DOI: 10.1002/cbic.200800360

Replication of Genetic Information with Self-Encoded Replicase in Liposomes

Hiroshi Kita,[a] Tomoaki Matsuura,[b] Takeshi Sunami,[b] Kazufumi Hosoda,[c] Norikazu Ichihashi,[a] Koji Tsukada,[d] Itaru Urabe,[d] and Tetsuya Yomo*[a, b, c]

In all living systems, the genome is replicated by proteins that are encoded within the genome itself. This universal reaction is essential to allow the system to evolve. Here, we have constructed a simplified system involving encapsulated macromolecules termed a "self-encoding system", in which the genetic information is replicated by self-encoded replicase in liposomes. That is, the universal reaction was reconstituted within a microcompartment bound by a lipid bilayer. The system was assembled by using one template RNA sequence as the information molecule and an in vitro translation system reconstituted from purified translation factors as the machinery for decoding the information. In this system, the catalytic subunit of Qβ replicase is synthesized from the template RNA that encodes the protein. The replicase then replicates the template RNA that was used for its production. This in-liposome self-encoding system is one of the simplest such systems available; it consists of only 144 gene products, while the information and the function for its replication are encoded on different molecules and are compartmentalized into the microenvironment for evolvability.

Introduction

Even the simplest living cells display enormous complexity, and exhibit three basic properties making life different from other types of networks: self-maintenance (metabolism), self-reproduction, and evolvability. This raises the question of whether such complexity is essential for a living system.[1] To address this question, researchers have been attempting to synthesize artificial cells from simple known substances.[1,2] Once such artificial cells have been developed, it will be possible to identify the precise number (or minimum number) of components that can constitute a living system. In addition, attempts to synthesize artificial cells can contribute to a better understanding of the origin of life because it will provide a physically possible path that could have led to primitive living cells. Furthermore, experimentally increasing the complexity of the artificial cell by starting from the simplest one will provide an opportunity to simulate evolutionary processes[1,2] to the development of more complex organisms and eventually current organisms.

Although the experimental construction of artificial cells has been proposed many times over the last several years,[1,2] progress toward this goal has been proceeding in discrete steps with researchers assembling elements that partially fulfill the properties of a living system. For example, it was shown to be possible to generate artificial lipid vesicles (liposomes) of the same size as small bacteria from amphiphilic molecules.[3] Artificial vesicles were also shown to be capable of autocatalytic growth, and even to be able to undergo repeated cycles of growth and division.[4] Various types of biological reaction (nucleic acid and protein synthesis,[5] integration of pore proteins,[6] two-stage genetic cascade reaction,[7] and production of enzymes involved in the synthesis of membrane lipids)[8] have been performed successfully within the environment provided by liposomes. These studies represent significant steps toward assembly of an artificial cell. The next crucial step is to embed a universal property that is possessed by all living systems, that is, a genetic information replication system that is capable of conducting the replication reaction based on information encoded on its own genome. If such a multicomponent gene replication system in which the information unit (genotype) and the functional unit that replicates the information (phenotype) are encoded on different molecules can be compartmentalized, for example, by cell membranes, the system will fulfill the genotype–phenotype linkage and thus have the potential to evolve. Conversely, evolvability of the multicomponent gene replication system is maintained as long as the genetic information is replicated with the self-encoded enzyme, and such a system does not necessarily have to

[a] *H. Kita, Dr. N. Ichihashi, Prof. T. Yomo*
Exploratory Research for Advanced Technology (ERATO)
(Japan) Science and Technology Agency (JST)
2-1 Yamadaoka, Suita, Osaka, 565–0871 (Japan)
Fax: (+81) 6-6879–7433
E-mail: yomo@ist.osaka-u.ac.jp

[b] *Dr. T. Matsuura, T. Sunami, Prof. T. Yomo*
Department of Bioinformatics Engineering
Graduate School of Information Science and Technology, Osaka University
2–1 Yamadaoka, Suita, Osaka, 565–0871 (Japan)

[c] *K. Hosoda, Prof. T. Yomo*
Graduate School of Frontier Biosciences, Osaka University
University 2–1 Yamadaoka, Suita, Osaka, 565–0871 (Japan)

[d] *Dr. K. Tsukada, Prof. I. Urabe*
Department of Biotechnology
Graduate School of Engineering, Osaka University
2–1 Yamadaoka, Suita, Osaka, 565–0871 (Japan)

Supporting information for this article is available on the WWW under http://www.chembiochem.org or from the author.

encode other types of machinery, such as protein and lipid synthesis systems.

We assembled a simplified genetic replication reaction by self-encoded replicase in liposomes by using only defined components: one template RNA sequence as an information molecule and an in vitro translation system that was reconstructed from purified translation factors[9] as the machinery for decoding the information that was encoded on the RNA. While the compartmentalization of an in vitro translation system by using water-in-oil emulsions has been reported previously,[10] we used liposomes, which provide a biologically relevant environment. In the in vitro translation system, the β-subunit of Qβ replicase[11] is synthesized from the template RNA that encodes the protein. The replicase then replicates the template RNA used for its production. Thus, in this reaction, replicase is generated through translation of the self-encoded gene; this mimics current living systems. We designated this as self-encoding system. First, we report the kinetic behavior of the self-encoding system in vitro. Then, we describe expandability, in which a new functional gene is incorporated into the RNA sequence to allow the self-encoding system to exhibit additional phenotypes (functions). Finally, we report that the self-encoding system can be encapsulated and performed in liposomes. From the results of detailed statistical and kinetic analyses of the in-liposome reaction, we concluded that the system was functioning as designed. This system is one of the simplest artificial multicomponent self-encoding system composed of only 144 gene products in which the information unit (genotype) and the functional unit that replicates the information (phenotype) are encoded on different molecules, and importantly has the potential to evolve by being compartmentalized in liposomes.

