Chemistry & Biology
Article

Importance of Parasite RNA Species Repression for Prolonged Translation-Coupled RNA Self-Replication

Yohsuke Bansho,[1] Norikazu Ichihashi,[2,4] Yasuaki Kazuta,[4] Tomoaki Matsuura,[2,3,4] Hiroaki Suzuki,[2,4] and Tetsuya Yomo[1,2,4,*]
[1]Graduate School of Frontier Biosciences
[2]Graduate School of Information Science and Technology
[3]Graduate School of Engineering
Osaka University, Osaka, Suita, Yamada-oka 1-5, 565-0871, Japan
[4]Exploratory Research for Advanced Technology, Japan Science and Technology Agency, Osaka, Suita, Yamada-oka 1-5, 565-0871, Japan
*Correspondence: yomo@ist.osaka-u.ac.jp
DOI 10.1016/j.chembiol.2012.01.019

SUMMARY

Increasingly complex reactions are being constructed by bottom-up approaches with the aim of developing an artificial cell. We have been engaged in the construction of a translation-coupled replication system of genetic information from RNA and a reconstituted translation system. Here a mathematical model was established to gain a quantitative understanding of the complex reaction network. The sensitivity analysis predicted that the limiting factor for the present replication reaction was the appearance of parasitic replicators. We then confirmed experimentally that repression of such parasitic replicators by compartmentalization of the reaction in water-in-oil emulsions improved the duration of self-replication. We also found that the main source of the parasite was genomic RNA, probably by nonhomologous recombination. This result provided experimental evidence for the importance of parasite repression for the development of long-lasting genome replication systems.

INTRODUCTION

There is a large gap between chemistry and biology. All living organisms are driven by various types of chemical reactions catalyzed by enzymes, but it is largely unknown how these reactions are organized into the "living state." To bridge this gap, many groups are attempting to construct biological functions in vitro from nonliving molecules, including nucleotide polymerization (Chakrabarti et al., 1994; Mansy et al., 2008), replication of genetic information (Guatelli et al., 1990; Lincoln and Joyce, 2009; Mills et al., 1967; Walker et al., 1992; Wright and Joyce, 1997), membrane growth (Hanczyc et al., 2003; Kuruma et al., 2009; Takakura et al., 2003; Walde et al., 1994), cell division (Zhu and Szostak, 2009), transport (Noireaux and Libchaber, 2004), protein translation (Shimizu et al., 2001), reactions in membrane vesicles (Kita et al., 2008; Murtas et al., 2007; Nomura et al., 2003; Oberholzer et al., 1995a, 1995b, 1999; Yu et al., 2001), and growth-replication coupling (Chen et al., 2004; Kurihara et al., 2011), and to finally integrate these functions into an artificial or minimal cell (Deamer, 2005; Forster and Church, 2006; Jewett and Forster, 2010; Luisi, 2002; Noireaux et al., 2011; Pohorille and Deamer, 2002; Stano, 2011; Szostak et al., 2001). Some of these studies have utilized simple compounds that presumably existed in the prebiotic world, such as fatty acids or nucleotides, to gain insight into the origin of life. Other studies have utilized biological molecules, such as RNA and protein, to understand the molecular organization required to achieve present-day biological functions. The latter type of approach is known as the semisynthetic approach (Chiarabelli et al., 2009; Luisi et al., 2006) or in vitro synthetic biology (Forster and Church, 2007; Simpson, 2006).

Here we focused on the replication of genetic information and its construction by the semisynthetic approach. A number of types of in vitro replication systems have been constructed, including isothermal replication systems for DNA or RNA using exogenously supplied enzymes (Guatelli et al., 1990; Mills et al., 1967; Walker et al., 1992; Wright and Joyce, 1997) and an RNA replication system without any protein enzymes using cross-reactive self-ligating ribozymes (Lincoln and Joyce, 2009). These results clearly demonstrated that replication of genetic information can be achieved by a relatively small number of proteins or ribozymes. One of the next important challenges is to develop these simple replication systems into more complex systems closer to those occurring in nature, which is expected to produce knowledge regarding how individual reactions are organized into more complex biological functions (Benner and Sismour, 2005; Ichihashi et al., 2010b; Pohorille and Deamer, 2002).

In natural replication systems, such as a cell or virus, replication of genetic information is coupled with translation of the information, where an information molecule (DNA or RNA) is replicated by the replication enzyme translated from the information encoded in itself. This translation-replication coupling allows an information molecule to utilize various types of proteins for its replication, and is therefore considered a prerequisite to achieve open-ended evolution, one of the characteristics of living organisms (Ruiz-Mirazo et al., 2004; Szathmáry and Maynard Smith, 1995).

Recently, we constructed a translation-coupled replication system from artificial genomic RNA encoding an RNA replicase

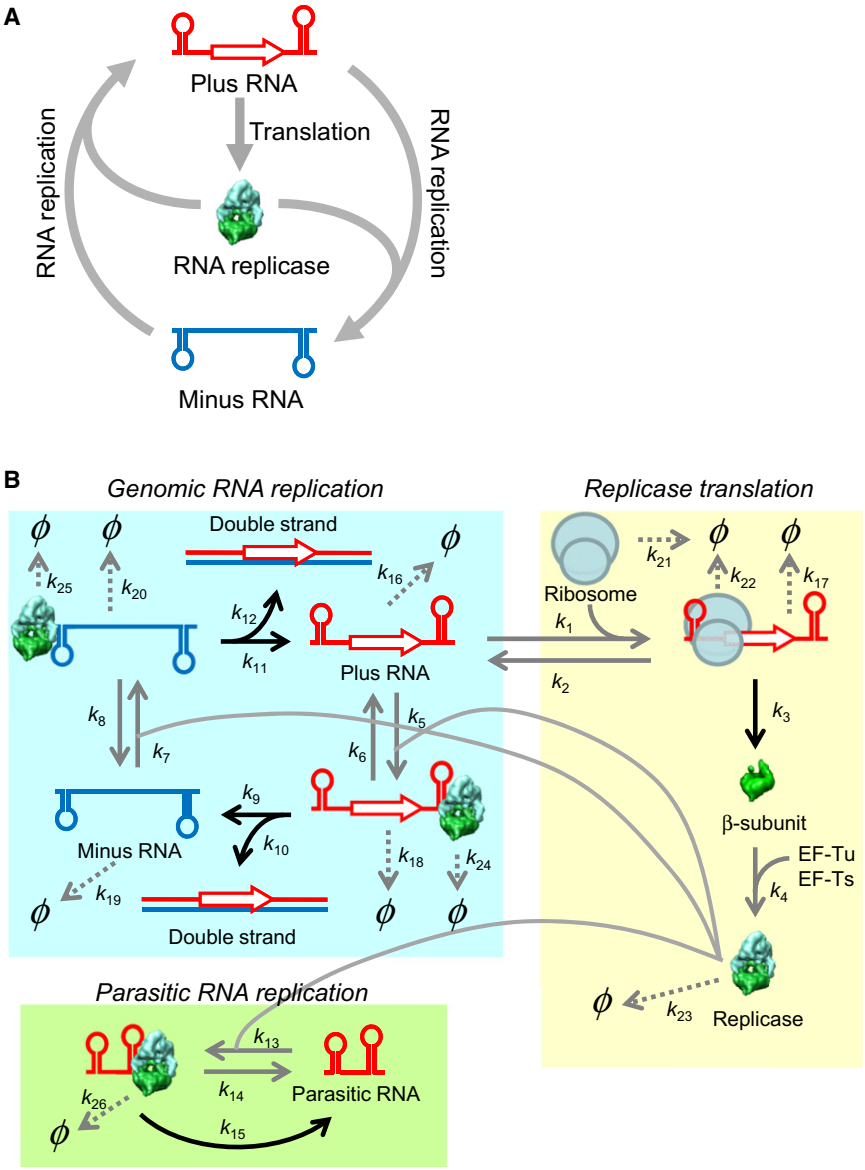

Figure 1. Schematic Drawing of the Translation-Coupled RNA Self-Replication System

(A) Simple scheme. The system contained genomic plus RNA and a modified cell-free translation system (PURE system). The plus RNA carried the gene encoding the coliphage Qβ catalytic subunit of RNA replicase (RNA-dependent RNA polymerase). The subunit is translated from genomic plus RNA and forms an active replicase with the other subunits, EF-Tu and Ts, which are components of the cell-free translation system. The heterotrimer replicase synthesizes the complementary strand, the minus RNA, using the plus RNA as a template. Then, using the minus RNA as a template, the replicase synthesizes complementary plus RNA.

