

Y. S. Kim, *et al.* [2009] *Med. Hypotheses Res.* 5: 1–18.

Genetic Code is Composed of Pyrimidine-Purine DNA Segments as a Basic Unit Displaying DNA Base Pair Polarity

Yeon Sook Kim^{4,5}, Dae Sung Lee⁴, Dae Gwan Lee⁴, Suk Keun Lee^{*1,4}, Sun Jin Choi⁴, In Sun Song¹, Sang Shin Lee¹, Young Joon Lee¹, Yun Woo Lee⁴, Sang Chul Park² and Je Geun Chi³

¹Department of Oral Pathology, College of Dentistry, Kangnung National University, Gangneung, Korea; ²Department of Biochemistry, and ³Department of Pathology, Seoul National University College of Medicine, Seoul, 110-799, Korea; ⁴Institute of Molecular Genetics, Gangneung, Korea; Department of dental Hygiene, College of Health, Cheongju University, Cheongju, Korea.

Abstract. DNA base pair hybridization composed of a pyrimidine (Py) and a purine (Pu) shows characteristic electrostatic polarity by hydrogen bonds between Py and Pu. As this electrostatic charge of hydrogen bond is not neutralized between hydrogen donor and acceptor, it can induce a base pair polarity of each nucleotide, and also affect a polarity from the 5' to 3' direction in each DNA strand. Although many Py and Pu sequences are arranged in DNA duplex in a complex manner, we hypothesize that a Py_nPu_n sequence is a basic pyrimidine-purine DNA segment (Pyu DNA segment), contrast to the Pu_nPy_n sequence (the Puy DNA segment). A series of short DNA segments, 6-30 bp different pairs of synthesized oligonucleotides, were analyzed by EtdBr-intercalated electrophoresis assay and reverse phase HPLC. The annealing of primers composed of Pyu or Puy sequences were also evaluated by the optimized PCR. Resultantly, the DNAs composed of Pyu DNA segment were more strongly intercalated by EtdBr than those composed of Puy DNA segment. And more, in PCR the primers composed of Pyu sequences produced more object DNA than those composed of Puy sequences. Here, we show that the Pyu DNA segment is a basic unit of the genetic code, which can produce more specific hybridization and more stable DNA duplex than Puy DNA segment.

Correspondence: Dr. Suk Keun Lee, Department of Oral Pathology, College of Dentistry, Kangnung National University, Gangneung Korea. Tel: +82 33-640-2228. Fax: +82 33-642-6410, E-mail: sklee@mail.kangnung.ac.kr

Abbreviations used: Py, pyrimidine; Pu, purine; Pyu, pyrimidine-purine; Puy, purine-pyrimidine; PCR, polymerase chain reaction; NMR, nuclear magnetic resonance.

Received on 04-10-2008; accepted on 05-18-2008.

1. Introduction

The genetic code containing all of the information for the viability and replication of cells uses only four nucleotides [1, 2]. However, the DNA chain has polarity in the 5'→3' direction, and using NMR, a family of self-complementary DNA decamers containing single alpha-anomeric nucleotides flanked by 3'-3' and 5'-5' phosphodiester linkages showed characteristic polarity that affected the structural and dynamic properties of these sequences [3]. Direct evidence of DNA polarity was also found in the three-dimensional structure of the hairpin formed by d(ATCCTA)_n-d(GTTA)_n- d(TAGGAT)_n by means of two-dimensional NMR, distance geometry and molecular dynamic calculations [3, 4]. The data indicate that the influence of polarity of a closing dA:dT pair is much less significant than that of a closing dC:dG base pair [5]. DNA polarity has also been reported at the binding sites for different proteins (enzymes), *i.e.*, *Escherichia coli* Rep helicase [6], yeast TATA box binding protein (yTBP) [7, 8], XerC, XerD, and FtsK proteins [9]. However, the molecular dielectrostatic polarity between donor and acceptor of hydrogen bond was not clearly defined up to date.

Understanding the DNA functions requires knowledge of the structure of local sequence-dependent conformations. Recent studies have shown that simple repetitive d(CA/TG) dinucleotide sequences adopt a left-handed non-B-DNA structure under negative superhelical stress [10]. Although the pattern of chemical reactivities and the helical parameters observed for these sequences differ significantly from those of standard Z-DNA, the supercoil-induced conformation in d(CA/TG)_n sequences mimics the left-handed Z-DNA [10]. In *Escherichia coli*, d(GC)_n repetitive sequences have been shown to be deletion prone, but the precise mechanism of this mutagenic pathway is still unknown. Interrupting the repeat of d(GC)_n with either a d(AT) or a d(GT) dinucleotide decreases the rate of deletion within these sequences, and introducing pyrimidine-purine alternating sequences beside the d(GC) insert results in an increased rate of deletion [11]. Instability within tracts of repeti-

tive sequences, *i.e.*, CpG islands, has recently been associated with several genetic disorders, including so-called triplet repeat diseases and hypermethylation in certain forms of colorectal cancers [12-14]. All these findings may indicate the existence of DNA base pair polarity, which occurs in a sequence-specific manner as an important signal for the functional elements of genetic codes. In this study, we hypothesize that the genetic code is organized in specially polarized patterns for its stability and function, and then a concept of DNA base pair polarity is demonstrated to develop a new method of gene analysis.

2. Materials and Methods

2.1. Ethidium bromide intercalating electrophoresis assay

To determine the difference of hybridization strength between P_y_nP_u_n•P_y_nP_u_n and P_u_n•P_u_nP_y_n (n = 3, 4, 5, 6, 7, 8, 10, 12) a unique experiment of DNA competition was done by an ethidium bromide (EtdBr) intercalating electrophoresis assay. Because the EtdBr migrates in the opposite direction from dsDNA in electrophoresis, and then only the strongly intercalated EtdBr in the DNA duplex remains after electrophoresis, contrast to the loosely intercalated or stacked EtdBr. The remained EtdBr can be detected by UV illumination. It is clear that more EtdBr are intercalated into the DNA duplex hybridized more strongly (Supplementary FIG. 2). This phenomenon is similar to the more intercalation of EtdBr to the supercoiled DNA compared to the B-type DNA [34]. However, EtdBr is strongly intercalated only into the DNA duplex, and is to be strongly bound to CG intercalation sites (Pyu sequence), *i.e.*, CGCG (2 CG sites) and CCGG (1 CG site), but weakly bound to GGCC (no CG site) [35]. The typical P_y_nP_u_n•P_y_nP_u_n and P_u_n•P_u_nP_y_n sequences, C_nA_n•T_nG_n and A_nC_n•G_nT_n (n = 3, 4, 5, 6, 7, 8, 10, 12), respectively, are not self-complementary in buffer solution. Each pairs of 10 μL oligonucleotides (1 μM) were dissolved in 0.2 M NaCl solution and subsequently hybridized by heating up to 90°C for 10 min, and then slowly cooled at room tempera-

ture. The hybridized DNAs were mixed with loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol, 30% glycerol in water containing EtdBr (10 $\mu\text{g}/\text{mL}$) and incubated for 10 min in room temperature to be intercalated by EtdBr. Then the DNA sample was electrophoresed using non-denaturing gel electrophoresis (20%; 19:1 acrylamide/bis- acrylamide) at 120 V in TBE buffer (89 mM Tris-borate, 2 mM EDTA, pH 8.0) for 60 min. The different pairs of oligonucleotides, *i.e.*, $C_nA_n \bullet T_nG_n$ and $T_nA_n \bullet T_nA_n$ ($n = 3, 4, 5, 6, 7, 8, 10, 12$), $T_nA_n \bullet T_nA_n$ and $A_nT_n \bullet A_nT_n$ ($n = 3, 4, 5, 6, 7, 8, 10, 12$), $GGG(T_nA_n)_{12/n}CCC$ ($n = 1, 2, 3, 4, 6, 12$), and $GGG(C_nA_n)_{12/n}GGG$ and $CCC(T_nG_n)_{12/n}CCC$ ($n = 1, 2, 3, 4, 6, 12$), were also applied in the same procedure for the EtdBr intercalating electrophoresis assay.

2.2. PCR amplification using different oligonucleotide primers

As the strength of DNA hybridization is directly affect the annealing of oligonucleotide primer to the template DNA in PCR, the PCR amplification was performed using each oligonucleotide of $(C_nA_n)_{12/n}$ and $(T_nG_n)_{12/n}$ ($n = 1, 2, 3, 4, 6, 12$) which did not form a self complementary hybridization. The PCR amplification until the exponential stage, 25 cycles, with optimized annealing temperature was carried out to elucidate the annealing efficiency of differently polarized oligonucleotide primers. Six Pyu DNA segments composed of $(C_nA_n)_{12/n} \bullet (T_nG_n)_{12/n}$ ($n = 1, 2, 3, 4, 6, 12$) were bidirectionally subcloned into pBluescript II KS+ (Stratagene, USA) through a EcoRI linker site. A Puy DNA segment composed of $A_{12}C_{12} \bullet G_{12}T_{12}$ was also bidirectionally subcloned in order to compare with $C_{12}A_{12} \bullet T_{12}G_{12}$ for their contrast DNA base pair polarity. These subcloned plasmids were purified and used as a template, and each insert DNA was used as a primer for the PCR. The annealing temperature of C_nA_n and T_nG_n ($n = 1, 2, 3, 4, 6, 12$) was 57 and 59°C, respectively, calculated by a DNA tool program (<http://biology.semo.edu/cgi-bin/dnatools.pl>, Southeast Missouri State University, USA). The experimental PCRs were done simultaneously in the optimized condition, briefly, annealing at 57°C for the C_nA_n primers or

at 59°C for the T_nG_n primers for 30 seconds, elongation at 72°C for 30 seconds, denaturation at 95°C for 30 seconds, and 25 cycles of PCR. However, the PCR was repeatedly performed preliminarily using primer pairs between vector and insert sequences to obtain the proper PCR condition producing the expected DNA consistently.

