CHEMBIOCHEM

Polymerase Amplification, Cloning, and Gene Expression of Benzo-Homologous "yDNA" Base Pairs

Jijumon Chelliserrykattil, Haige Lu, Alex H. F. Lee, and Eric T. Kool*[a]

A widened DNA base-pair architecture is studied in an effort to explore the possibility of whether new genetic system designs might possess some of the functions of natural DNA. In the "yDNA" system, pairs are homologated by addition of a benzene ring, which yields (in the present study) benzopyrimidines that are correctly paired with purines. Here we report initial tests of ability of the benzopyrimidines yT and yC to store and transfer biochemical and biological information in vitro and in bacterial cells. In vitro primer extension studies with two polymerases showed that the enzymes could insert the correct nucleotides opposite these yDNA bases, but with low selectivity. PCR amplifications with a thermostable polymerase resulted in correct pairings in 15–20% of the cases, and more successfully when yT or yC were situated within the primers. Segments of DNA containing one or two yDNA bases were then ligated into a plasmid and tested for their ability to successfully lead the expression of an active protein in vivo. Although active at only a fraction of the activity of fully natural DNA, the unnatural bases encoded the correct codon bases in the majority of cases when singly substituted, and yielded functioning green fluorescent protein. Although the activities with native polymerases are modest with these large base pairs, this is the first example of encoding protein in vivo by an unnatural DNA base pair architecture.

Introduction

A long-standing challenge for chemists inspired by natural systems has been to find new chemical designs for DNA base pairs.[1–6] One goal of such work is to evaluate whether such designs can mimic natural physical properties of DNA such as helix formation. Fewer studies have been directed to the biochemical and biological properties of such new pairs. However, some unnatural base pairs have recently been shown to be replicated by polymerase enzymes in vitro,[2,7–12] suggesting future applications in expansion of the genetic code.[13]

Until recently, designed base pairs have been constructed exclusively for function in the context of the natural DNA genetic system. However, one can ask whether other genetic systems could exist that operate with an entirely different base pairing architecture from that of natural DNA. In this light we recently reported the design and synthesis of size-expanded nucleosides (xDNA and yDNA) capable of forming base pairs and helices larger than those of natural DNA.[14–16] In these designs, the dimension of the natural nucleobases was expanded by 2.4 Å through addition of a benzene ring to the natural heterocyclic framework.[17] While the yDNA analogues studied here (Figure 1) differ from the xDNA analogues in the extension vector orientation, both could be synthetically incorporated into oligonucleotide strands and form stable expanded double helices. These expanded nucleotide analogues may serve as useful tools in better understanding the self assembly and replication of DNA, and its recognition by proteins. They also have the potential to lead to new applications in biotechnology as a result of their ability to bind DNA with high affinity,[14] and the fact that all the expanded bases are fluorescent.[18]
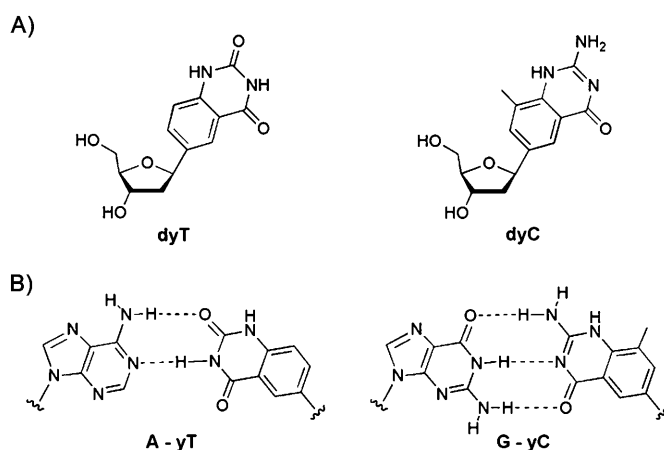

Figure 1. Widened (yDNA) nucleoside and base pair structures. A) Structures of benzopyrimidine nucleosides dyT and dyC. B) Proposed structures of widened base pairs in this study.