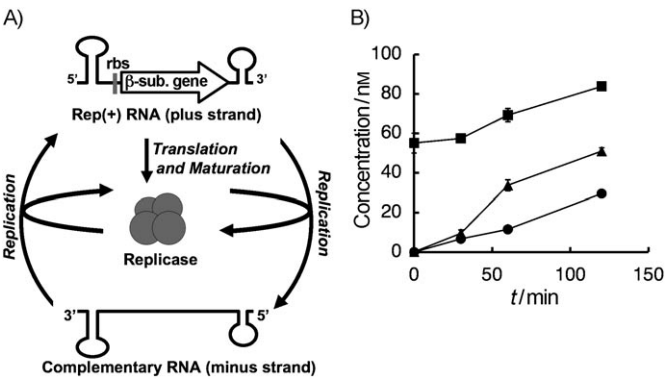

Figure 1. Replication of genetic information with self-encoded replicase. A) Scheme of the replication reaction with self-encoded protein. The template RNA (Rep(+) RNA) encoded the β-subunit of Qβ replicase within the sequence of a highly structured RNA (midivariant-RNA (MDV-RNA)[13]), which is a naturally occurring template for Qβ replicase. MDV-RNA provides loop structures at both ends of the RNA that are required for replication by Qβ replicase.[13b] rbs: ribosome binding site; β-sub. gene: β-subunit of the Qβ replicase gene. B) Time course of the replication reaction with self-encoded replicase in vitro at 37 °C. The concentrations of synthesized β-subunit (●), minus-strand RNA (▲), and plus-strand RNA (■) as a function of incubation time. Values represent mean ± s.d. from two independent experiments.

Results

Replication of genetic information with self-encoded replicase in vitro

A self-encoding system was assembled by using one RNA as an information molecule and an in vitro translation system reconstructed from purified translation factors[9] as the machinery for decoding the genetic information encoded on the RNA (Figure 1 A). In the in vitro translation system, all the components that were necessary for translation reactions were supplied individually in a highly purified form.[9] The RNA (Rep(+) RNA, Figure 1 A) was designed such that the replication reaction could be catalyzed by the protein that was translated from its own RNA sequence. For this purpose, the RNA was designed to fulfill two requirements. First, the RNA was designed to encode the β-subunit of Qβ replicase, an RNA-dependent RNA polymerase responsible for replicating the RNA genome of coliphage Qβ,[11] which can be decoded by the translational machinery. Qβ replicase is a heterotetramer composed of a β-subunit that is encoded on the phage genome and three host proteins of *Escherichia coli*: ribosomal protein S1, elongation factor Tu (EF-Tu), and Ts (EF-Ts). Because all three host proteins were originally included in the in vitro translation system as necessary factors for the translation reaction, it is sufficient to provide only the β-subunit to obtain the production of mature Qβ replicase (Figure 1 A).[12] Second, the RNA was designed to serve as a template for the replication reaction by Qβ replicase. For this requirement, midivariant-RNA (MDV-RNA),[13] which is known to act as a template for Qβ replicase, was used as a replication scaffold for the β-subunit sequence (Figure 1 A).

The time courses after the addition of Rep(+) RNA to the in vitro translation system is shown in Figure 1 B. Increases in the concentrations of β-subunit (circles) and the minus-strand of the template RNA (triangles) were observed. Furthermore, the plus-strand RNA concentration (squares) increased up to 1.5-fold after 120 min; this was attributed to the replication reaction of the minus-strand RNA with the translated replicase. When the same reaction was carried out in the absence of amino acids (Cys and Thr), neither production of β-subunit and minus-strand RNA, nor increases in plus-strand concentration were observed; this indicates the necessity of translation of the replicase for replication of its own RNA. From these results, we concluded that the RNA template was able to produce functional replicase that then effectively replicated itself.

Integration of an additional phenotype into the self-encoding system

We examined whether the self-encoding system could be expanded to express an additional phenotype. The antisense sequence of the β-galactosidase gene (*lacZ*) was inserted into the template RNA to construct Rep(+)Gal(−) RNA (Figure 2 A). This modified RNA was added to the in vitro translation system in the presence of fluorogenic substrate 5-chloromethylfluorescein di-β-D-galactopyranoside (CMFDG), and β-galac-

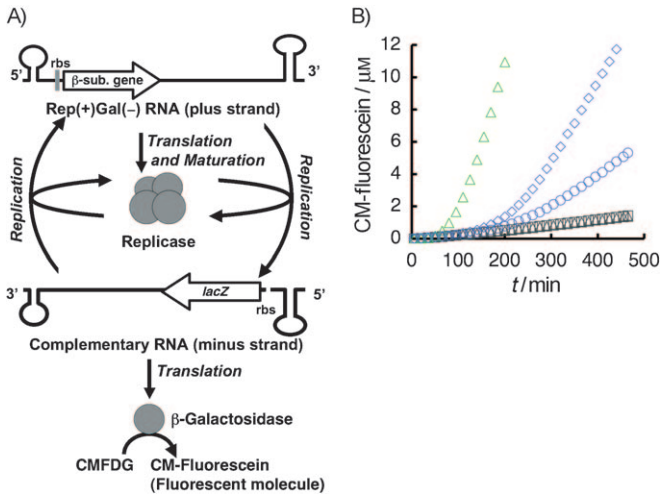

Figure 2. Integration of an additional phenotype into the self-encoding system. A) Schematic representation of the reaction with an additional phenotype that was generated by insertion of the *lacZ* gene. The Qβ replicase β-subunit was encoded on the plus-strand RNA, and β-galactosidase was encoded on the minus-strand RNA (complement of the plus-strand RNA). Nonfluorescent CMFDG was hydrolyzed by β-galactosidase to yield green fluorescent CM-fluorescein. B) Real-time detection of the increase in green fluorescence intensity. The reactions were carried out at 37 °C by using 70 nM (◇) or 30 nM (○) Rep(+)Gal(−) RNA, ΔRep(+)Gal(−) RNA (△), or ΔRep(+)-Gal(−) RNA supplemented with purified Qβ replicase (△) as template RNA or without template RNA (□).

tosidase activity (increase in green fluorescence intensity) was detected as an additional functionality of the self-encoding system (Figure 2B). Furthermore, by comparing the results with 30 nM or 70 nM Rep(+)Gal(−) RNA (blue circles and blue diamonds, respectively), the yield of 5-chloromethylfluorescein (CM-fluorescein) was found to be dependant on the RNA concentration, that is, 0.14 ± 0.01 μM CM-fluorescein per nM RNA at 400 min. Experiments performed in the absence of template RNA under the same reaction conditions showed only a slight increase in fluorescence (black squares), this is presumably due to a minute level of CMFDG hydrolysis activity of some component(s) in the translation system. A template RNA that lacks part of the β-subunit gene (ΔRep(+)Gal(−) RNA, black triangles) showed a slight fluorescence increase that was similar to that in the absence of template RNA, but with the assistance by externally added Qβ replicase, it gave rise to a detectable signal increase (green triangles). These results indicate that translation of functional replicase was necessary for the additional functionality. The fluorescence signal with Rep(+)Gal(−) RNA is produced through a reaction cascade (Figure 2A): 1) Rep(+)Gal(−) RNA produces the replicase; 2) the replicase generates the minus-strand from the plus-strand; 3) the minus-strand is translated into β-galactosidase; 4) β-galactosidase hydrolyzes CMFDG. Whereas synthesis of the replicase proceeded linearly over time (Figure S1 in the Supporting Information), the fluorescence increased initially in a higher-order reaction (Figure 2B, blue circles and blue diamonds). This is the consequence of the reaction cascade in which the rate of fluorescent molecule production is correlated with the β-galactosidase concentration, which increases over time. After 350 min, the reaction velocity became nearly constant, that is, the β-galactosidase concentration became constant, presumably due to inactivation of the translation machinery.[9]