(B) Detailed mathematical model. This reaction model is mainly composed of three parts: genomic RNA replication, replicase translation, and parasitic RNA replication. Black arrows represent the process in which RNA or replicase was synthesized. The gray arrows represent binding or dissociation processes. Dotted arrows and the symbol ϕ represent processes in which components are degraded or inactivated irreversibly. For simplicity, some of the pathways are not shown here, such as returning of the replicase from the RNA-replicase complex after complementary strand synthesis. In this model, the complementary strand, double-stranded RNA formation, and degradation of parasitic RNA were neglected. See also Table S1 and Figure S4.

and a reconstituted cell-free translation system of *Escherichia coli* (Kita et al., 2008). In this system, the RNA replication enzyme is translated from the genomic RNA and replicates the original genomic RNA (Figure 1A). This replication system has the same scheme as natural replicators such as RNA viruses, although the reaction efficiency, especially the duration of replication, was much lower than those seen in nature for an as yet unknown reason(s), suggesting that there still exist unknown conditions that must be satisfied to construct a long-lasting translation-coupled replication system.

Generally, the largest hurdle in the construction of in vitro systems with performance close to those in nature is their complexity, which makes their behavior unpredictable and thus their improvement difficult. This is the case in the translation-coupled replication system described above, which involves many components, such as genomic RNA, replicases, ribosomes, other translation proteins, tRNAs, NTPs, amino acids, and so forth, and these components participate in several types of reactions, including RNA synthesis, replicase translation, ATP reactivation, aminoacylation of tRNA, and so on. The concentrations of the components change dynamically over time; all of the components are consumed or degraded over time, whereas some components, such as replicases or genomic RNA, increase by translation or RNA replication, respectively. These dynamics cause the behavior of the system to be both nonlinear and unpredictable (Kita et al., 2008).

To overcome this unpredictability and to find conditions that must be satisfied to achieve long-lasting translation-coupled replication, we analyzed the replication system using a mathematical model that explained the system behavior quantitatively and predicted that the factor limiting the duration of replication is the appearance of parasitic replicators. Then, we experimentally demonstrated that encapsulation of the reaction into microscale compartments repressed the appearance of parasitic replicators and improved the duration of translation-coupled replication.

RESULTS

Modeling and Parameter Estimation

We previously constructed a translation-coupled replication system composed of artificial genomic RNA (plus RNA; 2,125

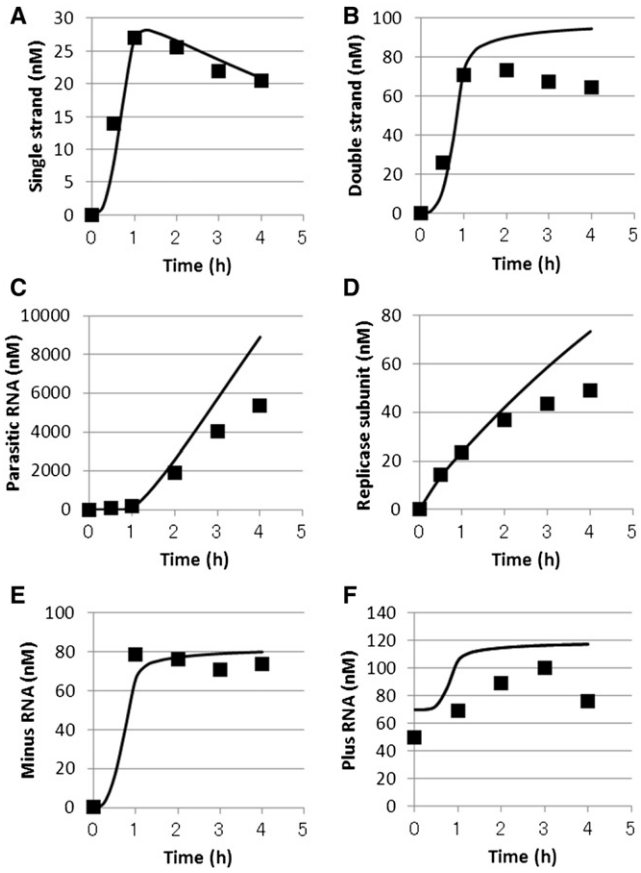

Figure 2. Simulation and Experimental Results of the Kinetics of Components in the Present Translation-Coupled RNA Self-Replication Reaction

The translation-coupled self-replication reaction was experimentally performed under standard conditions. The mixture of newly synthesized plus and minus RNA in single-stranded form (A) or double-stranded form (B) and parasitic RNA (C) was measured by [^{32}P]UTP incorporation, followed by agarose gel electrophoresis and autoradiography. The translated replicase β subunit (D) was measured by [^{35}S]methionine incorporation, followed by SDS-PAGE and autoradiography. More detailed methods of the radioisotope incorporation assay are described in Supplemental Information (Experiment 6). Minus RNA (E) and plus RNA (F) were measured by quantitative PCR after reverse transcription as described in Experimental Procedures. The squares show empirical data and the lines show the results of numerical simulation using the mathematical model and the parameters listed in Table S1.

bases) and a reconstituted cell-free translation system (Figure 1A) (Kita et al., 2008; Shimizu et al., 2001). This plus RNA encodes the catalytic β subunit of an RNA-dependent RNA polymerase (Qβ replicase), derived from coliphage Qβ, and has recognition sequences for the replicase at both termini. In this reaction, the subunit of the replicase is translated from genomic plus RNA and forms an active replicase complex with elongation factor EF-Tu and Ts contained in the translation system. The translated replicase then synthesizes the complementary minus RNA using the plus RNA as a template. As minus RNA can also be used as a template for the replicase, the replicase synthesizes plus RNA from the minus RNA. However, this replication reaction stopped by 1 hr, and the plus RNA showed less than 1.5-fold replication (Kita et al., 2008) (Figure 2F). Due to the complexity of the system, there are several possible explanations for the limited duration of the reaction, as described below.

Appearance of Parasitic RNA

We found that small (usually 200–300 bases) replicable RNAs appeared spontaneously in our translation-coupled replication system (Figure 2C). These small RNAs do not produce the replicase because they do not encode the enzyme. However, these RNAs can act as templates for replicase already present in the system and are replicated very rapidly due to their small size, resulting in competitive inhibition of genomic replication. According to these characteristics, the small RNAs can be regarded as parasites in this translation-coupled replication system.

Double-Stranded RNA Formation

Genomic plus and minus strands are active in their single-stranded forms. If they anneal and form double-stranded RNA, they become inert for both translation and replication (Nishihara et al., 1983). Therefore, if the double-stranded RNA accumulates at a sufficient rate, the translation-coupled replication reaction may stop. Indeed, we found that a substantial part of synthesized RNA became double stranded during the self-replication reaction (Figure 2B).

Inactivation of the Cell-free Translation System

The translation activity of the cell-free translation system decreases gradually during incubation at 37°C due to either inactivation of translational proteins or consumption of NTPs and amino acids (Shimizu et al., 2001).

