Available online at www.sciencedirect.com

SciVerse ScienceDirect

Current Opinion in Chemical Biology

The XNA world: progress towards replication and evolution of synthetic genetic polymers

Vitor B Pinheiro and Philipp Holliger

Life's diversity is built on the wide range of properties and functions that can be encoded in natural biopolymers such as polypeptides and nucleic acids. However, despite their versatility, the range of chemical functionalities is limited, particularly in the case of nucleic acids. Chemical modification of nucleic acids can greatly increase their functional diversity but access to the full phenotypic potential of such polymers requires a system of replication. Here we review progress in the chemical and enzymatic synthesis, replication and evolution of unnatural nucleic acid polymers, which promises to enable the exploration of a vast sequence space not accessible to nature and deliver ligands, catalysts and materials based on this new class of biopolymers.

Address
Laboratory of Molecular Biology, Medical Research Council, Hills Road, Cambridge CB2 0QH, UK

Corresponding author: Holliger, Philipp (ph1@mrc-lmb.cam.ac.uk)

Current Opinion in Chemical Biology 2012, **16**:1–8

This review comes from a themed issue on
Synthetic Biology
Edited by Jason W Chin and Lingchong You

1367-5931/$ – see front matter

http://dx.doi.org/10.1016/j.cbpa.2012.05.198

Introduction

DNA and RNA are the repositories of genetic information in biology and appear uniquely suited for information storage and replication [1]. However, modifications of the tripartite chemical structure of nucleic acids, with alternative nucleobases, backbone linkages and ribofuranose congeners, are possible. Some of these chemistries allow duplex formation, information storage and potentially evolution (Figures 1 and 2). In this review we will adopt the term XNA (*xeno*-nucleic acids), first proposed by Herdewijn and Marliere [2] to describe any such synthetic genetic polymer with a focus on those that have shown potential for either chemical and/or enzymatic replication — a prerequisite for evolution.

There have been many recent advances in this field and it is impossible to cover them all in this short review. We apologize to our colleagues whose works have not been included.

Base-modified XNAs

Chemical modifications to the nucleobases can be added to N7 in purines or C5 in pyrimidines [3,4]. They are usually well tolerated as they extend into the major groove and therefore do not give rise to steric clashes, at least at low multiplicity [5–7]. However, full substitution with large hydrophobic fluorophores even at one base [8] or at all bases with a variety of chemical groups (fDNA) can result in significantly altered physico-chemical properties including organic phase partition or a tendency to adopt non-standard helical conformations [9,10]. Such modifications can expand the functionality of nucleic acids as shown for DNAzymes [11] and for aptamer selections [12••]. C5 pyrimidine modifications are also widely used in biology [13,14]. Indeed, a novel chemostat system has allowed the evolution of a bacterium in which dT is completely replaced with the unnatural 5-chlorodeoxyuridine [15••].

Kool and colleagues explored nucleobase expansion by one [16,17] or two benzene units [18], which give rise to large, fluorescent analogues that assemble into expanded double helices. Enzymatic synthesis of xDNA and yDNA has, however, proven difficult [19,20] and may require a significant reshaping of the primer-template binding 'funnel' to accommodate the expanded xDNA or yDNA helices. Nevertheless, plasmids containing up to four consecutive xDNA insertions could be correctly replicated *in vivo* [21]. Such *in vivo* replication bypass of modified nucleic acids has also been reported for some sugar [22] and backbone modifications [23•] (see below).

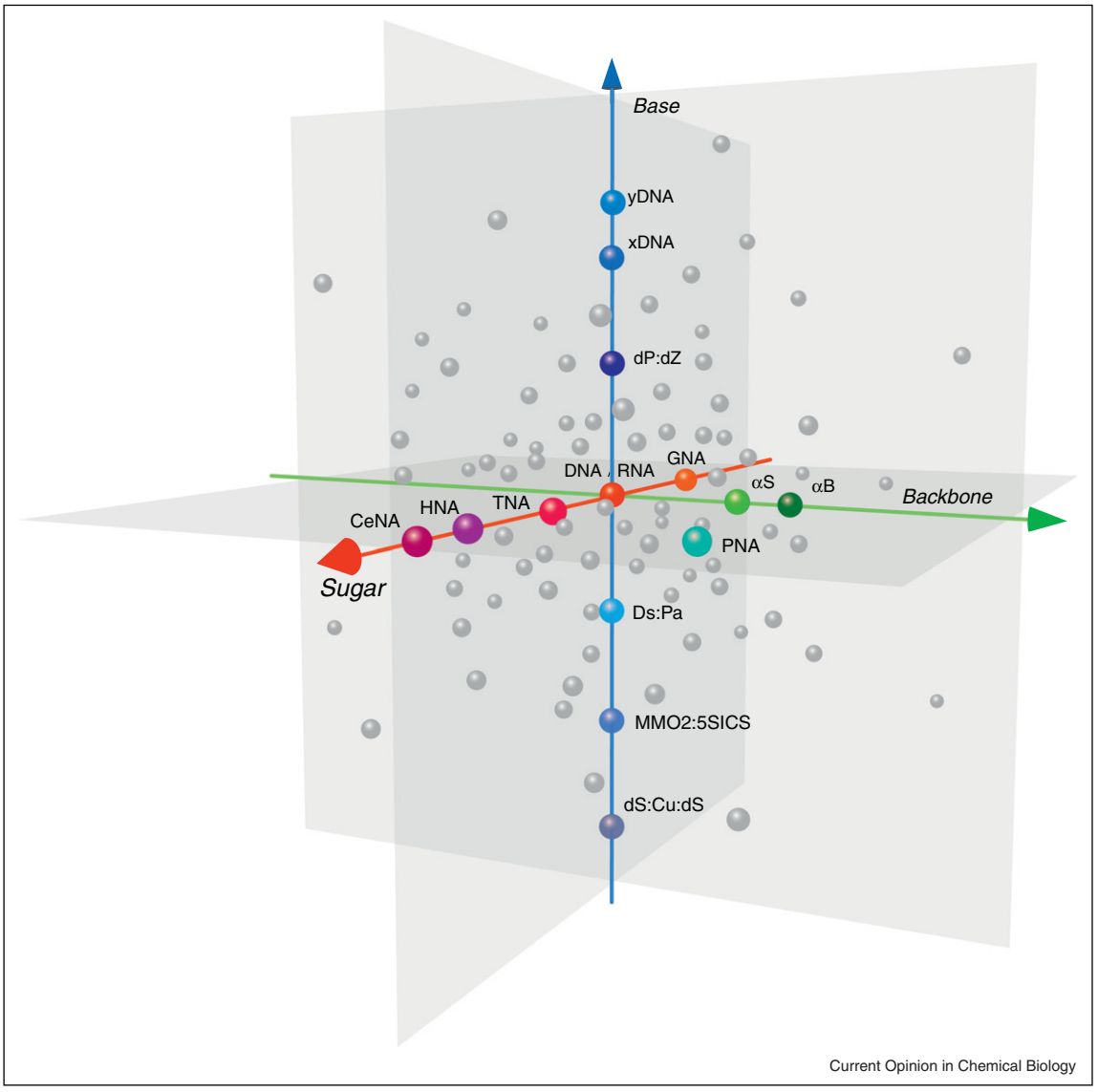
Modifications of nucleobase chemistry can give rise to novel base-pairs and expand the genetic alphabet [24]. Significant advances in this area include highly functional novel base-pairs based on reshuffling hydrogen-bond donor and acceptor patterns [25,26], an increased number of donor–acceptor pairs [27], hydrophobic base-pairs based on steric complementarity [28,29] or even metal ion complexation [30,31,32]. Such novel base-pairs are of great interest as they may allow the site-specific labelling of DNA and RNA.