Replication of genetic information with self-encoded replicase in liposomes

We investigated whether the self-encoding system carrying β-galactosidase activity as a phenotype can be embedded in cell-sized liposomes. Rep(+)Gal(−) RNA and the in vitro translation system were encapsulated into liposomes along with CMFDG. Subsequently, liposomes were diluted with the in vitro translation system that lacked EF-G, EF-Tu, and EF-Ts to prevent protein synthesis from occurring outside the liposomes. In this way, osmotic pressure differences that might cause leakage of the internal components were reduced, and the biochemical reaction could continue for up to 275 min as described below. Fluorescence microscopy was used for direct visualization of the performance of the reaction inside the liposomes (Figure 3A). The liposomes were clearly outlined by red fluorescence (Figure 3A, left), which was derived from the red fluorescent lipid used as the membrane marker,[14] while green fluorescence was observed inside the liposomes (Figure 3A, middle). Overlaying the images clearly indicated that the reaction occurred inside the liposomes (Figure 3A, right). Hence, we have successfully established an in-liposome self-encoding system in which the information of the replicase is decoded to replicate itself in cell-size lipid compartments.

Quantitative analysis of the in-liposome self-encoding system was carried out by using a fluorescence-activated cell sorter (FACS). Rep(+)Gal(−) RNA and the in vitro translation system were encapsulated into liposomes along with CMFDG and R-phycoerythrin (PE).[15] The replication reaction was carried out in liposomes, and the time course was analyzed by FACS (Figure 3B), which provided quantitative data for individual liposomes in the internal aqueous volume (vertical axes) from the PE red fluorescence[15] and hydrolysis of CMFDG from green fluorescence intensity (horizontal axes). From the vertical axis of Figure 3B, estimated internal volume was found to range from 1 to 100 fL, which is typical for multilamellar liposomes prepared by the freeze-dry method as described previously.[15] Furthermore, as shown in Figure 3B, the distribution of liposomes moved toward the right over time; this indicates an increase in the number of fluorescent products that are encapsulated inside the liposomes. This further evidenced the synthesis of functional β-galactosidase through the reaction cascade. The time course data could be used to conduct a kinetic analysis of the reaction in liposomes (Figure 4B).

To verify the occurrence of the in-liposome self-encoding reaction, the histograms (frequency distribution) of product concentration in each liposome after 420 min are shown in Figure 3C. When Rep(+)Gal(−) RNA was used as a template, two distinct populations were observed with peaks at 1 and 20 μM, respectively. On the other hand, a single peak with a value of 0.6 μM was observed in the absence of the RNA or in the presence of RNA that encodes a defective β-subunit gene, ΔRep(+)Gal(−) RNA. Note that because these liposomes