If designed genetic systems are to function like natural ones, they must be able to encode and transfer sequence information to copies. However, this presents a serious challenge for the widened yDNA pairs, because DNA polymerases can be highly sensitive to the size of base pairs.[19] While our long-term goal is to discover and develop new enzyme variants that can adapt to the large size of expanded pairs, we have begun to examine whether naturally occurring enzymes and replication systems have any ability to process the information en-

[a] Dr. J. Chelliserrykattil, Dr. H. Lu, Dr. A. H. F. Lee, Prof. Dr. E. T. Kool
Department of Chemistry, Stanford University
Stanford, CA 94305-5080 (USA)
Fax: (+1) 650-725-0259
E-mail: kool@stanford.edu

coded by expanded pairs. Here we report on the encoding of information in the yDNA system. We find that although such pairs are processed with relatively low efficiency and fidelity in vitro, native polymerases can in fact replicate such pairs successfully in some cases. Moreover, we demonstrate that single yDNA bases can correctly encode genetic information in living cells.

Results and Discussion

Our initial experiments explored the ability of the benzopyrimidine C-nucleosides[15b] dyT and dyC (Figure 1 A) to template the addition of natural nucleotides opposite them, using native polymerases. We studied the Klenow fragment of *Escherichia coli* DNA polymerase I (KF *exo⁻*), and a commonly used thermostable polymerase, the DNA polymerase from *Thermococcus litoralis* (Vent *exo⁻*)[20] to recognize dyT and dyC bases in a synthetic 28 mer DNA template. If insertion of dATP and dGTP were successful, the resulting A-yT and G-yC pairs (respectively) would be about 2.4 Å larger than natural Watson–Crick pairs. The single nucleotide insertion studies were performed under nonforcing conditions, such that inefficiently formed base pairs (like mismatches), would not be formed in the case of the natural base system. The data showed (Figure 2) that both the polymerases tested incorporated the natural nucleotides of dA opposite yT and dG opposite yC with qualitatively moderate efficiencies. The selectivity with both enzymes was low, as dTTP was also incorporated across from both yC and yT with comparable efficiencies. We hypothesize that this T-yT and T-yC mispairing is due to a wobble-type pair geometry (Figure 3). Interestingly, such a T-yT pair is reasonably close in size and orientation to a natural A-T pair. A similar structure can be drawn with a possible tautomer of yC.[21,22]

Encouraged by the ability of the thermostable Vent *exo⁻* polymerase to incorporate nucleotides opposite the expanded bases in a DNA template, we next explored the fidelity of this activity in a polymerase chain reaction (PCR)-based assay. Two oligonucleotides were synthesized and annealed, each containing two modifications (one dyT and one dyC), and having a 16nt region of overlap at their 3′ ends (Table 1). PCR was performed using this template and natural dNTPs. Note that the product after amplification in these reactions is natural DNA, as no modified nucleotides were used in the PCR. The resulting 79-base-pair (bp) PCR products were directly cloned into a PCR cloning vector, and transformed into *E. coli*. DNA was then extracted from a small population of the transformed clones, and the resulting sequences were analyzed. The obtained sequences (Table 1) provide information regarding the fidelity of replication of yT and yC bases in the initial template DNA. We observed that in a majority of the clones, the obtained sequence corresponded to the misincorporation of dT opposite both yC and yT in the template, consistent with the above single nucleotide incorporation experiments. The expanded pairs were formed relatively infrequently, with yC being read correctly (that is, incor-
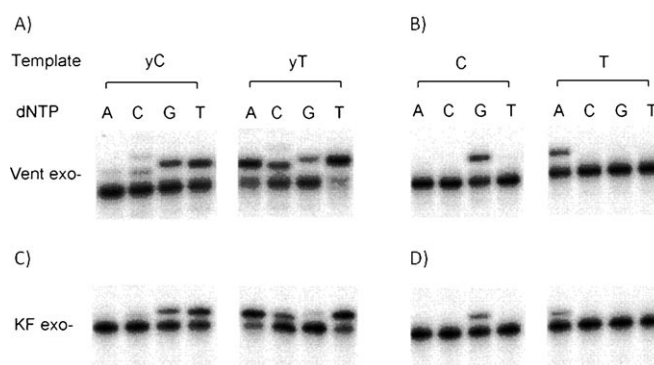