RNA Degradation

Genomic plus and minus RNAs are degraded during incubation at 37°C, probably by RNase contamination in the translation system (Figure S4A available online). Although all the components of the system are purified to almost homogeneity, undetectable amounts of RNase may be sufficient for degradation of the RNA.

Inactivation of the Replicase

We found previously that the purified replicase is unstable at 37°C (Ichihashi et al., 2010a). This is also the case for the replicase translated in the cell-free translation system, where the half-life of the replicase activity was about 50 min (Figure S4B).

One or various combinations of these negative factors may shorten the duration of the translation-coupled self-replication reaction to different extents. To evaluate the effects of each factor quantitatively, we constructed a mathematical model including all the negative factors listed above (shown schematically in Figure 1B; details are presented in Supplemental Information). This model includes 11 components (plus RNA, minus RNA, replicase, ribosome, their complexes, etc.) and 26 reactions (ribosome binding to plus RNA, translation of replicase, replicase binding to plus and minus RNA, synthesis of RNA, etc.). We estimated the rate constants of all the reactions (k_1–k_{26}) included in the model by measuring the concentrations of the components under various experimental conditions (see Supplemental Experimental Procedures; parameter values are presented in Table S1). Using this model and rate constants, we were able to simulate the dynamics of the components (plus RNA, minus RNA, single-stranded RNA, double-stranded RNA, replicase) during the translation-coupled self-replication reaction within 20% error (Figure 2).

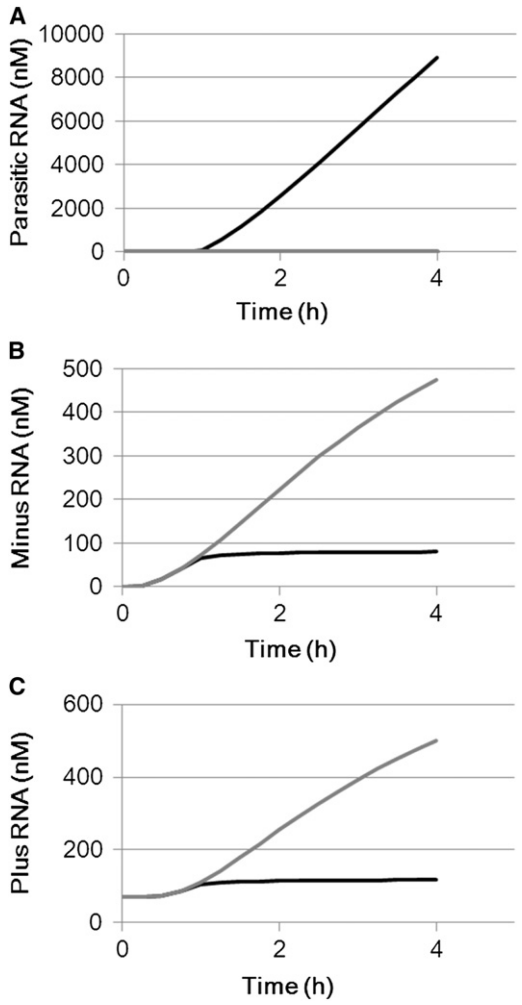

Figure 3. Computational Simulation of the Kinetics of Parasitic and Genomic RNAs without the Appearance of Parasite

The kinetics of each RNA, parasitic RNA (A), genomic minus RNA (B), and plus RNA (C), in the translation-coupled RNA self-replication reaction were simulated using the mathematical model (Figure 1B; Supplemental Information) and parameters shown in Table S1. Simulation of the present state, where about 2 fM parasite was assumed to appear immediately at time 0, is shown as the black lines, which showed a good fit with the experimental results (see Figure 2). The predicted results in the case in which the initial parasitic concentration was set at zero and the other parameters were the same as the present state are shown as gray lines.

See also Figure S1.

Using the model and parameters, we predicted the dynamics in the cases where each of the negative factors described above was eliminated or altered. For example, for negative factor 1, initial parasite concentration was set to zero, and for negative factor 2, the double-strand formation rate was decreased down to 30% (see the legend of Figure S1 for details). The simulation showed that improvement of negative factors 2–5 had only slight effects on the duration of plus RNA replication (Figure S1), whereas if factor 1 were eliminated (i.e., if parasite was assumed not to appear), plus and minus RNA replication was significantly prolonged (Figure 3). This computational result suggested that the main limiting factor for the present RNA self-replication system is the appearance of parasitic RNA.

Repression of Parasitic RNA by Compartmentalization

We experimentally examined whether repression of the parasitic RNA indeed enhanced genome replication. In previous studies on the origin of life, many theoretical reports indicated that spatial structures, such as compartments, repress the amplification of parasitic replicators by restricting their diffusion (Bresch et al., 1980; Maynard Smith, 1979; Szathmáry and Demeter, 1987). The basic concept of compartmentalization is that when the number of compartments is much larger than the number of parasites, the amplification of parasite is confined to a small number of compartments, whereas the other compartments are free from parasite where the genome replication continues.

We used water-in-oil emulsion as a compartmentalization method. This type of emulsion was first proposed as an artificial cell model by Tawfik and Griffiths (1998), and then utilized to repress nontemplate PCR amplification (Ghadessy et al., 2001). Here we adapted this technique to the translation-coupled replication system, and examined whether compartmentalization represses the parasite and enhances the replication of genomic RNA as predicted by the simulation.

The standard reaction mixture containing plus RNA (70 nM) and the cell-free translation system was dispersed in the oil phase (4.5% Span 80, 0.5% Tween 80 in light mineral oil [v/v]) by mixing with various strengths to produce emulsions of various average sizes ranging from 6.1 to 2,700 μm (Figure S2). The emulsions were incubated for 3 hr at 37°C to allow the translation-coupled replication reaction to proceed. After collecting the water phase by centrifugation, we measured replicated genomic minus RNA and parasitic RNA. In the largest emulsion (2,700 μm), parasitic RNA was amplified to more than 15,000 nM, but was below the detection limit (<500 nM) in the smallest emulsion (6.1 μm) (Figure 4A). In contrast, the synthesis of genomic minus RNA was about 2-fold higher in the smallest emulsion than in the largest emulsion (Figure 4B), inversely correlated with parasitic replication.

To examine the duration of the reaction, we measured the kinetics of minus RNA and parasitic RNA in both the smallest (6.1 μm) and largest (2,700 μm) emulsions. Parasitic RNA appeared after 2 hr and showed amplification up to 8,000 nM in the largest emulsion, whereas the parasite was not detected even after 5 hr in the smallest emulsion (Figure 5A). In contrast, the synthesis of genomic minus RNA stopped after 1 hr in the largest emulsion, whereas it continued up to 3 hr in the smallest emulsion (Figure 5B). For genomic plus RNA, no significant replication was observed in the largest emulsion, whereas the replication continued up to 4 hr and finally the plus RNA replicated to more than double the initial concentration in the smallest emulsion (Figure 5C). These results indicated that the appearance of parasite indeed limited the genomic plus and minus RNA replication in the translation-coupled self-replication system, consistent with our predictions based on the mathematical model and simulation (Figure 3).

