Medicinal and Bioorganic Chemistry of Nucleosides and Nucleotides

Review

Unnatural Base Pairs for Synthetic Biology

Noriko Saito-Tarashima and Noriaki Minakawa*

*Graduate School of Pharmaceutical Science, Tokushima University;
1-78-1 Shomachi, Tokushima 770-8505, Japan.*

Received August 29, 2017

In this review, we have summarized the research effort into the development of unnatural base pairs beyond standard Watson–Crick (WC) base pairs for synthetic biology. Prior to introducing our research results, we present investigations by four outstanding groups in the field. Their research results demonstrate the importance of shape complementarity and stacking ability as well as hydrogen-bonding (H-bonding) patterns for unnatural base pairs. On the basis of this research background, we developed unnatural base pairs consisting of imidazo[5',4':4.5]pyrido[2,3-*d*]pyrimidines and 1,8-naphthyridines, *i.e.*, Im:Na pairs. Since Im bases are recognized as ring-expanded purines and Na bases are recognized as ring-expanded pyrimidines, Im:Na pairs are expected to satisfy the criteria of shape complementarity and enhanced stacking ability. In addition, these pairs have four non-canonical H-bonds. Because of these preferable properties, ImN^N:NaO^O, one of the Im:Na pairs, is recognized as a complementary base pair in not only single nucleotide insertion, but also the PCR.

Key words unnatural base pair; hydrogen-bonding; shape complementarity; DNA polymerase; PCR

1. Introduction

Recently, work by the Human Genome Project-Write, which focuses on synthesizing human genomes, has started.¹⁾ Rewriting entire human genomes will deepen our understanding of the genetic code and have an impact on human health. In this manner, synthetic biology is a bottom-up-type research field that deals with the preparation of materials that comprise life systems. As only two base pairs have been selected during the evolution of life, *i.e.*, adenine (A):thymine (T) and guanine (G):cytosine (C) pairs, these represent ideal genetic polymers. The specific formation of hydrogen bonds (H-bonds) in the A:T pair (two H-bonds) and G:C pair (three H-bonds) is the most fundamental rule of genetic information. In 1962, with surprising foresight, Rich proposed the possibility of an extra artificial base pair, *i.e.*, isoguanine (**isoG**, 6-amino-2-oxopurine) and isocytosine (**isoC**, 2-amino-4-oxopurine), representing fifth and sixth DNA nucleobases.²⁾ The artificially designed **isoG:isoC** pair has three H-bonds with the specific proton donor (D) and proton acceptor (A) geometry [DDA:AAD], which is different from those in the A:T pair ([DA:AD]) and the G:C pair ([ADD:DAA]) (Fig. 1). If an extra base pair can function selectively in replication, transcription, and translation alongside natural Watson–Crick (WC) base pairs, it could potentially allow expansion of the genetic code. Thus, the creation of unnatural base pairs is a challenging and ideal research theme in synthetic biology. Herein, research into the development of unnatural base pairs and their applications are described.

2. Typical Strategy for Designing New Unnatural Base Pairs

Prior to presenting our unnatural base pair studies, the work of four famous and pioneering groups focusing on unnatural base pairs is introduced.

2.1. Unnatural Base Pairs with Non-standard H-Bonding Geometries; Benner's Group In 1989, Benner and colleagues synthesized **isoG** and **isoC** nucleosides and their triphosphates with the goal of expanding the genetic alphabet³⁾ (Fig. 1). The **isoG:isoC** pair was recognized as a complementary base pair by polymerases both in *in vitro* replication and transcription systems.⁴⁾ They also designed other unnatural base pairs with different H-bonding patterns, such as the **X:κ** pair.⁵⁾ Additionally, they succeeded in incorporating the unnatural amino acid 3-iodotyrosine into a peptide by using a pair of 54-mer mRNA comprising **isoC** and tRNA with an **isoGUC** anticodon in *in vitro* translation systems.⁶⁾ These were the first studies to succeed in artificially rebuilding the central dogma using unnatural base pairs, indicating that the alteration of H-bonding geometries in base pairs is a promising strategy for creating a new unnatural base pairs. However, the selectivity of the **isoG:isoC** pair in enzymatic replication was unsatisfactory. This is because **isoG** has a problem with tautomerism, in that the enol form of **isoG** has a [DAD] H-bonding pattern that is complementary to that of T.⁴⁾ In 2005, to address this drawback of **isoG**, they replaced natural T with 2-thioT (**T^s**).⁷⁾ Because of the bulkiness and H-bonding properties (weak proton acceptability) of the thione, **T^s** is less likely to mispair with the tautomer of **isoG** than natural T. Fidelity per doubling of the **isoG:isoC** pair along with the A:**T^s** pair in the PCR was improved by around 98%, although