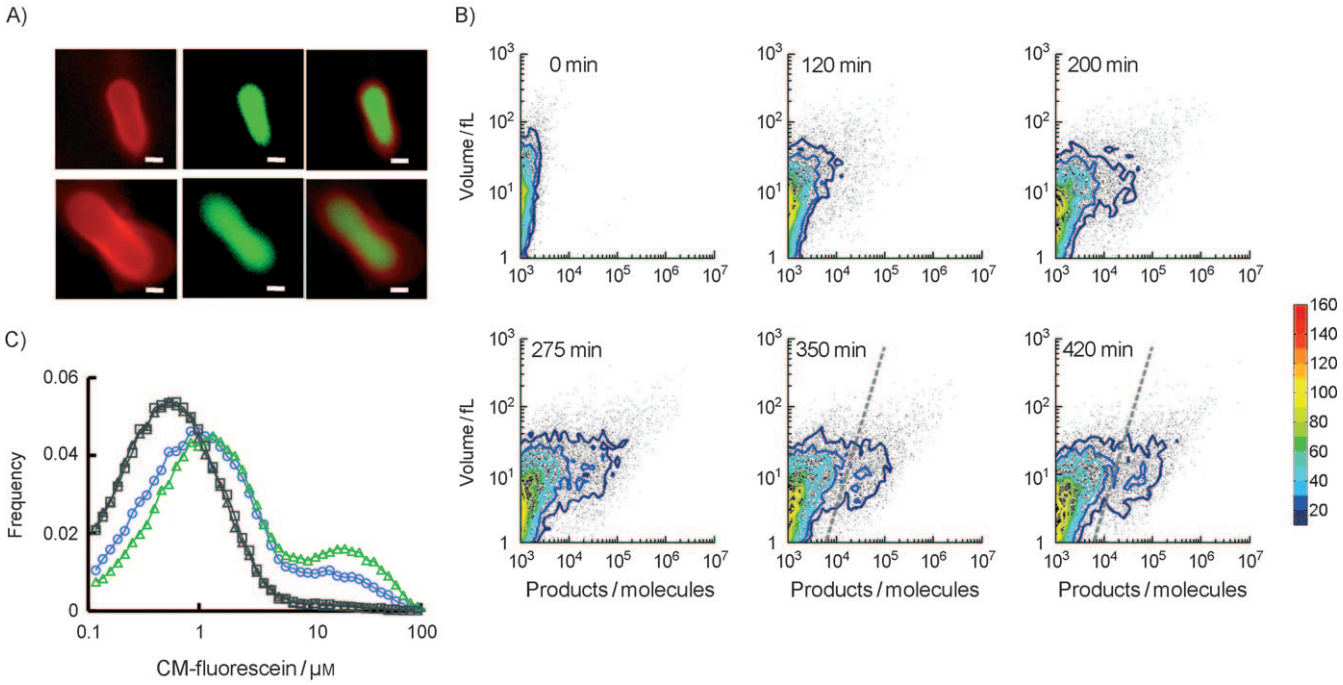

Figure 3. Replication of genetic information with self-encoded replicase in liposomes. A) The reaction within the liposomes by using Rep(+)Gal(−) RNA as a template observed under fluorescence microscopy. Left, red image (membranes); middle, green image (hydrolyzed CMFDG); right, overlay of left and middle images. Scale bar indicates 1 μm. B) Time course of the reaction analyzed by FACS. The results of 15 000 liposomes by using Rep(+)Gal(−) RNA. The results of FACS analysis of product (horizontal) and internal aqueous volume (vertical) of each liposome are shown. Dots represent the data of individual liposomes. Contour maps are overlaid. The frequency is depicted in color code. At 350 and 420 min, the number of reacted liposomes defined as those with the CM-fluorescein molecule (M_{CMF}) was $M_{CMF} \geq 6036\,V^{0.42}$ (right of the dashed lines), where V is the liposome internal volume (fL). The results with ΔRep(+)Gal(−) RNA supplemented with purified Qβ replicase and without template RNA are shown in Figure S2. C) Histogram (frequency distribution) of the CM-fluorescein concentration in liposomes after reaction at 37 °C for 420 min with Rep(+)Gal(−) RNA (○), ΔRep(+)Gal(−) RNA (△), ΔRep(+)Gal(−) RNA supplemented with purified Qβ replicase (△), or without template RNA (□). Distribution of concentration of CM-fluorescein in individual liposomes was estimated by using Equations (5) and (6).

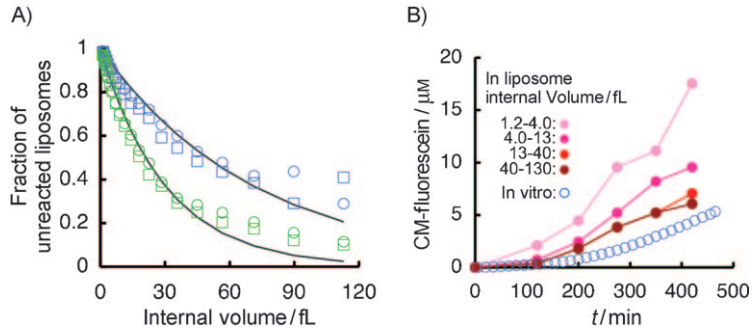

Figure 4. Statistical and kinetic analyses of the reaction in liposomes. A) The fractions of liposomes below the background (i.e., unreacted) at each liposome internal volume, by using Rep(+)Gal(−) RNA (blue) or ΔRep(+)Gal(−)RNA supplemented with purified Qβ replicase (green) at 350 (circles) and 420 min (squares) are shown. The solid line shows the results of curve fitting with Equation (1) when $P(A) = 1$ (Supporting Information). Note that the curve did not differ when $P(A) < 1$. B) Time courses of the reaction in liposomes with different internal volumes and in vitro. Concentration of the product (vertical axis) in liposomes was calculated as described in the Experimental Section.

cannot produce β-galactosidase, these peaks can be considered to be background signals. Nevertheless, when Qβ replicase was encapsulated together with ΔRep(+)Gal(−) RNA, two distinct populations appeared as with Rep(+)Gal(−) RNA. As shown in Figure 2B, when ΔRep(+)Gal(−) RNA was supplemented with Qβ replicase, β-galactosidase was produced, and thus an increase in fluorescence was observed (Figure 2B, green triangles). Therefore, the population with a peak at 20 μM was Qβ replicase dependent. These results suggested that the population with a peak at 20 μM, which was obtained with Rep(+)Gal(−) RNA was likely to be those that had succeeded in carrying out the gene replication reaction, whereas the other at around 0.6 to 1 μM was those that had not. Whereas there were differences in the background signal (peaks at 0.6 and 1 μM) among the samples, these differences were very small compared to the populations that contained synthesized β-galactosidase, and are thus negligible. These results indicated that within liposomes that exhibit β-galactosidase activity, the RNA was replicated by the replicase that was translated from it.

Reaction efficiency of the self-encoding system in liposomes

In this section, we describe the estimation of the reaction efficiency in liposomes. The reacted liposomes are defined as those on the right of the dashed lines of Figure 3B and Figure S2, reactions with Rep(+)Gal(−) RNA and ΔRep(+)Gal(−) RNA that was supplemented with Qβ replicase occurred in

only 13% and 22% of liposomes, respectively. Furthermore, the fractions of unreacted liposomes ($P_{unreact}$) at 350 and 420 min were dependent on the liposome internal volume (Figure 4A). We also found that $P_{unreact}$ did not differ between 350 and 420 min (Figure 4A). This was attributed to inactivation of the translation machinery, which has been reported previously for the in vitro translation system[9] that was used in this study and also observed in Figure 2B. Note that this does not indicate inactivation of β-galactosidase—in fact, β-galactosidase was still functional up to 420 min (Figure 4B)—but this means that there is no newly synthesized β-galactosidase after 350 min. It is also important to note that $P_{unreact}$ of ΔRep(+)Gal(−) RNA supplemented with Qβ replicase (green) was smaller than that of Rep(+)Gal(−) RNA (blue) in all liposome internal volumes (Figure 4A).

For the gene replication reaction in liposomes to be detected as β-galactosidase activity, the liposomes should contain more than one molecule of each of the components necessary for the cascade reaction. Under such conditions, these components should then produce more than one molecule of β-galactosidase from the template RNA. Note that FACS analysis allows us to detect the presence of a single molecule of β-galactosidase (T. Sunami, K. Hosoda, H. Suzuki, T. Matsuura, T. Yomo, unpublished results). By assuming that all stochastic processes are Poisson processes,[16] $P_{unreact}$ can be written as in Equation 1 (for details see the Supporting Information):

$$P_{nonreact} = 1 - P_{react} = \begin{cases} e^{-\gamma V} & (P(A) = 1) \\ 1 - (1 - e^{-\gamma V}) \times \prod_{i=1}^{n} (1 - e^{-C_i V}) & (P(A) < 1) \end{cases} \quad (1)$$

where n is the number of components necessary to be encapsulated in liposomes for the reaction, C_i (molecules/fL) is the average concentration of the i-th component ($i = 1, 2, \ldots, n$), V (fL) is the liposome internal volume, $P(A)$ is the probability of all the components that are necessary for the translation and the replication reaction being encapsulated in liposomes with an internal volume V (Supporting Information), and γ (events/fL) is the average frequency of successful replication reaction per fL. A successful replication reaction is equivalent to the production of β-galactosidase. Equation (1) suggests that $P_{unreact}$ is likely to exhibit dependence on the liposome internal volume, which was in fact the case (Figure 4A).

