13

SELF-REPLICATION IN CHEMISTRY AND BIOLOGY

Philipp Holliger and David Loakes

MRC Laboratory of Molecular Biology, Cambridge CB2 2QH, UK

13.1 INTRODUCTION: SELF-REPLICATION, FIDELITY, AND HEREDITY

Self-replication involves the product-directed assembly of components to form a new product; in its simplest form, it involves the joining of just two components. The product acts as a template both to correctly position the two components and to allow for efficient joining of them. The newly formed product can then dissociate to provide a new template for further replication [1,2] (Fig. 13-1).

Important concepts for consideration are fidelity and heredity. Self-replication may be perfect, in which case all products (all "offspring") are identical to the template (the parents), or imperfect, in which case, offspring may differ from their parents. The degree of perfection (or imperfection) of self-replication is called the fidelity, which varies greatly among self-replicating systems. High-fidelity replication denotes a system where only few alterations are introduced into the offspring molecules while low-fidelity replicators will produce a great deal of variation in their offspring. Depending on the system these variations can again be transmitted through the next self-replication cycle. In such a case, the self-replication system will display heredity. Self-replication with heredity is a fundamental property of life and a prerequisite for evolution.

The most widely studied self-replicating systems involve nucleic acids and these are most relevant to extant or plausible primordial biological systems, as nucleic acids are uniquely suited for self-replication, heredity, and evolution. We also briefly discuss chemical replicators based on autocatalytic networks or template-driven replicators
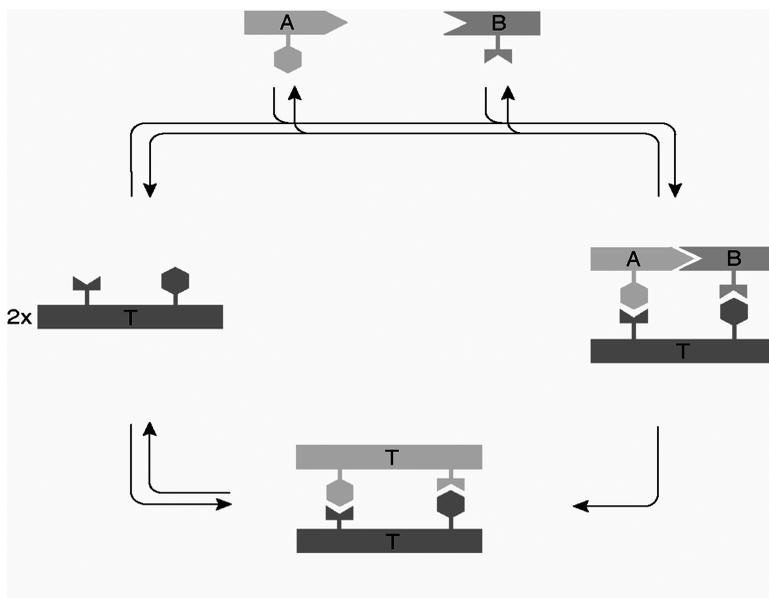
Figure 13-1 Scheme of a simple self-replication system (after [1,2]). A and B, building blocks; T, template.

with limited heredity (e.g., peptides, prions). However, replicators based on digital (e.g., computer viruses) or cultural heredity (e.g., memes), or self-replicating macroscopic machines are beyond the scope of this review, despite striking progress in the latter field [3]. The review will consider primarily recent literature referring to older literature only when necessary, and is not meant to be exhaustive.

13.2 CHEMICAL SYSTEMS CAPABLE OF SELF-REPLICATION

For a number of years chemists have explored a host of molecular systems with autocatalytic and dynamic combinatorial properties. Some of these are capable of templating and catalyzing their own synthesis, that is, catalyze the product-directed synthesis of more products from its constituent parts. Due to the limited complexity of such systems, either self-replication is perfect or alterations (side reactions) are nonhereditary as they would interfere with the self-replication ability. One of the earliest reported is a nucleoside-based system, described by Rebek and colleagues, and involves hydrogen-bonding and stacking interactions in organic media [4–8] (Fig. 13-2).

There have been a number of reports of autocatalytic Diels–Alder reactions. The bicyclic transition state that is formed offers a basis for efficient self-replication, that is, for the transfer of chemical information coded in terms of both regio- and stereoselectivity. In this example, the diene is also chiral, thus allowing for the transfer of diastereomeric information. Following this first published report of
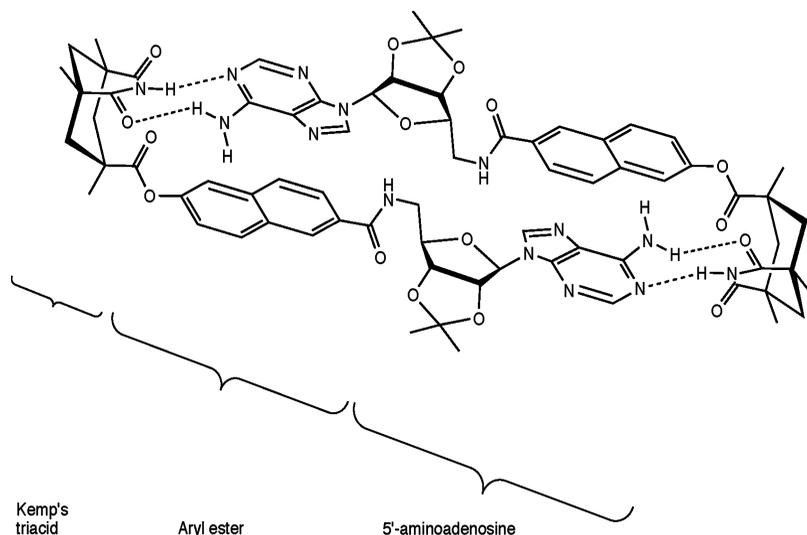
Figure 13-2 One of the earliest self-replicating nucleoside-based systems involves hydrogen-bonding and aryl-stacking interactions in organic media. 5′-Aminoadenosine reacts with the aryl-pentafluorophenyl ester derivative of Kemp's triacid, which then acts as catalyst for further coupling reactions. The reaction involves hydrogen bonding of adenosine to the imine, as blocking of the imine NH group leads to a 10-fold drop in catalytic rate.

the autocatalytic Diels–Alder reaction, it has been demonstrated that using a chiral-starting material only one of the four possible diastereoisomers were formed [9]. Other groups have examined self-replicating Diels–Alder reactions [10,11] including von Kiedrowski [9]. In addition to the advantages of being capable of self-replicating by homochiral autocatalysis and heterochiral cross-catalysis, they are much more efficient replicators than nucleic acids or peptides, giving rise to almost exponential replication.

13.3 PEPTIDE SELF-REPLICATION

Peptides of a certain length fold spontaneously into three-dimensional structures defined by their sequence. These in turn may specifically associate with other peptides in defined oligomeric complexes. Peptide self-replication is generally based on a peptide A acting as a template and promoting the template-directed ligation of two smaller isomorphic peptides. In its simplest form, the two smaller peptides are fragments of A (A′, A″) and ligation thus produces further copies of the parent peptide in a homodimeric complex (Fig. 13-3).

The first self-replicating peptide described was a 32-residue α-helical coiled-coil peptide based on the leucine-zipper domain of the yeast transcription factor GCN4 [12]. It has been shown to promote thioester-mediated amide bond formation

A

A'   O   SH   A"
        SBn
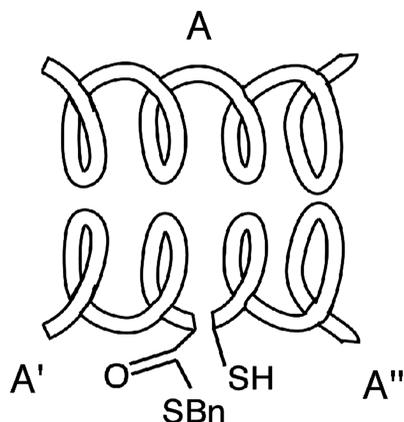
Figure 13-3 Coiled-coil peptide self-replicators are able to promote template-directed ligation of two smaller peptides (A′, A″) producing a further copy of the parent peptide (A). Typical coupling reactions occur with carbodiimide or thioester (shown) chemistry.

between a 15- and 17mer fragments in neutral aqueous conditions. It was subsequently shown that this system was also able to distinguish between native and mutant precursor peptides. When mutant peptides are used there is a catalytic step in which mutant progeny are produced, but due to changes in the hydrophobic core of the precursor peptides the correct peptide may be preferentially produced [13,14]. This peptide replicator is also capable of chiral selectivity by efficiently amplifying homochiral products from a racemic mixture of peptide fragments [15] and is capable of discriminating between structures containing only a single chiral mutation. The system demonstrates a dynamic stereochemical editing function whereby heterochiral sequences promote the production of homochiral products. Thus, the peptide replicator system demonstrates the emergence of fidelity of replication.

The self-replicating peptide described by Ghadiri has been computationally analyzed where it was found that the dynamics are governed principally by two reversible hydrophobic interactions between the template and a peptide fragment and between two template molecules [16]. The association of two template molecules was found to be most favorable leading to a build up of the inactive template dimer in the autocatalytic step, thus limiting the self-replication. Analysis of the heterochiral system described by Ghadiri [15] indicated that cross-catalytic processes involving D- and L-species play a significant role. Chiral amplification is mainly due to the formation of *meso*-like species, leading to an enantiomeric excess in the final product [17,18].

Chmielewski has developed a self-replicating peptide that is pH-dependent. The sequence contains glutamic acid side chains such that at physiological pH the peptide is a random coil. However, under acidic conditions the peptide adopts a coiled-coil structure, similar to that developed by Ghadiri, which is then able to promote self-replication [19]. As noted above, the self-replication of peptides is limited as the most

stable species is the template dimer. Chmielewski has therefore designed a self-replicating peptide containing a proline residue in place of one of the glutamic acid residues in the pH-dependent replicator. The presence of the proline causes a kink in the coiled-coil structure, which allows for more efficient separation of the template dimer species, thus improving the efficiency of self-replication [20].