* To whom correspondence should be addressed. e-mail: minakawa@tokushima-u.ac.jp

Fig. 1. Unnatural Base Pairs Developed by Benner's Group

Fig. 2. Unnatural Base Pairs Developed by Kool's Group

that with natural A:T was 93%.⁷⁾ However, when using an unnatural base pair with 98% replication fidelity, the retention of the unnatural base pair in its amplified DNA fragment after a 20-cycle PCR is decreased to 67% (*i.e.*, $0.98^{20} \approx 0.67$). Because the error rate for natural WC pairing in replication is *ca.* 10^{-6} errors/bp, highly exclusive selectivity of unnatural base pairs is required. Thus, they also created another unnatural base pair comprising 2-aminoimidazo[1,2-*a*]-1,3,5-triazin-4(8*H*)-one (**P**) and 6-amino-5-nitro-2(1*H*)-pyridone (**Z**).⁸⁾ The **P:Z** pair, which has [AAD:DDA] H-bonding geometry, exhibits up to 99.8% fidelity per doubling without using the A:T^S pair because, unlike **isoG**, **Z** does not tautomerize.⁹⁾ Recently, they applied the six-letter genetic system with the **P:Z** pair to the cell-systematic evolution of ligands by exponential enrichment (SELEX) system and succeeded in obtaining highly active aptamers against HepG2 liver cancer cells.¹⁰⁾

2.2. Non-hydrogen-Bonded Unnatural Base Pairs;

Kool's Group During the same decade as Benner's pioneering works, Kool *et al.* have explored the possibility of non-H-bonded unnatural base pairs. In 1998, they created an unnatural base pair comprising 4-methylbenzimidazole (**Z**)¹¹⁾ and 2,4-difluorotoluene (**F**) as steric isosteres of the natural A:T pair^{12,13)} (Fig. 2A). In an *in vitro* replication system, **Z** and **F** were equally replaced with natural A and T but not G and C, demonstrating the importance of shape complementary and stacking interactions in addition to H-bonding in base pairing. Additionally, they designed a modified **Z** base, 9-methyl-1*H*-imidazo[4,5-*b*]pyridine (**Q**), that has a proton acceptor corresponding to the N3 atom.¹⁴⁾ Because the incorporation efficiency of **Q** by Klenow fragment (KF) DNA polymerase is superior to that of **Z**, the importance of proton acceptors in the minor groove for unnatural base pair design is also demonstrated.

To further evaluate the importance of shape complementar-

ity in base pairing, they also created size-expanded (benzo-fused) WC-like base pairs, such as **xA:T** and **A:xT** pairs (termed **xDNA**), that have the same H-bonding geometry as natural WC base pairs but with their pairing edges shifted outward by 2.4 Å (*i.e.*, the width of benzene)^{15,16} (Fig. 2B). KF polymerase incorporated natural nucleoside triphosphate (dNTP) opposite **xDNA** bases in a DNA template with an efficiency *ca.* 1000-fold lower than that of natural pairs,¹⁶ and endogenous *Escherichia coli* (*E. coli*) enzymes accurately transcribed **xDNA** to encode the bacteria phenotype.¹⁷

2.3. Creation of a Semi-synthetic Organism with an Unnatural Base Pair; Romesberg's Group Romesberg and colleagues have also developed various kinds of non-H-bonded unnatural base pairs. In 1999, they reported the self-complementary 7-propynylisocarbostyril (**PICS**):**PICS** pair¹⁸ (Fig. 3). When **PICS:PICS** base pairs are incorporated into DNA, the resulting duplex shows high thermal stability, and KF polymerase recognizes the **PICS:PICS** pair as a complementary base pair. However, further replication reactions after **PICS:PICS** base pairing are terminated because

Fig. 3. Unnatural Base Pairs Developed by Romesberg's Group

PICS bases overlap with each other, indicating structural change in the DNA duplex. Consequently, they explored more than 100 kinds of unnatural base pairs^{19–25} and succeeded in developing **5SICS:MMO2** and **5SICS:NaM** pairs, which are replicable unnatural base pairs in the PCR.^{26–28} In 2014, they reported the creation of a semi-synthetic organism containing the **5SICS:NaM** base pair. In this work, an exogenously expressed nucleoside triphosphate transporter imported **d5SICS** and **dNaM** triphosphates efficiently into *E. coli*, and an endogenous replication system used them in the genetic codes.²⁹ This report had a great impact on synthetic biology, and some researchers consider the created organism to be “alien.”

