Review in Advance first posted online on March 3, 2014. (Changes may still occur before final publication online and in print.)

Progress Toward Synthetic Cells

J. Craig Blain and Jack W. Szostak

Howard Hughes Medical Institute, Department of Molecular Biology, and Center for Computational and Integrative Biology, Massachusetts General Hospital, Boston, Massachusetts 02114; email: craig@jcraigblain.com, szostak@molbio.mgh.harvard.edu

Annu. Rev. Biochem. 2014. 83:11.1–11.26

The *Annual Review of Biochemistry* is online at biochem.annualreviews.org

This article's doi: 10.1146/annurev-biochem-080411-124036

Keywords

protocells, artificial life, origin of life, self-replication, evolution

Abstract

The complexity of even the simplest known life forms makes efforts to synthesize living cells from inanimate components seem like a daunting task. However, recent progress toward the creation of synthetic cells, ranging from simple protocells to artificial cells approaching the complexity of bacteria, suggests that the synthesis of life is now a realistic goal. Protocell research, fueled by advances in the biophysics of primitive membranes and the chemistry of nucleic acid replication, is providing new insights into the origin of cellular life. Parallel efforts to construct more complex artificial cells, incorporating translational machinery and protein enzymes, are providing information about the requirements for protein-based life. We discuss recent advances and remaining challenges in the synthesis of artificial cells, the possibility of creating new forms of life distinct from existing biology, and the promise of this research for gaining a deeper understanding of the nature of living systems.

Contents

INTRODUCTION

Models of the first primitive cells, or protocells, are being created as a way of studying the origin and early evolution of life. Similarly, simplified models of modern cells are being used to test our understanding of the requirements for cellular function, including growth and division. Such divergent goals are inspiring an increasing number of laboratories to work toward the synthesis of artificial cells of widely varying composition and complexity. These efforts are united by the goals and challenges of assembling nonliving components into living cells.

In this review, we focus on bottom-up approaches to the construction of artificial cells from molecular components or subsystems (**Figure 1**). Such approaches are well suited to purposes ranging from modeling the earliest life forms to developing minimal protein-based cells to exploring the creation of cells based on nonbiological components. We begin by considering the genomes of artificial cells, which range from short oligonucleotides for model protocells to much longer RNAs for artificial cells that incorporate viral polymerases to even larger DNA genomes that approach the size of bacterial genomes. We then review recent work on compartments, the supramolecular structures that define the cellular nature of life. Most compartment boundaries are vesicles composed of bilayer membranes, but they range widely in composition from simple fatty acids to more complex lipids and to various synthetic nonbiological lipids. Finally, we discuss efforts to integrate replicating genomes with replicating compartments to generate artificial cells that are capable of sustained reproduction and Darwinian evolution. We also discuss efforts to generate cell-like structures based on nonbiological genetic materials and nonbiological forms of compartmentalization. We do not discuss the top-down approach of systematically removing genes from extant organisms to achieve a minimal genome (1), nor do we consider the creation of new types of cells by genome synthesis or extensive genome editing (2, 3).

Protocell: a simple cell-like entity (with a compartment and genetic material) capable of self-replication, metabolism, and Darwinian evolution

Ribozyme: an RNA molecule that can fold into a three-dimensional structure and catalyze reactions, similarly to a protein enzyme

GENOMES AND GENOME REPLICATION

To propagate indefinitely, a synthetic cell must be able to replicate its own genetic material. In this section, we first discuss nonenzymatic and ribozyme-catalyzed RNA replication because these are the two proposed processes for genome replication in RNA-based protocells. We then discuss approaches to protein-catalyzed genome replication in artificial cells, which typically take advantage of viral RNA or DNA polymerases because of their simplicity in comparison to cellular DNA replication systems. **Supplemental Figure 1** (follow the **Supplemental Material link** from the Annual Reviews home page at **http://www.annualreviews.org**) shows the chemical structures of the different nucleotides, base pairs, and nucleic acids discussed here.

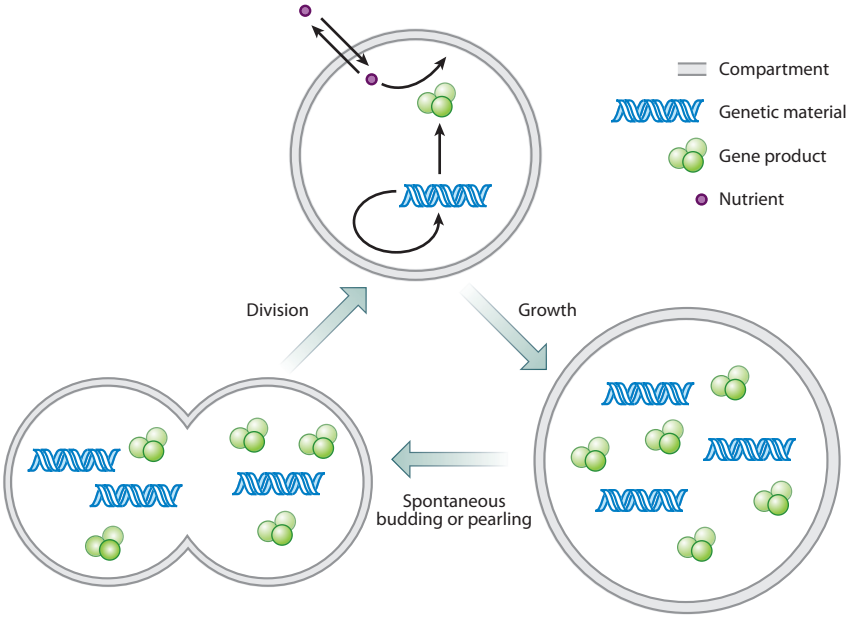

Figure 1

A cycle of growth and division for a minimal cell. The genetic material, typically a nucleic acid, encodes or directs the synthesis of gene products, which can be translated proteins or a folded version of the genetic material in the case of ribozyme-based cells. The gene products catalyze replication of the genetic material and perform basic cellular metabolism, for example, synthesis of compartment components. With highly activated nucleotides, catalysis of genetic replication may not be necessary. Meanwhile, the compartment, which is typically a phospholipid or a fatty acid vesicle, grows through the internal synthesis or external addition of lipids or by competing with other cells for their lipids. As the cell grows, the system becomes unstable and the cell divides into small daughter cells that repeat the process.

Nonenzymatic Template-Directed Replication

Nucleic acids are by far the most promising candidates for the genetic material of protocells because they can direct their own replication through template copying (**Figure 2*a***). The question of how nucleic acids could replicate before the evolution of genetically encoded protein enzymes is central to research on the origin of life (4–7). After briefly reviewing early work on this topic, we discuss recent advances in nonenzymatic RNA replication and consider the approaches being taken to overcome the remaining problems with what appears to be the simplest form of genetic replication. Finally, we review recent research in which, by dropping the constraints of prebiotic plausibility, investigators have used chemically modified nucleic acids to rapidly copy mixed sequence templates containing all four bases (8–10).

The template-enhanced nonenzymatic synthesis of nucleic acids was first demonstrated in the 1960s (11–13), but the reactions were slow and inefficient and, in the case of RNA synthesis, led to a predominance of nonnatural 2′–5′-linked material. Extensive subsequent research by Orgel and colleagues (14–17) showed that ribonucleotides activated with 2-methylimidazole (2-MeIm) as a leaving group can be used to copy short C-rich RNA templates on a timescale of hours or days without enzymes. Although the products were predominantly 3′–5′ linked, the reaction depended strongly on sequence and base composition, and A and U residues were copied particularly slowly (18) and with poor fidelity (19). Even all-G:C sequences could not be

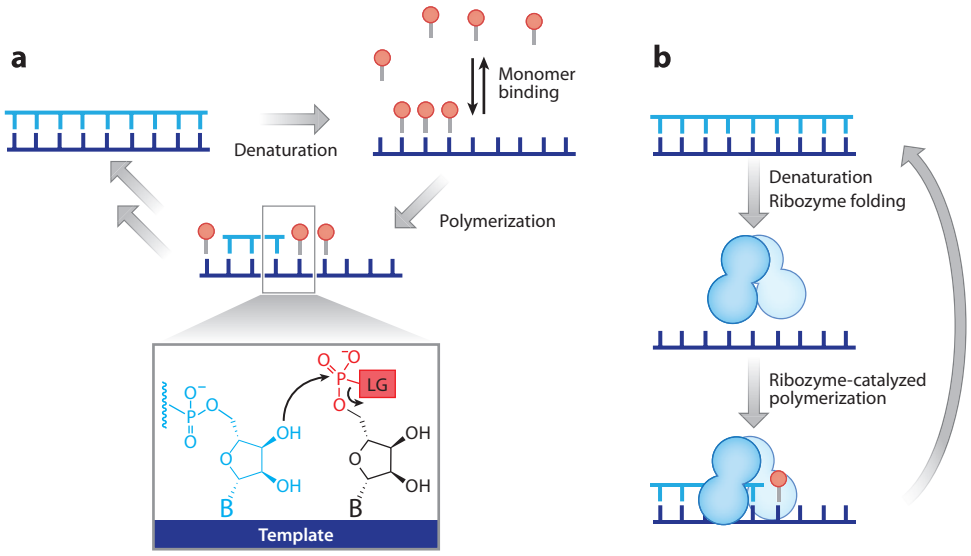

Figure 2

Pathways for the self-replication of nucleic acids. (*a*) In nonenzymatic nucleic acid self-replication, activated nucleotides or short oligonucleotides bind to a complementary base in a single-stranded template. The 3′-hydroxyl group of a bound residue attacks the 5′-phosphate of an adjacent residue to displace a leaving group, typically 2-methylimidazole, and forms a new phosphodiester bond. This reaction is repeated at each position to copy the entire sequence. Chemical modifications to the natural RNA structure can enhance the rate and fidelity of this process. (*b*) In ribozyme-catalyzed replication, a single-stranded RNA molecule folds into a tertiary structure that can catalyze the template-copying reaction. In both models, the stability of double-stranded RNA, hydrolysis of RNA and activated nucleotides, and lack of sequence-general copying limit the potential for self-replication.

copied efficiently; for example, the best yield obtained in efforts to copy a 14-mer all-G:C template was 2%. For these and other reasons, including incomplete regiospecificity; the need for primers; the problem of strand separation; and the lack of mild, specific chemistry for monomer activation (and reactivation following hydrolysis), Orgel and colleagues, and indeed most researchers in the field, eventually became convinced that nonenzymatic RNA replication was not chemically realistic (20, 21).

Nucleotide activation. The 2′–3′ cyclic phosphate nucleotides, although attractive from the point of view of prebiotic synthesis, are insufficiently activated to lead to efficient template copying (22, 23), whereas the nucleoside triphosphates (NTPs) used in biology are too unreactive for nonenzymatic polymerization (24). Nucleotides activated with 2-MeIm, known as phosphor-2-methylimidazolides, react relatively quickly to preferentially provide the natural and more stable (25) 3′–5′ linkage (26). Interestingly, other imidazole derivatives do not react as quickly (26), and activation with imidazole itself preferentially yields 2′–5′ linkages (27). To date, we have no mechanistic understanding of why 2-MeIm-activated nucleotides react so quickly and regiospecifically with the 3′-hydroxyl group, although sterics, nucleophile and activating-group acid dissociation constants (pK_as), and interactions between the activating groups of stacked monomers (17) may play a role. The rate of addition of guanosine 5′-phosphor-2-methylimidazolide to a primer on a poly(C) template depends on the length of the template, the monomer concentration, and the magnesium concentration, but it is typically ∼1 h^{-1} (28), whereas the rate of monomer hydrolysis is ∼0.02–0.05 h^{-1} (29).

Although 2-MeIm is the most widely used activating group, the search for replacements with superior properties continues. Richert and colleagues (30–32) have extensively used 1-hydroxy-7-azabenzotriazole (HOAt), which was originally developed for peptide coupling (33). They achieved higher rates with this leaving group than with 2-MeIm when adding single nucleotides to a 3′-amino-modified primer on a DNA template (30, 31) and in an RNA system (32). It is not yet clear how the two activating groups compare for the addition of multiple nucleotides, or how faster monomer hydrolysis and the requirement for higher pH for HOAt-activated nucleotides (32) would affect their utility in a self-replicating system.

