

Review

# Synthetic Life with Alternative Nucleic Acids as Genetic Materials

Peng Nie, Yanfen Bai and Hui Mei \* 

Shenzhen Key Laboratory of Synthetic Genomics, Guangdong Provincial Key Laboratory of Synthetic Genomics, CAS Key Laboratory of Quantitative Engineering Biology, Shenzhen Institute of Synthetic Biology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China; peng.nie@siat.ac.cn (P.N.); yf.bai@siat.ac.cn (Y.B.)

\* Correspondence: hui.mei@siat.ac.cn

Academic Editor: Katherine Seley-Radtke

Received: 14 July 2020; Accepted: 29 July 2020; Published: 31 July 2020

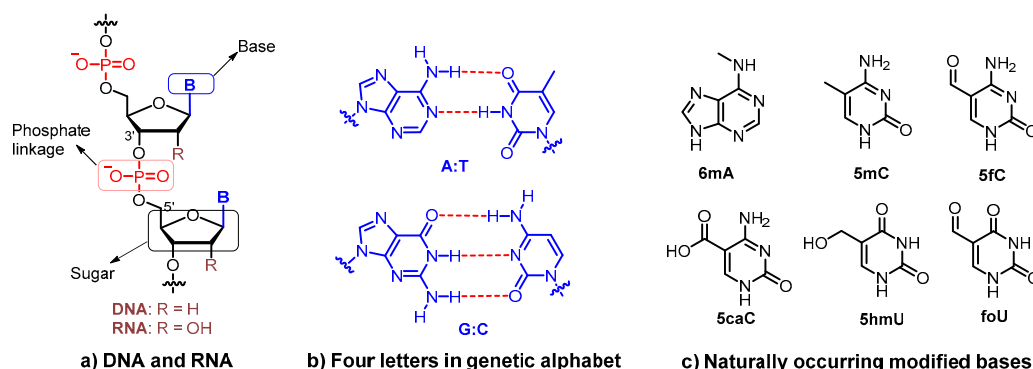


**Abstract:** DNA, the fundamental genetic polymer of all living organisms on Earth, can be chemically modified to embrace novel functions that do not exist in nature. The key chemical and structural parameters for genetic information storage, heredity, and evolution have been elucidated, and many xenobiotic nucleic acids (XNAs) with non-canonical structures are developed as alternative genetic materials *in vitro*. However, it is still particularly challenging to replace DNAs with XNAs in living cells. This review outlines some recent studies in which the storage and propagation of genetic information are achieved *in vivo* by expanding genetic systems with XNAs.

**Keywords:** synthetic biology; xenobiotic nucleic acids (XNAs); genetic material; synthetic life

## 1. Introduction

DNA is the fundamental genetic material of all living organisms on Earth that stores and propagates genetic information. It is widely acknowledged that DNA, with a very uniform structure, is composed of four nucleobases (A, adenine; G, guanine; C, cytosine; T, thymine), 2'-deoxyribose, and charged phosphate backbones (Figure 1a). Four letters in the genetic alphabet form two base pairs (A:T and G:C) following the complementary rule, which is essential for the formation of a double helix structure and genetic information transmission (Figure 1b) [1].



**Figure 1.** Structures of genetic materials. (a) Structures of DNA and RNA containing sugar, base, and phosphate moieties; (b) Watson–Crick base pairs formed by A:T, and G:C in DNA; (c) Examples of modified bases observed in nature.

Nevertheless, in nature, considerable chemical variations of the nucleobase, normally referred to as epigenetic markers, are observed in both prokaryotic and eukaryotic genomic DNA [2–6].

Modified bases, including but are not limited to N6-methyladenine (m6A), 5-methylcytosine (5mC), 5-formylcytosine (5fC), 5-carboxycytosine (5caC), 5-hydroxymethyl-uracil (5hmU), and 5-formyluracil (foU), have been identified to convey epigenetic information [5] (Figure 1c). In eukaryotes, such modifications function as a regulatory element, representing a second layer of regulatory information beyond the essential information encoded in the base sequence [4]. In mammals, 5mC and 5hmC are abundant in the brain and play an important role in organizing the chromatin structure and regulating gene expression [7]. Methylated nucleobases, such as m6A and 5mC, are observed in bacteria to protect cellular DNA from endonuclease-mediated cleavage that destroys the invading bacteriophage or viral DNA [8]. Moreover, m6A and 5mC are also active in cellular processes, such as DNA replication, transcription, transposition, and DNA repair [6]. The discovery of novel DNA modifications with different biological functionalities has become an attractive research area.

In most living organisms, only a minor portion of the nucleobases are modified in genomic DNA to bring in regulatory or protective functionalities. However, certain modified nucleobases can completely replace the canonical bases in some bacteriophages [6,9]. For example, in the DNAs of *Bacillus* phages SPO1, SP8, H1, 2C, and SP82, thymidine is completely replaced by 5hmdU [10]. Thymidine is substituted for deoxyuracil (dU) in the whole genome of *Bacillus subtilis* bacteriophages PBS1 and PBS2 [11,12]. Specific metabolic pathways are discovered in these phages that produce altered intracellular dNTP pools to synthesize modified DNA that are capable of escaping from host DNA repair systems, providing a possible method of using living cells with engineered genes that mimic phage genes for the incorporation of synthetic nucleotides into artificially built genetic systems [13,14].

To investigate the key chemical and structural parameters for genetic information storage, heredity, and evolution in vitro, a series of xenobiotic-nucleic acids (XNAs) are synthesized by replacing natural bases, sugars, and phosphate linkages with their unnatural counterpart. Modification of three subunits of nucleotides leads to sugar-modified XNAs, phosphate-modified XNAs, as well as base-modified XNAs, or their combination. Some of these XNAs can mimic natural nucleic acids to form a stable double helix between DNA/RNA or themselves following Watson–Crick base-pairing rules. The biological evaluation of these XNAs gives inspiring insights into the question of why nature chooses DNA/RNA as genetic materials rather than other chemicals, a fundamental question of the origin of life.

To further understand the potential of XNAs for genetic heredity, XNA replication should be accomplished. Genetic information transfer is achieved through DNA replication mediated by DNA polymerases in nature. However, in most cases, XNA building blocks, unnatural nucleotides, are not good substrates for natural DNA polymerases due to their high specificity [15,16]. Under such circumstances, laboratory-evolved XNA polymerases have been adopted to transmit and propagate the genetic information stored in XNAs with improved efficiency and fidelity [17,18]. Recent achievements proved that some XNAs function as alternative genetic materials in vitro with the function of information storage and propagation [19–21].

The following question is whether these XNAs can replace DNAs, and function as genetic materials inside living cells. Due to the complex biological and chemical environment, the invasion of XNA into a living cell is far more difficult to achieve. In this review, we will briefly introduce different kinds of XNAs with various chemical modifications and specific biological functions, and summarize some initial studies on the in vivo implementation of building xenobiotic life with XNAs. We hope that this review will encourage more systematic research on the exploration of XNA-based synthetic biology.

## 2. XNAs in Synthetic Biology

### 2.1. Sugar-Modified XNAs

One of the motivations to study XNAs is to develop modified oligonucleotides with improved chemical and biological properties, which can enable them to function more effectively over DNA/RNA for biomedical, biotechnology, and nanotechnology applications. Sugar-modified XNAs, such as bicyclo-DNA [22], LNA (locked nucleic acid) [23,24], and HNA (hexitol-nucleic acid) (Figure 2) [25],

have been developed and display better chemical stability, superior nuclease resistance, and improved pharmacokinetic properties. Another goal is to explore possible explanations for why nature chooses ribofuranosyl sugar-based nucleic acids as information carriers. Thus, XNAs with “potentially natural” sugar moieties are developed and examined in the Eschenmoser group [26]. Sugars, including arabinose, lyxose, threose, allose, mannose, and glucose, are postulated to coexist with ribose under prebiotic circumstances.

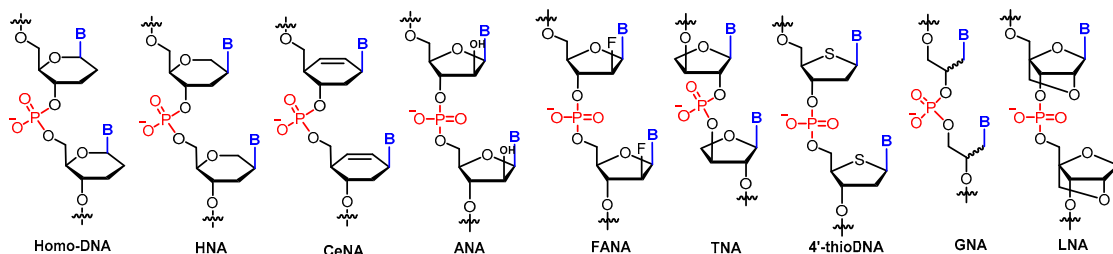


Figure 2. XNAs with artificial backbones.

Modified nucleic acids with various cyclic or acyclic sugar moieties were synthesized and systematically studied regarding their duplex formation between themselves or with DNA/RNA [26]. It was found that the stability of these sugar-modified XNA duplexes varies significantly. For example, the stabilities of allose-NA and altrose-NA duplexes are very low [26]. In contrast, homo-DNA, an oligonucleotide composed of 2',3'-dideoxyallopribose instead of 2'-deoxyribofuranose, forms a more stable Watson–Crick duplex than the corresponding DNA [27]. Interestingly, homo-DNA does not cross-pair with DNA or RNA as homo-DNA shows a ladder-like geometry but not the typical B/A-form helical geometry observed in DNA or RNA, indicating that homo-DNA is not able to transfer genetic information to DNA [27]. Only a small number of XNAs can form a stable antiparallel duplex with DNA/RNA following the canonical base-pairing rules, with examples including HNA [25,28], TNA (threose nucleic acid) [29], CeNA (cyclohexenyl nucleic acid) [30], ANA (arabino nucleic acid) [31], FANA (2'-fluoro-arabino nucleic acid) [32], GNA (glycol nucleic acid) [33], and LNA [34] (Figure 2).

