

In vitro characterisation of the MS2 RNA polymerase complex reveals novel host factors that modulate leviviral replicase activity

Alexander Wagner

Max Planck Institute of Biochemistry

Laura Weise

Max Planck Institute of Biochemistry

Hannes Mutschler (✉ hannes.mutschler@tu-dortmund.de)

TU Dortmund University <https://orcid.org/0000-0001-8005-1657>

Article

Keywords: MS2, RNA, replicase

Posted Date: April 12th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-378476/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Communications Biology on March 25th, 2022. See the published version at <https://doi.org/10.1038/s42003-022-03178-2>.

Abstract

The coliphage MS2 is a well-established model organism that has helped to reveal a variety of fundamental concepts for RNA-based translational control and viral RNA packaging. Despite the comprehensive characterisation of the MS2 life cycle, the macromolecular composition of its RNA replicase remained obscure. Here, we sought to identify the missing host proteins required for the assembly of the active replicase complex. By combining a purified, inactive MS2 replicase sub-complex with selected modules of an *in vitro* translation system, we confirmed that the three suspected host factors EF-Ts, EF-Tu and ribosomal protein S1 form an active core-enzyme with the viral replicase subunit. Unexpectedly, we also found that the translation initiation factors IF1 and IF3 directly modulate MS2 replicase activity. While IF1 enhances replicase activity in a template-independent manner, IF3 acts as an inhibitor that prevents polymerase initiation and / or elongation. Both observations suggest a previously unknown role of these host proteins during the phage life cycle. Finally, we demonstrate the *in vitro* formation of small RNAs that contain minimal motifs required for MS2 replicase-dependent amplification. Our work sheds new light on the architecture of the MS2 replication machinery while also providing the basis for a new cell-free evolution platform.

Introduction

The leviviral bacteriophage MS2 is one of the smallest viral pathogens known with its single-stranded 3.6 kilobase (kb) RNA genome encoding for only four proteins (1, 2). Due to its simple architecture, MS2 finds extensive use as a model organism in various biological research fields and helped to reveal a large number of fundamental biological processes such as translational control, feedback inhibition and key aspects of viral life cycles (3–6). Moreover, MS2 is used as a surrogate for human pathogenic RNA viruses to study the properties, stability and detectability of RNA viruses under different environmental conditions (7–9). Apart from their relevance to RNA biochemistry and molecular biology, protein and RNA components of ssRNA coliphages such as Q β and MS2 serve as attractive platforms for RNA imaging (10–12), RNA packaging and delivery (13–15), and molecular evolution and gene expression in cell-free systems (16–20).

Infection of male *Escherichia coli* cells by MS2 is dependent on the maturation protein (mp), which is present as a single copy in the MS2 capsid and enables binding of MS2 virions to the host's F-pili (6). Following infection, the host translation machinery immediately begins synthesizing the MS2 coat protein (cp), whose main function is to encapsulate newly synthesised viral (+) strands. Ribosomal read-through events of this gene provide a brief opportunity for initiation of translation of the replicase (rep) gene, whose ribosome-binding site is otherwise sequestered in an "operator" hairpin RNA structure (21) and inhibited by long-range RNA-RNA interactions (22). Similarly, timing and level of mp expression are controlled by the RNA folding kinetics of an untranslated leader sequence (23). To initiate genome replication, the MS2rep subunit forms an active holocomplex, which likely consists of the *E. coli* elongation factors EF-Tu, EF-Ts, ribosomal protein S1 and a yet undefined host protein (Fig. 1) (1). This complex synthesises new genomic (+) strand using antigenomic (-) strands intermediates as template. At

later stages of the life cycle, cp dimers bind and stabilize the operator hairpin which ultimately leads to a complete suppression of the rep translation (21). Finally, expression of the lysis gene (*lys*) leads to the release of mature virions from the infected host, enabling new cycles of infection (6).

Leviviral replicases are amongst the most ancient RNA-dependent polymerase enzymes and share key characteristics with other viral RNA polymerases (24), which makes them attractive model systems for drug discovery and mutational analyses. However, although MS2 is one of the main model systems for RNA viruses, fundamental aspects of its genome replication remain unclear to date. In particular, a host factor postulated almost 50 years ago (25, 26), that is required for RNA replicase activity remains unknown. Here, using our previously established reporter system for MS2 replicase activity in PURE systems (20), we are able to identify new host factors that modulate the activity of the phage enzyme. Through stepwise reduction of PURE proteins, we identified initiation factor 1 (IF1) as a co-factor that stimulates the activity of the MS2 replicase core complex consisting of MS2rep, EF-Tu, EF-Ts, and ribosomal protein S1. Furthermore, we found that initiation factor 3 (IF3) acts as an inhibitor for replicase activity, which has potential implications for the viral life cycle. Finally, we demonstrate that the MS2 replicase complex can evolve small replicable RNA species during the *in vitro* replication of the full-length genome. These new RNA scaffolds may be well-suited for the generation of robust self-replicating synthetic RNA systems in cell-free *in vitro* systems.

Material And Methods

Preparation of RNAs and plasmids

The preparation of all RNAs and the cloning of all plasmids are described in detail in the Supplementary Methods. Supplementary Table 1 contains all primers used for RNA preparation and cDNA synthesis. Supplementary Table 2 lists all DNA templates used for IVT. Supplementary Table 3 contains the sequences of all plasmids used in this study.

Preparation of Proteins

The preparation of proteins followed an adapted version of the protocol described by Shepherd *et al.* (27). In short, Top10 *E. coli* cells (Thermo Fisher Scientific) were transformed with each plasmid by electroporation. Proteins were overexpressed at 16°C overnight in Lysogeny Broth (LB Lennox) with 0.2 % L-Arabinose for all pBAD33-based constructs or 1 mM IPTG for pLD1-3, respectively. Subsequently, proteins were purified over HisPur™ Ni-NTA resin (Thermo Fisher Scientific) using HEPES buffer (50 mM HEPES·KOH pH 7.5, 250 mM NH₄Cl, 10 mM MgCl₂, 5 mM DTT) and stored in HEPES/glycerol buffer (50 mM HEPES·KOH pH 7.5, 100 mM KCl, 10 mM MgCl₂, 7 mM DTT, 30 % Glycerol). Preparations from pLD1, pLD2 and pLD3 were adjusted to ten-fold stocks with the final protein content per reaction equalling concentrations as described previously (27). A more detailed protocol is included in the Supplementary Methods.

Real time fluorescence measurements

All readout constructs are based on designs by Weise *et al.* (20). For experiments using the PURExpress® In Vitro Protein Synthesis Kit (NEB), standard reactions were supplemented with 10 µM DFHBI-1T, MS2rep and F30-Bro(-) or MS2-Bro(+), respectively, as described for the individual experiments:

Transcription of F30-Bro(-) in PURExpress

300 nM MS2rep, 150 nM F30-Bro(-) and 10 nM DFHBI-1T

Transcription of MS2-Bro(+) in PURExpress

1 µM MS2rep, 15 nM MS2-Bro(+) and 10 nM DFHBI-1T

All experiments based on the homemade PURE system (PURE 3.0/ PUREred) were assembled including 0.5 mM each ATP, GTP, CTP and UTP, 10 mM DTT, 15 µM EF-Tu, 1.5 µM S1 and all other components as followed.

Transcription of F30-Bro(-) in PURE 3.0

300 nM MS2rep, 150 nM F30-Bro(-), 1x LD1, LD2 and LD3, and 10 mM DFHBI-1T

LD2 depletion experiments

300 nM MS2rep, 50 nM *F30-Bro(-)*, 5 µM LD2 components, and 10 mM DFHBI-1T

Single LD2 component

300 nM MS2rep, 50 nM *F30-Bro(-)*, 15 µM EF-Ts, 5 µM LD2 components, and 10 mM DFHBI-1T

Co-factor titration

300 nM MS2rep, 50 nM *F30-Bro(-)*, 15 µM EF-Ts, and 10 mM DFHBI-1T

MS2-Bro replication

1 µM nM MS2rep, 50 nM *F30-Bro*, 15 µM EF-Ts, 10 mM DFHBI-1T, and 0.5 U per µL RNase inhibitor (moloX)

Untemplated replication

200 µM nM MS2rep or Qβrep, 15 µM EF-Ts, and 1x SYBR™ Green II Nucleic Acid gel stain (Thermo Fisher Scientific)

The final concentration of MgCl₂ was 6 mM, with HEPES, KCl and glycerol supplemented in 50 mM, 100 mM and 18 %, respectively. All reactions were prepared in MicroAmp Fast 8-TubeStrips (Thermo Fisher Scientific) and incubated at 37°C in a StepOne Real-Time PCR System (Thermo Fisher Scientific). Fluorescence signals were recorded every 60 s over a total of four hours or every 120 s for six hours for

experiments involving F30-Bro(+/-) or MS2-Bro(+/-), respectively. Unless stated otherwise, all experiments were performed in independent triplicates assembled from the same stock solutions.