Finally, there are two examples of cross-replication between peptides and nucleic acids. In the first example, L-α-amino-γ-nucleobase-butyric acids (NBAs) were substituted into peptides adopting coiled-coil structures to enhance peptide recognition. Templates and fragments were then synthesized containing complementary adenine–thymine or guanine–cytosine sequences at various positions within the peptide. While it was found that the effect of NBAs in the peptide was sequence dependent, it was shown that the increased recognition architecture could be used to design more efficient self-replicating peptides [21]. In the second example, Ellington has examined the ligation of short oligonucleotides by a peptide [22]. Using a 17mer arginine-rich motif (ARM), a 35mer anti-REV RNA aptamer was developed for ligation studies. Aptamer half-molecules bearing a 5'-iodine and a 3'-phosphorothioate could be chemically ligated by cyanogen bromide in the presence of the ARM peptide.

The systems described in this section may also have relevance for the prebiotic synthesis of peptides. It has been demonstrated that amino acids can adsorb to mineral surfaces where they undergo chemical ligation to form random polypeptide species. Together with self-replication, this may provide a process for the selective enrichment of a defined set of peptide sequences. All of the self-replicating peptides described so far adopt α-helical coiled-coil structures, but Ghadiri has speculated that self-replication through β-sheet motifs are also likely [23,24].

Interesting examples of peptide self-replication are provided by prions. These are a number of metastable proteins, which can be converted to a misfolded insoluble form. The insoluble form is capable of catalyzing the conversion of soluble prion protein into the insoluble form. In some cases the insoluble form is infectious and can be transmitted within and across species giving rise to so-called transmissible spongiform encephalopathies (TSEs), of which "mad cow disease" (BSE) is the best known. Intriguingly, prions display heredity in the form of strain and species specific characteristics, which appear to be encoded in the conformation of the prion protein [25]. In yeast, these can provide diverse, heritable phenotypes that are beneficial under certain circumstances. Indeed it has been proposed that prions may act as an epigenetic switch in yeast and fungi or even as a form of molecular long-term memory in the nervous system of *A. californica* [26].
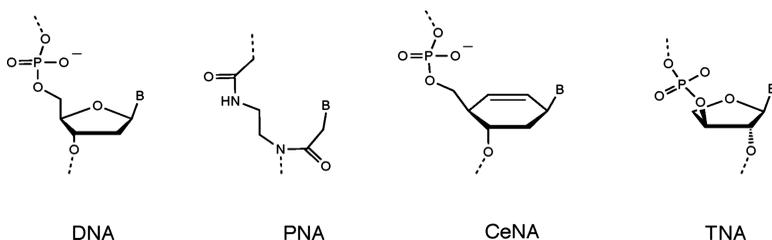
13.4 NUCLEIC ACIDS

In 1953, Watson and Crick published their seminal article on the structure of DNA, which ends with the now famous understatement "It has not escaped our notice that the specific pairing we have postulated immediately suggests a copying mechanism for the genetic material" [27]. Indeed, another 50 years of research into the structure and

function of DNA (and RNA) have not only confirmed the proposed semiconservative mode of replication, whereby one strand of DNA acts as a template for synthesis of the opposing strand (and vice versa), but also has revealed that DNA and RNA are singularly suited as molecules for information storage and transmission, for replication and heredity [28]. For one, in nucleic acids, the polyanionic phosphate backbone dominates the physicochemical properties of the molecule (e.g., solubility) to such an extent that changes to neither base composition nor sequence have much effect. In other words, in sharp contrast to, for example, proteins, nucleic acids display similar properties (e.g., solubility) regardless of the sequence, that is, the information encoded within. Furthermore, charge repulsion along the polyphosphate backbone favors an extended conformation of nucleic acid polymers facilitating their templating function in replication and read-out of the hydrogen-bonding pattern at the Watson–Crick face of the bases. Finally, nucleic acid polymorphism is constrained to essentially just two apomorphic classes, A and B (there is also a left-handed helix system, Z, which occurs only under certain conditions and is restricted to alternating purine–pyrimidine (GC) sequences).

13.4.1   Altered Backbones

A- and B-form nucleic acids arise as a result of the restricted spectrum of furanose (ribose or deoxyribose) sugar conformations. This relative inflexibility provides a stable scaffold for the nucleobases and is essential for duplex stability. It is therefore not surprising that many modifications to backbone chemistry have led to nucleic acids that are no longer capable of forming stable duplex structures with either DNA or RNA or themselves. A notable exception is peptide nucleic acids (PNAs), in which the ribofuranose-phosphate backbone of DNA/RNA is replaced by N-(2-aminoethyl)-glycine (Fig. 13-4). PNAs can hybridize specifically and extraordinarily strongly to DNA and RNA making them of significant use in both antisense and antigene strategies [29]. However, longer PNAs can be poorly soluble. Nevertheless, PNA can be used in information transfer to DNA and RNA and it has been proposed that PNA may have been involved in prebiotic evolution [34–36] (see later).

DNA   PNA   CeNA   TNA

Figure 13-4   Various nucleic acid backbone modifications have been examined as alternative genetic systems. These include peptide nucleic acid, cyclohexene nucleic acid (CeNA) and α-L-threose nucleic acid (TNA).

Another example is morpholino nucleosides that are neutral analogues of DNA in which the sugar is substituted by a morpholine ring [30]. They have been shown to bind well with RNA, and have therefore been a subject for investigation in the field of antisense therapy [31]. However, they are poor substrates for enzymes, including RNase-H, and act by a steric-blocking mechanism.

13.4.2 Altered Sugars

Nucleic acid chemists have also synthesized a variety of modifications to the ribofuranose sugars in DNA and RNA in an attempt to modify hybridization properties and study the determinants of Watson–Crick-directed duplex formation. For example, Orgel, Herdewijn, and colleagues have investigated the properties of hexose sugars and have demonstrated nonenzymatic information transfer from nucleic acids derived from 1,5-anhydrohexitol nucleosides (HNA) [32–34] and altritol nucleosides (ANA) [35]. A number of other hexopyranosyl- and pentopyranosyl-nucleoside systems have been studied by Eschenmoser [36]. These systems show a remarkable spectrum of hybridization properties in not only self-pairing systems but also cross-pairing with DNA and RNA.

Cyclohexane- and cyclohexene-nucleic acid systems are conformationally flexible nucleic acid mimics that can hybridize with themselves, DNA and RNA (Fig. 13-4) [37–39]. The self-pairing system is more stable than that with DNA, but is most stable with RNA. While they are yet to be shown to be of use in information transfer they are recognized by some enzymes as they can RNase-H activity when hybridized with RNA, and are therefore of use as antisense agents [40].

Another sugar modification that displays interesting properties is the tetrofuranose α-L-threose nucleic acids (TNA) (Fig. 13-4) [41], which forms specific base pairs with itself, DNA, and RNA. Furthermore, TNA has been shown to be functional in replacing RNA as part of an RNA cleaving ribozyme, albeit with somewhat reduced activity [42]. Finally, TNA templates and TNA triphosphates have been shown to be reasonable substrates for various DNA and RNA polymerises [43–46] and like PNA has been proposed to be involved in prebiotic evolution (see Section 13.6).

13.4.3 Altered Bases

For the reasons discussed above, the introduction of alternative base pairing systems into nucleic acids is much less problematic than alterations to the sugar–phosphate backbone structure. Such systems would not give rise to alternative nucleic acid structures, but may be used to introduce alternative or additional information content into nucleic acids without altering their overall structure. As a result, nucleic acids comprising modified bases are often better substrates for enzymatic replication. The challenge here lies in devising alternative systems, which are both orthogonal to the canonical bases as well as specific in recognition.

There are limited ways in which such novel base-pairing schemes can be devised. Nucleobases can be designed that will display altered recognition based on alternative hydrogen-bonding patterns [47,48], hydrophobic interactions [49], or chelation of a metal ion [50]. Alternatively, specificity and orthogonality can be achieved using size and/or steric effects [51,52].

One attractive strategy for forming an alternative base pair is to use two nucleosides that can form a specific base pair without pairing with any of the natural nucleobases. One of the first such base pair to be described was that between isocytidine (iC) and isoguanine (iG), in which the hydrogen-bonding groups of cytosine and guanosine are inverted. This allows for the formation of a specific base pair with a different donor and acceptor pattern from that of the natural base pairs (Fig. 13-5). The iC–iG pair has been shown to be replicated by both DNA and RNA polymerases [53], including in PCR [54]. One problem associated with this new base pair is that the iG exists in two different tautomeric forms, the minor of which specifically pairs with thymidine, leading to a loss of fidelity in replication reactions. Different strategies have been developed to avoid this. For example, Benner has used 2-thiothymidine instead of thymidine to prevent mispairing with the minor iG tautomer, as the 2-thio-group does not hydrogen bond effectively [55], while Seela has shown that the 7-deaza analogue of iG does form tautomers to a much reduced extent ($>10^3$-fold less) [56].

Benner and coworkers have devised a complete set of alternative hydrogen-bonded base pairs, each of which are held together by three hydrogen bonds and retain the size and geometry of the canonical base pairs [57]. One of the more advanced pairs is
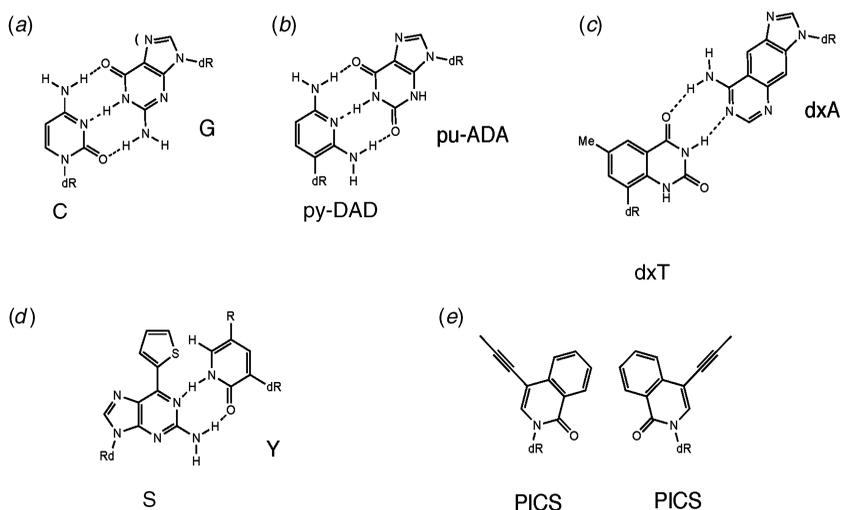
Figure 13-5   Many novel base pairing systems have been examined as an alternative genetic coding system. Specific alternative hydrogen-bonding systems can be used in conjunction with the native base pairs (a) such as py-DAD/pu-ADA (b) and an expanded version of native base pairs (c). Other systems use hydrogen-bonding and steric effects (d) or non-hydrogen-bonding self-pairs (e).

that between 2,4-diaminopyrimidine (py-DAD (py: pyrimidine; D: H-bond donor (e.g., NH_2 group); A: H-bond acceptor (e.g., C=O group)) and xanthine (pu-ADA) (Fig. 13-5) [58,59]. The py-DAD/pu-ADA base pair has been shown to be a substrate for a mutant HIV-1 reverse transcriptase, which replicates the new base pair with good fidelity in the presence of native DNA nucleotides [60].