In order to evaluate the feasibility of using XNAs for heredity and evolution, a replication system should be developed. Non-enzymatic replication of nucleic acids has been studied to understand the chemical origin of genetic materials [35,36]. It was found that chemical replication suffers with insufficient efficiency, low fidelity, and limited length of polymerization. In nature, replication is catalyzed by DNA polymerase with high efficiency and fidelity. Due to the structural difference between XNAs and DNA/RNA, normally, XNAs are not good substrates for natural DNA polymerases. Even though genetic information is stored in XNA sequences, it is still invisible to the natural system. With the help of directed evolution, specific XNA polymerases were obtained to transfer the genetic information back and forth between XNA and DNA with reasonable fidelity and efficiency [17,18]. Using a compartmentalized self-tagging (CST) selection strategy, specific XNA polymerases that can synthesize HNA/CeNA from DNA templates were obtained by Holliger and coworkers [21]. XNA replication systems with high fidelity were built for six XNAs (HNA, CeNA, LNA, ANA, FANA, and TNA). The replications between FANA and FANA, CeNA and CeNA, or HNA and CeNA were also realized based on these engineered polymerases.

Next, the capacity of XNAs for Darwinian evolution was evaluated by *in vitro* selection of functional aptamers and catalysts. Aptamers are single-stranded oligonucleotides that fold into complex 3-D structures and bind to specific targets, and are commonly produced by systematic evolution of ligands by exponential enrichment (SELEX), also referred to as *in vitro* selection or *in vitro* evolution. To select XNA aptamers, the XNA library used in SELEX has to be amplified by XNA polymerases, and selected XNAs should be able to be transcribed into cDNA. An HNA aptamer binding to the HIV TAR RNA motif was obtained with the dissociation constant (KD) between 28 and 67 nM by using the above-mentioned HNA replication system [21]. Similarly, FANA aptamers that bind to HIV-1

integrase or reverse transcriptase (RT) were selected with picomolar affinities [37,38]. Computational design and sampling and pooling of beneficial mutations were utilized by Chaput's group to generate TNA polymerase Kod-RI with high efficiency [39,40]. With this enzyme, a biologically stable TNA aptamer was discovered with a high binding affinity to HIV RT [41]. Starting from random XNA oligomer pools, functional XNAs (ANA, FANA, HNA, and CeNA) with trans-RNA endonuclease and ligase activities (XNAzymes) were obtained by in vitro selection [42]. Recently, a general RNA-cleaving FANA enzyme (FANAzyme) that functions with enhanced activity was evolved in the laboratory by Chaput's group [43].

In summary, like DNA/RNA, these sugar-modified XNAs are also capable of undergoing Darwinian evolution to acquire specific target binding affinities or catalytic activities [21,41–44]. These results reveal that some XNAs, which are capable of heredity and evolution, can be used as alternative genetic materials in vitro.

The in vivo implementation of sugar-modified XNA, such as HNA and CeNA, into living systems was carried out by Herdewijn and coworkers. HNA nucleoside analogues were initially developed as potential antiviral agents due to their potent activities against DNA viruses [45]. Later on, HNA oligonucleotides were prepared and found to be chemically and enzymatically stable nucleic acids that hybridized with DNA, RNA, and itself based on its structural characteristics [28]. CeNA is composed of carbocyclic nucleotides that contain a cyclohexenyl sugar moiety with a C=C double bond, which makes the cyclohexene system more flexible than furanose ring. Interestingly, it is reported that similar conformation changes between CeNA and natural DNA were observed, indicating the possibility of using CeNA as a biocompatible polymer [30].

Up to six continuous HNA or CeNA nucleotides in the template were read by the *E. coli* cell machinery to give ThyA-produced colonies [46]. The *thyA* gene encoded thymidylate synthase, which is necessary for *E. coli* cells to grow in thymidine-deficient medium. To introduce HNA into the plasmid, restriction enzymes NheI and NsiI were utilized to digest the *thyA* gene, and six codons around the cysteine at position 146 were deleted to create a gap for the ligation of XNAs (Figure 3). After ligation, the plasmids were transformed into a *thyA*-deletion strain to test their biological activity. The results showed that HNA served as a short template to produce the active *thyA* gene and restore the function of thymidylate synthase. The relatively poor acceptance of HNA for the endogenous replication machinery led to a decreased number of *thyA*<sup>+</sup> colonies. Even though T and G were replaced by uracil and hypoxanthine in the these XNAs, ANA and HNA still succeeded in restoring the functional *thyA* gene and generating survival colonies. Similar results were obtained when in vivo genetic selection screening was performed with CeNA and ANA (Figure 3). Unfortunately, mosaic template containing GNA did not yield any of the *thyA* colonies even though one codon was substituted by GNA.



**Figure 3.** Genetic information transfer from XNAs to DNAs in vivo.

Although genetic information encoded in XNAs can be transliterated to DNAs, it is difficult for these XNAs to directly function as a template for transcription. However, for 4'-thioDNA, a natural-like XNA in which the 4'-oxygen atom is substituted by a sulfur atom, the direct transcription can be performed in human cells. As reported, 4'-thioDNA is resistant against nuclease and hybridizes like RNA molecules [47]. 4'-ThioDNA containing 4'-thio-dT and 4'-thio-dC was amplified by DNA polymerase from *Pyrococcus kodakaraensis* (KOD dash DNA polymerase) and used as template for transcription both in vitro and in vivo to produce RNA with gene-silencing activity (Figure 4) [48].

Further exploration of the transcription was carried out by using 4'-thioDNA templates with 4'-thioA and 4'-thioG nucleotides to afford luciferase in murine cells, showing that 4'-thioDNA is a potential candidate for the development of a synthetic genetic system [49].



Figure 4. 4'-ThioDNA with gene-silencing activity in mammalian cells.

Chemical alteration of the ribofuranose ring induces conformation changes that consequently lead to a lower level of acceptance of synthetic nucleotides by natural DNA/RNA polymerases. Directed evolution affords engineered polymerases that can transfer the genetic information to promote the application of XNAs in living cells [19]. Significant achievements have been achieved for the development of engineered polymerases that can transfer genetic information between XNAs and DNA to date. The next challenge is to develop XNA-templated XNA polymerases for XNA replication. Engineered RNA polymerase that can transcribe XNAs into RNAs is needed for expanding the central dogma with XNAs. The discovery of these new enzymes may allow us to create artificial life forms based on XNA genetic systems.

## 2.2. Phosphate-Modified XNAs

The classic phosphate diester linkages are substituted by different functional groups in sulfone-DNA [50], PS-DNA [51], PN-DNA [52], NP-DNA [53], triazole-DNA [54], and dPhoNA [55] (Figure 5).

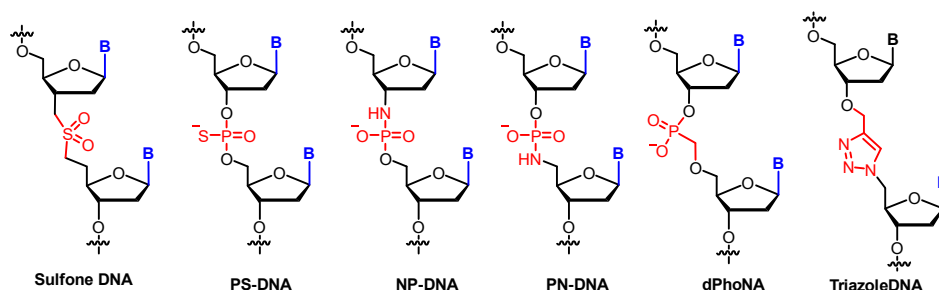


Figure 5. Examples of phosphate linkage-modified XNAs.

In natural DNA/RNA, phosphate diester is the linker between nucleosides and is negatively charged at physiological pH. It is demonstrated that the repeating negative charges retain DNA/RNA within lipid membranes and stabilize the diester linkages against hydrolysis [56]. Investigations on an uncharged DNA analogue in which the phosphate diester was replaced by a dimethylenesulfone group [50] showed that charged phosphate linkage is also important for base-pairing and Darwinian evolution [57]. Even though the short oligomer of sulfone-DNA exhibits a similar geometry to natural DNA with G:C and C:G pairing [58], self-aggregation was observed when the oligomers were long enough. Single nucleobase alternation leads to significant changes in the solubility, aggregation, and chemical reactivities of sulfone-DNA, suggesting that it is not a suitable genetic polymer in water.

PS-DNA, a modified XNA in which the non-bridging oxygen of the phosphate diester is replaced by a sulphur atom, was synthesized, aiming to develop antiviral agents against HIV in the 1980s [59]. The phosphorothioate motif shares a similar structure with the commonly seen phosphate diester linkage and offers considerable advantages, such as enhanced resistance to nucleases [60]. It is also reported that naturally occurring phosphorothioation was discovered in bacterial DNA and played

an important role in many cellular processes [60]. The replacement of oxygen by sulphur takes place under the catalysis of the gene products of the *dndABCDE* cluster. PS-DNA, together with DndFGH, constitutes a defense system that can distinguish and attack non-PS-modified foreign DNA [60–62]. PS-DNA is an alternative genetic material that can be potentially used for the construction of artificial life forms.