We then fit the data (Figure 4A) with Equation (1) to obtain the value γ. For the fit with Equation (1), we considered C_i to be a common parameter between the two sets of data (reaction with Rep(+)Gal(−) RNA and ΔRep(+)Gal(−) RNA supplemented with Qβ replicase) and γ to be a parameter that was different between the two sets. This was because the reaction does not require Qβ replicase to be encapsulated in the liposomes at the initial stage; instead it is synthesized during the reaction. Therefore, addition of Qβ replicase is unrelated to the C_i value, but rather can affect the reaction efficiency γ by increasing the concentration of the intermediate of the reaction. By fitting the experimental results by using Equation (1), we obtained $\gamma = 0.014$ (events per fL) for the reaction with Rep(+)Gal(−) RNA, regardless of the value of $P(A)$, and this low reaction efficiency is the predominant determinant of the stochastic observation found in Figure 4A (see Supporting Information for details). Note that we obtained values of C_i that were always significantly larger than γ (for example, $C_1 = 191$ when $n = 1$). This observation was consistent with the fact that sufficiently high concentrations of all components that are required for the in-liposome self-replication reaction were used for encapsulation (see Experimental Section).

The value γ is equivalent to the number of template RNA molecules that produced the β-galactosidase because template RNA is the rate-limiting parameter in this reaction (Figure 2B). That is, γ represents the number of RNA molecules among the initially encapsulated RNA per femtoliter that gave rise to the production of β-galactosidase. Because we used 30 nM (approximately 20 molecules/fL) RNA as a template for the in-liposome reaction, the γ value indicates that only one in 1400 RNA molecules (0.014/20 = 1/1400) could produce functional β-galactosidase in the liposomes.

Kinetic analysis of in-liposome replication reaction of genetic information with self-encoded replicase

Liposomes provide a reaction environment that is very different from that in the test tube because the volume is typically of the femtoliter order, and the molecules that are encapsulated in liposomes are in the presence of millimolar concentrations of lipids. Therefore, we compared the performance of the replication reaction of the genetic information with self-encoded replicase between in liposome and in vitro (that is, in the test tube) systems (Figure 4B). The reaction was conducted in liposomes with a wide range of internal volumes (Figure 3B). To investigate the effects of the volume on the internal reactions, the time courses of the reaction in liposomes with different volumes were plotted (Figure 4B). The time course was obtained by using the median values of the CM-fluorescein concentration that was obtained for each liposome internal volume (1.2–4.0, 4.0–13, 13–40, 40–130 fL) at different reaction times.

We first found that the reaction proceeds faster in smaller liposomes (Figure 4B). By using $\gamma = 0.014$ (events/fL), the average number of reactions that occurred in reacted liposomes (N_v) can be written as in Equation 2:

$$N_v = \frac{\gamma V}{1 - e^{-\gamma V}} \quad (2)$$

From Equation (2), the number of reactions that occurred in liposomes that were smaller than 13 fL was almost 1. Therefore, the number of β-galactosidase molecules that was produced should be the same in these liposomes. As a consequence, the same number of CM-fluorescein molecules was produced, smaller liposomes will give higher concentrations of CM-fluorescein, and thus faster production of fluorescence was observed, whereas the reaction kinetics in larger liposomes

became increasingly similar to those in vitro. These observations indicated that the in liposome self-encoding system was functioning as designed.

Thus, the statistical and kinetic properties of the self-encoding system were elucidated based on the fluorescent signal in liposomes that was measured by FACS. From the results of statistical analysis, we were able to estimate the reaction efficiency in liposomes. The results of the kinetic analysis indicated that the reaction proceeded faster in smaller liposomes, and the reaction kinetics in larger liposomes became increasingly similar to those in vitro; this indicates that our system was functioning as designed.

Discussion

Since the pioneering work of Spiegelman[13a,17] and Biebricher and Eigen[18] that showed the emergence of evolved RNA sequences through the RNA replication reaction with Qβ replicase, externally added purified Qβ replicase and its specific template RNA have been used for in vitro replication as a minimal model of life.[4a,13a,17] As opposed to using externally added replicase for the RNA replication reaction, in the present study the RNA was replicated by the Qβ replicase encoded on the RNA molecule itself, and can therefore be designated as a self-encoding system. In this regard, our system is distinct from the replicating systems that were reported previously.[4a,13a,17] Encapsulation of our self-encoding system in liposomes is essential for its evolvability, which is one of the basic properties of living systems.[1] Evolvability requires two processes: genetic diversification and selection based on phenotype. The Qβ replicase used here has a high error rate,[19] and thus the self-encoding system can produce genetic diversity directly. Moreover, compartmentalization by liposomes fulfills the linkage between the genotype and phenotype for selectability. Hence, this is the first report of an artificial multicomponent gene replication system with a self-encoded replicase in which the information unit (genotype) and the functional unit that replicates the information (phenotype) are encoded on different molecules, and importantly has the potential to evolve by being compartmentalized in liposomes, a biologically relevant environment. In this regard, our system is distinct from single-component replication systems that have been reported previously[17,20] in which both the genotype and phenotype are encoded on the same molecule, which requires the assistance of an exogenous mutator to produce genetic diversity, but does not require compartmentalization for genotype–phenotype linkage.

How many components and how much complexity were required for the in-liposome self-encoding system? We can provide the precise number of components in this system: one RNA sequence, 36 proteins, the ribosome, tRNAs, small chemical compounds (e.g., NTPs, amino acids), and three lipids. The schematic of the system is shown in Figure S3. The total number of 144 gene products (3 rRNAs,[21] 46 tRNAs,[9] 55 ribosomal proteins,[21] 38 proteins for protein synthesis (counting heterodimers as two[9]), one reporter protein, and one RNA replicase) is comparable with that of the proposed minimal cellular life (approximately 150), which can grow with only small-molecule nutrients.[1,2c] Our system is composed of only the defined components, which provides the possibility of modulating the system as desired, and is thus distinct from the self-encoding system by using bacterial cells.[22]

Therefore, our simplified system might be developed further by merging with self-replicating liposome technology[4] to add the basic characteristics of life to our system –- self-maintenance and self-reproduction, that is, duplication of the whole system[1,2] –- and simultaneously overcome the limit of our system that is caused by substrate depletion or system inactivation. These can be achieved by first sorting the liposomes with a higher reaction efficiency by FACS, and then fusing these with liposomes that carry components that are necessary for the reaction (for example, amino acids and proteins), thereby providing resources, then duplicating the system by division. Note that the liposomes that were used in this study were prepared by the freeze-dry method, which generates multilamellar vesicles,[15a] and further modification of liposome preparation might be required to merge our system with the self-replicating liposome technology. In this way, our system can be evolved and sustained continuously.