2.3. Reverse phase HPLC analysis

In order to elucidate the conformational change of each DNA segment the reverse phase HPLC analysis (Agilent Chemstation, Agilent Technologies, USA) using ODS-AQ column (YMC, Waters Corp.) under 0.2 M NaCl solution, flow rate of 0.5 ml/min at 260 nm, was performed using the double strand oligodeoxynucleotides (dsDNAs) of $A_{12} \bullet T_{12}$, $A_6T_6 \bullet A_6T_6$, and $T_6A_6 \bullet T_6A_6$, which contained same number of dA and dT, and $G_{12} \bullet C_{12}$, $G_6C_6 \bullet G_6C_6$, and $C_6G_6 \bullet C_6G_6$, which contained same number of dG and dC.

3. Results and Discussion

3.1. Pyu DNA segment defined by DNA base pair polarity hypothesis

The hypothesis of DNA base pair polarity arises from the polarized electrostatic status of hydrogen bonds between a pyrimidine (Py) and a purine (Pu), *i.e.*, thymine:adenine (dT:dA) and cytosine:guanine (dC:dG), when each hybridization occurs (FIG. 1). Regarding to the electrostatic charges of each base pairs, the pairs of dT:dA and dC:dG showed the characteristic electric status of intermolecular hydrogen bonds. Although each of thymine and adenine has one hydrogen donor and one hydrogen acceptor for two hydrogen bonds between them, the purine ring of adenine can absorb the electron elicited by a hydrogen acceptor (N1). Thus the dT:dA pair can be polarized by a single hydrogen bond between oxygen at thymine C4 and amine (NH₂) at adenine C6, producing $-e$ charge in thymine. While the dC:dG pair is formed by three hydrogen bonds, one hydrogen donor from cytosine and two hydrogen donor from guanine. The hy-

Fig. 1

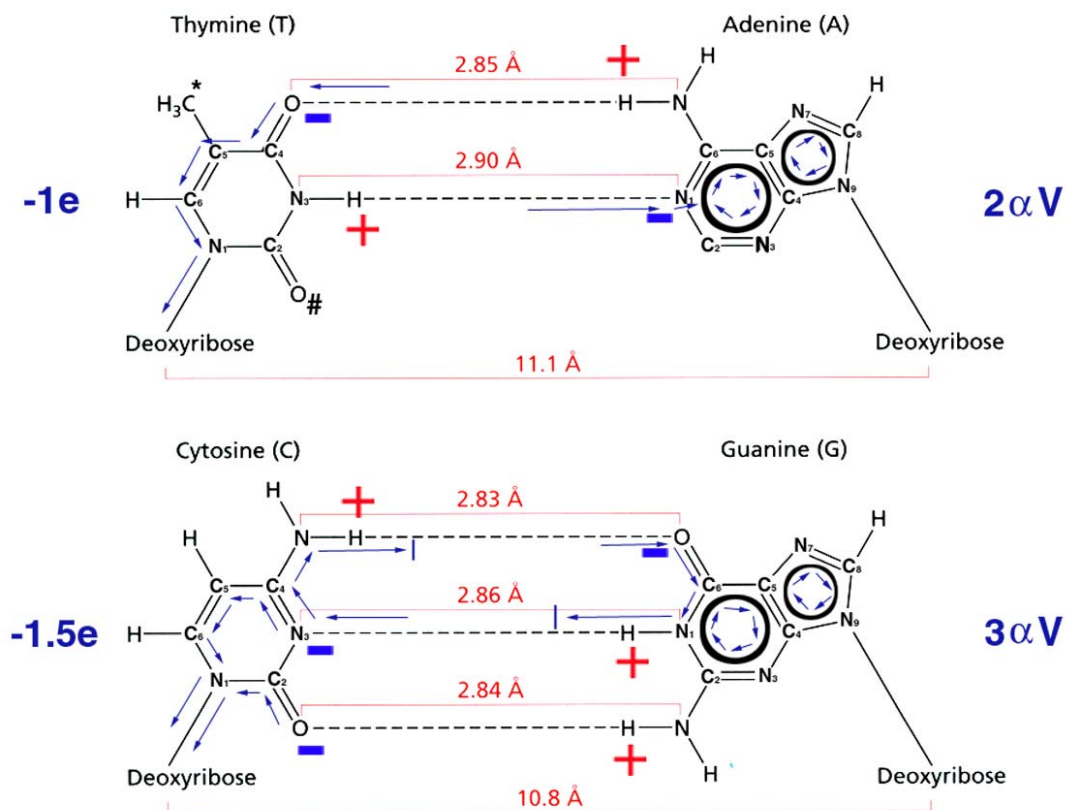


Figure 1. Schemes for the putative induction of electrostatic charge between thymine and adenine, and between cytosine and guanine. The intramolecular electric circuit is illustrated by blue arrows. *Methyl group increases the hydrophobicity of pyrimidine ring in thymine. # Carbonyl group can maintain the negative electrostatic charge of thymine.

drogen donor, an amine at cytosine C₄, can induce $-e$ charge at oxygen of guanine C₆, but almost absorbed into the purine ring of guanine, but the elicited $-e$ charge at cytosine N₃ is partly transferred into the empty space of the hydrogen donor at cytosine C₄ and also transferred into deoxyribose through its pyrimidine ring. Because there is no electron transfer in the hydrogen bond, the loss and gain of electrostatic charge should be antagonistic between the two sites of the hydrogen bond, resulted in the polarized charge of about $-1/2e$ in cytosine. Furthermore the hydrogen donor, an amine at guanine C₂ can induce $-e$ charge at the oxygen of cytosine C₂, which is directly transferred into the deoxyribose. Thus, the dC:dG pair is electrostatically

charged about $-1.5e$ in a cytosine. For the calculation of base pair polarity energy we simplified the voltage of dT:dA pair as $2\alpha V$ and the voltage of dC:dG as $3\alpha V$ (FIG. 1).

On the other hand, although the ribose-phosphate backbone is same in every nucleotide, the oxygen of the ribose ring partly inhibits electric flow. Thus, the electric charge elicited from a pyrimidine base can easily flow through C₃ of ribose and phosphate diester, producing electric flow from the 5' to 3' direction. In another sense, the ribose ring acts like a diode, inducing 5'-3' polarity, and the transferred electron is accumulated in the phosphate group, acting as a condenser. Between the ribose ring and the phosphate group a CH₂ group is hydrophobic and

Fig. 2

Scheme of potential electric flow in a deoxynucleotide

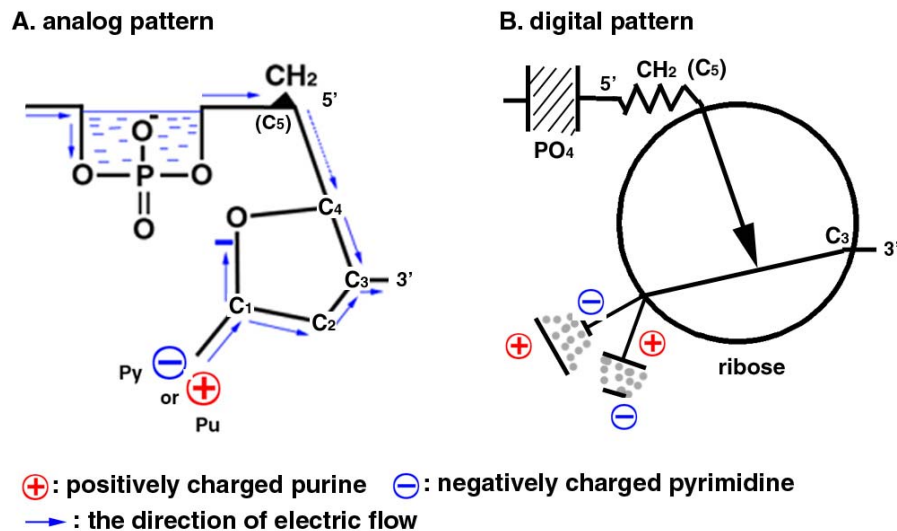


Figure 2. Scheme of intramolecular electric status of a deoxynucleotide and its potential electric flow.

relatively resistant to the electric flow. Therefore, the ribose-phosphate backbone of a nucleotide can be illustrated as a simple electric circuit composed of a diode (ribose ring), a condenser (phosphate group), and a resistance (CH₂ group) (FIG. 2).

