simple RNA viruses like MS2 to mediate some of the early stages of nucleic acid transfer. Furthermore, this work rules out the possibility that the MS2 genomic RNA determines the overall virus shell arrangement. Rather, our data are the first to suggest that the A protein acts to effect global structural rearrangement of the MS2 protein shell during infection.

This SANS solution study reveals two general types of phenomenon related to the structure of the MS2 virus coat. The first, which was discussed above, is related to the role of the A protein on the organization of the virus coat and its implications for virus assembly and RNA ejection. The second is the differences observed between the crystal and solution structures of MS2, both in the absence and in the presence of the A protein, which is discussed below.

Two observations were made with respect to the relationship between the crystal and solution structure of MS2 virus: (a) the dimensions of the WT MS2 virus particle in crystalline form is smaller than that measured in solution; and (b) the structural differences seen in solution between A protein-deficient MS2 protein shell and WT MS2 are not seen in the crystal structure derived data.

In general, there is remarkably good agreement between the overall shape and dimensions of proteins measured using crystallographic and solution scattering-based methods.[62] What few differences there are in the literature represent no more than a 5–12% deviation between crystalline and solution-based measurements such as in the case of the xylanase enzyme from *Trichoderma longibrachiatum* and *E. coli* asparaginase II.[63,64] This is the scale of difference (~5%) between the dimensions of WT MS2 protein shell measured in solution *versus* the crystal structure (with the crystal structure measurements being smaller than the solution).

However, an approximately 50% increase in the thickness of virus protein shells that lack the A protein compared to WT MS2 in solution is much larger than these typical differences. Furthermore, the observation that a difference of this magnitude is not seen in similar samples when measured by X-ray crystallography is reason for pause, because it provides important information about the physical properties of the virus shell. For instance, in the case of immunoglobulin IgG molecules, the discrepancies between the dimensions of the molecule, as measured in the crystal structure and solution are an exception to the above rule and are of the magnitude seen in this study. The radius of gyration, R_g (a measure of the size) of IgG molecules based on the crystal structure is 46 Å, while in solution R_g values range from 60–74 Å. This represents a ~30–50% difference.[62] In general, this difference is thought to occur because the IgG molecules are very flexible. In solution, IgG molecules are clearly extended. However, in the crystalline state they are severely compacted.[62]

Like IgG molecules, the protein shell of versions of MS2 that lack A protein are clearly thicker in solution, but thinner after being crystallized. Taken together, these results suggest that the native MS2 virus shell is somewhat flexible and, in the absence of the A protein, this flexibility is enhanced. Furthermore, in the absence of the A protein, the transition between the thick (post-infection) and thin (pre-infection) protein shell conformational states can be reversible under crystallization conditions.

A number of mechanisms might explain the differences between solution and crystal structure results described herein. The use of deuterated water (^2H_2O) in the SANS experiments and crystal contacts (packing effects) or solvation effects in the crystallography experiments may alter the structure of the MS2 coat protein shell. Deuterated water has been shown to alter the *in vitro* properties of some proteins by either increasing the conformational stability of some proteins and/or by promoting the assembly of macromolecular complexes. For example, the polymerization of actin, flagellin, and recA have all been shown to be promoted and stabilized by ^2H_2O.[65–67] Also, the analysis of tubulin polymers using SANS, revealed that ^2H_2O altered the morphology of the tubulin polymer significantly at pH values other than 7.0.[68]

To our knowledge, only two viruses have been reported to exhibit structural changes in deuterated water. Assembled poliovirus capsid has been shown to be protected against heat or high pH-induced dissociation by ^2H_2O and $MgCl_2$.[69] Also, the polymerization of the tobacco mosaic virus protein is promoted and stabilized by ^2H_2O.[70] However, our contrast variation analysis of native MS2 virus particles showed no evidence of structural changes related to deuterated solvent, as determined by the consistent inner and outer

radii (R_1 and R_2) values obtained at all contrasts, except 10% ^2H_2O, where the RNA dominated the scattering.[38] In 10% ^2H_2O, the structure of the RNA is uniquely resolved. But at all other contrasts (or deuterated water levels), the structure of the coat protein shell is constant and predominates the scattering. Likewise, the recombinant MS2-HCV and MS2-λ samples were measured under the same contrast conditions as WT MS2 and no structural changes in the protein shell were observed for these samples. Thus, we have ruled out the possibility that the effects that we see in solution might be due to an artifact of the use of deuterated water.

On the other hand, it has been shown that crystal contacts, or crystal packing, can have an effect on protein structure.[71–74] Such structural differences have been observed between the basic pancreatic trypsin inhibitor (BPTI) protein, which forms crystals of a different type and space group, with the most variability being associated with loop structures.[72] Such differences can be further enhanced in solution, in the absence of crystal contacts. Furthermore, by comparison of the X-ray and NMR structures of interleukin 8, it was found that the loop structures in dimers and other higher-order structures showed larger structural differences between crystal and solution than did loops in monomeric forms of the protein.[71] It is plausible that the structural differences we observe between the crystal and solution structures of MS2 are due to changes in the more flexible loop regions of the protein as a result of physical crystal contacts.