DNA with N3'→P5' phosphoramidate linkages (NP-DNA) was first developed by Gryaznov and co-workers for therapeutic applications [53,63]. Although capable of forming stable duplexes with both RNA/DNA strands, NP-DNA is functionally more similar to RNA because its structure closely resembles the classical A-form RNA structure. It is used as a model to study the nonenzymatic template-directed replication, which may provide insight into the chemical replication of genetic materials [64]. NP-DNA can be synthesized from RNA template and extended up to 25 nucleotides inside a model protocell [65]. Enzymatic replication or reverse transcription of NP-DNA were performed by Szostak and coworkers, showing that NP-DNA might be a good candidate for the construction of synthetic life [52,66]

Recently, Herdewijn's group reported the *in vivo* study of a synthetic genetic polymer bearing the P3'→N5' phosphoramidate linkages, which is denoted as PN-DNA [67]. Compared to DNA, PN-DNA is more stable under basic conditions and more acid-labile, such that the phosphoramidate linkage can be cleaved under acid conditions [68]. Enzymatical synthesis of PN-DNA was successfully achieved by employing Taq DNA polymerase together with KF (Klenow fragment) and Vent (*exo*<sup>-</sup>) polymerases. Multiple NH-dCTPs were incorporated into the R67DHFR gene sequence to give modified plasmids with trimethoprim resistance, which was then transformed into *E. coli* cells on trimethoprim-containing medium. With sufficient accuracy, the *in vivo* transliteration of PN-DNA to natural DNA was achieved, thus providing viable colonies with antibiotic resistance.

dPhoNA represents a family of nucleic acids that contain a phosphonate linkage rather than the natural phosphate diester. Phosphonate is an isostere of phosphate diester that contains a stable P-C bond, which is resistant against chemical and enzymatical degradation. Terminator polymerase is demonstrated as a catalyst for the condensation of the phosphonate derivatives of adenine to afford dPhoNA with enhanced stability against nucleolytic degradation [69]. However, the polymerases are not able to catalyze the synthesis of a sequence longer than four different nucleotide phosphonates [70]. Directed evolution may provide better polymerases with improved recognition functions for the development of dPhoNA-based genetic systems.

Another interesting and successful example of phosphate-modified XNAs is triazoleDNA, which comprises a triazole ring between nucleotides instead of natural phosphate diesters. As presented in Figure 6, treatment of the pRSET-mCherry plasmid with NdeI and EcoRI leads to the cleavage of the *mCherry* gene, furnishing pRSET fragments after gel purification. The azide-alkyne click reaction was employed for the assembling of the *iLOV* gene, which was then ligated to the pRSET plasmid in the presence of T4 DNA ligase. The biocompatibility of triazole linkers in the *iLOV* gene was examined by *in vivo* replication and transcription in *E. coli* cells to produce a green fluorescent protein successfully [71]. As the first example of a non-natural linker being functional in human cells, triazoleDNA was directly transcribed into mRNA and encodes the fluorescent protein [72]. Notably, the overall physical and chemical properties of DNA were maintained after the insertion of a limited number of triazole linkers into DNA.



Figure 6. Transcription and translation of triazoleDNA *in vivo*.

### 2.3. Sugar- and Phosphate-Modified XNAs

Both the sugar moieties and the phosphate are chemically modified in some nucleoside analogues. The combination of phosphonate linkages and artificial sugar rings gave rise to the discovery of two orthogonal XNAs: tPhoNA (3'-2' phosphonomethyl-threosyl nucleic acid) [73] and ZNA (an XNA analogue with an acyclic methylphosphonate backbone) [74]. The former contains a threose ring while the latter one bears an acyclic backbone (Figure 7). Initially, the repeating nucleoside units of both tPhoNA and ZNA are developed for antiviral drug discovery [75,76]. tPhoNA successfully served as a xenobiotic genetic material in vitro in the presence of engineered polymerases. Further in vivo studies showed that tPhoNA was not a good substrate for *E. coli* DNA polymerases, indicating that tPhoNA was not accessible to the natural system. Based on the obtained results, an orthogonal genetic system could be built with tPhoNA and laboratory-evolved tPhoNA polymerases.

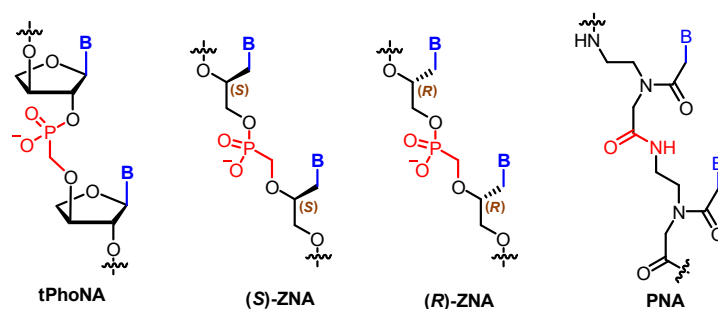


Figure 7. Structures of tPhoNA, ZNA, and PNA.

Due to the existence of a chiral carbon center in the backbone, ZNA is obtained as two enantiomers, (S)- and (R)-ZNA. (S)-ZNA strongly hybridizes with itself and interacts weakly with the complementary single-stranded DNA, showing that ZNA is chemically orthogonal. Biologically, the diphosphates of (S)-ZNA monomers can be recognized as substrates by *E. coli* polymerase I and the algal nucleotide transporter. Successful in vivo transliteration of ZNA to DNA was conducted inside *E. coli*, which proved that ZNA can be potentially used for genetic information propagation.

Protein, which is composed of amino acid residues linked by amide bonds, is another biological polymer in living organisms other than DNA/RNA. Whether polyamide linkers could be used instead of phosphate diesters in DNAs has been investigated towards the development of alternative genetic materials. Invented by Nielsen, Egholm, and their collaborators, peptide nucleic acid (PNA) is a neutral oligomer in which the whole deoxyribose phosphodiester backbone is replaced by N-(2-aminoethyl)glycine [77] (Figure 7). It forms very stable duplex structures with complementary DNA, RNA, or PNA. PNA has been used in the areas of gene therapy, genetic diagnostics, and nanotechnology [78].

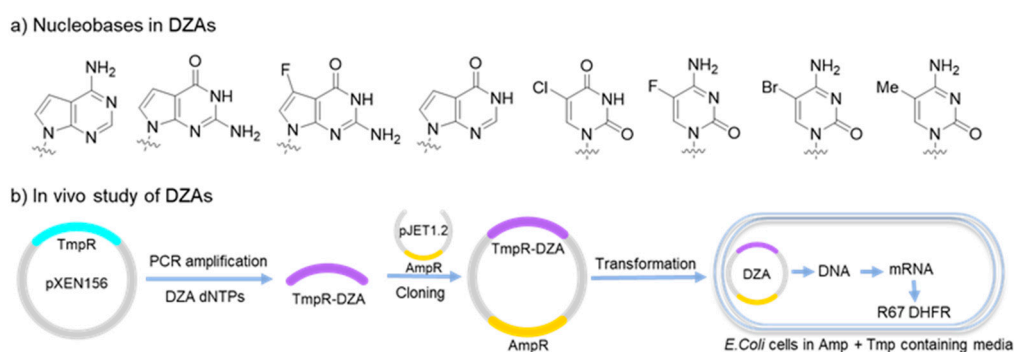
As a molecule that lies between protein and DNA, PNA was studied to test the origin of life since it might be synthesized and used in the early Earth. The synthesis of PNA building blocks was achieved by Nielson under simulated prebiotic conditions [79]. Orgel and coworker demonstrated that the genetic information transfer between PNA and RNA was successful [80]. The description of the PNA self-replicating system and non-templated PNA replication prompted the speculation that PNA may be a prebiotic genetic material [81,82]. According to Benner, the repeating charge in DNA might be a universal feature of genetic molecules in water, including those in alien life throughout the cosmos [57]. Without negative charges, the self-aggregation and precipitation of PNA in water were observed [83]. The possibility of using PNA as genetic materials or prebiotic genetic materials is still unclear, thus further investigation is required.

#### 2.4. Base-Modified XNAs

Genetic information is stored in nucleic acid sequences and transmitted following Watson–Crick base-pairing rules during DNA replication and transcription. In nature, some nucleobases are damaged by oxidation, depurination, deamination, or other mutagens, leading to decreased accuracy of information transfer [84,85]. Base-modified XNAs with improved properties are considered and used instead of natural DNAs.

The functionalities of DNAs have been widely expanded by chemical manipulation of nucleobases over the past few decades. These unnatural nucleobases are used in SELEX for the development of modified aptamers and deoxyribozymes with desired functions [86]. For example, selections from DNA pools containing 2'-deoxyuridine derivatives with hydrophobic groups at the 5-position of the nucleobase generate aptamers that have more accessible epitopes on protein targets and broader scope of applications in diagnostics and therapeutics [87]. In addition, SELEX and related in vitro selections are utilized to obtain DNAzymes, which are highly specific catalytic DNAs bearing modified bases [88]. Some RNA-cleaving DNAzymes having protein-like side chains attached to the bases were found to be active in the absence of divalent metal ions [89] despite the fact that most of the selected DNAzymes require multivalent metal cations for decent activities.

DZA, described as 'a fully morphed DNA containing all four non-conical nucleotides', was developed by Herdewijn's group. Unnatural bases used in DZAs share similar skeletons with their natural counterparts (Figure 8a), including 5-chloro-2'-deoxyuridine (5CIU), 5-methyl-2'-deoxycytidine (5MeC), 5-fluoro-2'-deoxycytidine (5FC), 7-deaza-2'-deoxyadenosine (7dA), 7-deaza-2'-deoxyguanosine (7dG), 7-fluoro-7-deaza-2'-deoxyguanosine (7FG), and 2'-deoxyinosine (dI). Taq or Vent (*exo*<sup>-</sup>) DNA polymerases served as catalyst for the PCR amplification of DZAs. DZA fragments efficiently block the restriction sites from enzymatic cleavage, representing a better property as alternative genetic materials. In vivo studies were performed to examine the biological functions of DZAs. A pXEN156 plasmid was employed as parent template to synthesize DZA-containing genes by PCR amplification, followed by ligation to a pJET1.2 plasmid with ampicillin resistance (AmpR) by T4 DNA ligase. The modified plasmid was then transformed into *E. coli* cells cultured in media with Amp and trimethoprim (Tmp). Transliteration of DZAs was successfully achieved in *E. coli* cells to generate the R67DHFR protein, which is resistant to Tmp (Figure 8b) [90,91]. The results suggest that DZAs are biocompatible genetic materials with new functionality and can be used for the construction of synthetic life forms. Additionally, such DZA libraries can be potentially used for the selection of functional aptamers and DNAzymes.