2.4. Unnatural Base Pair as a Powerful Tool for Creating Highly Functional Nucleic Acids; Hirao's Group Hirao *et al.* have also focused on the creation of unnatural base pairs that function in replication, transcription, and translation in the same way as natural WC base pairs. Their unnatural base pairs were developed by exploiting the concept of steric hindrance. In their 2-amino-6-(2-thienyl)purine (**s**):2-oxo-1*H*-pyridine (**y**) pair, the purine-like **s** has a bulky substituent at the major groove side^{30,31} (Fig. 4). Thus, the **s:y** pair is selectively recognized as a complementary base pair by KF polymerase in *in vitro* replication systems. Furthermore, in 2002 they succeeded in synthesizing the Ras protein modified with 3-iodotyrosine from a DNA template containing the **s** base by combining T7 polymerase transcription and *E. coli* *in vitro* translation systems.³² They also developed the **Ds:Pa** base pair, in which H-bonding atoms and substituents located at the base-pairing side are excluded.³³ The replication selectivity of the **Ds:Pa** pair is superior to that of the **s:y** pair, and the **Ds:Pa** pair can be amplified in the PCR with over 99% fidelity per doubling using γ -amino triphosphates of **Ds** and **A**.³³ Concerning selectivity in replication, their unnatural base pairs exhibit the best performances among the reported unnatural base pairs. The low misincorporation rate of the recently developed **Ds:diol1-Px** pair (5×10^{-5} errors/bp) is close to the mispairing error rate of natural WC pairs (2×10^{-5} errors/bp).^{34,35} By making use of this superior property of the

Fig. 4. Unnatural Base Pairs Developed by Hirao's Group

Fig. 5. a) First Generation of Im:Na Pairs ($\text{ImN}^{\text{O}}:\text{NaO}^{\text{N}}$ and $\text{ImO}^{\text{N}}:\text{NaN}^{\text{O}}$); b) $\text{A}:\text{NaO}^{\text{O}}$ Mispair; c) Second Generation of Im:Na Pair ($\text{ImN}^{\text{N}}:\text{NaO}^{\text{O}}$); d) $\text{A}:\text{NaO}^{\text{O}}$ and $\text{G}:\text{NaO}^{\text{O}}$ Mispairs

Ds:diol1-**Px** pair, they succeeded in obtaining a DNA aptamer containing the **Ds** base against human protein target, vascular endothelial cell growth factor-165 (VEGF-165).³⁶⁾ Because the affinities of aptamers that have **Ds** bases are >100-fold improved over those of aptamers containing only natural bases, the potential of genetic alphabet expansion as a powerful tool for creating highly functional nucleic acids is demonstrated.

3. Four H-Bonding Unnatural Base Pairs; Our Group

In contrast to the research described above, we began our unnatural base pair studies to address the simple question: why did WC base pairs come to contain two or three H-bonds during the evolution of life? To answer this, we have explored four H-bonding base pairs.^{37–41)} As purine-type nucleobases, a series of imidazo[5',4':4.5]pyrido[2,3-*d*]pyrimidines (Im) were designed,³⁷⁾ while 1,8-naphthyridines (Na) were designed as their complementary pyrimidine nucleobases.³⁸⁾ For the first generation of our four-H-bonding unnatural base pairs, two Im:Na pairs, *i.e.*, $\text{ImN}^{\text{O}}:\text{NaO}^{\text{N}}$ and $\text{ImO}^{\text{N}}:\text{NaN}^{\text{O}}$, which have alternate H-bonding geometries, were developed. As can be seen in Fig. 5a, these pairs have four non-canonical H-bonds and expanded aromatic surfaces, and they satisfy the shape complementarity criterion like WC base pairs. Because of the contributions of these effects, DNA duplexes containing these pair(s) are significantly thermally stabilized (*ca.* +8°C/pair).³⁸⁾ In addition, both pairs are recognized by KF polymerase as complementary in single nucleotide insertion. However, the kinetic parameters determined for their 5'-triphosphates revealed that the efficiencies of incorporation for $\text{ImN}^{\text{O}}:\text{NaO}^{\text{N}}$ and $\text{ImO}^{\text{N}}:\text{NaN}^{\text{O}}$ pairs are 1–2 orders of magnitude lower than those of natural A:T and G:C pairs. Furthermore, misincorporation of natural dNTP, for example, that of 2'-deoxyadenosine 5'-triphosphate (dATP) against NaN^{O} in the template was clearly observed at the same efficiency as that of $\text{ImO}^{\text{N}}\text{TP}$ against NaN^{O} in the template owing to the possible formation of an $\text{A}:\text{NaN}^{\text{O}}$ pair with two H-bonds^{42,43)} (Fig. 5b).