Kool has examined the effect of size as an alternative genetic base pairing system. In the most striking example, the natural nucleosides have been redesigned with an expanded size by incorporation of a phenyl group between the sugar and the hydrogen-bonding ring [51,61–63]. This size expanded system, termed xDNA (x for expanded) (Fig. 13-5) or yDNA (y for wide), retain the features of regular DNA, such as Watson–Crick base pairing and right-handed helicity, but possess an expanded diameter when in a double helix [64]. When xDNA nucleosides are incorporated into regular duplex DNA there is distortion of the backbone due to the increased size of the base pair (2.4 Å) [65], but a duplex comprised solely of xDNA shows enhanced stability compared to DNA due to enhanced stacking interactions [66,67]. xDNA and yDNA represent two novel genetic systems, possessing many of the features found in regular DNA, but their expanded sizes should make them distinct from DNA.

A further method for developing a new base pair is an analogue that preferentially forms a self-pair. This class of analogue tends to be planar, aromatic, and non-hydrogen bonding, yet they can still be recognized by cellular enzymes, such as polymerases. There are a number of such analogues reported. Romesberg, Schultz, and coworkers have synthesized a number of analogues such as 7-azaindole (7-AI), propynylisocarbostyrile (PICS) (Fig. 13-5) as well as some fluoroaromatic analogues [68] and evaluated them as potential self-pairing nucleosides [69,70]. Various of the analogues prepared are also recognized with reasonable selectivity by DNA polymerases [49,68,71].

One of the most highly developed systems of novel, specific base pairing has been designed by Hirao and Yokoyama, who used steric effects to design various novel base pairs, two of which were found to be compatible with various cellular events. The systems they devised replaced the pyrimidine base with a pyridone and the purine with a C6-modified diaminopurine (Fig. 13-5), and retained hydrogen-bonding capability [72,73]. The pyridone will not form stable base pairs with the natural purines while if the purine base pairs with thymine it will be destabilized by a steric clash between the pyrimidine O4 and the purine C6 modification (Fig. 13-5). These analogues have been shown to form specific base pairs and to be recognized by DNA polymerises [74,75], RNA polymerises [52,76–79] and in translation, allowing the site-specific introduction of an unnatural amino acid *in vitro* [80].

13.5 NUCLEIC ACID SELF-REPLICATION

13.5.1 De Novo Synthesis of Nucleic Acid Polymers

In 1954, physicist George Gamow founded the RNA-tie club with a group of 20 scientists (one for each of the naturally occurring amino acids) who were interested in

the function of RNA. One of the original members of this group, Leslie Orgel, has carried out significant studies in the field of self-replicating systems, and aspects of his work are discussed in this chapter. Over the following years the relationship between nucleic acids and proteins became better understood, but scientists such as Orgel started to ask questions about the origins of life and in particular about the prebiotic synthesis of nucleic acids.

While self-replication requires a template molecule to start from, these had to be first generated *de novo* from precursor molecules. Ferris *et al.* [81–84] have investigated the *de novo* synthesis of RNA polynucleotides on common clay minerals such as montmorillonite as a model for prebiotic synthesis. It was shown that mononucleotides activated as phosphoro-imidazolides would react with other nucleotide polyphosphates, for example, triphosphates, to form predominantly 3′, 5′-linked oligonucleotides in the presence of montmorillonite clay with a rate enhancement of 1000-fold compared to the absence of montmorillonite. It has also been shown that oligonucleotide 5′-polyphosphates (including triphosphates) can be formed from polynucleotide monophosphates and sodium trimetaphosphate [85]. Thus, a feasible mechanism for the synthesis of the original RNA polynucleotides has been described. Much of the further work carried out to investigate template-directed self-replication nevertheless makes use of 5′-imidazole-activated nucleotides as they are more reactive derivatives for the synthesis of oligo- and polynucleotides.

### 13.5.2   Template-Directed Synthesis of Nucleic Acids

Orgel et al. [86] have been involved in a majority of the work in the field of nonenzymatic template-directed synthesis of oligonucleotides. Early work from this group demonstrated that random copolymer RNA templates could be used to replicate RNA in solution without the need for an enzyme or catalyst over several days and at high Mg^{2+} concentrations (Fig. 13-6). These reactions are template-dependent, and under the reaction conditions AT base pairs are formed much less efficiently than GC pairs. Under these conditions, oligonucleotides in the range of 20–30 nucleotides can be produced over a period of 1 week. Analysis of the products demonstrated that there is a mixture of 2′,5′- and 3′,5′-linkages, with the 2′,5′-linkages predominating. This is probably due to the fact that the 2′-hydroxyl group is six to nine times more reactive than the 3′-hydroxyl group [87,88]. Synthesis of DNA using a DNA template and activated deoxynucleotides is much less efficient, and it has been reported that some sequences cannot be copied [89].

Szostak has studied the nonenzymatic template-directed ligation of oligoribonucleotides and shown that there is a dependence for binding to metal ions before ligation can occur [90]. A series of metal ions were assayed and Mn^{2+} and Mg^{2+} ions are most efficient for catalysis while Pb^{2+} and Zn^{2+} ions do not. They also demonstrated that the nonenzymatic ligation proceeds with a preference for 3′–5′ phosphodiester linkages in preference to 2′–5′, though it is dependent on the ligation chemistry (imidazolide or triphosphate) [91]. The preference for 3′–5′ linkages is in contrast to that reported by Orgel, who reported a preference for 2′–5′ linkages for template-directed replication, suggesting that the type of linkage obtained may be sensitive to reaction conditions.
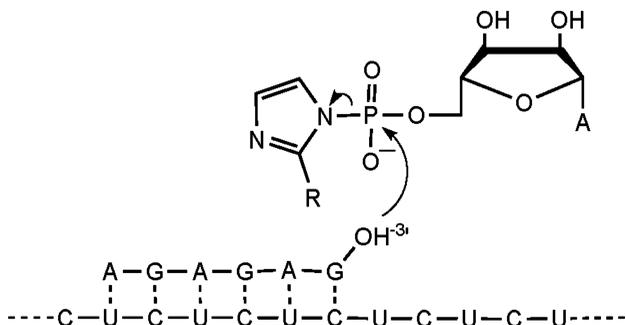
Figure 13-6 Nonenzymatic template-directed synthesis of RNA involving activated nucleoside monophosphates.

The above examples describe nonenzymatic synthesis of the complementary (−) strand of oligonucleotides templated by the (+) strand. For autocatalysis to occur, it is required that the two strands ((+) and (−)) separate and the (−) strand templates the re-synthesis of the (+) strand. The first example of such a truly self-replicating system was described by von Kiedrowski [92]. In this work, two trideoxynucleotides leading to a hexameric palindromic template were used, each trideoxynucleotide was 3′-protected to prevent elongation beyond a hexamer sequence. Initial coupling was carried out using a water-soluble carbodiimide (EDC) under conditions that led to the hexamer template rather than pyrophosphate dimer. Once formed, the product serves as template for further self-replication, and being palindromic, both (−) and (+) strands are formed in the same reaction (Fig. 13-7). As noted above, a possible product from the EDC-mediated coupling reaction is an oligodeoxynucleotide with an internal pyrophosphate linkage. Such
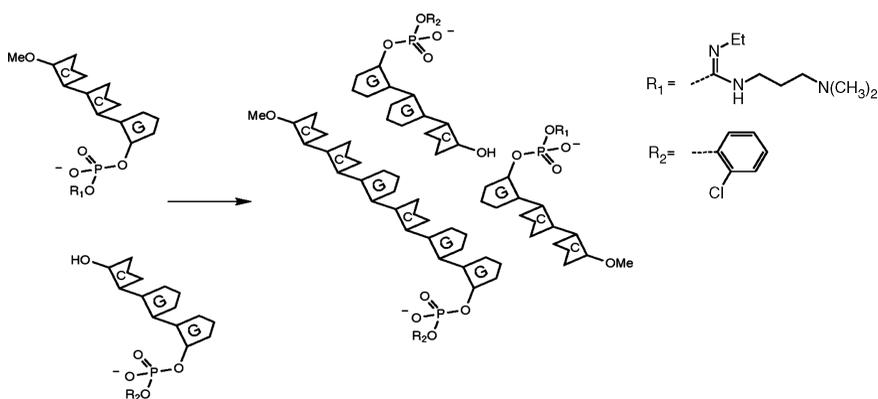
Figure 13-7 The first example of an autocatalytic system was described by Günter von Kiedrowski involved carbodiimide coupling of two trideoxynucleotides leading to a hexameric palindromic template. Once formed, the product serves as template for further self-replication, and being palindromic both (−) and (+) strands are formed in the same reaction.

modified oligodeoxynucleotides have been examined as substrates for self-replication and shown to still carry out sequence-dependent autocatalysis despite the phosphate modification [93].

Similar autocatalysis was observed for the synthesis of palindromic oligodeoxynucleotides using EDC-mediated formation of a $3'-5'$-phosphoramidate linkage between the two trimer building blocks [94,95]. The kinetics of self-replication has also been studied using fluorescently labeled tetramers by measurement of FRET [96]. Another autocatalytic system has been described by Nicolaou [97] for the synthesis of longer (24 mer) duplex palindromic polypurine/polypyrimidine DNA.