Nucleophilic catalysts have been employed to generate highly reactive nucleotides in situ from 2-MeIm- or HOAt-activated monomers. In particular, *N*-alkylated imidazole derivatives (8, 9) and pyridine (34) have been used to accelerate nonenzymatic primer extension reactions with amino-terminated primers, but similar catalysis has not been demonstrated with an RNA system. Adenine derivatives such as 1-methyladenine are also good potential activating groups (35), but their use in template copying has not been explored. A thorough understanding of the mechanism of nonenzymatic primer extension would allow for the rational design and optimization of improved activating groups, as opposed to the ad hoc screening that is currently employed.

Base-pairing and fidelity. Not only do polymerases accelerate the phosphoryl transfer reaction; they also increase the affinity and specificity of the nucleotide–template interaction (36; but see Reference 37 for a possible simple mechanism suggesting that this task may not be difficult to accomplish without enzymes). Without polymerases, fidelity and sequence context become significant challenges to replication. G and C residues are copied much more quickly than A and U residues due to stronger base-pairing, and the presence of multiple A or U residues in a row in a template drastically reduces copying efficiency (18). Furthermore, the comparable strength of a G:U wobble pair versus an A:U base pair (38) leads to significant misincorporation of G across from U (39).

Investigators have used two approaches to overcome these problems. In the first, modified base pairs were used in place of A:U. For example, the 5-propynyluracil (U^P)–diaminopurine (D) base pair is similar to A:U but has three hydrogen bonds and improved base stacking. This base pair has been used to achieve comparable rates of nonenzymatic polymerization to the G:C base pair in multiple systems (8, 40). However, because the propynyl group lowers the pK_a at position N3 of the U ring (41) it may stabilize the enol tautomer and thus strengthen Watson–Crick-like G:U mispairing, similar to the 5-bromo substitution (42). Conversely, 2-thio substitution on U or T strengthens base-pairing to A but weakens wobble pairing to G due because sulfur is a weak H-bond acceptor (43, 44) and may therefore improve both rate and fidelity. Many other alternate base pairs are also possible (45). The second approach is to use short oligonucleotides instead of or in addition to monomers (46, 47), given that the cooperativity of base-pairing helps stabilize individual A:U pairs. High fidelity can be achieved in nonenzymatic oligonucleotide ligation at elevated temperatures (48), and oligonucleotide ligation is selective for 3′–5′ linkages (49). Investigators have used activated trimers to demonstrate replication of a hexamer template (50), but the poor efficiency of ligation has so far prevented this approach from being extended to longer templates.

After the incorporation of a mismatched nucleotide in a template-copying reaction, both the rate (19) and the fidelity (51) of the next addition decrease. This stalling effect causes accurate template copies to be produced more quickly than mutant sequences, thereby increasing the apparent fidelity of the replication process and the length of sequences that can be maintained through replication. Conversely, mutant products tend to incorporate stretches of mismatches and may therefore cause "leaps" through sequence space, which could facilitate the evolution of novel structures (51).

Backbone modifications. Changing the hydroxyl nucleophile of a ribo- or deoxyribonucleotide to a more reactive amine is an effective way to enhance the rate of nonenzymatic polymerization. The Orgel group (52–55) initially explored both the 2′-amino and 3′-amino modifications of ribonucleotides in template-copying reactions and found that they were more reactive than their hydroxyl counterparts. More recently, our group and Richert's group (56, 57) studied the 2′- and 3′-amino-dideoxyribonucleotides in primer extension reactions. Using primers with a 3′-amino terminus and so-called helper oligonucleotides that assist monomer binding, Richert and colleagues achieved rapid and accurate template-directed single-nucleotide primer extension. By applying these lessons to an all RNA system, these authors efficiently and accurately copied short mixed-sequence RNA templates (58), although this process was slow.

Both 2′-amino-modified (8) and 3′-amino-modified (9) nucleotides can copy an RNA C_4 template in >80% yield in ∼10 min. The reaction rates are sensitive to the template structure; A-form helices [RNA and locked nucleic acid (LNA)] yield faster rates than do DNA templates in both cases. In the 2′-amino system, A:U base pairs were not copied efficiently, so U^P:D pairs were used instead. Surprisingly, in the 3′-amino system, both A_4 and T_4 templates were copied efficiently. In a further step toward self-replication, we used 3′-amino nucleotides to copy mixed sequence templates composed of phosphoramidate DNA (3′-NP-DNA), the product of 3′-amino nucleotide polymerization (59). The G:T wobble base pair yielded significant misincorporation in this system, but the use of the 2-thio modification of T, which stabilizes pairing to A and weakens pairing to G (43), improved both the fidelity and the rate of copying (59). Although 3′-NP-DNA is attractive as a possible genetic polymer for protocells, the copying of longer mixed sequence templates has yet to be demonstrated.

Many modified-backbone nucleic acids that are capable of Watson–Crick base-pairing have been synthesized (60), but only a few have been examined from the point of view of their ability to support self-replication (61). Threose nucleic acids (TNAs) form stable Watson–Crick duplexes with complementary TNA, RNA, and DNA strands (62) and have been used as templates in the nonenzymatic polymerization of ribonucleotides (63). Furthermore, an aptamer made of TNA has been isolated by in vitro selection (64). Pyranosyl-RNA (p-RNA) has been used as a template for the nonenzymatic ligation of p-RNA oligonucleotides (23). Interestingly, hexitol nucleic acid (HNA) (65) and altritol nucleic acid (ANA) (66), both of which form single strands that are preorganized into an A-form geometry, are superior to RNA as templates for the nonenzymatic polymerization of activated ribonucleotides (67). However, the corresponding activated HNA and ANA monomers do not polymerize efficiently, indicating that subtle differences in monomer structure and reactivity can affect the reaction (67).

Peptide nucleic acid (PNA) is based on a peptide backbone (68) but can form Watson–Crick base pairs with many nucleic acids (69). The peptide backbone means that the whole range of peptide chemistry can be employed to demonstrate template copying. For example, PNA dimers have been used in DNA template–directed polymerization through 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDAC)-mediated peptide bond formation (70), and reductive amination is also an effective way to polymerize PNA pentamers with aldehyde and amino termini on a DNA template (71). PNA monomers, however, are prone to cyclization (70). Conversely, a PNA C_{10} oligomer is an effective template for the polymerization of guanosine 5′-phosphorimidazolide (72). Interestingly, PNA also enables the template-directed attachment of nucleobase units onto a preformed backbone. Investigators have demonstrated this process by using reductive amination to make the process irreversible (73), and dynamically by using thioester linkages, wherein the base composition changed in response to alterations in template sequences (74). Because it is uncharged, PNA based on the original aminoethylglycyl backbone is prone

to aggregation (75, 76); however, a version of PNA based on a repeating dipeptide motif in which alternating amino acids have charged side chains (e.g., Asp) and nucleobase side chains can form stable, soluble duplexes (77).

Other challenges. There are many challenges on the path to nonenzymatic self-replication other than the basic problems of rate and fidelity (78). In particular, strand separation (and thus continuing cycles of replication) can be difficult. Even the simplest ribozymes are typically more than 30 nt long (79); however, an RNA duplex of that length may have a melting temperature higher than 100°C (80), and 3′-NP-DNA is even more stable (81). Therefore, thermal denaturation alone may not suffice for the replication of functional nucleic acids. Possible solutions to this problem are the use of denaturants such as formamide to destabilize the helix, the replication of short fragments that can assemble into larger functional structures, and the use of heterogeneous backbones. The mixture of 2′–5′ and 3′–5′ linkages that results from nonenzymatic RNA polymerization is usually considered undesirable. However, our laboratory recently showed that both an aptamer and a ribozyme can maintain function with up to 25% of 2′–5′ linkages randomly interspersed in the sequence (82). Because 2′–5′ linkages significantly destabilize a duplex (25) and can be copied nonenzymatically (28), they could help replicate RNA systems by facilitating strand separation.

Given that a continuous supply of activated monomers is required for indefinite self-replication, the tendency of activated monomers to hydrolyze or cyclize is also a serious problem. Furthermore, hydrolysis of activated monomers generates nucleotides that act as competitive inhibitors of template-directed polymerization (58). A recent study by Deck et al. (58) underscored the significance of these issues. These authors efficiently copied an RNA template containing all four natural bases by immobilizing the primer–template complex on magnetic beads and repeatedly exchanging the solution of activated monomers. We recently obtained a similar improvement in a vesicle-encapsulated RNA copying system by using dialysis to refresh the external solution with fresh monomers while removing hydrolyzed monomers (83). Ultimately, a more effective way to solve this problem would be through efficient and selective chemistry for in situ reactivation of hydrolyzed monomers. Various reagents that have been used to activate nucleotides in situ include carbodiimides (13), *N*-cyanoimidazole (84), and cyanogen bromide (85); however, these reagents also alkylate nucleobases and other nucleophiles (86, 87). A general, mild, and efficient way to reactivate monomers in situ would no doubt lead to significant improvements in the efficiency of nonenzymatic RNA replication.

In summary, many of the seemingly insurmountable problems with nonenzymatic nucleic acid replication have recently been overcome. Although several difficult problems remain, potential solutions to all of them have been proposed, so there is now a sense of optimism that nonenzymatic template-directed replication may be demonstrated in the not-too-distant future.

Ribozyme-Catalyzed Replication

In the 1960s, Woese (5), Orgel (6), and Crick (7) first predicted the existence of RNA catalysts, or ribozymes, as a solution to the chicken-and-egg problem of the origin of DNA and protein-based life. The discovery, some 15 years later, of ribozymes in biology (88, 89) immediately generated great interest in the possibility of ribozyme-catalyzed RNA replication (**Figure 2*b***) (90). Given substrates with a 5′-terminal guanosine residue, in order to mimic a splicing intermediate, self-splicing introns add nucleotides to a single-stranded oligonucleotide (91), ligate multiple oligonucleotides together on a template (92), and extend a primer in a template-dependent manner (93). Indeed, the *sunY* ribozyme can be broken into a multisubunit complex that can catalyze the template-directed assembly of one of its subunits from a

set of oligonucleotides (94). Despite the development of a shorter and more active version of the *sunY* ribozyme (95), it was never able to achieve self-replication for at least two reasons: The complementary oligonucleotide substrates bound to and inhibited the ribozyme, and the isoenergetic transesterification reaction used by the ribozyme provided insufficient driving force for the formation of long products.

Due to the inherent limitations of systems based on self-splicing introns, subsequent efforts shifted to the in vitro selection of novel ribozymes that could use activated nucleotides to copy templates (96). One of the first ribozymes isolated by in vitro selection catalyzed the attack of the 3′-hydroxyl of one oligonucleotide on the 5′-triphosphate of a second oligonucleotide, when both were aligned on a template (96). This ribozyme, known as the class I ligase, can catalyze the addition of a single nucleotide onto a primer by using an NTP substrate (97). McGinness et al. (98) used continuous in vitro evolution to select for a variant of this ribozyme that could catalyze three successive nucleotidyl transfer reactions in both the 5′→3′ and 3′→5′ directions. However, the polymerase activity conferred by the evolved point substitutions was modest, and it was the evolution of a novel accessory domain that first transformed the class I ligase into a true polymerase. The resulting R18 ribozyme can catalyze the template-directed elongation of a primer by up to 14 nt, again by using NTPs as substrates (99). The R18 ribozyme is a proof of principle that ribozymes can catalyze RNA polymerization, an assumption at the heart of the RNA world hypothesis. It has also been a useful starting point for further in vitro selection experiments with the ultimate goal of achieving RNA self-replication (100).

To date, the best RNA polymerase ribozyme is the tC19Z variant of the R18 ribozyme (101), which was developed through a combination of in vitro selection and engineering. It contains several base substitutions that improve activity and, critically, was engineered to bind the 5′ end of its substrate through sequence complementarity to a region of the ribozyme. The R18 ribozyme has very low affinity for its primer–template substrate but catalyzes the reaction rapidly and with some processivity once bound (102). By building in a sequence-specific interaction with the template, Wochner et al. (101) observed a dramatic increase in ribozyme activity. On a template composed of repeats of an optimized 11-nt sequence, the tC19Z ribozyme can extend a primer by up to 91 nt (albeit with a yield of only 0.035%). The mutation rate, as assessed by sequencing of full-length products, was 8.8×10^{-3}, although this value is probably an underestimate because it does not include products that may have stalled after a misincorporation event (19). This ribozyme polymerase was used to synthesize a catalytically active 24-nt portion of the mini-Hammerhead ribozyme, the first example of a ribozyme being synthesized by another ribozyme. Despite this major advance, R18 is close to 200 nt in length, so a significant increase in activity, processivity, and sequence generality would be required to approach the copying of its own sequence. Furthermore, R18 is highly structured and has many self-complementary regions that are strong blocks to continued template copying. Finally, R18 is active only in the presence of high concentrations of Mg^{2+}, which catalyzes RNA hydrolysis, including degradation of the ribozyme itself.