To illustrate the potentially flexible regions of the MS2 protein shell, the crystal structures of the WT MS2 virus, 2MS2, and A protein-deficient mutant, 1BMS, are shown in Figure 8. Recall that 1BMS is a version of MS2 that contains a point mutation at a conserved proline residue that is thought to interact with the A protein.[30,42] The structures are color-coded according to *B*-factor, which can be a measure of the flexibility of the structure. In this case, the colors, red, green and blue represent regions of high, medium and low mobility or flexibility, respectively. Note, that the most flexible regions of both MS2 structures are the loops that extend from its surface, followed by the FG β-loop regions that form the vertices at the 5-fold and quasi 6-fold regions. Comparing regions in the circles of similar color in Figure 8(a) and (b), note the extended green regions in the quasi 6-fold vertex in 1BMS (green means more mobility, blue is the least mobility and red is the highest mobility, which is found only in the surface loops). Also, note there is some green color in the 5-fold vertex region of 1BMS that is not there in the 2MS2. These small differences in the crystal structure can be enhanced in solution, where there is no crystal packing effect.

The idea that, in general, the MS2 virus particle might be flexible could explain why the observed internal radius of the virus (R_1) and even the external radius or diameter (R_2) appear to be larger in solution for the A protein-containing WT MS2 and MS2-capsid samples, while the shell thickness remains constant when measured in the solution or crystalline form. Also, since there is flexibility in the FG β-loops that make up the 5-fold and quasi 6-fold vertices, it is possible that the rearrangement of the protein shell structure that we observe in solution for the recombinant MS2 samples (which lack the A protein) is inhibited by crystal contacts in the crystal structures of similar systems. This would explain why mutations and even deletions of the FG β-loop regions that have been shown genetically to be important for assembly and were originally postulated to have a large effect on the structure of the MS2 protein shell display only minor changes in structure when compared to WT MS2 when examined by crystallography.

Our work described here suggests that the WT MS2 protein shell is flexible (and by extension that this inherent flexibility in the virus protein coat might be increased in the absence of the A protein). Recently, the mechanical properties of an RNA-free capsid of the bacteriophage ϕ29 were examined by scanning force microscopy.[75] In these studies, they

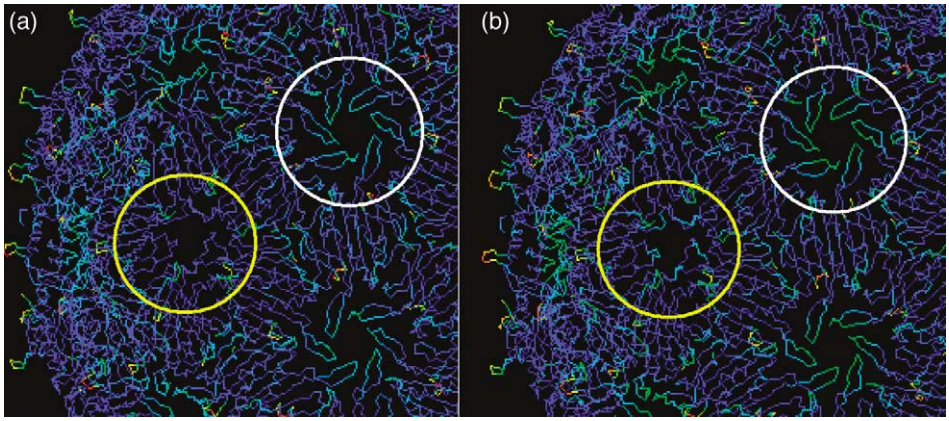

Figure 8. Magnified illustration of the crystal structures (a) 2MS2 and (b) 1BMS, colored by temperature factor from red (highest) to blue (lowest). Yellow circles are drawn around the FG β-loops at one of the 5-fold axes, whereas white circles are drawn around the FG β-loops at one of the quasi 6-fold axes.

found that, although the ϕ29 capsid is extremely tough with a Young's modulus of ~1.8 GPa (which is on the scale of hard plastic), the capsids are nevertheless extremely flexible. The capsids can be indented up to 30% of their height while withstanding nano-Newton forces. Furthermore, they observed that the elasticity of the protein shell varied across the surface, presumably reflecting differences in local protein shell arrangement.[75] Experiments of this type would shed light on the magnitude of difference in flexibility that exists between the WT MS2 and A protein-deficient recombinant MS2 samples.

Finally, our experiments do not formally rule out the possibility that salt, glycerol or ionic strength, and not crystal packing forces, may be responsible for the compaction of the WT MS2 virus and the conversion of the structure of the A protein-deficient coat protein shell from the thick form found in solution to the thin form observed in the final crystal structure. In general, the crystals were grown by vapor diffusion from hanging drops containing 0.6 mg/ml of MS2 capsids. The typical solvent conditions used for crystallization are 1.25% (w/v) polyethylene glycol 8000 in 100–400 mM sodium phosphate (pH 7.4), 0.02% (w/v) sodium azide.[31,76] The SANS buffer conditions were 10 mM Tris–HCl (pH 7.0), 100 mM NaCl, 1 mM $MgCl_2$ as given in Materials and Methods and our concentrations were as given in Table 1. Since SANS is a solution method, it is possible to repeat our experiments using solvation condition like those used for crystallization, so that the influence of solvation can be measured. These experiments are in progress.