Sugar-modified and Backbone-modified XNAs

Novel base-pairs expand the genetic alphabet and increase the potential density of stored genetic information but in doing so, lose the ability to communicate with extant biology. XNAs, in which sugar moieties and/or the phosphodiester backbone linkages are modified,

Please cite this article in press as: Pinheiro VB, Holliger P. The XNA world: progress towards replication and evolution of synthetic genetic polymers, Curr Opin Chem Biol (2012), http://dx.doi.org/10.1016/j.cbpa.2012.05.198

Figure 1



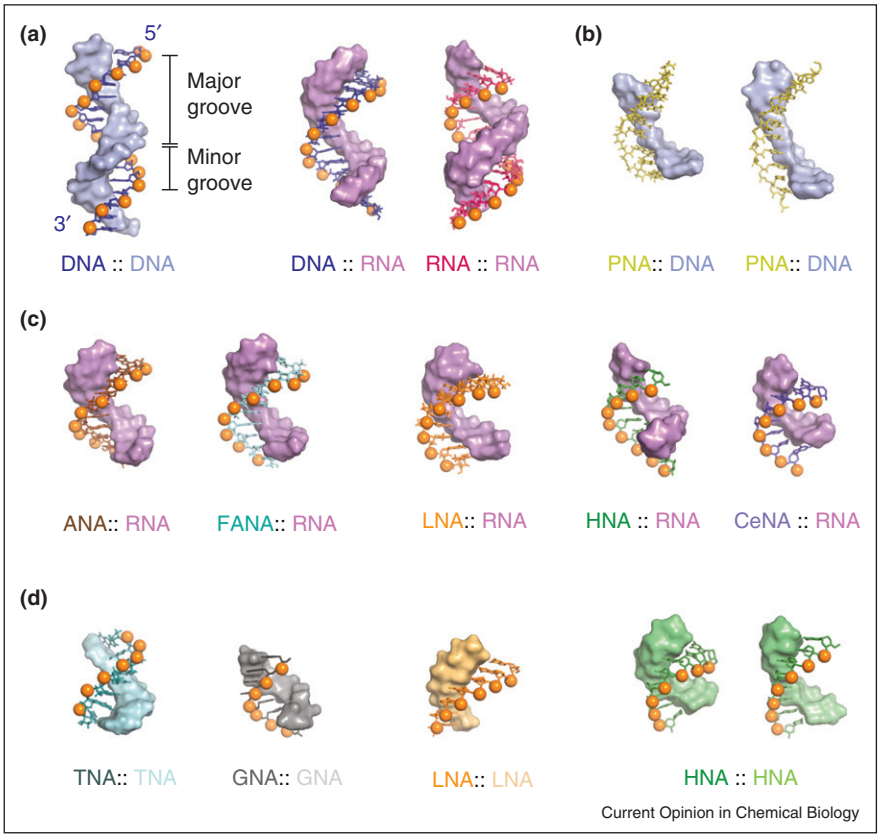
XNA chemical space. The *x*-, *y*- and *z*-axes representing sugar, base and backbone modifications, respectively. Most currently described XNAs (except PNA) lie on the axes, comprising only modifications to one of the three 'variables'. More divergent phenotypes should become accessible through a fuller exploration of the XNA space, that is, the replication and evolution of XNAs comprising a combination of modifications to base, sugar and backbone. Grey spheres represent these possible variants with multiple modifications not yet described in the XNA space.

can retain this key capacity. However, even minor chemical changes, such as the 2′-OH group in RNA, can greatly affect that interaction, resulting in a wide variety of helical conformations (Figure 2), duplex stabilities and altered base-pairing preferences (e.g. G·U wobble pairs).

Conservative variations entail modification of the 2′ position of ribofuranose to chemical moieties such as 2′O-methyl (2′Ome), 2′-azido (2′N_3) or 2′-fluoro (2′F). While incorporation of these modifications by DNA polymerases is generally poor, variants of T7 RNA polymerase can tolerate 2′ substitutions on one or several nucleobases [33–35]. This has enabled aptamer selection with 2′Ome at near full substitution [36] and with 2′F modified pyrimidines. Sullenger and colleagues used the latter substitution, which also greatly enhances the biostability of resulting nucleic acid polymers, to select aptamers directly *in vivo* isolating tumour homing aptamers [37•]. Recent work from our lab has enabled the synthesis and replication of fully substituted 2′F and 2′N_3 polymers (see below) [38•]. Marx and colleagues have also reported incorporation of 2′Se-methyl triphosphates into RNA for MAD (multi-wavelength anomalous dispersion) phasing in crystallography [39].

Please cite this article in press as: Pinheiro VB, Holliger P. The XNA world: progress towards replication and evolution of synthetic genetic polymers, Curr Opin Chem Biol (2012), http://dx.doi.org/10.1016/j.cbpa.2012.05.198

Figure 2



XNA structures. **(a)** Natural duplexes DNA::DNA (PDB ID: 3BSE), RNA::DNA (1EFS) and RNA::RNA (3ND4). DNA is blue, RNA is magenta or lilac. To better illustrate helical parameters, structures are shown with one strand in surface and the other in stick representation, with backbone phosphates shown as orange spheres. **(b)** XNA::DNA hybrid structures. NMR structure of aminoethylglycine PNA::DNA (1PDT; left) and crystal structure of D-lysine-based PNA::DNA (1NR8; right). PNA is yellow. **(c)** XNA::RNA hybrid structures (RNA: lilac). ANA (2KP3): brown, FANA (2KP4): teal, LNA (1H0Q): orange, HNA (2BJ6): green and CeNA (3KNC): light blue. **(D)** XNA::XNA structures. TNA [99]: cyan, GNA (2XC6): grey, LNA (2X2Q): tan and two crystal structures of HNA (481D and 1D7Z): green. Duplexes vary in length from eight (e.g. GNA::GNA) to 16 bases (e.g. DNA::DNA).