It is notable that the emulsion preparation strategy did not affect the replication reaction (Urabe et al., 2010) or the translation reaction (Figure S3). This is reasonable, because all components except the parasitic RNA should be present at identical

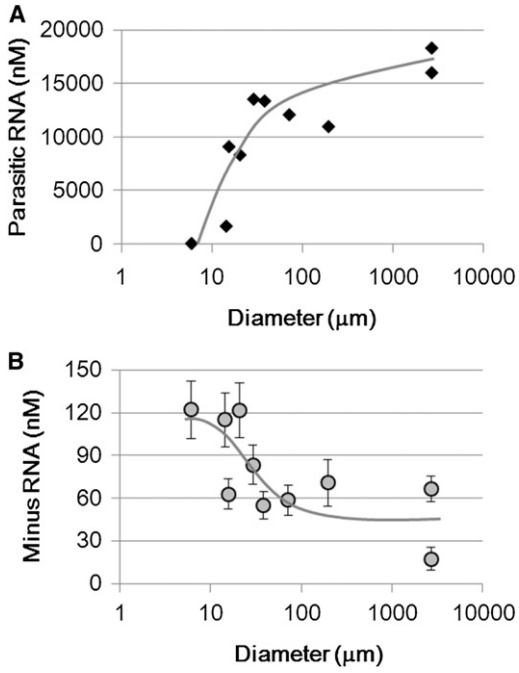

Figure 4. Effects of Compartment Size on the Internal Self-Replication Reaction

The translation-coupled self-replication reaction was performed in emulsions with various average sizes prepared by several mixing methods as described in Experimental Procedures. The size distributions were determined by measuring the sizes of more than 100 droplets for each preparation under a microscope (Figure S2). After 3 hr of incubation, all the emulsions were collected by centrifugation and the concentrations of parasitic RNA (A) and genomic minus RNA (B) were measured as described in Experimental Procedures. Error bars show standard deviations (n = 3). These data of parasitic RNA were used to estimate the rate of parasitic RNA appearance in Figure S6.

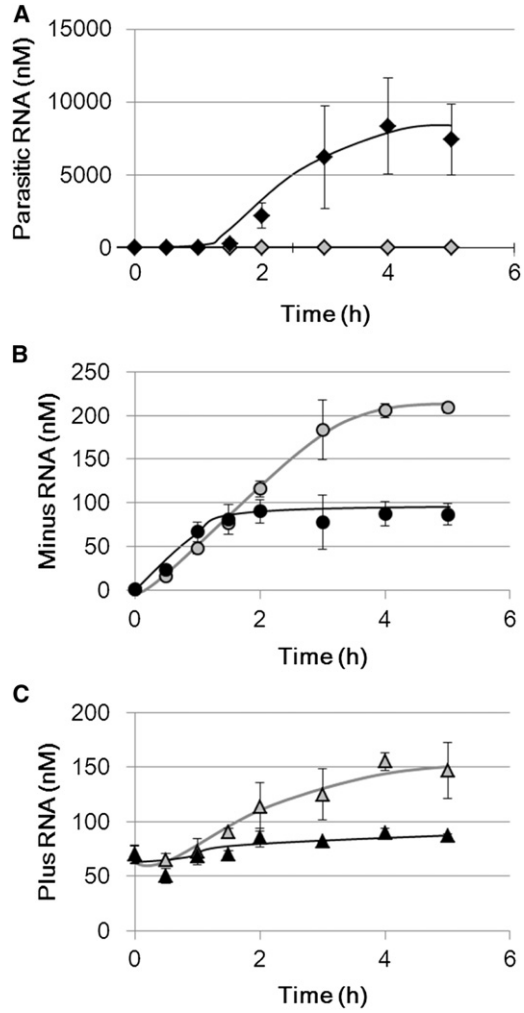

Figure 5. Effects of Compartment Size on the Kinetics of Parasitic and Genomic RNA Replication

The translation-coupled self-replication reaction was performed in emulsions with average sizes of 6.1 μm (open symbols) and 2,700 μm (filled symbols). The 6.1 μm emulsion was prepared by vigorous mixing with a vortex mixer and the sizes of more than 100 droplets were measured under a microscope. The 2,700 μm emulsion consisted of a single droplet. After incubation for the indicated times, all the emulsions were collected by centrifugation and the concentrations of parasitic RNA (A), genomic minus RNA (B), and plus RNA (C) were measured as described in Experimental Procedures. Error bars show standard deviations (n = 3).
See also Figure S3.

concentrations under our experimental conditions: the copy numbers of genomic plus RNA and all the translational factors were more than 1,000 molecules even in the smallest emulsion (6.1 μm). Therefore, compartmentalization affects only the parasitic RNA amplification but not the internal genomic replication or translation reactions.

Frequency of Parasite Appearance and Its Origin

In the above experiments, parasitic RNA was sufficiently repressed in emulsions with average sizes of less than 10 μm. To understand what determines the sufficient size, we next investigated the origin of the parasitic RNA and its frequency of appearance. We cloned some of the parasitic RNAs that appeared in the 15 μm emulsion during the self-replication reaction and analyzed the sequences. All of the clones were about 220 bases in length and showed high degrees of similarity to a previously reported RNA template for Qβ replicase, MDV-1 (Mills et al., 1967), except that relatively high levels of variation were observed in the region around base 160 (Figure S5).

MDV-1 RNA is one of the fastest replicable RNAs, which is known to appear frequently in the presence of Qβ replicase from many sources, such as de novo synthesis from NTPs (Biebricher et al., 1986) and selection and mutation from rRNA or tRNA (Kacian et al., 1971; Munishkin et al., 1988). Therefore, the parasitic RNA may also be derived from these sources in the cell-free translation system. Another possible source is recombination of the genomic plus RNA because the plus RNA was originally constructed by inserting the β subunit gene of Qβ replicase into the MDV-1 derivative, MDV-poly (Kita et al., 2008). Therefore, nonhomologous recombination of plus RNA at appropriate sites can produce MDV-1 derivatives. Nonhomologous recombination has been reported to occur spontaneously at 37°C, although at a low rate (10^{-7}/hr) (Chetverin et al., 1997). These observations suggest two models for the origin of the

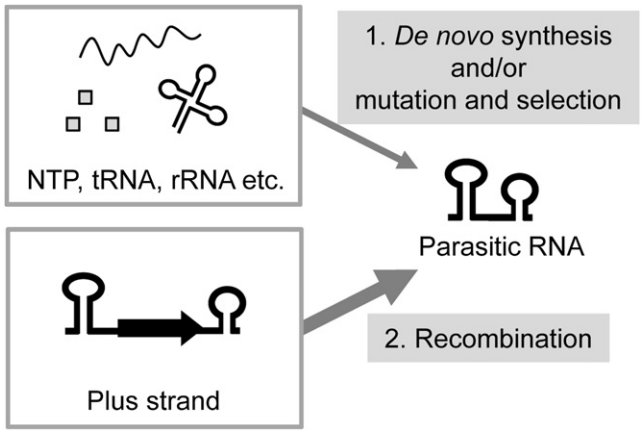

Figure 6. Two Possible Routes of Parasitic RNA Production
In the first route, parasitic RNA can be produced by de novo synthesis from NTPs included in the reaction mixture, or produced as a result of mutation and selection from RNAs present in the cell-free translation system such as tRNA and rRNA. In the second route, parasitic RNA can be produced by nonhomologous recombination from genomic plus RNA because plus RNA was originally constructed by inserting the replicase subunit gene into a replicable RNA, MDV-poly, which is known to replicate itself in the presence of the replicase.
See also Figure S5.

parasitic RNA: (1) de novo synthesis or synthesis from the RNAs in the cell-free translation system; and (2) nonhomologous recombination from genomic plus RNA (Figure 6).