Whatever the mechanism, be it crystal packing or solvation effects, it is clear that some force is required to cause the conversion of the A protein-deficient MS2 samples to undergo the conformational shift from the thick post-infection (seen in solution) to the thin (wild-type like pre-infection) protein shell conformation. We believe this observation is further evidence in support of the idea that energy is stored in the MS2 virus coat in the form of tension or strain. Thus, this work supports the idea that, in addition to playing a role in RNA packing and host recognition during infection, the A protein is involved in organizing the packing arrangement of the protein shell during assembly and may mediate crucial conformational changes in the protein shell necessary for the release of the genomic RNA attached to the inner surface of the virus coat during RNA ejection. Furthermore, we propose a model that supports the idea that the presence of the A protein permits the storage of tension or strain energy in the virus coat protein shell that is used to provide the mode of force for RNA ejection.

Materials and Methods

Certain commercial materials, instruments, and equipment are identified here in order to specify the experimental procedure as completely as possible. In no case does such identification imply a recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials, instruments, or equipment identified are necessarily the best available for the purpose.

Bacteriophage, hosts, and medium

MS2 bacteriophage strain 15597-B1 and its *E. coli* host 15597 were purchased from the American Type Culture Center (Manassas, VA). *E. coli* strain 15597 was grown on MS2 broth, which contains, per liter: 10 g of Tryptone, 8 g of NaCl and 1 g of Bacto-yeast. After autoclaving: 10 ml of sterile 10% (w/v) glucose, 2 ml of 1 M $CaCl_2$ and 10 mg/ml of thiamine hydrochloride were added per liter.[77] WT MS2, MS2-HCV, MS2-λ, and MS2-capsid were stored in TSM buffer (10 mM Tris–HCl (pH 7.0), 100 mM NaCl, 1 mM $MgCl_2$) unless stated otherwise. MS2 phage were grown and purified by ultracentrifugation as described.[38] RNA-free MS2 capsids were produced as described,[39] using purified WT MS2 virus as starting material.

In general, to produce MS2-HCV and MS2-λ virus particles, an expression vector was genetically engineered to express only the MS2 coat protein and either the HCV or λ RNA sequences.[33] Expression of the MS2 coat protein is followed by the spontaneous assembly and incorporation of the HCV or λ RNA. Note that these virus particles do not contain the genetic information to encode the A protein. Therefore, there is no pleiotropic effect associated with chemical removal of the A protein because the A protein was never present at any stage of the virus production or assembly process. Specifically, MS2-HCV virus particles were produced from the *E. coli* expression vector pAR-HCV-2b, which contains a 412 nt sequence from the 5′ non-coding core region of HCV subtype2b (Genbank Accession no. M62321).[33] MS2-λ virus particles were produced from the *E. coli* expression vector pAR-l–1.0, which contains 908 nt (1329-421) of l sequence (Genbank Accession no. M17233) (Ambion Diagnostics, personal communication). Both vectors were generous gifts from Ambion Diagnostics, Inc. (Austin, TX) and were grown as described,[78] and purified by cesium gradient ultracentrifugation as described.[38] The measured density of the WT MS2 particles was 1.38(± 0.01) g/cm^3, which is the same as the value reported by Strauss & Sinsheimer.[79] The corresponding measured densities for MS2-capsid, MS2-HCV and MS2-λ from these studies were 1.32(±0.02) g/cm^3, 1.36(±0.02) g/cm^3 and 1.37(±0.02) g/cm^3, respectively.

SDS/PAGE

SDS/PAGE was performed as described.[80] Commercially available pre-cast SDS/18% (w/v) polyacrylamide gels (Tris–glycine gels) for the Novex gel apparatus system were purchased from Invitrogen (Carlsbad, CA) and used according to the manufacturer's instructions. The Tris–glycine SDS/PAGE running buffer and sample buffers were either purchased from Invitrogen (Carlsbad, CA) or made according to the manufacturer's instructions. Samples were diluted by 50% (v/v) in 2× Tris–glycine sample buffer, incubated at 85 °C for 2 min, and then loaded directly onto the gels. Electrophoresis was carried out for 2–3 h at 30–40 mA/gel. The gels were stained in Coomassiwe brilliant blue R250 solution (Sigma, St. Louis, MO) according to the manufacturer's instructions and photographed.

SANS measurements

SANS measurements were performed on the 30 m SANS instruments at the NIST Center for Neutron Research in Gaithersburg, MD.[81] The neutron wavelength, λ, was 6 Å, with a wavelength spread, Δλ/λ, of 0.15. The neutron intensity at the sample was between 9.8×10^5 cts/s and 5.3×10^6 cts/s, depending upon the instrument configuration. Scattered neutrons were detected with a 64 cm×64 cm two-dimensional, position-sensitive ^{3}He detector, with 128×128 pixels at a resolution of 0.5 cm/pixel, and CF_3 as the stopping gas.[82] Raw counts were normalized to a common monitor count and corrected for empty cell counts, ambient room background counts and non-uniform detector response. Data were placed on an absolute scale by normalizing the scattered intensity to the incident beam flux. Finally, the data were radially-averaged to produce scattered intensity, $I(Q)$, *versus* Q curves, where $Q=4\pi\sin(\theta)/\lambda$ and 2θ is the scattering angle. Sample-to-detector distances of 12 m and 2.5 m were used in order to cover the range $0.005\ \text{Å}^{-1}\leq Q\leq 0.17\ \text{Å}^{-1}$. The scattered intensities from the samples were then further corrected for buffer scattering and incoherent scattering from hydrogen in the samples.

