

Complete Genome Sequence of the Broad Host Range Single-Stranded RNA Phage PRR1 Places It in the *Levivirus* Genus with Characteristics Shared with Alloleviviruses

Tanja M. Ruokoranta, A. Marika Grahn,† Janne J. Ravantti, Minna M. Poranen, and Dennis H. Bamford*

Department of Biological and Environmental Sciences and Institute of Biotechnology, Viikki Biocenter,
P.O. Box 56, 00014 University of Helsinki, Helsinki, Finland

Received 16 May 2006/Accepted 21 June 2006

Single-stranded RNA (ssRNA) bacteriophages of the family *Leviviridae* infect gram-negative bacteria. They are restricted to a single host genus. Phage PRR1 is an exception, having a broad host range due to the promiscuity of the receptor encoded by the IncP plasmid. Here we report the complete genome sequence of PRR1. Three proteins homologous with those of other ssRNA phages, i.e., maturation, coat, and replicase proteins, were identified. A fourth protein has a lysis function. Comparison of PRR1 with other members of the *Leviviridae* family places PRR1 in the genus *Levivirus* with some characteristics more similar to those of members of the genus *Allolevivirus*.

The discovery of single-stranded RNA (ssRNA) phages dates back to the early 1960s, when Tim Loeb and Norton Zinder identified male-specific coliphages in New York sewage (35). Characterization of these viruses has helped the understanding of RNA virus replication, translation, virus assembly, and protein-RNA interactions. All ssRNA phages that have been isolated so far are icosahedral (diameter, ~270 nm) and adsorb to the pili of gram-negative bacteria. These phages form the family *Leviviridae*, which has two genera, *Levivirus* and *Allolevivirus*. Leviviruses contain a genome of ~3,500 nucleotides (nt), with genes encoding replicase, coat, and maturation proteins and a protein with lysis function, and phage MS2 is representative of leviviruses. Alloleviviruses have a genome of ~4,200 nt. The prototype virus for this genus is Q β , and the major difference compared to leviviruses is the presence of an additional read-through gene. The complete genome sequences of 11 *Leviviridae* phages have been deposited in GenBank (<http://www.ncbi.nlm.nih.gov>; 10 March 2006). Phages MX1 and SP were not included here, since they are identical to Q β and FI, respectively.

The ssRNA phage PRR1 was isolated in Michigan in the early 1970s (21). PRR1 is unique because it adsorbs to a pilus encoded by broad-host-range incompatibility group P plasmids, thus extending the host range of PRR1 to a number of gram-negative bacteria (22). Herein, the genome sequence of PRR1, along with comparisons to other ssRNA phages, is reported.

PRR1 was propagated in *Pseudomonas aeruginosa* PAO1 (12). Virus particles were purified as described previously (3) and concentrated by differential centrifugation. The genomic RNA was isolated by two successive phenol extractions followed by ethanol precipitation. cDNA was generated (nt re-

gion 30 to 3519) using random hexamers and Thermoscript reverse transcriptase (Pharmacia), amplified by PCR with Dynazyme (Finnzymes), and inserted into plasmid pDrive, which was transferred to *Escherichia coli* K-12 Ez (QIAGEN). The termini were captured by rapid amplification of cDNA ends methodology (25) using terminal transferase (Finnzymes). Both strands from several overlapping clones were sequenced. The start codon of the putative lysis protein was changed to AUG, and the gene (nt region 1780 to 1944) was inserted into plasmid pJ2 (20), which was transferred to *E. coli* BL21(DE3) pLysS (Novagen).

The genome sequence of PRR1 (3,573 nt) allowed the identification (BLAST) (2) of four genes (Fig. 1A) organized similarly to members of the genus *Levivirus*. The overall nt sequence similarity to other ssRNA phage genomes ranged from 43% to 48%. The unweighted-pair group method using average linkages (UPGMA) tree of the entire genomes of sequenced ssRNA phages (Fig. 1B) positions PRR1 and AP205 outside of the allolevivirus Q β and levivirus MS2 clusters, indicating broad sequence variability among ssRNA phages.

To confirm the putative lysis gene location, nt region 1780 to 1944 was inserted into an expression vector. Cell lysis occurred upon induction (Fig. 1C). There are two successive (6 nt in between) unusual start codons, AUU and GUG, in the PRR1 lysis gene, which supports the proposed “scanning model” (1) for the initiation of lysis gene translation. The lysis protein of PRR1 has one predicted transmembrane helix (residues 24 to 46).