Reagents and conditions: (a) $\text{dba}_3\text{Pd}_2\cdot\text{CHCl}_3$, DMF, 100°C; (b) Na_2CO_3 , aq. EtOH, 80°C; (c) $\text{NH}_4\text{OH}/1,4\text{-dioxane}$, 100°C.

Chart 1

To improve efficiency and selectivity, a new Im:Na pair, *i.e.*, $\text{ImN}^{\text{N}}:\text{NaO}^{\text{O}}$, has been envisioned^{39,44)} (Fig. 5c). This pair has a [DAAD:ADDA] H-bonding geometry, and thus is expected to avoid the misincorporation of natural A and G (Fig. 5d). The chemistry and enzymatic behavior of the $\text{ImN}^{\text{N}}:\text{NaO}^{\text{O}}$ pair is described below.

3.1. Synthesis of the Nucleoside Units for the $\text{ImN}^{\text{N}}:\text{NaO}^{\text{O}}$ Pair The most straightforward synthesis of ImN^{N} nucleoside **1** is thought to be through intramolecular cyclization of the 5-pyrimidinylimidazole nucleoside, which can be prepared *via* Stille coupling between the 5-iodoimidazole nucleoside **2** and (tributylstannyl)pyrimidine **3** (Chart 1). When a mixture of **2** prepared from 2'-deoxyinosine and **3** prepared from 2,4-dichloropyrimidine is

heated in *N,N*-dimethylformamide (DMF) in the presence of tris(dibenzylideneacetone)dipalladium(0)-chloroform adduct ($\text{dba}_3\text{Pd}_2 \cdot \text{CHCl}_3$), a mixture of coupling product **4** and a spontaneously cyclized tricyclic product **5** is obtained. Subsequent treatment of the mixture under basic conditions converges the mixture to the tricyclic product **5**. Finally, treatment of **5** with a mixture of 1,4-dioxane and NH_4OH gives the desired ImN^{N} nucleoside **1** in good yield.³⁷⁾ The resulting **1** is then converted into the corresponding phosphoramidite unit and 5'-triphosphate under the appropriate conditions.^{44,45)}

For the synthesis of NaO^{O} nucleoside **6**, which is an unusual C-nucleoside, the palladium-catalyzed Heck reaction was envisioned. As illustrated in Chart 2, 3-iodo-1,8-naphthyridine derivative **7** prepared from 2-amino-7-hydroxy-1,8-naphthyridine and glycol **8** are prepared. Then, Heck coupling of **7** with **8** in the presence of palladium acetate and triphenylarsine followed by deprotection and stereoselective reduction affords 1,8-naphthyridine C-nucleoside **9**. After protection of the hydroxyl groups with silyl groups to give **10**, the substituent at the 2-position is converted into an acetoxy group *via* **11**. Fi-

Reagents and conditions: (a) $\text{Pd}(\text{OAc})_2$, AsPh_3 , Bu_3N , DMF, 60°C ; (b) TBAF, THF; (c) $\text{NaBH}(\text{OAc})_3$, AcOH , CH_3CN ; (d) TIPSCl , imidazole, DMF, 55°C ; (e) NH_3/MeOH , 80°C ; (f) NaNO_2 , AcOH ; (g) NH_3/MeOH , 80°C .