From the electric circuit pattern of nucleotide shown in FIG. 2 a single strand of DNA duplex, for example 5'-TCAGC-3', can be illustrated as FIG. 3. The 5'-TCAGC-3' has one Pyu DNA segment, 5'-TCAG-3' and an overhanging sequence of C charged negatively. The complementary sequence of 5'-TCAG-3' will be a 5'-CTGA-3, which also forms a Pyu DNA segment (FIG. 3).

The above hypothetical concept of DNA base pair polarity, the Pyu DNA segment, and the DNA intramolecular electric flow in DNA duplex can provide the mathematical calculation of DNA hybridization strength energy as shown in the supplementary data.

3.2. Symbolization of DNA base pair polarity

From the concept of base pair polarity, we

symbolize the alphabet of A, C, G, and T deoxynucleotides into the polarity signals, *i.e.*, a pyrimidine is symbolized as a lower semicircle in blue under the horizontal line, and a purine is symbolized as an upper semicircle in red above horizontal line. A and T can be distinguished by empty semicircles in contrast to filled semicircles for G and C (FIG. 4). A representative feature of DNA base pair polarity program was demonstrated in the supplementary data (Supplementary FIG. 1).

As a DNA strand is polarized from the 5' to 3' direction according to the common mode of transcription from the 5' to 3' direction, the electrostatic charge of each base pair can easily run from the 5' to 3' direction through the deoxyribose-phosphate axis. The negative charge arising from Py goes readily into the 3' direction when a Pu is next. Whereas the positive charge arising from Pu is unlikely to go in the 3' direction when Py is arranged next. If pyrimidines (Pys) were arranged in a repeating fashion, the negative charges would be continuously accumulated in the phosphate groups and work as a negative-

Fig. 3

Scheme of electric flow in a strand of DNA duplex and its DNA base pair polarity

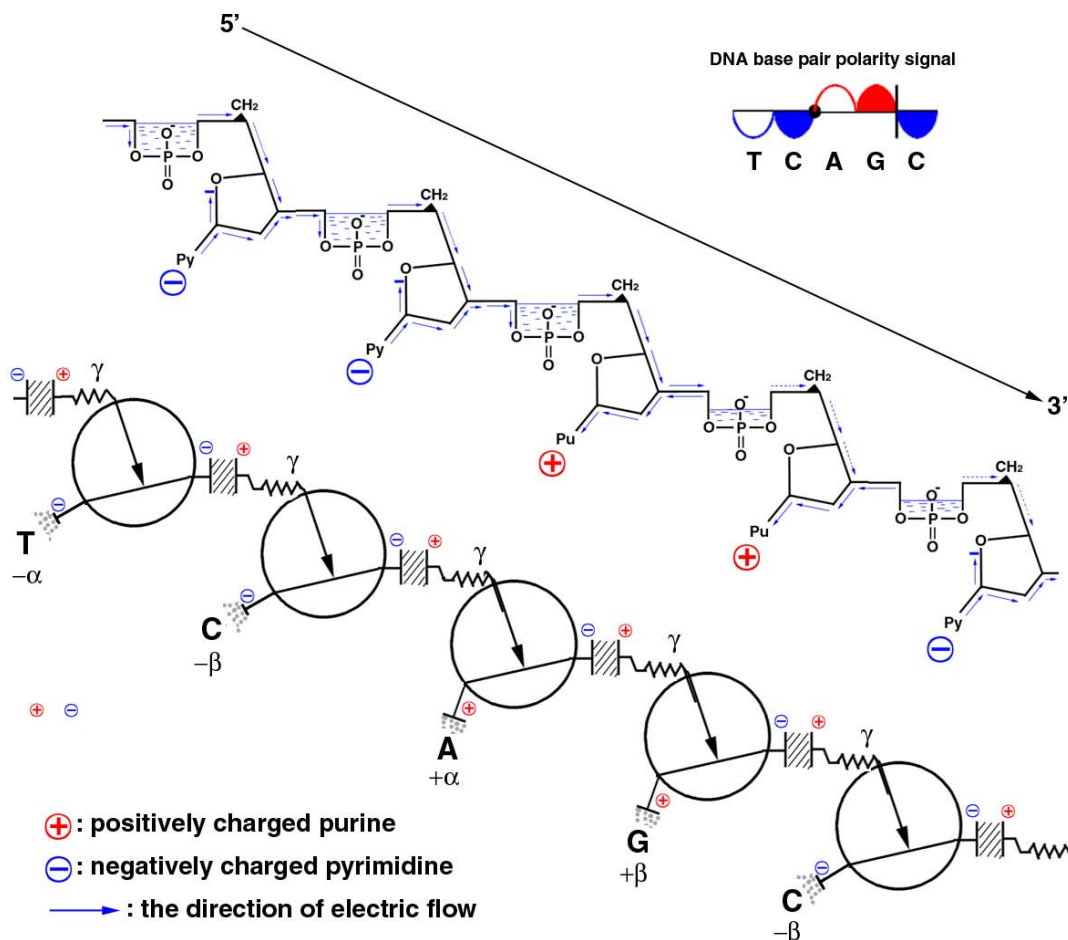


Figure 3. Scheme of electric flow in a strand of DNA duplex and its DNA base pair polarity.

charge group. And then if purines (Pus) were located after Pys, the accumulated electrons would overflow to the empty phosphate groups of Pus, which are positively charged. On the other hand, if Pys were located after Pus, no electric flow would be expected between the former and the latter due to the strand polarity threshold of deoxyribose-phosphate axis from 5' to 3' direction. Therefore, every border between Pu and Py becomes a limiting boundary to resist the electric flow of the DNA polarity, resulted in a division for a Pyu DNA segment, Py_nPu_n (n is a natural number). In the Pyu DNA segment the electric

flow is unique, centered at the border between Py and Pu, named a central point of DNA segment. Thus, it is presumed that their molecular interactions between DNA base pairs are grouped and their polarity charges are reciprocally amplified between both strands of the DNA segment.

3.3. Hybridization of $C_nA_n \bullet T_nG_n$ and $T_nA_n \bullet T_nA_n$ ($n = 3, 4, 5, 6, 7, 8, 10, 12$)

$C_nA_n \bullet T_nG_n$ and $T_nA_n \bullet T_nA_n$ ($n = 3, 4, 5, 6, 7, 8, 10, 12$) showed higher efficiency of EtdBr in-

Fig. 4

Scheme of DNA base pair polarity

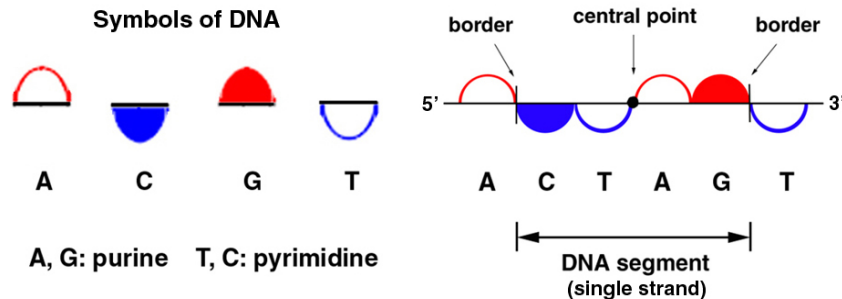


Figure 4. A scheme of DNA base pair polarity illustrating the characteristic symbols for the polarity of each nucleotide (left) and a single strand of a Pyu DNA segment (right).

tercalation than the $A_n C_n \bullet G_n T_n$ and $A_n T_n \bullet A_n T_n$ ($n = 3, 4, 5, 6, 7, 8, 10, 12$), respectively, as the number n increased. $C_n A_n \bullet T_n G_n$ and $A_n C_n \bullet G_n T_n$ ($n = 3, 4, 5, 6$) showed too weak signal of EtdBr intercalation in the gel, but $C_n A_n \bullet T_n G_n$ and $A_n C_n \bullet G_n T_n$ ($n = 7, 8, 10, 12$) showed comparable levels of EtdBr intercalation. $C_7 A_7 \bullet T_7 G_7$ showed higher EtdBr intercalation than $A_7 C_7 \bullet G_7 T_7$ by 90%, $C_8 A_8 \bullet T_8 G_8$ showed higher EtdBr intercalation than $A_8 C_8 \bullet G_8 T_8$ by 68.9%, $C_{10} A_{10} \bullet T_{10} G_{10}$ showed higher EtdBr intercalation than $A_{10} C_{10} \bullet G_{10} T_{10}$ by 25%, and $C_{12} A_{12} \bullet T_{12} G_{12}$ showed higher EtdBr intercalation than $A_{12} C_{12} \bullet G_{12} T_{12}$ by 6% (FIG. 5).