The clustering of the individual putative PRR1 proteins (Fig. 2) revealed an intriguing grouping. The maturation protein was closest to type I leviviruses (MS2), whereas the coat protein was closer to alloleviviruses (Q β and SP). The replicase and lysis proteins were closest to type II leviviruses (GA). Conserved regions in the replicase gene reported in other ssRNA phages (16, 23) also were found in PRR1. Nearly uniform-sized coat proteins (Fig. 1A) ranged from 128 amino acids (aa) in PP7 to 133 aa in Q β , with PRR1 having 132 aa. The

* Corresponding author. Mailing address: Viikki Biocenter 2, P.O. Box 56 (Viikinkaari 5), FIN-00014 University of Helsinki, Helsinki, Finland. Phone: 358-9-19159100. Fax: 358-9-19159098. E-mail: dennis.bamford@helsinki.fi.

† Present address: Roal Oy, P.O. Box 57, 05201 Rajamäki, Finland.

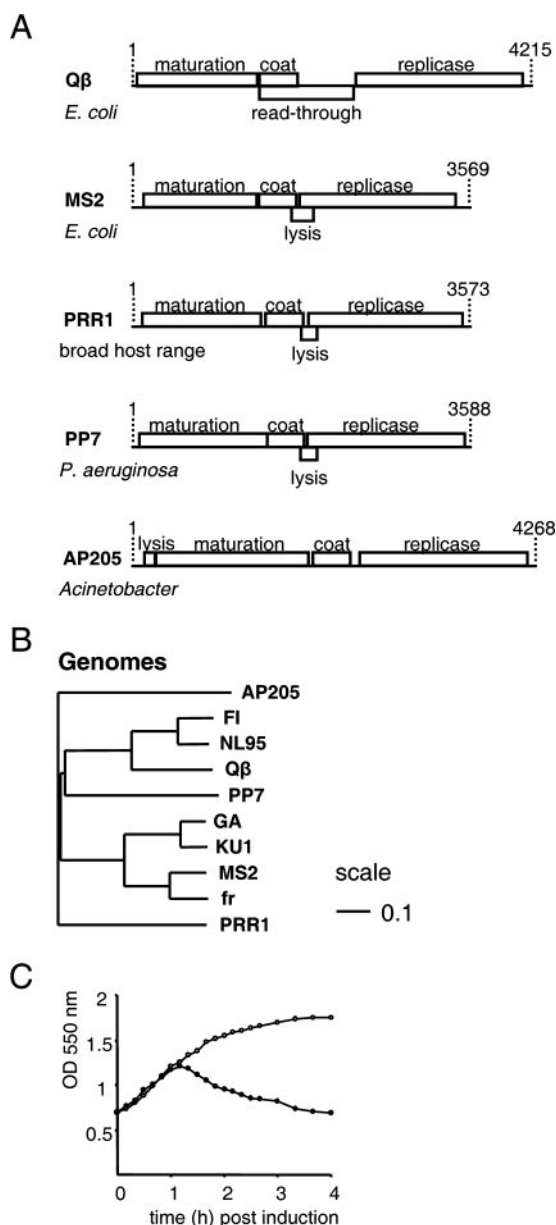


FIG. 1. (A) Genetic maps of representative ssRNA phages of the *Leviviridae* family. The map of AP205 is a correction of that published in reference 16. The correction is presented in reference 32. (B) An UPGMA tree showing relationships between ssRNA phages based on nucleotide sequence comparison of the complete genomes constructed by ClustalW (30) and PHYLIP (9). Scale, 0.1 units. (C) Induction of bacteria with (filled circles) and without (open circles) the putative lysis gene (nt 1780 to 1944). The expression of the lysis gene was induced with 1 mM IPTG (isopropyl- β -D-thiogalactopyranoside) at a cell density of 3×10^8 CFU/ml. OD 550 nm, optical density at 550 nm.

protein sequence obtained differed by two residues (V-89 \rightarrow I, and an additional K located at position 116) from that reported previously (8).