Systematic investigation of the chemical neighbourhood of the natural ribofuranose sugars by Eschenmoser and colleagues has shown that only a limited set of congeners retains the capacity to cross-hybridize with DNA and/or RNA. An interesting case is the tetrose-based TNA (α-L-threofuranosyl nucleic acids), which is capable not only of self-pairing but also of cross-hybridization with DNA and RNA [40,41]. Up to 80 consecutive TNA triphosphates can be incorporated by 'Therminator^{TM}' DNA polymerase (9°N*exo*$^{-}$: A485L) [42]. Another congener, 3′-deoxyapionucleic acids (apioNA), with the natural C3 and C4 substituents from DNA displaced to C3 and C2, has been recently reported [43]. Triphosphates are well incorporated by Therminator^{TM} DNA polymerase but the potential of apioDNA for reverse transcription and functional selection remains unexplored [43].

Of great biotechnological importance is LNA (locked nucleic acids; 2′-O,4′-C-methylene-β-D-ribonucleic acids) in which the ribofuranose ring is 'locked' in a single conformation by a methylene bridge between the 2′-O and C4, leading to considerable duplex stabilization [44]. Chimeric polymers containing low-density substitution (or spiking) of LNA bases in a natural DNA context have been used to great effect on a variety of applications, including PCR, aptamers and aptazymes [45]. However, at higher multiplicity, LNA triphosphates generally are poor polymerase substrates [45,46].

A range of other interesting analogues in which ribofuranose is replaced by six-membered ring structures, have also been investigated [47]. Among these, of great interest are HNA (hexitol nucleic acids) [48] and CeNA (cyclohexenyl nucleic acids) [49]. Both HNA and CeNA retain the ability to pair with DNA and RNA [50,51], they are non-toxic to cells [22] but are poor substrates for incorporation [52,53] and as templates for reverse transcription [52,54].

Please cite this article in press as: Pinheiro VB, Holliger P. The XNA world: progress towards replication and evolution of synthetic genetic polymers, Curr Opin Chem Biol (2012), http://dx.doi.org/10.1016/j.cbpa.2012.05.198

In a more radical departure from the natural structures, a wide range of acyclic moieties to link phosphate to base have been explored, notably GNA (glycerol nucleic acids) [55] and FNA (flexible nucleic acids) [56]. Surprisingly, despite poor (or no) cross-hybridization to DNA, FNA triphosphates and S-GNA triphosphates could be incorporated for example by Therminator^{TM} DNA polymerase [57,58]. GNA could serve as template for DNA synthesis by Bst DNA polymerase [59]. Interestingly, GNA can self-pair quite stably [55], so like dXyNA (xylose nucleic acids) [60], might allow the development of genetic systems orthogonal to DNA or RNA.

Simple modifications to the phosphodiester chemistry are of great importance in a number of applications. Among these are the widely used phosphorothioates [61] and boranophosphates [62,63], in which the alpha oxygen is replaced with a sulphur or a borano group. Both are good substrates for DNA polymerases allowing aptamers selection at single base substitution [63,64] and PCR amplification at full substitution [65].

A much more drastic modification to backbone chemistry is implemented in peptide nucleic acids (PNA) in which nucleobases are displayed on an aminoethylglycine backbone. The resulting polymer is charge neutral and capable of strong canonical base-pairing, to itself and to natural nucleic acids, with a wide range of important applications in biotechnology [6,66]. The backbone itself can be modified to alter the polymers' physico-chemical properties, for example, D-lysine [67], arginine [67,68] and even cysteine [69]. While they remain challenging substrates for enzymatic incorporation, the non-enzymatic template-dependent synthesis of PNA has been realized [70].

Azide-alkyne Huisgen cycloaddition or 'click' chemistry has also been exploited for the synthesis of new backbone linkages [23•,71,72]. Nakamura and colleagues generated all triazole-linked DNA (^{TL}DNA) [71], while Brown and colleagues have focused on modifications that are traversed by natural polymerases as template and that are tolerated *in vivo* [23•] with manifold applications, for example in gene assembly.

Non-enzymatic synthesis and replication of synthetic polymers

Purely chemical synthesis of XNAs on a DNA template is a potentially powerful approach as a wide variety of chemical structures should in principle be accessible through this strategy. Following on from the pioneering work of Orgel [73,74], Szostak and colleagues used 2′-amino-phosphoroimidazolides to synthesize a N2′-P5′ phosphoroamidate DNA. Short mixed sequences could be synthesized quickly but at relatively modest fidelities [75].

Reductive amination, pioneered by Lynn and colleagues [76], was adapted by Liu and Rosenbaum [70] and enabled efficient polymerization of PNA pentamers on a DNA template as well as model selections [77••]. A remarkable modification introduced by Ghadiri and colleagues is the thioester PNA (tPNA). In tPNA, cysteines are introduced into the peptide repeating unit to present thiol groups to which nucleobase thioesters attach reversibly [69], providing a dynamic adaptability of the tPNA sequence as determined by complementary tPNA, DNA or RNA templates.

Chemical XNA reverse transcription has not been systematically explored, but pioneering experiments by Orgel using HNA [78] and AtNA (altritol nucleic acids) [79] templates suggest that templates preorganized into A-form promote more efficient RNA synthesis by promoting a sugar pucker switch to the C3-endo conformation, which is optimal for efficient copying [80].

Towards the enzymatic synthesis and replication of synthetic polymers

Another potentially powerful strategy for XNA synthesis (and reverse transcription) is the use of engineered polymerases. The challenge here is to overcome the stringent substrate specificity of polymerase enzymes, while maintaining activity and fidelity. As discussed above, some commercially available polymerases, notably Therminator^{TM}, display a capacity for the synthesis of some XNAs such as TNA [42]. But most analogues are poor polymerase substrates at full substitution, both as nucleotides for polymer synthesis and as templates for reverse transcription. Nevertheless, substantial progress has been made recently (summarized in Table 1).