To distinguish between these two models, we examined the plus RNA dependency of the appearance of parasitic RNA, because model 2 is dependent on plus RNA whereas model 1 is not. We measured the amount of amplified parasitic RNA after incubation of the cell-free translation system with or without plus RNA in the presence of exogenously added purified replicase in the 20 μm emulsion. In addition, to evaluate the effects of internal translation of the replicase, we used mutant RNA lacking the 100 base internal region of the replicase gene that has therefore lost the ability to produce active replicase (Δrep plus RNA). We found that a much larger amount of parasitic RNA appeared in the reaction with both plus RNA and Δrep plus RNA than in that with no plus RNA (Figure 7), indicating that the appearance of parasitic RNA in the replication system is mostly dependent on plus RNA, probably through nonhomologous recombination. Based on this model, the relatively high degree of variation in the region around base 160 among the parasite clones described above can be explained by the variation in recombination sites. Note that we observed a small amount of parasitic RNA even without plus RNA, indicating that the parasitic RNA also has a plus RNA-independent origin.

We next estimated the frequency of the appearance of parasitic RNA from the size dependency shown in Figure 4A based on the recombination model. Assuming that the rate of parasite appearance by recombination is constant over time, we simulated parasitic RNA amplification in emulsions of various sizes using the mathematical model. The results indicated that the size dependency of the appearance of parasitic RNA could be well explained when the rate of appearance was set to 10^{-8}/hr

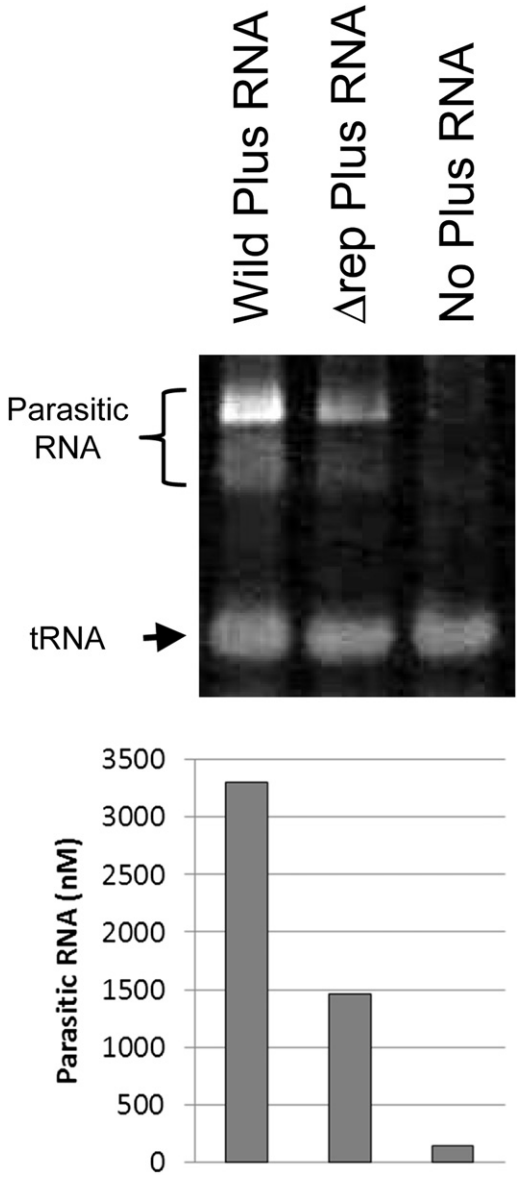

Figure 7. Dependency of Parasitic RNA Appearance on Genomic Plus RNA
In the presence of exogenously added replicase (200 nM), the standard reaction mixture was incubated for 2 hr with or without normal or mutant plus RNA in emulsions with an average size of 20 μm prepared by gentle mixing with a vortex mixer. All of the emulsions were collected and subjected to polyacrylamide gel electrophoresis, followed by staining with SYBR Green II (Invitrogen). The mutant plus RNA lacked part of the replicase subunit coding region and thus did not produce active replicase. The quantitative results are shown at the bottom.
See also Figure S6.

per plus RNA molecule (Figure S6). This value is close to that reported previously (10^{-7}/hr and 10^{-9}/hr), although both were rough estimates (Chetverin et al., 1997; Chetverina et al., 1999). Based on this value, the effects of compartmentalization on the parasite replication in Figures 4 and 5 can be explained as follows. The average copy numbers of parasite appearing in

1 hr were 0.001 and 1 molecules in 10 μm and 100 μm compartments, respectively. Therefore, in the 10 μm compartment, the majority of the emulsion was free from parasite and genome replication continued for a longer period.

DISCUSSION

Recently, several groups have proposed the construction of a minimal or artificial cell from scratch, and many types of cell functions have been constructed as described in the Introduction. One of the next challenges is the integration of these individual functions into higher-order biological functions, such as the translation-coupled replication shown here. Generally, the integration of different types of reactions sometimes causes unexpected interactions among components and reactions, resulting in unpredicted behavior of an integrated system. To overcome this problem, we first established a mathematical model of the system that includes most of the potentially problematic pathways, and obtained reliable parameter values experimentally (Figure 1B; Table S1). Using this model and parameters, we analyzed the behavior of the system under various conditions (i.e., at different parameter values or different initial concentrations) and found conditions under which the translation-coupled replication lasts for a longer time (Figure 3). The computational predictions were further supported by experimental evidence (Figure 5). Although this type of quantitative simulation has not been fully utilized for the in vitro construction with few exceptions (Kim and Winfree, 2011; Montagne et al., 2011), the results of this study suggested that it could be a useful and probably indispensable tool for further construction of integrated and thus more complex biological systems, such as an artificial or minimal cell.

The importance of a compartmental structure for parasite repression has been studied for decades theoretically using models of primordial life, including the "hypercycle" model, a hypothetical primitive genetic information replication system (Bresch et al., 1980; Eigen and Schuster, 1978; Niesert et al., 1981; Takeuchi and Hogeweg, 2009). The translation-coupled RNA replication system used here is also categorized as a minimum hypercycle (Eigen et al., 1991). As these previous studies were mainly theoretical, a number of questions remain. Do parasites actually appear and act as a limiting factor in replication systems composed of polynucleotides and proteins? From where and at what rate do such parasites appear? The results of this study experimentally answer these questions. We found that parasitic replicators actually appeared from genomic RNA at a rate of about 10^{-8}/hr and became the limiting factor for replication of genomic RNA. This result provides experimental evidence for the importance of repressing parasitic replicators by using a compartmental structure to achieve long-lasting self-replication.

Although this was a case study of a translation-coupled replication system of genetic information, the insights obtained here are of potential importance for other replication systems, including hypothetical primitive replicators in the origin of life, recently proposed artificial cell models, and artificial replication systems constructed in vitro. For example, in the hypothetical RNA world, genome recombination may have produced parasitic replicators by deleting the replication ribozyme encoded on the genome but maintaining the recognition sequence for the ribozyme. Similar selfish replicators were reported to appear in some artificial replication systems constructed in vitro (Breaker and Joyce, 1994; Hanczyc and Dorit, 1998). Once this type of parasite appeared, it would have replicated much faster than the genome because of its small size and eventually inhibit replication of the genome. Therefore, the frequency of the appearance of such parasites must generally be low to achieve long-lasting self-replication of the genome. This restricts the possible conditions under which primitive replication may have emerged on the ancient Earth. It also restricts the possible design of an artificial cell. To date, various schemes for the construction of artificial cell models have been proposed (Deamer, 2005; Forster and Church, 2006; Jewett and Forster, 2010; Luisi, 2002; Noireaux et al., 2011; Pohorille and Deamer, 2002; Stano, 2011; Szostak et al., 2001), but the importance of parasite repression has not been fully appreciated. The results of this study suggest that it is important to evaluate the frequency of parasite appearance and to design the cell size in which parasite appearance is negligible to construct an artificial or minimum cell harboring a genome replication system.