SANS data analysis

The Guinier approximation, $I(Q)=I(0)\exp(-Q^2R_g^2/3)$, was used on the low-$Q$ portions of the data to obtain initial values for the radius of gyration, R_g, and the forward scattering intensity, $I(0)$, of the samples. This analysis is valid only in the region where $QR_g\sim1$. The GNOM program,[83] which makes use of all of the data, rather than a limited data set at small Q values, was used to determine the distance distribution function, $P(r)$, the radius of gyration, R_g, and the forward scattering intensity, $I(0)$. This analysis requires the stipulation of a maximum dimension, D_{max}, beyond which $P(r)=0$. Typically, several values of D_{max} are explored in order to find the range over which the $P(r)$ function is stable. Since all of the data are used, this approach typically leads to more accurate determinations of R_g and $I(0)$ that are less influenced by possible aggregation effects.

The MS2-capsid scattered intensity was obtained in 100% ^2H_2O TSM buffer. The molecular mass, M, of the MS2-capsid sample was calculated from the forward scattering, $I(0)$, using the equation:

$$I(0) = n(\Delta\rho V)^2 \tag{1}$$

where $\Delta\rho=(\rho-\rho_s)$ is the contrast, or the difference between the scattering length density of the molecule (ρ) and the solvent (ρ_s), n is the number density of MS2-capsid particles and V is the particle volume. The number density can be written as:

$$n = cN_A/M$$

where c is the concentration, and N_A is Avogadro's number. The volume can be written as:

$$V = M/N_A d$$

where d is the mass density. Now, equation (1) can be rewritten as:

$$\frac{I(0)}{c} = \frac{(\Delta\rho)^2}{N_A d^2}M \tag{2}$$

The only unknown parameter in equation (2) is M, since all other parameters can be measured or calculated. The $I(0)$ value is generally taken from the GNOM[83] analysis of the data. Both c and d can be measured directly during sample preparation, and $\Delta\rho$ can be calculated from the chemical composition of the sample and solvent. It is important to note that $I(0)$ must be on an absolute scale, usually in cm^{-1}, in order to obtain accurate M values from equation (1) or (2).

Since the WT MS2, MS2-HCV and MS2-λ samples contain both a protein and an RNA component, the scattered intensities from the coat protein shells were obtained by decomposition of the scattered intensities from the protein/RNA complexes as described.[38] Specifically, the scattered intensities from the three MS2 protein/RNA complexes were decomposed into the scattering from their components, $I_{PROT}(Q)$ and $I_{RNA}(Q)$ using the equation:

$$\begin{aligned} I(Q) = {} & \Delta\rho_{PROT}^2\, I_{PROT}(Q) \\ & + \Delta\rho_{PROT}\Delta\rho_{RNA}\, I_{PROTRNA}(Q) \\ & + \Delta\rho_{RNA}^2\, I_{RNA}(Q) \end{aligned} \tag{3}$$

The cross-term, $I_{PROTRNA}(Q)$, represents the interference function between the protein and RNA components. The known quantities in equation (1) are $\Delta\rho_{PROT}$ and $\Delta\rho_{RNA}$, and the unknowns are $I_{PROT}(Q)$, $I_{RNA}(Q)$ and $I_{PROTRNA}(Q)$. Since measurements were made at five different contrasts, or $^2H_2O/H_2O$ buffer conditions, there is sufficient information to solve for the three unknown component intensities from the set of simultaneous equations for $I(Q)$ at each contrast. Only the $I_{PROT}(Q)$ results are reported here for WT MS2, MS2-HCV and MS2-λ.

The M values of the protein and RNA components of the MS2 complexes were calculated in a manner similar to equation (1), using the relation:[38]

$$I(0) = n(\Delta\rho_{PROT}\, V_{PROT} + \Delta\rho_{RNA}\, V_{RNA})^2 \tag{4}$$

Equation (4) can be rewritten as:[38]

$$\left[\frac{I(0)}{n}\right]^{1/2} = \left(\frac{|\Delta\rho_{PROT}|}{N_A d_{PROT}}\right)M_{PROT} + \left(\frac{|\Delta\rho_{RNA}|}{N_A d_{RNA}}\right)M_{RNA}, \tag{5}$$

Now, there are only two unknowns, M_{PROT} and M_{RNA} using $V_{PROT}=M_{PROT}/(N_A\, d_{PROT})$ and $V_{RNA}=M_{RNA}/(N_A\, d_{RNA})$. Note that equation (5) requires that the number density of particles at each contrast be known independent of the measured sample concentration. This is because the number density itself is a function of concentration and M, as noted above in equation (2). If such independent number densities cannot be measured, then the total M cannot be separated into the M_{PROT} and M_{RNA} components. However, given an independent measure of the number densities, the $I(0)$ values obtained from the GNOM[83] analysis of the data for each contrast are the used, along with the measured number densities, to solve the set of simultaneous equations in equation (5) to obtain the M values for the protein and RNA components separately in the MS2 complexes. Independent measurements of number density at each contrast were obtained for the WT MS2 samples[38] using the Integrated Virus Detection System, a specialized instrument developed for counting virus particles‡.[38,84] Similar

‡ Wick, C. H. (2002). Method and apparatus for counting submicron sized particles. US patent 6,485,686, USA; Wick, C. H. (2002). Method and system for detecting and recording submicron sized particles. US patent 6,491,872, USA.

independent number density information is available at each contrast for MS2-HCV, but not for MS2-λ. Thus, we report only the M_{PROT} results for WT MS2 and MS2-HCV in this work.