To extend beyond this narrow range of 'permissive' substrates, polymerase function has been engineered by design, screening and directed evolution strategies [81,82]. Despite several successes (e.g. all αS-DNA [65], fDNA [9], 2′Ome-DNA [36]), engineering polymerases for the synthesis of nucleic acid polymers in which all four canonical nucleotides are replaced by unnatural analogues has remained challenging. The difficulties are illustrated by the paradigmatic case of engineering RNA polymerase activity from a DNA polymerase scaffold.

Rational design [83], screening [84,85] and directed evolution by phage display [86] or compartmentalized self-replication (CSR) [87] have all been used with some success to endow DNA polymerases with basal RNA polymerase activity, allowing NTP incorporation with catalytic efficiencies (k_{cat}/K_m) approaching those of dNTP incorporation. However, none of these polymerases is able to efficiently synthesize longer RNA polymers [84,88,89], suggesting that a second determinant of polymerase substrate specificity remained to be discovered that could allow long RNA synthesis.

Please cite this article in press as: Pinheiro VB, Holliger P. The XNA world: progress towards replication and evolution of synthetic genetic polymers, Curr Opin Chem Biol (2012), http://dx.doi.org/10.1016/j.cbpa.2012.05.198

Table 1

Enzymatic XNA synthesis, reverse transcription and replication.

XNA	Synthesis (DNA → XNA)	Reverse transcription (XNA → DNA)	Replication (XNA → XNA)	Refs.
αS	TaqM1	TaqM1	TaqM1	[65]
CyDNA[e]	Pfu E10	Pfu E10	Pfu E10	[8]
fDNA	Pwo[a]	Pwo[a]	Pwo[a]	[9,10]
RNA	T7RP/TGK/M1-3	RTs	RdRPs	[38•,85]
2′F[e]	T7RP: Y639F, H784A[b]/TGK	RT521	n.d.	[34]
2′Ome[e]	T7RP: Y639F	Thermoscript™ [c]	n.d.	[36]
2′Seme[e]	T7RP: Y639V, H784G, E593G, V685A	n.d.	n.d.	[33,39]
$2'N_3$	TGK	RT521	n.d.	[38•]
ANA	D4K	RT521(K)	n.d.	[90••]
FANA	D4K	RT521(K)	D4K/RT521K	[90••]
HNA	6G12	RT521(K)	RT521K[f]	[90••]
CeNA	6G12/C7	RT521(K)	6G12/RT521K[f]	[90••]
TNA	Therminator™ [d]/RT521	RT521	n.d.	[42,90••]
LNA	C7	RT521K	n.d.	[90••]

Commercially available polymerases: [a](Roche), [b]DuraScribe® (Epicentre), [c](Life technologies), [d](New England Biolabs). Only >50 nt incorporations and full substitution considered except [e](single (CyDNA, SeMe), double (2′F) or nearly full (2′Ome) nucleotide substitution) and [f](<6 nts).

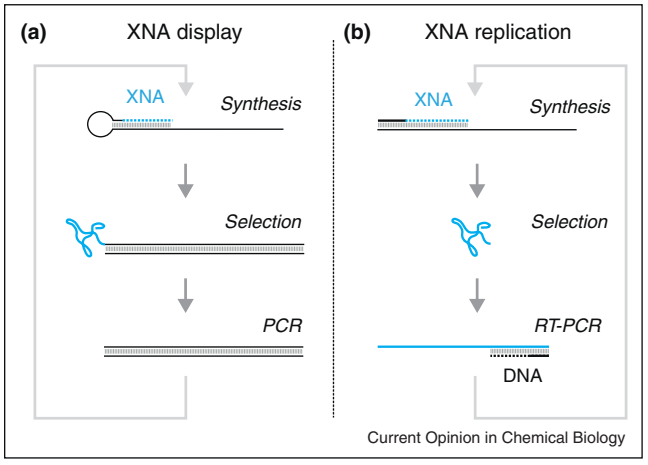
Recently, our lab discovered a plausible candidate motif for such a determinant. Working with Tgo, the replicative DNA polymerase of the archaeon *Thermococcus gorgonarious*, we identified a motif in the thumb subdomain, and within it a single residue mutation, capable of relieving the synthetic block and enabling processive synthesis of protein-coding RNAs of over 1.7 kb in length primed from either DNA or RNA primers [38•]. This motif in the thumb subdomain was discovered using a novel polymerase selection system (CST) and appears to be a generally enabling feature for XNA synthesis [90••]. Using CST selection we readily isolated a number of variants of Tgo DNA polymerase with an ability to processively synthesize a range of XNAs including HNA, CeNA, LNA, TNA, ANA (arabinonucleic acids) and FANA (2′-fluoro-arabinonucleic acid).

However, XNA synthesis (although potentially useful for the bulk production of single stranded XNA oligomers) is a 'dead-end' without a decoding mechanism since genetic information remains locked in XNA, precluding both analysis and evolution.

An elegant approach to circumvent the necessity for an XNA reverse transcriptase (RT) has been the use of a 'display' system, in which the original DNA genotype remains intact and bound to the functional XNA under selection (Figure 3A). This nucleic acid display methodology has been demonstrated for both TNA-based and PNA-based systems as a viable strategy to isolate XNA functional molecules [42,77••,91,92••]. Potential drawbacks of the system include the presence of a potentially interfering DNA tag during aptamer selection and the fact that systematic errors of XNA synthesis from the DNA template (e.g. such as slippage at homopolymeric runs) remain undetected as the XNA polymer is never analyzed by itself.

To enable direct XNA analysis and evolution, it is therefore desirable to develop a capacity for XNA reverse transcription. A number of natural reverse transcriptases (RNA-dependent DNA polymerases) are tolerant of some 2′-modifications enabling information recovery and evolution, for example, in 2′Ome-DNA [36]. However, for the majority of XNAs, no natural RT activity is available. As RNA reverse transcriptase (RT) activity has been engineered in DNA polymerase scaffolds by screening and rational design [93–95], it should, by analogy, be possible to engineer XNA RTs in the same way. Our lab has used Statistical Coupling Analysis (SCA) [96,97], a bioinformatics method that scores phylogenetic covariation, and rational design to engineer an XNA RT *de novo*. Starting from a mutation known to confer some RT activity to a DNA polymerase, we reasoned that any allosteric network identified by SCA that could be involved in template recognition would have to lie in

Figure 3



XNA selection schemes. **(a)** XNA display [91] as used in the selection of three-base anti-Thrombin TNA aptamers [92••] and PNA model selection [77••]. **(b)** XNA replication as used in the selection of anti-TAR and anti-HEL HNA aptamers [90••].