Further efforts to optimize R18-derived ribozymes may be facilitated by the recently determined structure of the class I ligase core (103, 104). Globally, the class I ligase resembles a tripod with three helical legs converging on the active site (103). Similarly to most proteinaceous polymerases, the class I ligase uses an Mg^{2+} ion to coordinate the 3′-hydroxyl group of the primer and the α-phosphate; however, it does not use the second ion of the canonical two–metal ion mechanism. Instead, the pyrophosphate leaving group is oriented back into the major groove, and nucleobase and hydroxyl groups are thought to stabilize the transition state electrostatically through hydrogen bonding. The accessory domain is less well characterized, but Wang et al. (105) used mutational analysis to identify critical secondary structures

and tertiary interactions. They suggested that the domain is probably draped over the vertex of the class I ligase tripod structure.

Because the point mutations identified by selection experiments on the R18 ribozyme have resulted in only modest improvements, the ribozyme may represent a local maximum of activity in sequence space. If so, the significant advance required to achieve self-replication may come either from the evolution of one or more additional accessory domains or from an entirely different polymerase core. After developing the R18 polymerase, Lawrence & Bartel (106) selected for eight other accessory domains for the class I ligase. Although none of the new ribozymes are better than the R18 polymerase, they offer different starting points for further selection. Accessory domains conferring enhanced processivity seem to offer the greatest potential for improvement. Similarly, several independent core catalytic domains are now available, and some have been transformed into nascent polymerases. McGinness & Joyce (107) evolved the hc ligase into the 18-2 ribozyme, which can catalyze the addition of single nucleotides to a primer by using NTPs as substrates. Other natural (108) and selected (109) ligase ribozymes could be evolved into polymerases; however, the direct selection for polymerase ribozymes with high affinity for the primer–template substrate would avoid the difficulty of evolving an RNA that was initially selected to perform one function into a variant with a different function.

Polymerase activity has the advantage of allowing for an open-ended exploration of sequence space; however, ribozyme-specific systems with more limited evolutionary potential have been used to explore aspects of RNA self-replication. Joyce and colleagues (110–112) developed a system based on the R3C ligase ribozyme that can ligate together two halves of another R3C ligase (109), which can then ligate together two halves of the original ligase ribozyme. This system shows the exponential growth that is characteristic of self-replication (113). These authors incorporated limited evolutionary dynamics into this system by coupling the particular sequence of a variable 4-nt "genotype" in the substrate recognition arm to a particular sequence of a variable region in the ligase active site (114). After continuous in vitro evolution, the population was dominated by sequences with a more active ligase "phenotype" that preferentially replicated ligases with the same genotype, as well as inactive parasitic sequences that were recognized by the active ligases through stable mispairing at the genotype site. However, the diversity of this system is inherently limited by the length of the genotype and phenotype sequences, which must be provided externally.

Lehman and colleagues (115–118) developed another approach to ribozyme self-replication that is based on fragment ligation but in this case uses splicing chemistry. This three-ribozyme system amplified itself through a cooperative network of cross-catalysis. Although the compartmentalization of such self-replicating systems could in principle be coupled to cell-level evolutionary selection (119), achieving continued replication inside protocells would be difficult because of the need to feed the system with large oligonucleotide substrates while retaining ligated products in the cell. Furthermore, the evolutionary capabilities of such systems are inherently limited by the large, defined RNA substrates that must be provided, whereas polymerase-based replication systems can more readily evolve novel functions through open-ended exploration of sequence space.

Replicase: a ribozyme or protein enzyme that can accelerate RNA replication, either directly or indirectly

Protein-Catalyzed Replication

When proteins are introduced into the system, the issue of nucleic acid genome replication becomes relatively simple. Many enzyme-based methods, including the polymerase chain reaction (PCR) (120), the Qβ replicase (121), and reverse transcriptase/RNA polymerase systems (122, 123), have been used to replicate nucleic acid sequences.

PCR has been used to model genome replication in several artificial cell systems. Although most polymerases are strongly inhibited by fatty acids, PCR proceeds well in phospholipid

vesicles (120, 124). Deamer and colleagues (125) found that although phospholipid membranes are normally impermeable to NTP substrates, temperature cycling across the lipid melting transition creates defects that allow NTPs to enter vesicles, thus allowing for continued RNA replication. Phospholipid vesicles are also quite robust to the temperature fluctuations required for PCR. Sugawara and colleagues (124) have used PCR to replicate DNA within more complex replicating vesicles composed of phospholipids and cationic lipids.

The Qβ replicase system is particularly interesting because it uses the RNA-dependent RNA polymerase from the Qβ bacteriophage, which can replicate its genomic RNA isothermally without the use of primers (126). This system was the first one used to demonstrate the evolution of a nucleic acid sequence in vitro (127), and it has been successfully incorporated into vesicles (121). Yomo and colleagues (128–131) have extensively studied the activity of the Qβ replicase in model protocells.

With protein-based replication, the problem of genome replication is deferred to the problem of replicating the protein polymerase, which necessitates incorporation of complex translational machinery. In vitro translation systems have successfully been incorporated into vesicles, resulting in the efficient synthesis of proteins (132, 133), including the Qβ replicase itself (128). However, replicating the entire system, including DNA and RNA polymerases, transcription and translation factors, transfer RNAs, the ribosome, and other components, is still a distant goal, and proposals to do so involve more than 100 genes (134, 135). Properly coordinating the replication of such a large synthetic system may require additional regulatory components.

COMPARTMENTS AND COMPARTMENT REPLICATION

Research on developing self-replicating compartments for model protocells is relatively advanced, in that complete replication cycles of growth and division have been demonstrated. Fatty acid–based vesicles are prebiotically plausible and have many physical properties that are appropriate for a protocell with minimal (or no) evolved biochemical machinery. For more advanced artificial cell models, more complex membranes are required. Phospholipid-based membranes seem, so far at least, to be essential for avoiding polymerase inhibition by fatty acids; however, the use of phospholipid membranes raises a host of problems, ranging from the need to import NTPs and other small-molecule substrates through these much less permeable membranes to the need for mechanisms to drive membrane growth and division. Generally, these problems are solved through the use of additional proteins, further increasing the complexity of the system, although in some cases novel physical processes can provide alternative solutions. Alternative forms of encapsulation or segregation, such as emulsion (**Figure 3*a***), coacervate, and aqueous two-phase systems (ATPS), are also being explored.

Coacervate droplets formed by separation of organic molecules from bulk water due to electrostatic and/or hydrophobic effects

Aqueous two-phase system(s) (ATPS): droplets formed by the spontaneous separation of two water-soluble polymers in water, for example, polyethylene glycol and dextran

Fatty Acid–Based Vesicles

The protocell membrane is a compartment boundary that serves two critical functions. First, it provides a selective barrier between the cell and its environment, allowing the influx of nutrients and the efflux of waste while stably encapsulating the macromolecular contents of the protocell. Second, it provides the physical link between functional gene products and the genome from which they were expressed; this physical link is necessary for Darwinian evolutionary selection. Compartmentalization also allows for the selection of cooperative gene networks. Without compartmentalization, self-replicating systems are prone to parasitism from species that are replicated efficiently (136). For a protocell membrane to be viable as part of a model for the origin of cellular life, it should be able to assemble, grow, and divide spontaneously without the assistance of any gene products.

Fatty acids form bilayer membranes when the solution pH is approximately the pK_a of the carboxylate headgroups in the membrane,

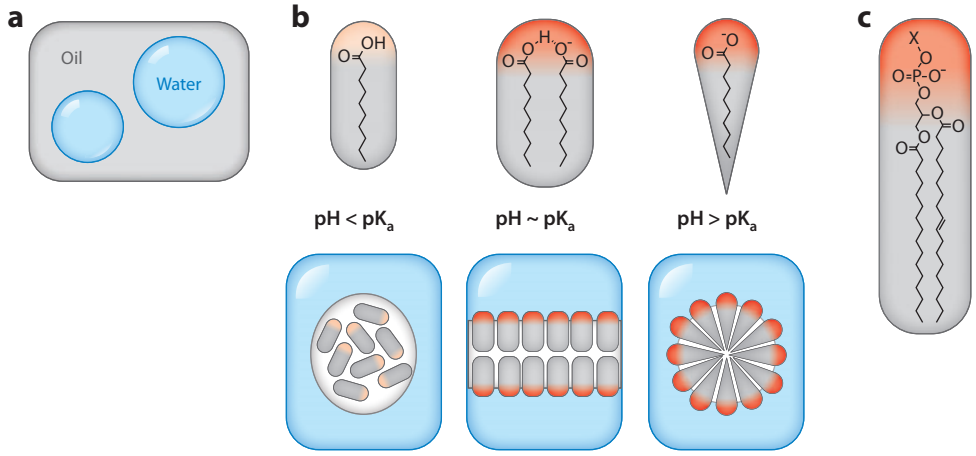

Figure 3

Common components of protocell compartments. (*a*) Nucleic acids, proteins, and other hydrophilic species can be compartmentalized in water-in-oil emulsions through the stabilization of the water droplets with surfactants. Droplets can fuse and divide by vigorous agitation. (*b*) Fatty acids spontaneously aggregate above a certain concentration. At a pH below their acid dissociation constant (pK_a) of ~7 to 9, these aggregates are amorphous precipitates or oils, and above their pK_a, repulsion of the charged carboxylate groups leads to the formation of small micelles. However, near their pK_a, the fatty acids are partially ionized and form hydrogen-bonded dimers that stabilizes the formation of extended bilayers. These bilayers can curve onto themselves to form enclosed vesicles. (*c*) Phospholipids have two fatty acid chains per charged headgroup and therefore form stable bilayers under a wider range of conditions than do fatty acids.

allowing them to form transient hydrogen-bonded dimers (**Figure 3*b***) (137, 138). The formation of vesicles at this pH is highly cooperative and is characterized by a critical aggregate concentration of fatty acid monomers above which vesicles spontaneously self-assemble (139). Unlike phospholipids, fatty acids rapidly exchange between vesicle membranes, micelles, and dissolved monomers. This dynamic exchange makes fatty acids particularly well suited as components of protocell membranes because such vesicles can grow (*a*) through the external addition of fatty acids (140), (*b*) by absorbing material from other vesicles (119, 141), or (*c*) through the generation of fatty acids in situ from precursors (142). Furthermore, as multilamellar vesicles (i.e., vesicles with multiple membranes) grow, they form threadlike shapes that readily divide into spherical daughter vesicles with the application of shear forces (143) or following photochemically induced pearling (144). These cycles of growth and division can be repeated without the loss of vesicle contents (143). We recently showed that this process can be driven simply by the concentration of a lipid solution by solvent evaporation (**Figure 4**) (145). Although the division of spherical vesicles can be induced by extrusion through narrow pores, it causes partial loss of contents and is therefore less desirable than the loss-free division of filamentous vesicles (146).

Fatty acid membranes are much more permeable to small charged molecules than are phospholipid membranes. In particular, activated nucleotides can pass through the membrane and take part in template-directed polymerization inside the protocell to produce oligonucleotide products that remain trapped inside (147). To a certain extent, one can tune the permeability (147, 148), stability (149, 150), and dynamics (141, 151, 152) of fatty acid vesicles by changing the acyl-chain length and saturation state of the fatty acid components or by adding, for example, fatty alcohols, fatty esters, and polycyclic aromatic hydrocarbons.

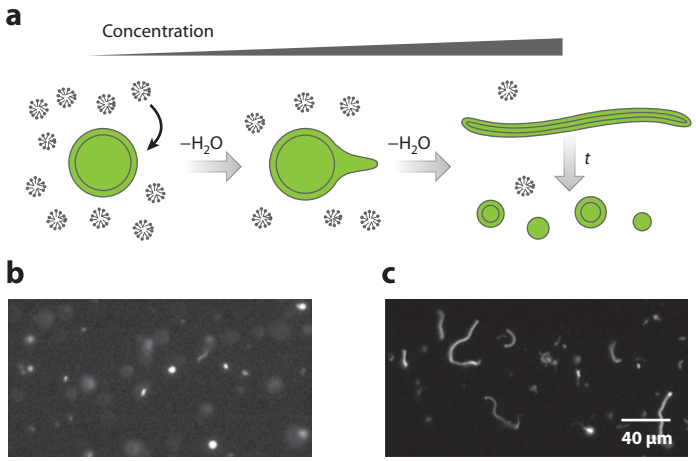

Figure 4

(*a*) Demonstration of fatty acid vesicle growth and division. Concentration of spherical multilamellar fatty acid vesicles (*b*) by evaporating the water solvent causes the membrane to grow into long tubules (*c*), which can divide into small vesicles through simple agitation. The contents of the vesicle are maintained during this growth and division process. Similar processes may have been an important mechanism of replication for protocells on the early Earth. Modified with permission from Reference 145.