Since MS2 can be approximated very well by a spherical shell at the resolution level of the SANS measurements, the data were also fit to a core-shell sphere model in order to obtain the radii of the protein shell and the solvent core.[23] The neutron-scattering length density of the core was treated as an additional fitting parameter that allowed the amount of water, *versus* RNA, in the core to be calculated using the relation:

$$\rho_{CORE} = X\rho_{RNA} + (1-X)\rho_{SOLVENT} \quad (6)$$

where X is the mass fraction of RNA in the core, ρ_{CORE} is the fitted scattering length density of the core portion of the core-shell model and ρ_{RNA} and $\rho_{SOLVENT}$ are the known scattering length densities of the RNA and the solvent, respectively. Thus, if there is no RNA in the core, $\rho_{CORE} = \rho_{SOLVENT}$. The core-shell model fits take into account the resolution function of the SANS instruments.

For comparison to the various crystal structures of related icosahedral viruses, model SANS data were calculated using the program CRYSON.[24] These model data assume perfect instrumental resolution, so they do not overlap the experimental data exactly. However, they can be fit to the same core-shell model as the experimental data, without the correction for the resolution function of the SANS instruments. Thus, the fitted parameters from the model SANS intensities can then be compared directly to those from the measured SANS data.

Acknowledgements

We acknowledge the support of the National Institute of Standards and Technology (NIST), US Department of Commerce, and the US Army Aberdeen Proving Grounds in providing facilities used in this work. This material is based upon activities supported by the National Science Foundation under agreement no. DMR-9986442. D.A.K. was supported by the NIST-National Research Council Fellowship at the NIST Biotechnology Division. S.K. and D.A.K. thank Mr Ali Deyhim, a NIST/NSF Summer Undergraduate Research Fellow, for help with a portion of the SANS data analysis. We thank Drs Breck Byers, Matthias Frank, Janet Huie and Thomas Silhavy for careful reading of the manuscript.

References

1. Stockley, P. G., Stonehouse, N. J. & Valegard, K. (1994). Molecular mechanism of RNA phage morphogenesis. *Int J. Biochem.* **26**, 1249–1260.
2. Ni, C.-Z., Syed, R., Kodandapani, R., Wickersham, J., Peabody, D. S. & Ely, K. R. (1995). Crystal structure of the MS2 coat protein dimer: implications for RNA binding and virus assembly. *Structure*, **3**, 255–263.
3. Golmohammadi, R., Valegard, K., Fridborg, K. & Liljas, L. (1993). Refined structure of bacteriophage MS2 at 2.8 Å resolution. *J. Mol. Biol.* **234**, 620–639.
4. Valegard, K., Liljas, L., Fridborg, K. & Unge, T. (1991). Structure determination of the bacteriophage MS2. *Acta Crystallog. sect. B*, **47**, 949–960.
5. Valegard, K., Liljas, L., Fridborg, K. & Unge, T. (1990). Three dimensional structure of the bacterial virus MS2. *Nature*, **345**, 36–41.
6. Curtiss, L. K. & Krueger, R. G. (1974). Localization of coliphage MS2 A-protein. *J. Virol.* **14**, 503–508.
7. O'Callaghan, R., Bradley, R. & Paranchych, W. (1973). Controlled alterations in the physical and biochemical properties of R17 bacteriophage induced by guanidine hydrochloride. *Virology*, **54**, 476–494.
8. Fiers, W., Contreras, R., Duerinck, F., Haegman, G., Iserentant, D., Merregeart, J. *et al*. (1976). Complete nucleotide sequence of bacteriophage MS2 RNA: primary and secondary structure of the replicase gene. *Nature*, **260**, 500–507.
9. Rowswell, S., Stonehouse, N. J., Convery, M. A., Adams, C., Ellington, A. D., Hirao, I. *et al*. (1998). Crystal structures of a series of RNA aptamers complexed to the same target. *Nature Struct. Biol.* **5**, 970–975.
10. Fiers, W., Contreras, R., Duerinck, F., Haegman, G., Merregeart, J., Jou, W. M. *et al*. (1975). A-Protein gene of bacteriophage MS2. *Nature*, **256**, 273–278.
11. Min Jou, W., Haegeman, G., Ysebaert, M. & Fiers, W. (1972). Nucleotide sequence of the gene coding for the bacteriophage MS2 coat protein. *Nature*, **237**, 82–88.
12. Carey, J., Cameron, V., de Haseth, P. L. & Uhlenbeck, O. C. (1983). Sequence-specific interaction of R17 coat protein with its ribonucleic acid binding site. *Biochemistry*, **22**, 2601–2610.
13. Kozak, M. & Nathans, D. (1971). Fate of maturation protein during infection by coliphage MS2. *Nature New Biol.* **234**, 209–211.
14. Shiba, T. & Miyake, T. (1975). New type of infectious complexx of *E. coli* RNA phage. *Nature*, **254**, 157–158.
15. Shiba, T. & Suzuki, Y. (1981). Localization of A Protein in the RNA-A complex of RNA Phage MS2. *Biochim. Biophys. Acta*, **654**, 249–255.
16. Argetsinger, J. E. & Gussin, G. N. (1966). Intact ribonucleic acid from defective particles of bacteriophage R17. *J. Mol. Biol.* **21**, 421–434.
17. Heisenberg, M. (1966). Formation of defective bacteriophage particles by Fr. amber mutants. *J. Mol. Biol.* **17**, 136–144.
18. Konig, R., van den Worm, S., Plaiser, J., Van Duin, J., Abrahams, J. P. & Koerten, H. (2003). Visualization by cryo-electron microscopy of genomic RNA that binds to the protein capsid inside bacteriophage MS2. *J. Mol. Biol.* **332**, 415–422.
19. Krueger, S. (1998). SANS provides unique information on the structure and function of biological macromolecules in solution. *Physica B*, **241**, 1131–1137.
20. Jacrot, B. & Zaccai, G. (1981). Determination of molecular weight by neutron scattering. *Biopolymers*, **20**, 2413–2426.
21. Struhrmann, H. B. & Miller, A. (1978). Small-angle neuton scattering of biological structures. *J. Appl. Crystrallog.* **11**, 325–345.
22. Koch, M. H. J., Vachette, P. & Svergun, D. I. (2003). Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution. *Quart. Rev. Biophys.* **36**, 147–227.
23. Hayter, J. B. (1983). Physics of amphiphiles, micelles, vesicles, and microemulsions. In *Proceedings of the International School of Physics "Enrico Fermi", Course XC* (DeGiorgio, V. & Corti, M., eds), pp. 59–93, Elsevier Science Publishing Co., North-Holland.

