VARIATION OF GENETIC ALPHABETS OF NUCLEOBASES. M. Kimoto[1,2] and I. Hirao[1,2], [1]RIKEN Center for Life Science Technologies (CLST), 1-7-22 Suehiro-cho, Tsurumi-ku, Kanagawa, Yokohama, 230-0045, Japan, relies on[2]TagCyx Biotechnologies, 1-6-126 Suehiro-cho, Tsurumi-ku, Kanagawa, Yokohama, 230-0045, Japan.

Introduction: Nucleic acids are unique biopolymers, which work as both genetic information materials and functional molecules, such as a catalyst and ligands. The present DNA and RNA molecules are composed of four subunits called nucleotides, containing A, G, C, and T(U), as a nucleobase. Two sets of complementary base pairs, A–T(U) and G–C, play a critical role in storing genetic information and replicating genetic materials. Hypothetically, the number of subunits composing biopolymers might be associated with their replication competence and their funcional capability. Accordingly, the original material on the early Earth has been considered to contain fewer than four different subunits, like a precursor involving only a single base-pairing unit, such as adenine and inosine [1].

Through *in vitro* evolution experiments, Joyce and his colleagues demonstrated that ribozyme activities, which have been found in current life, can be also exhibited even with macromolecules comprising of three or only two different nuleotides [2–4]. Although the activities were much less than those of current RNA moleucles composing of four nucleotides, their results showed that the minimum number of distinct subunits allowing to develop functional informational macromolecules is two. With just only one subunit, it would be difficult to have information and basis for Darwinian evolutions [4], but it might be possible to retain the ability to replicate as a self pairing.

The expansion of the genetic alphabet by unnatural base pairs: Then, how about more than four nucleotides for nucleic acids, if not as much as 20, like amino acids for proteins? While several theoretical approaches have speculated that the optimized number of base types is four, for replicative genetic information storage, computational analysis also suggested the possibilities of six and eight under high-fidelity replicative conditions [5]. In addition, Alexander Rich already imagined a new base pair system including a third base pair as early as 1962 [6], and pioneering studies of unnatural base pairs were started in the late 1980s [7,8]. To date, three research groups, including our own, have reported their own unnatural base pairs, which can be replicable by DNA polymerases with appreciable fidelity as a third base pair [9–14].

To examine the hypothesis that increasing the number of bases augments nucleic acid functionality, we have been studying the development of unnatural base pairs. Recentrly, we developed a hydrophobic **Ds**–**Px** pair (**Ds**: 7-(2-thienyl)imidazo[4,5-b]pyridine; **Px**: 4-propynyl-2-nitropyrrole), which exhibits extremely high selectivity in replication: the **Ds**–**Px** pair retains more than 97% even after 100-cycle (10 times of 10-cycle) PCR amplification, allowing a practical use in further applications [11,12].

By using PCR involvings the **Ds**–**Px** pair, we designed a new method to generate high-affinity DNA aptamers containing a hydrophobic **Ds** base as a fifth base [15]. We obtained anti-VEGF165 aptamer (47-mer) containing two **Ds** bases and anti-IFNγ aptamer (49-mer) containing three **Ds** bases. Their binding affinities, the K_d values, to each target protein were 0.65 pM and 38 pM, respectively, which were more than one hundred times smaller than those of the existing DNA aptamers containing natural bases only. In addition, the binding abilities were largely dependent on the **Ds** bases: replacement of the **Ds** bases to the natural A bases in the aptamer significantly reduced their binding affinities, indicating the **Ds** bases actually involves their improved binding abilities. This was the first example that increasing the number of the components of nucleic acids, by adding the hydrophobic **Ds** bases to four natural bases, significantly augmented the functionalities of nucleic acids.

References: [1] Crick F. H. C. (1968) *J. Mol. Biol., 38,* 367–379. [2] Gesteland R. F. et al. (eds) *The RNA World* 2nd edn (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1999). [3] Rogers J. and Joyce G. F. (1999) *Nature, 402,* 323–325. [4] Reader J. S. and Joyce G. F. (2002) *Nature, 420,* 841–844. [5] Gardner P. P. et al. (2003) *Proc. Biol. Sci. 270,* 1177–1182. [6] Rich A. in *Horizions in Biochemistry* (eds Kasha, M. and Pullman, B.) 103–126 (Academic, New York, 1962). [7] Switzer C. et al (1989) *J. Am. Chem. Soc. 111,* 8322–8323. [8] Rappaport H. P. (1988) *Nucleic Acids Res. 16,* 7253–7267. [9] Yang Z. et al. (2011) *J. Am. Chem. Soc,. 133,* 15105–15112. [10] Malyshev D. A. et al. (2012) *Proc. Natl. Acad. Sci. U S A. 109,* 12005–12010. [11] Kimoto M. et al. (2009) *Nucleic Acids Res. 37*, e14. [12] Yamashige R. et al. (2012) *Nucleic Acids Res. 40,* 2793–2806. [13] Hirao I. and Kimoto M. (2012) *Proc. Jpn. Acad., Ser. B, Phys. Biol. Sci. 88*, 3453–3467. [14] Hirao I. et al. (2012) *Acc. Chem. Res., 45,* 2055–2065. [15] Kimoto M. et al. (2013) *Nat. Biotechnol., 31,* 453–457.