Please cite this article in press as: Pinheiro VB, Holliger P. The XNA world: progress towards replication and evolution of synthetic genetic polymers, Curr Opin Chem Biol (2012), http://dx.doi.org/10.1016/j.cbpa.2012.05.198

its vicinity. Investigating the SCA hits in the neighbouring residues, we discovered a single residue proximal to the conserved C-motif in the polymerase active site, mutation of which conferred broad XNA RT activity on at least six different XNAs (HNA, CeNA, LNA, TNA, ANA, FANA) as well as RNA, 2′F-DNA, 2′N_3-DNA and 2′-Ome-DNA. Together with XNA polymerases, this allowed both synthesis and reverse transcription (Figure 3B) and hence enabled evolution of specific anti-HIV TAR RNA and anti-hen egg lysozyme aptamers in the HNA system [90••].

Conclusion

Significant progress has been made in developing strategies for the replication of XNAs. We anticipate that in future both chemical and enzymatic methodologies will enable the synthesis, replication and evolution of XNA polymers of increasingly divergent chemistry and information content. This will give rise to a new field of 'synthetic genetics' [98], that is, the exploration of the informational, structural and catalytic potential of genetic polymers.

Acknowledgement

Our work is supported by the Medical Research Council: MRC U105178804.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Benner SA: **Understanding nucleic acids using synthetic chemistry**. *Acc Chem Res* 2004, **37**:784-797.

2. Herdewijn P, Marliere P: **Toward safe genetically modified organisms through the chemical diversification of nucleic acids**. *Chem Biodivers* 2009, **6**:791-808.

3. Hasegawa T, Shoji A, Kuwahara M, Ozaki H, Sawai H: **Synthesis and property of DNA labeled with fluorescent acridone**. *Nucleic Acids Symp Ser (Oxf)* 2006, **50**:145-146.

4. Kuwahara M, Suto Y, Minezaki S, Kitagata R, Nagashima J, Sawai H: **Substrate property and incorporation accuracy of various dATP analogs during enzymatic polymerization using thermostable DNA polymerases**. *Nucleic Acids Symp Ser (Oxf)* 2006, **50**:31-32.

5. Weisbrod SH, Marx A: **Novel strategies for the site-specific covalent labelling of nucleic acids**. *Chem Commun* 2008, **44**:5675-5685.

6. Brudno Y, Liu DR: **Recent progress toward the templated synthesis and directed evolution of sequence-defined synthetic polymers**. *Chem Biol* 2009, **16**:265-276.

7. Hocek M, Fojta M: **Nucleobase modification as redox DNA labelling for electrochemical detection**. *Chem Soc Rev* 2011, **40**:5802-5814.

8. Ramsay N, Jemth AS, Brown A, Crampton N, Dear P, Holliger P: **CyDNA: synthesis and replication of highly Cy-dye substituted DNA by an evolved polymerase**. *J Am Chem Soc* 2010, **132**:5096-5104.

9. Jager S, Famulok M: **Generation and enzymatic amplification of high-density functionalized DNA double strands**. *Angew Chem Int Ed Engl* 2004, **43**:3337-3340.

10. Jager S, Rasched G, Kornreich-Leshem H, Engeser M, Thum O, Famulok M: **A versatile toolbox for variable DNA functionalization at high density**. *J Am Chem Soc* 2005, **127**:15071-15082.

11. Hollenstein M, Hipolito CJ, Lam CH, Perrin DM: **A self-cleaving DNA enzyme modified with amines, guanidines and imidazoles operates independently of divalent metal cations (M2+)**. *Nucleic Acids Res* 2009, **37**:1638-1649.

12. •• Vaught JD, Bock C, Carter J, Fitzwater T, Otis M, Schneider D, Rolando J, Waugh S, Wilcox SK, Eaton BE: **Expanding the chemistry of DNA for in vitro selection**. *J Am Chem Soc* 2010, **132**:4141-4151.
The authors report on the isolation of aptamers against previously refractory targets such as TNFRSF9 using a blend of dNTPs in which dTTP was 100% replaced by deoxyuridine analogues C5-modified with various aliphatic 'side-chains'. This is a compelling demonstration of the benefits of expanded chemical diversity in aptamer selections.

13. Gommers-Ampt JH, Borst P: **Hypermodified bases in DNA**. *FASEB J* 1995, **9**:1034-1042.

14. Ku CS, Naidoo N, Wu M, Soong R: **Studying the epigenome using next generation sequencing**. *J Med Genet* 2011, **48**:721-730.

15. •• Marliere P, Patrouix J, Doring V, Herdewijn P, Tricot S, Cruveiller S, Bouzon M, Mutzel R: **Chemical evolution of a bacterium's genome**. *Angew Chem Int Ed Engl* 2011, **50**:7109-7114.
The authors used *in vivo* continuous evolution to completely replace the natural thymine in DNA with 5-chlorouridine in a bacterial genome. This is the first report of a wholesale replacement of one of the canonical bases *in vivo*. Future analysis of the genome of the new 5-Cl-dU dependent organism promises to illuminate the adaptive pathways involved and may aid the establishment of *in vivo* XNA replication.

16. 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.

17. 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.

18. Lee AH, Kool ET: **Exploring the limits of DNA size: naphtho-homologated DNA bases and pairs**. *J Am Chem Soc* 2006, **128**:9219-9230.

19. Chelliserrykattil J, Lu H, Lee AH, Kool ET: **Polymerase amplification, cloning, and gene expression of benzo-homologous ''yDNA'' base pairs**. *Chembiochem* 2008, **9**:2976-2980.

20. Lu H, Lynch SR, Lee AH, Kool ET: **Structure and replication of yDNA: a novel genetic set widened by benzo-homologation**. *Chembiochem* 2009, **10**:2530-2538.

21. Krueger AT, Peterson LW, Chelliserry J, Kleinbaum DJ, Kool ET: **Encoding phenotype in bacteria with an alternative genetic set**. *J Am Chem Soc* 2011, **133**:18447-18451.

22. Pochet S, Kaminski PA, Van Aerschot A, Herdewijn P, Marliere P: **Replication of hexitol oligonucleotides as a prelude to the propagation of a third type of nucleic acid in vivo**. *C R Biol* 2003, **326**:1175-1184.

