Asymmetric cryo-EM reconstruction of phage MS2 reveals genome structure in situ

Roman I. Koning1,2, Josue Gomez-Blanco3, Inara Akopjana4, Javier Vargas3, Andris Kazaks4, Kaspars Tars4, José María Carazo3 & Abraham J. Koster1,2

In single-stranded ribonucleic acid (RNA) viruses, virus capsid assembly and genome packaging are intertwined processes. Using cryo-electron microscopy and single particle analysis we determined the asymmetric virion structure of bacteriophage MS2, which includes 178 copies of the coat protein, a single copy of the A-protein and the RNA genome. This reveals that in situ, the viral RNA genome can adopt a defined conformation. The RNA forms a branched network of stem-loops that almost all allocate near the capsid inner surface, while predominantly binding to coat protein dimers that are located in one-half of the capsid. This suggests that genomic RNA is highly involved in genome packaging and virion assembly.

1 Department of Molecular Cell Biology, Leiden University Medical Center, P.O. Box 9600, 2300 RC Leiden, The Netherlands. 2 Netherlands Centre for Electron Nanoscopy, Institute of Biology Leiden, Leiden University, Einsteinweg 55, 2333 CC Leiden, The Netherlands. 3 Biocomputing Unit, Centro Nacional de Biotecnología, Consejo Superior de Investigaciones Científicas (CNB-CSIC), Darwin 3, Campus Universidad Autónoma, Cantoblanco, Madrid 28049, Spain. 4 Biomedical Research and Study Centre, Ratsupites 1, LV-1067 Riga, Latvia. Correspondence and requests for materials should be addressed to R.I.K. (email: r.i.koning@lumc.nl).
bacteriophage MS2 (ref. 1) is a species of the *Leviviridae* family of small, positive-sense, single-stranded ribonucleic acid (RNA) bacteriophages that infect their host via adsorption to bacterial pili. The MS2 virion consists of an RNA genome, encapsidated by a $T = 3$ shell, containing coat protein (CP) and a single copy of the maturation or so-called A-protein (AP), which attaches to the viral RNA and binds the host receptor. The MS2 genome is one of the smallest known, comprising just 3,569 nucleotides, and was the first genome—an RNA–protein reconstruction of the MS2 virion (Fig. 1a) at 8.7 Å (EMD-3403) is similar to the previously reported capsid protein X-ray structure (protein data bank (pdb) entry: 2MS2) (refs. 15 and 16). X-ray crystallography, revealing sequence-specific RNA–protein reconstructions of the MS2 virion suggested that the genome of the MS2 virion (Fig. 1a–d, Supplementary Movie 1) and replaces one CC conformer-type CP2 in the capsid, which thus contains 178 copies of the CP. The AP forms a 9 nm long ‘handle’ (∼1.8 nm in diameter) extending outwards with a ∼30° angle from the surface of a two-fold symmetry axis position in the capsid, extending over a neighbouring hole at a three-fold axis.