24. Svergun, D. I., Richards, S., Koch, M. H. J., Kuprin, S. & Zaccai, G. (1998). Protein hydration in solution: experimental observation by X-ray and neutron scattering. *Pro. Natl Acad. Sci. USA*, **95**, 2267–2272.
25. Sokolova, A., Malfois, M., Caldentey, J., Svergun, D. I., Koch, M. H. J., Bamford, D. H. & Tuma, R. (2001). Solution structure of bacteriophage PRD1 vertex complex. *J. Biol. Chem.* **276**, 46187–46195.
26. Tomita, M., Hasegawa, T., Tsukihara, T., Miyajima, S., Nagao, M. & Sato, M. (1999). Two concentric protein shell structure with spikes of silkworm *Bombyx mori* cytoplasmic polyhedrosis virus revealed by small-angle neutron scattering using the contrast variation method. *J. Biochem. (Tokyo)*, **125**, 916–922.
27. Torbet, J. (1979). Neutron scattering study of the solution structure of bacteriophages PF1 and Fd. *FEBS Letters*, **108**, 61–65.
28. Peabody, D. S. & Al-Bitar, L. (2001). Isolation of viral coat protein mutants with altered assembly and aggregation properties. *Nucl. Acids Res.* **29**, e113.
29. Stonehouse, N. J. & Stockley, P. G. (1993). Effects of amino acid substitutions on the thermal stability of MS2 capsids lacking genomic RNA. *FEBS Letters*, **334**, 355–389.
30. Stonehouse, N. J., Valegard, K., Golmohammadi, R., van den Worm, S., Walton, C., Stockley, P. G. & Liljas, L. (1996). Crystal structures of MS2 capsid with mutations in the subunits FG loop. *J. Mol. Biol.* **256**, 330–339.
31. Valegard, K., Murrary, J., Stonehouse, N. J., Van den Worm, S., Stockely, P. G. & Liljas, L. (1997). Three-dimensional structures of two complexes between recombinant MS2 capsids and RNA operator fragments reveal sequence-specific protein–RNA interactions. *J. Mol. Biol.* **270**, 724–738.
32. Liljas, L., Fridborg, K., Valegard, K., Bundle, M. & Pumpens, P. (1994). Crystal Structure of bacteriophage fr capsids at 3.5 Å resolution. *J. Mol. Biol.* **244**, 279–290.
33. Walkerpeach, C. R., Winkler, M., DuBois, D. B. & Pasloske, B. L. (1999). Ribonuclease-resistant RNA controls (armored RNA) for reverse transcription-PCR, branched DNA, and genotyping assays for hepatitis C virus. *Clinical Virology*, **45**, 2079–2085.
34. Pasloske, B. L., Walkerpeach, C. R., Obermoeller, R. D., Winkler, M., DuBois, D. B. & Armored, R. N. A. (1998). technology for production of ribonuclease-resistant viral RNA controls and standards. *J. Clin. Microbiol.* **36**, 3590–3594.
35. Pickett, G. G. & Peabody, D. S. (1993). Encapsidation of heteologous RNAs by bacteriophage MS2 coat protein. *Nucl. Acids Res.* **21**, 4621–4626.
36. Mastico, R. A., Talbot, S. J. & Stockley, P. G. (1993). Multiple presentation of foreign peptides on the surface of an RNA-free spherical bacteriophage capsid. *J. Gen. Virol.* **74**, 541–548.
37. Thomas, G. J. & Prescott, B. (1976). Studies of virus structure by laser-Raman spectroscopy. *J. Mol. Biol.* **102**, 103–124.
38. Kuzmanovic, D. A., Elashivili, I., Wick, C. H., O'Connell, C. & Krueger, S. (2003). Bacteriophage MS2: molecular weight and spatial distribution of the protein and RNA Components by small-angle neutron scattering and virus counting. *Structure*, **11**, 1339–1348.