3.4. Hybridization of $T_n A_n \bullet T_n A_n$ and $A_n T_n \bullet A_n T_n$ ($n = 3, 4, 5, 6, 7, 8, 10, 12$)

$T_n A_n \bullet T_n A_n$ and $A_n T_n \bullet A_n T_n$ ($n = 3, 4, 5, 6, 7, 8, 10, 12$) are typical Pyu and Puy DNA segments, respectively, which are self-complementary in buffer solution. The same concentration of each $T_n A_n \bullet T_n A_n$ and $A_n T_n \bullet A_n T_n$ was dissolved in 0.2 M NaCl solution and heated at 90°C for 10 min, and slowly cooled in room temperature to produce a DNA duplex of the $T_n A_n \bullet T_n A_n$ and $A_n T_n \bullet A_n T_n$, respectively. These DNA samples were also examined by the same EtdBr intercalation electrophoresis assay as the above. In the competition between the self complementary monomeric hybridization and the self comple-

mentary dimeric hybridization of $T_n A_n$ or $A_n T_n$ oligonucleotides ($n = 3, 4, 5, 6, 7, 8, 10, 12$), the oligonucleotides which have strong monomeric hybridization will less likely form the dimeric hybridization. The self complementary monomeric hybridization is composed of base pairs with opposite polarity, but because it has a same ribose-phosphate axis, the opposite electrostatic charges of sense and antisense strands can be easily neutralized. However, the retained EtdBr intercalation after electrophoresis was not observed in short DNAs, $T_n A_n \bullet T_n A_n$ ($n = 3, 4, 5, 6, 7, 8$) and $A_n T_n \bullet A_n T_n$ ($n = 3, 4, 5, 6, 7$), but as the DNA size increased, it appeared from $T_{10} A_{10} \bullet T_{10} A_{10}$ of the $T_n A_n \bullet T_n A_n$ group, while appeared from $A_8 T_8 \bullet A_8 T_8$ of the $A_n T_n \bullet A_n T_n$ group. Although the retained EtdBr intercalation after electrophoresis was too weak to detect in the gel, the additional EtdBr staining after electrophoresis disclosed the dimeric hybridization occurred from $T_{10} A_{10} \bullet T_{10} A_{10}$ in the $T_n A_n \bullet T_n A_n$ group, while from $A_6 T_6 \bullet A_6 T_6$ in the $A_n T_n \bullet A_n T_n$ group, as the number of n increased. The monomeric hybridization was more strongly intercalated by EtdBr than the dimeric hybridization until $T_8 A_8 \bullet T_8 A_8$ in the $T_n A_n \bullet T_n A_n$ group, while until $A_5 T_5 \bullet A_5 T_5$ group in the $A_n T_n \bullet A_n T_n$ group. And the dimeric hybridization of $T_n A_n \bullet T_n A_n$ ($n = 10, 12$) showed stronger EtdBr intercalation than the dimeric hybridization of $A_n T_n \bullet A_n T_n$ ($n = 6, 7, 8, 10, 12$) (FIG. 6). These findings suggest that as

Fig. 5

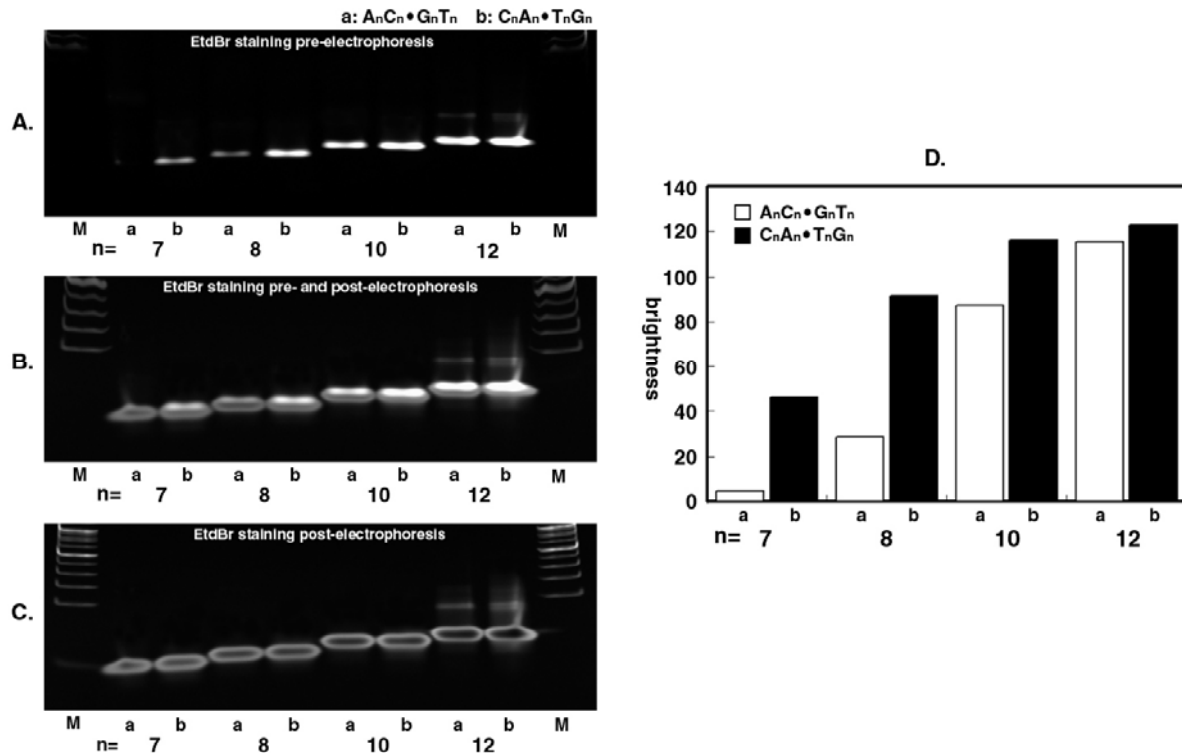


Figure 5. Hybridization of $A_n C_n \bullet G_n T_n$ and $C_n A_n \bullet T_n G_n$ oligonucleotides ($n = 7, 8, 10, 12$) detected by EtdBr intercalation. A: EtdBr staining pre-electrophoresis, B: EtdBr staining pre- and post-electrophoresis, C: EtdBr staining post-electrophoresis, D: a graph for the comparisons of EtdBr intercalation between $A_n C_n \bullet G_n T_n$ group (a) and $C_n A_n \bullet T_n G_n$ group (b) from panel A. *This figure is a representative one from the repeated experiments more than three times.

the EtdBr intercalation is stronger in the $T_n A_n \bullet T_n A_n$ group than in the $A_n T_n \bullet A_n T_n$ group ($n = 3, 4, 5, 6, 7, 8, 10, 12$), both of the monomeric and dimeric hybridization of the $T_n A_n \bullet T_n A_n$ are more strongly interacted with EtdBr than those of the $A_n T_n \bullet A_n T_n$.

3.5. Self complementary hybridization of $GGG(T_n A_n)_{12/n}CCC$ ($n = 1, 2, 3, 4, 6, 12$)

In the self-complementary hybridization experiment using $GGG(T_n A_n)_{12/n}CCC$ ($n = 1, 2, 3, 4, 6, 12$) oligonucleotides, which have all the same number of nucleotides, 30 nt, and all the same kinds of deoxynucleotides, *i.e.*, 3dG, 3dC, 12dT, and 12dA. The short Pyu DNA segment, $GGG(T_n A_n)_{12/n}CCC$ ($n = 1$ and 2) showed higher level of monomeric hybridization than the

dimeric hybridization. This phenomenon is gradually shifted to greater dimeric hybridization by increasing heat treatment from 20°C to 90°C. The monomeric hybridization of $T_n A_n$ ($n = 1, 2$) was almost overwhelmed by the dimeric hybridization when the temperature increased to 90°C (FIG. 7). Therefore, it is presumed that the competition between the monomeric and dimeric hybridization is thermodynamic and directly affected by the base pair polarity of DNA segment. However, as the size of DNA segment increased, $T_n A_n$ (n is from 3 to 12), their dimeric hybridization became more prominent than the monomeric hybridization, because the bigger DNA segment had more polarized nucleotides strong enough to produce dimeric hybridization. This result may indicate that the significant differences of hybridization exist depend on the

Fig. 6

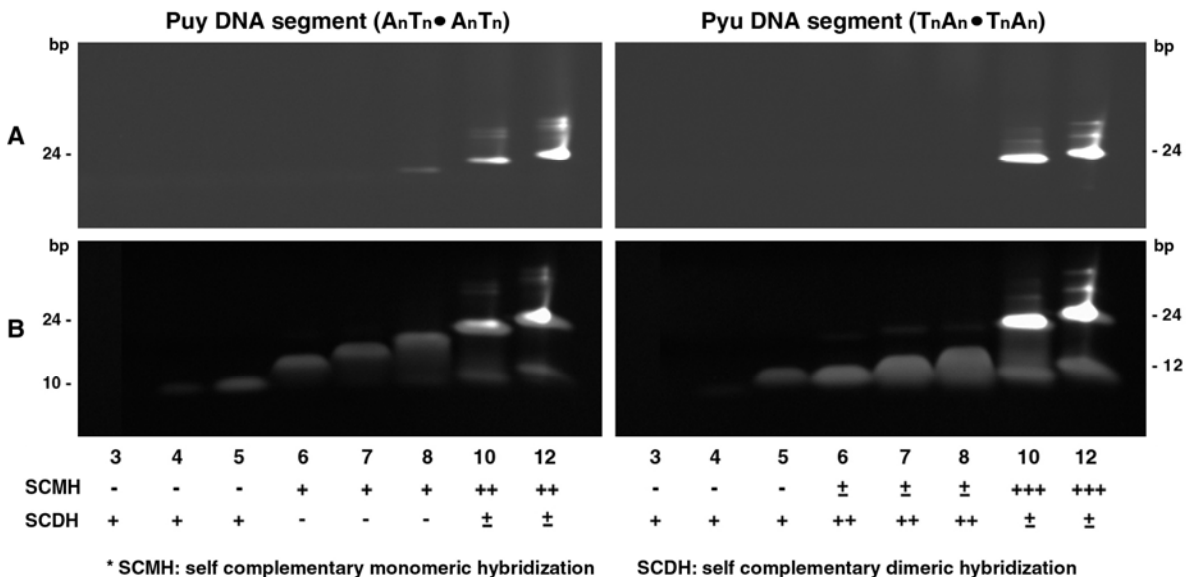


Figure 6. Hybridization of $A_nT_n \bullet A_nT_n$ and $T_nA_n \bullet T_nA_n$ oligonucleotides ($n = 3, 4, 5, 6, 7, 8, 10, 12$) detected by EtdBr intercalation. A: EtdBr staining pre-electrophoresis, B: EtdBr staining post-electrophoresis. SCDH: self-complementary dimeric hybridization, SCM: self-complementary monomeric hybridization. *This figure is a representative one from the repeated experiments more than three times.