Figure 2. Survey of selectivity of enzymatic nucleotide incorporation opposite yDNA bases. A) and C) show yDNA template bases; B) and D) show natural bases as controls. Enzymes are *Thermococcus litoralis* DNA polymerase (Vent *exo⁻*) and Klenow fragment of DNA pol I (KF *exo⁻*). Standard 10 μL reactions were performed at 37 °C (KF *exo⁻*) or 68 °C (Vent *exo⁻*) and contained 25 nM radiolabeled primer-template (primer, 5′-TAA TAC AAC TCA CTA TAG GGA GA-3′; template, 5′-ACT GXT CTC CCT ATA GTG AGT CGT ATT A-3′, where X = yT,T,yC, or C). A) Reaction solution contained 0.4 units μL⁻¹ Vent *exo⁻*, 500 μM dNTP; reaction time 20 min. B) Reaction contained 0.05 units μL⁻¹ Vent *exo⁻*, 50 μM dNTP; reaction time 3 min. C) Reaction contained 0.02 units μL⁻¹ KF *exo⁻*, 50 μM dNTP; reaction time 4 min. D) Reaction containe 0.005 units μL⁻¹ KF *exo⁻*,5 μM dNTP; reaction time 3 min. Products were resolved by denaturing polyacrylamide gel electrophoresis (7.6 M urea, 20% acrylamide) and visualized by phosphorimaging.
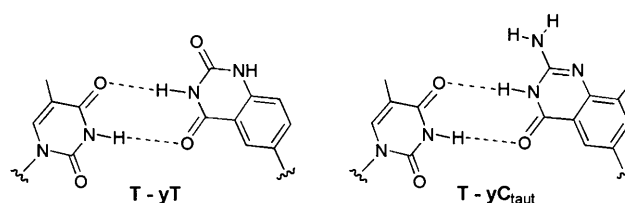

Figure 3. Proposed structures of mismatched T-yT and T-yC wobble-type pairs. The similarity in geometry to standard Watson–Crick pairs might explain the common T-yT and T-yC mispairing observed in vitro.

poration of G) in 15% of clones at one of the two modified positions and similarly, yT being read correctly (incorporation of A) in 20% of clones at one of the two positions. Any other mispairings were rare.

Table 1. PCR amplification of templates containing yT and yC bases with natural dNTPs.

Y2F/Y2R template
primer–ATG<u>T</u>A<u>C</u>AAGAACGCACGATATCTTTCAAA-3′
                                  3′-GTGCTATAGAAAGTTTCTAC<u>T</u>GC<u>C</u>TTT–primer
Expected sequence
    primer–ATG**T**AC AAGAACGCACGATATCTTTCAAAGATG**A**CG**G**AAA–primer
Observed sequence
    primer–ATG*AA*AAAGAACGCACGATATCTTTCAAAGATG*TCG*T*AAA–primer (55%)
    primer–ATG*AA*AAAGAACGCACGATATCTTTCAAAGATG**A**CG*T*AAA–primer (15%)
    primer–ATG*AA*C AAGAACGCACGATATCTTTCAAAGATG*TCG*T*AAA–primer (15%)
    primer–ATG*C*A AAGAACGCACGATATCTTTCAAAGATG**A**CG*T*AAA–primer (5%)
     primer–ATG*AA*AAGAACGCACGATATCTTTCAAAGATG*G*CG*T*AAA–primer (5%)
    primer–ATG*AA*AAAGAACGCACGATATCTTTCAAAGATG*CC*GAAA–primer (5%)

Observed sequences were obtained after bacterial cloning and sequencing of amplified duplexes. The yDNA template bases are underlined. Expected bases replacing these are shown in bold; mutated results are shown in italics. Represented percentages of sequences cloned are as shown.

To test yDNA base pair replication in a second context, we next performed PCR using primers that each contained two yDNA bases. The modified primers were designed such that the y-bases were incorporated towards the 5′-end, so that their ability to anneal with the template would not be compromised (Table 2). In these experiments, replication is more challenging since the polymerase would have to continually use DNA containing yDNA bases as the template during each step of PCR. As in the previous assay, the PCR products were directly cloned and sequenced. We observed again that in the region downstream from the primer, the observed sequence corresponded to the misincorporation of T opposite both yC and yT in the template in the majority of cases. "Correct" incorporation of A opposite yT occurred in 20% of cases on average (10% at one position and 30% at the second position), and of G opposite yC, in 10% of cases on average (15% at one position and 5% at the second position) (Table 2). However, in the primer binding regions, we observed far fewer misincorporations. All cases tested (at two sites in each of 20 clones) showed correct replacement of yT by T, indicating A-yT pairing. Similarly, yC was replaced by C on an average in 88% of the cases (75% at one position, 100% at the second position). In this experiment, it should be noted that the products from each PCR cycle are modified DNAs containing yC and yT modifications near the ends. This could potentially pose an additional difficulty for ligation and cloning into the TA vector. Thus, in this experiment we cannot rule out some possible bias of the bacterial replication machinery during plasmid replication for certain sequence outcomes over others, especially near the ends of the amplified duplexes. Nevertheless, the results show that a native polymerase can read and replicate the information stored by yDNA pairs to a substantial extent.