The RNA. The inside of our asymmetric MS2 map reveals that the genome within the MS2 capsid has a defined tertiary structure, forming an intricate three-dimensional (3D) branched network of interconnected SLs (Fig. 1c–e, Supplementary Movie 2). The RNA density accounts for 95% of the calculated mass of the full genome. A less ordered region in the map can explain the missing 5%. In total, 59 SL structures were discriminated in the genome, which were all connected to each other, of which the majority (53, 90%) ended near the capsid (Supplementary Fig. 3), while 6 ends were centrally located. To quantify and assess the PSs, the interactions of the RNA SL with CP2, TR–CP2 X-ray model (PDB entry: 1ZDH) was fitted into the EM density map. In total, 44 SLs (83% of the SLs near the capsid and 75% of all identified SLs) ended at a CP2 RNA binding sites. 2 SLs interacted with the AP (Fig. 1d) and 7 were not located near a CP2 interactions. Close examination of these 44 binding sites showed that at 23 sites the RNA from the X-ray SLs were fitted into the EM density map. In total, 44 SLs (83% of the SLs near the capsid and 75% of all identified SLs) ended at a CP2 RNA binding site, 2 SLs interacted with the AP (Fig. 1d) and 7 were not located near a CP2 interactions. Close examination of these 44 binding sites showed that at 23 sites the RNA from the X-ray model fitted into the EM density map (Supplementary Fig. 4), with some of them showing in detail the interactions between nucleotides A − 4 and A − 10 in the 19-nucleotide RNA chain of the RNA hairpin loop (1ZDH) that both make contact to Thr45 and Ser47 (not shown) located in the beta-sheets of different CP molecules (Fig. 1f). The observed amounts of SLs (59) and PSs (between 23 and 44) match earlier estimates, predicting 53 SLs and 35 PSs, that were based on the predicted RNA secondary structure, obtained by phylogenetic analysis and experimental probing of RNA in solution, and analysis of potential binding SLs. Our results show that the majority of the many SL RNA structures that are present in the genome of MS2 bind the CP2 in the capsid. Since these SLs all have different predicted sequences, as a result this suggests that a wide variety of different RNA sequences actually bind to CP2 in situ in an overall similar conjunction as the TR, similar to what was observed with several aptamer–CP2 interactions. At several sites RNA hairpin–EM densities were also observed near the capsid but in deviating conformations or sites, often rotated (7 cases) or shifted (7 cases), compared with the TR structure (data not shown). In addition, several stem structures were sideways associated to CP2 sites. Therefore alternative, non-sequence specific, binding modes of the RNA to the CP2 likely exist, including binding to other amino acids in the capsid.

The obtained resolution of the EM map is sufficient to visualize the double-stranded (ds) RNA SLs, including its helical nature; however, single RNA strands are not clearly resolved and therefore it was not possible to trace predicted secondary structures into the 3D map and to allocate predicted SLs in the MS2 genome into the density. To investigate all individual SL–CP2 interactions in the genome to the observed SLs in the EM map a higher resolution structure would be

Results

Asymmetric structure. The outside of our asymmetric EM map of the MS2 virion (Fig. 1a) at 8.7 Å (EMD-3403) is similar to the
required. Nevertheless, the two regions known to connect to the AP could putatively be allocated in the EM map. The 3′ RNA end, including nucleotides 3,510–3,527, forms a very specific repeat of short SLs that binds adjacently along the AP, while near the 5′ RNA the SL formed by nucleotides 388–414 binds the AP (Fig. 2a, Supplementary Movie 3).

RNA–protein interactions. While the AP ensures specific packaging of its coding RNA and TR–CP2 interactions promote capsid formation, the question is how the SL–CP2 interactions that we observed here influence virion assembly. To explore a potential role of the RNA genome structure in this process we investigated the distribution of PSs, interactions of SLs with CP2, over the capsid. From the 89 CP2, 44 (49%) have a connecting SL, 33 (37%) has crossing (ds) RNA density, while 9 (10%) does not have any adjacent density. Notably, these 44 PSs are distributed unevenly over the capsid, being localized predominantly on one side of the capsid (Supplementary Fig. 3). This uneven distribution was even more pronounced for the 23 RNA SLs of which the density matched the TR X-ray model (Supplementary Fig. 4). Of these, 19 (82%) were bound to dimers that were located on one of half of the capsid in only three CP2 pentamers (Fig. 2b, Supplementary Movie 4). These pentamers were adjacent to the AP and its two binding SLs.