number of Pyu DNA segment. However, both of the self-complementary monomeric and dimeric hybridizations are all made by hydrogen bonds interacting between nucleotides, and may function collaterally in a DNA duplex.

When the hybridization of $GGG(T_nA_n)_{12/n}CCC \bullet GGG(T_nA_n)_{12/n}CCC$ ($n = 1, 2, 3, 4, 6, 12$) pretreated at 70°C for 10 min was examined, it became apparent that as the number of n increased the dimeric hybridization increased and became maximum in $GGG(T_4A_4)_3CCC \bullet GGG(T_4A_4)_3CCC$, and thereafter the dimeric hybridization decreased greatly. Contrast to the hybridization experiment of $T_nA_n \bullet T_nA_n$ ($n = 3, 4, 5, 6, 7, 8, 10, 12$), disclosing the increased dimeric hybridization as the number of n increased (FIG. 1 and 2), the dimeric hybridization of $GGG(T_nA_n)_{12/n}CCC \bullet GGG(T_nA_n)_{12/n}CCC$ was affected by the number of Pyu DNA segment, $(T_nA_n)_{12/n}$, resulted in the maximum in three Pyu DNA segments, $(T_4A_4)_3$, contrast to a single Pyu DNA segment, $T_{12}A_{12}$ (FIG. 8). This phenomenon may indicate that the multiple DNA segments may have strong influence in the

formation of dimeric hybridization as well as the increased size of DNA segment. This finding is clear in a graph of FIG. 8C plotted by the EtdBr-intercalated bands of the dimeric hybridization, which is almost identical to the graph based on the calculation by the DINAMelt server (Rensselaer Polytechnic Institute) programmed for dimeric hybridization (FIG. 8E). However, when the sum of the self-complementary monomeric and dimeric hybridization was plotted, the maximum EtdBr intercalation was found in $GGG(TA)_{12}CCC \bullet GGG(TA)_{12}CCC$ ($n = 1$), and then the EtdBr intercalation was gradually reduced as the number of n increased (FIG. 8D). This graph is similar to the plot by the total DNA hybridization energy of DNA segments based on the DNA base pair polarity of this study (FIG. 8E). Thus, we think the DNA hybridization energy obtained from the DNA base pair polarity program may represent the whole hybridization strength, while the calculation by the DINAMelt server program may show only the value for the dimeric hybridization.

Fig. 7

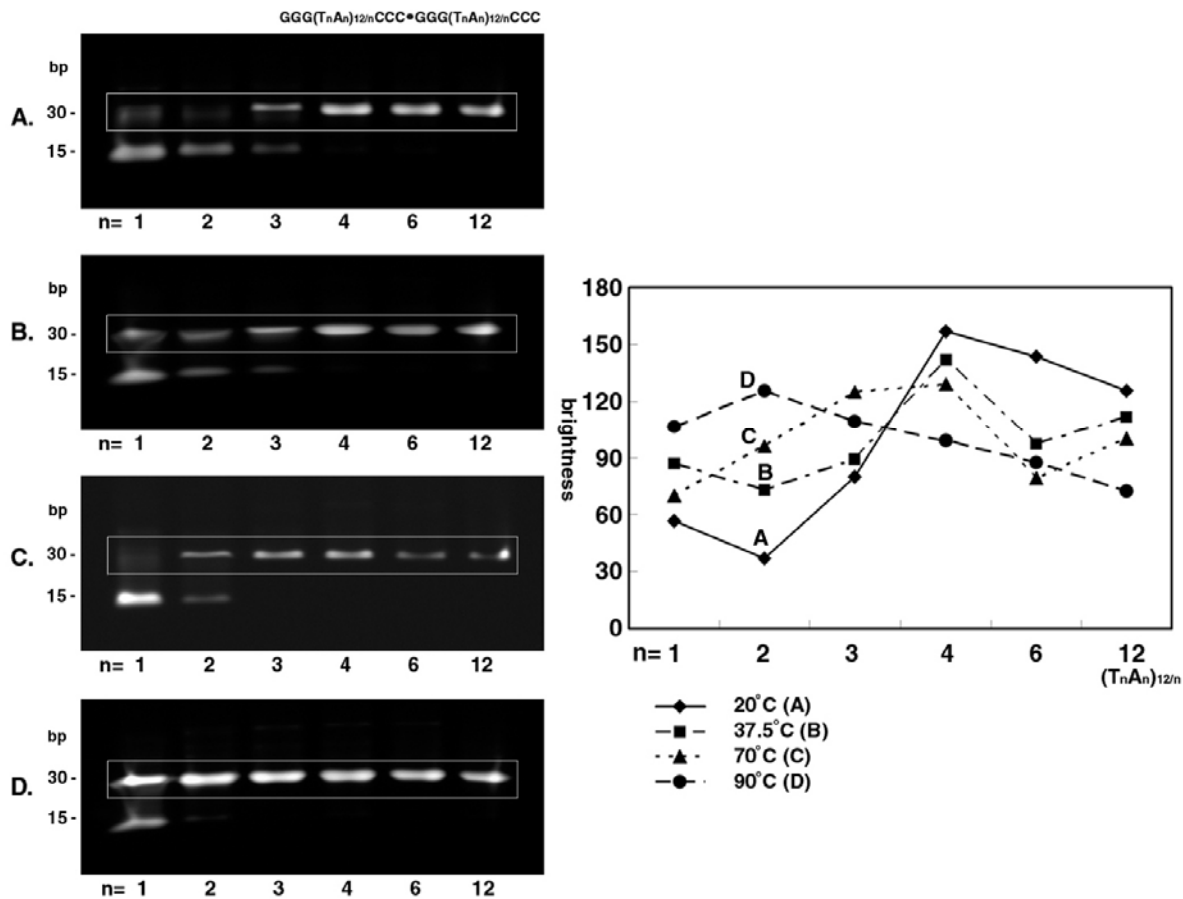


Figure 7. Hybridization of $GGG(T_nA_n)_{12/n}CCC$ and $GGG(T_nA_n)_{12/n}CCC$ oligonucleotides ($n = 1, 2, 3, 4, 6, 12$) depend on heat treatment increased from $20^\circ C$ to $90^\circ C$, then the self-complementary monomeric hybridization reduced while the self-complementary dimeric hybridization increased. Heat treatment at $20^\circ C$ (A), $37.5^\circ C$ (B), $70^\circ C$ (C), and $90^\circ C$ (D). *This figure is a representative one from the repeated experiments more than three times.

3.6. Hybridization of $GGG(C_nA_n)_{12/n}GGG$ and $CCC(T_nG_n)_{12/n}CCC$ oligonucleotides ($n = 1, 2, 3, 4, 6, 12$)

The hybridization between $GGG(C_nA_n)_{12/n}$ and GGG and $CCC(T_nG_n)_{12/n}$ and CCC oligonucleotides ($n = 1, 2, 3, 4, 6, 12$) showed the heterodimeric hybridization without self-complementary hybridization. The EtdBr intercalation retained after electrophoresis at 120V for 60 min in TBE buffer was most intense in $GGG(CA)_{12}GGG \bullet CCC(TG)_{12}CCC$ and gradually decreased until in $GGG(C_{12}A_{12})GGG \bullet CCC(T_{12}G_{12})CCC$ (FIG. 9A and C). Whereas the total dimeric hybridization

stained with EtdBr post-electrophoresis was almost the same in all $GGG(C_nA_n)_{12/n}GGG \bullet CCC(T_nG_n)_{12/n}CCC$ ($n = 1, 2, 3, 4, 6, 12$) (FIG. 9B and D). However, the 30 base pair DNA composed of the shortest DNA segment or the greatest number of DNA segments, $GGG(CA)_{12}GGG \bullet CCC(TG)_{12}CCC$, shows the strongest intercalation of EtdBr compared to the others, $GGG(C_nA_n)_{12/n}GGG \bullet CCC(T_nG_n)_{12/n}CCC$ ($n = 2, 3, 4, 6, 12$) (FIG. 9A and C). While the ensemble enthalpy (ΔH) of $GGG(C_nA_n)_{12/n}GGG \bullet CCC(T_nG_n)_{12/n}CCC$ ($n = 1, 2, 3, 4, 6, 12$), calculated as $\Delta H = \Delta G - TdG/dT$ using the DINAMelt server (Rensselaer Polytechnic Institute) was plotted