Three-dimensional structures determined for virions Q β (10), MS2 (31), fr (17), GA (28), and PP7 (29) allowed threading of the PRR1 coat protein sequence to available structures. With the exceptions of MS2 and fr, the overall sequence sim-

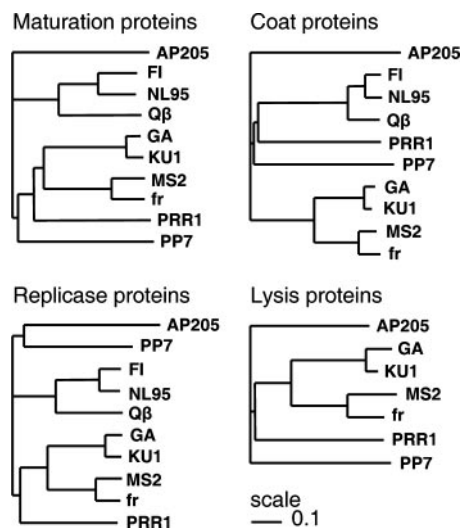


FIG. 2. UPGMA trees showing relationships between ssRNA phage proteins based on amino acid sequence comparisons constructed by ClustalW (30) and PHYLIP (9).

ilarity between different coat proteins was <30%. Consequently, each coat protein was modeled separately. For the final analysis, only models derived from the Q β , MS2, and PP7 structures were used. Despite the low sequence similarity, the models (Fig. 3) strongly resembled each other. The β -sheets E, F, and G (RNA binding region) (10, 13) and the interacting C-terminal helices especially were conserved. The greatest differences were in the flexible loop regions. Interestingly, the β -sheets A and B on the modeled PRR1 structures consistently resembled similarly located MS2 β -sheets, although the coat protein sequence clustered more closely to those of alloviruses (e.g., Q β) (Fig. 2).

ssRNA phages regulate vital functions by specific secondary and tertiary RNA structures (32). Secondary structure predictions for the PRR1 genome resulted in ambiguous results for the 3' terminus and the coat start regions. However, interpretable results were obtained with the 5' terminus and replicase start regions (Fig. 4). The PRR1 5' terminus (Fig. 4A) resembled that of the inactive conformation of phage MS2 (11) except for the absence of the "south" hairpin. Interestingly, a distant sequence (coordinates 1251 to 1269) complementary to the maturation gene Shine-Dalgarno sequence (SD) was found in PRR1, resembling the regulation system described for Q β (4, 6, 7). If both aforementioned features operate in PRR1, the maturation gene translation is very tightly controlled.

Unlike in other ssRNA phages, the start codon for the PRR1 replicase gene is located upstream of the predicted operator hairpin instead of being part of the hairpin (Fig. 4B) (16) and the SD site is part of an extensive long-range interaction designated Min Jou (19). Long-distance interactions in this region were more extensive in PRR1 than in other ssRNA viruses. These observations support the proposal that the operator hairpin is not the only element for the primary control of the replicase gene translation initiation (5, 33). As in the case for *Acinetobacter* phage AP205 (16), the operator hairpin in PRR1 has characteristics of both of the ssRNA phage gen-