The above experiments demonstrated that small numbers of yDNA base pairs can be amplified in vitro and replicated in vivo. However, the cloned segments were in noncoding parts of the DNA, and the yDNA bases were processed prior to reaching cells. Thus, the question remained as to whether yDNA bases could encode genetic information (as observed by a phenotype) to any extent in a living system. To directly probe this question, we examined the in vivo replication of synthetic DNAs containing yDNA bases without amplification. A 46-base-pair segment within the green fluorescent protein (GFP) gene in pGFPuv plasmid vector, a plasmid with a high copy number of expression in cells, (see Experimental Section) was replaced with an identical synthetically prepared DNA containing yDNA modifications (Table 3). For this, two complementary oligonucleotides, both containing increasing numbers of yT and yC bases within the gene segment, were synthesized.

Table 3. In vivo replication of gene segments containing yDNA base pairs.

y-Gene1[a]
5′-GTACAGGAGCGCACTATTTCTT<u>T</u>CAAAGATGACGGAAACTACAAGA
3′-TCCTCGCGTGATAAAGAAAGT<u>T</u>TCTACTGCCTTTGATGTTCTGCGC
Expected sequence
GTACAGGAGCGCACTATTTCTTT**TCAA**AGATGACGGAAACTACAAGA
Observed sequence
GTACAGGAGCGCACTATTTCTTT**TCAA**AGATGACGGAAACTACAAGA (60%)
GTACAGGAGCGCACTATTTCTTT**TCA**TAGATGACGGAAACTACAAGA (40%)
y-Gene2
5′-GTACAGGAGCGCACTATTTCT**TTC**AAGGATGACGGAAACTACAAGA-3′
3′-TCCTCGCGTGATAAAGAAAG**TTC**CTACTGCCTTTGATGTTCTGCGC-5′
Expected sequence
GTACAGGAGCGCACTATTTCT**TTCAAG**GATGACGGAAACTACAAGA
Observed sequence
GTACAGGAGCGCACTATTTCT**TTCAAG**GATGACGGAAACTACAAGA (15%)
GTACAGGAGCGCACTATTTCT**TTCAA**TGATGACGGAAACTACAAGA (60%)
GTACAGGAGCGCACTATTTCT**TTC**TATGATGACGGAAACTACAAGA* (15%)
y-Gene3
5′-GTACAGGAGCGCACTATTTCT**TTC**AAGGATGACGGAAACTACAAGA-3′
3′-TCCTCGCGTGATAAAGAAAG**TTC**CTACTGCCTTTGATGTTCTGCGC-5′
Observed sequence
no colonies observed

The yDNA bases are underlined in the synthetic GFP gene segments. Expected bases replacing these are shown in bold; mutated results are shown in italics. Percentages of clones are given in brackets. Observed sequences were obtained after bacterial cloning and sequencing of synthetic duplexes in a GFP-encoding plasmid. [a] Silent mutation sites (used as markers) are pictured in gray.

Table 2. PCR amplification with both primers and templates containing yT and yC bases (using natural dNTPs).

caagag<u>t</u>g<u>c</u>catgcccgaaggt (forward primer)
(reverse primer) gatgttctgcgca<u>c</u>gac<u>t</u>tca
Expected sequences
caagag**t**g**c**catgcccgaaggttATG**TAC**AAGAACGCACGATATCTTTCAAAGATG**A**CG**G**AAActacaagacgcgt**g**ctg**a**agt
Observed sequences
caagag**t**g**c**catgcccgaaggttATG*AAAA*AGAACGCACGATATCTTTCAAAGATG*TCG*TAAActacaagacgcgt**g**ctg**a**agt (35%)
caagag**t**g**c**catgcccgaaggttATG*AAAA*AGAACGCACGATATCTTTCAAAGATG**A**CG*T*AAActacaagacgcgt**g**ctg**a**agt (20%)
caagag**t**g*a*catgcccgaaggttATG*AAAA*AGAACGCACGATATCTTTCAAAGATG*TCG*TAAActacaagacgcgt**g**ctg**a**agt (10%)
caagag**t**g**c**catgcccgaaggttATG*AA*T*AA*GAACGCACGATATCTTTCAAAGATG*TCG*TAAActacaagacgcgt**g**ctg**a**agt (5%)
caagag**t**g**c**catgcccgaaggttATG*A*GA**C**AAGAACGCACGATATCTTTCAAAGATG*TCG*TAAActacaagacgcgt**g**ctg**a**agt (5%)
caagag**t**g*a*catgcccgaaggttATG*A*GA**C**AAGAACGCACGATATCTTTCAAAGATG*TCG*TAAActacaagacgcgt**g**ctg**a**agt (5%)
caagag**t**g**c**catgcccgaaggttATG*AAAA*AGAACGCACGATATCTTTCAAAGATG**A**CG**G**AAActacaagacgcgt**g**ctg**a**agt (5%)
caagag**t**g**c**catgcccgaaggttATG**T**AAAAGAACGCACGATATCTTTCAAAGATG*TCG*TAAActacaagacgcgt**g**ctg**a**agt (5%)
caagag**t**g*a*catgcccgaaggttATG*AAAA*AGAACGCACGATATCTTTCAAAGATG**A**CG*T*AAActacaagacgcgt**g**ctg**a**agt (5%)
caagag**t**g*a*catgcccgaaggttATG**TAC**AAGAACGCACGATATCTTTCAAAGATG*TCG*TAAActacaagacgcgt**g**ctg**a**agt (5%)