Fig. 8

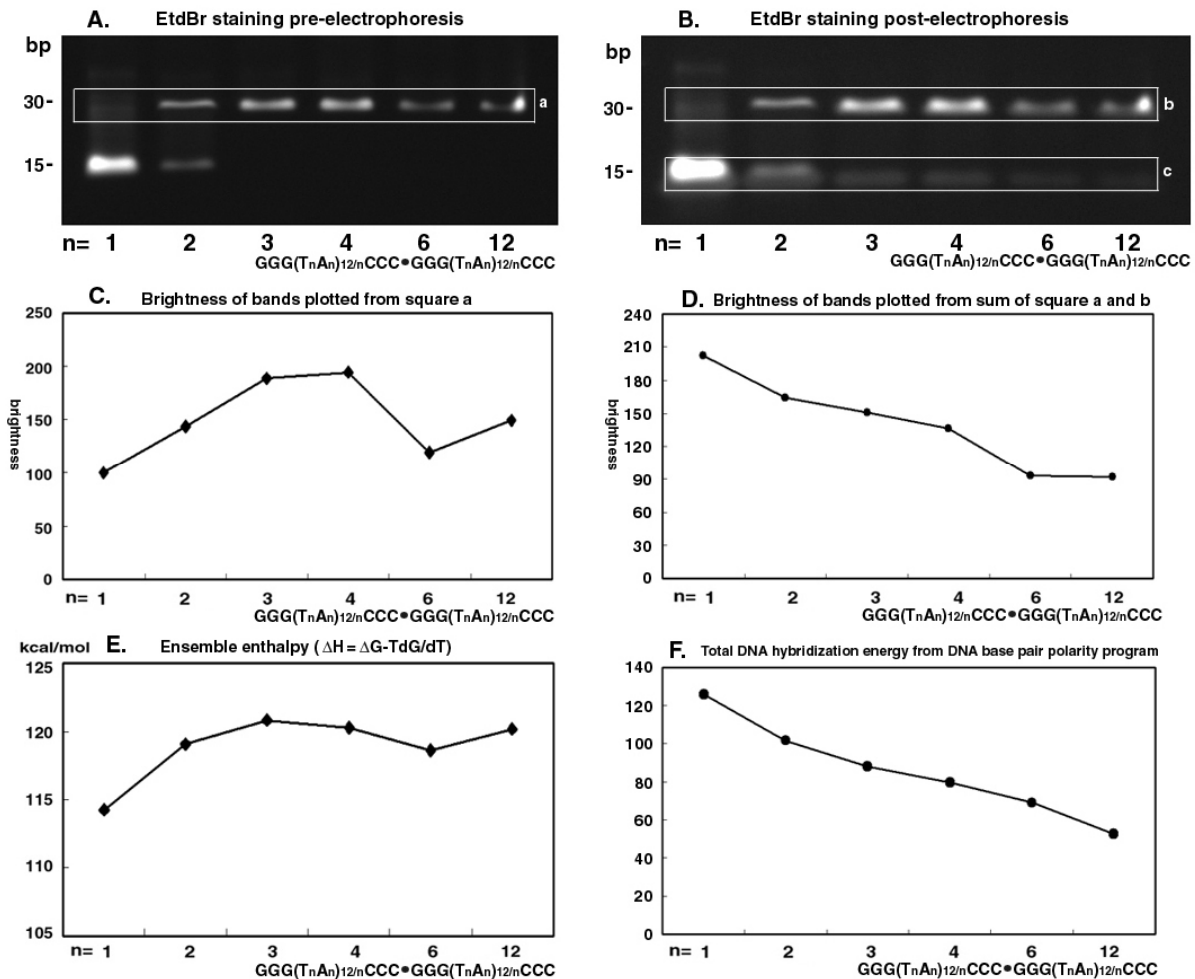


Figure 8. Hybridization of $GGG(T_nA_n)_{12/n}CCC$ and $GGG(T_nA_n)_{12/n}CCC$ oligonucleotides ($n = 1, 2, 3, 4, 6, 12$) and EtdBr intercalation. A: EtdBr staining pre-electrophoresis, B: EtdBr staining post-electrophoresis, C: brightness of bands plotted from square a in panel A, D: brightness of bands plotted from sum of square a and b, E: Ensemble enthalpy (ΔH) of $GGG(T_nA_n)_{12/n}CCC \bullet GGG(T_nA_n)_{12/n}CCC$ at each temperature, calculated as $\Delta H = \Delta G - TdG/dT$ using the DINAMelt server (Rensselaer Polytechnic Institute), F: Total DNA hybridization energy calculated by DNA base pair polarity program of this study. *This figure is a representative one from the repeated experiments more than three times.

differently from the graph of FIG. 9C when the n increased to 6 and 12 (FIG. 9F).

Taken together, the facts that the $C_nA_n \bullet T_nG_n$ and $T_nA_n \bullet T_nA_n$ are more strongly intercalated by EtdBr than the $A_nC_n \bullet G_nT_n$ and $A_nT_n \bullet A_nT_n$, respectively, may support the DNA base pair polarity concept, suggesting the $C_nA_n \bullet T_nG_n$ and $T_nA_n \bullet T_nA_n$ can induce stronger electrostatic charge to form more stable DNA segments than the $A_nC_n \bullet G_nT_n$ and $A_nT_n \bullet A_nT_n$.

3.7. PCR primers composed of different DNA segments

In the PCRs using the pairs between the insert primer of $(C_nA_n)_{12/n}$ or $(T_nG_n)_{12/n}$ ($n = 1, 2, 3, 4, 6, 12$) and vector sequence primer of pBS1 or pBS2 the DNA products were plotted a double peak curve, disclosing the first peak by the insert primers of $(C_3A_3)_4$ and $(T_2G_2)_6$, a minimum by the insert primers of $(C_4A_4)_3$ and $(T_3G_3)_4$, and the sec-