23. • El-Sagheer AH, Sanzone AP, 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 U S A* 2011, **108**:11338-11343.
The authors characterize two modified nucleotides that can undergo Cu-catalyzed Huisgen cycloaddition ('click' chemistry) to generate a modified inter-nucleotide linkage that is well-tolerated by natural polymerases. There are many potential applications for example in gene synthesis and gene shuffling approaches.

24. Hirao I, Kimoto M, Yamashige R: **Natural versus artificial creation of base pairs in DNA: origin of nucleobases from the perspectives of unnatural base pair studies**. *Acc Chem Res* 2012 http://dx.doi.org/10.1021/ar200257x. Epub ahead of print.

25. Hoshika S, Chen F, Leal NA, Benner SA: **Artificial genetic systems: self-avoiding DNA in PCR and multiplexed PCR**. *Angew Chem Int Ed Engl* 2008, **49**:5554-5557.

Please cite this article in press as: Pinheiro VB, Holliger P. The XNA world: progress towards replication and evolution of synthetic genetic polymers, Curr Opin Chem Biol (2012), http://dx.doi.org/10.1016/j.cbpa.2012.05.198

26. Yang Z, Hutter D, Sheng P, Sismour AM, Benner SA: **Artificially expanded genetic information system: a new base pair with an alternative hydrogen bonding pattern**. *Nucleic Acids Res* 2006, **34**:6095-6101.

27. Ogata S, Takahashi M, Minakawa N, Matsuda A: **Unnatural imidazopyridopyrimidine:naphthyridine base pairs: selective incorporation and extension reaction by Deep Vent (exo-) DNA polymerase**. *Nucleic Acids Res* 2009, **37**:5602-5609.

28. Kimoto M, Kawai R, Mitsui T, Yokoyama S, Hirao I: **An unnatural base pair system for efficient PCR amplification and functionalization of DNA molecules**. *Nucleic Acids Res* 2009, **37**:e14.

29. Seo YJ, Malyshev DA, Lavergne T, Ordoukhanian P, Romesberg FE: **Site-specific labeling of DNA and RNA using an efficiently replicated and transcribed class of unnatural base pairs**. *J Am Chem Soc* 2011, **133**:19878-19888.

30. Kaul C, Muller M, Wagner M, Schneider S, Carell T: **Reversible bond formation enables the replication and amplification of a crosslinking salen complex as an orthogonal base pair**. *Nat Chem* 2011, **3**:794-800.

31. Chiba J, Inouye M: **Exotic DNAs made of nonnatural bases and natural phosphodiester bonds**. *Chem Biodivers* 2010, **7**:259-282.

32. Clever GH, Shionoya M: **Alternative DNA base pairing through metal coordination**. *Met Ions Life Sci* 2012, **10**:269-294.

33. Chelliserrykattil J, Ellington AD: **Evolution of a T7 RNA polymerase variant that transcribes 2′-O-methyl RNA**. *Nat Biotechnol* 2004, **22**:1155-1160.

34. Padilla R, Sousa R: **Efficient synthesis of nucleic acids heavily modified with non-canonical ribose 2′-groups using a mutantT7 RNA polymerase (RNAP)**. *Nucleic Acids Res* 1999, **27**:1561-1563.

35. Padilla R, Sousa R: **A Y639F/H784A T7 RNA polymerase double mutant displays superior properties for synthesizing RNAs with non-canonical NTPs**. *Nucleic Acids Res* 2002, **30**:e138.

36. Burmeister PE, Lewis SD, Silva RF, Preiss JR, Horwitz LR, Pendergrast PS, McCauley TG, Kurz JC, Epstein DM, Wilson C *et al.*: **Direct in vitro selection of a 2′-O-methyl aptamer to VEGF**. *Chem Biol* 2005, **12**:25-33.

37. • Mi J, Liu Y, Rabbani ZN, Yang Z, Urban JH, Sullenger BA, Clary BM: **In vivo selection of tumor-targeting RNA motifs**. *Nat Chem Biol* 2010, **6**:22-24.
The authors report the first *in vivo* selection of aptamers in mice using 2′F pyrimidine modified RNA. The authors isolate an aptamer that localizes to hepatic tumor cells and display strong affinity (~14 nM) against an RNA helicase tumor marker (p68). Such *in vivo* selections promise the generation of tumor-homing aptamers optimally adapted to the physiological environment.

38. • Cozens C, Pinheiro VB, Vaisman A, Woodgate R, Holliger P: **A short adaptive path from DNA to RNA polymerases**. *Proc Natl Acad Sci U S A* 2012, **109**:8067-8072 http://dx.doi.org/10.1073/pnas.1120964109.
The authors report the discovery of a substrate specificity checkpoint in the thumb subdomain of a DNA polymerase. In conjunction with the previously reported steric gate residue, mutation of this checkpoint enables the synthesis of long (>1.5 kb) RNAs from a wide range of primers enabling a free choice of 5′-UTR and 5′-UTR chemistry.

39. Siegmund V, Santner T, Micura R, Marx A: **Enzymatic synthesis of 2′-methylseleno-modified RNA**. *Chem Sci* 2011, **2**:2224-2231.

40. Eschenmoser A: **Chemical etiology of nucleic acid structure**. *Science* 1999, **284**:2118-2124.

41. Schoning K, Scholz P, Guntha S, Wu X, Krishnamurthy R, Eschenmoser A: **Chemical etiology of nucleic acid structure: the alpha-threofuranosyl-(3′ → 2′) oligonucleotide system**. *Science* 2000, **290**:1347-1351.

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

43. Kataoka M, Kouda Y, Sato K, Minakawa N, Matsuda A: **Highly efficient enzymatic synthesis of 3′-deoxyapionucleic acid (apioNA) having the four natural nucleobases**. *Chem Commun (Camb)* 2011, **47**:8700-8702.

44. Singh SK, Koshkin AA, Wengel J, Nielsen P: **LNA (locked nucleic acids): synthesis and high-affinity nucleic acid recognition**. *Chem Commun* 1998, **34**:455-456.

45. Veedu RN, Wengel J: **Locked nucleic acids: promising nucleic acid analogs for therapeutic applications**. *Chem Biodivers* 2010, **7**:536-542.

46. Veedu RN, Vester B, Wengel J: **Enzymatic incorporation of LNA nucleotides into DNA strands**. *Chembiochem* 2007, **8**:490-492.

47. Herdewijn P: **Nucleic acids with a six-membered 'carbohydrate' mimic in the backbone**. *Chem Biodivers* 2010, **7**:1-59.

48. Verheggen I, Van Aerschot A, Toppet S, Snoeck R, Janssen G, Balzarini J, De Clercq E, Herdewijn P: **Synthesis and antiherpes virus activity of 1,5-anhydrohexitol nucleosides**. *J Med Chem* 1993, **36**:2033-2040.