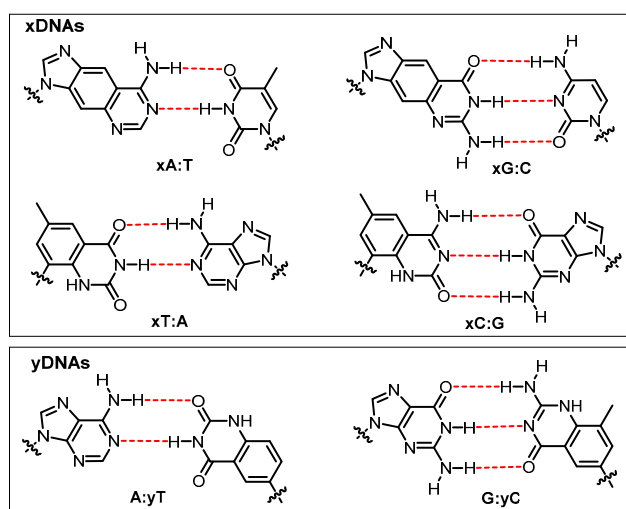


**Figure 8.** DZAs as alternative genetic materials. (a) Non-canonical nucleobases used in DZAs; (b) DZAs used as genetic template in bacterial cells.

Modified bases that consist of fused aromatic rings for the construction of synthetic genetic system were synthesized and evaluated by Kool's group. Embedded with size-expanded artificial bases (xDNA and yDNA, Figure 9), these DNAs are fluorescent and display high stacking affinity compared to natural DNA [92]. They were utilized to build a synthetic genetic set through cloning and expression



in *E. coli* cells [93,94]. The pGFPuv plasmid was digested by the restriction enzymes BsrGI and MluI followed by the removal of 5'-phosphate by antarctic phosphatase to afford DNA fragments ready for the ligation of xDNA segments. After ligation under T4 DNA ligase catalysis, the xDNA-containing plasmid was transformed into *E. coli* to yield green fluorescent protein successfully. It was demonstrated that a maximum of four consecutive modified base pairs can be read accurately by cellular enzymes in *E. coli* to generate native DNA and express functional proteins [93]. Further explorations of sized-expanded yDNAs were executed by testing their ability to store and transfer genetic information in vitro and in vivo. Primer extension assays were performed with KF  $exo^-$  and the thermostable DNA polymerase Vent  $exo^-$ , showing that the enzymes incorporated the natural nucleotides' opposite yDNA bases correctly, but the selectivity was unsatisfactory (T-yT and T-yC mispairing was observed). Plasmids with one or two yDNA bases were evaluated in vivo, leading to the generation of a functional protein [94].



**Figure 9.** Structures of sized-expanded bases in xDNA and yDNA.

Complete replacement of natural nucleobases with non-canonical ones in the genomes of living organisms is challenging. Peter G. Schultz and coworkers reported that engineered *E. coli* cells were obtained by replacing thymidine with 5-hydroxymethyluridine (5hmU) [13], or the replacement of cytidine with 5-hydroxymethylcytosine (5hmC) [14]. For 5hmU incorporation, the metabolic pathway of pyrimidine nucleotides in *E. coli* was engineered to mimic the metabolic approach used by SPO1 bacillus phage. A series of biochemical reactions that converted deoxyuridine monophosphate (dUMP) into 5hmdUTP took place in engineered *E. coli* under the catalysis of cellular enzymes, including hydroxymethylase, mononucleotide kinase, and nucleotide diphosphate kinase. The synthesis of TTP was inhibited by the disruption of the *thyA*, *deoA*, and *trmA* genes of *E. coli* as TTP is the competitor of 5hmdUTP during DNA replication. As a result, 5hmdUTP was successfully incorporated into the *E. coli* DNA genome and approximately 75% of the thymidine was replaced by 5hmU [13]. Similarly, the replacement of natural dC with 5hmC residues was studied. The biosynthetic pathway of *E. coli* was manipulated using bacteriophage T4 genes to incorporate 5hmC into the genome. Having these modified genes, the intracellular synthesis of 5hmdCTP was successfully achieved, offering 5hmdCTP as competitive substrates with dCTP for DNA synthesis. The results showed that about 63% of 2'-deoxycytidine was displaced by 5hmC in the *E. coli* genome [14]. Further experiments may be required to investigate the key physiological regulators that limit entire genomic replacement.

Surprisingly, mutated *E. coli* strains containing chimeric DNA–RNA sequences (40–50% ribonucleotide content) in their genomic DNA were observed during the exploration to improve 5hmC content [95]. It has been described that DNA polymerases can incorporate ribonucleotides and DNA–RNA chimeras are eligible for transcription. Recent studies postulate that homogeneous DNA

genomes may have evolved from heterogeneous chimeric RNA–DNA templates [96]. The detection of this DNA–RNA chimera may provide further insights into the origin of life [95].

Substitution of thymidine by bromodeoxyuridine (BrdU) in the DNA of the *E. coli* strain TAU-bar [97] or a hamster cell line [98] was performed decades ago. Even though the van der Waals radii of the bromine atom and methyl group are almost the same, the higher electronegativity of bromine affects the electron distribution of the aromatic ring, which may result in mispairing in DNA owing to a keto–enol tautomerism [97]. The replacement was achieved smoothly in that BrdU was quantitatively incorporated into *E. coli* DNA. As demonstrated by Davidson, the level of BrdU substitution in the DNA of the hamster cell line was higher than 99.8%, and possibly 100%. However, in both cases, the presence of high levels of BrdU in the genomic DNA raised some problems, such as a similar phenomenon to thymineless death and mutagenesis caused by base transitions.

Recently, investigations on the evolution of genomic DNA of *E. coli* strains with 5-chlorouracil (5-CIU) were conducted by Rupert Mutzel and coworkers, showing that 98.4% of thymine was displaced [99]. Despite the presence of chlorine in various natural products, it is not found in any natural nucleobase. The reasons for choosing 5-CIU as a substitute for thymine include (i) the structure of 5-CIU closely resembles thymine, which allows it to form a stable base pair with adenine; (ii) it can be converted into chlorodeoxyuridine nucleotides by a panel of cellular enzymes; and (iii) it is chemically more stable than fluoro, bromo, and iodo analogues, which enables it to avoid being reduced to uracil in the cytoplasm and ambiguous pairing caused by tautomerism [99]. The bacterial strain THY1 used for the transliteration of 5-chlorouracil was prepared by reconstruction of a wild-type *E. coli* strain MG1655, including deletion of the *thyA* and *udp* genes and the *deoCABD* operon, and addition of a plasmid containing the *Lactobacillus leichmannii* gene *ntd*. After cultivation and evolution for 25 weeks, the modified strains were able to grow with chlorouracil and the thymidine residual fraction was below 2%. As a result, 1502 genetic mutations were observed in this artificially constructed strain due to the genomic changes, revealing the difficulty of replacing DNA with XNA in the bacterial genome.

### 2.5. Unnatural Base Pairs (UBPs)

Although the nucleobase, sugar, or backbone moieties have been reconstructed in the above mentioned DZAs or XNAs, Watson–Crick base pairing geometry is still adopted and no new genetic codon is created. In contrast, insertion of a third base pair (UBP) other than A:T and G:C into DNA strands greatly expands the genetic code, offering an opportunity to create more life forms that do not exist in nature [100–102]. Numerous unnatural base pairs (UBPs), also known as artificially expanded genetic alphabets, have been developed ever since the first UBP (isoG: isoC) was synthesized and incorporated successfully into DNA/RNA by Steven Benner in 1989 [103]. Other UBPs based on altered hydrogen bonding were created, such as dB:dS and P:Z base pairs (Figure 10) [104,105], and subsequently used for the construction of a genetic system containing eight letters. Compared with the natural genetic system, eight-letter DNA and RNA possess an enhanced ability to store data and may provide insights into the search for extraterrestrial life [105].

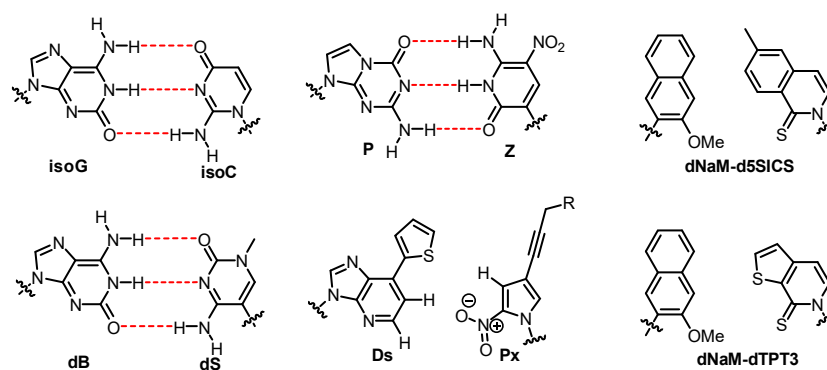


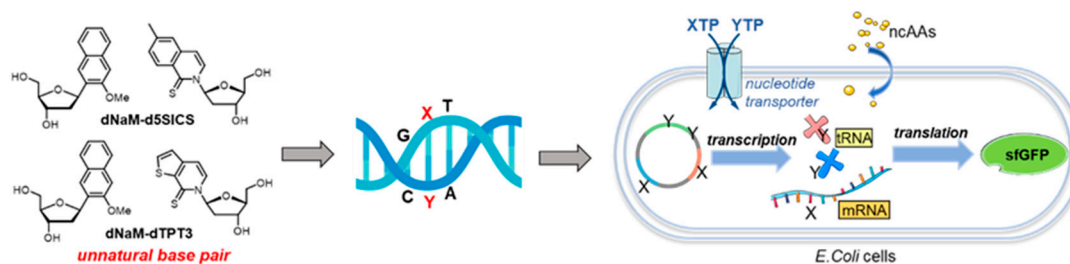
Figure 10. Structures of UBPs.