The early work by von Kiedrowski involved the replication of self-complementary sequences while natural replication involves the replication of complementary sequences. Using the previous system of chemical ligation of trimers, a minimal system for the synthesis of complementary replication has been described based on cross-catalytic template-directed synthesis using phosphoramidate linkages [98,99]. Two self-complementary and two complementary templates compete for four common trimeric precursors, and evidence was obtained to show that cross-catalytic self-replication of complementary sequences occurs with an equal efficiency to autocatalysis of the self-complementary sequence.

A common problem with replication by these systems is product inhibition, whereby the product dimer does not efficiently dissociate. As a result of this, there is parabolic rather than exponential amplification, and exponential amplification is a dynamic prerequisite for Darwinian selection. Using a system-denoted SPREAD (surface-promoted replication and exponential amplification of DNA analogues) exponential amplification was achieved by using a step to liberate the daughter strands from the template and cycling the amplification process [100]. More recent work by von Kiedrowski describes the self-assembly of three-dimensional DNA nanoscaffolds as a step toward artificially self-replicating systems on a nanometer scale [101,102] (see Section 13.10).

13.6   RNA SELF-REPLICATION: THE RNA WORLD

The emergence of a polymer (such as RNA) capable of self-replication, mutation, and hence evolution toward more efficient self-replication, represents an attractive and plausible concept for the origin of life. Several strands of evidence support the concept of such an "RNA" world, whereby RNA would serve as both genetic material as well as catalyst, preceding modern biology. These include aspects of modern metabolism (such as nucleotide cofactors, genetic control (self-splicing introns [103], riboswitches [104]), and most strikingly protein synthesis [105–107] that involve RNA and may thus represent relics from the "RNA world." The versatility of RNA to serve as both a receptor and catalyst has been further underlined by the wide range of activities documented in naturally occurring RNA receptors and ribozymes as well as in the ready evolution of novel activities using *in vitro* evolution methods like SELEX [108].

Despite its catalytic and conformational versatility RNA seems a somewhat perverse choice as the primordial genetic material, because it appears to be both

difficult to synthesize and extremely unstable under presumed prebiotic conditions. This has led some to propose a "pre-RNA world," which utilized other polymers such as PNA (in which the ribofuranose-phosphate backbone is replaced by an achiral peptide backbone) or TNA (in which the ribose is replaced by a tetrofuranose), which were superseded by RNA at a later stage. Both PNA and TNA can form stable helices with RNA (and DNA) and interpolymer genetic information transfer should thus be possible. Indeed, it has been shown that information can be transferred non-enzymatically between PNA and DNA [109], DNA and PNA [110] and PNA to RNA [111]. Using "Therminator" polymerase, it has been shown that TNA strands up to 80-nucleotides long can be synthesized from a DNA template with good fidelity [46,112]. Orgel and coworkers have also examined other nucleic acid systems and found that nucleosides containing 1,5-anhydrohexital (HNA) can be used in place of ribose to carry out templated nonenzymatic replication [32,33]. The information transfer of HNA to RNA requires the formation of an A-form product and therefore information transfer to DNA is inefficient [34]. The templating of information with hexose sugars is even more efficient when the 1,5-anhydrohexital sugar is replaced by altritol (ANA, HNA that has an additional hydroxyl group) [35]. TNA appears the most attractive pre-RNA polymer as long PNA strands suffer from solubility problems due to the uncharged nature of the polypeptide backbone. Nevertheless, it remains to be seen if TNA displays similar versatility as a receptor and catalyst as RNA.

The case for RNA has recently been further strengthened by the discovery of long RNA polymers in eutectic ice phases [113,114], the stabilization of ribose by borate evaporates [115], the selective uptake of ribose (compared to other aldopentoses) by phospholipid and fatty acid vesicles [116] and the sequestration of enatiomerically pure D-ribose from a prebiotic mixture [117]. The latter is especially significant as the presence of small amounts of L-enantiomers of nucleosides effectively poison chain elongation in templated nonenzymatic RNA synthesis using the natural D-enantiomer [118].

13.6.1   The Search for an RNA Replicase

A cornerstone of the "RNA world" hypothesis is that there exists somewhere in sequence space a ribozyme replicase capable of self-replication. Indeed a number of naturally occurring as well as selected ribozymes display some ability for self-replication, most notably through assembly and enzymatic ligation of oligonucleotides [119]. Indeed, recently a self-replicating ligase ribozyme was described that directed its own assembly from constituent parts, and in an initial phase displayed true exponential growth [120]. This report demonstrates the potential of the approach toward a self-replicating system. However, because of the need to provide presynthesized oligonucleotide substrates and the need to retain substantial base-pairing with the ligase, the ability of such system to evolve is restricted. More complex, multicomponent self-ligation networks [95,121] may allow the inclusion of sufficient molecular diversity for some evolution to proceed.

A more general self-replication capability may be achieved by the use of shorter oligonucleotide substrates, ideally activated nucleotide precursors such as the

nucleotide triphosphates (NTPs) utilized by modern polymerases. Intriguingly, both natural as well as evolved ribozymes have been shown to display weak primer extension ability using NTPs as substrates [122]. In ground-breaking work, Bartel and colleagues have evolved the primer extension capability of one such ribozyme, the R18 replicase, to the point where template-directed replication of up 14 nucleotides is possible [123]. As the R18 ribozyme is about 180 nucleotides long, an increase of processivity of a little more than one order of magnitude, should bring true self-replication within reach.

However, self-replication must proceed with a degree of fidelity, as defined by the "error threshold," above which genetic information encoded in the replicase would be irretrievably corrupted. An extensive theoretical framework on error threshold has been developed but it is unclear to what extent these can be applied to the practical case of an RNA replicase ribozyme. For example, the R18 replicase does appear to display fairly substantial template-dependent differences in processivity and fidelity [124,125], making it difficult to assign a meaningful overall mutation rate. While a recent study indicates that ribozymes in general may have an "relaxed error threshold" and thus be able to tolerate higher mutation rates than previously assumed [126], a number of *in vitro* evolution studies suggest that the class I ligase core (on which the R18 replicase is based) is rather resistant to mutation [108,127]. This may indicate that it represents a structure close to an evolutionary optimum, suggesting that at least half of the R18 replicase might be rather sensitive to poor fidelity in self-replication.

13.7 COMPARTMENTALIZATION: TOWARD THE DESIGN OF A SIMPLE CELL

For Darwinian evolution to proceed a putative replicase needs a form of "genetic packaging" such as confinement inside a compartment or at the very least spatial colocalization, for example, on the surface of mineral grains. Without such diffusion-limitation a replicase would fruitlessly replicate unrelated (and most likely inactive) sequences and eventually disappear from the sequence pool. Theoretical studies have also shown that limited diffusion aids replicase evolution by limiting the spread of replication parasites [128]. Physical proximity of a replicase to its "offspring" thus ensures both the growth and spread of the self-replicating entity as well as preventing takeover by fast-replicating "parasites."

13.7.1 Vesicles

Compartmentalization can potentially occur in many forms. An attractive format is vesicles comprising a bilayer of amphiphilic lipids. Such vesicles form spontaneously upon mixing of the constituent lipids with an aqueous solution. Some clay minerals, which promote the synthesis of polynucleotides from activated precursors, have been found to also catalyze the formation of vesicles. Szostak, Luisi, and colleagues [129,130] in particular have shown that vesicles comprising fatty acids as their main
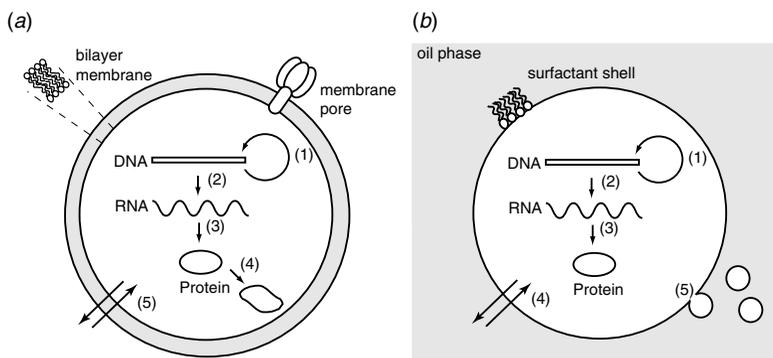
Figure 13-8 Compartmentalization in (a) vesicles and (b) water-in-oil emulsions. Both systems support DNA replication (1), transcription (2) and translation (3), which in the case of vesicles can be made to last for days using a membrane pore (hemolysin) to feed the reaction from surrounding medium. Bilayer membranes as well as W/O are partially permeable (5) to small molecules (e.g., H^+, Mg^{2+}, NTP). Specific molecules can be targeted to emulsion compartments using nanoemulsions (6).

constituents can display both autocatalytic growth as well as multiple cycles of growth and division. The fluid bilayer membrane also allows ready exchange of small molecules across membranes and this can drive competition for limited resources, as vesicles containing a larger amount of an osmotically-active compound (e.g., RNA) grow in size at the expense of others. Permeability is related *inter alia* to the length of the aliphatic chain and is thus in principle, controllable, with longer aliphatic chains leading to progressively less fluid, less permeable membranes [131]. Ribozyme activity [132], DNA as well as RNA replication [131,133], long-lasting transcription and translation [134] and even a two-stage genetic cascade have been demonstrated in vesicles (Fig. 13-8). Vesicles are therefore potentially attractive formats for a synthetic protocell.

From a synthetic biology perspective of engineering a suitable protocell, one of the problems that remain to be solved is that of cell reproduction. While vesicles made from hydrolyzable surfactants can be made to reproduce (whereby hydrolysis generates building blocks to form new vesicles) and vesicle fission and budding can be induced by application of physical and chemical forces [130], such replication is largely independent of vesicle content.

13.7.2 Emulsions

Although unlikely to have been relevant in prebiotic evolution, from a synthetic biology perspective an alternative format for a protocell may be based on emulsions. Emulsions are heterogeneous and, in general, metastable mixtures of two immiscible liquid phases with one of the phases dispersed in the other as droplets of microscopic size. Emulsions may be produced from any suitable combination of immiscible liquids by stirring, homogenization, or through microfluidic methods [135]. For the construction of a protocell so-called "water-in-oil" (W/O) emulsions are preferable,

in which the disperse, internal phase forms a suspension of cell-like, aqueous "droplets" within an inert hydrophobic liquid matrix. Nevertheless, a O/W design for a protocell has been proposed [136], in which the genetic material (made from the neutral DNA analogue, PNA) is contained within lipid droplets suspended in an aqueous phase.