39. Kuzmanovic, D. A., Elashivili, I., Wick, C. H., O'Connell, C. & Krueger, S. (2005). Quantification of RNA in bacteriophage MS2-like viruses in solution by small angle X-ray scattering. *Radiat. Phys. Chem. J.* In the press.
40. Tars, K., Bundle, M., Fridborg, K. & Liljas, L. (1997). Crystal structure of bacteriophage GA and a comparison of bacteriophages belonging to the major groups of Escherichia Leviviruses. *J. Mol. Biol.* **271**, 759–773.
41. Golmohammadi, R., Fridborg, K., Bundle, M., Valegard, K. & Liljas, L. (1996). Crystal structure of bacteriophage QB at 3.5 Å resolution. *Structure*, **4**, 543–554.
42. Hill, H. R., Stonehouse, N. J., Fonseca, S. & Stockely, P. G. (1997). Analysis of phage MS2 coat protein mutants expressed from a reconstituted phagemid reveals that proline 78 is essential for viral infectivity. *J. Mol. Biol.* **266**, 1–7.
43. Stonehouse, N. J., Scott, D. J., Fonseca, S., Murrary, J., Adams, C. & Clarke, A. R. (1996). Molecular Interactions in the RNA bacteriophage MS2. *Biochem. Soc. Trans.* **24**, 412S.
44. Axblom, C., Tars, K., Fridborg, K., Orna, L., Bundle, M. & Liljas, L. (1998). Structure of phage fr capsids with a deletion in the FG loop: implications for viral assembly. *Virology*, **249**, 80–88.
45. Caspar, D. L. & Klug, A. (1962). Physical Properties in the construction of regular viruses. *Cold Spring Harbor Symp. Quant. Biol.* **27**, 1–24.
46. Krahn, P. M., O'Callaghan, R. J. & Paranchych, W. (1972). Stages in phage R17 infection. VI. Injection of a protein and RNA into the host cell. *Virology*, **47**, 628–637.
47. Silverman, P. M. & Valentine, R. C. (1969). The RNA injection step of bacteriophage f2 infection. *J. Gen. Virol.* **4**, 111–124.
48. Van Duin, J. (1988). Single-stranded RNA bacteriophages. In *The Bacteriophages* (Calendar, R., ed.), vol. 1, pp. 117–167, Plenum Press, New York.
49. Steitz, J. A. (1968). Isolation of the A protein from bacteriphage R17. *J. Mol. Biol.* **33**, 937–945.
50. Osborn, M., Weiner, A. M. & Weber, K. (1970). Large scale purification of A- protein from bacterior17. *Eur. J. Biochem.* **17**, 63–67.
51. Moody, M. F. (1999). Geometry of phage head construction. *J. Mol. Biol.* **293**, 401–433.
52. Verbraeken, E. & Fiers, W. (1972). Studies on the bacteriophage MS2 XX. Expansion of the virion in low salt. *Virology*, **50**, 690–700.
53. Steitz, J. A. (1968). A slowly sedimenting form, infective form of bacteriophage R17. *J. Mol. Biol.* **33**, 947–951.
54. Hershey, A. D. & Chase, M. (1952). Independent functions of viral protein and nucleic acid in growth of bacteriophage. *J. Gen. Physiol.* **36**, 39–56.
55. Evilevitch, A. L., Lavelle, L., Knobler, C. M., Raspaud, E. & Gelbart, W. M. (2003). Osmotic pressure inhibition of DNA from phage. *Proc. Natl Acad. Sci. USA*, **100**, 9292–9295.
56. Smith, D. E., Tans, S. J., Smith, S. B., Grimes, S., Anderson, D. L. & Bustamante, C. (2001). The bacteriophage phi29 portal motor can package DNA against a large internal force. *Nature*, **413**, 1122–1126.
57. Earnshaw, W. C. & Harrison, S. C. (1977). DNA arrangement in isometric phage heads. *Nature*, **268**, 598–602.
58. Zavriev, S. K. & Shemyakin, M. F. (1982). RNA polymerase-dependant mechanism for the stepwise T7 phage DNA transport from the virion into *E. coli*2. *Nucl. Acids Res.* **10**, 1635–1652.