Fig. 9

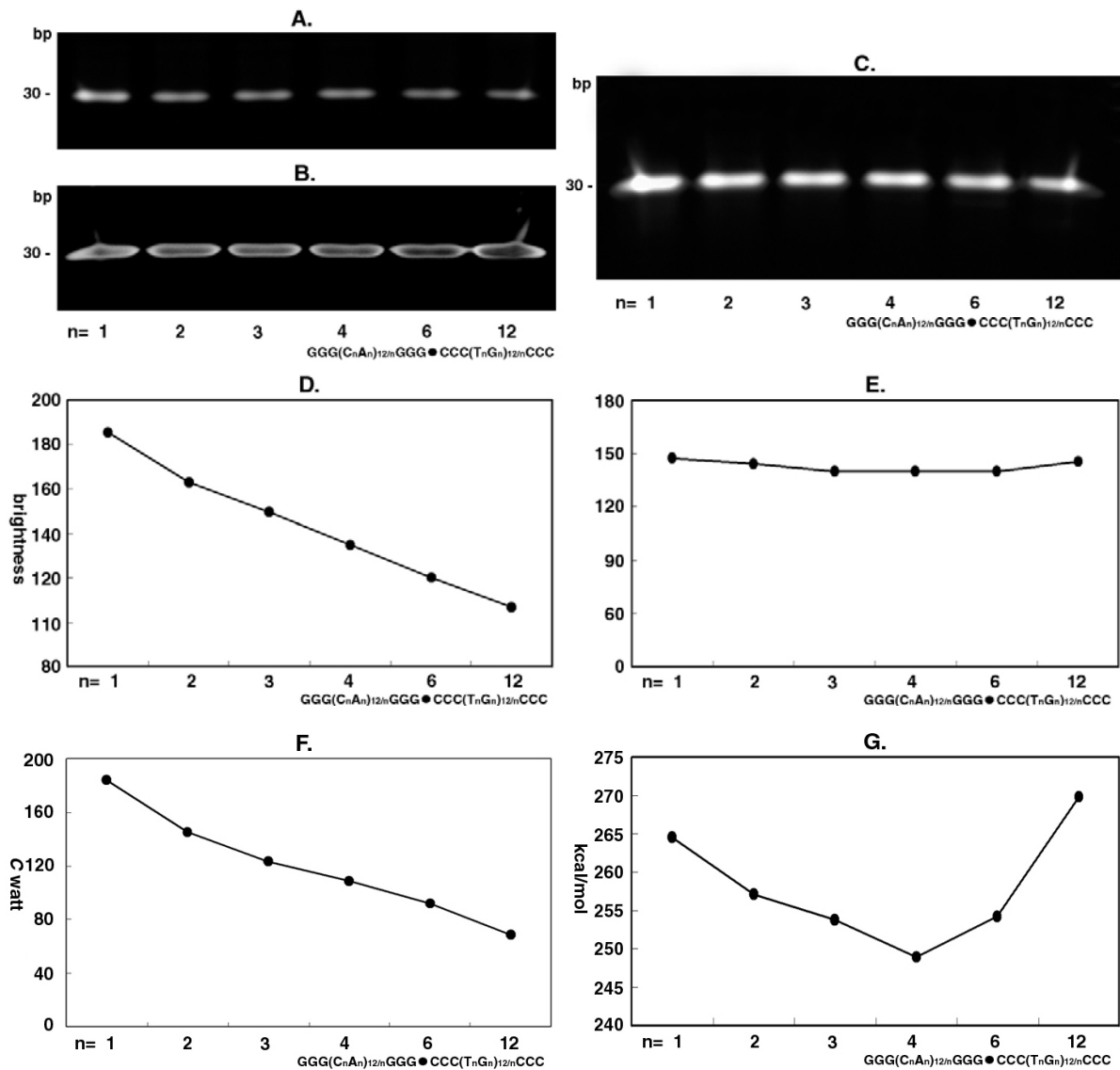


Figure 9. Hybridization of GGG(C_nA_n)_{12/n}GGG and CCC(T_nG_n)_{12/n}CCC oligonucleotides (n = 1, 2, 3, 4, 6, 12) detected by EtdBr intercalation. A: EtdBr staining pre-electrophoresis, B: EtdBr staining post-electrophoresis, C: Merging of panel A and B, D: A graph for the brightness of the bands from panel A, E: A graph for the brightness of the bands from panel B, F: total DNA hybridization energy (DHE, C watt) calculated by DNA base pair polarity program of this study, G: Ensemble enthalpy (ΔH) of GGG(C_nA_n)_{12/n}GGG•CCC(T_nG_n)_{12/n}CCC (n = 1, 2, 3, 4, 6, 12), calculated as ΔH = ΔG-TdG/dT using the DINAMelt server (Rensselaer Polytechnic Institute). *This figure is a representative one from the repeated experiments more than three times.

ond peak by the insert primers of C₁₂A₁₂ and T₁₂G₁₂ (FIG. 6). Because the insert primers of (CA)₁₂ and (TG)₁₂ usually produced non-specific extra-bands in PCR, it was thought that (CA)₁₂

and (TG)₁₂ primers were composed of too small and too many DNA segments to be annealed specifically to the target DNA. On the other hand, the C₁₂A₁₂ or T₁₂G₁₂ primer was composed of a

single DNA segment containing 24 nucleotides polarized greatly, so that a nonspecific annealing could be occurred if they were applied to the genomic DNA containing a lot of genes rather than to the single plasmid DNA used in this study. Therefore, the insert primers producing the first peak may have an implication for the strong and specific annealing upon the target DNA. However, the PCR using (C₃A₃)₄ or (T₃G₃)₄ primer produced the maximum amount of object DNA independently on the vector primers (FIG. 10). However, in the comparison between insert primers of P_y_nP_u_n and P_u_nP_y_n, the P_y_nP_u_n primers, C₁₂A₁₂ and T₁₂G₁₂, produced more amplification of object DNA than the P_u_nP_y_n primers, A₁₂C₁₂ and G₁₂T₁₂ (FIG. 10). Thus, we presume that the C₁₂A₁₂ and T₁₂G₁₂ primers (P_y_nP_u_n) are more annealed on the template DNA than the A₁₂C₁₂ and G₁₂T₁₂ primers (P_u_nP_y_n), because the P_y_nP_u_n, a Pyu DNA segment, forms stronger hybridization of the DNA duplex than the P_u_nP_y_n, a Puy DNA segment.

3.8. DNA segment conformation by reverse phase HPLC analysis

The reverse phase HPLC analysis (Agilent Chemstation, Agilent Technologies, USA) using ODS-AQ column (YMC, Waters Corp.) under 0.2 M NaCl solution, flow rate of 0.5 ml/min at 260 nm, was performed with the double strand oligodeoxynucleotides of A₁₂•T₁₂, A₆T₆•A₆T₆, and T₆A₆•T₆A₆, which contained same number of dA and dT, and with those of G₁₂•C₁₂, G₆C₆•G₆C₆, and C₆G₆•C₆G₆, which contained same number of dG and dC. The A₁₂•T₁₂, a long A-tract DNA, showed a much reduced peak compared to the A₆T₆•A₆T₆ and T₆A₆•T₆A₆, and then the T₆A₆•T₆A₆, Pyu DNA segment, showed a reduced peak than the A₆T₆•A₆T₆, Puy DNA segment, by 22.6% ($p=0.00018$). And the G₁₂•C₁₂, a long G-tract DNA forming the quadruplex, showed much reduced peak, however, the C₆G₆•C₆G₆, Pyu DNA segment, showed greater peak than the G₁₂•C₁₂, by 63.3% ($p=0.00013$), while the G₆C₆•G₆C₆, Puy DNA segment, showed less peak than the C₆G₆•C₆G₆ (FIG. 11b). This result indicated that the Pyu DNA segment can form a unique conformation, resulted in the

more stable DNA duplex by enhancing the A-tract hybridization and by preventing the quadruplex formation of G-tract DNA.

3.9. Evaluation of Pyu DNA segment

Based on the Pyu DNA segment hypothesized in this study the hybridization and PCR experiment using various oligonucleotide pairs demonstrated that the Pyu DNA segment showed stronger hybridization than Puy DNA segment, so that it can maintain its characteristic DNA structure more stably compared to the Puy DNA segment. Thereby, the entire genetic code is able to be divided into Pyu DNA segments, which are basic units of DNA double helix holding different electrostatic polarity. The average number of base pairs per a Pyu DNA segment is slightly variable depend on the regions of gene. In the human, mouse, rat, and chick genomes the average number of base pairs per a Pyu DNA segment is about 4.73 bps in the promoter region of 154,371 bps, about 4.72 bps in the intron of 46,351 bps, about 4.79 bps in the 5' untranslated region (UTR) of 23,186 bps, about 5.07 bps in the 3' UTR of 73,150 bps, and about 4.56 bps in the open reading frame (ORF) of 186,591 bps. The average number of base pairs per a Pyu DNA segment in total genome analyzed in this study was about 4.72 bps in 483,649 bps, however, the ORF was usually composed of the shortest Pyu DNA segment in the genomic regions (Supplementary Table 2).

All the hydrogen bonds of DNA base pairs produce the electrostatic charge which can be directly transformed into the hybridization energy of DNA duplex. The linear composition of DNA hybridization may display as a DNA base pair polarity in each strand of DNA duplex. Thus we think that the total DNA hybridization energy (DHE) of Pyu DNA segment in this study includes the sum of self complementary monomeric and dimeric hybridization, and the two strand dimeric hybridization, while the ensemble enthalpy of two strand DNA hybridization (Rensselaer Polytechnic Institute) may represent only the dimeric hybridization. However, although the DNA polarity has not been clearly examined due to its rapidly changing elec-