Chart 2

nally, treatment of the resulting **12** with methanolic ammonia at 60°C in a sealed tube gives the desired NaO^{O} nucleoside **6** in good yield. In a similar manner as for **1**, **6** is converted into the corresponding phosphoramidite unit and 5'-triphosphate for enzymatic evaluation.^{39,45)}

3.2. Investigation of Single Nucleotide Insertion with the $\text{ImN}^{\text{N}}:\text{NaO}^{\text{O}}$ Pair To investigate the efficiency and selectivity of the newly designed $\text{ImN}^{\text{N}}:\text{NaO}^{\text{O}}$ pair in *in vitro* replication systems, we examined single nucleotide insertion using KF polymerase, and the kinetic parameters, such as the Michaelis constant (K_m), the maximum rate of the enzyme reaction (V_{max}), and the incorporation efficiency (V_{max}/K_m), for the $\text{ImN}^{\text{N}}:\text{NaO}^{\text{O}}$ pair were determined and compared with those of the two previous Im:Na pairs⁴⁵⁾ (Fig. 6). As discussed above, the values of V_{max}/K_m for the $\text{ImN}^{\text{O}}:\text{NaO}^{\text{N}}$ pair are 1–2 orders of magnitude lower than those of the natural A:T pair (6.0×10^7 – $9.0 \times 10^7 \% \text{min}^{-1} \text{M}^{-1}$), as presented in the first row of Fig. 6. This result is thought to be due to the fact that the NaO^{N} base lacks a proton acceptor corresponding to the O2 atom of the natural pyrimidine base. For the $\text{ImO}^{\text{N}}:\text{NaO}^{\text{O}}$ pair, the V_{max}/K_m values are better than those for the $\text{ImN}^{\text{O}}:\text{NaO}^{\text{N}}$ pair (second row in Fig. 6). However, as well as the desired ImO^{N} , undesired A is incorporated against NaN^{O} in the template with a comparable V_{max}/K_m value (Fig. 5b).

Concerning incorporation efficiencies, the V_{max}/K_m values for the $\text{ImN}^{\text{N}}:\text{NaO}^{\text{O}}$ pair are superior to those for the $\text{ImN}^{\text{O}}:\text{NaO}^{\text{N}}$ and $\text{ImO}^{\text{N}}:\text{NaN}^{\text{O}}$ pairs because the ImN^{N} and NaO^{O} bases have proton acceptors at positions corresponding to the N3 of a purine and the O2 of a pyrimidine, respectively. In addition, the $\text{ImN}^{\text{N}}:\text{NaO}^{\text{O}}$ pair has higher thermal stability than the two previous Im:Na pairs owing to the [DAAD:ADDA] H-bonding pattern.³⁹⁾ The preferable base-pairing properties of the $\text{ImN}^{\text{N}}:\text{NaO}^{\text{O}}$ pair lead to it having the highest incorporation efficiency among the three Im:Na pairs. With respect to specificity, misincorporations of natural A and/or G against ImN^{N} in the template are controlled by the [DAAD:ADDA] H-bonding pattern of the $\text{ImN}^{\text{N}}:\text{NaO}^{\text{O}}$ pair. The efficiency of $\text{ImN}^{\text{N}}\text{TP}$ incorporation against NaO^{O} is at least ten-times higher than those of natural dATP and 2'-deoxyguanosine 5'-triphosphate (dGTP) incorporations. Thus, as expected from Fig. 5d, formation of both A: NaO^{O} and G: NaO^{O} should be negligible owing to the NH proton repulsion between the 6-amino group of A and N8 of NaO^{O} , and that between N1 of G and N1 of NaO^{O} , respectively.

3.3. PCR Amplification with $\text{ImN}^{\text{N}}:\text{NaO}^{\text{O}}$ Pair To apply the newly developed $\text{ImN}^{\text{N}}:\text{NaO}^{\text{O}}$ pair to synthetic biology research like that reported by the four aforementioned

Fig. 6. Graphs of Incorporation Efficiency (V_{max}/K_m) Values

a) Incorporation of dYTP against a series of Im bases in the template. b) Incorporation of dYTP against a series of Na bases in the template. a; n.d.=not determined.

Fig. 7. Fifteen Cycles of PCR Involving an ImN^N:NaO^O Pair

(a) Schematics of the template and primers, and the resulting amplicon. Gel electrophoresis of PCR products obtained using Taq DNA polymerase (b), Deep Vent exo⁻ DNA polymerase (c), Deep Vent exo⁺ DNA polymerase (d), and Pfx⁵⁰ DNA polymerase (e) under different dNTP conditions.

groups, this pair should be viable in PCR amplification. Thus, according to the method reported by Hirao *et al.*,³⁴⁾ PCR involving the ImN^N:NaO^O pair was examined under various dNTP conditions (Fig. 7a).