Fatty acids are particularly interesting due to their probable involvement in the origin of life. They were almost certainly available on the early Earth, given that they have been found in meteorites (albeit in low levels) (153, 154); they can be synthesized abiotically (155); and they readily self-assemble into membranes under potentially prebiotic conditions (146, 156). Vesicle dynamics may have provided an early opportunity for selection on the basis of competition for fatty acids (119, 141). However, vesicles composed of fatty acids are not as robust as those made of phospholipids. In particular, they precipitate in the presence of millimolar concentrations of divalent cations, whereas higher concentrations are required for nonenzymatic RNA polymerization (150) and for many ribozymes.

Phospholipid and Synthetic Lipid Vesicles

Phospholipid vesicles are the compartments of choice for more complex artificial cells; their biophysical properties have been studied extensively (**Figure 3*c***) (139). They are stable under a wide range of conditions and can be formed by various techniques (157). However, because phospholipid molecules are essentially permanently anchored within membranes, and therefore do not exchange between membranes, the growth of phospholipid vesicles must occur through processes that are different from those that lead to the growth of fatty acid vesicles. Under certain conditions, phospholipid vesicles can grow through fusion with other vesicles (158–161), but vesicle–vesicle fusion mediated by fusogenic lipids or peptides is often associated with significant contents leakage. In principle, in situ enzymatic lipid synthesis (162, 163) should allow unlimited growth as long as appropriate substrates are supplied, and the necessary enzymes are generated internally by translation. Division by budding can be induced by the depletion volume effect (159) or phase separation (164) of encapsulated polymers, or through enzymatic activity (165, 166).

Sugawara and colleagues (124, 167, 168) have explored membrane vesicle growth and division by using nonbiological lipids and small-molecule catalysts. In one system, an amphiphilic aldehyde is generated in situ through the action of a membrane-localized catalyst on a protected precursor; the aldehyde then condenses with an amphiphilic amine to form a cationic bola-amphiphile (124). In related experiments, the in situ synthesis of membrane-forming amphiphiles perturbed the structure of multilamellar vesicles to cause either vesicle division or the assembly and release of new vesicles from the interior of a preformed vesicle. The combination of catalyzed lipid synthesis and subsequent vesicle division is an exciting advance toward the self-replication of systems of compartments that are chemically distinct from anything observed in biology.

Alternative Approaches to Compartmentalization

Although phospholipid and fatty acid vesicles are the most-studied models for the membranes of artificial cells, other systems have been

developed. As reviewed by Kamat et al. (169), various polymers have been used to create vesicles known as polymersomes, including polyethylene derivatives (170), polypeptides (171), and dendrimers (172). Polymersomes can be readily engineered to have particular desirable properties and functionality. Because their stability makes growth and division difficult (173–175), artificial cells based on polymersomes are typically used to model aspects of nongrowing cells. The design of subunits that would generate polymersomes with the dynamic properties required for growth (either continuous or via fusion events) and division is an interesting challenge for future research. Moving even further away from standard biology, artificial cells based on nonmembrane types of compartmentalization have recently been studied. Water-in-oil emulsions have been used to encapsulate the Qβ replicase system (130) and complete transcription–translation systems (176, 177). Remarkably, the aqueous droplets of emulsions can be "fed" by fusion with other droplets and can be forced to divide by shear forces or extrusion, allowing for a cycle of growth and division (178). In a completely different nonbiological system, eutectic phase channels formed in frozen water have been used to spatially partition a ribozyme polymerase (179) and nonenzymatic RNA polymerization reactions (180), although such systems are not amenable to cycles of replication.

Interestingly, the components of a protocell can also be localized through selective partitioning in multiphase systems. Functional ribozymes can be enriched in the dextran-rich droplets that form spontaneously in aqueous mixtures of polyethylene glycol and dextran (181). Nucleotides and cationic peptides can assemble into coacervate microdroplets that can sequester enzymes and other components (182). An important question is whether the lack of a low-permeability membrane barrier in these systems causes rapid exchange of RNAs (or other genetic materials) between droplets, which could diminish or abolish the spatial partitioning of genetic materials that is required for Darwinian evolution. The combination of coacervate systems or ATPS with lipid membranes provides an opportunity for compartmentalization within a protocell, analogous to that obtained in biological cells by organelles; furthermore, such combined systems provide a novel pathway for protocell division in response to osmotic and surface tension effects (164).

INTEGRATED ARTIFICIAL CELLULAR SYSTEMS

Interesting new challenges and opportunities arise when self-replicating genetic systems and compartments are brought together. In this section, we discuss the distinct issues that arise in the design of integrated protocell models, versus more complex protein-based artificial cells.

Simple Protocells

In an attempt to resolve the genetics-first versus compartments-first controversy in the origin-of-life field, one of us (J.W.S.), together with Bartel and Luisi (183), proposed an integrated protocell model based on a self-replicating nucleic acid genome within a self-replicating membrane boundary. That model was based on ribozyme-catalyzed RNA replication because, at the time, there was no plausible path toward nonenzymatic RNA replication. The model also proposed that the emergence of a second ribozyme that carried out some cellular function, such as synthesis of a lipid component of a vesicle compartment, would result in an integrated cell capable of Darwinian evolution. Since we proposed that early model, advances in both nucleic acid and compartment replication have brought the field to the point at which various protocell models are being constructed and evaluated.

A fundamental aspect of any viable protocell model is that the physicochemical conditions required for stability and replication of the nucleic acid and membrane components must be compatible. Fatty acid vesicles are stable at a pH ranging from ~7 to 9, depending on the particular fatty acid used (137, 138, 184), and this range can be extended by the incorporation

of fatty alcohols and glycerol monoesters (185). Fortunately, this pH range is favorable for both nonenzymatic RNA polymerization (58, 186) and RNA polymerase ribozyme activity (99, 101). Fatty acid–based vesicles can also be stable up to 100°C, allowing for thermal denaturation of encapsulated nucleic acids and enhanced permeation of nucleotide substrates (149).

Unfortunately, fatty acid membranes are unstable above low millimolar concentrations of divalent cations (150), whereas nonenzymatic RNA polymerization (186) and ribozyme polymerase activity (99) typically require at least 50 mM Mg^{2+}. There are several potential solutions to this apparent incompatibility. Nonenzymatic template-directed copying with aminosugar nucleotides does not require Mg^{2+} and has successfully been performed in fatty acid vesicles (147), suggesting that the creation of artificial cells based on phosphoramidate nucleic acids may be possible. However, a protocell with an RNA genome and ribozyme catalysts would require alternative solutions: either a replacement for divalent cations that does not destroy fatty acid membranes or a means of complexing the essential Mg^{2+} ion so as to protect the membrane while allowing RNA replication to proceed. We recently found that the tricarboxylic acid citrate chelates Mg^{2+} such that fatty acid membranes are preserved while nonenzymatic template copying is only minimally affected (83). This observation has allowed us to encapsulate an RNA primer–template complex inside fatty acid vesicles, add activated ribonucleotides to the vesicles, and observe primer extension as the nucleotides spontaneously diffuse across the membrane and then copy the template strands inside the vesicle. This advance is sufficient to allow the further development of model protocells based on RNA, but it also raises the exciting prospect that chelators with catalytic activity could facilitate the copying of long mixed-sequence templates. An alternative approach to the compatibility problem would be the isolation of a ribozyme replicase that functions at very low levels of Mg^{2+} ions, which may be feasible because biological ribozymes function at low intracellular Mg^{2+} levels. Finally, it may be possible to replace the fatty acids of the protocell membranes with alternative lipids, such as nonionic amphiphiles, that do not interact strongly with Mg^{2+} ions.

Other requirements for a fully integrated protocell are that nutrients, especially activated nucleotides, must be able to enter the cell and waste must be able to leave, but the self-replicating genetic system must be permanently trapped inside. Fatty acid vesicles have an advantage in this respect because they are permeable to small charged molecules, including nucleotides, but not oligonucleotides (149). Raising the temperature or adding low millimolar concentrations of Mg^{2+} further increases their permeability (149). Phospholipid vesicles are impermeable to even monovalent cations; however, small charged molecules can travel through transient defects in the bilayer membrane (187) that form at the gel phase–to–liquid phase transition temperature (188). This property has been exploited to enable feeding of external NTPs to encapsulated enzymatic reactions (125), although this method may require thermocycling between the membrane phase transition temperature and the temperature optimum of the enzyme (189). Alternatively, protein pores (133, 190), shorter-chain phospholipids (191), or detergents (192) can be used to permeabilize the membrane. Another attractive approach is to use vesicle fusion to deliver nutrients encapsulated in feeder vesicles (161); unfortunately, most simple vesicle fusion methods are inefficient and induce considerable contents leakage (158–160, 193).

By directly coupling the protocell genome to cell-level phenotypes, compartmentalization may drive the evolution of more complex functions in a way that could emulate the early evolution of life. Perhaps the simplest such coupling results from the osmotic pressure due to encapsulated RNA (119). The membranes of osmotically swollen vesicles are under tension, but this high-energy state can relax through the absorption of fatty acids from surrounding vesicles that are less swollen because they contain less RNA. Thus, any mutations that enhance RNA replication, leading to

the faster accumulation of internal RNA, will cause faster membrane growth. Although osmotically driven competitive growth is attractive because of the simplicity of the physically mediated linkage between genome replication and cell growth, osmotically swollen vesicles are very difficult to divide.

An alternative but still simple way to link genomic function to cell growth would involve a genomically encoded ribozyme with phospholipid synthase activity. We recently showed that fatty acid vesicles that contain even small amounts of phospholipids (141) grow at the expense of surrounding vesicles with less phospholipid content. This effect is independent of osmotic pressure; instead, it results from the ordering effect of phospholipids on the membrane, which causes a decreased rate of dissociation of fatty acid molecules from the membrane. The net effect is that genomic sequences leading to a low level of phospholipid synthesis (e.g., by condensation of a lysophospholipid with a fatty acyl thioester) would confer a strong growth advantage on the protocell; moreover, the resulting growth would follow the filamentous growth pathway, so subsequent division is simple. A very similar growth advantage results from the internal synthesis of hydrophobic peptides (194), which could also be catalyzed either directly by a ribozyme or indirectly by a catalytic peptide that, in turn, is synthesized by a ribozyme. These two processes would lead to a strong selective advantage for genomically encoded and thus heritable phospholipid or peptide synthesis activity, given the availability of appropriate substrates.

For any artificial cellular system to be capable of indefinite propagation, the replication of the genome, biochemical contents, and compartment must be coordinated. Interestingly, this task may pose less of a problem in very simple protocells than in more complex artificial cells that contain many components. In a simple protocell that uses either nonenzymatic or ribozyme-catalyzed genome replication, with postsynthesis strand separation driven by high temperature excursions, genome replication would be self-limiting because strand reannealing is a second-order process. Thus, increasing concentrations of genomic RNA would lead to faster reannealing until the reannealing occurs faster than strand copying; as a result, a steady-state level of genomic RNA would be attained. However, if compartment replication were much faster than genome replication, the population could be overrun by empty protocells. Although this outcome could be avoided by the controlled addition of limited quantities of micelles to fatty acid vesicles (143), if growth is driven by an internal process such as ribozyme-promoted phospholipid synthesis (141), then excess cell membrane growth would cause ribozyme dilution, which would decrease the rate of ribozyme-promoted growth. Thus, the combination of ribozyme-promoted growth and concentration-dependent strand reannealing could, in principle, lead to a steady-state level of genomic RNA in a population of replicating protocells, without additional regulatory signals.

Sugawara and colleagues (124) recently demonstrated an interesting version of a similar process at the DNA level. These authors showed that PCR-mediated DNA amplification can drive the growth and division of cationic vesicles. In this system, the vesicles replicate through the adsorption of a cationic lipid precursor, which is hydrolyzed by a catalyst to generate the final cationic membrane component. Although the mechanism of the DNA-enhanced growth and division is not entirely clear, it is probably a consequence of (*a*) the interaction between the anionic DNA and the cationic membrane, which enhanced adsorption of the cationic precursor, and possibly (*b*) DNA-induced curvature changes that lead to enhanced vesicle division.