Serial transfer and sequencing

For serial transfer experiments, reactions were setup by mixing components in the following composition: 15 μ M EF-Tu, 15 μ M EF-Ts, 1.5 μ M S1, 1 μ M MS2rep, 0.5 mM of each ATP, GTP, CTP and UTP, 10 mM DTT and 0.5 U / μ L RNase inhibitor (moloX). IF1 was supplemented at a final concentration of 15 μ M. The final concentration of MgCl₂ was 6 mM, with HEPES, KCl and glycerol supplemented in 50 mM, 100 mM and 18 %, respectively. For the first reaction, MS2wt RNA was provided at 50 nM final concentration. Reactions were incubated at 37°C for three hours, frozen in liquid nitrogen, and stored at 80°C. New reactions were mixed as described above, except that a 1/5 reaction volume (corresponding to MS2wt RNA in H₂O) was replaced with an aliquot from the previous reaction.

To obtain sequencing data, RNAs from serial transfer experiments were purified using Monarch® RNA Cleanup Kit (NEB) then polyadenylated using *E. coli* Poly(A) Polymerase (NEB) and re-purified with the same clean-up kit. Double-stranded cDNAs were synthesized using Template Switching RT Enzyme Mix (NEB) in combination with primers CDSII-T₂₄VN and TSO-CDSII (Supplementary Table 1) following the manufacturer's 2nd Strand cDNA Synthesis protocol. DNAs were purified using Monarch® PCR & DNA Cleanup Kit (NEB) and subsequently prepared for cloning with NEB® PCR Cloning Kit (NEB). For transformation, 2 μ L of the reactions were transformed into chemically competent Top10 *E. coli* cells (Thermo Fisher Scientific).

The replication capability of MSRP-22 was analysed by gel electrophoresis. The samples for the replication time course were obtained by preparing following reaction mixes and incubating at 37°C: 15 μ M EF-Tu, 15 μ M EF-Ts, 1.5 μ M S1, 0.5 mM of each ATP, GTP, CTP and UTP, and 10 mM DTT. The final concentration of MgCl₂ was 6 mM, with HEPES, KCl and glycerol supplemented at 50 mM, 100 mM and 18 %, respectively. MS2rep was supplemented at 500 nM concentration and RNAs in the concentrations described in Fig. 7.

Preparation of gel electrophoresis samples

All samples for gel electrophoresis were prepared by mixing equal volumes of reaction mixture and 2x RNA Gel loading dye (Thermo Fisher Scientific). Samples were heated to 70°C for five minutes and then slowly cooled down back to room temperature. This allowed bound proteins to denature and dissociate from RNAs, as well as the complementary RNA strands to denature and anneal.

Results

Purified MS2 replicase is functional in recombinant *in vitro* translation systems

To characterise the activity requirements of the MS2 replicase complex, we overexpressed the MS2rep subunit in *E. coli* and purified it by single-step immobilized metal ion affinity chromatography (IMAC). In the one-step purification protocol, MS2rep co-eluted with the two potential host factors ribosomal protein

S1 and the translation factor (TF) EF-Ts, both of which are components of the related Q β replicase complex (Supplementary Fig. 1). Surprisingly, we noticed that EF-Tu, which is an essential and tightly-binding host factor for the Q β replicase complex that readily co-purifies with Q β rep, did not co-elute with the His-tagged MS2rep subunit. Thus, the assembly properties of the replicase core complexes between the genera *Allolevivirus* (Q β) and *Levivirus* (MS2) show unexpected differences.

Next, we sought to explore if the MS2·S1·EF-Ts heterocomplex could be used to initiate transcription of a genuine MS2 template. To this end, we made use of our previously established MS2 RNA polymerase assay for the detection of MS2 replicase activity in recombinant *in vitro* transcription translation (PURE) systems (20, 28). In this assay, a fluorescence readout is produced by (+) strand synthesis of the broccoli aptamer from a (-) strand template that is fused with the 3'-end of genomic MS2 (-) strand (F30-Bro(-)), in the presence of the fluorogen DFHBI-1T (29) (Fig. 2A). In agreement with literature reports for *in situ* expressed MS2rep (20), we observed a strong fluorogenic readout when both F30-Bro(-) and MS2·S1·EF-Ts were incubated in the commercially available PURExpress® system. We also observed F30-Bro(+) synthesis, albeit at a lower level, in a homemade PURE system (PURE 3.0), which was prepared according to a previously established protocol (27) (Fig. 2B). Based on this result, we inferred that the PURE systems contained the host-factors required to assemble the active MS2 holoenzymes.

To further narrow down the range of possible *E. coli* proteins that are required for MS2 replicase activity, we investigated the activity in the presence of different PURE protein fractions. In PURE 3.0, 30 of the 31 *E. coli* TFs are obtained after co-expression and purification of the TF genes from three large expression plasmids resulting in three protein fractions (LD1, LD2 and LD3) (27). An additional enzyme mix contains 70S ribosomes as well as the elongation factor EF-Tu and the enzymes necessary to reconstitute a NTP regeneration system based on creatine phosphate (Supplementary Table 4). Initially, we tested MS2 replicase activity in reduced PURE 3.0 reactions (PUREred) based on the LD1-LD3 protein fractions. We also used a simplified enzyme mix containing only ribosomal protein S1 and EF-Tu because we did not expect that ribosomes or kinase enzymes to be the missing co-factors based on the genetic similarities between MS2 and Q β (30). As anticipated, MS2 showed transcription of F30-Bro(+) in the PUREred environment (Fig. 2C). Next, we sought to narrow down the range of potential host factors by further omitting protein components from the PUREred setup. Hereby, we confirmed that TFs EF-Tu and the ribosomal protein S1 are critical co-factors required for full MS2 replicase activity, similarly to Q β , as depletion of both proteins would lead to a drastic loss in F30-Bro(+) synthesis (Fig. 2C). While depletion of EF-Tu had a drastic effect on F30-Bro(+) synthesis, we still observed significant transcription in the absence of added S1 protein presumably due to the presence of S1 protein in the purified complex (Supplementary Fig. 1).

In addition to the expected dependency of the replicase on S1 and EF-Tu, we also observed an unforeseen impact on F30-Bro(+) synthesis upon depletion of the individual LD protein fractions. While depletion of both LD1 and LD3 appeared to weakly stimulate MS2rep activity, omission of LD2 caused a drastic loss of F30-Bro(+) transcription (Fig. 2C). LD2 contains eight enzymes, four tRNA synthetases (AlaRS, AsnRS, IleRS, PheRS1 + 2), *E. coli* Methionyl-tRNA formyltransferase (MTF), the translation elongation factor EF-

Ts (which co-purifies with the MS2rep subunit), and the two translation initiation factors IF1 and IF3 (27). To identify TFs responsible for this marked effect on replicase activity, we performed selective depletion experiments with all LD2 proteins (Fig. 3A). Omitting the added EF-Ts led to a complete loss of activity, revealing that the excess amount of EF-Ts in the LD2 fractions is essential for transcription of the unnatural F30-Bro(-) template. We further observed an unexpected alteration of transcription activity by omitting the initiation factors IF1 and IF3. While depletion of IF1 caused a ~ 50% reduction in F30-Bro(+) synthesis, the omission of IF3 led to a more than 400% increase in transcription activity compared to the positive control reaction containing the full 1x LD2 protein fraction. These findings indicated that IF1 stimulates MS2 replicase activity, whereas IF3 acts as an inhibitor. This hypothesis was further corroborated in additional experiments, in which an excess of 5 μ M of each of the eight individually purified LD2 proteins was added to the reaction mixture (Fig. 3B). Whereas excess IF1 led to an increased transcription activity, adding an excess of IF3 completely abolished F30-Bro(+) synthesis. In contrast, the omission or supplementation of the four tRNA synthetases had no impact on MS2 replicase activity. Both the inhibitory effect of IF3 and stimulating effect of IF1 showed clear dose-dependencies with observable effects already at low micromolar concentrations (Fig. 3C), supporting the notion that they are based on a direct functional interaction with the MS2 replicase core complex. In contrast, control experiments using MTF or PEG8000 at increasing concentrations showed that neither non-specific protein-protein interactions nor excluded volume effects are responsible for the observed modulation of replicase activity induced by IF1 and IF3. Furthermore, we found evidence that inhibition by IF3 is based on a direct competition between IF3 and the replicase for RNA binding as only an excess of the MS2rep complex could rescue transcriptional activity (Supplementary Fig. 2).

