Efforts toward creating unnatural base pairs for an expanded genetic code

Ichiro Hirao, Tsuneo Mitsui, Tsuyoshi Fujiwara, Michiko Kimoto, Taiko To, Taeko Okuni, Akira Sato, Yoko Harada and Shigeyuki Yokoyama

Abstract
A series of unnatural base pairs was designed and examined for the expansion of the genetic alphabet and for a better understanding of the mechanism of nucleic acid biosyntheses. To improve the shape complementarity of the previously developed unnatural base pairs, 2-amino-6-(N,N-dimethylamino)purine (x) - pyridon-2-one (y) and 2-amino-6-(2-thienyl)purine (s) - y, the pyrimidine analogue, y, was replaced by a five-member ring, 4-imidazolin-2-one (z), and the s-z pairing in replication was examined. Unnatural bases based on the five-member ring were also applied to the development of non-hydrogen-bonded base pairs.

Introduction
Creating unnatural base pairs that can be recognized by polymerases should lead to new advancements in genetic engineering. The expansion of the genetic codes enables specific incorporation of unnatural amino acids into proteins in vitro and in vivo. Furthermore, the introduction of unnatural nucleotides into DNA and RNA allows for increased functionality of nucleic acids. Thus, many efforts have been made to develop unnatural bases that are specifically incorporated into DNA and RNA opposite the unnatural pairing bases in DNA templates. However, the precedents, such as the nonstandard hydrogen-bonded or hydrophobic base pairs, are still insufficient for specific and efficient site-directed incorporation.

Based on the concepts of the hydrogen-bonding patterns in combination with the shape complementarity, we have developed unnatural base pairs of 2-amino-6-(N,N-dimethylamino)purine (x) - pyridon-2-one (y) and 2-amino-6-(2-thienyl)purine (s) - y, the pyrimidine analogue, y, was replaced by a five-member ring, 4-imidazolin-2-one (z), and the s-z pairing in replication was examined. Unnatural bases based on the five-member ring were also applied to the development of non-hydrogen-bonded base pairs.

Fig. 1. Unnatural base pairs
RESULTS AND DISCUSSION

The nucleoside of \( z, l-(2\text{-deoxy-}\beta\text{-D-ribofuranosyl})\text{-}4\text{-imidazolin-2-one} \), was synthesized via trimethylsilylation of \( 4\text{-imidazolin-2-one} \), glycosidation with \( 2\text{-deoxy-3,5-di-O-p-toluoyl-\alpha\text{-D-erythropentofuranosyl chloride} \), and deprotection of the toluoyl groups. The nucleoside was then converted to the triphosphate for the substrate and to the amidite for the chemical synthesis of DNA templates.

Primer extension after the incorporation of \( z \) opposite \( y \) was efficiently continued by the exonuclease proficient Klenow fragment, but the extension after the non-cognate A-\( z \) pairing was much less efficient. However, single-nucleotide insertion experiments by the exonuclease deficient polymerase showed that the efficiency of the s-\( z \) pairing was similar to that of the s-\( y \) pairing, although the incorporation efficiency of \( z \) opposite A was remarkably reduced as compared with that of \( y \) opposite A. This may be because the hydrophilicity of \( z \) is unfavorable for the interaction with the polymerase. Thus, we designed more hydrophobic bases with the five-member ring, and applied the unnatural bases to non-hydrogen-bonded base pairs.

Kool and co-workers have developed a hydrophobic base pair of 9-methyl-imidazo[4,5-b]pyridine (Q) as an A analogue and 2,4-difluorotoluene (F) as a T analogue to prove the importance of the shape fitting between pairing bases in replication fidelity. The unnatural Q-F, Q-T, and A-F pairings efficiently function in replication. However, the hydrogen of F, corresponding to the H3 of pyrimidines, clashes with the hydrogen of Q on the pairing surface. In addition, F is not recognized by some polymerases, because it lacks the 2-keto group of pyrimidines. Thus, we newly designed the hydrophobic pyrrole-2-carboxaldehyde (Pa) to replace F. The five-member ring of Pa would avoid the steric collision with Q, and the aldehyde group of Pa could interact with polymerases. Thus, the nucleotides of Pa were synthesized, and the Q-Pa pairing in replication was examined with various polymerases.

Single-nucleotide insertion experiments using the Klenow fragment showed the efficient incorporation of the Q-Pa pairing. For example, the efficiency of the incorporation of Pa opposite Q \( (V_{\text{max}}/K_M = 1.3 \times 10^5) \) was similar to that of F opposite Q \( (V_{\text{max}}/K_M = 2.2 \times 10^5) \). Furthermore, Pa was incorporated opposite Q by the AMV reverse transcriptase, but F was not incorporated by the same polymerase.

Here, we have shown that the efficiency and specificity of unnatural base pairings in replication can be increased by fine-tuning the shape complementarity and the polymerase recognition sites of unnatural base pairs. Further analyses of polymerase reactions using the s-\( z \) and Q-Pa pairs are in progress.

References