49. Herdewijn P, De Clercq E: **The cyclohexene ring as bioisostere of a furanose ring: synthesis and antiviral activity of cyclohexenyl nucleosides**. *Bioorg Med Chem Lett* 2001, **11**:1591-1597.

50. Wang J, Verbeure B, Luyten I, Froeyen M, Hendrix C, Rosemeyer H, Seela F, Van Aerschot A, Herdewijn P: **Cyclohexene nucleic acids (CeNA) form stable duplexes with RNA and induce RNase H activity**. *Nucleosides Nucleotides Nucleic Acids* 2001, **20**:785-788.

51. Boudou V, Kerremans L, De Bouvere B, Lescrinier E, Schepers G, Busson R, Van Aerschot A, Herdewijn P: **Base pairing of anhydrohexitol nucleosides with 2,6-diaminopurine, 5-methylcytosine and uracil as base moiety**. *Nucleic Acids Res* 1999, **27**:1450-1456.

52. Kempeneers V, Renders M, Froeyen M, Herdewijn P: **Investigation of the DNA-dependent cyclohexenyl nucleic acid polymerization and the cyclohexenyl nucleic acid-dependent DNA polymerization**. *Nucleic Acids Res* 2005, **33**:3828-3836.

53. Vastmans K, Pochet S, Peys A, Kerremans L, Van Aerschot A, Hendrix C, Marliere P, Herdewijn P: **Enzymatic incorporation in DNA of 1,5-anhydrohexitol nucleotides**. *Biochemistry* 2000, **39**:12757-12765.

54. Vastmans K, Froeyen M, Kerremans L, Pochet S, Herdewijn P: **Reverse transcriptase incorporation of 1,5-anhydrohexitol nucleotides**. *Nucleic Acids Res* 2001, **29**:3154-3163.

55. Schlegel MK, Peritz AE, Kittigowittana K, Zhang L, Meggers E: **Duplex formation of the simplified nucleic acid GNA**. *Chembiochem* 2007, **8**:927-932.

56. Joyce GF, Schwartz AW, Miller SL, Orgel LE: **The case for an ancestral genetic system involving simple analogues of the nucleotides**. *Proc Natl Acad Sci U S A* 1987, **84**:4398-4402.

57. Heuberger BD, Switzer C: **A pre-RNA candidate revisited: both enantiomers of flexible nucleoside triphosphates are DNA polymerase substrates**. *J Am Chem Soc* 2008, **130**:412-413.

58. Chen JJ, Tsai CH, Cai X, Horhota AT, McLaughlin LW, Szostak JW: **Enzymatic primer-extension with glycerol-nucleoside triphosphates on DNA templates**. *PLoS One* 2009, **4**:e4949.

59. Tsai CH, Chen J, Szostak JW: **Enzymatic synthesis of DNA on glycerol nucleic acid templates without stable duplex formation between product and template**. *Proc Natl Acad Sci U S A* 2007, **104**:14598-14603.

60. Maiti M, Siegmund V, Abramov M, Lescrinier E, Rosemeyer H, Froeyen M, Ramaswamy A, Ceulemans A, Marx A, Herdewijn P: **Solution structure and conformational dynamics of deoxyxylonucleic acids (dXNA): an orthogonal nucleic acid candidate**. *Chemistry* 2012, **18**:869-879.

61. Eckstein F: **Nucleoside phosphorothioates**. *Annu Rev Biochem* 1985, **54**:367-402.

62. Porter KW, Briley JD, Shaw BR: **Direct PCR sequencing with boronated nucleotides**. *Nucleic Acids Res* 1997, **25**:1611-1617.

Please cite this article in press as: Pinheiro VB, Holliger P. The XNA world: progress towards replication and evolution of synthetic genetic polymers, Curr Opin Chem Biol (2012), http://dx.doi.org/10.1016/j.cbpa.2012.05.198

63. Shaw BR, Dobrikov M, Wang X, Wan J, He K, Lin JL, Li P, Rait V, Sergueeva ZA, Sergueev D: **Reading, writing, and modulating genetic information with boranophosphate mimics of nucleotides, DNA, and RNA**. *Ann N Y Acad Sci* 2003, **1002**:12-29.

64. Somasunderam A, Thiviyanathan V, Tanaka T, Li X, Neerathilingam M, Lokesh GL, Mann A, Peng Y, Ferrari M, Klostergaard J *et al.*: **Combinatorial selection of DNA thioaptamers targeted to the HA binding domain of human CD44**. *Biochemistry* 2010, **49**:9106-9112.

65. 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.

66. Nielsen PE: **Peptide nucleic acids (PNA) in chemical biology and drug discovery**. *Chem Biodivers* 2010, **7**:786-804.

67. Nielsen PE, Haaima G, Lohse A, Buchardt O: **Peptide nucleic acids (PNAs) containing thymine monomers derived from chiral amino acids: hybridization and solubility properties of D-lysine PNA**. *Angew Chem Int Ed Engl* 1996, **35**:1939-1942.

68. Zhou P, Dragulescu-Andrasi A, Bhattacharya B, O'Keefe H, Vatta P, Hyldig-Nielsen JJ, Ly DH: **Synthesis of cell-permeable peptide nucleic acids and characterization of their hybridization and uptake properties**. *Bioorg Med Chem Lett* 2006, **16**:4931-4935.

69. Ura Y, Beierle JM, Leman LJ, Orgel LE, Ghadiri MR: **Self-assembling sequence-adaptive peptide nucleic acids**. *Science* 2009, **325**:73-77.

70. Rosenbaum DM, Liu DR: **Efficient and sequence-specific DNA-templated polymerization of peptide nucleic acid aldehydes**. *J Am Chem Soc* 2003, **125**:13924-13925.

71. Isobe H, Fujino T, Yamazaki N, Guillot-Nieckowski M, Nakamura E: **Triazole-linked analogue of deoxyribonucleic acid ((TL)DNA): design, synthesis, and double-strand formation with natural DNA**. *Org Lett* 2008, **10**:3729-3732.

72. El-Sagheer AH, Brown T: **Synthesis and polymerase chain reaction amplification of DNA strands containing an unnatural triazole linkage**. *J Am Chem Soc* 2009, **131**:3958-3964.

73. Wu T, Orgel LE: **Nonenzymatic template-directed synthesis on hairpin oligonucleotides. 3. Incorporation of adenosine and uridine residues**. *J Am Chem Soc* 1992, **114**:7963-7969.