IF1 stimulates synthesis of the full-length MS2 genome

While the F30-Bro(+) synthesis enables monitoring (+) strand synthesis from an artificial (-) template, it provides no information on the complete replication cycle of the natural ~ 3600 nt MS2 genome. To probe full-length genome replication by the *in vitro* reconstituted MS2 replicase complex, we integrated the broccoli aptamer into the (+)-strand of MS2 wild type (MS2wt) genome at an amenable site downstream of the open reading frame for the maturation protein (MS2(+)-Bro(+)) (Fig. 4A), where its insertion should only minimally interfere with replication (31–33). Using this construct, we were able to observe a continuous increase in DFHBI-1T fluorescence when the MS2rep·S1·EF-Ts complex was incubated in PURExpress system, which suggests processive genome replication (Fig. 4B). Comparison with reference inputs of MS2(+)-Bro(+) showed an estimated sixfold amplification, corresponding to an increase of MS2(+)-Bro(+) from 15 nM to approximately 90 nM over a 6-hour time course.

Having shown that the MS2 replicase complex can replicate genomic MS2 RNA in the PURExpress system, we further sought to dissect the influence of the individual co-factors on the ability of the replicase to synthesise the genomic (+) and (-) strands (Fig. 5A, B). Synthesis of genomic (+) and (-) strands from the corresponding MS2-Bro(-) template produced only a weak fluorescence output compared to synthesis of F30-Bro(+) from F30-Bro(-) template (Fig. 5C, D). Notably, the omission of an excess of S1 in these experiments did not significantly affect genome synthesis unlike for the shorter

unnatural F30-bro construct used previously (Fig. 2). This finding indicates that the bound S1 present in the purified complex is sufficient for effective replication of the natural replicase substrate.

The fluorescence output of the broccoli aptamer domain during genomic (-) strand synthesis was seemingly not affected by supplementation of IF1 (Fig. 5E). At the same time, however, an in-gel fluorescence analysis revealed a drastic increase in RNA-synthesis in the presence of the co-factor, with the majority of product migrating at the expected size of a full-length duplex (Supplementary Fig. 3). This suggests that only very little newly synthesised (-) strand RNA was present as single strand. In contrast, a stimulation of both overall RNA synthesis as well as broccoli fluorescence was observed during synthesis of MS2(+)Bro(+) from MS2(-)Bro(+) template in the presence of IF1 (Supplementary Fig. 3, Fig. 5F). Thus, IF1 stimulated MS2 replicase activity independent from the polarity of the template. In this minimal *in vitro* environment, the protein either caused a direct reduction of non-fluorescent inert duplex product during genomic (+) strand synthesis, or enhanced folding of the aptamer reporter domain.

Spontaneous formation of replicable RNA species

The purified replicase complex of phage Q β is well known for its spontaneous *in vitro* synthesis of rapidly amplifying RNA species of different length and nucleotide sequence, even in the absence of externally added template molecules (34–36). To test if MS2rep is capable of a similar spontaneous generation of short amplifiable RNA species, we compared the activity of both purified Q β and reconstituted MS2 core complex in template-free reactions supplemented with NTPs and SYBR Green nucleic acid stain. As expected, we observed a rapid increase in fluorescence after a brief lag phase of ~ 5–10 min in the presence of the Q β heterotetramer (Fig. 6A), suggestive of the rapid formation of small amplifying RNA species (“RNA parasites”) described in previous studies (36). We verified the efficient formation of small replicable RNAs by the Q β replicase by gel electrophoresis (Supplementary Fig. 4). Surprisingly, we observed no such spontaneous formation of RNAs when the MS2 replicase complex in the presence of IF1 was incubated for 75 min under the same conditions (Fig. 6A). As the MS2 enzyme did not show a similar strong background activity as the Q β core complex, we asked whether the enzyme was able to produce short amplifying RNAs as by-products during MS2 genome replication. To test this hypothesis, we performed serial transfer experiments with the MS2 replicase core complex (MS2rep·S1·EF-Ts·EF-Tu) in the absence and presence of IF1 and MS2(+) RNA (Fig. 6B). In reactions with the full-length genome, we observed a rapid degeneration of the ~ 3600 nt RNA molecule during the first two dilutions concomitant with the emergence of smaller RNA species with a broad size distribution and a dominant RNA band migrating at ~ 200 nt. The emergence of aberrant RNA products was strongly increased when the MS2 core complex was further supplemented with IF1. Intriguingly, in the presence of IF1, the small RNA species emerged even in the absence of input MS2(+) after the first serial transfer (Fig. 6B). To obtain more information about the sequence properties of the newly evolved RNA replicators, we reverse transcribed, sequenced, and analysed the reaction products. Notably, we obtained only a single clonal sequence from these experiments (MSRP-22), which showed an almost perfect homology with the first 118 nt of the 5'UTR and 105 nt of the 3' UTR of MS2wt (Fig. 7A, B, Supplementary Table 5). We confirmed that MSRP-22 is a genuine RNA template for MS2 replicase since it was specifically amplified in an input concentrations-dependent manner in batch reactions (Fig. 7C). In contrast, MS2rep was not able to

amplify RQ135, a typical RNA parasite of Q β rep (37) within the same time window. This finding confirmed the differential template requirements of both phage replicases.

Discussion

Previous attempts to characterise leviviral replicases *in vitro* were hampered by the general lability of the enzymes and the dependency of replicase activity on an unknown, yet easily dissociable, host factor (25, 26). Using the recombinant heterotrimeric MS2rep·S1·EF-Ts complex as a starting point, we elucidated the enigmatic properties of this archetypical class of RNA polymerase enzymes. Compared to the stable Q β rep·EF-Ts·EF-Tu·S1 heterotetramer, which is capable of replicating a number of non-genomic RNAs as well as genomic Q β (+) RNA *in vitro* (38), the MS2 replicase complex shows a number of different characteristics. Firstly, EF-Tu, which is an integral part of the active complex for both replicases, only weakly binds to the stable MS2rep·S1·EF-Ts core trimer. Secondly, the previously unknown host factor required for full activity was identified to be the translation initiation factor IF1. IF1 is a S1 domain protein that contains an oligomer binding (OB) fold found in a variety of RNA chaperones (39). Indeed, it has been reported that IF1 acts as a transcription anti-termination factor, which can destabilize strong secondary structure elements and thereby facilitating RNA polymerase read-through (40). Furthermore, IF1 showed RNA chaperone activity during trans-splicing assays both *in vivo* and *in vitro* (41). These properties suggest that IF1 may support MS2 genome replication by destabilizing secondary structure elements of single-stranded templates and exposing terminal and internal binding sites for the MS2 replicase and / or through facilitating replication initiation and product release. The influence of IF1 on RNA synthesis resembles the role of the OB-proteins S1 and Hfq for (-) strand synthesis during the the Q β replication cycle. Here, Hfq facilitates the access of replicase to the 3'-end of the genomic (+) strand (42, 43), while S1 appears to contribute to both termination of replication and re-initiation after product release (44). Thus, a similar yet MS2-specific role of IF1 for the remodelling of MS2 templates seems plausible.

The strong inhibitory effect of IF3 on MS replicase activity seems counterintuitive at first and rather indicative of an *in vitro* artefact. However, early studies reported a specific interaction of IF3 with the 3'-terminus of MS2 (45–47). Our results suggest a direct competition for the 3'-end between the replicase and IF3 because increasing concentrations of the replicase compensated for the inhibitory effect of IF3. Notably, MS2 gene expression was also shown to be completely independent from IF3, unlike for *E. coli* host proteins (48). Thus, replication and translation of MS2 genomes seem to be well adapted to conditions under which IF3 levels are minimal, such as it was reported for *E. coli* cells that have reached stationary phase (49). While phage virions are still produced under these conditions, protein synthesis rates are usually insufficient to trigger cell lysis and phage release (50). Therefore, we speculate that the ability of the MS2 replication system to continuously produce infectious but host-contained virions in slowly growing cells may enable phage persistence within starved bacterial populations. This state may last until growth conditions and therefore a lytic reproduction cycle can be restored. Inhibition of genome replication by IF-3 might also play a role in the well-orchestrated infection cycle of MS2 and requires further *in vivo* studies.