Despite the fact that hydrogen bonding is very important for the formation and stability of the DNA structure, it is not the only driving force. It has been addressed that base stacking, the neighboring interaction between aromatic  $\pi$ -electron systems of the bases, is another factor that affects the stability of base pairing. In 1995, the first synthesis and in vitro evaluation of nonpolar base pairs that pair through shape complementary were presented by Kool's group [106], providing evidence that a hydrogen bond is not absolutely required for the synthesis of DNA. The corresponding hydrophobic base pairs were incorporated into oligonucleotides to form stable self-pairing patterns, raising the possibility of adopting hydrophobic base pairs to expand genetic alphabet. Some hydrophobic UBPs are comprised of heterocycles with totally different structures, for instance, Ds:Px [107], dNaM:d5SICS [108], and dNaM:dTPT3 [109] base pairs (Figure 10). The Ds:Px base pair is a representative hydrophobic base pair developed by Hirao's group. Initially, the research of Hirao's group focused on the development of hydrogen-bonded UBPs. However, they started the investigation on hydrophobic UBPs because the selectivity of their hydrogen-bonded UBPs is insufficient during replication, and small hydrophilic bases are not good substrates for polymerases [110]. It has been illustrated that these UBPs are biologically functional as third base pairs during the replication, transcription, and/or translation process in vitro [100–102].

The application of UBPs in the field of drug therapy and synthetic biology has been investigated. Developed by Benner and coworkers, aptamers containing a third base pair (P:Z, Figure 10) were applied to recognize HepG2 liver cancer cells [104]. Aptamers targeting the human vascular endothelial growth factor (VEGF165) with pM affinity were selected from DNA pools containing Ds nucleotides [111]. dNaM:d5SICS and dNaM:dTPT3 base pairs are successful UBPs created by Romesberg's group that have been used for the creation of semi-synthetic organisms.

In synthetic biology, to create artificial life with UBPs, challenges that have to be overcome include the transport of unnatural triphosphates into cells, faithful replication of UBP, and maintenance of UBP-containing DNAs during cell division. In 2014, Romesberg's group succeeded in constructing semi-synthetic strains of *E. coli*, in which a genome that contains a hydrophobic base pair can be stably propagated [108]. The algal transporter PtNTT2 was found to be an effective transporter for the importing of d5SICSTP and dNaMTP into *E. coli* and pol I polymerase was believed to replicate a plasmid containing UBPs faithfully and efficiently. This UBP is fairly stable against DNA repair pathways, affording the first semi-synthetic organism (SSO) with an expanded genetic alphabet.

Based on the above-mentioned results, further investigations on the in vivo transcription and translation of UBPs were performed [109]. A new base pair dNaM-dTPT3 was prepared and then transcribed into mRNA and tRNA to form unnatural codons and the corresponding unnatural anticodons, respectively (Figure 11). The uptake and incorporation of non-canonical amino acids (ncAAs) were realized to yield a green fluorescent protein, which means the successful creation of SSO with new functions. Besides, DNAs that lost the UBP can be eliminated by CRISPR (clustered regularly interspaced short palindromic repeats)-Cas9 (associated protein 9) technology to further promote the health of synthetic organisms, allowing more attractive prospects in the development of artificial life with UBPs [112]. Lately, it was described that nine unnatural codons were identified with the ability to produce unnatural protein through incorporation of an encoded ncAA. Three of the codons, which are orthogonal and can be decoded, were utilized to construct the first SSO with 67 codons, allowing the development of proteins and new life forms with novel functions [113].



**Figure 11.** Construction of semi-synthetic organisms with UBPs.

### 3. Conclusions and Outlook

A variety of unnatural nucleic acids have been synthesized by replacing natural bases, sugars, and phosphate linkages with artificial structures to investigate their potential as alternative genetic materials. The key chemical and structural parameters for genetic information storage, heredity, and evolution have been elucidated. Engineered polymerases promoted the application of XNAs as alternative genetic materials in vitro. However, it is still challenging to use XNAs instead of DNAs in living cells. In order to know whether this idea is feasible, exploration of xenobiotic life with XNA as genetic material is necessary.

Initial studies on the in vivo implementation of XNAs have been performed, and it is demonstrated that plasmids containing sugar, backbone, or nucleobase-modified XNAs can be read by the cellular machinery to generate natural DNA with correct genetic information. The central dogma is expanded through the insertion of UBPs, resulting in the successful construction of semi-synthetic organisms with an extended genome. Current achievements are encouraging for the integration of unnatural genetic materials into the living system to create synthetic life forms [114–116]. The construction of synthetic life forms is helpful to reveal the fundamental principles of living systems and enlighten our understanding of current life on Earth.

However, as alternative genetic materials, XNAs should be able to replace DNA in living organisms with complete genome substitution. Up to now, only few examples of artificial life whose whole genome was reconstructed is reported, in which the natural base thymine is replaced by 5-position-modified uracil after chemical evolution. As a consequence of this replacement in DNA, significant mutations in the resulting bacterial genome were detected. Although some promising results have been obtained in this research area, the construction of XNA-based synthetic life is still in its infancy, thus more effort is needed to achieve this goal.

**Author Contributions:** Conceptualization and writing—review and editing, P.N. and H.M.; writing—original draft preparations, P.N., Y.B. and H.M.; funding acquisition, H.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Key R&D Program of China, grant number 2019YFA0904001; Guangdong Provincial Key Laboratory of Synthetic Genomics, grant number 2019B030301006; Shenzhen Key Laboratory of Synthetic Genomics, grant number ZDSYS201802061806209; Shenzhen Institute of Synthetic Biology Scientific Research Program, grant number ZTXM20190005 and ZTXM20190020.

**Conflicts of Interest:** The authors declare no conflict of interest.

### References

1. Watson, J.D.; Crick, F.H. Genetical implications of the structure of deoxyribonucleic acid. *Nature* **1953**, *171*, 964–967. [[CrossRef](#)] [[PubMed](#)]
2. Ehrlich, M.; Gama-Sosa, M.A.; Carreira, L.H.; Ljungdahl, L.G.; Kuo, K.C.; Gehrke, C.W. DNA methylation in thermophilic bacteria: N<sup>4</sup>-methylcytosine, 5-methylcytosine, and N<sup>6</sup>-methyladenine. *Nucleic Acids Res.* **1985**, *13*, 1399–1412. [[CrossRef](#)] [[PubMed](#)]
3. Portela, A.; Esteller, M. Epigenetic modifications and human disease. *Nat. Biotechnol.* **2010**, *28*, 1057. [[CrossRef](#)] [[PubMed](#)]

4. Carell, T.; Kurz, M.Q.; Müller, M.; Rossa, M.; Spada, F. Non-canonical bases in the genome: The regulatory information layer in DNA. *Angew. Chem. Int. Ed.* **2018**, *57*, 4296–4312. [[CrossRef](#)]
5. Bilyard, M.K.; Becker, S.; Balasubramanian, S. Natural, modified DNA bases. *Curr. Opin. Chem. Biol.* **2020**, *57*, 1–7. [[CrossRef](#)]
6. Weigele, P.; Raleigh, E.A. Biosynthesis and function of modified bases in bacteria and their viruses. *Chem. Rev.* **2016**, *116*, 12655–12687. [[CrossRef](#)]
7. Jin, S.-G.; Wu, X.; Li, A.X.; Pfeifer, G.P. Genomic mapping of 5-hydroxymethylcytosine in the human brain. *Nucleic Acids Res.* **2011**, *39*, 5015–5024. [[CrossRef](#)]
8. Sánchez-Romero, M.A.; Cota, I.; Casadesús, J. DNA methylation in bacteria: From the methyl group to the methylome. *Curr. Opin. Microbiol.* **2015**, *25*, 9–16. [[CrossRef](#)]
9. Warren, R. Modified bases in bacteriophage DNAs. *Annu. Rev. Microbiol.* **1980**, *34*, 137–158. [[CrossRef](#)]
10. Stewart, C.R.; Casjens, S.R.; Cresawn, S.G.; Houtz, J.M.; Smith, A.L.; Ford, M.E.; Peebles, C.L.; Hatfull, G.F.; Hendrix, R.W.; Huang, W.M. The genome of *Bacillus subtilis* bacteriophage SPO1. *J. Mol. Biol.* **2009**, *388*, 48–70. [[CrossRef](#)]
11. Takahashi, I.; Marmur, J. Replacement of thymidylic acid by deoxyuridylic acid in the deoxyribonucleic acid of a transducing phage for *Bacillus subtilis*. *Nature* **1963**, *197*, 794–795. [[CrossRef](#)] [[PubMed](#)]
12. Price, A.R. Bacteriophage PBS2-induced deoxycytidine triphosphate deaminase in *Bacillus subtilis*. *J. Virol.* **1974**, *14*, 1314–1317. [[CrossRef](#)] [[PubMed](#)]
13. Mehta, A.P.; Li, H.; Reed, S.A.; Supekova, L.; Javahishvili, T.; Schultz, P.G. Replacement of Thymidine by a Modified Base in the *Escherichia coli* Genome. *J. Am. Chem. Soc.* **2016**, *138*, 7272–7275. [[CrossRef](#)] [[PubMed](#)]
14. Mehta, A.P.; Li, H.; Reed, S.A.; Supekova, L.; Javahishvili, T.; Schultz, P.G. Replacement of 2'-Deoxycytidine by 2'-Deoxycytidine Analogues in the *E. coli* Genome. *J. Am. Chem. Soc.* **2016**, *138*, 14230–14233. [[CrossRef](#)]
15. Ghadessy, F.J.; 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. [[CrossRef](#)] [[PubMed](#)]
16. Kunkel, T.A.; Bebenek, K. DNA replication fidelity. *Annu. Rev. Biochem.* **2000**, *69*, 497–529. [[CrossRef](#)]
17. Chen, T.; Romesberg, F.E. Directed polymerase evolution. *FEBS Lett.* **2014**, *588*, 219–229. [[CrossRef](#)]
18. Houlihan, G.; Arangundy-Franklin, S.; Holliger, P. Exploring the chemistry of genetic information storage and propagation through polymerase engineering. *Acc. Chem. Res.* **2017**, *50*, 1079–1087. [[CrossRef](#)]
19. Pinheiro, V.B.; Holliger, P. The XNA world: Progress towards replication and evolution of synthetic genetic polymers. *Curr. Opin. Chem. Biol.* **2012**, *16*, 245–252. [[CrossRef](#)]
20. Ereemeeva, E.; Herdewijn, P. Reprint of: Non Canonical Genetic Material. *Curr. Opin. Biotechnol.* **2019**. [[CrossRef](#)]
21. Pinheiro, V.B.; Taylor, A.I.; Cozens, C.; Abramov, M.; Renders, M.; Zhang, S.; Chaput, J.C.; Wengel, J.; Peak-Chew, S.-Y.; McLaughlin, S.H.; et al. Synthetic genetic polymers capable of heredity and evolution. *Science* **2012**, *336*, 341–344. [[CrossRef](#)] [[PubMed](#)]
22. Tarköy, M.; Leumann, C. Synthesis and Pairing Properties of Decanucleotides from (3' S, 5' R)-2'-Deoxy-3', 5'-ethano-β-D-ribofuranosyladenine and-thymine. *Angew. Chem. Int. Ed.* **1993**, *32*, 1432–1434. [[CrossRef](#)]
23. Koshkin, A.A.; Nielsen, P.; Meldgaard, M.; Rajwanshi, V.K.; Singh, S.K.; Wengel, J. LNA (locked nucleic acid): An RNA mimic forming exceedingly stable LNA: LNA duplexes. *J. Am. Chem. Soc.* **1998**, *120*, 13252–13253. [[CrossRef](#)]
24. Obika, S.; Nanbu, D.; Hari, Y.; Andoh, J.; Morio, K.; Doi, T.; Imanishi, T. Stability and structural features of the duplexes containing nucleoside analogues with a fixed N-type conformation, 2'-O, 4'-C-methyleneribonucleosides. *Tetrahedron Lett.* **1998**, *39*, 5401–5404. [[CrossRef](#)]
25. Augustyns, K.; Van Aerschot, A.; Urbanke, C.; Herdewijn, P. Influence of the Incorporation of 1-(2, 3-Dideoxy-β-D-Erythro-Hexopyranosyl)-Thymine on the Enzymatic Stability and Base-Pairing Properties of Oligodeoxynucleotides. *Bull. Soc. Chim. Belg.* **1992**, *101*, 119–130. [[CrossRef](#)]
26. Eschenmoser, A. Etiology of potentially primordial biomolecular structures: From vitamin B12 to the nucleic acids and an inquiry into the chemistry of life's origin: A retrospective. *Angew. Chem. Int. Ed.* **2011**, *50*, 12412–12472. [[CrossRef](#)]