Artificial Cells Based on Proteins

Because many of the applications proposed for artificial cells (134, 195) require proteins, many researchers have attempted to develop systems for efficient compartmentalized transcription and translation. Early experiments by Oberholzer et al. (132) showed that polyuridylic

acid can be translated into polyphenylalanine in phospholipid vesicles. Yu et al. (196) and Nomura et al. (197) synthesized green fluorescent protein in greater yield by using a cell-free *Escherichia coli* extract encapsulated in giant phospholipid vesicles (>1 μm in diameter), but yields were still limited by the impermeability of the phospholipid membrane, which made it impossible to replenish nucleotides and amino acids. Noireaux & Libchaber (133) solved the problem of vesicle impermeability by expressing the α-hemolysin pore protein in their vesicles, which allowed them to maintain translation for up to 4 days by adding nutrients externally. In a further step toward a totally synthetic cell, Murtas et al. (198) incorporated the PURE (protein synthesis using recombinant elements) translation system (199), which is reconstituted from purified components, into phospholipid vesicles. Nevertheless, even this system has a limited lifetime because critical components degrade or become inactivated.

One of the ultimate goals of this line of research is for the encapsulated transcription–translation machinery to regenerate all of the cell components, which would prolong synthetic activity and ultimately lead to a complete self-replicating system—in effect, the reconstitution or semisynthesis of a simplified bacterial cell. However, many challenges must be addressed before this goal can be achieved.

For complex artificial cells containing elaborate biochemical networks with larger numbers of components, the problem of maintaining the proper balance between replication of the genome, the biochemical machinery, and the compartment itself becomes more complex. Even in relatively simple systems, regulation can quickly become challenging. For example, in a recent study the Qβ replicase was translated from Qβ genomic RNA, which was concurrently replicated by the Qβ replicase (128). In this system, the ribosome and the Qβ replicase inhibited one another's activity by competing for binding to the Qβ RNA. As a result, efficient genome amplification was obtained only at an optimized ratio of ribosomes to replicase (129).

When an artificial cell divides, its contents randomly partition into the two daughter cells; the more components the system has, the more copies are required to ensure that both daughter cells acquire at least one copy of each essential component. Thus, it may become important to incorporate simple regulatory feedback mechanisms to keep multicomponent systems, such as the translation apparatus, internally balanced during growth. Reliance on the statistical segregation of many components may impose a minimum cell size, such that all components can be synthesized at a high copy number prior to cell division. Alternatively, incorporating mechanisms for the nonrandom segregation of key components may increase the efficiency of cell division.

Yomo and colleagues (178) demonstrated an interesting way to avoid this problem by developing an emulsion droplet model of a replicating artificial cell. In this system, the droplets contain a transcription–translation system. Droplet growth is controlled by fusion with smaller droplets containing fresh transcription–translation mixture, and division is controlled by periodic passage of the emulsion through a filter with pores of a defined size. The cycle of growth and division is thus manually controlled but can be continued indefinitely, and indeed the system was taken through almost 200 generations of growth and division. A genomic RNA coding for the Qβ replicase was encapsulated; the RNA was replicated by its encoded replicase. In a remarkable example of the spontaneous emergence of Darwinian behavior, the RNA and the protein replicase coevolved enhanced binding affinity and specificity. This process led, over many generations, to more efficient synthesis of the replicase and thus more efficient replication of the RNA. This system is, in principle, open ended and could be used to evolve additional RNA-encoded proteins, as long as a means of imposing an appropriate selective pressure can be found. The demonstration of Darwinian evolution in this droplet replication system is impressive and makes this artificial cell model clearly the most advanced so far.

Although gene expression in phospholipid vesicles is now routine, efforts to translate proteins that function in the growth or division of the artificial cell are still in the early stages. A step in this direction is the expression in vesicles of two enzymes from a pathway that produces phospholipids from glycerol phosphate and fatty acyl-CoA (162). Although the expected lipid products were detectable, the efficiency of the synthesis was far too low to result in measurable membrane growth. In a step toward internal energy metabolism, a single subunit of the F_1F_0-ATPase was expressed in vesicles containing the remaining subunits to complete the formation of a functional ATPase complex (200). Although very crude cell division may not require any cytoskeletal machinery, as evidenced by the growth of L-form mutant bacteria (201, 202), more controlled cell division requires the actin and tubulin homologs mreB and ftsZ. The expression of the actin homolog mreB in liposomes results in the assembly of fibril bundles on the membrane surface, but only if the membrane protein mreC is coexpressed. Interestingly, the fibril bundles can perturb the shape of the liposome, suggesting that effects on division may not be too difficult to reconstitute. Additional pathways that must be fully reconstituted to enable an artificial cell to replicate include the replication of a moderately large (~100-kB) DNA genome and the assembly of functional ribosomes.

FUTURE OF THE FIELD

For nucleic acid–based protocells, the development of a self-replicating genome remains the most significant challenge. Enhancing nonenzymatic template copying through chemical modifications (8–10) and/or the use of a mixture of activated mononucleotides and short oligonucleotides (46) and developing novel RNA-dependent RNA polymerase ribozymes are promising avenues of research. Once such a replication system is developed, it should become possible to construct integrated replicating protocells and then to follow the emergence of Darwinian evolution in a protocell population seeded with random-sequence templates. Both the selection of metabolic ribozymes with useful activities, such as phospholipid synthesis (141), and the evolution of structural RNAs that modulate cell division are possible outcomes, but the most interesting possibility may be the spontaneous emergence of a completely unexpected function. Protocells should be especially good at selecting for complexes of oligonucleotides that assemble and function together as ribozymes, given that their relationship will be maintained as they replicate; such experiments are essentially impossible in the absence of compartmentalization. Once any RNA that confers an advantage on its host cell emerges, there should be strong selection for the emergence of ribozymes that enhance replication efficiency and accuracy. Such experiments may eventually shed light on the emergence of metabolism and translation and may address the question of the cellular complexity attainable in an RNA-based cell. Protocell research has already contributed to the development of plausible pathways and geochemical scenarios for the origin of life (78, 203), and an improved understanding of the requirements for protocell reproduction should further constrain such models. Nucleic acid–based protocells are generally considered to have fewer practical applications than those capable of expressing proteins. However, given the wide variety of aptamers and ribozymes that have been isolated (204) and the rapid developments in the fields of biosensing (205), DNA nanotechnology (206), and computation (207), the functional potential of nucleic acids should not be underestimated.

For artificial cellular systems based on internal protein expression, an entirely different set of challenges lies ahead. In essence, all pathways required for the growth and division of a minimal bacterial cell must be fully reconstituted to function with an efficiency approaching that observed in extant biological cells. Thus, DNA replication, transcription, and translation; ribosome assembly; and membrane synthesis must all be reconstituted and built into the design of the artificial cell. In the process of building and combining these substrates, many

interesting questions can be addressed. For example, can all of metabolism be avoided simply by incorporating membrane channels and supplying nutrients externally? With respect to self-assembly, are there components other than the genome and the membrane that can be made only in an autocatalytic or self-templated manner? Candidates for such materials include the peptidoglycan cell wall, which grows by addition of new material to old, and membrane protein chaperones that are required for their own folding (208). Only the complete reconstitution of a living cell from purified components will prove that all such self-templating materials and processes have been identified. The synthesis of a living artificial cell from components will open the door to many more adventurous lines of research, such as the synthesis of so-called enantio-life, in which all components have reversed chirality (209), and the synthesis of cells with further-reduced complexity, which would approximate the nature of intermediates in the evolutionary path from protocells to modern life. It may even become possible to assemble cells in which the standard biopolymers are replaced with altered versions, so as to explore the range of molecular diversity that is compatible with cellular life. Finally, there is no shortage of proposed applications for artificial cells (134, 195, 210); they range from the biosynthesis of pharmaceuticals (211) or biofuels (212) to advanced drug-delivery technology, in which artificial cells would be engineered to control the timing and site of drug release. Artificial cells could be designed to continuously synthesize and deliver a drug within a human host (210); signal to and modify the behavior of natural cells; or carry out functions usually performed by healthy human cells, such as the transport of oxygen through the blood (213). The exciting challenge of the synthesis of life from nonliving components, combined with the potential for transformative applications, will undoubtedly drive progress in this field for many years to come.

FUTURE ISSUES

1. Investigators should overcome the barriers to nonenzymatic nucleic acid replication, including the slow rate of A:U copying, G:U wobble pairing, strand separation after copying, and in situ nucleotide activation.
2. There is a need to develop a smaller polymerase ribozyme with a lower requirement for Mg^{2+}, which would increase its stability to hydrolysis.
3. Simpler and more robust methods of phospholipid vesicle growth and division are required.
4. Experiments should aim to improve compatibility between nucleic acid replication and vesicle replication, especially the instability of fatty acid vesicles in the presence of divalent cations, which are required for RNA replication.
5. Investigators should achieve replication of the translational machinery for increased autonomy of protein-based protocells.
6. Nucleic acid replication must be templated by preexisting nucleic acids. Are there other cellular structures that require a preformed template for their replication, for example, components of the cell wall?

DISCLOSURE STATEMENT

J.W.S. is a Co-Chair of the Steering Committee of the Simons Collaboration on the Origin of Life and a Co-Chair of the Steering Committee of the Harvard Origins of Life Initiative.

LITERATURE CITED

1. Hutchison C, Peterson S, Gill S, Cline R, White O, et al. 1999. Global transposon mutagenesis and a minimal mycoplasma genome. *Science* 286:2165–69
2. Gibson DG, Benders GA, Andrews-Pfannkoch C, Denisova EA, Baden-Tillson H, et al. 2008. Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science* 319:1215–20
3. Isaacs FJ, Carr PA, Wang HH, Lajoie MJ, Sterling B, et al. 2011. Precise manipulation of chromosomes in vivo enables genome-wide codon replacement. *Science* 333:348–53
4. Rich A. 1962. On the problems of evolution and biochemical information transfer. In *Horizons in Biochemistry: Albert Szent-Györgyi Dedicatory Volume*, ed. M Kasha, B Pullman, pp. 103–26. San Diego: Academic
5. Woese CR. 1967. *The Genetic Code: The Molecular Basis for Genetic Expression*. New York: Harper & Row
6. Orgel LE. 1968. Evolution of the genetic apparatus. *J. Mol. Biol.* 38:381–93
7. Crick FHC. 1968. The origin of the genetic code. *J. Mol. Biol.* 38:367–79
8. Schrum JP, Ricardo A, Krishnamurthy M, Blain JC, Szostak JW. 2009. Efficient and rapid template-directed nucleic acid copying using 2′-amino-2′,3′-dideoxyribonucleoside-5′-phosphorimidazolide monomers. *J. Am. Chem. Soc.* 131:14560–70
9. Zhang S, Zhang N, Blain JC, Szostak JW. 2013. Synthesis of N3′-P5′-linked phosphoramidate DNA by nonenzymatic template-directed primer extension. *J. Am. Chem. Soc.* 135:924–32
10. Kaiser A, Richert C. 2013. Nucleotide-based copying of nucleic acid sequences without enzymes. *J. Org. Chem.* 78:793–99
11. Schramm G, Grotsch H, Pollman DW. 1962. Non-enzymatic synthesis of polysaccharides, nucleosides and nucleic acids and the origin of self-reproducing systems. *Angew. Chem. Int. Ed.* 1:1–64
12. Naylor R, Gilham PT. 1966. Studies on some interactions and reactions of oligonucleotides in aqueous solution. *Biochemistry* 5:2722–28
13. Sulston J, Lohrmann R, Orgel LE, Miles HT. 1968. Nonenzymatic synthesis of oligoadenylates on a polyuridylic acid template. *Proc. Natl. Acad. Sci. USA* 59:726–33
14. Inoue T, Joyce GF, Grzeskowiak K, Orgel LE, Brown JM, Reese CB. 1984. Template-directed synthesis on the pentanucleotide CpCpGpCpC. *J. Mol. Biol.* 178:669–76
15. Haertle T, Orgel LE. 1986. Template-directed synthesis on the oligonucleotide d(C_7-G-C_7). *J. Mol. Biol.* 188:77–80
16. Joyce GF, Orgel LE. 1986. Non-enzymic template-directed synthesis on RNA random copolymers. Poly(C,G) templates. *J. Mol. Biol.* 188:433–41
17. Wu TF, Orgel LE. 1992. Nonenzymatic template-directed synthesis on hairpin oligonucleotides. 2. Templates containing cytidine and guanosine residues. *J. Am. Chem. Soc.* 114:5496–501
18. Wu TF, Orgel LE. 1992. Nonenzymatic template-directed synthesis on hairpin oligonucleotides. 3. Incorporation of adenosine and uridine residues. *J. Am. Chem. Soc.* 114:7963–69
19. Rajamani S, Ichida JK, Antal T, Treco DA, Leu K, et al. 2010. Effect of stalling after mismatches on the error catastrophe in nonenzymatic nucleic acid replication. *J. Am. Chem. Soc.* 132:5880–85
20. Orgel LE. 2004. Prebiotic chemistry and the origin of the RNA world. *Crit. Rev. Biochem. Mol.* 39:99–123
21. Joyce G. 2002. The antiquity of RNA-based evolution. *Nature* 418:214–21
22. Renz M, Lohrmann R, Orgel LE. 1971. Catalysts for the polymerization of adenosine cyclic 2′,3′-phosphate on a poly(U) template. *Biochim. Biophys. Acta* 240:463–71
23. Pitsch S, Krishnamurthy R, Bolli M, Wendeborn S, Holzner A, et al. 1995. Pyranosyl-RNA ("p-RNA"): base-pairing selectivity and potential to replicate. *Helv. Chim. Acta* 78:1621–35
24. Rohatgi R, Bartel DP, Szostak JW. 1996. Kinetic and mechanistic analysis of nonenzymatic, template-directed oligoribonucleotide ligation. *J. Am. Chem. Soc.* 118:3332–39
25. Wasner M, Arion D, Borkow G, Noronha A, Uddin A, et al. 1998. Physicochemical and biochemical properties of 2′,5′-linked RNA and 2′,5′-RNA:3′,5′-RNA "hybrid" duplexes. *Biochemistry* 37:7478–86
26. Inoue T, Orgel LE. 1981. Substituent control of the polyC-directed oligomerization of guanosine 5′-phosphorimidazolide. *J. Am. Chem. Soc.* 103:7666–67
27. Weimann BJ, Lohrmann R, Orgel LE, Schneider-Bernloehr H, Sulston JE. 1968. Template-directed synthesis with adenosine-5′-phosphorimidazolide. *Science* 161(3839):387