While the isolated MS2 replicase holocomplex synthesizes genomic RNA from both (+) and (-) strand templates in batch reactions, both products readily anneal to form double-stranded RNA, which can no longer be used as a template for replication (51, 52). In contrast, replication of the full-length genome in PURE systems enabled sustained synthesis of the genomic (+) strand. Under these more *in vivo*-like conditions, ribosome binding and translation to the MS2(+) RNA might counteract a direct annealing with newly synthesised (-) strand thereby supporting continuous replication similar to that described for Q β (53).

Q β replicase is notorious for the spontaneous generation of short, exponentially amplifying “RNA parasites” from trace amounts of contaminating RNA even in single batch reactions (34, 36). In contrast, the reconstituted MS2 replicase core complex required additional serial transfers before replicative RNA species were enriched in significant amounts. While the emerging RNA population was dominated by a small sequence migrating at \sim 300 nt in agarose gels, it retained an overall broad size distribution even after five serial transfers. The lack of convergence to a discrete number of small, dominant replicators (such as for Q β replicase, Supplementary Fig. 4) implies either that initiation, rather than elongation is critical for sequence replication, or that longer sequences are generated by mechanisms such as template switching or non-templated terminal transferase activity (54).

The generally lower tendency of MS2 replicase to rapidly generate and amplify non-genomic RNA species in comparison with Q β replicase remains unclear. One explanation is an overall lower *in vitro* activity and / or lifetime of the holoenzyme similar to pioneering reports that used enriched protein fractions for activity assays (55). Alternatively, MS2 replicase might be less promiscuous than Q β replicase towards non-genomic templates and therefore less prone to generate short parasitic sequences. In agreement with this conjecture, the clonal sequence derived from the serial transfer experiment consisted exclusively of major parts of the 5' and 3'UTR of the wild-type genome without additional sequence elements, suggesting that these motifs form the core of replicable units. Thus, both UTR-elements may enable the design of replicable mRNA species, which could be used for dynamic *in vitro* evolution studies (18, 19, 34, 35) or the generation of self-amplifying mRNAs under *in vivo* conditions (56).

Declarations

ACKNOWLEDGEMENT

We thank K. Libicher for help with preparing the pLD protein fractions. We thank the MPIB mass spectrometry core facility for the help with protein identification and analysis. We thank E. Y. Song for critically proofreading the manuscript. HM is grateful for funding by the European Research Council (ERC starting grant, RiboLife) under 802000 and the MaxSynBio consortium, which is jointly funded by the Federal Ministry of Education and Research of Germany and the Max Planck Society.

CONFLICT OF INTEREST

The authors declare no competing interests.

References

1. Fiers,W., Contreras,R., Duerinck,F., Haegeman,G., Iserentant,D., Merregaert,J., Min Jou,W., Molemans,F., Raeymaekers,A., Van den Berghe,A., *et al.* (1976) Complete nucleotide sequence of bacteriophage MS2 RNA: Primary and secondary structure of the replicase gene. *Nature*, **260**, 500–507.
2. Koning,R.I., Gomez-Blanco,J., Akopjana,I., Vargas,J., Kazaks,A., Tars,K., Carazo,J.M. and Koster,A.J. (2016) Asymmetric cryo-EM reconstruction of phage MS2 reveals genome structure in situ. *Nat. Commun.*, **7**, 1–6.
3. Kastelein,R.A., Berkhout,B., Overbeek,G.P. and van Duin,J. (1983) Effect of the sequences upstream from the ribosome-binding site on the yield of protein from the cloned gene for phage MS2 coat protein. *Gene*, **23**, 245–254.
4. Kozak,M. and Nathans,D. (1972) Translation of the genome of a ribonucleic acid bacteriophage. *Bacteriol. Rev.*, **36**, 109–134.
5. Stock-Ley,P.G., Stonehouse,N.J. and Valegård,K. (1994) Molecular mechanism of RNA phage morphogenesis. *Int. J. Biochem.*, **26**, 1249–1260.
6. Tars,K. (2020) ssRNA phages: Life cycle, structure and applications. In *Biocommunication of Phages*. Springer International Publishing, pp. 261–292.
7. Dawson,D.J., Paish,A., Staffell,L.M., Seymour,I.J. and Appleton,H. (2005) Survival of viruses on fresh produce, using MS2 as a surrogate for norovirus. *J. Appl. Microbiol.*, **98**, 203–209.
8. Mikel,P., Vasickova,P., Tesarik,R., Malenovska,H., Kulich,P., Vesely,T. and Kralik,P. (2016) Preparation of MS2 phage-like particles and their use as potential process control viruses for detection and quantification of enteric RNA viruses in different matrices. *Front. Microbiol.*, **7**, 1911.
9. Miranda,J.A. and Steward,G.F. (2017) Variables influencing the efficiency and interpretation of reverse transcription quantitative PCR (RT-qPCR): An empirical study using Bacteriophage MS2. *J. Virol. Methods*, **241**, 1–10.
10. Bertrand,E., Chartrand,P., Schaefer,M., Shenoy,S.M., Singer,R.H. and Long,R.M. (1998) Localization of ASH1 mRNA particles in living yeast. *Mol. Cell*, **2**, 437–445.
11. Golding,I., Paulsson,J., Zawilski,S.M. and Cox,E.C. (2005) Real-time kinetics of gene activity in individual bacteria. *Cell*, **123**, 1025–1036.
12. Chubb,J.R., Trcek,T., Shenoy,S.M. and Singer,R.H. (2006) Transcriptional Pulsing of a Developmental Gene. *Curr. Biol.*, **16**, 1018–1025.
13. Pasloske,B.L., Walkerpeach,C.R., Obermoeller,R.D., Winkler,M. and Dubois,D.B. (1998) Armored RNA technology for production of ribonuclease-resistant viral RNA controls and standards. *J. Clin. Microbiol.*, **36**, 3590–3594.
14. Ashley,C.E., Carnes,E.C., Phillips,G.K., Durfee,P.N., Buley,M.D., Lino,C.A., Padilla,D.P., Phillips,B., Carter,M.B., Willman,C.L., *et al.* (2011) Cell-specific delivery of diverse cargos by bacteriophage MS2 virus-like particles. *ACS Nano*, **5**, 5729–5745.

15. Pan,Y, Zhang,Y, Jia,T, Zhang,K, Li,J. and Wang,L. (2012) Development of a microRNA delivery system based on bacteriophage MS2 virus-like particles. *FEBS J.*, **279**, 1198–1208.
16. Spiegelman,S., Haruna,I., Holland,I.B., Beaudreau,G. and Mills,D. (1965) The synthesis of a self-propagating and infectious nucleic acid with a purified enzyme. *Proc. Natl. Acad. Sci. U. S. A.*, **54**, 919–27.
17. Kita,H., Matsuura,T., Sunami,T., Hosoda,K., Ichihashi,N., Tsukada,K., Urabe,I. and Yomo,T. (2008) Replication of genetic information with self-encoded replicase in liposomes. *Chembiochem*, **9**, 2403–10.
18. Ichihashi,N., Usui,K., Kazuta,Y., Sunami,T., Matsuura,T. and Yomo,T. (2013) Darwinian evolution in a translation-coupled RNA replication system within a cell-like compartment. *Nat. Commun.*, **4**, 2494.
19. Mizuuchi,R. and Ichihashi,N. (2018) Sustainable replication and coevolution of cooperative RNAs in an artificial cell-like system. *Nat. Ecol. Evol.*, **2**, 1654–1660.
20. Weise,L.I., Heymann,M., Mayr,V. and Mutschler,H. (2019) Cell-free expression of RNA encoded genes using MS2 replicase. *Nucleic Acids Res.*, **47**, 10956–10967.
21. Peabody,D.S. (1997) Role of the coat protein-RNA interaction in the life cycle of bacteriophage MS2. *Mol. Gen. Genet.*, **254**, 358–364.
22. Licis,N., van Duin,J., Balklava,Z. and Berzins,V. (1998) Long-range translational coupling in single-stranded RNA bacteriophages: an evolutionary analysis. *Nucleic Acids Res.*, **26**, 3242–3246.
23. Groeneveld,H., Thimon,K. and Van Duin,J. (1995) Translational control of maturation-protein synthesis in phage MS2: A role for the kinetics of RNA folding? *RNA*, **1**, 79–88.
24. Černý,J., Černá Bolfíková,B., Valdés,J.J., Grubhoffer,L. and Růžek,D. (2014) Evolution of tertiary structure of viral RNA dependent polymerases. *PLoS One*, **9**, e96070.
25. Fedoroff,N. V. and Zinder,N.D. (1973) Factor requirement of the bacteriophage f2 Replicase. *Nat. New Biol.*, **241**, 105–108.
26. Yonesaki,T., Furuse,K., Haruna,I. and Watanabe,I. (1982) Relationships among four groups of RNA coliphages based on the template specificity of GA replicase. *Virology*, **116**, 379–381.
27. Shepherd,T.R., Du,L., Liljeruhm,J., Samudyata, Wang,J., Sjödin,M.O.D., Wetterhall,M., Yomo,T. and Forster,A.C. (2017) De novodesign and synthesis of a 30-cistron translation-factor module. *Nucleic Acids Res.*, **45**, 10895–10905.
28. Shimizu,Y., Inoue,A., Tomari,Y., Suzuki,T., Yokogawa,T., Nishikawa,K. and Ueda,T. (2001) Cell-free translation reconstituted with purified components. *Nat. Biotechnol.*, **19**, 751–755.
29. Song,W., Strack,R.L., Svensen,N. and Jaffrey,S.R. (2014) Plug-and-play fluorophores extend the spectral properties of spinach. *J. Am. Chem. Soc.*, **136**, 1198–1201.
30. Kidmose,R.T., Vasiliev,N.N., Chetverin,A.B., Andersen,G.R. and Knudsen,C.R. (2010) Structure of the Q β replicase, an RNA-dependent RNA polymerase consisting of viral and host proteins. *Proc. Natl. Acad. Sci. U. S. A.*, **107**, 10884–10889.