We showed that most of the parasitic RNA was produced from plus RNA, probably through RNA recombination (Figure 7), and postulated that the recombination site is located around base 160, as the parasitic RNAs have a relatively high degree of variation in this region (Figure S5). These observations raise questions regarding the possible mechanism of RNA recombination that can explain these results. It is known that RNA recombination can occur by several mechanisms, broadly classified as homologous or nonhomologous. Template switching is a well-known mechanism of homologous recombination, in which the complex of RNA replicase and nascent chain dissociates from the template RNA during replication, and then binds to the homologous region of another template RNA and restarts the replication reaction (Biebricher and Luce, 1992; Simon-Loriere and Holmes, 2011). As this mechanism requires homologous sequence overlap in the RNA fragments for recombination, it cannot explain the results observed here; plus RNA does not have a homologous sequence around the recombined region, and the existence of variation in the region of base 160 is not compatible with this mechanism. Two mechanisms of nonhomologous recombination have been proposed by Chetverin's group. The first is *trans*-esterification between the 3′ end of an RNA fragment and the internal region of another RNA fragment catalyzed by Qβ replicase (Chetverin et al., 1997). This mechanism also seems irrelevant to our results because the 3′ end does not participate in recombination to produce parasitic RNA. The second mechanism of nonhomologous recombination is called self-recombination (Chetverin, 1999), in which RNA recombines between internal regions almost randomly in the presence of magnesium ions without any protein, although the precise details of this mechanism are unknown. This type of recombination occurs also within an RNA molecule, resulting in deletion of an internal region, and produces variation in sequence at the recombination site, consistent with our observations. The frequency of recombination was reported to be 10^{-9}/hr, similar to our recombination rate (10^{-8}/hr). These observations suggest that this type of self-recombination is the most plausible mechanism to explain the production of parasitic RNA from plus RNA observed in our experiments.

Although we improved the duration of translation-coupled RNA replication in this study, replication eventually stopped after about 4 hr (Figure 5). Simulation using our mathematical model predicts that the next problem is double-stranded RNA formation between plus and minus RNA. As double-stranded RNA is inert for both translation and replication, it is a dead-end product of the self-replication system. This has been indicated previously as an inevitable problem of the system using single-stranded RNA as a genome, and must be resolved to achieve recursive replication (Forster and Church, 2007; Szostak et al., 2001). Recently, the 3D structure of the replicase was determined (Kidmose et al., 2010; Takeshita and Tomita, 2010). According to the structure, newly synthesized RNA and template RNA extrude from different parts of the replicase as single-stranded molecules. Thus, double-stranded RNA is thought to be formed by reassociation between plus and minus RNA after replication. This reassociation can be controlled by the RNA sequence, given that the ratio of the double-stranded form to the total replication product of the phage Qβ genomic RNA is low (about 30%) (Urabe et al., 2010), suggesting that the phage genome RNA contains sequences that inhibit reassociation. A more precise understanding of the reassociation mechanism will facilitate the development of new artificial genomic RNA with a lower tendency to form double-stranded RNA.

SIGNIFICANCE

The cell is composed of chemical molecules and is driven by chemical reactions among them. However, it is still not possible to assemble a living cell from nonliving molecules, indicating a lack of knowledge regarding the organization of molecules into a living state. To obtain such knowledge, many groups are attempting to construct each biological function individually, and integrate them into an artificial or minimal cell (Deamer, 2005; Forster and Church, 2006; Jewett and Forster, 2010; Luisi, 2002; Noireaux et al., 2011; Pohorille and Deamer, 2002; Stano, 2011; Szostak et al., 2001). This is called the bottom-up approach or in vitro synthetic biology. Here we attempted to construct one of the fundamental biological functions, replication of genetic information, from identified molecules. We found that the appearance of parasitic replicators is a critical problem for the replication system, and successfully overcame this problem by encapsulating the reaction into microcompartments. These results experimentally demonstrated that a mechanism for repression of parasitic replicators is required to achieve a long-lasting genome replication system, providing important insights for the possible design of an artificial cell.

The knowledge obtained here has broad significance for all researchers interested in not only artificial cell construction but also the origin of life. The importance of parasite repression for primitive replication systems was first proposed in the 1970s by using Manfred Eigen's hypercycle model (Maynard Smith, 1979) and has attracted the interest of many researchers over the intervening decades (Takeuchi and Hogeweg, 2007), but experimental demonstration has not been reported. This report indicates that parasite, the replication of which depends on other replicators, actually appears and is a limiting factor for a genome replication system composed of nucleotides and polypeptides.

EXPERIMENTAL PROCEDURES

Materials

Qβ replicase was purified as described previously (Kita et al., 2006). Plus and minus RNAs were prepared by in vitro transcription following SmaI digestion of the plasmids pUCmdv(–)β(+) and pUCmdv(+)β(–), respectively (Kita et al., 2008). The plasmid for mutant plus RNA used in Figure 7 was constructed by digestion of pUCmdv(–)β(+) with SalI and self-ligation of the larger fragment.

Translation-Coupled RNA Self-Replication Reaction

The standard reaction mixture contained the plus RNA (70 nM) and the reconstituted cell-free translation system (PURE system; Shimizu et al., 2001) modified as described below (Kazuta et al., 2008; Matsuura et al., 2009). All of the protein components were purified in our laboratory and mixed in the compositions described previously (Hosoda et al., 2008) except for the omission of T7 RNA polymerase. The compositions of low-molecular-weight compounds were optimized for the RNA replication reaction as follows: tyrosine (0.3 mM), cysteine (0.3 mM), 18 other amino acids (0.36 mM), 0.39 mg/ml tRNA mix (Roche), ATP (3.75 mM), GTP (2.5 mM), CTP (1.25 mM), UTP (1.25 mM), HEPES-KOH (pH 7.6, 100 mM), potassium glutamate (70 mM), spermidine (0.375 mM), magnesium acetate (18 mM), creatine phosphate (25 mM), dithiothreitol (6 mM), and 5-formyl-5,6,7,8-tetrahydrofolic acid (10 μg/ml). All of the reactions were performed at 37°C.

Assay for Replicated RNA

Plus and minus RNA was measured by quantitative PCR followed by reverse transcription using the complementary RNA as a standard as follows. Reaction mixtures containing plus or minus RNA were diluted more than 10,000-fold with 1 mM EDTA (pH 8.0) and heated at 95°C for 5 min to dissociate double-stranded RNA. The heated sample was subjected to reverse transcription (RT) (PrimeScript; Takara) with RT primer at 50°C for 30 min according to the manufacturer's instructions, followed by 5-fold dilution and mixing with the solution for quantitative PCR (SYBR Premix Ex Taq; Takara) containing PCR primers. The RT primers (5′-GCAAGTGACTCAGGATTCGTACATAATA TCGTCTCCGTAAACAGTG-3′) or (5′-TAAGCGAATGTTGCGAGCACGGCCCA TTCTGTGTACCTCAAG-3′) and PCR primer sets (5′-GGTAGTGTTGTTACCT ACGAGAAG-3′ and 5′-GCAAGTGACTCAGGATTCGTAC-3′) or (5′-GATCCA CCCGCGGTTTTTC-3′ and 5′-TAAGCGAATGTTGCGAGCAC-3′) were used for plus RNA and minus RNA detection, respectively. To measure the amounts of parasitic RNA, we separated the parasitic RNA from other RNAs by 8% PAGE with 0.1% SDS in TBE buffer (pH 8.4) containing Tris(hydroxymethyl) aminomethane (100 mM), boric acid (90 mM), and EDTA (1 mM), followed by staining with SYBR Green II (Invitrogen). The intensities of the bands corresponding to 100–300 bases were determined. Note that because staining with SYBR Green II is dependent on RNA length and sequence, the parasitic RNA concentrations shown in Figures 4, 5, and 7 are qualitative rather than quantitative.