28. Prakash T, Roberts C, Switzer C. 1997. Activity of 2′,5′-linked RNA in the template-directed oligomerization of mononucleotides. *Angew. Chem. Int. Ed.* 36:1522–23
29. Kanavarioti A, Bernasconi C, Doodokyan DL, Alberas DJ. 1989. Magnesium-ion catalyzed P–N bond hydrolysis in imidazolide-activated nucleotides—relevance to template-directed synthesis of polynucleotides. *J. Am. Chem. Soc.* 111:7247–57
30. Stütz JAR, Kervio E, Deck C, Richert C. 2007. Chemical primer extension: individual steps of spontaneous replication. *Chem. Biodivers.* 4(4):784–802
31. Hagenbuch P, Kervio E, Hochgesand A, Plutowski U, Richert C. 2005. Chemical primer extension: efficiently determining single nucleotides in DNA. *Angew. Chem. Int. Ed.* 44:6588–92
32. Vogel SR, Deck C, Richert C. 2005. Accelerating chemical replication steps of RNA involving activated ribonucleotides and downstream-binding elements. *Chem. Commun.* 2005:4922–24
33. Carpino LA. 1993. 1-Hydroxy-7-azabenzotriazole—an efficient peptide coupling additive. *J. Am. Chem. Soc.* 115:4397–98
34. Röthlingshöfer M, Kervio E, Lommel T, Plutowski U, Hochgesand A, Richert C. 2008. Chemical primer extension in seconds. *Angew. Chem. Int. Ed.* 47:6065–68
35. Prabahar KJ, Ferris JP. 1997. Adenine derivatives as phosphate-activating groups for the regioselective formation of 3′,5′-linked oligoadenylates on montmorillonite: possible phosphate-activating groups for the prebiotic synthesis of RNA. *J. Am. Chem. Soc.* 119:4330–37
36. Loeb LA, Kunkel TA. 1982. Fidelity of DNA synthesis. *Annu. Rev. Biochem.* 51:429–57
37. Petruska J, Sowers LC, Goodman MF. 1986. Comparison of nucleotide interactions in water, proteins, and vacuum: model for DNA polymerase fidelity. *Proc. Natl. Acad. Sci. USA* 83:1559–62
38. Chen JL, Dishler AL, Kennedy SD, Yildirim I, Liu B, et al. 2012. Testing the nearest neighbor model for canonical RNA base pairs: revision of GU parameters. *Biochemistry* 51:3508–22
39. Leu K, Obermayer B, Rajamani S, Gerland U, Chen IA. 2011. The prebiotic evolutionary advantage of transferring genetic information from RNA to DNA. *Nucleic Acids Res.* 39:8135–47
40. Chaput J, Sinha S, Switzer C. 2002. 5-Propynyluracil-diaminopurine: an efficient base-pair for nonenzymatic transcription of DNA. *Chem. Commun.* 2002:1568–69
41. Budow S, Eickmeier H, Reuter H, Seela F. 2009. 2′-Deoxy-5-propynyluridine: a nucleoside with two conformations in the asymmetric unit. *Acta Crystallogr. C* 65:645–48
42. Katritzky AR, Waring AJ. 1962. Tautomeric azines. Part I. The tautomerism of 1-methyluracil and 5-*bromo*-1-methyluracil. *J. Chem. Soc.* 1962:1540–44
43. Testa SM, Disney MD, Turner DH, Kierzek R. 1999. Thermodynamics of RNA–RNA duplexes with 2- or 4-thiouridines: implications for antisense design and targeting a group I intron. *Biochemistry* 38:16655–62
44. Sintim H, Kool E. 2006. Enhanced base pairing and replication efficiency of thiothymidines, expanded-size variants of thymidine. *J. Am. Chem. Soc.* 128:396–97
45. Wojciechowski F, Leumann CJ. 2011. Alternative DNA base-pairs: from efforts to expand the genetic code to potential material applications. *Chem. Soc. Rev.* 40:5669–79
46. Szostak JW. 2011. An optimal degree of physical and chemical heterogeneity for the origin of life? *Philos. Trans. R. Soc. B* 366:2894–901
47. Lohrmann R, Bridson PK, Orgel LE. 1981. Condensation of activated diguanylates on a poly(C) template. *J. Mol. Evol.* 17:303–6
48. James KD, Ellington AD. 1997. Surprising fidelity of template-directed chemical ligation of oligonucleotides. *Chem. Biol.* 4:595–605
49. Rohatgi R, Bartel DP, Szostak JW. 1996. Nonenzymatic, template-directed ligation of oligoribonucleotides is highly regioselective for the formation of 3′–5′ phosphodiester bonds. *J. Am. Chem. Soc.* 118:3340–44
50. Sievers D, Kiedrowski von G. 1994. Self-replication of complementary nucleotide-based oligomers. *Nature* 369:221–24
51. Leu K, Kervio E, Obermayer B, Turk-MacLeod RM, Yuan C, et al. 2013. Cascade of reduced speed and accuracy after errors in enzyme-free copying of nucleic acid sequences. *J. Am. Chem. Soc.* 135:354–66
52. Lohrmann R, Orgel LE. 1976. Template-directed synthesis of high molecular weight polynucleotide analogues. *Nature* 261:342–44

53. Zielinski WS, Orgel LE. 1985. Oligomerization of activated derivatives of 3′-amino-3′-deoxyguanosine on poly(C) and poly(dC) templates. *Nucleic Acids Res.* 13:2469–84
54. Zielinski WS, Orgel LE. 1987. Oligoaminonucleoside phosphoramidates. Oligomerization of dimers of 3′-amino-3′-deoxy-nucleotides (GC and CG) in aqueous solution. *Nucleic Acids Res.* 15:1699–715
55. Tohidi M, Zielinski WS, Chen CH, Orgel LE. 1987. Oligomerization of 3′-amino-3′-deoxyguanosine-5′-phosphorimidazolidate on a d(CpCpCpCpC) template. *J. Mol. Evol.* 25:97–99
56. Kaiser A, Spies S, Lommel T, Richert C. 2012. Template-directed synthesis in 3′- and 5′-direction with reversible termination. *Angew. Chem. Int. Ed.* 51:8299–303
57. Kervio E, Hochgesand A, Steiner UE, Richert C. 2010. Templating efficiency of naked DNA. *Proc. Natl. Acad. Sci. USA* 107:12074–79
58. Deck C, Jauker M, Richert C. 2011. Efficient enzyme-free copying of all four nucleobases templated by immobilized RNA. *Nat. Chem.* 3:603–8
59. Zhang S, Blain JC, Zielinska D, Gryaznov S, Szostak JW. 2013. Fast and accurate non-enzymatic copying of an RNA-like synthetic genetic polymer. *Proc. Natl. Acad. Sci. USA* 110:17732–37
60. Hill D, Mio M, Prince R, Hughes T, Moore J. 2001. A field guide to foldamers. *Chem. Rev.* 101:3893–4011
61. Eschenmoser A. 1999. Chemical etiology of nucleic acid structure. *Science* 284:2118–24
62. Schöning K, Scholz P, Guntha S, Wu X, Krishnamurthy R, Eschenmoser A. 2000. Chemical etiology of nucleic acid structure: the α-threofuranosyl-(3′→2′) oligonucleotide system. *Science* 290:1347–51
63. Heuberger BD, Switzer C. 2006. Nonenzymatic oligomerization of RNA by TNA templates. *Org. Lett.* 8:5809–11
64. Yu H, Zhang S, Chaput JC. 2012. Darwinian evolution of an alternative genetic system provides support for TNA as an RNA progenitor. *Nat. Chem.* 4:183–87
65. Hendrix C, Rosemeyer H, De Bouvere B, Van Aerschot A, Seela F, Herdewijn P. 1997. 1′,5′-Anhydrohexitol oligonucleotides: hybridisation and strand displacement with oligoribonucleotides, interaction with RNase H and HIV reverse transcriptase. *Chem. Eur. J.* 3:1513–20
66. Allart B, Khan K, Rosemeyer H, Schepers G, Hendrix C, et al. 1999. D-Altritol nucleic acids (ANA): hybridisation properties, stability, and initial structural analysis. *Chem. Eur. J.* 5:2424–31
67. Kozlov I, Zielinski M, Allart B, Kerremans L, Van Aerschot A, et al. 2000. Nonenzymatic template-directed reactions on altritol oligomers, preorganized analogues of oligonucleotides. *Chem. Eur. J.* 6:151–55
68. Nielsen PE. 1999. Peptide nucleic acid. A molecule with two identities. *Acc. Chem. Res.* 32:624–30
69. Wittung P, Nielsen PE, Buchardt O, Egholm M, Nordén B. 1994. DNA-like double helix formed by peptide nucleic acid. *Nature* 368:561–63
70. Schmidt JG, Christensen L, Nielsen PE, Orgel LE. 1997. Information transfer from DNA to peptide nucleic acids by template-directed syntheses. *Nucleic Acids Res.* 25:4792–96
71. Kleiner RE, Brudno Y, Birnbaum ME, Liu DR. 2008. DNA-templated polymerization of side-chain-functionalized peptide nucleic acid aldehydes. *J. Am. Chem. Soc.* 130:4646–59
72. Schmidt JG, Nielsen PE, Orgel LE. 1997. Information transfer from peptide nucleic acids to RNA by template-directed syntheses. *Nucleic Acids Res.* 25:4797–802
73. Heemstra JM, Liu DR. 2009. Templated synthesis of peptide nucleic acids via sequence-selective base-filling reactions. *J. Am. Chem. Soc.* 131:11347–49
74. Ura Y, Beierle JM, Leman LJ, Orgel LE, Ghadiri MR. 2009. Self-assembling sequence-adaptive peptide nucleic acids. *Science* 325:73–77
75. Nielsen PE, Egholm M, Berg RH, Buchardt O. 1991. Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science* 254:1497–500
76. Egholm M, Buchardt O, Nielsen PE, Berg RH. 1992. Peptide nucleic acids (PNA). Oligonucleotide analogues with an achiral peptide backbone. *J. Am. Chem. Soc.* 114:1895–97
77. Mittapalli GK, Reddy KR, Xiong H, Munoz O, Han B, et al. 2007. Mapping the landscape of potentially primordial informational oligomers: oligodipeptides and oligodipeptoids tagged with triazines as recognition elements. *Angew. Chem. Int. Ed.* 46:2470–77
78. Szostak JW. 2012. The eightfold path to non-enzymatic RNA replication. *J. Syst. Chem.* 3:2