First, when Taq DNA polymerase, which is a standard thermophilic DNA polymerase for routine PCR, is used, a 75 base-pair amplicon in the presence of ImN^NTP and NaO^OTP along with all four kinds of dNTPs is successfully obtained (Fig. 7b, lane 4). However, similar PCR products are observed under the conditions lacking ImN^NTP (lane 3), indicating that inaccurate amplification occurs, presumably owing to misincorporation of natural A and/or G against NaO^O in the resulting DNA fragment. Thus, we screened suitable thermophilic DNA polymerases, and typical results are shown in Figs. 7c–e. Exonuclease-deficient Deep Vent (Deep Vent exo⁻) DNA polymerase gives the full-length amplicon in both the presence and absence of NaO^OTP (Fig. 7c). Conversely, the same polymerase with 3'→5' exonuclease activity (Deep Vent exo⁺) preferentially affords the PCR product in the presence of all 5'-triphosphates (Fig. 7d), suggesting that the proofreading activity identifies mismatched base pairs with natural nucleobases and corrects them to the ImN^N:NaO^O pair. It has been reported that the proofreading activity of DNA polymerases improves the accuracy of incorporating unnatural base pair analogs,^{32,33)} and the benefits of this activity are apparent in our case.

To further evaluate the fidelity of the ImN^N:NaO^O pair in PCR amplification, we sequenced the resulting PCR product according to methods reported by the groups of Benner²⁶⁾ and Hirao.³⁴⁾ As a result, the lowest total mutation rates of the ImN^N:NaO^O pair is observed when using Pfx⁵⁰ DNA polymerase, and it is estimated to be *ca.* 6% after 15 PCR cycles (fidelity ≈0.995 per doubling) (the analysis of PCR products by gel electrophoresis is shown in Fig. 7d). Although the replication fidelity of the ImN^N:NaO^O pair is slightly inferior to those of other unnatural base-pair analogs,^{8,9,26,28,34,46,47)} it is strongly indicated that the ImN^N:NaO^O pair acts as an orthogonal base pair for WC base pairs during PCR amplification.

4. Conclusion

In this review, we have described our research efforts as well as those of four outstanding groups in the development of unnatural base pairs beyond WC base pairs for synthetic biology. Unlike other groups, we focused on the number of H-bonds together with H-bonding patterns, and thus developed unnatural base pairs consisting of imidazo[5',4':4.5]pyrido[2,3-*d*]pyrimidines and 1,8-naphthyridines, *i.e.*, Im:Na pairs, having four non-canonical H-bonds. Among the Im:Na pairs, the ImN^N:NaO^O pair was recognized by KF polymerase with high specificity and efficiency in the single nucleotide insertion reaction because of the [DAAD:ADDA] H-bonding pattern. Furthermore, accurate PCR amplification was achieved using DNA polymerases possessing 3'→5' exonuclease activity (*ca.* 99.5% per doubling). As is well known, WC base pairs consist of A:T with two H-bonds and G:C with three H-bonds. Thus, it is worth noting that faithful *in vitro* replication of the unnatural ImN^N:NaO^O pair with four H-bonds was achieved, similar to that of WC base pairs. The biological applications of this work,^{48,49)} including the synthetic biology of Im:Na pairs, are currently under development and will be reported elsewhere.

Acknowledgments We thank all of our colleagues, especially Prof. A. Matsuda, Dr. N. Kojima, Dr. S. Hikishima, Mr. K. Kuramoto, and Mr. S. Ogata (Hokkaido University), who contributed to early part of our studies described here. This work was supported by Grants-in-Aid for Scientific Research from the Japan Society for the Promotion of Science (JSPS). N.S.T. thanks the research program for the development of intelligent Tokushima artificial exosome (iTEX) from Tokushima University.

Conflict of Interest The authors declare no conflict of interest.