Experimental Section

Chemicals, biological materials, and plasmids: The plasmid pUC-MDV-LR[13c] was kindly provided by Dr. Y. Inokuchi (Teikyo University). The purified Qβ replicase was prepared as described previously.[12] 1-Palmitoyl-2-oleoyl-sn-phosphatidylcholine (POPC) was purchased from Avanti Polar Lipids (Alabaster, AL, USA). Cholesterol was purchased from Nacalai Tesque(Kyoto, Japan), and distearoyl phosphatidylethanolamine-polyethylene glycol 5000 (DSPE-PEG5000) was kindly provided by the NOF Corporation (Tokyo, Japan). BODIPY-labeled PC (**2a** in ref. [14]) was a kind gift from Dr. Sugawara (University of Tokyo). R-Phycoerythrin (PE) and CMFDG were purchased from Molecular Probes (Invitrogen). With the exception of the ribosomal protein S1, the proteins that were used for the in vitro translation system were purchased from Post Genome Institute Co., Ltd (Tokyo, Japan). S1 protein was overexpressed and purified from *E. coli* cells as described previously.[23]

Construction of plasmids and preparation of template RNAs: The plasmid pUC-MDV-LR[13c] contained the sequence of MDV-poly RNA downstream of the T7 promoter sequence, a SmaI restriction site at the 3′-end of the MDV-poly sequence, and a BglII site within the MDV-poly RNA sequence for cloning.[13b] The plasmid pUC-MDVminus, which contained the complementary MDV-poly sequence downstream of the T7 promoter, was constructed by PCR amplification of the MDV-poly sequence from the plasmid pUC-MDV-LR, and then ligated between the BamHI and EcoRI restriction sites of the plasmid pUC-MDV-LR. Plasmids encoding Rep(+) RNA and its complementary strand were constructed by inserting the fragment that contained the Qβ replicase β-subunit sequence into the BglII sites of pUC-MDVminus and pUC-MDV-LR, respectively. These plasmids were then used to generate the standards for quantitative real-time PCR (RT-QPCR) by in vitro transcription as described previously.[13b] A plasmid encoding Rep(+)Gal(−) RNA was constructed as follows. The sequence of *lacZ* was amplified by PCR from the plasmid pSV-β-Galactosidase Vector (Promega), and the amplified product and the Qβ replicase β-subunit gene were inserted into the BglII restrictions site of pUC-MDV-LR. A plasmid

encoding ΔRep(+)Gal(−) RNA was constructed by deleting the 126-nt segment between the two SacI sites of the β-subunit gene of the plasmid that encoded Rep(+)Gal(−) RNA. DNA sequences of all plasmids were confirmed by DNA sequencing. Template RNAs were prepared by in vitro transcription as described previously.[13b] All template RNA sequences are given in Supporting Information.

Preparation of liposomes: Liposomes were prepared by the freeze-dried empty liposome (FDEL) method as described previously.[15] Briefly, the lipid mixture (1.2 μmol, mixed with a molar ratio of POPC/cholesterol/DSPE-PEG5000 = 58:39:3) was dissolved in CH_2Cl_2/Et_2O (1:1, *v/v*) and subjected to rotary evaporation in a pear-shaped flask under vacuum to yield a thin lipid film. Then deionized H_2O (100 μL) was added to the film under argon gas. After 15 min, the lipid film was vortexed to disperse the liposomes. The liposome dispersion was homogenized on ice by sonication with an ultrasonic disrupter (Tomy Seiko, Tokyo, Japan) and was extruded through a polycarbonate filter with a pore size of 0.4 μm (Whatman). The solution was then transferred to a tube and lyophilized in a freeze dryer (Labconco, Kansas City, MO, USA). The freeze-dried empty liposomes were stored at 20 °C under argon gas. Liposomes that were prepared by this method had the multilamellar membrane structures and complex structures as reported previously.[15a] Liposomes that were observed under fluorescence microscopy were prepared with a lipid mixture that contained BODIPY-labeled PC (about 1 mol%, compound **2a** in ref. [14]) to visualize the membrane. Note that liposomes that were used for FACS analysis did not contain BODIPY-labeled PC, but instead encapsulated R-phycoerythrin (PE), which was used to estimate the internal volume of individual liposomes.

Replication of genetic information with the self-encoded replicase: The gene replication reaction with self-encoded replicase was carried out by using template RNA (55 nM) and an in vitro translation system that had been reconstituted, with incubation at 37 °C. Unless otherwise specified, the standard in vitro translation systems contained $Mg(OAc)_2$ (13 mM), potassium glutamate (100 mM), spermidine (2 mM), dithiothreitol (1 mM), ATP (2 mM), GTP (2 mM), CTP (1 mM), UTP (1 mM), creatine phosphate (20 mM), 48 A260 units of tRNA mix, 10-formyl-5,6,7,8-tetrahydrofolic acid (10 ng mL^{-1}), each amino acid at 0.3 mM, 2-[4-(2-hydroxyethyl)-1-piperazinyl]ethanesulfonic acid (HEPES) (50 mM, pH 7.6), and enzyme mix. The enzyme mix contained ribosomes (1.2 μM), IF1 (2.5 μM), IF2 (0.21 μM), IF3 (0.95 μM), EF-G (3.2 μM), EF-Tu (12 μM), EF-Ts (8.2 μM), RF1 (0.25 μM), RF2 (0.24 μM), RF3 (0.17 μM), RRF (0.48 μM), AlaRS (725 nM), ArgRS (31 nM), AsnRS (380 nM), AspRS (127 nM), CysRS (24 nM), GlnRS (60 nM), GluRS (233 nM), GlyRS (87 nM), HisRS (8 nM), IleRS (396 nM), LeuRS (42 nM), LysRS (113 nM), MetRS (27 nM), PheRS (676 nM), ProRS (165 nM), SerRS (39 nM), ThrRS (85 nM), TrpRS (28 nM), TyrRS (7 nM), ValRS (17 nM), MTF (588 nM), creatine kinase (0.47 μM), myokinase (0.93 μM), nucleoside-diphosphate kinase (1.3 μM), pyrophosphatase (0.62 μM), ribosomal protein S1 (4.6 μM), and RNasin Plus RNase Inhibitor (1 U μL^{-1}; Promega). A liposome of a typical size (4 fL) is expected to carry 17 TyrSR molecules (the minimum among protein components), 28800 EF-Tu molecules (the maximum among protein components), 72 RNA templates, 2880 ribosomes, etc. It is worth noting that all components that were required for the reactions were likely to be encapsulated into liposomes (> 1 fL) as described in the Results section.