27. Eschenmoser, A.; Dobler, M. Warum Pentose- und nicht Hexose-Nucleinsäuren?? Teil I. Einleitung und Problemstellung, Konformationsanalyse für Oligonucleotid-Ketten aus 2', 3'-Dideoxyglucopyranosyl-Bausteinen ('Homo-DNS') sowie Betrachtungen zur Konformation von A- und B-DNS. *Helv. Chim. Acta* **1992**, *75*, 218–259. [[CrossRef](#)]
28. Hendrix, C.; Rosemeyer, H.; Verheggen, I.; Van Aerschot, A.; Seela, F.; Herdewijn, P. 1', 5'-Anhydrohexitol oligonucleotides: Synthesis, base pairing and recognition by regular oligodeoxyribonucleotides and oligoribonucleotides. *Chem. Eur. J.* **1997**, *3*, 110–120. [[CrossRef](#)]
29. Schöning, K.-U.; Scholz, P.; Guntha, S.; Wu, X.; Krishnamurthy, R.; Eschenmoser, A. Chemical etiology of nucleic acid structure: The  $\alpha$ -threofuranosyl-(3'  $\rightarrow$  2') oligonucleotide system. *Science* **2000**, *290*, 1347–1351. [[CrossRef](#)]
30. Wang, J.; Verbeure, B.; Luyten, I.; Lescrinier, E.; Froeyen, M.; Hendrix, C.; Rosemeyer, H.; Seela, F.; Van Aerschot, A.; Herdewijn, P. Cyclohexene nucleic acids (CeNA): Serum stable oligonucleotides that activate RNase H and increase duplex stability with complementary RNA. *J. Am. Chem. Soc.* **2000**, *122*, 8595–8602. [[CrossRef](#)]
31. Noronha, A.M.; Wilds, C.J.; Lok, C.-N.; Viazovkina, K.; Arion, D.; Parniak, M.A.; Damha, M.J. Synthesis and Biophysical Properties of Arabinonucleic Acids (ANA): Circular Dichroic Spectra, Melting Temperatures, and Ribonuclease H Susceptibility of ANA:RNA Hybrid Duplexes. *Biochemistry* **2000**, *39*, 7050–7062. [[CrossRef](#)] [[PubMed](#)]
32. Wilds, C.J.; Damha, M.J. 2'-Deoxy-2'-fluoro- $\beta$ -D-arabinonucleosides and oligonucleotides (2' F-ANA): Synthesis and physicochemical studies. *Nucleic Acids Res.* **2000**, *28*, 3625–3635. [[CrossRef](#)]
33. Zhang, L.; Peritz, A.; Meggers, E. A simple glycol nucleic acid. *J. Am. Chem. Soc.* **2005**, *127*, 4174–4175. [[CrossRef](#)] [[PubMed](#)]
34. Singh, S.K.; Koshkin, A.A.; Wengel, J.; Nielsen, P. LNA (locked nucleic acids): Synthesis and high-affinity nucleic acid recognition. *Chem. Commun.* **1998**, 455–456. [[CrossRef](#)]
35. Joyce, G.; Inoue, T.; Orgel, L. Non-enzymatic template-directed synthesis on RNA random copolymers: Poly (C, U) templates. *J. Mol. Biol.* **1984**, *176*, 279–306. [[CrossRef](#)]
36. Adamala, K.; Szostak, J.W. Nonenzymatic template-directed RNA synthesis inside model protocells. *Science* **2013**, *342*, 1098–1100. [[CrossRef](#)]
37. Rose, K.M.; Ferreira-Bravo, I.A.; Li, M.; Craigie, R.; Ditzler, M.A.; Holliger, P.; DeStefano, J.J. Selection of 2'-Deoxy-2'-Fluoroarabino Nucleic Acid (FANA) Aptamers That Bind HIV-1 Integrase with Picomolar Affinity. *ACS Chem. Biol.* **2019**, *14*, 2166–2175. [[CrossRef](#)]
38. Ferreira-Bravo, I.A.; Cozens, C.; Holliger, P.; DeStefano, J.J. Selection of 2'-deoxy-2'-fluoroarabinonucleotide (FANA) aptamers that bind HIV-1 reverse transcriptase with picomolar affinity. *Nucleic Acids Res.* **2015**, *43*, 9587–9599.
39. Larsen, A.C.; Dunn, M.R.; Hatch, A.; Sau, S.P.; Youngbull, C.; Chaput, J.C. A general strategy for expanding polymerase function by droplet microfluidics. *Nat. Commun.* **2016**, *7*, 1–9. [[CrossRef](#)]
40. Dunn, M.R.; Otto, C.; Fenton, K.E.; Chaput, J.C. Improving polymerase activity with unnatural substrates by sampling mutations in homologous protein architectures. *ACS Chem. Biol.* **2016**, *11*, 1210–1219. [[CrossRef](#)]
41. Mei, H.; Liao, J.-Y.; Jimenez, R.M.; Wang, Y.; Bala, S.; McCloskey, C.; Switzer, C.; Chaput, J.C. Synthesis and evolution of a threose nucleic acid aptamer bearing 7-deaza-7-substituted guanosine residues. *J. Am. Chem. Soc.* **2018**, *140*, 5706–5713. [[CrossRef](#)] [[PubMed](#)]
42. Taylor, A.I.; Pinheiro, V.B.; Smola, M.J.; Morgunov, A.S.; Peak-Chew, S.; Cozens, C.; Weeks, K.M.; Herdewijn, P.; Holliger, P. Catalysts from synthetic genetic polymers. *Nature* **2015**, *518*, 427–430. [[CrossRef](#)] [[PubMed](#)]
43. Wang, Y.; Ngor, A.K.; Nikoomezar, A.; Chaput, J.C. Evolution of a general RNA-cleaving FANA enzyme. *Nat. Commun.* **2018**, *9*, 1–10. [[CrossRef](#)] [[PubMed](#)]
44. Yu, H.; Zhang, S.; Chaput, J.C. Darwinian evolution of an alternative genetic system provides support for TNA as an RNA progenitor. *Nat. Chem.* **2012**, *4*, 183–187. [[CrossRef](#)] [[PubMed](#)]
45. Verheggen, I.; Van Aerschot, A.; Van Meervelt, L.; Rozenski, J.; Wiebe, L.; Snoeck, R.; Andrei, G.; Balzarini, J.; Claes, P. Synthesis, biological evaluation, and structure analysis of a series of new 1, 5-anhydrohexitol nucleosides. *J. Med. Chem.* **1995**, *38*, 826–835. [[CrossRef](#)]
46. Pezo, V.; Liu, F.W.; Abramov, M.; Froeyen, M.; Herdewijn, P.; Marlière, P. Binary Genetic Cassettes for Selecting XNA-Templated DNA Synthesis In Vivo. *Angew. Chem. Int. Ed.* **2013**, *52*, 8139–8143. [[CrossRef](#)]