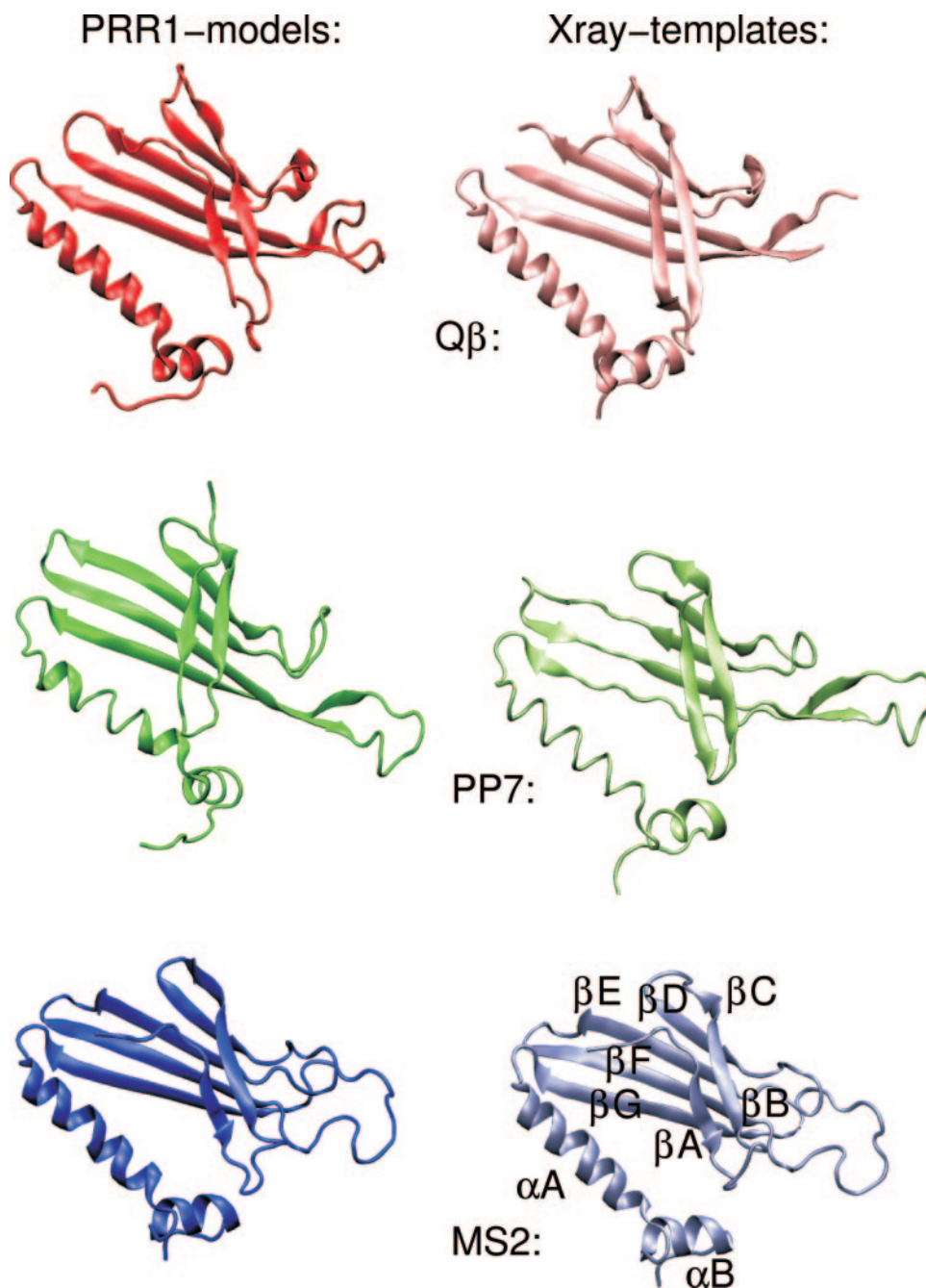


FIG. 3. Three-dimensional models of the PRR1 coat protein. Each model was built separately using the MODELLER program (18). The structural superpositioning of the models (left) and known structures for Q β , PP7, and MS2 coat proteins (right) was performed using the STAMP algorithm (24), and the results were visualized with the VMD program (14). β -Sheets A and B always are modeled similarly despite the underlying template structure. The secondary elements are depicted in the MS2 X-ray structure.

era. PRR1 has a bulged nucleotide C in the operator hairpin as in alloviruses, but it is positioned on the “left” side as in leviviruses. Leviviruses have a 4-nt loop in the operator hairpin, while alloviruses have a 3-nt loop (16). *Pseudomonas* phage PP7 is an exception, with a 6-nt loop (23), as is PRR1, with its 5-nt loop. Coat protein gene terminator hairpins are similar in all ssRNA phages studied so far (23). However, the PRR1 hairpin equivalent differs considerably from the others,

since the coat protein gene stop codon is positioned in the lysis gene hairpin.