79. McCall MJ, Hendry P, Mir AA, Conaty J, Brown G, Lockett TJ. 2000. Small, efficient hammerhead ribozymes. *Mol. Biotechnol.* 14:5–17
80. Xia T, SantaLucia J Jr, Burkard M, Kierzek R, Schroeder S, et al. 1998. Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson–Crick base pairs. *Biochemistry* 37:14719–35
81. Gryaznov S, Lloyd DH, Chen JK, Schultz RG, DeDionisio LA, et al. 1995. Oligonucleotide N3′-P5′ phosphoramidates. *Proc. Natl. Acad. Sci. USA* 92:5798–802
82. Engelhart AE, Powner MW, Szostak JW. 2013. Functional RNAs exhibit tolerance for non-heritable 2′–5′ versus 3′–5′ backbone heterogeneity. *Nat. Chem.* 5:390–94
83. Adamala K, Szostak JW. 2013. Non-enzymatic template-directed RNA synthesis inside model protocells. *Science* 342:1098–100
84. Ferris JP, Huang CH, Hagan WJ. 1989. *N*-Cyanoimidazole and diimidazole imine: water-soluble condensing agents for the formation of the phosphodiester bond. *Nucleosides Nucleotides* 8:407–14
85. Dolinnaya NG, Sokolova NI, Ashirbekova DT, Shabarova ZA. 1991. The use of BrCN for assembling modified DNA duplexes and DNA–RNA hybrids: comparison with water-soluble carbodiimide. *Nucleic Acids Res.* 19:3067–72
86. Gilham PT. 1962. An addition reaction specific for uridine and guanosine nucleotides and its application to the modification of ribonuclease action. *J. Am. Chem. Soc.* 84:687–88
87. Chu BC, Wahl GM, Orgel LE. 1983. Derivatization of unprotected polynucleotides. *Nucleic Acids Res.* 11:6513–29
88. Kruger K, Grabowski PJ, Zaug AJ, Sands J, Gottschling DE, Cech TR. 1982. Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of tetrahymena. *Cell* 31:147–57
89. Guerrier-Takada C, Gardiner K, Marsh T, Pace N, Altman S. 1983. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* 35:849–57
90. Gilbert W. 1986. The RNA world. *Nature* 319:618–18
91. Been MD, Cech TR. 1988. RNA as an RNA polymerase: net elongation of an RNA primer catalyzed by the tetrahymena ribozyme. *Science* 239:1412–16
92. Doudna J, Szostak JW. 1989. RNA-catalyzed synthesis of complementary-strand RNA. *Nature* 339:519–22
93. Bartel DP, Doudna JA, Usman N, Szostak JW. 1991. Template-directed primer extension catalyzed by the tetrahymena ribozyme. *Mol. Cell. Biol.* 11:3390–94
94. Doudna J, Couture S, Szostak JW. 1991. A multisubunit ribozyme that is a catalyst of and template for complementary strand RNA synthesis. *Science* 251:1605–8
95. Green R, Szostak JW. 1992. Selection of a ribozyme that functions as a superior template in a self-copying reaction. *Science* 258:1910–15
96. Bartel DP, Szostak JW. 1993. Isolation of new ribozymes from a large pool of random sequences. *Science* 261:1411–18
97. Ekland E, Bartel DP. 1996. RNA-catalysed RNA polymerization using nucleoside triphosphates. *Nature* 382:373–76
98. McGinness K, Wright M, Joyce G. 2002. Continuous in vitro evolution of a ribozyme that catalyzes three successive nucleotidyl addition reactions. *Chem. Biol.* 9:585–96
99. Johnston W, Unrau P, Lawrence M, Glasner M, Bartel DP. 2001. RNA-catalyzed RNA polymerization: accurate and general RNA-templated primer extension. *Science* 292:1319–25
100. Zaher HS, Unrau PJ. 2007. Selection of an improved RNA polymerase ribozyme with superior extension and fidelity. *RNA* 13:1017–26
101. Wochner A, Attwater J, Coulson A, Holliger P. 2011. Ribozyme-catalyzed transcription of an active ribozyme. *Science* 332:209–12
102. Lawrence M, Bartel DP. 2003. Processivity of ribozyme-catalyzed RNA polymerization. *Biochemistry* 42:8748–55
103. Shechner DM, Grant RA, Bagby SC, Koldobskaya Y, Piccirilli JA, Bartel DP. 2009. Crystal structure of the catalytic core of an RNA-polymerase ribozyme. *Science* 326:1271–75
104. Shechner DM, Bartel DP. 2011. The structural basis of RNA-catalyzed RNA polymerization. *Nat. Struct. Mol. Biol.* 18:1036–42

105. Wang QS, Cheng LKL, Unrau PJ. 2011. Characterization of the B6.61 polymerase ribozyme accessory domain. *RNA* 17:469–77
106. Lawrence M, Bartel DP. 2005. New ligase-derived RNA polymerase ribozymes. *RNA* 11:1173–80
107. McGinness KE, Joyce GF. 2002. RNA-catalyzed RNA ligation on an external RNA template. *Chem. Biol.* 9:297–307
108. Vicens Q, Cech TR. 2009. A natural ribozyme with 3′,5′ RNA ligase activity. *Nat. Chem. Biol.* 5:97–99
109. Rogers J, Joyce GF. 2001. The effect of cytidine on the structure and function of an RNA ligase ribozyme. *RNA* 7:395–404
110. Paul N, Joyce GF. 2002. A self-replicating ligase ribozyme. *Proc. Natl. Acad. Sci. USA* 99:12733–40
111. Kim D-E, Joyce GF. 2004. Cross-catalytic replication of an RNA ligase ribozyme. *Chem. Biol.* 11:1505–12
112. Lincoln TA, Joyce GF. 2009. Self-sustained replication of an RNA enzyme. *Science* 323:1229–32
113. Ferretti AC, Joyce GF. 2013. Kinetic properties of an RNA enzyme that undergoes self-sustained exponential amplification. *Biochemistry* 52:1227–35
114. Sczepanski JT, Joyce GF. 2012. Synthetic evolving systems that implement a user-specified genetic code of arbitrary design. *Chem. Biol.* 19:1324–32
115. Riley CA, Lehman N. 2003. Generalized RNA-directed recombination of RNA. *Chem. Biol.* 10:1233–43
116. Hayden EJ, Lehman N. 2006. Self-assembly of a group I intron from inactive oligonucleotide fragments. *Chem. Biol.* 13:909–18
117. Hayden EJ, von Kiedrowski G, Lehman N. 2008. Systems chemistry on ribozyme self-construction: evidence for anabolic autocatalysis in a recombination network. *Angew. Chem. Int. Ed.* 47:8424–28
118. Vaidya N, Manapat ML, Chen IA, Xulvi-Brunet R, Hayden EJ, Lehman N. 2012. Spontaneous network formation among cooperative RNA replicators. *Nature* 491:72–77
119. Chen I, Roberts R, Szostak JW. 2004. The emergence of competition between model protocells. *Science* 305:1474–76
120. Oberholzer T, Albrizio M, Luisi PL. 1995. Polymerase chain reaction in liposomes. *Chem. Biol.* 2:677–82
121. Oberholzer T, Wick R, Luisi PL, Biebricher CK. 1995. Enzymatic RNA replication in self-reproducing vesicles: an approach to a minimal cell. *Biochem. Biophys. Res.* 207:250–57
122. Guatelli JC, Whitfield KM, Kwoh DY, Barringer KJ, Richman DD, Gingeras TR. 1990. Isothermal, in vitro amplification of nucleic acids by a multienzyme reaction modeled after retroviral replication. *Proc. Natl. Acad. Sci. USA* 87:1874–78
123. Wright M, Joyce G. 1997. Continuous in vitro evolution of catalytic function. *Science* 276:614–17
124. Kurihara K, Tamura M, Shohda K-I, Toyota T, Suzuki K, Sugawara T. 2011. Self-reproduction of supramolecular giant vesicles combined with the amplification of encapsulated DNA. *Nat. Chem.* 3:775–81
125. Chakrabarti AC, Breaker RR, Joyce GF, Deamer DW. 1994. Production of RNA by a polymerase protein encapsulated within phospholipid vesicles. *J. Mol. Evol.* 39:555–59
126. Haruna I, Spiegelman S. 1965. Autocatalytic synthesis of a viral RNA in vitro. *Science* 150:884–86
127. Mills D, Peterson R, Spiegelman S. 1967. An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule. *Proc. Natl. Acad. Sci. USA* 58:217–24
128. Kita H, Matsuura T, Sunami T, Hosoda K, Ichihashi N, et al. 2008. Replication of genetic information with self-encoded replicase in liposomes. *ChemBioChem* 9:2403–10
129. Ichihashi N, Matsuura T, Kita H, Hosoda K, Sunami T, et al. 2008. Importance of translation-replication balance for efficient replication by the self-encoded replicase. *ChemBioChem* 9:3023–28
130. Urabe H, Ichihashi N, Matsuura T, Hosoda K, Kazuta Y, et al. 2010. Compartmentalization in a water-in-oil emulsion repressed the spontaneous amplification of RNA by Qβ replicase. *Biochemistry* 49:1809–13
131. Bansho Y, Ichihashi N, Kazuta Y, Matsuura T, Suzuki H, Yomo T. 2012. Importance of parasite RNA species repression for prolonged translation-coupled RNA self-replication. *Chem. Biol.* 19:478–87
132. Oberholzer T, Nierhaus KH, Luisi PL. 1999. Protein expression in liposomes. *Biochem. Biophys. Res.* 261:238–41
133. Noireaux V, Libchaber A. 2004. A vesicle bioreactor as a step toward an artificial cell assembly. *Proc. Natl. Acad. Sci. USA* 101:17669–74
134. Forster AC, Church GM. 2006. Towards synthesis of a minimal cell. *Mol. Syst. Biol.* 2:45