This uneven distribution of SL–CP2 interactions supports a two-step encapsidation model in which RNA condensation proceeds full capsid formation. The role of the asymmetric distribution of RNA–protein interactions could be two-fold. Multiple SLs could increase efficient capsid formation by CP2 recruitment from the surroundings to form the first CP2 pentamers, adding both efficiency and localization to CP2–CP2 interactions that drive coat formation. Alternatively, CP2 could induce SLs formation and condensation of the MS2 RNA, thereby inducing genome compaction during encapsidation. These two

Figure 1 | The asymmetric reconstruction of bacteriophage MS2. Asymmetric structure of bacteriophage MS2 (green-blue radially coloured) shows the AP (yellow) (a), which replaces one CP dimer (b). Inside the protein capsid a structured genome (grey) is present (c) that is connected to the AP (d). The reconstruction shows the double-stranded helices in the stem loop structures (e). At some positions individual NA’s connecting to the capsid are resolved, as shown by fitting of the X-ray structure of the 19-nucleotide TR (magenta) bound to the capsid (blue) (pdb:1ZDH) in the EM density (grey) (f). Scale bar is 100 Å.
possible functions are not mutually exclusive and also support the suggested formation of new SLs in the MS2 genome.

**Discussion**

We reconstructed the MS2 virion, revealing the AP and the single conformation of the RNA genome in situ. The genome is intimately and asymmetrically linked to the capsid. The presence of a single AP, breaking the symmetry in the capsid, in MS2 is exceptional among viruses and might play an important role in the uniquely structured genome, which might not appear in other (small) viruses. Even so structures of several other viruses have shown hints of (partly) ordered (ds)RNA. It remains to be seen whether asymmetric cryo-EM single particle reconstructions of other viruses would reveal similar genome ordering inside virions. The potential of cryo-EM to explore asymmetric structure determination of complete virions, including their genome, would provide unprecedented details on viral RNA structures, RNA–protein interactions and insight into viral assembly, which is valuable knowledge for drug design by targeting disruptions of viral genome folding and packaging.

**Methods**

**Purification of phages.** An overnight culture of *E. coli* strain XL1 blue was grown at 37 °C in LB medium until an OD₆₀₀ of 0.5. Calcium chloride was added to a final concentration of 2 mM and the cells were infected with phage MS2 at a multiplicity of 10 and incubated for another three hours. Lysates with phage titres of approximately 1 × 10⁸ were used for purification. Cellular debris was removed by centrifugation and phage concentrated by ultrafiltration using 15 ml 100 kDa cutoff Amicon concentrators. Phage was further purified by gel-filtration...
Section preparation. Aliquots of purified MS2 were applied to glow-discharged holey carbon film supported by copper grids (Quantifoil R2/2) after glow discharging with negative polarity for 1 min at 30 mA using a K950X carbon coating (Emitech). Grids were vitrified by plunging into a liquid propane/ethane mixture (2:1 v/v), which was cooled by liquid nitrogen. Samples were plunged using a Leica EM GP from room temperature and blotted for 1–2 s using filter paper. After vitrification, the grids were stored in liquid nitrogen until use.

Data collection. Data acquisition was performed on a Titan Krios transmission electron microscope (FEI) operated with Cs correction at 300 keV using EPU automated single particles acquisition software (FEI). Seven frames per images were recorded on a back-thinned Falcon II detector at a nominal magnification of 59,000 × with a sampling size of 1.14 Å per pixel.

Image processing. Image processing was performed using Scipion platform (http://scipion.cnb.csic.es), which is an integrative image framework that currently mainly uses Xnipp (http://xmipp.cnb.csic.es/)3, Relion (http://www2.mrc-lmb.cam.ac.uk/reliqu/index.php/Main_Page), Spider (http://spider.wadsworth.org)5 and EMAN (http://blake.bcm.edu/emanwiki/ EMAN2)6 packages. Graphics were produced by UCSF Chimera (http://www.cgl.ucsf.edu/chimera)7. All movies were aligned using Optical Flow approach8, while contrast transfer functions (CTFs) were estimated using CTFIND3 (ref. 7), and were used to select the best quality micrographs. A total of 22,441 particles were picked automatically using Xmipp 8, which were further CTFFIND3 (ref. 7), and were used to select the best quality micrographs. A total of 59,000 × with a sampling size of 1.14 Å per pixel.