The stem-loop structures of the 3' untranslated region in the members of the two *Leviviridae* genera differ. PRR1 takes an intermediary position between the complex 3' untranslated region structures of phage MS2 and the simpler ones of phages SP, PP7, and AP205 (not shown). Conserved U1 stem-loops (16) have an invariable UGCUU sequence in

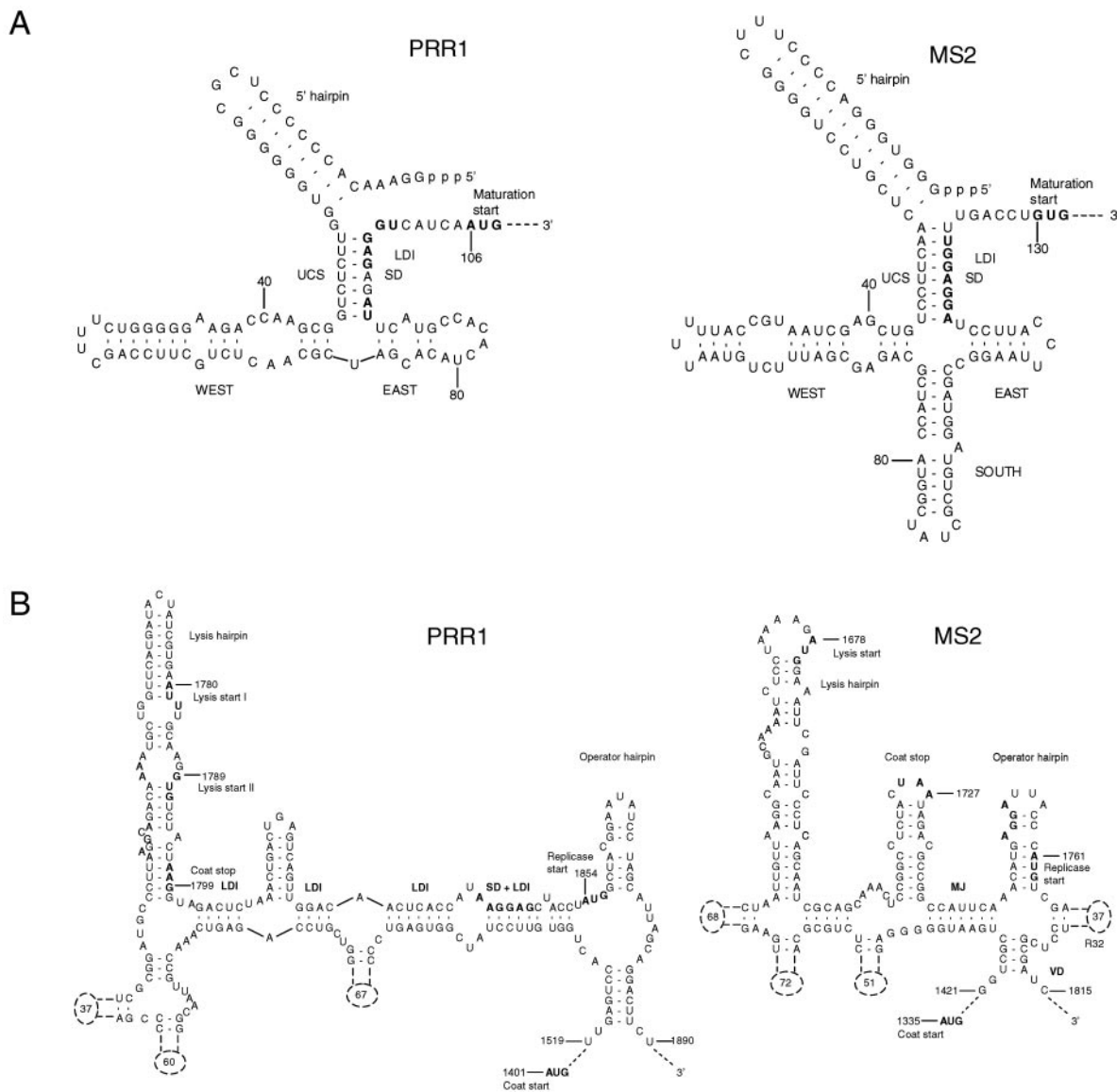


FIG. 4. (A) Equilibrium RNA structures in the 5' untranslated region of PRR1 and MS2. The MS2 structure was modified from that reported previously (inactive maturation gene) (34). The PRR1 RNA structure was attained by forcing the long-range interactions to pair (MFOLD) (36). 5', west, east, and south hairpins are indicated. (B) RNA structure around the start of the replicase gene. The model for MS2 has been developed by phylogenetic sequence comparison, enzymatic and chemical experiments, and comparison of phenotypes of constructed mutants (27). The two putative start sites for the PRR1 lysis gene are marked with I and II. Numbers in dashed hairpins indicate the number of nucleotides not drawn. Start and stop codons of the genes and the ribosome-binding sites (SD) are marked in bold. LDI, long-distance interaction; UCS, upstream complementary sequence. Min Jou (MJ) (19), R32, and VD are long-distance interactions. Structure predictions for PRR1 were generated by the MFOLD program (36).

the loop, except for that of PP7, in which it is CGCUC. PRR1 differed from them with a CGCUU sequence, which forms a long-range pseudoknot important for replication regulation (15, 26). In the PRR1 sequence, there was a fully complementary sequence within the replicase gene (coordinates 2828 to 2832).