59. Moffatt, B. A. & Studier, F. W. (1988). Entry of bacteriophage T7 DNA into the cell and escape from host restriction. *J. Bacteriol.* **170**, 2095–2105.
60. Garcia, L. R. & Molineux, I. J. (1995). Rate of translocation of bacteriophage T7 DNA across the membranes of *Escherichia coli*. *J. Bacteriol.* **177**, 4066–4076.
61. Kemp, P., Gupta, M. & Molineux, I. J. (2004). Bacteriophage T7 DNA ejection into cells is initiated by an enzyme-like mechanism. *Mol. Microbiol.* **53**, 1251–1265.
62. Kratky, O. & Pilz, I. (1978). A comparison of X-ray small-angle scattering results to crystal structure analysis and other physical techniques in the field of biological macromolecules. *Quart. Rev. Biophys.* **11**, 39–70.
63. Kozak, M. (2004). Direct comparison of the crystal and solution structure of xylanase from *Trichoderma Longibrachiatum*. *Protein Pept. Letters*, **11**, 301–306.
64. Kozak, M. & Jurga, S. (2002). A comparison between the crystal and solution structures of *Escherichia coli* asparaginase II. *Acta Biochim. Polonica*, **49**, 509–513.
65. Omori, H., Kuroda, M., Naora, H., Takeda, H., Nio, Y., Otani, H. & Tamura, K. (1997). Deuterium oxide (heavy water) accelerates actin assembly *in vitro* and changes microfilament distribution in cultured cells. *Eur. J. Cell Biol.* **74**, 273–280.
66. Uratani, Y. (1974). Polymerization of *Salmonella* flagellin in water and deuterium oxide media. *J. Biochem. (Tokyo)*, **75**, 1143–1151.
67. Ruigrok, R. & DiCapua, E. (1991). On the polymerization state of recA in the absence of DNA. *Biochimie*, **73**, 191–198.
68. Sackett, D. L., Chernomordik, V., Krueger, S. & Nossal, R. (2003). Use of small-angle neutron scattering to study tubulin polymers. *Biomacromolecules*, **4**, 461–467.
69. Chen, C.-H., Wu, R., Roth, L. G., Guillot, S. & Crainic, R. (1997). Elucidating Mechanisms of thermostabilization of Poliovirus by D_2O and $MgCl_2$. *Arch. Biochem. Biophys.* **342**, 108–116.
70. Khalil, M. T. & Lauffer, M. A. (1967). Polymerization–depolymerization of Tobacco mosaic virus protein. X. Effect of D_2O. *Biochemistry*, **6**, 2474–2480.
71. Baldwin, E. T., Weber, I. T., St. Charles, R., Xuan, J.-C., Appella, E., Yamada, M. *et al*. (1991). Crystal structure of interleukin 8:Symbiosis of NMR and crystallography. *Proc. Natl Acad. Sci. USA*, **88**, 502–506.
72. Kossiakoff, A. A., Randal, M., Guernot, J. & Eigenbrot, C. (1992). Variability of conformations at crystal contacts in BPTI represent true low-energy structures: correspondence among lattice packing and molecular dynamic studies. *Proteins: Struct. Funct. Genet.* **14**, 65–74.
73. Jacobson, M. P., Friesner, R. A., Xiang, Z. & Honig, B. (2002). On the role of the crystal environment in determining protein side chain conformation. *J. Mol. Biol.* **320**, 597–608.
74. Rapp, C. S. & Pollack, R. M. (2005). Crystal packing effects on protein loops. *Proteins: Struct. Funct. Bioinfomatics*, **60**, 103–109.
75. Ivanovska, I. L., de Pablo, P. J., Ibarra, B., Sgalari, G., MacKintosh, F. C., Carrascosa, J. L. *et al*. (2004). Bacteriophage capsids: tough nanoshells with complex elastic properties. *Proc. Natl Acad. Sc. USA*, **101**, 7600–7605.
76. Valegard, K., Unge, T., Montelius, I., Strandberg, B. & Fiers, W. (1986). Purification, crystallization and preliminary X-ray data of the bacteriophage MS2. *J. Mol. Biol.* **190**, 587–591.
77. Davis, J. E. & Sinsheimer, R. L. (1963). The replication of bacteriophage MS2. *J. Mol. Biol.* **6**, 203–207.
78. DuBois, D. B., Winkler, M. M. & Pasloske, B. L. (1997). U.S. patent 5,677,124.
79. Strauss, J. H. & Sinsheimer, R. L. (1963). Purification and properties of bacteriophage MS2 and of its ribonucleic acid. *J. Mol. Biol.* **7**, 43–54.
80. Laemmli, U. K. (1970). Cleavage of structural protein during the assembly of the head of bacteriphage T4. *Nature*, **227**, 680–685.
81. Glinka, C. J., Barker, J. B., Hammouda, B., Krueger, S., Moyer, J. J. & Orts, W. J. (1998). The 30 m small-angle neutron scattering instruments at the National Institute of Standards and Technology. *J. Appl. Crystallog.* **31**, 430–445.
82. Kopp, M. K., Valentine, K. H., Christophorou, L. G. & Carter, J. G. (1982). New gas mixtures improves performance of ^{3}He neutron counters. *Nuclear Instrum. Methods*, **201**, 395–401.
83. Semenyuk, A. V., Svergun, D. & I, . (1991). GNOM—a program package for small-angle scattering data processing. *J. Appl. Crystallog.* **24**, 537–545.
84. Wick, C. H. & McCubbin, P. E. (1999). Characterization of purified MS2 bacteriophage by the physical counting methodology used in the integrated virus detection system (IVDS). *Toxicol. Methods*, **9**, 245–252.

Edited by M. Moody

(Received 18 August 2005; received in revised form 8 November 2005; accepted 12 November 2005)
Available online 1 December 2005