For the reaction with additional phenotype that was generated by insertion of the *lacZ* gene, the in vitro translation system that is described above was supplemented with template RNA (30 nM), CMFDG (100 μM), Alexa Fluor 647 (50 nM; Invitrogen), and $Mg(OAc)_2$ (11 mM). When the in vitro translation system was supplemented with purified Qβ replicase, 25 nM was used.

Reactions in liposomes were carried out as described previously.[15] Briefly, aliquots (25 μL) of the reaction mixture, which consisted of the in vitro translation system, CMFDG (100 μM), PE (400 nM), and template RNA (30 nM), were added to the lyophilized lipids at 4 °C. Subsequently, the liposomes were diluted with the in vitro translation system that lacked EF-G, EF-Tu, and EF-Ts to prevent protein synthesis from occurring outside the liposomes. To initiate the translation reaction, liposomes were incubated at 37 °C.

Characterization of RNA replication and protein synthesis reaction during the gene replication reaction with self-encoded replicase: Plus and minus-strand RNAs were measured by quantitative real-time PCR (RT-QPCR). Reverse transcription was performed as described below. The samples were diluted 10000-fold with H_2O, heated at 95 °C for 5 min to denature the double-stranded regions, and mixed with reverse transcriptase (PrimeScript RTase; Takara) and the primer, 5′-TAA GCG AAT GTT GCG AGC ACC TTG TAT GGT CCG TAA TCA C or 5′-TAA GCG AAT GTT GCG AGC ACG CTG CAA CGT AAT ACT ATA C for plus or minus-strand synthesis, respectively. Reactions were carried out in accordance with the manufacturer's instructions. cDNA samples were then diluted 100-fold with H_2O and mixed with qPCR MasterMix (Eurogentec, Seraing, Belgium), a dual-labeled probe, 5′-FAM-TGC CCT CGT CGG ATC GGT CCT AAT-BHQ-1 (Sigma), and primers 5′-TAA GCG AAT GTT GCG AGC AC and 5′-TGC CTA AAC AGC TGC AAC GT or 5′-CGC TCT CGG TCC CTT GTA TG for sense and antisense strand quantification, respectively. Reactions were carried out in accordance with the manufacturer's instructions by using a real-time PCR system (Mx3005P; Stratagene). The amount of RNA was measured by using the in-vitro-transcribed complementary strand as a standard.

The amounts of synthesized proteins were measured by incorporation of [^{35}S]-methionine during synthesis and measurement of the band intensity on SDS-PAGE. The band intensity that corresponded to the replicase β-subunit was quantified from the results of autoradiography. The amount of methionine that was incorporated into the β-subunit was calculated by comparing the band intensity to those of known concentrations of methionine, and the amount of synthesized β-subunit could then be calculated.

Reactions in liposomes were analyzed by using a FACSAria cytofluorometer (BD Biosciences, San Jose, CA, USA). The reaction inside the liposomes was observed by using a Nikon TE2000-PFS microscope that was equipped with an oil-immersion objective (Plan Apo 100×, NA 1.4). Fluorescence images were collected by using an Andor iXon DV887 EMCCD camera.

Quantitative analysis of the reaction in liposomes: The quantitative data for individual liposomes in the internal aqueous volume and the hydrolysis of CMFDG were estimated essentially as described previously.[15] Briefly, red ($FI_{R\text{-}obs}$) and green ($FI_{G\text{-}obs}$) fluorescence, which were derived from PE and CM-fluorescein, respectively, were obtained for individual liposomes by FACS. $FI_{R\text{-}obs}$ and $FI_{G\text{-}obs}$ were converted to FI_R and FI_G by using Equations (3) and (4), where FI_R and FI_G are the fluorescence intensities after correcting for the fluorescence spectra overlap. FI_R and FI_G were then converted to liposome internal volume (V (fL)) and amount of CM-fluorescein (M_{CMF} (molecules)) by using Equations (5) and (6), respectively, where C_{PE} is the concentration of PE (M).

$$FI_R = \frac{FI_{R\text{-}obs} - 0.25\,FI_{G\text{-}obs}}{1 - 0.25 \times 0.005} \quad (3)$$

$$FI_G = \frac{FI_{G\text{-obs}} - 0.005\,FI_{R\text{-obs}}}{1 - 0.25 \times 0.005} \quad (4)$$

$$V = \frac{FI_R}{7 \times 10^8\,C_{PE}} \quad (5)$$

$$M_{CMF} = 35.01\,FI_G \quad (6)$$

Time course of the reaction in liposomes: Median values of the CM-fluorescein concentration were first obtained for each liposome internal volume (1.2–4.0, 4.0–13, 13–40, 40–130 fL) from the data that is shown in Figure 3B, that is, the data after 420 min in which the translation reaction had stopped primarily due to inactivation of the translation system (Figure 2B). The rank in CM-fluorescein concentration (rank 1 = highest CM-fluorescein concentration) of the liposomes exhibiting the median value was then obtained. Then, the CM-fluorescein concentrations of the liposomes of identical rank to those above were obtained for each liposome internal volume at different reaction times.

Acknowledgements

We thank Drs. Y. Inokuchi and T. Ueda for helpful discussions and technical support; T. Nakaishi, K. Sato, J. Cho, and K. Obinata for help at the initial stage of the project; S. Tsuru for help in preparation of Figures; J. Ichinose for advice on microscopy; and Dr. M. Hanczyc for comments on the manuscript. This research was supported in part by "Special Coordination Funds for Promoting Science and Technology: Yuragi Project" and "Global COE (Centers of Excellence) Program" of the Ministry of Education, Culture, Sports, Science, and Technology, Japan.