47. Inoue, N.; Minakawa, N.; Matsuda, A. Synthesis and properties of 4'-ThioDNA: Unexpected RNA-like behavior of 4'-ThioDNA. *Nucleic Acids Res.* **2006**, *34*, 3476–3483. [[CrossRef](#)]
48. Inoue, N.; Shionoya, A.; Minakawa, N.; Kawakami, A.; Ogawa, N.; Matsuda, A. Amplification of 4'-ThioDNA in the presence of 4'-Thio-dTTP and 4'-Thio-dCTP, and 4'-ThioDNA-directed transcription in vitro and in mammalian cells. *J. Am. Chem. Soc.* **2007**, *129*, 15424–15425. [[CrossRef](#)]
49. Maruyama, H.; Furukawa, K.; Kamiya, H.; Minakawa, N.; Matsuda, A. Transcription of 4'-thioDNA templates to natural RNA in vitro and in mammalian cells. *Chem. Commun.* **2015**, *51*, 7887–7890. [[CrossRef](#)]
50. Huang, Z.; Schneider, K.C.; Benner, S.A. Building blocks for oligonucleotide analogs with dimethylene sulfide, sulfoxide, and sulfone groups replacing phosphodiester linkages. *J. Org. Chem.* **1991**, *56*, 3869–3882. [[CrossRef](#)]
51. Lackey, D.B.; Patel, J. Biochemical synthesis of chirally pure Rp oligonucleotide phosphorothioates. *Biotechnol. Lett.* **1997**, *19*, 475–478. [[CrossRef](#)]
52. Lelyveld, V.S.; Zhang, W.; Szostak, J.W. Synthesis of phosphoramidate-linked DNA by a modified DNA polymerase. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 7276–7283. [[CrossRef](#)] [[PubMed](#)]
53. Gryaznov, S.; Chen, J.-K. Oligodeoxyribonucleotide N3'. fwdarw. P5' Phosphoramidates: Synthesis and Hybridization Properties. *J. Am. Chem. Soc.* **1994**, *16*, 3143–3144. [[CrossRef](#)]
54. El-Sagheer, A.H.; Sanzone, A.P.; Gao, R.; Tavassoli, A.; Brown, T. Biocompatible artificial DNA linker that is read through by DNA polymerases and is functional in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 11338–11343. [[CrossRef](#)]
55. Rejman, D.; Snášel, J.; Liboska, R.; Točík, Z.; Pačes, O.; Králíková, Š.; Rinnová, M.; Koiš, P.; Rosenberg, I. Oligonucleotides with isopolar phosphonate internucleotide linkage: A new perspective for antisense compounds? *Nucleosides Nucleotides Nucleic Acids* **2001**, *20*, 819–823. [[CrossRef](#)]
56. Westheimer, F.H. Why nature chose phosphates. *Science* **1987**, *235*, 1173–1178. [[CrossRef](#)]
57. Benner, S.A.; Hutter, D. Phosphates, DNA, and the search for nonterrean life: A second generation model for genetic molecules. *Bioorg. Chem.* **2002**, *30*, 62–80. [[CrossRef](#)]
58. Roughton, A.L.; Portmann, S.; Benner, S.A.; Egli, M. Crystal structure of a dimethylene sulfone-linked ribodinucleotide analog. *J. Am. Chem. Soc.* **1995**, *117*, 7249–7250. [[CrossRef](#)]
59. Matsukura, M.; Shinozuka, K.; Zon, G.; Mitsuya, H.; Reitz, M.; Cohen, J.S.; Broder, S. Phosphorothioate analogs of oligodeoxynucleotides: Inhibitors of replication and cytopathic effects of human immunodeficiency virus. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 7706–7710. [[CrossRef](#)]
60. Wang, L.; Jiang, S.; Deng, Z.; Dedon, P.C.; Chen, S. DNA phosphorothioate modification—A new multi-functional epigenetic system in bacteria. *Fems Microbiol. Rev.* **2019**, *43*, 109–122. [[CrossRef](#)]
61. Zhou, X.; He, X.; Liang, J.; Li, A.; Xu, T.; Kieser, T.; Helmann, J.D.; Deng, Z. A novel DNA modification by sulphur. *Mol. Microbiol.* **2005**, *57*, 1428–1438. [[CrossRef](#)] [[PubMed](#)]
62. Xu, T.; Yao, F.; Zhou, X.; Deng, Z.; You, D. A novel host-specific restriction system associated with DNA backbone S-modification in Salmonella. *Nucleic Acids Res.* **2010**, *38*, 7133–7141. [[CrossRef](#)] [[PubMed](#)]
63. Gryaznov, S.M.; Lloyd, D.H.; Chen, J.-K.; Schultz, R.G.; DeDionisio, L.A.; Ratmeyer, L.; Wilson, W.D. Oligonucleotide N3'→ P5' phosphoramidates. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 5798–5802. [[CrossRef](#)] [[PubMed](#)]
64. Zhang, S.; Blain, J.C.; Zielinska, D.; Gryaznov, S.M.; Szostak, J.W. Fast and accurate nonenzymatic copying of an RNA-like synthetic genetic polymer. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 17732–17737. [[CrossRef](#)] [[PubMed](#)]
65. O'Flaherty, D.K.; Zhou, L.; Szostak, J.W. Nonenzymatic template-directed synthesis of mixed-sequence 3'-NP-DNA up to 25 nucleotides long inside model protocells. *J. Am. Chem. Soc.* **2019**, *141*, 10481–10488. [[CrossRef](#)]
66. Lelyveld, V.S.; O'Flaherty, D.K.; Zhou, L.; Izgu, E.C.; Szostak, J.W. DNA polymerase activity on synthetic N3'→ P5' phosphoramidate DNA templates. *Nucleic Acids Res.* **2019**, *47*, 8941–8949. [[CrossRef](#)]
67. Nguyen, H.; Abramov, M.; Ereemeeva, E.; Herdewijn, P. In Vivo Expression of Genetic Information from Phosphoramidate-DNA. *ChemBioChem* **2020**, *21*, 272–278. [[CrossRef](#)]
68. Wolfe, J.L.; Kawate, T.; Belenky, A.; Stanton Jr, V. Synthesis and polymerase incorporation of 5'-amino-2', 5'-dideoxy-5'-N-triphosphate nucleotides. *Nucleic Acids Res.* **2002**, *30*, 3739–3747. [[CrossRef](#)]

69. Renders, M.; Emmerechts, G.; Rozenski, J.; Krecmerová, M.; Holý, A.; Herdewijn, P. Enzymatic synthesis of phosphonomethyl oligonucleotides by terminator polymerase. *Angew. Chem.* **2007**, *119*, 2553–2556. [[CrossRef](#)]
70. Renders, M.; Lievrouw, R.; Krecmerová, M.; Holý, A.; Herdewijn, P. Enzymatic polymerization of phosphonate nucleosides. *ChemBioChem* **2008**, *9*, 2883–2888. [[CrossRef](#)]
71. Kukwikila, M.; Gale, N.; El-Sagheer, A.H.; Brown, T.; Tavassoli, A. Assembly of a biocompatible triazole-linked gene by one-pot click-DNA ligation. *Nat. Chem.* **2017**, *9*, 1089. [[CrossRef](#)] [[PubMed](#)]
72. Birts, C.N.; Sanzone, A.P.; El-Sagheer, A.H.; Blaydes, J.P.; Brown, T.; Tavassoli, A. Transcription of click-linked DNA in human cells. *Angew. Chem. Int. Ed.* **2014**, *53*, 2362–2365. [[CrossRef](#)] [[PubMed](#)]
73. Liu, C.; Cozens, C.; Jaziri, F.; Rozenski, J.; Marechal, A.; Dumbre, S.; Pezo, V.; Marlière, P.; Pinheiro, V.B.; Groaz, E.; et al. Phosphonomethyl oligonucleotides as backbone-modified artificial genetic polymers. *J. Am. Chem. Soc.* **2018**, *140*, 6690–6699. [[CrossRef](#)]
74. Luo, M.; Groaz, E.; Froeyen, M.; Pezo, V.; Jaziri, F.; Leonczak, P.; Schepers, G.; Rozenski, J.; Marlière, P.; Herdewijn, P. Invading *Escherichia coli* Genetics with a Xenobiotic Nucleic Acid Carrying an Acyclic Phosphonate Backbone (ZNA). *J. Am. Chem. Soc.* **2019**, *141*, 10844–10851. [[CrossRef](#)] [[PubMed](#)]
75. Holý, A. Antiviral acyclic nucleoside phosphonates structure activity studies. *Antivir. Res.* **2006**, *71*, 248–253. [[CrossRef](#)] [[PubMed](#)]
76. Wu, T.; Froeyen, M.; Kempeneers, V.; Pannecouque, C.; Wang, J.; Busson, R.; De Clercq, E.; Herdewijn, P. Deoxythreosyl phosphonate nucleosides as selective anti-HIV agents. *J. Am. Chem. Soc.* **2005**, *127*, 5056–5065. [[CrossRef](#)] [[PubMed](#)]
77. Nielsen, P.E.; Egholm, M.; Berg, R.H.; Buchardt, O. Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science* **1991**, *254*, 1497–1500. [[CrossRef](#)]
78. Sharma, C.; Awasthi, S.K. Versatility of peptide nucleic acids (PNA s): Role in chemical biology, drug discovery, and origins of life. *Chem. Biol. Drug Des.* **2017**, *89*, 16–37. [[CrossRef](#)]
79. Nelson, K.E.; Levy, M.; Miller, S.L. Peptide nucleic acids rather than RNA may have been the first genetic molecule. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 3868–3871. [[CrossRef](#)]
80. Böhrer, C.; Nielsen, P.E.; Orgel, L.E. Template switching between PNA and RNA oligonucleotides. *Nature* **1995**, *376*, 578–581. [[CrossRef](#)]
81. Singhal, A.; Nielsen, P.E. Cross-catalytic peptide nucleic acid (PNA) replication based on templated ligation. *Org. Biomol. Chem.* **2014**, *12*, 6901–6907. [[CrossRef](#)] [[PubMed](#)]
82. Rosenbaum, D.M.; Liu, D.R. Efficient and sequence-specific DNA-templated polymerization of peptide nucleic acid aldehydes. *J. Am. Chem. Soc.* **2003**, *125*, 13924–13925. [[CrossRef](#)] [[PubMed](#)]
83. Benner, S.A. Understanding nucleic acids using synthetic chemistry. *Acc. Chem. Res.* **2004**, *37*, 784–797. [[CrossRef](#)] [[PubMed](#)]
84. Cadet, J.; Wagner, J.R. DNA base damage by reactive oxygen species, oxidizing agents, and UV radiation. *Cold Spring Harb. Perspect. Biol.* **2013**, *5*, a012559. [[CrossRef](#)]
85. Kunkel, T.A. DNA replication fidelity. *J. Biol. Chem.* **2004**, *279*, 16895–16898. [[CrossRef](#)]
86. Meek, K.N.; Rangel, A.E.; Heemstra, J.M. Enhancing aptamer function and stability via in vitro selection using modified nucleic acids. *Methods* **2016**, *106*, 29–36. [[CrossRef](#)]
87. Rohloff, J.C.; Gelinis, A.D.; Jarvis, T.C.; Ochsner, U.A.; Schneider, D.J.; Gold, L.; Janjic, N. Nucleic acid ligands with protein-like side chains: Modified aptamers and their use as diagnostic and therapeutic agents. *Mol. Ther. Nucleic Acids* **2014**, *3*, e201. [[CrossRef](#)]
88. Hollenstein, M. Expanding the catalytic repertoire of DNAzymes by modified nucleosides. *J. Chim.* **2011**, *65*, 770–775. [[CrossRef](#)]
89. Sidorov, A.V.; Grasby, J.A.; Williams, D.M. Sequence-specific cleavage of RNA in the absence of divalent metal ions by a DNAzyme incorporating imidazolyl and amino functionalities. *Nucleic Acids Res.* **2004**, *32*, 1591–1601. [[CrossRef](#)]
90. Ereemeeva, E.; Abramov, M.; Margamuljana, L.; Rozenski, J.; Pezo, V.; Marlière, P.; Herdewijn, P. Chemical morphing of DNA containing four noncanonical bases. *Angew. Chem. Int. Ed.* **2016**, *55*, 7515–7519. [[CrossRef](#)] [[PubMed](#)]
91. Ereemeeva, E.; Abramov, M.; Margamuljana, L.; Herdewijn, P. Base-modified nucleic acids as a powerful tool for synthetic biology and biotechnology. *Chem. Eur. J.* **2017**, *23*, 9560–9576. [[CrossRef](#)] [[PubMed](#)]