31. Olsthoorn,R.C.L. (2014) A recombinant RNA bacteriophage system to identify functionally important nucleotides in a self-cleaving ribozyme. *Viol. J.*, **11**, 2–5.
32. Klovins,J., Van Duin,J. and Olsthoorn,R.C.L. (1997) Rescue of the RNA phage genome from RNase III cleavage. *Nucleic Acids Res.*, **25**, 4201–4208.
33. Van Meerten,D., Zelwer,M., Régnier,P. and Van Duin,J. (1999) In vivo oligo(A) insertions in phage MS2: Role of Escherichia coli poly(A) polymerase. *Nucleic Acids Res.*, **27**, 3891–3898.
34. Furubayashi,T., Ueda,K., Bansho,Y., Motooka,D., Nakamura,S., Mizuuchi,R. and Ichihashi,N. (2020) Emergence and diversification of a host-parasite RNA ecosystem through Darwinian evolution. *Elife*, **9**, 1–15.
35. Bansho,Y., Furubayashi,T., Ichihashi,N. and Yomo,T. (2016) Host-parasite oscillation dynamics and evolution in a compartmentalized RNA replication system. *Proc. Natl. Acad. Sci. U. S. A.*, **113**, 4045–4050.
36. Chetverin,A.B., Chetverina,H. V. and Munishkin,A. V. (1991) On the nature of spontaneous RNA synthesis by Q β replicase. *J. Mol. Biol.*, **222**, 3–9.
37. AV,M., LA,V., VI,U., LA,B., HV,C. and AB,C. (1991) Efficient templates for Q beta replicase are formed by recombination from heterologous sequences. *J. Mol. Biol.*, **221**.
38. Vasilyev,N.N., Kutlubaeva,Z.S., Ugarov,V.I., Chetverina,H. V. and Chetverin,A.B. (2013) Ribosomal protein S1 functions as a termination factor in RNA synthesis by Q β phage replicase. *Nat. Commun.*, **4**, 1781.
39. Rajkowitsch,L., Chen,D., Stampfl,S., Semrad,K., Waldsich,C., Mayer,O., Jantsch,M.F., Konrat,R., Bläsi,U. and Schroeder,R. (2007) RNA chaperones, RNA annealers and RNA helicases. *RNA Biol.*, **4**, 118–130.
40. Phadtare,S., Kazakov,T., Bubunenko,M., Court,D.L., Pestova,T. and Severinov,K. (2007) Transcription antitermination by translation initiation factor IF1. *J. Bacteriol.*, **189**, 4087–4093.
41. Croitoru,V., Semrad,K., Prenninger,S., Rajkowitsch,L., Vejen,M., Laursen,B.S., Sperling-Petersen,H.U. and Isaksson,L.A. (2006) RNA chaperone activity of translation initiation factor IF1. *Biochimie*, **88**, 1875–1882.
42. Schuppli,D., Georgijevic,J. and Weber,H. (2000) Synergism of mutations in bacteriophage Q β RNA affecting host factor dependence of Q β replicase. *J. Mol. Biol.*, **295**, 149–154.
43. Schuppli,D., Miranda,G., Tsui,H.C.T., Winkler,M.E., Sogo,J.M. and Weber,H. (1997) Altered 3'-terminal RNA structure in phage Q β adapted to host factor- less Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.*, **94**, 10239–10242.
44. Kamen,R., Kondo,M., Romer,W. and Weissmann,C. (1972) Reconstitution of Qbeta Replicase Lacking Subunit alpha with Protein-Synthesis-Interference Factor i. *Eur. J. Biochem.*, **31**, 44–51.
45. Johnson,B. and Szekely,M. (1977) Specific binding site of E. coli initiation factor 3 (IF3) at a 3'-terminal region of MS2 RNA. *Nature*, **267**, 550–552.
46. Fowler,M.J.F. and Szekely,M. (1979) The specific interaction of Escherichia coli initiation factor IF3 with coliphage-MS2 ribonucleic acid. *Biochem. Soc. Trans.*, **7**, 979–980.

47. Johnson,B. and Szekely,M. (1979) The Binding Site of IF-3 of MS2 RNA. *Methods Enzymol.*, **60**, 343–350.
48. Zipori,P, Bosch,L. and van Duin,J. (1978) Translation of MS2 RNA in vitro in the Absence of Initiation Factor IF-3. *Eur. J. Biochem.*, **92**, 235–241.
49. Scheps,R. and Revel,M. (1972) Deficiency in Initiation Factors of Protein Synthesis in Stationary-Phase Escherichia coli. *Eur. J. Biochem.*, **29**, 319–325.
50. Ricciuti,C.P. (1972) Host-Virus Interactions in Escherichia coli: Effect of Stationary Phase on Viral Release from MS2-Infected Bacteria. *J. Virol.*, **10**, 162–165.
51. Weissmann,C. and Borst,P. (1963) Double-stranded ribonucleic acid formation in vitro by MS 2 phage-induced RNA synthetase. *Science (80-)*, **142**, 1188–1191.
52. Usui,K., Ichihashi,N. and Yomo,T. (2015) A design principle for a single-stranded RNA genome that replicates with less double-strand formation. *Nucleic Acids Res.*, **43**, 8033–8043.
53. Usui,K., Ichihashi,N., Kazuta,Y., Matsuura,T. and Yomo,T. (2014) Effects of ribosomes on the kinetics of Q β replication. *FEBS Lett.*, **588**, 117–123.
54. Biebricher,C.K. and Luce,R. (1992) In vitro recombination and terminal elongation of RNA by Q β replicase. *EMBO J.*, **11**, 5129–5135.
55. Weissmann,C., Simon,L. and Ochoa,S. (1963) Induction by an RNA phage of an enzyme catalyzing incorporation of ribonucleotides into ribonucleic acid. *Proc. Natl. Acad. Sci. U. S. A.*, **49**, 407–414.
56. Yao,Y., Zhang,W., Zhang,M., Jin,S., Guo,Y., Zu,Y., Ren,K., Wang,K., Chen,G., Lou,C., *et al.* (2019) A Direct RNA-to-RNA Replication System for Enhanced Gene Expression in Bacteria. *ACS Synth. Biol.*, **8**, 1067–1078.
57. Lorenz,R., Bernhart,S.H., Höner zu Siederdisen,C., Tafer,H., Flamm,C., Stadler,P.F. and Hofacker,I.L. (2011) ViennaRNA Package 2.0. *Algorithms Mol. Biol.*, **6**, 26.
58. Gruber,A.R., Lorenz,R., Bernhart,S.H., Neuböck,R. and Hofacker,I.L. (2008) The Vienna RNA websuite. *Nucleic Acids Res.*, **36**, W70–W74.
59. Waterhouse,A.M., Procter,J.B., Martin,D.M.A., Clamp,M. and Barton,G.J. (2009) Jalview Version 2 - a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, **25**, 1189–1191.