Keywords: artificial cells · in vitro translation · liposomes · Qβ replicase · RNA

[1] a) P. L. Luisi, F. Ferri, P. Stano, *Naturwissenschaften* **2006**, *93*, 1–13; b) P. L. Luisi, T. Oberholzer, A. Lazcano, *Helv. Chim. Acta* **2002**, *85*, 1759–1777.
[2] a) J. W. Szostak, D. P. Bartel, P. L. Luisi, *Nature* **2001**, *409*, 387–390; b) A. C. Forster, G. M. Church, *Mol. Syst. Biol.* **2006**; DOI: 10.1038/msb4100090; c) D. Deamer, *Trends Biotechnol.* **2005**, *23*, 336–338.
[3] A. D. Bangham, M. M. Standish, J. C. Watkins, *J Mol Biol.* **1965**, *13*, 238–252.
[4] a) T. Oberholzer, R. Wick, P. L. Luisi, C. K. Biebricher, *Biochem. Biophys. Res. Commun.* **1995**, *207*, 250–257; b) K. Takakura, T. Sugawara, *Langmuir* **2004**, *20*, 3832–3834; c) M. M. Hanczyc, S. M. Fujikawa, J. W. Szostak, *Science* **2003**, *302*, 618–622.
[5] a) A. C. Chakrabarti, R. R. Breaker, G. F. Joyce, D. W. Deamer, *J. Mol. Evol.* **1994**, *39*, 555–559; b) T. Oberholzer, K. H. Nierhaus, P. L. Luisi, *Biochem. Biophys. Res. Commun.* **1999**, *261*, 238–241; c) W. Yu, K. Sato, M. Wakabayashi, T. Nakaishi, E. P. Ko-Mitamura, Y. Shima, I. Urabe, T. Yomo, *J. Biosci. Bioeng.* **2001**, *92*, 590–593; d) A. Fischer, A. Franco, T. Oberholzer, *ChemBioChem* **2002**, *3*, 409–417; e) T. Oberholzer, P. L. Luisi, *J. Biol. Phys.* **2002**, *28*, 733–744; f) S. M. Nomura, K. Tsumoto, T. Hamada, K. Akiyoshi, Y. Nakatani, K. Yoshikawa, *ChemBioChem* **2003**, *4*, 1172–1175; g) G. Murtas, Y. Kuruma, P. Bianchini, A. Diaspro, P. L. Luisi, *Biochem. Biophys. Res. Commun.* **2007**, *363*, 12–17.
[6] V. Noireaux, A. Libchaber, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 17669–17674.
[7] K. Ishikawa, K. Sato, Y. Shima, I. Urabe, T. Yomo, *FEBS Lett.* **2004**, *576*, 387–390.
[8] Y. Kuruma, *Origins Life Evol. Biospheres* **2007**, *37*, 409–413.
[9] Y. Shimizu, A. Inoue, Y. Tomari, T. Suzuki, T. Yokogawa, K. Nishikawa, T. Ueda, *Nat. Biotechnol.* **2001**, *19*, 751–755.
[10] a) A. D. Griffiths, D. S. Tawfik, *Trends Biotechnol.* **2006**, *24*, 395–402; b) O. J. Miller, K. Bernath, J. J. Agresti, G. Amitai, B. T. Kelly, E. Mastrobattista, V. Taly, S. Magdassi, D. S. Tawfik, A. D. Griffiths, *Nat. Methods* **2006**, *3*, 561–570.
[11] a) T. Blumenthal, G. G. Carmichael, *Annu. Rev. Biochem.* **1979**, *48*, 525–548; b) A. B. Chetverin, A. S. Spirin, *Prog. Nucleic Acid Res. Mol. Biol.* **1995**, *51*, 225–270.
[12] H. Kita, J. Cho, T. Matsuura, T. Nakaishi, I. Taniguchi, T. Ichikawa, Y. Shima, I. Urabe, T. Yomo, *J. Biosci. Bioeng.* **2006**, *101*, 421–426.
[13] a) D. L. Kacian, D. R. Mills, F. R. Kramer, S. Spiegelman, *Proc. Natl. Acad. Sci. USA* **1972**, *69*, 3038–3042; b) K. Hosoda, T. Matsuura, H. Kita, N. Ichihashi, K. Tsukada, T. Yomo, *J. Biol. Chem.* **2007**, *282*, 15516–15527; c) Y. Inokuchi, M. Kajitani, A. Hirashima, *J. Biochem.* **1994**, *116*, 1275–1280; d) Y. Wu, D. Y. Zhang, F. R. Kramer, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 11769–11773; e) L. Ryabova, E. Volianik, O. Kurnasov, A. Spirin, Y. Wu, F. R. Kramer, *J. Biol. Chem.* **1994**, *269*, 1501–1505.
[14] T. Toyota, H. Tsuha, K. Yamada, K. Takakura, K. Yasuda, T. Sugawara, *Langmuir* **2006**, *22*, 1976–1981.
[15] a) K. Sato, K. Obinata, T. Sugawara, I. Urabe, T. Yomo, *J. Biosci. Bioeng.* **2006**, *102*, 171–178; b) T. Sunami, K. Sato, T. Matsuura, K. Tsukada, I. Urabe, T. Yomo, *Anal. Biochem.* **2006**, *357*, 128–136.
[16] D. M. Rissin, D. R. Walt, *Nano Lett.* **2006**, *6*, 520–523.
[17] D. R. Mills, R. L. Peterson, S. Spiegelman, *Proc. Natl. Acad. Sci. USA* **1967**, *58*, 217–224.
[18] C. K. Biebricher, M. Eigen, W. C. Gardiner Jr. , *Biochemistry* **1985**, *24*, 6550–6560.
[19] J. W. Drake, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 4171–4175.
[20] a) T. Tjivikua, P. Ballester, J. Rebek Jr. , *J. Am. Chem. Soc.* **1990**, *112*, 1249–1250; b) D. Sievers, G. von Kiedrowski, *Nature* **1994**, *369*, 221–224; c) W. S. Zielinski, L. E. Orgel, *Nature* **1987**, *327*, 346–347; d) N. Paul, G. F. Joyce, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 12733–12740; e) D. H. Lee, J. R. Granja, J. A. Martinez, K. Severin, M. R. Ghadiri, *Nature* **1996**, *382*, 525–528; f) R. Issac, J. Chmielewski, *J. Am. Chem. Soc.* **2002**, *124*, 6808–6809.
[21] M. Kitagawa, T. Ara, M. Arifuzzaman, T. Ioka-Nakamichi, E. Inamoto, H. Toyonaga, H. Mori, *DNA Res.* **2005**, *12*, 291–299.
[22] a) F. J. Ghadessy, N. Ramsay, F. Boudsocq, D. Loakes, A. Brown, S. Iwai, A. Vaisman, R. Woodgate, P. Holliger, *Nat. Biotechnol.* **2004**, *22*, 755–759; b) F. J. Ghadessy, J. L. Ong, P. Holliger, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 4552–4557.
[23] A. R. Subramanian, *Prog. Nucleic Acid Res. Mol. Biol.* **1983**, *23*,101–142.

Received: May 28, 2008
Published online on September 11, 2008