The yDNA template bases are underlined. Expected bases replacing these are shown in bold; mutated results are shown in italics. Represented percentages of sequences cloned are as shown. Template sequence is given in Table 1. Observed sequences were obtained after bacterial cloning and sequencing of amplified duplexes.

Upon hybridization of these sequences, a double-stranded DNA fragment was yielded that had sticky ends similar to those generated by two restriction enzymes on the plasmid vector. Genetically silent mutations were also introduced within the synthetic gene to distinguish the input unnatural DNA from possible unmodified vector contamination. These "artificial gene" segments were cloned into the vector, replacing the original sequence, and then transformed into *E. coli*. Here, only if the bacterial DNA replication machinery is capable of correctly propagating the genetic information provided in the form of modified DNA content, would colonies with green fluorescence form on an LB plate. Note that the bacteria only need to make a few faithful copies of the modified plasmid input initially to foster further natural plasmid replication. A high copy number yielding plasmid would also help maintain the continued replication as compared with a low copy number plasmid.

We found that green colonies were in fact obtained when gene segments containing one yDNA base and two yDNA bases in each strand were present in the vector. Transformation efficiency was lowered by about 200-fold compared to a control reaction where unmodified DNA was used for ligation and transformation. However, upon sequencing the resulting plasmids extracted from the fluorescent colonies from modified DNAs, we observed that there were many clones where the genetic information was preserved by correct replication of the unnatural pairs. When there were two yDNA modifications in the cloned gene segment, 60% of the sequences obtained had the expected fidelity of replication with dA added opposite yT. The mutational tendency for the misinsertion of T opposite yT was also observed again, with 40% of the clones exhibiting the corresponding mutation (Table 3). In the more stringent case where there were two yDNA modifications in each strand, the sequences obtained were more diverse, with only 15% of the clones showing complete fidelity; a small population also showed a preference for the addition of A opposite yC. Interestingly, no misincorporations were observed while replicating the yDNA modifications in the top strand, thus showing complete fidelity in the clones tested. In the case where there were three consecutive yDNA base modifications in each strand, we did not observe any green transformants, thus indicating the limits of the *E. coli* replication machinery for yDNA replication.

Conclusions

Overall, our experiments show that despite the large size of yDNA base pairs, DNA partially modified with yDNA bases can be recognized and replicated to some extent in vitro even by natural polymerases. This suggests the future possibility of finding or developing enzymes that are more accepting of the large size of these pairs. More surprisingly, yDNA pairs could be successfully replicated in living bacterial cells to a limited extent, and we demonstrate for the first time that yDNA base pairs can help store and transfer information leading to an observable phenotype. The findings suggest that with further development of replicases that can better adapt to altered geometries genetic systems that use unnatural information-encoding architecture may one day be viable.

Experimental Section

Modified oligodeoxynucleotides: Modified nucleoside analogues dyT and dyC were synthesized and characterized following the methods described previously.[21] Suitably protected cyanoethyl phosphoramidite derivatives were prepared following literature procedures.[21] DNA sequences containing modified bases were prepared on an Applied Biosystems 394 instrument following the published methods (Foster City, CA, USA).[15b] Unmodified oligodeoxynucleotides were purchased from Integrated DNA technologies (IDT, Coralville, IA, USA).