References and Notes

- Boeke J. D., Church G., Hessel A., Kelley N. J., Arkin A., Cai Y., Carlson R., Chakravarti A., Cornish V. W., Holt L., Isaacs Farren J.,

- Kuiken T., Lajoie M., Lessor T., Lunshof J., Maurano M. T., Mitchell L. A., Rine J., Rosser S., Sanjana N. E., Silver P. A., Valle D., Wang H., Way J. C., Yang L., *Science*, **353**, 126–127 (2016).
- 2) Rich A., “Horizons in Biochemistry,” ed. by Kasha M., Pullman B., Academic Press, New York, 1962, pp. 103–126.
- 3) Switzer C., Moroney S. E., Benner S. A., *J. Am. Chem. Soc.*, **111**, 8322–8323 (1989).
- 4) Switzer C., Moroney S. E., Benner S. A., *Biochemistry*, **32**, 10489–10496 (1993).
- 5) Piccirilli J. A., Krauch T., Moroney S. E., Benner S. A., *Nature* (London), **343**, 33–37 (1990).
- 6) Bain J. D., Switzer C., Chamberlin R., Benner S. A., *Nature* (London), **356**, 537–539 (1992).
- 7) Sismour A. M., Benner S. A., *Nucleic Acids Res.*, **33**, 5640–5646 (2005).
- 8) Yang Z., Sismour A. M., Sheng P., Puskar N. L., Benner S. A., *Nucleic Acids Res.*, **35**, 4238–4249 (2007).
- 9) Yang Z., Chen F., Alvarado J. B., Benner S. A., *J. Am. Chem. Soc.*, **133**, 15105–15112 (2011).
- 10) Sefah K., Yang Z., Bradley K. M., Hoshika S., Jiménez E., Zhang L., Zhu G., Shanker S., Yu F., Turek D., Tan W., Benner S. A., *Proc. Natl. Acad. Sci. U.S.A.*, **111**, 1449–1454 (2014).
- 11) Structure of the Kool’s **Z** base is different from that of Benner’s one.
- 12) Morales J. C., Kool E. T., *Nat. Struct. Biol.*, **5**, 950–954 (1998).
- 13) Guckian K. M., Krugh T. R., Kool E. T., *Nat. Struct. Biol.*, **5**, 954–959 (1998).
- 14) Morales J. C., Kool E. T., *J. Am. Chem. Soc.*, **121**, 2323–2324 (1999).
- 15) Liu H., Gao J., Lynch S. R., Saito Y. D., Maynard L., Kool E. T., *Science*, **302**, 868–871 (2003).
- 16) Krueger A. T., Lu H., Lee A. H. F., Kool E. T., *Acc. Chem. Res.*, **40**, 141–150 (2007).
- 17) Krueger A. T., Peterson L. W., Chelliserry J., Kleinbaum D. J., Kool E. T., *J. Am. Chem. Soc.*, **133**, 18447–18451 (2011).
- 18) McMin D. L., Ogawa A. K., Wu Y., Liu J., Schultz P. G., Romesberg F. E., *J. Am. Chem. Soc.*, **121**, 11585–11586 (1999).
- 19) Tae E. L., Wu Y., Xia G., Schultz P. G., Romesberg F. E., *J. Am. Chem. Soc.*, **123**, 7439–7440 (2001).
- 20) Henry A. A., Yu C., Romesberg F. E., *J. Am. Chem. Soc.*, **125**, 9638–9646 (2003).
- 21) Henry A. A., Olsen A. G., Matsuda S., Yu C., Geierstanger B. H., Romesberg F. E., *J. Am. Chem. Soc.*, **126**, 6923–6931 (2004).
- 22) Leconte A. M., Matsuda S., Hwang G. T., Romesberg F. E., *Angew. Chem. Int. Ed.*, **45**, 4326–4329 (2006).
- 23) Matsuda S., Fillo J. D., Henry A. A., Rai P., Wilkens S. J., Dwyer T. J., Geierstanger B. H., Wemmer D. E., Schultz P. G., Spraggon G., Romesberg F. E., *J. Am. Chem. Soc.*, **129**, 10466–10473 (2007).
- 24) Leconte A. M., Hwang G. T., Matsuda S., Capek P., Hari Y., Romesberg F. E., *J. Am. Chem. Soc.*, **130**, 2336–2343 (2008).