Figures

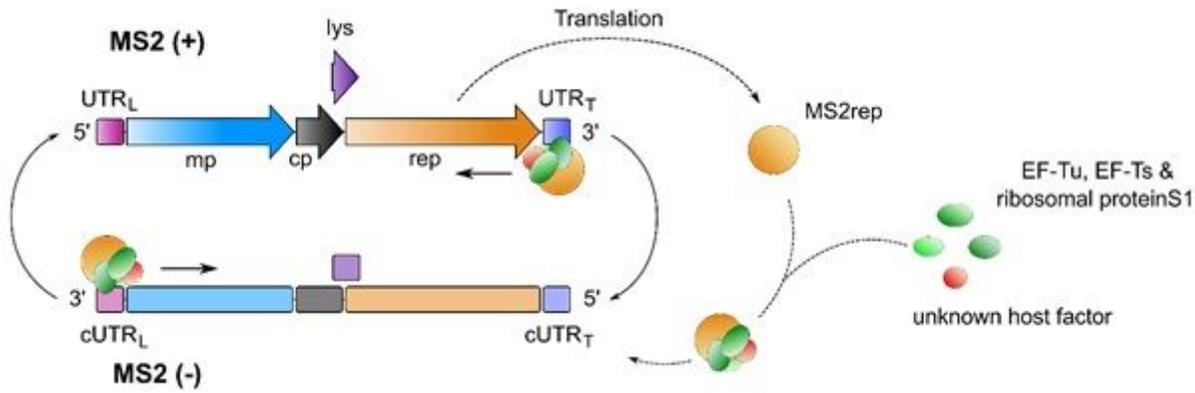


Figure 1

Scheme of replication for the ssRNA genome of MS2 bacteriophage. The 3569 nt genome encodes four proteins: the maturation protein (mp) for host infection, the coat protein (cp) for genome encapsulation, the lysis protein (lys) for cell lysis and the MS2replicase subunit (rep) for replication of the ssRNA genome. The four open reading frames on the plus strand (coloured arrows) are embedded between the leading untranslated region at the 5'-end (UTR_L) and the tailing UTR at the 3'-end (UTR_T). For the minus strand, the respective complementary sequences are depicted as rounded rectangles. Following translation of the MS2replicase subunit (MS2rep), formation of the holo-complex takes place. This complex is suspected to consist of host factors EF-Tu, EF-Ts, ribosomal protein S1, and a putative additional host factor. RNA replication is initiated at the 3'-ends of the respective template strands, leading to the synthesis of the complementary, new template strand.

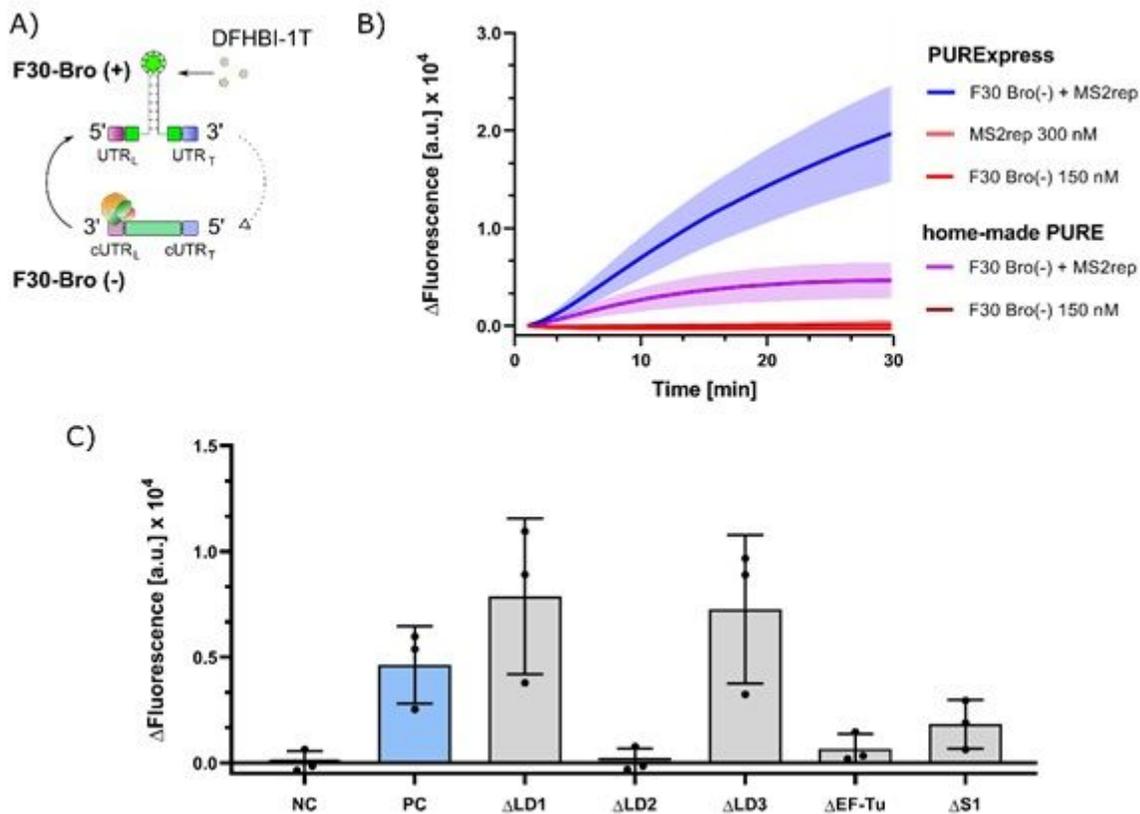


Figure 2

Effect of different PURE co-factors on F30-Bro(+) synthesis by MS2rep A) The broccoli aptamer (F30-Bro) for fluorescence readout is incorporated into a replication scaffold based on the MS2 UTRs. Using F30-Bro(-) as template, the MS2 replicase complex synthesises F30-Bro(+) and the broccoli aptamer therein binds fluorogenic DFHBI-1T. B) Replicase activity in commercially available PUREExpress as well as in a homemade PURE 3.0. Reactions were incubated at 37°C for 30 minutes and fluorescence was measured every minute. C) Endpoint fluorescence of depletion permutation assay for homemade PURE-based replication system, programmed with 50 nM F30-Bro(-) RNA. Reactions contained all components except MS2rep (NC), all components (PC), or all components except the indicated. Fluorescence was measured every minute over a 30-minute time course at 37°C. Error bars indicate standard deviation, based on three independent technical replicates.