PRR1 is clearly a levivirus, based on its genome organization (Fig. 1A) and the amino acid sequences of all proteins, but the coat protein more strongly resembles that of alloviruses than that of leviruses (Fig. 2). This suggests recombination between members of alloviruses and leviruses. PRR1 is a

broad-host-range phage that might promote horizontal gene transfer and recombination between different members of the *Leviviridae* family.

Nucleotide sequence accession number. The sequence has been deposited in GenBank (<http://www.ncbi.nlm.nih.gov>) under accession number DQ836063.

Petri Auvinen is acknowledged for valuable discussions. Research grants 162993, 164298, and 172621 (Finnish Center of Excellence Programme [2000-2005]) from the Academy of Finland supported this study.

REFERENCES

1. Adhin, M. R., and J. van Duin. 1990. Scanning model for translational reinitiation in eubacteria. *J. Mol. Biol.* **213**:811–818.
2. Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**:403–410.
3. Bamford, D. H., L. Rouhiainen, K. Takkinen, and H. Soderlund. 1981. Comparison of the lipid-containing bacteriophages PRD1, PR3, PR4, PR5 and L17. *J. Gen. Virol.* **57**:365–373.
4. Beekwilder, J., R. Nieuwenhuizen, R. Poot, and J. van Duin. 1996. Secondary structure model for the first three domains of Q β RNA. Control of A-protein synthesis. *J. Mol. Biol.* **256**:8–19.
5. Berkhout, B., and J. van Duin. 1985. Mechanism of translational coupling between coat and replicase genes in phage MS2. *Nucleic Acids Res.* **13**:6955–6967.
6. de Smit, M. H., and J. van Duin. 1990. Secondary structure of the ribosome binding site determines translational efficiency: a quantitative analysis. *Proc. Natl. Acad. Sci. USA* **87**:7668–7672.
7. de Smit, M. H., and J. van Duin. 1994. Control of translation by mRNA secondary structure in *Escherichia coli*: a quantitative analysis of literature data. *J. Mol. Biol.* **244**:144–150.
8. Dhase, P., J. S. Vandekerckhove, and M. C. van Montagu. 1979. The primary structure of the coat protein of the broad-host-range RNA bacteriophage PRR1. *Eur. J. Biochem.* **94**:375–386.
9. Felsenstein, J. 1993. PHYLIP (Phylogeny Inference Package) version 3.5c. Department of Genetics, University of Washington, Seattle.
10. Golmohammadi, R., K. Fridborg, M. Bundule, K. Valegård, and L. Liljas. 1996. The crystal structure of bacteriophage Q β at 3.5 Å resolution. *Structure* **4**:543–554.
11. Groeneveld, H., K. Thimon, and J. van Duin. 1995. Translational control of maturation protein synthesis in phage MS2: a role of kinetics of RNA folding. *RNA* **1**:79–88.
12. Holloway, B. W. 1969. Genetics of *Pseudomonas*. *Bacteriol. Rev.* **33**:419–443.
13. Horn, W. T., K. Tars, E. Grahn, C. Helgstrand, A. J. Baron, H. Lago, C. J. Adams, D. S. Peabody, S. E. Phillips, N. J. Stonehouse, L. Liljas, and P. G. Stockley. 2006. Structural basis of RNA binding discrimination between bacteriophages Q β and MS2. *Structure* **14**:487–495.
14. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**:33–38.
15. Klovin, J., and J. van Duin. 1999. A long-range pseudoknot in Q β RNA is essential for replication. *J. Mol. Biol.* **294**:875–884.
16. Klovin, J., G. P. Overbeek, S. H. E. van den Worm, H.-W. Ackermann, and J. van Duin. 2002. Nucleotide sequence of a ssRNA phage from *Acinetobacter*: kinship to coliphages. *J. Gen. Virol.* **83**:1523–1533.
17. Liljas, L., K. Fridborg, K. Valegård, M. Bundule, and P. Pumpers. 1994. Crystal structure of bacteriophage fr capsids at 3.5 Å resolution. *J. Mol. Biol.* **244**:279–290.