Polymerase reactions: 28 mer/23 mer template-primer duplexes with the sequence (5′-ACT GXT CTC CCT ATA GTG AGT CGT ATT A) · (5′-TAA TAC GAC TCA CTA TAG GGA GA), where X is dyT, dyC, dT, or dC, were used as polymerase substrates. The enzymes used were the exonuclease-deficient Klenow fragment of DNA polymerase I (Cat. No. M0212L from New England Biolabs, Ipswich, MA, USA) and the exonuclease-deficient Vent DNA polymerase from *T. litoralis* (Cat. No. M0257S from New England Biolabs). Primer 5′-termini were radiolabeled by using 5′-[γ-^{32}P]-ATP (Amersham Biosciences, Piscataway, NJ, USA) and T4 Polynucleotide Kinase (Invitrogen, Carlsbad, CA, USA). Labeled primers were purified by MicroSpin™ G-25 columns (Amersham Biosciences). The single nucleotide insertion reactions were performed at 37 °C (for KF *exo⁻*) or 68 °C (for Vent *exo⁻*). Reaction buffer (10×) contained Tris-HCl (500 mM, pH 7.4), MgCl$_2$ (100 mM), DTT (10 mM), BSA (0.5 mg mL⁻¹) for KF *exo⁻*; or KCl (100 mM), (NH$_4$)$_2$SO$_4$ (100 mM), Tris-HCl (200 mM, pH 8.8), MgSO$_4$ (20 mM) Triton X-100 (1%) for Vent *exo⁻*. duplex solution (2.5 μL) containing nonradiolabeled duplex (5 μM), radiolabeled duplex (25 nM) and reaction buffer (2×) was mixed with enzyme solution (2.5 μL) containing varied concentrations of enzyme and reaction buffer (2×), and was incubated for one minute. dNTP solution (5.0 μL) at various concentrations was then added and the mixture reacted for various times (see legend of Figure 2). Reactions were stopped by addition of stop buffer (15 μL) containing formamide (40%), xylene cyanol (0.05%) and bromophenol blue (0.05%). Products were resolved by denaturing polyacrylamide gel electrophoresis (7.6 M urea, 20% acrylamide) and visualized by phosphorimager (Amersham Biosciences, formerly Molecular Dynamics) and ImageQuant program.

Polymerase amplification: For the PCR amplification experiments, primers (0.4 μM each), templates (50 ng of annealed double-stranded DNA, Table 1), and dNTPs (200 μM each), were mixed with reaction buffer and Vent *exo⁻* polymerase (both from New England Biolabs, Ipswich, MA; Cat No: M0254S) and adjusted to a final reaction volume of 50 μL with water. The reaction conditions consisted of an initial denaturation step of 95 °C for 1 min, followed by 20 cycles of 94 °C for 30 s, 50 °C for 30 s, and 68 °C for 40 s. Products after PCR were visualized on an agarose gel. The conditions for the PCR reactions were initially optimized to determine the optimal annealing and extension temperatures for amplification using modified yDNA templates and Vent DNA polymerase. In general, with modified DNA as templates, it was observed that higher elongation temperature provided more robust amplifications.

Plasmid construction and cloning: The products from PCR assays using yDNA templates were directly cloned into the pCR2.1 vector using TA cloning kit (Invitrogen; cat no: K4575) and transformed into *E. coli*. Plasmid DNA was isolated from the transformants using

Miniprep DNA isolation kit (Qiagen, Valencia, CA, USA) and the amplified region was sequenced using appropriate primers. All DNA sequencing was done using the standard automated fluorescent sequencing techniques.

For cloning the GFP gene segment containing the yDNA bases to pGFPuv plasmid (Clontech, Mountain View, CA; Cat No:632312), which is a high copy number plasmid yielding ~500 copies per cell, the vector was first prepared by digesting the plasmid with the restriction enzymes, BsrGI and MluI (New England Biolabs); this removes a 46-base-pair region from within the GFP gene. The resulting linearized vector with cohesive ends was ligated with the annealed duplex DNA containing the yDNA bases (Table 3) using T4 DNA ligase (New England Biolabs). The ligated product was transformed to BL21 (DE3) cells (Stratagene, La Jolla, CA, USA) and plated on LB agar plates containing ampicillin and isopropyl thio-galactosidase (IPTG). The green fluorescent colonies obtained after transformation, identified by viewing under UV transilluminator, were picked, plasmid DNA isolated, and the cloned region within the gene was sequenced. Transformation efficiencies were compared with that of a control ligation reaction under identical conditions using unmodified duplex DNA, which was otherwise identical in sequence to the yDNA containing segment.