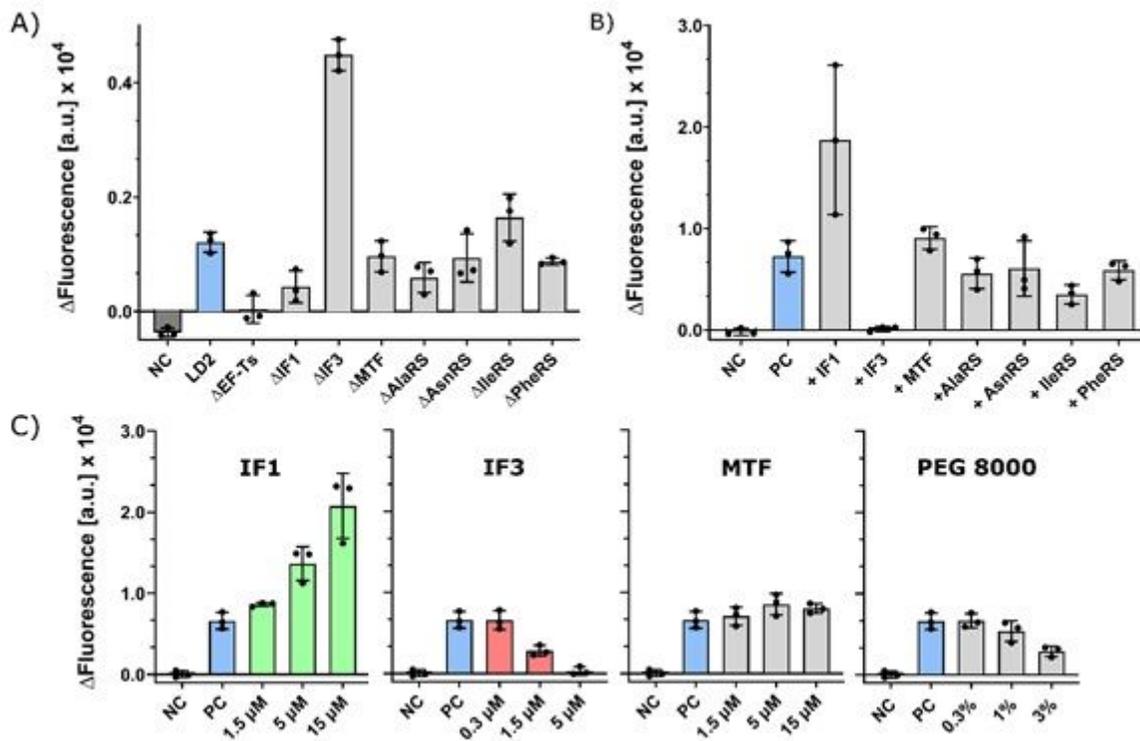


Figure 3

F30-Bro(+) transcription by MS2 replicase in the absence and presence of various PURE enzymes. A) DFHBI-1T fluorescence increase recorded after sample incubation at 37°C for one hour. Reactions contained 0.3 μM MS2rep, 50 nM F30-Bro(-) RNA, EF-Tu (15 μM), S1 (1.5 μM) and full LD2 (5 μM) except for the proteins indicated. Negative control (“NC”) contained no MS2rep while “LD2” contained the full set of LD2 proteins and 0.3 μM MS2rep. B) DFHBI-1T fluorescence as in A) but in the presence of MS2rep and individual LD2 proteins alone (5 μM). C) Co-factor titration assay. Reactions were carried out as in B) but with increasing concentrations of selected PURE proteins or PEG 8000 to test for excluded volume effects. Replication mixtures were programmed with 50 nM F30-Bro(-) RNA. Fluorescence was measured every minute over a 60-minute time course at 37°C. Error bars indicate standard deviation based on three independent replicates prepared using the same protein stock solutions. The negative values for NC in panel A resulted from a baseline drift during the experiment.

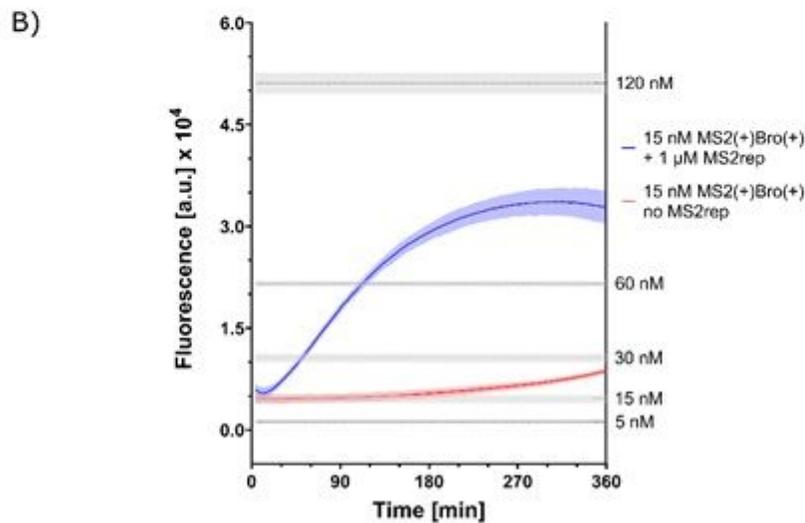
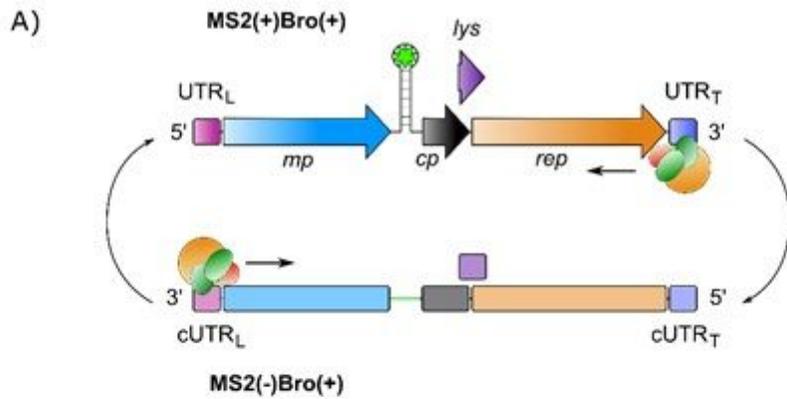


Figure 4

Full-length MS2 genome replication in PURExpress. A) Replication scheme of MS2-Bro(+) RNA by encoded MS2replicase subunit (MS2rep) in the commercial translational system PURExpress (NEB). B) Fluorescence change over the six-hour incubation at 37°C for replication analysis of 15 nM MS2(+)
Bro(+) in PURExpress with and without 1 μ M MS2rep. Grey lines indicate average fluorescence levels of reference concentrations of in vitro transcribed MS2(+)
Bro(+) in PURExpress. Standard deviations are from three independent replicates.

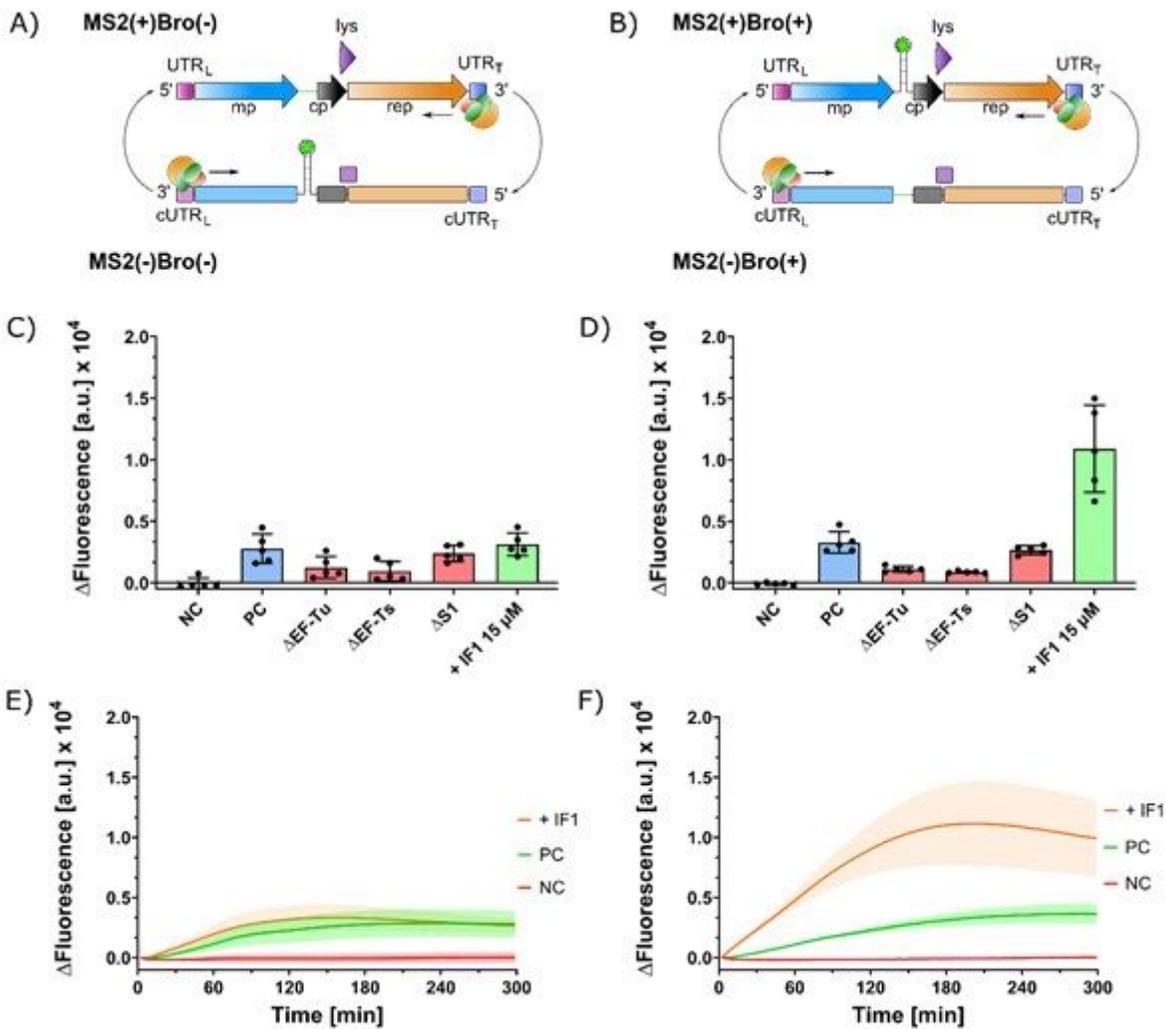


Figure 5

Replication of genomic (-) and (+) strand under isolated reaction conditions. Replication and fluorescence read-out scheme for A) MS2(-)Bro(-) or B) MS2(+)-Bro(+) synthesis. DFHBI-1T fluorescence of C) MS2(-)Bro(-) or D) MS2(+)-Bro(+) after 3 h incubation in reaction buffer supplemented with MS2rep (1 μ M) and in the presence or absence of EF-Tu and EF-Ts (15 μ M each), S1 (1.5 μ M), and IF (15 μ M). Positive controls (PC) contained all co-factors except IF1. Negative controls (NC) contained no MS2rep. Initial non-fluorescent template concentrations were 50 nM. E) and F) show the fluorescence time traces of both reactions in C) and D) recorded over five hours. Error bars indicate standard deviations calculated from independent replicates from the same protein stocks.