18. Marti-Renom, M. A., A. C. Stuart, A. Fiser, R. Sánchez, F. Melo, and A. Sali. 2000. Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* **29**:291–325.
19. Min Jou, W., G. Haegeman, M. Ysebaert, and W. Fiers. 1972. Nucleotide sequence of the gene coding for the bacteriophage MS2 coat protein. *Nature (London)* **237**:82–88.
20. Ojala, P. M., J. T. Juuti, and D. H. Bamford. 1993. Protein P4 of double-stranded RNA bacteriophage $\phi 6$ is accessible on the nucleocapsid surface: epitope mapping and orientation of the protein. *J. Virol.* **67**:2879–2886.
21. Olsen, R. H., and D. D. Thomas. 1973. Characteristics and purification of PRR1, an RNA phage specific for the broad host range *Pseudomonas* R1822 drug resistance plasmid. *J. Virol.* **12**:1560–1567.
22. Olsen, R. H., and P. Shipley. 1973. Host range and properties of the *Pseudomonas aeruginosa* R factor R1822. *J. Bacteriol.* **113**:772–780.
23. Olsthoorn, R. C., G. Garde, T. Dayhuff, J. F. Atkins, and J. van Duin. 1995. Nucleotide sequence of a single-stranded RNA phage from *Pseudomonas aeruginosa*: kinship to coliphages and conservation of regulatory RNA structures. *Virology* **206**:611–625.
24. Russell, R. B., and G. J. Barton. 1992. Multiple protein sequence alignment from tertiary structure comparison: assignment of global and residue confidence levels. *Proteins* **14**:309–323.
25. Sambrook, J., and D. W. Russell. 2001. Molecular cloning: a laboratory manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
26. Schuppli, D., I. Barrera, and H. Weber. 1994. Identification of recognition elements of bacteriophage Q β minus strand RNA that are essential for template activity with Q β replicase. *J. Mol. Biol.* **243**:811–815.
27. Skripkin, E. A., M. R. Adhin, M. H. de Smit, and J. van Duin. 1990. Secondary structure of the central region of bacteriophage MS2 RNA. *J. Mol. Biol.* **211**:447–463.
28. Tars, K., M. Bundule, K. Fridborg, and L. Liljas. 1997. The crystal structure of bacteriophage GA and a comparison of bacteriophages belonging to the major groups of *Escherichia coli* leviviruses. *J. Mol. Biol.* **271**:759–773.
29. Tars, K., K. Fridborg, M. Bundule, and L. Liljas. 2000. The three-dimensional structure of bacteriophage PP7 from *Pseudomonas aeruginosa* at 3.7-Å resolution. *Virology* **272**:331–337.
30. Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. ClustalW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**:4673–4680.
31. Valegård, K., L. Liljas, K. Fridborg, and T. Unge. 1990. The three-dimensional structure of bacterial virus MS2. *Nature* **345**:36–41.
32. van Duin, J., and N. Tsareva. 2004. Single-stranded RNA phages, p. 175–196. *In* R. Calendar and S. Abedon (ed.), *The bacteriophages*, 2nd ed. Oxford University Press, New York, N.Y.
33. van Himbergen, J., B. van Geffen, and J. van Duin. 1993. Translational control by a long range RNA-RNA interaction: a basepair substitution analysis. *Nucleic Acids Res.* **21**:1713–1717.
34. van Meerten, D., G. Genevieve, and J. van Duin. 2001. Translational control by delayed RNA folding: Identification of the kinetic trap. *RNA* **7**:483–494.
35. Zinder, N. D. 1975. RNA phages, p. v. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
36. Zuker, M., D. H. Mathews, and D. H. Turner. 1999. Algorithms and thermodynamics for RNA secondary structure prediction: a practical guide, p. 11–43. *In* J. Barciszewski and B. F. C. Clark (ed.), *RNA biochemistry and biotechnology*. NATO ASI series. Kluwer Academic Publishers, Dordrecht, The Netherlands.