92. Krueger, A.T.; Lu, H.; Lee, A.H.; Kool, E.T. Synthesis and properties of size-expanded DNAs: Toward designed, functional genetic systems. *Acc. Chem. Res.* **2007**, *40*, 141–150. [[CrossRef](#)] [[PubMed](#)]
93. Krueger, A.T.; Peterson, L.W.; Chelliserry, J.; Kleinbaum, D.J.; Kool, E.T. Encoding phenotype in bacteria with an alternative genetic set. *J. Am. Chem. Soc.* **2011**, *133*, 18447–18451. [[CrossRef](#)] [[PubMed](#)]
94. Chelliserrykattil, J.; Lu, H.; Lee, A.H.; Kool, E.T. Polymerase Amplification, Cloning, and Gene Expression of Benzo-Homologous “yDNA” Base Pairs. *ChemBioChem* **2008**, *9*, 2976–2980. [[CrossRef](#)]
95. Mehta, A.P.; Wang, Y.; Reed, S.A.; Supekova, L.; Javahishvili, T.; Chaput, J.C.; Schultz, P.G. Bacterial genome containing chimeric DNA–RNA sequences. *J. Am. Chem. Soc.* **2018**, *140*, 11464–11473. [[CrossRef](#)]
96. Gavette, J.V.; Stoop, M.; Hud, N.V.; Krishnamurthy, R. RNA–DNA chimeras in the context of an RNA world transition to an RNA/DNA world. *Angew. Chem. Int. Ed.* **2016**, *55*, 13204–13209. [[CrossRef](#)]
97. Hanawalt, P. [95] Preparation of 5-bromouracil-labeled DNA. *Methods Enzym.* **1967**, *12*, 702–708.
98. Bick, M.D.; Davidson, R.L. Total substitution of bromodeoxyuridine for thymidine in the DNA of a bromodeoxyuridine-dependent cell line. *Proc. Natl. Acad. Sci. USA* **1974**, *71*, 2082–2086. [[CrossRef](#)]
99. Marlière, P.; Patrouix, J.; Döring, V.; Herdewijn, P.; Tricot, S.; Cruveiller, S.; Bouzon, M.; Mutzel, R. Chemical evolution of a bacterium’s genome. *Angew. Chem. Int. Ed.* **2011**, *50*, 7109–7114. [[CrossRef](#)]
100. Hamashima, K.; Kimoto, M.; Hirao, I. Creation of unnatural base pairs for genetic alphabet expansion toward synthetic xenobiology. *Curr. Opin. Chem. Biol.* **2018**, *46*, 108–114. [[CrossRef](#)]
101. Dien, V.T.; Morris, S.E.; Karadeema, R.J.; Romesberg, F.E. Expansion of the genetic code via expansion of the genetic alphabet. *Curr. Opin. Chem. Biol.* **2018**, *46*, 196–202. [[CrossRef](#)] [[PubMed](#)]
102. Benner, S.A.; Karalkar, N.B.; Hoshika, S.; Laos, R.; Shaw, R.W.; Matsuura, M.; Fajardo, D.; Moussatche, P. Alternative Watson–Crick synthetic genetic systems. *Cold Spring Harb. Perspect. Biol.* **2016**, *8*, a023770. [[CrossRef](#)]
103. Switzer, C.; Moroney, S.E.; Benner, S.A. Enzymatic incorporation of a new base pair into DNA and RNA. *J. Am. Chem. Soc.* **1989**, *111*, 8322–8323. [[CrossRef](#)]
104. Zhang, L.; Yang, Z.; Sefah, K.; Bradley, K.M.; Hoshika, S.; Kim, M.-J.; Kim, H.-J.; Zhu, G.; Jimenez, E.; Cansiz, S.; et al. Evolution of functional six-nucleotide DNA. *J. Am. Chem. Soc.* **2015**, *137*, 6734–6737. [[CrossRef](#)] [[PubMed](#)]
105. Hoshika, S.; Leal, N.A.; Kim, M.-J.; Kim, M.-S.; Karalkar, N.B.; Kim, H.-J.; Bates, A.M.; Watkins, N.E.; SantaLucia, H.A.; Meyer, A.J.; et al. Hachimoji DNA and RNA: A genetic system with eight building blocks. *Science* **2019**, *363*, 884–887. [[CrossRef](#)] [[PubMed](#)]
106. Schweitzer, B.A.; Kool, E.T. Hydrophobic, non-hydrogen-bonding bases and base pairs in DNA. *J. Am. Chem. Soc.* **1995**, *117*, 1863–1872. [[CrossRef](#)]
107. Okamoto, I.; Miyatake, Y.; Kimoto, M.; Hirao, I. High fidelity, efficiency and functionalization of Ds–Px unnatural base pairs in PCR amplification for a genetic alphabet expansion system. *ACS Synth. Bio.* **2016**, *5*, 1220–1230. [[CrossRef](#)]
108. Malyshev, D.A.; Dhami, K.; Lavergne, T.; Chen, T.; Dai, N.; Foster, J.M.; Corrêa, I.R.; Romesberg, F.E. A semi-synthetic organism with an expanded genetic alphabet. *Nature* **2014**, *509*, 385–388. [[CrossRef](#)]
109. Zhang, Y.; Ptacin, J.L.; Fischer, E.C.; Aerni, H.R.; Caffaro, C.E.; San Jose, K.; Feldman, A.W.; Turner, C.R.; Romesberg, F.E. A semi-synthetic organism that stores and retrieves increased genetic information. *Nature* **2017**, *551*, 644–647. [[CrossRef](#)]
110. Hirao, I.; Kimoto, M. Unnatural base pair systems toward the expansion of the genetic alphabet in the central dogma. *roc. Jpn. Acad. Ser. B* **2012**, *88*, 345–367. [[CrossRef](#)]
111. Kimoto, M.; Yamashige, R.; Matsunaga, K.-I.; Yokoyama, S.; Hirao, I. Generation of high-affinity DNA aptamers using an expanded genetic alphabet. *Nat. Biotechnol.* **2013**, *31*, 453. [[CrossRef](#)] [[PubMed](#)]
112. Zhang, Y.; Lamb, B.M.; Feldman, A.W.; Zhou, A.X.; Lavergne, T.; Li, L.; Romesberg, F.E. A semisynthetic organism engineered for the stable expansion of the genetic alphabet. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 1317–1322. [[CrossRef](#)] [[PubMed](#)]
113. Fischer, E.C.; Hashimoto, K.; Zhang, Y.; Feldman, A.W.; Dien, V.T.; Karadeema, R.J.; Adhikary, R.; Ledbetter, M.P.; Krishnamurthy, R.; Romesberg, F.E. New codons for efficient production of unnatural proteins in a semisynthetic organism. *Nat. Chem. Biol.* **2020**, *16*, 570–576. [[CrossRef](#)] [[PubMed](#)]
114. Budisa, N.; Kubyshev, V.; Schmidt, M. Xenobiology: A Journey towards Parallel Life Forms. *ChemBioChem* **2020**, *21*, 1–5.

115. Schmidt, M.; Pei, L.; Budisa, N. Xenobiology: State-of-the-Art, Ethics, and Philosophy of New-to-Nature Organisms. In *Synthetic Biology–Metabolic Engineering*; Springer: Cham, Switzerland, 2017; pp. 301–315.
116. Schmidt, M. Xenobiology: A new form of life as the ultimate biosafety tool. *Bioessays* **2010**, *32*, 322–331. [[CrossRef](#)] [[PubMed](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).