As with vesicles, cell-like aqueous compartments formed in W/O emulsions support various enzymatic reactions including coupled *in vitro* transcription and translation [137,138], as well as DNA replication and PCR [139] (Fig. 13-8). The size of aqueous compartments can readily be controlled by varying emulsion composition and mechanical energy input (between 70 nM [140] and 150 μM [141]). Just like vesicles, emulsions are also remarkably permeable to small molecules such as solvated ions and (at high temperatures) even nucleoside triphosphates [139]. Reagents can also be delivered to emulsion compartments in a controlled way using nanoemulsions [142]. However, even after prolonged exposure to high temperatures there appears to be little, if any, exchange of polypeptides or nucleic acids (>30 bp) between compartments [139].

13.7.3 Compartmentalized Evolution

In vitro compartmentalization (IVC) in emulsions allows a stable linkage of genotype and phenotype [137] and this has been exploited for *in vitro* evolution. IVC has allowed the evolution of DNA methylases with altered substrate specificity [143], a super-fast phosphotriesterase [144] as well as novel ribozymes [127,140]. Emulsions can also be used to segregate self-replication reactions. Compartmentalized self-replication (CSR) exploits this for the directed evolution of polymerases [139]. In CSR, polymerases catalyze the replication of their own encoding gene. As a result, adaptive gains by the polymerases translate directly into more genetic "offspring" (i.e., more efficient self-replication). Due to this positive feedback loop, the genes encoding polymerases that are well adapted to the selection conditions (and therefore capable of efficient self-replication) will increase in copy number while genes encoding poorly adapted polymerases will disappear from the gene pool.

CSR has allowed the directed evolution of polymerases with increased thermostability, inhibitor tolerance or a generically expanded substrate spectrum [139,145]. CSR may also be regarded as a simple test bed for self-replication. For example, a classic outcome of *in vitro* replication experiments is an adaptation of the template sequence toward more rapid replication [146]. This typically takes the form of truncation as well as mutation and (in solution) invariably gives rise to (often heavily truncated) "replication parasites," which have lost much of the genetic information encoding the original phenotype but are optimized for replication speed. Such parasites arise frequently, for example in PCR amplifications (primer dimers) or *in vitro* evolution experiments [147]. While template evolution appears to occur in CSR through silent mutations reducing GC content facilitating strand separation and destabilizing secondary structures [139]. However, template truncation was not observed (despite the considerable size of the *Taq* gene (2.5 kb)). Presumably, this is due to both the strong phenotypic selection in CSR as well as the effect of

compartmentalization, which limits the spread of parasites to the compartment, where they occur (e.g., see Ref. 128). For a self-replicating RNA replicase, template evolution (e.g., through mutations that destabilize secondary structures) may be a mixed blessing and requires a trade-off between structural stability of the replicase structure itself and its replicability.

A wide range of other forms of compartmentalization (or diffusion limitation) are imaginable that may have played a role in prebiotic replication. These include the surface of fine particulate matter or in porous minerals (e.g., clays such as montmorillonite; see Section 13.5.2), eutectic ice phases [113], or aerosol droplets in the atmosphere [148].

13.8   *IN VIVO* REPLICATION

All of the examples discussed so far involve *ex vivo* designs of self-replicating entities. However, for various applications it may be desirable to consider invading present-day biological systems with self-replicating species. These already exist of course in biology in the form of plasmids, viruses, and so on; however, these interact extensively with the host organism. One concern when building synthetic devices is both their potentially toxic effect on host biology, as well as the possible interference of host cellular functions with the operation of the device.

In order to escape such interference one may ask, if it would be possible to build synthetic self-replicating circuits that are capable of operating independently from the rest of the cell. Such orthogonal episomes would carry their own polymerases for specific replication and transcription "on board" (in analogy to many viruses) but would still be subject to recombination, mutation, and degradation by the host genetic machinery. It might therefore be advantageous to consider synthetic episomes that are orthogonal in composition as well as replication. Such episomes would comprise unnatural nucleic acids (XNA) and thus would be isolated from the host genetic machinery by chemical, steric, or semantic differences.

The requirements for such a system parallel in many ways those for the polymers of a "pre-RNA" world, requiring the ability of cross-talk and mutual interconversion (transliteration) between "XNA" and DNA/RNA (Fig. 13-9). A design for an artificial genetic system might therefore be preferably based on an alternative backbone structure. This has the potential benefits of providing orthogonality, that is, synthetic and functional isolation within the cell (as altered backbone chemistry precludes utilization by the cellular genetic machinery) without altering the coding potential of the nucleic acids. In other words, a genetic entity constructed this way may be built from precursors that are sufficiently different from the natural nucleosides that they cannot be utilized by the pre-existing genetic machinery of the cell (replication/transcription/translation) and therefore do not give rise to toxicity while at the very same time are able to communicate with it. Interestingly, just such a scenario has recently been put forward for the origin of DNA. It proposes that such an orthogonal nucleic acid with a modified backbone (DNA) was "invented" by viruses infecting riboorganisms of the RNA world in order to avoid cellular defenses (e.g., RNAses) [149].
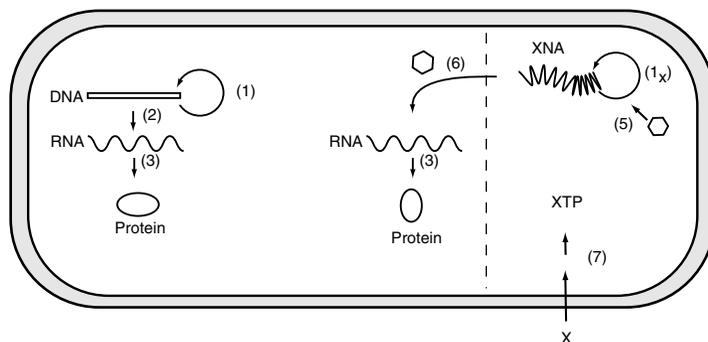
Figure 13-9 Invasion of the natural genetic system by an orthogonal episome made from unnatural nucleic acids (XNA). Because of its orthogonal design it is impervious to cellular degradation, recombination, or mutation mechanisms. A fully functional orthogonal episome must be capable of replication (1) through a dedicated polymerase (5) and transcription into RNA (or DNA) through a second dedicated polymerase (6). Thus, a specific set of proteins (or RNAs) could be encoded on the orthogonal episome to alter the cellular phenotype. Maintenance of the orthogonal episome will also require the engineering of efficient uptake (and possibly phosphorylation) pathways (7) for XNA precursors from the medium.

An alternative strategy, for which there are no known precedents in nature, would be steric orthogonality, which may be based around the expanded DNA (xDNA, yDNA) described by the Kool laboratory at Stanford (see Section 13.4.3). Because the base-pairing is not altered, x, y DNA are still capable of base-pairing with DNA and RNA but forming double helices with an expanded diameter and higher stability due to the increased stacking of the expanded bases [62,66]. Finally, orthogonality might simply be semantic in that genetic information encoded in such a way that it is "meaningless" to the cellular host, unless specific transliterases are provided.

A potentially important advantage of such systems could be safety. As these transgenes would be based around nucleic acid chemistry not present in nature its function and transmission will be entirely dependent and controlled by the supply of orthogonal precursors. Among other things, this will provide a novel and complete control of genetic safety issues as the propagation and inheritance of "foreign" genetic material in a transgenic organism can be simply turned off and the transgene excised by removing the supply of precursors.

13.9 MOLECULAR DEVICES AND AUTOMATA

Unlike any other molecule DNA affords ready control over intermolecular associations. DNA molecules associate according to well-understood rules of complementarity providing a diverse and programmable system, with known structures and a high degree of control over molecular interactions. DNA is thus increasingly recognized as a material of choice for self-assembling "bottom-up" nanostructures and nanodevices [150].

For example, using stable branched structure in conjunction with "sticky end" cohesion has allowed the generation of, for example, DNA cubes and octahedrons as well as two-dimensional DNA arrays some of which can be used for computation [150]. However, the conformational flexibility of DNA also lends itself to the construction of multistate devices with a flexible response to the input conditions. For example, Seeman and colleagues have constructed a nanomechanical device, which exploits the structural transition between the canonical right-handed B-DNA and left-handed Z-DNA in $[CG]_n$-rich sequences in response to high salt concentrations [151]. Thus, the device translates on input signal (ionic strength) into an observable fluorescence signal through fluorescence resonance energy transfer (FRET) differences of the two states. Other devices include DNA tweezers [152], rotary motors [153], walkers [154], and even a translation machine [155].

A different class of DNA devices has been built around cycles of ligation and cleavage with a type II restriction enzyme. Shapiro, Benenson and colleagues showed that this allowed the construction of DNA-based finite-state automata capable of autonomous computation at the molecular level [156]. One such automaton was shown to be able to analyze *in vitro* the levels of several RNAs involved in prostate cancer and compute an appropriate response, that is, release of an antisense molecule [157] offering the prospect of programmable, logical control of biological processes at the molecular scale.

These devices are assembled and sometimes powered by oligonucleotide fragments and thus their topologies are not amenable to replication. However, different topological designs are possible. In a striking example, Joyce and colleagues recently described a single-stranded 1.7-kb DNA sequence that folds into a octahedron in the presence of short DNA oligonucleotides [158]. In conjunction with engineered polymerases [159] this offers the future prospect of replicable nanostructures endowed with expanded chemical capabilities and amenable to iterative cycles of replication, mutation and selection, bringing directed evolution to nanotechnology and material science.

13.10   CONCLUSION

For synthetic biology, self-replication should be a long-term goal for the engineering of material devices. While the construction and implementation of circuits and devices *in vivo* (i.e., in extant biological systems) provides for self-replication as part of the reproduction of the organism, self-replication *ex vivo* or as part of a whole synthetic quasibiotic entity (e.g., a synthetic cell) may have a number of long-term advantages. For one, it may remove some of the unpredictability and instability that can be a consequence of integrating new functionalities into the cellular network [160]. Furthermore, the design and fabrication of synthetic conduits for self-replication promises significant insights and advances in understanding of the transition from prebiotic to biotic matter and the early evolution of life. Finally, self-replication as applied to the emergent technologies such as DNA nanotechnology and molecular automata [161] promises decisive reductions in manufacturing costs as well as bringing the potential of Darwinian evolution to nanosensors and molecular devices.