Acknowledgements

[1] a) J. A. Piccirilli, T. Krauch, S. E. Moroney, S. A. Benner, *Nature* **1990**, *343*, 33; b) C. Y. Switzer, S. E. Moroney, S. A. Benner, *Biochemistry* **1993**, *32*, 10489; c) J. Horlacher, M. Hottiger, V. N. Podust, U. Hubscher, S. A. Benner, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 6329; d) S. A. Benner, *Acc. Chem. Res.* **2004**, *37*, 784.

[2] H. P. Rappaport, *Nucleic Acids Res.* **1988**, *16*, 7253.

[3] a) B. A. Schweitzer, E. T. Kool, *J. Am. Chem. Soc.* **1995**, *117*, 1863; b) T. J. Matray, E. T. Kool, *J. Am. Chem. Soc.* **1998**, *120*, 6191; c) J. Lai, J. Qu, E. T. Kool, *J. Am. Chem. Soc.* **2004**, *126*, 3040; d) E. T. Kool, *Acc. Chem. Res.* **2002**, *35*, 936.

[4] E. L. Tae, Y. Wu, G. Xia, P. G. Schultz, F. E. Romesberg, *J. Am. Chem. Soc.* **2001**, *123*, 7439; b) A. A. Henry, F. E. Romesberg, *Curr. Opin. Chem. Biol.* **2003**, *7*, 727.

[5] a) T. Mitsui, M. Kimoto, A. Sato, S. Yokoyama, I. Hirao, *Bioorg. Med. Chem. Lett.* **2003**, *13*, 4515; b) T. Mitsui, A. Kitamura, M. Kimoto, T. To, A. Sato, I. Hirao, S. Yokoyama, *J. Am. Chem. Soc.* **2003**, *125*, 5298; c) I. Hirao, *Curr. Opin. Chem. Biol.* **2006**, *10*, 622.

[6] a) K. Tanaka, A. Tengeiji, T. Kato, N. Toyama, M. Shiro, M. Shionoya, *J. Am. Chem. Soc.* **2002**, *124*, 12494; b) M. Shionoya, K. Tanaka, *Curr. Opin. Chem. Biol.* **2004**, *8*, 592; c) N. Zimmermann, E. Meggers, P. G. Schultz, *J. Am. Chem. Soc.* **2002**, *124*, 13684; d) J. Parsch, J. W. Engels, *Nucleosides Nucleotides Nucleic Acids* **2001**, *20*, 815; e) S. Hikishima, N. Minakawa, K. Kuramoto, Y. Fujisawa, M. Ogawa, A. Matsuda, *Angew. Chem.* **2005**, *117*, 602; *Angew. Chem. Int. Ed.* **2005**, *44*, 596; f) Z. Johar, A. Zahn, C. J. Leumann, B. Jaun, *Chem. Eur. J.* **2008**, *14*, 1080; g) B. D. Heuberger, D. Shin, C. Switzer, *Org. Lett.* **2008**, *10*, 1091.

[7] a) A. M. Sismour, S. A. Benner, *Nucleic Acids Res.* **2005**, *33*, 5640; b) Z. Yang, D. Hutter, P. Sheng, A. M. Sismour, S. A. Benner, *Nucleic Acids Res.* **2006**, *34*, 6095.

[8] a) J. C. Morales, E. T. Kool, *Nat. Struct. Biol.* **1998**, *5*, 950; b) T. J. Matray, E. T. Kool, *Nature* **1999**, *399*, 704; c) J. S. Lai, E. T. Kool, *Chem. Eur. J.* **2005**, *11*, 2966.

[9] a) A. A. Henry, C. Yu, F. E. Romesberg, *J. Am. Chem. Soc.* **2003**, *125*, 9638; b) A. A. Henry, A. G. Olsen, S. Matsuda, C. Yu, B. H. Geierstanger, F. E. Romesberg, *J. Am. Chem. Soc.* **2004**, *126*, 6923; c) S. Matsuda, A. A. Henry, F. E. Romesberg, *J. Am. Chem. Soc.* **2006**, *128*, 6369; d) G. T. Hwang, F. E. Romesberg, *Nucleic Acids Res.* **2006**, *34*, 2037; e) S. Matsuda, J. D. Fillo, A. A. Henry, P. Rai, S. J. Wilkens, T. J. Dwyer, B. H. Geierstanger, D. E. Wemmer, P. G. Schultz, G. Spraggon, F. E. Romesberg, *J. Am. Chem. Soc.* **2007**, *129*, 10466; f) A. M. Leconte, G. T. Hwang, S. Matsuda, P. Capek, Y. Hari, F. E. Romesberg, *J. Am. Chem. Soc.* **2008**, *130*, 2336.