- 25) Seo Y. J., Hwang G. T., Ordoukhanian P., Romesberg F. E., *J. Am. Chem. Soc.*, **131**, 3246–3252 (2009).
- 26) Malyshev D. A., Seo Y. J., Ordoukhanian P., Romesberg F. E., *J. Am. Chem. Soc.*, **131**, 14620–14621 (2009).
- 27) Li L., Degardin M., Lavergne T., Malyshev D. A., Dhimi K., Ordoukhanian P., Romesberg F. E., *J. Am. Chem. Soc.*, **136**, 826–829 (2014).
- 28) Dhimi K., Malyshev D. A., Ordoukhanian P., Kubelka T., Hocek M., Romesberg F. E., *Nucleic Acids Res.*, **42**, 10235–10244 (2014).
- 29) Malyshev D. A., Dhimi K., Lavergne T., Chen T., Dai N., Foster J. M., Correa I. R. Jr., Romesberg F. E., *Nature* (London), **509**, 385–388 (2014).
- 30) Kawai R., Kimoto M., Ikeda S., Mitsui T., Endo M., Yokoyama S., Hirao I., *J. Am. Chem. Soc.*, **127**, 17286–17295 (2005).
- 31) Moriyama K., Kimoto M., Mitsui T., Yokoyama S., Hirao I., *Nucleic Acids Res.*, **33**, e129 (2005).
- 32) Hirao I., Ohtsuki T., Fujiwara T., Mitsui T., Yokogawa T., Okuni T., Nakayama H., Takio K., Yabuki T., Kigawa T., Kodama K., Yokogawa T., Nishikawa K., Yokoyama S., *Nat. Biotechnol.*, **20**, 177–182 (2002).
- 33) Hirao I., Kimoto M., Mitsui T., Fujiwara T., Kawai R., Sato A., Harada Y., Yokoyama S., *Nat. Methods*, **3**, 729–735 (2006).
- 34) Yamashige R., Kimoto M., Takezawa Y., Sato A., Mitsui T., Yokoyama S., Hirao I., *Nucleic Acids Res.*, **40**, 2793–2806 (2012).
- 35) Hirao I., Kimoto M., Yamashige R., *Acc. Chem. Res.*, **45**, 2055–2065 (2012).
- 36) Kimoto M., Yamashige R., Matsunaga K., Yokoyama S., Hirao I., *Nat. Biotechnol.*, **31**, 453–457 (2013).
- 37) Minakawa N., Kojima N., Hikishima S., Sasaki T., Kiyosue A., Atsumi N., Ueno Y., Matsuda A., *J. Am. Chem. Soc.*, **125**, 9970–9982 (2003).
- 38) Hikishima S., Minakawa N., Kuramoto K., Fujisawa Y., Ogawa M., Matsuda A., *Angew. Chem. Int. Ed.*, **44**, 596–598 (2005).
- 39) Kuramoto K., Tarashima N., Hiram Y., Kikuchi Y., Minakawa N., Matsuda A., *Chem. Commun.*, **47**, 10818–10820 (2011).
- 40) Hiram Y., Abe H., Minakawa N., Matsuda A., *Tetrahedron*, **66**, 8402–8406 (2010).
- 41) Hiram Y., Minakawa N., Matsuda A., *Bioorg. Med. Chem.*, **19**, 352–358 (2011).
- 42) Minakawa N., Ogata S., Takahashi M., Matsuda A., *J. Am. Chem. Soc.*, **131**, 1644–1645 (2009).
- 43) Ogata S., Takahashi M., Minakawa N., Matsuda A., *Nucleic Acids Res.*, **37**, 5602–5609 (2009).
- 44) Tarashima N., Higuchi Y., Komatsu Y., Minakawa N., *Bioorg. Med. Chem.*, **20**, 7095–7100 (2012).
- 45) Tarashima N., Komatsu Y., Furukawa K., Minakawa N., *Chem. Eur. J.*, **21**, 10688–10695 (2015).
- 46) Johnson S. C., Sherrill C. B., Marshall D. J., Moser M. J., Prudent J. R., *Nucleic Acids Res.*, **32**, 1937–1941 (2004).
- 47) Hirao I., Kimoto M., *Proc. Jpn. Acad. Ser. B*, **88**, 345–367 (2012).
- 48) Higuchi Y., Furukawa K., Miyazawa T., Minakawa N., *Bioconjug. Chem.*, **25**, 1360–1369 (2014).
- 49) Nomura Y., Kashiwagi S., Sato K., Matsuda A., *Angew. Chem. Int. Ed.*, **53**, 12844–12848 (2014).