ACKNOWLEDGMENTS

REFERENCES

1. Paul N, Joyce GF. Minimal self-replicating systems. *Curr Opin Chem Biol* 2004;8: 634–639.

2. Szathmary E. The evolution of replicators. *Philos Trans R Soc Lond B Biol Sci* 2000;355: 1669–1676.

3. Zykov V, Mytilinaios E, Adams B, Lipson H. Robotics: self-reproducing machines. *Nature* 2005;435:163–164.

4. Tjivikua T, Ballester P, Rebek J. A self-replicating system. *J Am Chem Soc* 1990;112: 1249–1250.

5. Nowick JS, Feng Q, Tjivikua T, Ballester P, Rebek J. Kinetic studies and modelling of a self-replicating system. *J Am Chem Soc* 1991;113:8831–8839.

6. Wintner EA, Tsao B, Rebek J. Evidence against an alternative mechanism for a self-replicating system. *J Org Chem* 1995;60:7997–8001.

7. Conn MM, Wintner EA, Rebek J. Template effects in new self-replicating molecules. *Angew Chem Int Ed* 1994;33:1577–1579.

8. Pieters RJ, Huc I, Rebek J. Reciprocal template effects in a replication cycle. *Angew Chem Int Ed* 1994;33:1579–1581.

9. Kindermann M, Stahl I, Reimold M, Pankau WM, von Kiedrowski G. Systems chemistry: kinetic and computational analysis of a nearly exponential organic replicator. *Angew Chem Int Ed* 2005;55:6750–6755.

10. Pearson RJ, Kassianidis E, Slawin AMZ, Philp D. Self-replication vs. reactive binary complexes—manipulating recognition-mediated cycloadditions by simple structural modifications. *Org Biomol Chem* 2004;2:3434–3441.

11. Kassianidis E, Pearson RJ, Philp D. Specific autocatalysis in diastereoisomeric replicators. *Org Lett* 2005;7:3833–3836.

12. Lee DH, Granja JR, Martinez JA, Severin K, Ghadiri MR. A self-replicating peptide. *Nature* 1996;382:525–528.

13. Lee DH, Severin K, Yokobayashi Y, Ghadiri MR. Emergence of symbiosis in peptide self-replication through a hypercyclic network. *Nature* 1997;390:591–594.

14. Severin K, Lee DH, Martinez JA, Vieth M, Ghadiri MR. Dynamic error correction in autocatalytic peptide networks. *Angew Chem Int Ed* 1998;37:126–128.

15. Saghatelian A, Yokobayashi Y, Soltani K, Ghadiri MR. A chiroselective peptide replicator. *Nature* 2001;409:797–801.

16. Islas JR, Pimienta V, Micheau J-C, Buhse T. Kinetic analysis of artificial peptide self-replication. Part I: The homochiral case. *Biophys Chem* 2003;103:191–200.

17. Islas JR, Pimienta V, Micheau J-C, Buhse T. Kinetic analysis of artificial peptide self-replication. Part II: The heterochiral case. *Biophys Chem* 2003;103:201–211.

18. Islas JR, Micheau J-C, Buhse T. Kinetic analysis of self-replicating peptides: possibility of chiral amplification in open systems. Orig Life Evol Biosph 2004;34:497–512.

19. Yao S, Ghosh I, Zutshi R, Chmielewski J. A pH modulated, self-replicating peptide. *J Am Chem Soc* 1997;119:10559–10560.

20. Li X, Chmielewski J. Peptide self-replication enhanced by a proline kink. *J Am Chem Soc* 2003;125:11820–11821.

21. Matsumura S, Takahashi T, Ueno A, Mihara H. Complementary nucleobase interaction enhances peptide-peptide and self-replicating catalysis. *Chem Eur J* 2003;9: 4829–4837.

22. Levy M, Ellington AD. Peptide-templated nucleic acid ligation. *J Mol Evol* 2003;56: 607–615.

23. Ghadiri MR, Granja JR, Milligan RA, McRee DE, Khazanovich N. Self-assembling organic nanotubes based on a cyclic peptide architecture. *Nature* 1993;366:324–327.

24. Kobayashi K, Granja JR, Ghadiri MR. Beta-sheet peptide architecture: measuring the relative stability of parallel vs antiparallel beta-sheets. *Angew Chem Int Ed* 1995;34:95–98.

25. Peretz D, Williamson RA, Legname G, Matsunaga Y, Vergara J, Burton DR, DeArmond SJ, Prusiner SB, Scott MR. A change in the conformation of prions accompanies the emergence of a new prion strain. *Neuron* 2002;34:921–932.

26. Shorter J, Lindquist S. Prions as adaptive conduits of memory and inheritance. *Nat Rev Genet* 2005;6:435–450.

27. Watson JD, Crick FHC. Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature* 1953;171:737–738.

28. Benner SA. Chemistry. Redesigning genetics. *Science* 2004;306:625–626.

29. Nielsen PE. The many faces of PNA. *Lett Peptide Sci* 2003;10:135–147.

30. Stirchak EP, Summerton JE, Weller DD. Uncharged stereoregular nucleic acid analogs: 2. Morpholino nucleoside oligomers with carbamate internucleoside linkages. *Nucleic Acids Res* 1989;17:6129–6141.

31. Summerton J. Morpholino antisense oligomers: the case for an RNase H-independent structural type. *Biochem Biophys Acta* 1999;1489:141–158.

32. Kozlov IA, Politis PK, Pitsch S, Herdewijn P, Orgel LE. A highly enantio-selective hexitol nucleic acid template for nonenzymatic oligoguanylate synthesis. *J Am Chem Soc* 1999;121:1108–1109.

33. Kozlov IA, De Bouvere B, Van Aerschot A, Herdewijn P, Orgel LE. Efficient transfer of information from hexitol nucleic acids to RNA during nonenzymatic oligomerization. *J Am Chem Soc* 1999;121:5856–5859.

34. Kozlov IA, Politis PK, Van Aerschot A, Busson R, Herdewijn P, Orgel LE. Nonenzymatic synthesis of RNA and DNA oligomers on hexitol nucleic acid templates: the importance of the A structure. *J Am Chem Soc* 1999;121:2653–2656.

35. Kozlov IA, Zielinski M, Allart B, Kerremans L, Van Aerschot A, Busson R, Herdewijn P, Orgel LE. Nonenzymatic template-directed reactions on altritol oligomers, preorganised analogues of oligonucleotides. *Chem Eur J* 2000;6:151–155.

36. Eschenmoser A. Chemical etiology of nucleic acid structure. *Science* 1999;284:2118–2124.

37. Maurinsh Y, Rosemeyer H, Esnouf R, Medvedovici A, Wang J, Ceulemans G, Lescrinier E, Hendrix C, Busson R, Sandra P, Seela F, Van Aerschot A, Herdewijn P. Synthesis and pairing properties of oligonucleotides containing 3-hydroxy-4-hydroxymethyl-1-cyclo-hexanyl nucleosides. *Chem Eur J* 1999;5:2139–2150.

38. Lescrinier E, Froeyen M, Herdewijn P. Difference in conformational diversity between nucleic acids with a six-membered 'sugar' unit and natural 'furnaose' nucleic acids. *Nucleic Acids Res* 2003;31:2975–2989.

39. Nauwelaerts K, Lescrinier E, Scelp G, Herdewijn P. Cyclohexenyl nucleic acids: conformationally flexible oligonucleotides. *Nucleic Acids Res* 2005;33:2452–2463.

40. Wang J, Verbeure B, Luyten I, Lescrinier E, Froeyen M, Hendrix C, Rosemeyer H, Seela F, Van Aerschot A, Herdewijn P. Cyclohexen nucleic acids (CeNA): serum stable oligo-nucleotides that activate RNase H and increase duplex stability with complementary RNA. *J Am Chem Soc* 2000;122:8595–8602.

41. Eschenmoser A. The TNA-family of nucleic acid systems: Properties and prospects. *Orig Life Evol Biosph* 2004;34:277–306.

42. Kempeneers V, Froeyen M, Vastmans K, Herdewijn P. Influence of threose nucleoside units on the catalytic activity of a hammerhead ribozyme. *Chem Biodivers* 2004;1: 112–123.

43. Kempeneers V, Vastmans K, Rozenski J, Herdewijn P. Recognition of threosyl nucleos-tides by DNA and RNA polymerases. *Nucleic Acid Res* 2003;31:6221–6226.

44. Ichida JK, Zou K, Horhota A, Yu B, McLaughlin LW, Szostak JW. An *in vitro* selection system for TNA. *J Am Chem Soc* 2005;127:2902–2903.

45. Horhota A, Zou K, Ichida JK, Yu B, McLaughlin LW, Szostak JW, Chaput JC. Kinetic analysis of an efficient DNA-dependent TNA polymerase. *J Am Chem Soc* 2005;127: 7427–7434.

46. Ichida JK, Horhota A, Zou K, McLaughlin LW, Szostak JW. High fidelity TNA synthesis by Therminator polymerase. *Nucleic Acids Res* 2005;33:5219–5225.

47. Piccirilli JA, Krauch T, Moroney SE, Benner SA. Enzymatic incorporation of a new base pair into DNA and RNA extends the genetic alphabet. *Nature* 1990;343:33–37.

48. Benner SA. Understanding nucleic acids using synthetic chemistry. *Acc Chem Res* 2004;37:784–797.

49. Yu C, Henry AA, Romesberg FE, Schultz PG. Polymerase recognition of unnatural base pairs. *Angew Chem Int Ed* 2002;41:3841–3844.

50. Atwell S, Meggers E, Spraggon G, Schultz PG. Structure of a copper-mediated base pair in DNA. *J Am Chem Soc* 2001;123:12364–12367.

51. Liu H, Gao J, Lynch SR, Saito YD, Maynard L, Kool ET. A four-base paired genetic helix with expanded size. *Science* 2003;302:868–871.

52. Ohtsuki T, Kimoto M, Ishikawa M, Mitsui T, Hirao I, Yokoyama S. Unnatural base pairs for specific transcription. *Proc Natl Acad Sci USA* 2001;98:4922–4925.