135. Luisi PL, Ferri F, Stano P. 2005. Approaches to semi-synthetic minimal cells: a review. *Naturwissenschaften* 93:1–13
136. Szathmáry E, Smith JM. 1995. The major evolutionary transitions. *Nature* 374:227–32
137. Haines TH. 1983. Anionic lipid headgroups as a proton-conducting pathway along the surface of membranes: a hypothesis. *Proc. Natl. Acad. Sci. USA* 80:160–64
138. Cistola DP, Hamilton JA, Jackson D, Small DM. 1988. Ionization and phase behavior of fatty acids in water: application of the Gibbs phase rule. *Biochemistry* 27:1881–88
139. Tanford C. 1980. *The Hydrophobic Effect: Formation of Micelles and Biological Membranes*. New York: Wiley
140. Chen IA, Szostak JW. 2004. A kinetic study of the growth of fatty acid vesicles. *Biophys. J.* 87:988–98
141. Budin I, Szostak JW. 2011. Physical effects underlying the transition from primitive to modern cell membranes. *Proc. Natl. Acad. Sci. USA* 108:5249–54
142. Walde P, Goto A, Monnard PA, Wessicken M, Luisi PL. 1994. Oparin's reactions revisited: enzymatic synthesis of poly(adenylic acid) in micelles and self-reproducing vesicles. *J. Am. Chem. Soc.* 116:7541–47
143. Zhu TF, Szostak JW. 2009. Coupled growth and division of model protocell membranes. *J. Am. Chem. Soc.* 131:5705–13
144. Zhu TF, Adamala K, Zhang N, Szostak JW. 2012. Photochemically driven redox chemistry induces protocell membrane pearling and division. *Proc. Natl. Acad. Sci. USA* 109:9828–32
145. Budin I, Debnath A, Szostak JW. 2012. Concentration-driven growth of model protocell membranes. *J. Am. Chem. Soc.* 134:20812–19
146. Hanczyc MM, Fujikawa SM, Szostak JW. 2003. Experimental models of primitive cellular compartments: encapsulation, growth, and division. *Science* 302:618–22
147. Mansy SS, Schrum JP, Krishnamurthy M, Tobé S, Treco DA, Szostak JW. 2008. Template-directed synthesis of a genetic polymer in a model protocell. *Nature* 454:122–25
148. Sacerdote M, Szostak JW. 2005. Semipermeable lipid bilayers exhibit diastereoselectivity favoring ribose. *Proc. Natl. Acad. Sci. USA* 102:6004–8
149. Mansy SS, Szostak JW. 2008. Thermostability of model protocell membranes. *Proc. Natl. Acad. Sci. USA* 105:13351–55
150. Monnard PA, Apel CL, Kanavarioti A, Deamer DW. 2002. Influence of ionic inorganic solutes on self-assembly and polymerization processes related to early forms of life: implications for a prebiotic aqueous medium. *Astrobiology* 2:139–52
151. Bruckner RJ, Mansy SS, Ricardo A, Mahadevan L, Szostak JW. 2009. Flip-flop-induced relaxation of bending energy: implications for membrane remodeling. *Biophys. J.* 97:3113–22
152. Apel CL, Deamer DW, Mautner MN. 2002. Self-assembled vesicles of monocarboxylic acids and alcohols: conditions for stability and for the encapsulation of biopolymers. *Biochim. Biophys. Acta* 1559:1–9
153. Yuen GU, Kvenvolden KA. 1973. Monocarboxylic acids in Murray and Murchison carbonaceous meteorites. *Nature* 246:301–3
154. Naraoka H, Shimoyama A, Harada K. 1999. Molecular distribution of monocarboxylic acids in Asuka carbonaceous chondrites from Antarctica. *Orig. Life Evol. B* 29:187–201
155. Rushdi AI, Simoneit BR. 2001. Lipid formation by aqueous Fischer–Tropsch-type synthesis over a temperature range of 100 to 400 degrees C. *Orig. Life Evol. B* 31:103–18
156. Deamer DW, Barchfield GL. 1982. Encapsulation of macromolecules by lipid vesicles under simulated prebiotic conditions. *J. Mol. Evol.* 18:203–6
157. Jesorka A, Orwar O. 2008. Liposomes: technologies and analytical applications. *Annu. Rev. Anal. Chem.* 1:801–32
158. Müller M, Zschörnig O, Ohki S, Arnold K. 2003. Fusion, leakage and surface hydrophobicity of vesicles containing phosphoinositides: influence of steric and electrostatic effects. *J. Membr. Biol.* 192:33–43
159. Terasawa H, Nishimura K, Suzuki H, Matsuura T, Yomo T. 2012. Coupling of the fusion and budding of giant phospholipid vesicles containing macromolecules. *Proc. Natl. Acad. Sci. USA* 109:5942–47
160. Johnson JM, Ha T, Chu S, Boxer SG. 2002. Early steps of supported bilayer formation probed by single vesicle fluorescence assays. *Biophys. J.* 83:3371–79
161. Sunami T, Caschera F, Morita Y, Toyota T, Nishimura K, et al. 2010. Detection of association and fusion of giant vesicles using a fluorescence-activated cell sorter. *Langmuir* 26:15098–103

162. Kuruma Y, Stano P, Ueda T, Luisi PL. 2009. A synthetic biology approach to the construction of membrane proteins in semi-synthetic minimal cells. *Biochim. Biophys. Acta* 1788:567–74
163. Murtas G. 2010. Internal lipid synthesis and vesicle growth as a step toward self-reproduction of the minimal cell. *Syst. Synth. Biol.* 4:85–93
164. Andes-Koback M, Keating CD. 2011. Complete budding and asymmetric division of primitive model cells to produce daughter vesicles with different interior and membrane compositions. *J. Am. Chem. Soc.* 133:9545–55
165. Staneva G, Angelova MI, Koumanov K. 2004. Phospholipase A2 promotes raft budding and fission from giant liposomes. *Chem. Phys. Lipids* 129:53–62
166. Takakura K, Sugawara T. 2004. Membrane dynamics of a myelin-like giant multilamellar vesicle applicable to a self-reproducing system. *Langmuir* 20:3832–34
167. Takakura K, Toyota T, Sugawara T. 2003. A novel system of self-reproducing giant vesicles. *J. Am. Chem. Soc.* 125:8134–40
168. Takahashi H, Kageyama Y, Kurihara K, Takakura K, Murata S, Sugawara T. 2009. Autocatalytic membrane-amplification on a pre-existing vesicular surface. *Chem. Commun.* 46:8791–93
169. Kamat NP, Katz JS, Hammer DA. 2011. Engineering polymersome protocells. *J. Phys. Chem. Lett.* 2:1612–23
170. Discher BM, Won YY, Ege DS, Lee JC, Bates FS, et al. 1999. Polymersomes: tough vesicles made from diblock copolymers. *Science* 284:1143–46
171. Holowka EP, Pochan DJ, Deming TJ. 2005. Charged polypeptide vesicles with controllable diameter. *J. Am. Chem. Soc.* 127:12423–28
172. Percec V, Wilson DA, Leowanawat P, Wilson CJ, Hughes AD, et al. 2010. Self-assembly of Janus dendrimers into uniform dendrimersomes and other complex architectures. *Science* 328:1009–14
173. Robbins GP, Jimbo M, Swift J, Therien MJ, Hammer DA, Dmochowski IJ. 2009. Photoinitiated destruction of composite porphyrin-protein polymersomes. *J. Am. Chem. Soc.* 131:3872–74
174. Mabrouk E, Cuvelier D, Brochard-Wyart F, Nassoy P, Li M-H. 2009. Bursting of sensitive polymersomes induced by curling. *Proc. Natl. Acad. Sci. USA* 106:7294–98
175. Kamat NP, Robbins GP, Rawson JS, Therien MJ, Dmochowski IJ, Hammer DA. 2010. A generalized system for photo-responsive membrane rupture in polymersomes. *Adv. Funct. Mater.* 20:2588–96
176. Tawfik DS, Griffiths AD. 1998. Man-made cell-like compartments for molecular evolution. *Nat. Biotechnol.* 16:652–56
177. Pietrini AV, Luisi PL. 2004. Cell-free protein synthesis through solubilisate exchange in water/oil emulsion compartments. *ChemBioChem* 5:1055–62
178. Ichihashi N, Usui K, Kazuta Y, Sunami T, Matsuura T, Yomo T. 2013. Darwinian evolution of a translation-coupled RNA self-replication system in a cell-like compartment. *Nat. Commun.* 4:2494
179. Attwater J, Wochner A, Pinheiro VB, Coulson A, Holliger P. 2010. Ice as a protocellular medium for RNA replication. *Nat. Commun.* 1:76
180. Monnard PA, Kanavarioti A, Deamer DW. 2003. Eutectic phase polymerization of activated ribonucleotide mixtures yields quasi-equimolar incorporation of purine and pyrimidine nucleobases. *J. Am. Chem. Soc.* 125:13734–40
181. Strulson CA, Molden RC, Keating CD, Bevilacqua PC. 2012. RNA catalysis through compartmentalization. *Nat. Chem.* 4:941–46
182. Koga S, Williams DS, Perriman AW, Mann S. 2011. Peptide–nucleotide microdroplets as a step towards a membrane-free protocell model. *Nat. Chem.* 3:720–24
183. Szostak JW, Bartel DP, Luisi PL. 2001. Synthesizing life. *Nature* 409:387–90
184. Fukuda H, Goto A, Yoshioka H, Goto R, Morigaki K, Walde P. 2001. Electron spin resonance study of the pH-induced transformation of micelles to vesicles in an aqueous oleic acid/oleate system. *Langmuir* 17:4223–31
185. Maurer SE, Deamer DW, Boncella JM, Monnard PA. 2009. Chemical evolution of amphiphiles: glycerol monoacyl derivatives stabilize plausible prebiotic membranes. *Astrobiology* 9:979–87
186. Wu TF, Orgel LE. 1992. Nonenzymatic template-directed synthesis on oligodeoxycytidylate sequences in hairpin oligonucleotides. *J. Am. Chem. Soc.* 114:317–22

187. Paula S, Volkov AG, Van Hoek AN, Haines TH, Deamer DW. 1996. Permeation of protons, potassium ions, and small polar molecules through phospholipid bilayers as a function of membrane thickness. *Biophys. J.* 70:339–48
188. Kanehisa MI, Tsong TY. 1978. Cluster model of lipid phase transitions with application to passive permeation of molecules and structure relaxations in lipid bilayers. *J. Am. Chem. Soc.* 100:424–32
189. Monnard PA, Luptak A, Deamer DW. 2007. Models of primitive cellular life: polymerases and templates in liposomes. *Philos. Trans. R. Soc. B* 362:1741–50
190. Yoshimoto M, Wang S, Fukunaga K, Fournier D, Walde P, et al. 2005. Novel immobilized liposomal glucose oxidase system using the channel protein OmpF and catalase. *Biotechnol. Bioeng.* 90:231–38
191. Nourian Z, Roelofsen W, Danelon C. 2012. Triggered gene expression in fed-vesicle microreactors with a multifunctional membrane. *Angew. Chem. Int. Ed.* 51:3114–18
192. Treyer M, Walde P, Oberholzer T. 2002. Permeability enhancement of lipid vesicles to nucleotides by use of sodium cholate: basic studies and application to an enzyme-catalyzed reaction occurring inside the vesicles. *Langmuir* 18:1043–50
193. Sunami T, Hosoda K, Suzuki H, Matsuura T, Yomo T. 2010. Cellular compartment model for exploring the effect of the lipidic membrane on the kinetics of encapsulated biochemical reactions. *Langmuir* 26:8544–51
194. Adamala K, Szostak JW. 2013. Competition between model protocells driven by an encapsulated catalyst. *Nat. Chem.* 5:495–501
195. Pohorille A, Deamer DW. 2002. Artificial cells: prospects for biotechnology. *Trends Biotechnol.* 20:123–28
196. Yu W, Sato K, Wakabayashi M, Nakaishi T, Ko-Mitamura EP, et al. 2001. Synthesis of functional protein in liposome. *J. Biosci. Bioeng.* 92:590–93
197. Nomura SM, Tsumoto K, Hamada T, Akiyoshi K, Nakatani Y, Yoshikawa K. 2003. Gene expression within cell-sized lipid vesicles. *ChemBioChem* 4:1172–75
198. Murtas G, Kuruma Y, Bianchini P, Diaspro A, Luisi PL. 2007. Protein synthesis in liposomes with a minimal set of enzymes. *Biochem. Biophys. Res.* 363:12–17
199. Shimizu Y, Inoue A, Tomari Y, Suzuki T, Yokogawa T, et al. 2001. Cell-free translation reconstituted with purified components. *Nat. Biotechnol.* 19:751–55
200. Kuruma Y, Suzuki T, Ono S, Yoshida M, Ueda T. 2012. Functional analysis of membranous F_o-*a* subunit of F_1F_o-ATP synthase by in vitroprotein synthesis. *Biochem. J.* 442:631–38
201. Leaver M, Domínguez-Cuevas P, Coxhead JM, Daniel RA, Errington J. 2009. Life without a wall or division machine in *Bacillus subtilis*. *Nature* 457:849–53
202. Allan EJ, Hoischen C, Gumpert J. 2009. Bacterial L-forms. *Adv. Appl. Microbiol.* 68:1–39
203. Ricardo A, Szostak JW. 2009. Origin of life on earth. *Sci. Am.* 301:54–61
204. Lee JF, Hesselberth JR, Meyers LA, Ellington AD. 2004. Aptamer database. *Nucleic Acids Res.* 32:D95–100
205. Liu J, Cao Z, Lu Y. 2009. Functional nucleic acid sensors. *Chem. Rev.* 109:1948–98
206. Ke Y, Ong LL, Shih WM, Yin P. 2012. Three-dimensional structures self-assembled from DNA bricks. *Science* 338:1177–83
207. Qian L, Winfree E. 2011. Scaling up digital circuit computation with DNA strand displacement cascades. *Science* 332:1196–201
208. Hagan CL, Westwood DB, Kahne D. 2013. Bam lipoproteins assemble BamA in vitro. *Biochemistry* 52:6108–13
209. Bohannon J. 2010. Mirror-image cells could transform science—or kill us all. *Wired*, Nov. 29. **http://www.wired.com/magazine/2010/11/ff_mirrorlife/**
210. Leduc PR, Wong MS, Ferreira PM, Groff RE, Haslinger K, et al. 2007. Towards an in vivo biologically inspired nanofactory. *Nat. Nanotechnol.* 2:3–7
211. Ro D, Paradise E, Ouellet M, Fisher K, Newman K, et al. 2006. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* 440:940–43
212. Rabinovitch-Deere CA, Oliver JWK, Rodriguez GM, Atsumi S. 2013. Synthetic biology and metabolic engineering approaches to produce biofuels. *Chem. Rev.* 113:4611–32
213. Chang TM. 1999. Future prospects for artificial blood. *Trends Biotechnol.* 17:61–67