Emulsification and Recovery

A mixture of mineral oil (Sigma) and surfactants, 4.5% (v/v) Span80 (Sigma) and 0.5% (v/v) Tween80 (Sigma), was used as the oil phase. The oil phase was saturated with the saturation buffer as follows. The saturation buffer, which had the same composition as the mixture of low-molecular-weight compounds except for the omission of NTPs and tRNA, was mixed vigorously with the oil phase (150 μl per 1 ml of oil phase) and then incubated for 10 min at 37°C. After centrifugation at 22,000 × *g* for 5 min, the upper oil phase was collected. Water-in-oil emulsion was prepared by mixing 500 μl of the saturated oil phase and 20 μl of the standard reaction mixture. The average volume of emulsion was roughly controlled by changing the strength of mixing. For example, the smallest emulsion (6.1 μm) was prepared by vigorous mixing with a vortex mixer at maximum intensity for 30 s, and the larger emulsion was prepared by gentle mixing with a pipette. The diameters of more than 100 droplets in each emulsion were measured under an optical microscope. After the reaction, 1 μl of 0.5 M EDTA (pH 8.0) was added to stop the reaction.

The water phase of the emulsion was recovered as described previously (Tawfik and Griffiths, 1998). The absorbance at 260 nm was measured to correct the differences in recovery among samples.

Modeling and Simulation

Reactions in the translation-coupled RNA self-replication system were modeled as shown schematically in Figure 1B, which is composed of 26 reaction formulas as described in Supplemental Information. Based on these formulas, the deviations of all the components are represented in the form of 11 differential equations (not shown). The differential equations were solved numerically using the NDSolve algorithm in Mathematica (Wolfram Research) for simulation of the kinetics of each component. Methods regarding modeling, simulation, and parameter estimation are described in more detail in Supplemental Information (see also Figure S4).

SUPPLEMENTAL INFORMATION

Supplemental Information includes six figures, one table, and Supplemental Experimental Procedures and can be found with this article online at doi:10.1016/j.chembiol.2012.01.019.

ACKNOWLEDGMENTS

We are grateful to Ms. Naoko Miki, Ms. Hitomi Komai, and Ms. Ryoko Otsuki for their technical assistance. This research was supported in part by Special Coordination Funds for Promoting Science and Technology: Yuragi Project and the Global COE (Centers of Excellence) Program of the Ministry of Education, Culture, Sports, Science, and Technology, Japan.

Received: December 10, 2011
Revised: January 18, 2012
Accepted: January 19, 2012
Published: April 19, 2012

REFERENCES

Benner, S.A., and Sismour, A.M. (2005). Synthetic biology. Nat. Rev. Genet. *6*, 533–543.

Biebricher, C.K., and Luce, R. (1992). In vitro recombination and terminal elongation of RNA by Qβ replicase. EMBO J. *11*, 5129–5135.

Biebricher, C.K., Eigen, M., and Luce, R. (1986). Template-free RNA synthesis by Qβ replicase. Nature *321*, 89–91.

Breaker, R.R., and Joyce, G.F. (1994). Emergence of a replicating species from an in vitro RNA evolution reaction. Proc. Natl. Acad. Sci. USA *91*, 6093–6097.

Bresch, C., Niesert, U., and Harnasch, D. (1980). Hypercycles, parasites and packages. J. Theor. Biol. *85*, 399–405.

Chakrabarti, A.C., Breaker, R.R., Joyce, G.F., and Deamer, D.W. (1994). Production of RNA by a polymerase protein encapsulated within phospholipid vesicles. J. Mol. Evol. *39*, 555–559.

Chen, I.A., Roberts, R.W., and Szostak, J.W. (2004). The emergence of competition between model protocells. Science *305*, 1474–1476.

Chetverin, A.B. (1999). The puzzle of RNA recombination. FEBS Lett. *460*, 1–5.

Chetverin, A.B., Chetverina, H.V., Demidenko, A.A., and Ugarov, V.I. (1997). Nonhomologous RNA recombination in a cell-free system: evidence for a transesterification mechanism guided by secondary structure. Cell *88*, 503–513.

Chetverina, H.V., Demidenko, A.A., Ugarov, V.I., and Chetverin, A.B. (1999). Spontaneous rearrangements in RNA sequences. FEBS Lett. *450*, 89–94.

Chiarabelli, C., Stano, P., and Luisi, P.L. (2009). Chemical approaches to synthetic biology. Curr. Opin. Biotechnol. *20*, 492–497.

Deamer, D. (2005). A giant step towards artificial life? Trends Biotechnol. *23*, 336–338.

Eigen, M., and Schuster, P. (1978). The hypercycle: a principle of natural self-organization. Part B: the abstract hypercycle. Naturwissenschaften *65*, 7–41.

Eigen, M., Biebricher, C.K., Gebinoga, M., and Gardiner, W.C. (1991). The hypercycle. Coupling of RNA and protein biosynthesis in the infection cycle of an RNA bacteriophage. Biochemistry *30*, 11005–11018.

Forster, A.C., and Church, G.M. (2006). Towards synthesis of a minimal cell. Mol. Syst. Biol. *2*, 45.

Forster, A.C., and Church, G.M. (2007). Synthetic biology projects in vitro. Genome Res. *17*, 1–6.

Ghadessy, F.J., Ong, J.L., and Holliger, P. (2001). Directed evolution of polymerase function by compartmentalized self-replication. Proc. Natl. Acad. Sci. USA *98*, 4552–4557.

Guatelli, J.C., Whitfield, K.M., Kwoh, D.Y., Barringer, K.J., Richman, D.D., and Gingeras, T.R. (1990). Isothermal, in vitro amplification of nucleic acids by a multienzyme reaction modeled after retroviral replication. Proc. Natl. Acad. Sci. USA *87*, 1874–1878.

Hanczyc, M.M., and Dorit, R.L. (1998). Experimental evolution of complexity: in vitro emergence of intermolecular ribozyme interactions. RNA *4*, 268–275.

Hanczyc, M.M., Fujikawa, S.M., and Szostak, J.W. (2003). Experimental models of primitive cellular compartments: encapsulation, growth, and division. Science *302*, 618–622.

Hosoda, K., Sunami, T., Kazuta, Y., Matsuura, T., Suzuki, H., and Yomo, T. (2008). Quantitative study of the structure of multilamellar giant liposomes as a container of protein synthesis reaction. Langmuir *24*, 13540–13548.

Ichihashi, N., Matsuura, T., Hosoda, K., and Yomo, T. (2010a). Identification of two forms of Qβ replicase with different thermal stabilities but identical RNA replication activity. J. Biol. Chem. *285*, 37210–37217.

Ichihashi, N., Matsuura, T., Kita, H., Sunami, T., Suzuki, H., and Yomo, T. (2010b). Constructing partial models of cells. Cold Spring Harb. Perspect. Biol. *2*, a004945.

Jewett, M.C., and Forster, A.C. (2010). Update on designing and building minimal cells. Curr. Opin. Biotechnol. *21*, 697–703.

Kacian, D.L., Mills, D.R., and Spiegelman, S. (1971). The mechanism of Q replication: sequence at the 5′ terminus of a 6-S RNA template. Biochim. Biophys. Acta *238*, 212–223.

Kazuta, Y., Adachi, J., Matsuura, T., Ono, N., Mori, H., and Yomo, T. (2008). Comprehensive analysis of the effects of *Escherichia coli* ORFs on protein translation reaction. Mol. Cell. Proteomics *7*, 1530–1540.

Kidmose, R.T., Vasiliev, N.N., Chetverin, A.B., Andersen, G.R., and Knudsen, C.R. (2010). Structure of the Qβ replicase, an RNA-dependent RNA polymerase consisting of viral and host proteins. Proc. Natl. Acad. Sci. USA *107*, 10884–10889.

Kim, J., and Winfree, E. (2011). Synthetic in vitro transcriptional oscillators. Mol. Syst. Biol. *7*, 465.