53. Switzer C, Moroney SE, Benner SA. Enzymatic incorporation of a new base pair into DNA and RNA. *J Am Chem Soc* 1989;111:8322–8323.

54. Johnson SC, Marshall DJ, Harms G, Miller CM, Sherrill CB, Beaty EL, Lederer SA, Roesch EB, Madsen G, Hoffman GL, Laessig RH, Kopish GJ, Baker MW, Benner SA, Farrell PM, Prudent JR. Multiplexed genetic analysis using an expanded genetic alphabet. *Clin Chem* 2004;50:2019–2027.

55. Sismour AM, Benner SA. The use of thymidine analogs to improve the replication of an extra DNA base pair: a synthetic biological system. *Nucleic Acids Res* 2005;33:5640–5646.

56. Seela F, Wei C. The base-pairing properties of 7-deaza-2'-deoxyisoguanosine and 2'-deoxyisoguanosine in oligonucleotide duplexes with parallel and antiparallel chain orientation. *Helv Chim Acta* 1999;82:726–745.

57. Benner SA. Redesigning genetics. *Science* 2004;306:625–626.

58. Geyer CR, Battersby TR, Benner SA. Nucleobase pairing in expanded Watson–Crick-like genetic information systems. *Structure* 2003;11:1485–1498.

59. von Krosigk U, Benner SA. Expanding the genetic alphabet: pyrazine nucleosides that support a donor-donor-acceptor hydrogen-bonding pattern. *Helv Chim Acta* 2004;87:1299–1324.

60. Sismour AM, Lutz S, Park J-H, Lutz MJ, Boyer PL, Hughes SH, Benner SA. PCR amplification of DNA containing non-standard base pairs by variants of reverse transcriptase from human immunodeficiency virus-1. *Nucleic Acids Res* 2004;32:728–735.

61. Liu H, Gao J, Maynard L, Saito YD, Kool ET. Toward a new genetic system with expanded dimensions: size-expanded analogues of deoxyadenosine and thymidine. *J Am Chem Soc* 2004;126:1102–1109.

62. Lu H, He K, Kool ET. yDNA: a new geometry for size-expanded base pairs. *Angew Chem Int Ed* 2004;43:5834–5836.

63. Lee AH, Kool ET. A new four-base genetic helix, yDNA, composed of widened benzopyrimidine-purine pairs. *J Am Chem Soc* 2005;127:3332–3338.

64. Liu H, Lynch SR, Kool ET. Solution structure of xDNA: a paired genetic helix with increased diameter. *J Am Chem Soc* 2004;126:6900–6905.

65. Gao J, Liu H, Kool ET. Expanded-size bases in naturally sized DNA: evaluation of steric effects in Watson–Crick pairing. *J Am Chem Soc* 2004;126:11826–11831.

66. Gao J, Liu H, Kool ET. Assembly of the complete eight-base artificial genetic helix, xDNA, and its interaction with the natural genetic system. *Angew Chem Int Ed* 2005;44:3118–3122.

67. Liu H, Gao J, Kool ET. Helix-forming properties of size-expanded DNA, an alternative four-base genetic form. *J Am Chem Soc* 2005;127:1396–1402.

68. Henry AA, Olsen AG, Matsuda S, Yu C, Geierstanger BH, Romesberg FE. Efforts to expand the genetic alphabet: identification of a replicable unnatural DNA self-pair. *J Am Chem Soc* 2004;126:6923–6931.

69. Matsuda S, Henry AA, Schultz PG, Romesberg FE. The effects of minor-groove hydrogen-bond acceptors and donors on the stability and replication of four unnatural base pairs. *J Am Chem Soc* 2003;125:6134–6139.

70. Matsuda S, Romesberg FE. Optimization of interstrand hydrophobic packing interactions within unnatural DNA base pairs. *J Am Chem Soc* 2004;126:14419–14427.

71. Henry AA, Yu C, Romesberg FE. Determinants of unnatural nucleobase stability and polymerase recognition. *J Am Chem Soc* 2003;125:9638–9646.

72. Ishikawa M, Hirai I, Yokoyama S. Synthesis of 3-(2-deoxy-β-D-ribofuranosyl)pyridin-2-one and 2-amino-6-(*N*,*N*-dimethylamino)-9-(2-deoxy-β-D-ribofuranosyl)purine derivatives for an unnatural base pair. *Tetrahedron Lett* 2000;41:3931–3934.

73. Fujiwara T, Kimoto M, Sugiyama H, Hirao I, Yokoyama S. Synthesis of 6-(2-thienyl)purine nucleoside derivatives that form unnatural base pairs with pyridin-2-one nucleosides. *Bioorg Med Chem Lett* 2001;11:2221–2223.

74. Hirao I, Fujiwara T, Kimoto M, Yokoyama S. Unnatural base pairs between 2- and 6-substituted purines and 2-oxo(1*H*)pyridine for expansion of the genetic code. *Bioorg Med Chem Lett* 2004;14:4887–4890.

75. Kimoto M, Yokoyama S, Hirao I. A quantitative, non-radioactive single-nucleotide insertion assay for analysis of DNA replication fidelity by using an automated DNA sequencer. *Biotechnol Lett* 2004;26:999–1005.

76. Hirao I, Harada Y, Kimoto M, Mitsui T, Fujiwara T, Yokoyama S. A two-unnatural-base-pair system toward the expansion of the genetic code. *J Am Chem Soc* 2004;126: 13298–13305.

77. Endo M, Mitsui T, Okuni T, Kimoto M, Hirao I, Yokoyama S. Unnatural base pairs mediate the site-specific incorporation of an unnatural hydrophobic component into RNA transcripts. *Bioorg Med Chem Lett* 2004;14:2593–2596.

78. Kimoto M, Endo M, Mitsui T, Okuni T, Hirao I, Yokoyama S. Site-specific incorporation of a photo-crosslinking component into RNA by T7 transcription mediated by unnatural base pairs. *Chem Biol* 2004;11:47–55.

79. Kawai R, Kimoto M, Ikeda S, Mitsui T, Endo M, Yokoyama S, Hirao I. Site-specific fluorescent labeling of RNA molecules by specific transcription using unnatural base pairs. *J Am Chem Soc* 2005;127:17286–17295.

80. Hirao I, Ohtsuki T, Fujiwara T, Mitsui T, Yokogawa T, Okuni T, Nakayama H, Takio K, Yabuki T, Kigawa T, Kodama K, Yokogawa T, Nishikawa K, Yokoyama S. An unnatural base pair for incorporating amino acid analogs into proteins. *Nature Biotechnol* 2002;20: 177–182.

81. Ferris JP, Ertem G. Montmorillonite catalysis of RNA oligomer formation in aqueous solution. A model for the prebiotic formation of RNA. *J Am Chem Soc* 1993;115: 12270–12275.

82. Kawamura K, Ferris JP. Kinetic and mechanistic analysis of dinucleotide and oligonucleotide formation from the 5′-phosphorimidazolide of adenosine on Na^+-montmorillonite. *J Am Chem Soc* 1994;116:7564–7572.

83. Prabahar KJ, Ferris JP. 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* 1997;119:4330–4337.

84. Huang W, Ferris JP. Synthesis of 35-40 mers of RNA oligomers from unblocked monomers. A simple approach to the RNA world. *Chem Commun* 20031458–1459.

85. Gao K, Orgel LE. Phosphorylation and non-enzymatic template-directed ligation of oligonucleotides. *Orig Life Evol Biosph* 2000;30:45–51.

86. Joyce GF, Inoue T, Orgel LE. Non-enzymatic template-directed synthesis on RNA random copolymers. *J Mol Biol* 1984;176:279–306.

87. Usher DA. RNA double helix and the evolution of the 3′,5′ linkage. *Nat New Biol* 1972;235: 207–208.

88. Lohrmann R, Orgel LE. Preferential formation of (2′-5′)-linked internucleotide bonds in non-enzymatic reactions. *Tetrahedron* 1978;34:853–855.

89. Zielinski M, Kozlov IA, Orgel LE. A comparison of RNA with DNA in template-directed synthesis. *Helv Chim Acta* 2000;83:1678–1684.

90. Rohatgi R, Bartel DP, Szostak JW. Kinetic and mechanistic analysis of non-enzymatic, template-directed oligoribonucleotide ligation. *J Am Chem Soc* 1996;118: 3332–3339.

91. Rohatgi R, Bartel DP, Szostak JW. Nonenzymatic, template-directed ligation of oligoribonucleotides is highly selective for the formation of 3′-5′ phosphodiester bonds. *J Am Chem Soc* 1996;118:3340–3344.

92. von Kiedrowski G. A self-replicating hexadeoxynucleotide. *Angew Chem Int Ed* 1986;25: 932–935.

93. von Kiedrowski G, Wlotzka B, Helbing J. Sequence dependence of template-directed synthesis of hexadeoxynucleotide derivatives with 3,-5′ pyrophosphate linkage. *Angew Chem Int Ed* 1989;28:1235–1237.

94. von Kiedrowski G, Wlotzka B, Helbing J, Matzen M, Jordan S. Parabolic growth of a self-replicating hexadeoxynucleotide bearing a 3′-5′-phosphoramidate linkage. *Angew Chem Int Ed* 1991;30:423–426.

95. Achilles T, von Kiedrowski G. A self-replicating system from three starting materials. *Angew Chem Int Ed* 1993;32:1198–1201.

96. Schöneborn H, Bülle J, von Kiedrowski G. Kinetic monitoring of self-replicating systems through measurement of Fluorescence resonance energy transfer. *ChemBioChem* 2001;922–927.

97. Li T, Nicolaou KC. Chemical self-replication of palindromic duplex DNA. *Nature* 1994;369:218–221.

98. Sievers D, von Kiedrowski G. Self-replication of complementary nucleotide-based oligomers. *Nature* 1994;369:221–224.

99. Sievers D, von Kiedrowski G. Self-replication of hexadeoxynucleotide analogues: autocatalysis versus cross-catalysis. *Chem Eur J* 1998;4:629–641.

100. Luther A, Brandsch R, von Kiedrowski G. Surface-promoted replication and exponential amplification of DNA analogues. *Nature* 1998;396:245–248.