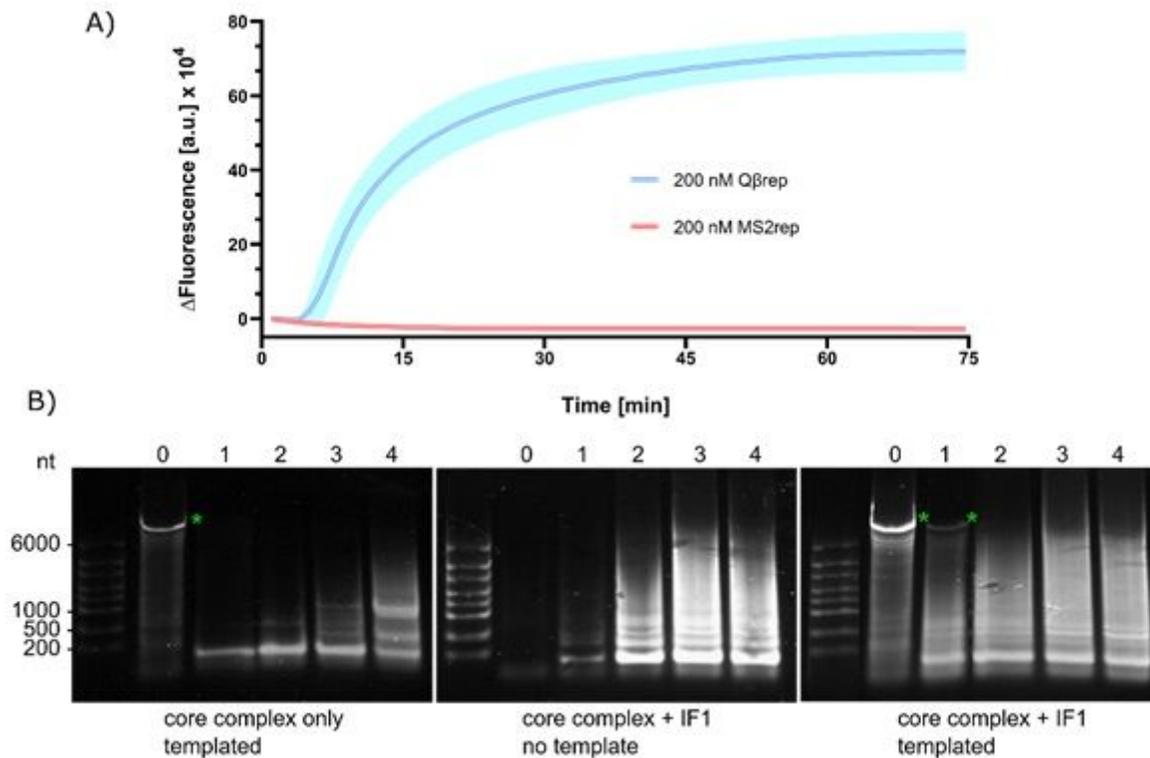


Figure 6

Generation of non-genomic replicable RNAs with recombinant Q β and MS2 replicase. A) Fluorescence time traces reactions lacking dedicated template for either reconstituted MS2 replicase (0.2 μ M MS2rep, 1.5 μ M S1, and 15 μ M of EF-Tu, EF-Ts, and IF1) or recombinant Q β replicase in the presence of 0.4 mM NTPs. Error bars indicate standard deviations from three independent technical replicates. B) Gel electrophoresis of samples from serial transfer experiments after 3 h of incubation at 37°C, run on a 1% TAE agarose gel. Lanes indicate number of transfers; asterisks indicate double stranded MS2wt RNA. Reactions contained EF-Tu and EF-Ts (15 μ M), S1 (1.5 μ M), MS2rep (1 μ M), and MS2wt plus strand RNA (50 nM) (left); no RNA and IF1 (15 μ M) (middle); or IF1 (15 μ M) and MS2wt plus strand RNA (50 nM) (right).

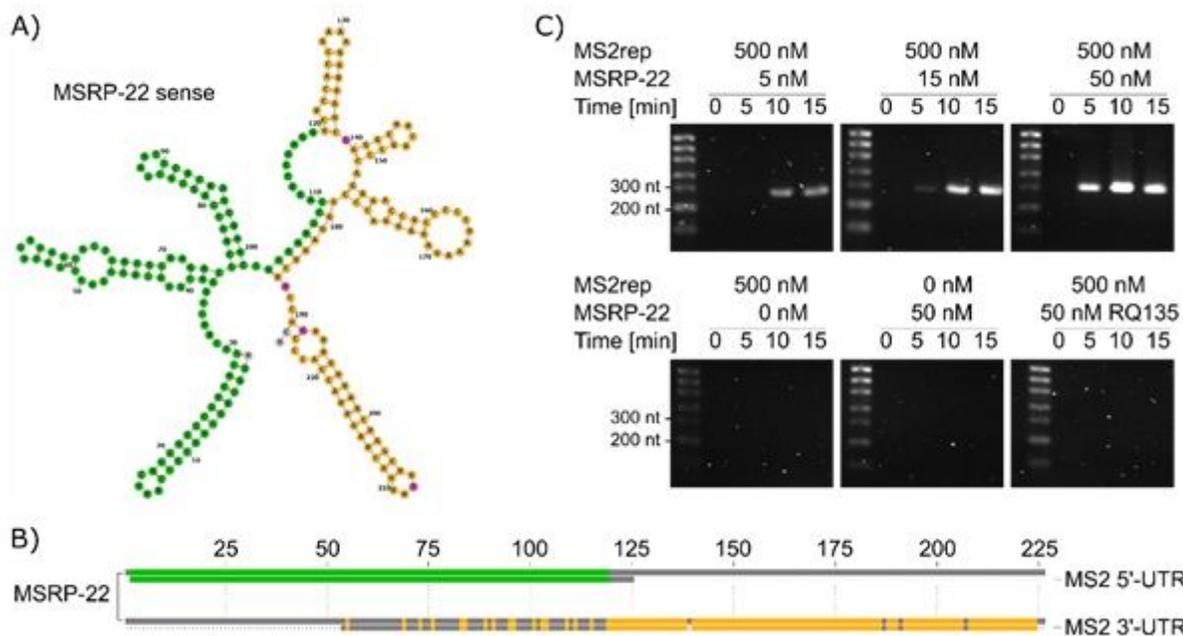


Figure 7

Properties of a selected small replicating RNA species. A) Minimum Free Energy structure of MSRP-22 was predicted by RNAfold (default parameters) (57). Colours indicate sequences aligning to the nucleotides 1 to 118 within the 5'-UTR (green) and nucleotides 3464 to 3568 within the 3'-UTR (yellow) of MS2wt, respectively. Nucleotides annotated in pink indicate mismatches and grey nucleotides could not be aligned to MS2wt (57, 58). Annotation as sense or antisense strand follows the orientation of the matching MS2 genomic or antigenomic strand, respectively. B) Sequence alignments of the isolated MS2 RNA parasite MSRP-22 with the 5'-UTR of MS2(+) (top, green) and 3'-UTR of MS2(+) (bottom, yellow), respectively. Alignments were made with Jalview (Version 2.11.1.3) (59). C) Time course of replication of MSRP-22, analysed by gel electrophoresis on a 2 % TAE agarose gel. Reaction mixes contained varying concentrations of input MSRP-22 anti RNA (5/ 15/ 50 nM) with 500 nM MS2rep core complex (top row). Controls contained either no input RNA, no MS2rep, or MS2rep with RQ135 RNA as non-specific RNA input (bottom row).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [WagneretalSI.docx](#)