Kita, H., Cho, J., Matsuura, T., Nakaishi, T., Taniguchi, I., Ichikawa, T., Shima, Y., Urabe, I., and Yomo, T. (2006). Functional Qβ replicase genetically fusing essential subunits EF-Ts and EF-Tu with β-subunit. J. Biosci. Bioeng. *101*, 421–426.

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. Chem. Bio. Chem. *9*, 2403–2410.

Kurihara, K., Tamura, M., Shohda, K., Toyota, T., Suzuki, K., and Sugawara, T. (2011). Self-reproduction of supramolecular giant vesicles combined with the amplification of encapsulated DNA. Nat. Chem. *3*, 775–781.

Kuruma, Y., Stano, P., Ueda, T., and Luisi, P.L. (2009). A synthetic biology approach to the construction of membrane proteins in semi-synthetic minimal cells. Biochim. Biophys. Acta *1788*, 567–574.

Lincoln, T.A., and Joyce, G.F. (2009). Self-sustained replication of an RNA enzyme. Science *323*, 1229–1232.

Luisi, P.L. (2002). Toward the engineering of minimal living cells. Anat. Rec. *268*, 208–214.

Luisi, P.L., Ferri, F., and Stano, P. (2006). Approaches to semi-synthetic minimal cells: a review. Naturwissenschaften *93*, 1–13.

Mansy, S.S., Schrum, J.P., Krishnamurthy, M., Tobé, S., Treco, D.A., and Szostak, J.W. (2008). Template-directed synthesis of a genetic polymer in a model protocell. Nature *454*, 122–125.

Matsuura, T., Kazuta, Y., Aita, T., Adachi, J., and Yomo, T. (2009). Quantifying epistatic interactions among the components constituting the protein translation system. Mol. Syst. Biol. *5*, 297.

Maynard Smith, J. (1979). Hypercycles and the origin of life. Nature *280*, 445–446.

Mills, D.R., Peterson, R.L., and Spiegelman, S. (1967). An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule. Proc. Natl. Acad. Sci. USA *58*, 217–224.

Montagne, K., Plasson, R., Sakai, Y., Fujii, T., and Rondelez, Y. (2011). Programming an in vitro DNA oscillator using a molecular networking strategy. Mol. Syst. Biol. *7*, 466.

Munishkin, A.V., Voronin, L.A., and Chetverin, A.B. (1988). An in vivo recombinant RNA capable of autocatalytic synthesis by Qβ replicase. Nature *333*, 473–475.

Murtas, G., Kuruma, Y., Bianchini, P., Diaspro, A., and Luisi, P.L. (2007). Protein synthesis in liposomes with a minimal set of enzymes. Biochem. Biophys. Res. Commun. *363*, 12–17.

Niesert, U., Harnasch, D., and Bresch, C. (1981). Origin of life between Scylla and Charybdis. J. Mol. Evol. *17*, 348–353.

Nishihara, T., Mills, D.R., and Kramer, F.R. (1983). Localization of the Qβ replicase recognition site in MDV-1 RNA. J. Biochem. *93*, 669–674.

Noireaux, V., and Libchaber, A. (2004). A vesicle bioreactor as a step toward an artificial cell assembly. Proc. Natl. Acad. Sci. USA *101*, 17669–17674.

Noireaux, V., Maeda, Y.T., and Libchaber, A. (2011). Development of an artificial cell, from self-organization to computation and self-reproduction. Proc. Natl. Acad. Sci. USA *108*, 3473–3480.

Nomura, S.M., Tsumoto, K., Hamada, T., Akiyoshi, K., Nakatani, Y., and Yoshikawa, K. (2003). Gene expression within cell-sized lipid vesicles. Chem. Bio. Chem. *4*, 1172–1175.

Oberholzer, T., Albrizio, M., and Luisi, P.L. (1995a). Polymerase chain reaction in liposomes. Chem. Biol. *2*, 677–682.

Oberholzer, T., Wick, R., Luisi, P.L., and Biebricher, C.K. (1995b). Enzymatic RNA replication in self-reproducing vesicles: an approach to a minimal cell. Biochem. Biophys. Res. Commun. *207*, 250–257.

Oberholzer, T., Nierhaus, K.H., and Luisi, P.L. (1999). Protein expression in liposomes. Biochem. Biophys. Res. Commun. *261*, 238–241.

Pohorille, A., and Deamer, D. (2002). Artificial cells: prospects for biotechnology. Trends Biotechnol. *20*, 123–128.

Ruiz-Mirazo, K., Peretó, J., and Moreno, A. (2004). A universal definition of life: autonomy and open-ended evolution. Orig. Life Evol. Biosph. *34*, 323–346.

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.

Simon-Loriere, E., and Holmes, E.C. (2011). Why do RNA viruses recombine? Nat. Rev. Microbiol. *9*, 617–626.

Simpson, M.L. (2006). Cell-free synthetic biology: a bottom-up approach to discovery by design. Mol. Syst. Biol. *2*, 69.

Stano, P. (2011). Minimal cells: relevance and interplay of physical and biochemical factors. Biotechnol. J. *6*, 850–859.

Szathmáry, E., and Demeter, L. (1987). Group selection of early replicators and the origin of life. J. Theor. Biol. *128*, 463–486.

Szathmáry, E., and Maynard Smith, J. (1995). The major evolutionary transitions. Nature *374*, 227–232.

Szostak, J.W., Bartel, D.P., and Luisi, P.L. (2001). Synthesizing life. Nature *409*, 387–390.

Takakura, K., Toyota, T., and Sugawara, T. (2003). A novel system of self-reproducing giant vesicles. J. Am. Chem. Soc. *125*, 8134–8140.

Takeshita, D., and Tomita, K. (2010). Assembly of Qβ viral RNA polymerase with host translational elongation factors EF-Tu and -Ts. Proc. Natl. Acad. Sci. USA *107*, 15733–15738.

Takeuchi, N., and Hogeweg, P. (2007). The role of complex formation and deleterious mutations for the stability of RNA-like replicator systems. J. Mol. Evol. *65*, 668–686.

Takeuchi, N., and Hogeweg, P. (2009). Multilevel selection in models of prebiotic evolution II: a direct comparison of compartmentalization and spatial self-organization. PLoS Comput. Biol. *5*, e1000542.

Tawfik, D.S., and Griffiths, A.D. (1998). Man-made cell-like compartments for molecular evolution. Nat. Biotechnol. *16*, 652–656.

Urabe, H., Ichihashi, N., Matsuura, T., Hosoda, K., Kazuta, Y., Kita, H., and Yomo, T. (2010). Compartmentalization in a water-in-oil emulsion repressed the spontaneous amplification of RNA by Qβ replicase. Biochemistry *49*, 1809–1813.

Walde, P., Wick, R., Fresta, M., Mangone, A., and Luisi, P.L. (1994). Autopoietic self-reproduction of fatty acid vesicles. J. Am. Chem. Soc. *116*, 11649–11654.

Walker, G.T., Little, M.C., Nadeau, J.G., and Shank, D.D. (1992). Isothermal in vitro amplification of DNA by a restriction enzyme/DNA polymerase system. Proc. Natl. Acad. Sci. USA *89*, 392–396.

Wright, M.C., and Joyce, G.F. (1997). Continuous in vitro evolution of catalytic function. Science *276*, 614–617.

Yu, W., Sato, K., Wakabayashi, M., Nakaishi, T., Ko-Mitamura, E.P., Shima, Y., Urabe, I., and Yomo, T. (2001). Synthesis of functional protein in liposome. J. Biosci. Bioeng. *92*, 590–593.

Zhu, T.F., and Szostak, J.W. (2009). Coupled growth and division of model protocell membranes. J. Am. Chem. Soc. *131*, 5705–5713.