Resolution. Fifty percent of the particles were randomly split in two halves. Each subset was refined with the same conditions as the initial refinement was used, filtered to 25 Å. The first four iterations were performed as global refinement, and the RNA mixture (2:1 v/v), which was cooled by liquid nitrogen. Samples were plunged using a Leica EM GP from room temperature and blotted for 1–2 s using filter paper. After vitrification, the grids were stored in liquid nitrogen until use.

Data availability. Density maps of the icosaedral reconstruction, the asymmetric reconstruction, and the RNA + A protein, a difference map from the symmetric reconstruction and the dimer depleted capsid structure, are available from the Electron Microscopy Data Bank with accession codes: EMD-3402, EMD-3403, and EMD-3404, respectively. The authors declare that all data supporting the findings of this study are available from the corresponding author upon request.

References

1. Davis, J. E., Strauss, J. H. & Sinsheimer, R. Bacteriophage MS2: another RNA phage. Science 134, 1427 (1961).
2. Fiers, W. et al. Complete nucleotide sequence of bacteriophage MS2 RNA: primary and secondary structure of the replicase gene. Nature 260, 500–507 (1976).
3. Olschoorn, R. J. & van Duijn, J. Bacteriophages with ssRNA. Encycl. Life Sci. http://dx.doi.org/10.1002/9780470159029.a0000778.pub3 (2011).
4. Licis, N., Van Duijn, J., Baklava, Z. & Berzinis, V. Long-range translational coupling in single-stranded RNA bacteriophages: an evolutionary analysis. Nucleic Acids Res. 26, 3242–3246 (1998).
5. Poot, R., Tsareva, N. V., Boni, I. V. & van Duijn, J. RNA folding kinetics regulates translation of phage MS2 maturation gene. Proc. Natl Acad. Sci. 94, 10110–10115 (1997).
6. Adhin, M. R. & van Duijn, J. Scanning model for translational reinitiation in bacteriophage MS2. J. Mol. Biol. 231, 811–818 (1990).
7. Van Humberg, J., van Geffen, R. & van Duijn, J. Translational control by a long range RNA-RNA interaction: a basepair substitution analysis. Nucleic Acids Res. 21, 1713–1717 (1993).
8. Jou, W. M., Haegeman, G., Ysebaert, M. & Fiers, W. Nucleotide sequence of the gene coding for the bacteriophage MS2 coat protein. Nature 237, 82–88 (1972).
9. Oh, S. & Fersht, A. R. Nucleotide sequence at the binding site for coat protein on RNA of bacteriophage R17. Proc. Natl Acad. Sci. 69, 3033–3037 (1972).
10. Sugiyama, T., Hebert, R. R. & Hartman, K. A. Ribonucleoprotein complexes formed between bacteriophage MS2 RNA and MS2 protein in vitro. J. Mol. Biol. 25, 455–463 (1967).
11. Hwang, P. P., Ling, C. M. & Overby, L. R. Self-assembly of Q-beta and MS2 phage particles: possible function of initiation complexes. Science 166, 1638–1640 (1969).
12. Shiba, T. & Suzuki, Y. Localization of A protein in the RNA-A protein complex of RNA phage MS2. Biochim. Biophys. Acta 654, 249–255 (1981).
13. Borodavka, A., Tuma, R. & Stockley, P. G. A two-stage mechanism of viral RNA condensation revealed by single molecule fluorescence. RNA Biol. 10, 481–489 (2013).
14. Dykenava, E. C., Stockley, P. G. & Twarock, R. Packaging signals in two single-stranded RNA viruses imply a conserved assembly mechanism and geometry of the packaged genome. J. Mol. Biol. 425, 3325–3349 (2013).
15. Valegård, K., Liljas, L., Eri said, K. & Unge, T. The three-dimensional structure of the bacterial virus MS2. Nature 345, 36–41 (1990).
16. Ni, C. Z. et al. Crystal structure of the MS2 coat protein dimer: implications for RNA binding and virus assembly. Structure 3, 255–263 (1995).
17. Valegård, K., Murray, J. B., Stockley, P. G., Stonehouse, N. J. & Liljas, L. Crystal structure of an RNA bacteriophage coat protein-operator complex. Nature 371, 623–626 (1994).
18. Convery, M. A. et al. Crystal structure of an RNA aptamer-protein complex at 2.8A resolution. Nat. Struct. Biol. 5, 133–139 (1998).
19. Horn, W. T. et al. The crystal structure of a high affinity RNA stem-loop complexed with the bacteriophage MS2 capsid: further challenges in the modeling of ligand-RNA interactions. RNA 10, 1776–1782 (2004).
20. Roswell, S. et al. Crystal structures of a series of RNA aptamers complexed to the same protein target. Nat. Struct. Biol. 5, 970–975 (1998).
21. Valegård, K. et al. The 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 (1997).
22. Van Den Worm, S. H. E. et al. Crystal structures of MS2 coat protein mutants in complex with wild-type RNA operator fragments. Nucleic Acids Res. 26, 1345–1351 (1998).
23. Toropova, K., Basnak, G., Twarock, R., Stockley, P. G. & Ranson, N. A. The three-dimensional structure of genomic RNA in bacteriophage MS2: implications for assembly. J. Mol. Biol. 375, 824–836 (2008).
24. Koning, R. et al. Visualization by cryo-electron microscopy of genomic RNA that binds to the protein capsid inside bacteriophage MS2. J. Mol. Biol. 332, 415–422 (2003).
25. Dent, K. C. et al. The asymmetric structure of an icosaedral virus bound to its receptor suggests a mechanism for genome release. Structure 21, 1225–1234 (2013).
26. Van Den Worm, S. RNA-protein interactions in the ss-RNA phages (Leiden University, 2004).
27. Borodavka, A., Tuma, R. & Stockley, P. G. Evidence that viral RNAs have evolved for efficient, two-stage packaging. Proc. Natl Acad. Sci. 109, 15769–15774 (2012).
28. Larson, S. B. et al. Double-helical RNA in satellite tobacco mosaic virus. Nature 361, 179–182 (1993).
29. Zhang, X. et al. In situ structures of the segmented genome and RNA polymerase complex inside a dsRNA virus. Nature 527, 531–534 (2015).