101. Eckardt LH, Naumann K, Pankau WM, Rein M, Schweitzer M, Windhab N, von Kiedrowski G. Chemical copying of connectivity. *Nature* 2002;420:286.

102. Scheffler M, Dorenbeck A, Jordan S, Wüstefeld M, von Kiedrowski G. Self-assembly of trisoligonucleotidyls: the case for nano-acetylene and nano-cyclobutadiene. *Angew Chem Int Ed* 1999;38:3312–3315.

103. Haugen P, Simon DM, Bhattacharya D. The natural history of group I introns. *Trends Genet* 2005;21:111–119.

104. Tucker BJ, Breaker RR. Riboswitches as versatile gene control elements. *Curr Opin Struct Biol* 2005;15:342–348.

105. Noller HF, Hoffarth V, Zimniak L. Unusual resistance of peptidyl transferase to protein extraction procedures. *Science* 1992;256:1416–1419.

106. Nissen P, Hansen J, Ban N, Moore PB, Steitz TA. The structural basis of ribosome activity in peptide bond synthesis. *Science* 2000;289:920–930.

107. Ban N, Nissen P, Hansen J, Moore PB, Steitz TA. The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. *Science* 2000;289:905–920.

108. Joyce GF. Directed evolution of nucleic acid enzymes. *Annu Rev Biochem* 2004;73: 791–836.

109. Böhler C, Nielsen PE, Orgel LE. Template switching between PNA and RNA oligonu-cleotides. *Nature* 1995;376:578–581.

110. Schmidt JG, Christensen L, Nielsen PE, Orgel LE. Information transfer from DNA to peptide nucleic acids by template-directed syntheses. *Nucleic Acids Res* 1997;25: 4792–4796.

111. Schmidt JG, Nielsen PE, Orgel LE. Information transfer from peptide nucleic acids to RNA by template-directed synthesis. *Nucleic Acids Res* 1997;25:4797–4802.

112. Chaput JC, Szostak JW. TNA synthesis by DNA polymerases. *J Am Chem Soc* 2003;125: 9274–9275.

113. Monnard PA, Kanavarioti A, Deamer DW. Eutectic phase polymerization of activated ribonucleotide mixtures yields quasi-equimolar incorporation of purine and pyrimidine nucleobases. *J Am Chem Soc* 2003;125:13734–13740.

114. Trinks H, Schroder W, Biebricher CK. Ice and the origin of life. *Orig Life Evol Biosph* 2005;35:429–445.

115. Ricardo A, Carrigan MA, Olcott AN, Benner SA. Borate minerals stabilize ribose. *Science* 2004;303:196.

116. Sacerdote MG, Szostak JW. Semipermeable lipid bilayers exhibit diastereoselectivity favoring ribose. *Proc Natl Acad Sci USA* 2005;102:6004–6008.

117. Springsteen G, Joyce GF. Selective derivatization and sequestration of ribose from a prebiotic mix. *J Am Chem Soc* 2004;126:9578–9583.

118. Kozlov IA, Pitsch S, Orgel LE. Oligomerization of activated D- and L-guanosine mononucleotides on templates containing D- and L-deoxycytidylate residues. *Proc Natl Acad Sci USA* 1998;95:13448–13452.

119. Doudna JA, Couture S, Szostak JW. A multisubunit ribozyme that is a catalyst of and template for complementary strand RNA synthesis. *Science* 1991;251: 1605–1608.

120. Paul N, Joyce GF. A self-replicating ligase ribozyme. *Proc Natl Acad Sci USA* 2002;99: 12733–12740.

121. Kim D-E, Joyce GF. Cross-catalytic replication of an RNA ligase ribozyme. *Chem Biol* 2004;11:1505–1512.

122. McGinness KE, Joyce GF. In search of an RNA replicase ribozyme. *Chem Biol* 2003;10:5–14.

123. Johnston WK, Unrau PJ, Lawrence MS, Glasner ME, Bartel DP. RNA-catalyzed RNA polymerization: accurate and general RNA-templated primer extension. *Science* 2001;292:1319–1325.

124. Lawrence MS, Bartel DP. Processivity of ribozyme-catalyzed RNA polymerization. *Biochemistry* 2003;42:8748–8755.

125. Lawrence MS, Bartel DP. New ligase-derived RNA polymerase ribozymes. *RNA* 2005;11:1173–1180.

126. Kun A, Santos M, Szathmary E. Real ribozymes suggest a relaxed error threshold. *Nat Genet* 2005;37:1008–1011.

127. Levy M, Griswold KE, Ellington AD. Direct selection of trans-acting ligase ribozymes by *in vitro* compartmentalization. *RNA* 2005;11:1555–1562.

128. Szabo P, Scheuring I, Czaran T, Szathmary E. *In silico* simulations reveal that replicators with limited dispersal evolve towards higher efficiency and fidelity. *Nature* 2002;420: 340–343.

129. Szostak JW, Bartel DP, Luisi PL. Synthesizing life. *Nature* 2001;409:387–390.

130. Hanczyc MM, Szostak JW. Replicating vesicles as models of primitive cell growth and division. *Curr Opin Chem Biol* 2004;8:660–664.

131. Chakrabarti AC, Breaker RR, Joyce GF, Deamer DW. Production of RNA by a polymerase protein encapsulated within phospholipid vesicles. *J Mol Evol* 1994;39: 555–559.

132. Chen IA, Salehi-Ashtiani K, Szostak JW. RNA catalysis in model protocell vesicles. *J Am Chem Soc* 2005;127:13213–13219.

133. Oberholzer T, Wick R, Luisi PL, Biebricher CK. Enzymatic RNA replication in self-reproducing vesicles: an approach to a minimal cell. *Biochemical and biophysical research communications* 1995;207:250–257.

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

135. Atencia J, Beebe DJ. Controlled microfluidic interfaces. *Nature* 2005;437:648–655.

136. Rasmussen S, Chen L, Nilsson M, Abe S. Bridging nonliving and living matter. *Artif Life* 2003;9:269–316.

137. Tawfik DS, Griffiths AD. Man-made cell-like compartments for molecular evolution. *Nat Biotechnol* 1998;16:652–656.

138. Ghadessy FJ, Holliger P. A novel emulsion mixture for *in vitro* compartmentalization of transcription and translation in the rabbit reticulocyte system. *Protein Eng Des Sel* 2004;17:201–204.

139. Ghadessy FJ, Ong JL, Holliger P. Directed evolution of polymerase function by compartmentalized self-replication. *Proc Natl Acad Sci USA* 2001;98:4552–4557.

140. Agresti JJ, Kelly BT, Jaschke A, Griffiths AD. Selection of ribozymes that catalyse multiple-turnover Diels-Alder cycloadditions by using *in vitro* compartmentalization. *Proc Natl Acad Sci USA* 2005;102:16170–16175.

141. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer ML, Jarvie TP, Jirage KB, Kim JB, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF, Rothberg JM. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 2005;437:376–380.

142. Bernath K, Magdassi S, Tawfik DS. Directed evolution of protein inhibitors of DNA-nucleases by *in vitro* compartmentalization (IVC) and nano-droplet delivery. *J Mol Biol* 2005;345:1015–1026.

143. Cohen HM, Tawfik DS, Griffiths AD. Altering the sequence specificity of HaeIII methyltransferase by directed evolution using *in vitro* compartmentalization. *Protein Eng Des Select* 2004;17:3–11.

144. Griffiths AD, Tawfik DS. Directed evolution of an extremely fast phosphotriesterase by *in vitro* compartmentalization. *EMBO J* 2003;22:24–35.

145. Ghadessy FJ, Ramsay N, Boudsocq F, Loakes D, Brown A, Iwai S, Vaisman A, Woodgate R, Holliger P. Generic expansion of the substrate spectrum of a DNA polymerase by directed evolution. *Nat Biotechnol* 2004;22:755–759.

146. Spiegelman S. An approach to the experimental analysis of precellular evolution. *Q Rev Biophys* 1971;4:213–253.

147. Breaker RR, Joyce GF. Emergence of a replicating species from an *in vitro* RNA evolution reaction. *Proc Natl Acad Sci USA* 1994;91:6093–6097.

148. Dobson CM, Ellison GB, Tuck AF, Vaida VV. Atmospheric aerosols as prebiotic chemical reactors. *Proc Natl Acad Sci USA* 2000;97:11864–11868.

149. Forterre P. The two ages of the RNA world, and the transition to the DNA world: a story of viruses and cells. *Biochimie* 2005;87:793–803.

150. Seeman NC. DNA in a material world. *Nature* 2003;421:427–431.

151. Mao C, Sun W, Shen Z, Seeman NC. A nanomechanical device based on the B-Z transition of DNA. *Nature* 1999;397:144–146.

152. Yurke B, Turberfield AJ, Mills AP Jr, Simmel FC, Neumann JL. A DNA-fuelled molecular machine made of DNA. *Nature* 2000;406:605–608.

153. Yan H, Zhang X, Shen Z, Seeman NC. A robust DNA mechanical device controlled by hybridization topology. *Nature* 2002;415:62–65.

154. Shin JS, Pierce NA. A synthetic DNA walker for molecular transport. *J Am Chem Soc* 2004;126:10834–10835.

155. Liao S, Seeman NC. Translation of DNA signals into polymer assembly instructions. *Science* 2004;306:2072–2074.

156. Benenson Y, Paz-Elizur T, Adar R, Keinan E, Livneh Z, Shapiro E. Programmable and autonomous computing machine made of biomolecules. *Nature* 2001;414:430–434.

157. Benenson Y, Gil B, Ben-Dor U, Adar R, Shapiro E. An autonomous molecular computer for logical control of gene expression. *Nature* 2004;429:423–429.

158. Shih WM, Quispe JD, Joyce GF. A 1.7-kilobase single-stranded DNA that folds into a nanoscale octahedron. *Nature* 2004;427:618–621.

159. Henry AA, Romesberg FE. The evolution of DNA polymerases with novel activities. *Curr Opin Biotechnol* 2005;16:370–377.

160. Sprinzak D, Elowitz MB. Reconstruction of genetic circuits. *Nature* 2005;438:443–448.

161. Seeman NC. From genes to machines: DNA nanomechanical devices. *Trends Biochem Sci* 2005;30:119–125.