[10] a) I. Hirao, M. Kimoto, T. Mitsui, T. Fujiwara, R. Kawai, A. Sato, Y. Harada, S. Yokoyama, *Nat. Methods* **2006**, *3*, 729; b) I. Hirao, Y. Harada, M. Kimoto, T. Mitsui, T. Fujiwara, S. Yokoyama, *J. Am. Chem. Soc.* **2004**, *126*, 13298.

[11] a) N. Paul, V. C. Nashine, G. Hoops, P. Zhang, J. Zhou, D. E. Bergstrom, V. J. Davisson, *Chem. Biol.* **2003**, *10*, 815; b) O. Adelfinskaya, V. C. Nashine, D. E. Bergstrom, V. J. Davisson, *J. Am. Chem. Soc.* **2005**, *127*, 16000.

[12] a) H. P. Rappaport, *Biochem. J.* **2004**, *381*, 709; b) S. C. Johnson, C. B. Sherrill, D. J. Marshall, M. J. Moser, J. R. Prudent, *Nucleic Acids Res.* **2004**, *32*, 1937; c) K. Kincaid, J. Beckman, A. Zivkovic, R. L. Halcomb, J. W. Engels, R. D. Kuchta, *Nucleic Acids Res.* **2005**, *33*, 2620; d) X. Zhang, I. Lee, A. J. Berdis, *Biochemistry* **2005**, *44*, 13101.

[13] J. D. Bain, C. Switzer, A. R. Chamberlin, S. A. Benner, *Nature* **1992**, *356*, 537.

[14] a) H. Liu, J. Gao, S. Lynch, L. Maynard, D. Saito, E. T. Kool, *Science* **2003**, *302*, 868; b) H. Liu, J. Gao, E. T. Kool, *J. Am. Chem. Soc.* **2005**, *127*, 1396; c) J. Gao, H. Liu, E. T. Kool, *Angew. Chem.* **2005**, *117*, 3178; *Angew. Chem. Int. Ed.* **2005**, *44*, 3118; d) S. R. Lynch, H. Liu, J. Gao, E. T. Kool, *J. Am. Chem. Soc.* **2006**, *128*, 14704.

[15] a) H. Lu, K. He, E. T. Kool, *Angew. Chem.* **2004**, *116*, 5958; *Angew. Chem. Int. Ed.* **2004**, *43*, 5834; b) A. H. F. Lee, E. T. Kool, *J. Am. Chem. Soc.* **2005**, *127*, 3332.

[16] A. Krueger, H. Lu, A. H. F. Lee, E. T. Kool, *Accounts Chem. Rev.* **2007**, *40*, 141.

[17] a) D. I. Scopes, J. R. Barrio, N. J. Leonard, *Science* **1977**, *195*, 296; b) N. J. Leonard, *Acc. Chem. Res.* **1982**, *15*, 128.

[18] a) H. Liu, J. Gao, D. Saito, L. Maynard, E. T. Kool, *J. Am. Chem. Soc.* **2004**, *126*, 1102; b) A. H. F. Lee, E. T. Kool, *J. Org. Chem.* **2005**, *70*, 132; c) H. Liu, J. Gao, E. T. Kool, *J. Org. Chem.* **2005**, *70*, 639.

[19] a) T. W. Kim, J. C. Delaney, J. M. Essigmann, E. T. Kool, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 15803; b) T. W. Kim, L. G. Brieba, T. Ellenberger, E. T. Kool, *J. Biol. Chem.* **2006**, *281*, 2289; c) H. O. Sintim, E. T. Kool, *Angew. Chem.* **2006**, *118*, 2008; *Angew. Chem. Int. Ed.* **2006**, *45*, 1974.

[20] P. Mattila, J. Korpela, T. Tenkanen, K. Pitkanen, *Nucleic Acids Res.* **1991**, *19*, 4967.

[21] A. H. F. Lee, E. T. Kool, *J. Org. Chem.* **2005**, *70*, 132.

[22] M. Fuentes-Cabrera, B. G. Sumpter, P. Lipkowski, J. C. Wells, *J. Phys. Chem. B* **2006**, *110*, 6379.