Acknowledgements

We thank Sjoerd van den Worm, Rene Olschoorn, Janis Rummink and Christoph Diebold for discussions and reading of the manuscript. This research has been executed with support of NoCEN, the Netherlands Centre for Electron Microscopy, Leiden, NL and NWO, the Netherlands Organisation for Scientific Research and is partly financed by the European Regional Development Fund of the European Commission. We acknowledge the Spanish Ministry of Economy and Competitiveness through Grants AIC-A-2011-0638, BIO2013-44647-R. This work was funded by Instruct (Grant reference PID: 1114), part of the European Strategy Forum on Research Infrastructures (ESFRI) in collaboration with the Instruct Image Processing Center I2PC. This work was supported by NanoNextNL of the Government of the Netherlands and the 130 partners.

Author contributions

R.K. conceived the project, collected and analysed the data and wrote the manuscript, J.G.B and J.V. performed image processing, analysed the data, and wrote the manuscript I.A. and A.K. purified the MS2, K.T, J.M.C. and A.K.J. supervised the studies.

Additional information

Supplementary Information accompanies this paper at http://www.nature.com/ncomms
Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Koning, R. I. et al. Asymmetric cryo-EM reconstruction of phage MS2 reveals genome structure in situ. Nat. Commun. 7:12524 doi: 10.1038/ncomms12524 (2016).

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/

© The Author(s) 2016