74. Wu T, Orgel LE: **Nonenzymatic template-directed synthesis on hairpin oligonucleotides. 2. Templates containing cytidine and guanosine residues**. *J Am Chem Soc* 1992, **114**:5496-5501.

75. Schrum JP, Ricardo A, Krishnamurthy M, Blain JC, Szostak JW: **Efficient and rapid template-directed nucleic acid copying using 2′-amino-2′,3′-dideoxyribonucleoside-5′-phosphorimidazolide monomers**. *J Am Chem Soc* 2009, **131**:14560-14570.

76. Li X, Zhan ZY, Knipe R, Lynn DG: **DNA-catalyzed polymerization**. *J Am Chem Soc* 2002, **124**:746-747.

77. •• Brudno Y, Birnbaum ME, Kleiner RE, Liu DR: **An in vitro translation, selection and amplification system for peptide nucleic acids**. *Nat Chem Biol* 2010, **6**:148-155.
The authors describe a display approach (Figure 3) that allows template-directed chemical synthesis of PNA using PNA pentamer units, PNA displacement through DNA second strand synthesis and a model selection of displayed PNA. This is a potentially general approach for the chemical synthesis and evolution of XNA polymers.

78. 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.

79. Kozlov IA, Zielinski M, Allart B, Kerremans L, Van Aerschot A, Busson R, Herdewijn P, Orgel LE: **Nonenzymatic template-directed reactions on altritol oligomers, preorganized analogues of oligonucleotides**. *Chemistry* 2000, **6**:151-155.

80. Zhang N, Zhang S, Szostak JW: **Activated ribonucleotides undergo a sugar pucker switch upon binding to a single-stranded RNA template**. *J Am Chem Soc* 2012, **134**:3691-3694.

81. Loakes D, Holliger P: **Polymerase engineering: towards the encoded synthesis of unnatural biopolymers**. *Chem Commun (Camb)* 2009, **45**:4619-4631.

82. Pinheiro VB, Ong JL, Holliger P: **Polymerase engineering: from PCR and sequencing to synthetic biology**. In *Protein Engineering Handbook*, vol III. Edited by Lutz S. Wiley-VCH; 2012.

83. Yang G, Franklin M, Li J, Lin TC, Konigsberg W: **A conserved Tyr residue is required for sugar selectivity in a Pol alpha DNA polymerase**. *Biochemistry* 2002, **41**:10256-10261.

84. Patel PH, Loeb LA: **Multiple amino acid substitutions allow DNA polymerases to synthesize RNA**. *J Biol Chem* 2000, **275**:40266-40272.

85. Staiger N, Marx A: **A DNA polymerase with increased reactivity for ribonucleotides and C5-modified deoxyribonucleotides**. *Chembiochem* 2010, **11**:1963-1966.

86. Xia G, Chen L, Sera T, Fa M, Schultz PG, Romesberg FE: **Directed evolution of novel polymerase activities: mutation of a DNA polymerase into an efficient RNA polymerase**. *Proc Natl Acad Sci U S A* 2002, **99**:6597-6602.

87. Ong JL, Loakes D, Jaroslawski S, Too K, Holliger P: **Directed evolution of DNA polymerase, RNA polymerase and reverse transcriptase activity in a single polypeptide**. *J Mol Biol* 2006, **361**:537-550.

88. McCullum EO, Chaput JC: **Transcription of an RNA aptamer by a DNA polymerase**. *Chem Commun (Camb)* 2009, **45**:2938-2940.

89. Shinkai A, Patel PH, Loeb LA: **The conserved active site motif A of *Escherichia coli* DNA polymerase I is highly mutable**. *J Biol Chem* 2001, **276**:18836-18842.

90. •• Pinheiro VB, Taylor AI, Cozens C, Abramov M, Renders M, Zhang S, Chaput JC, Wengel J, Peak-Chew S-Y, McLaughlin SH *et al.*: **Synthetic genetic polymers capable of heredity and evolution**. *Science* 2012, **336**:341-344.
The authors describe the evolution and design of polymerases capable of synthesising and reverse transcribing a range of XNAs (HNA, CeNA, LNA, TNA, ANA, FANA) and measure the fidelity of genetic information transfer between DNA and the various XNAs. Using XNA synthesis and reverse transcription, they describe the evolution and characterization of HNA aptamers against the HIV TAR RNA motif and hen egg lysozyme.

91. 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**:2802-2803.

92. •• Yu H, Zhang S, Chaput JC: **Darwinian evolution of an alternative genetic system provides support for TNA as an RNA progenitor**. *Nat Chem* 2012, **4**:183-187.
The authors describe the isolation of an anti-thrombin TNA aptamer composed of three of the four TNA bases. TNA was enzymatically synthesized using Therminator from templates lacking guanosines and selected using a display approach (Figure 3).

93. Jozwiakowski SK, Connolly BA: **A modified family-B archaeal DNA polymerase with reverse transcriptase activity**. *Chembiochem* 2011, **12**:35-37.

94. Sano S, Yamada Y, Shinkawa T, Kato S, Okada T, Higashibata H, Fujiwara S: **Mutations to create thermostable reverse transcriptase with bacterial family A DNA polymerase from Thermotoga petrophila K4**. *J Biosci Bioeng* 2012, **113**:315-321.

95. Sauter KB, Marx A: **Evolving thermostable reverse transcriptase activity in a DNA polymerase scaffold**. *Angew Chem Int Ed Engl* 2006, **45**:7633-7635.

96. Lockless SW, Ranganathan R: **Evolutionarily conserved pathways of energetic connectivity in protein families**. *Science* 1999, **286**:295-299.

97. Suel GM, Lockless SW, Wall MA, Ranganathan R: **Evolutionarily conserved networks of residues mediate allosteric communication in proteins**. *Nat Struct Biol* 2003, **10**:59-69.

98. Benner SA: **Redesigning genetics**. *Science* 2004, **306**:625-626.

99. Ebert MO, Mang C, Krishnamurthy R, Eschenmoser A, Jaun B: **The structure of a TNA–TNA complex in solution: NMR study of the octamer duplex derived from alpha-(L)-threofuranosyl-(3′-2′)-CGAATTCG**. *J Am Chem Soc* 2008, **130**:15105-15115.

Please cite this article in press as: Pinheiro VB, Holliger P. The XNA world: progress towards replication and evolution of synthetic genetic polymers, Curr Opin Chem Biol (2012), http://dx.doi.org/10.1016/j.cbpa.2012.05.198