Fig. 10

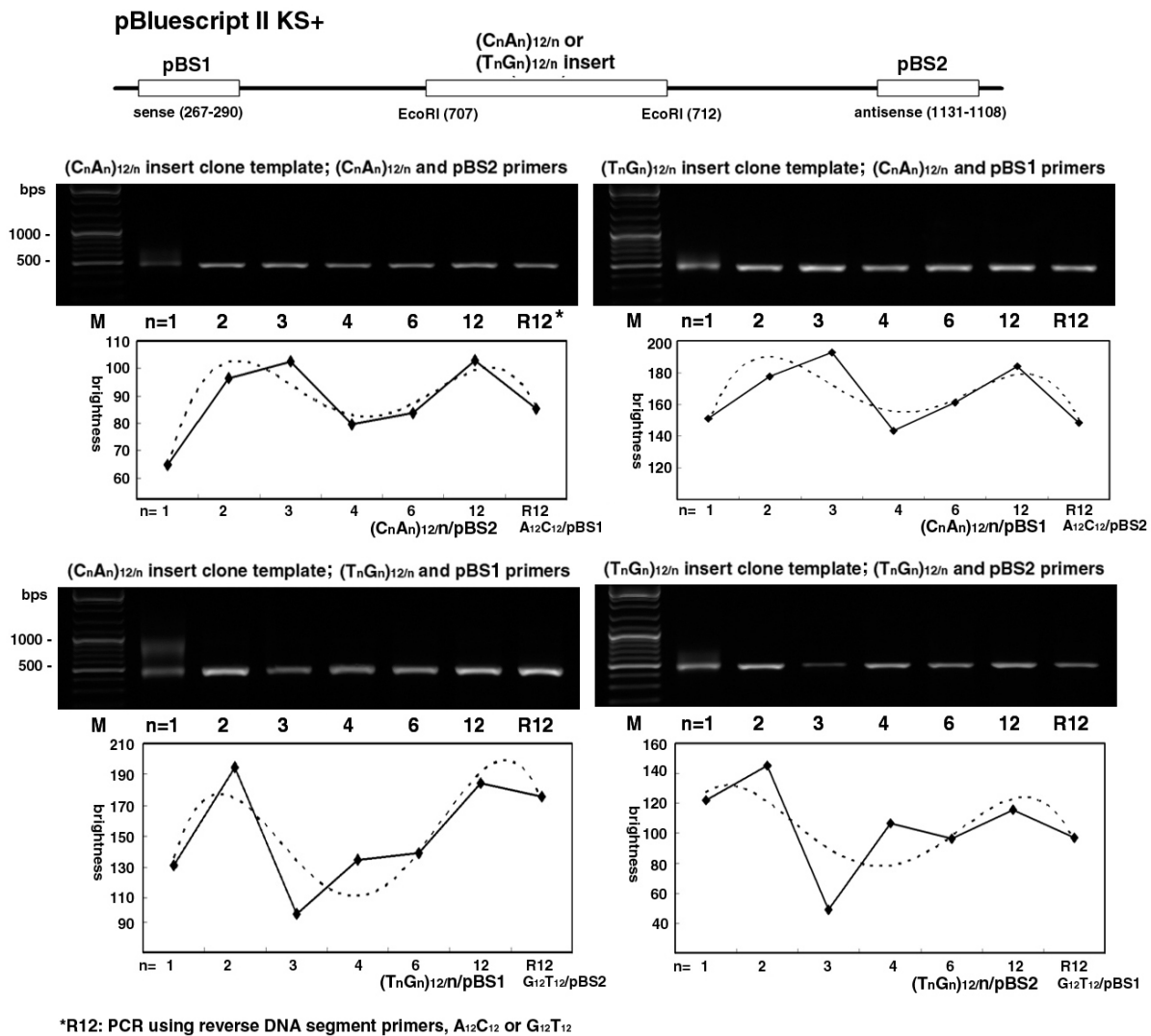


Figure 10. Location of primer sequences in the pBluescript II KS+ vector (A) and target DNA production by the optimized PCR using different pairs of primers (B and C). B: (C_nA_n)_{12/n} (n = 1, 2, 3, 4, 6, 12) primers paired by pBS1 or pBS2 primer. C: (T_nG_n)_{12/n} (n = 1, 2, 3, 4, 6, 12) primers paired by pBS1 or pBS2 primer. *This figure is a representative one from the repeated experiments more than three times.

trostatic charge of DNA hybridization, the prediction of DNA melting profile by the DINAMelt server, Rensselaer Polytechnic Institute, consistently showed the decreased $T_m(\text{Conc})$ and increased $\Delta H(\text{kcal/mol})$ in the C_nA_n•T_nG_n compared to the A_nC_n•G_nT_n (n = 3, 4, 5, 6, 7, 8, 10, 12) (Supplementary Table 3), indicating that the C_nA_n•T_nG_n, the Pyu DNA segments, showed the more unique and stable hybridization than the

A_nC_n•G_nT_n, the Puy DNA segments. The finding that the Pyu DNA segments of a DNA duplex have more stable hybridization than the Puy DNA segments can be ascribed to the continuous reciprocal induction of electrostatic charge resonance between both strands of Pyu DNA segments in DNA duplex. Therefore, the Pyu DNA segments are able to maintain its specific DNA conformation more stably than the Puy DNA

Fig. 11

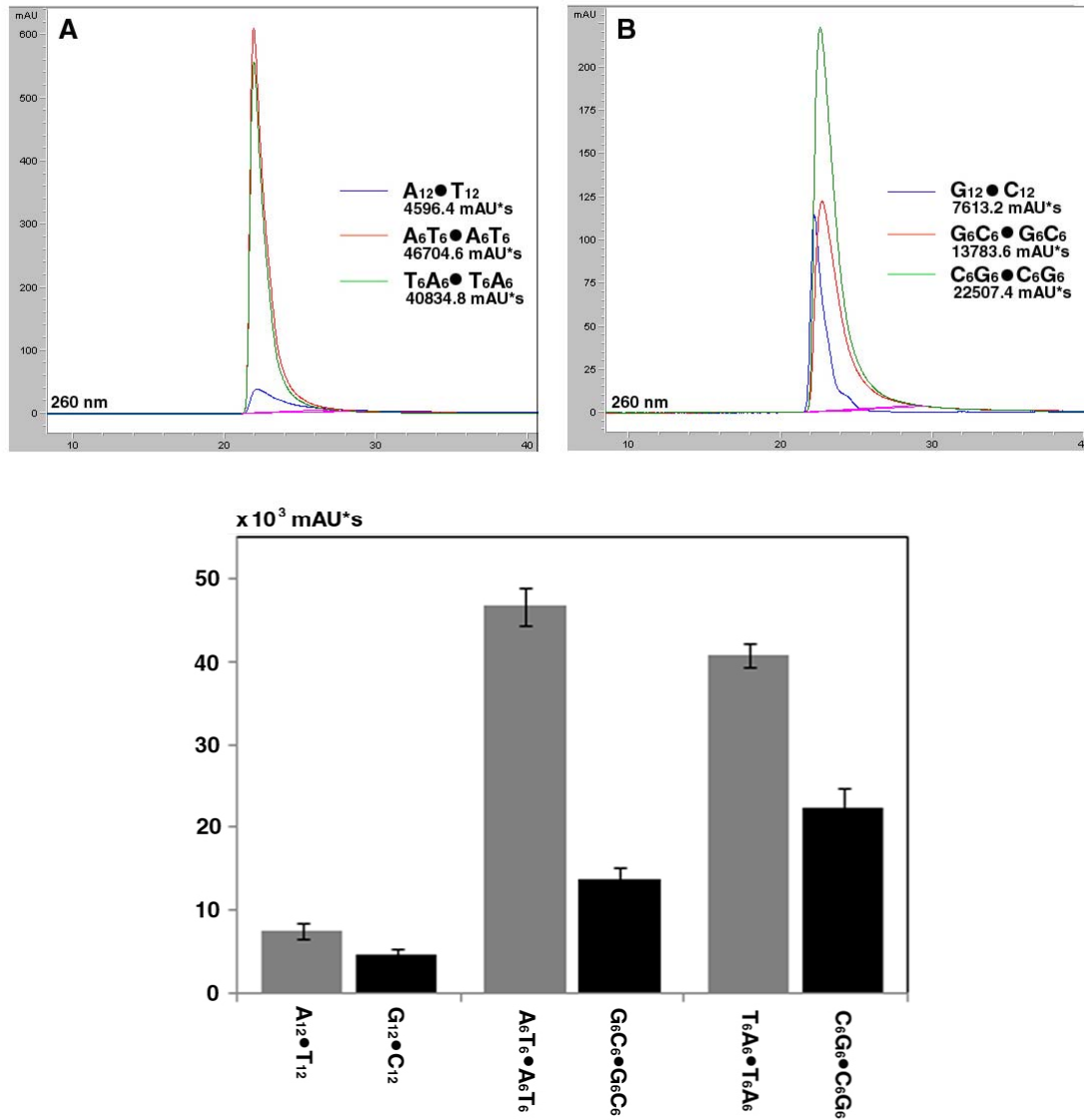


Figure 11. Reverse phase HPLC analysis of single and double strand DNAs, ODS-AQ column, running with 0.2 M NaCl solution at 260 nm. The area of the DNA peak (mAU*s) was integrated. A: A₁₂•T₁₂, a long A-tract, showed the smallest peak contrast to A₆T₆•A₆T₆ and T₆A₆•T₆A₆. The T-A step of Pyu DNA segment can enhance the bending conformation of A-tract. B: G₁₂•C₁₂, a long G-tract, showed the least peak contrast to G₆C₆•G₆C₆. And C₆G₆•C₆G₆ showed greater peak than G₆C₆•G₆C₆, preventing the quadruplex of poly dG. C: A graphic comparison of the above peaks in the same scale (mAU*s).

segments, and it is also presumed that the frequent Pyu DNA segments may tend to be more supercoiled than the Puy DNA segments. Actually in the human genome the CpG islet; (CG)_n•(CG)_n, trinucleotide repeat;

(CGG)_n•(CCG)_n [15], (CAG)_n•(CTG)_n, tetrameric repeat; (CCTG)_n•(CAGG)_n [16], pentameric repeat; (AATCT)_n•(AGATT)_n, and even dodecameric repeat; (C₄GC₄GCG)_n•(CGCG₄CG₄)_n [17] are mainly composed of one or several Pyu

DNA segments, which may produce a stable DNA structure in a limited number of the repeats in normal genetic code, and if the number of repeated Pyu DNA segment increased, the DNA polarity increased too much in a single DNA section, shifted to negative or positive polarity. This unbalanced DNA polarity may induce a conformational change of the DNA duplex which becomes irregular and supercoiled, eventually producing different genetic diseases. Also for the normal functions and conformations of genetic codes the binding sites of different transcription factors and enzymes are also composed of a basic polarity pattern of a Pyu DNA segment, *i.e.*, TATA box; (TA)₂•(TA)₂, E-box; (CA)(CG)(TG)•(CA) (CG)(TG), AP-1 binding site; (TGAG)(TCA)•(TGA)(CTCA), telomere sequences; (TTAGGG)_n•(CCCTAA)_n. Anthracyclines bind preferentially at pyrimidine-purine step, dCG in dCGATCGG which is composed of two Pyu DNA segments, dC-GA and dTC-GG. It is known that the 4,5-cationic group of the divalent anthracycline usually interacts with the G-tract major groove, and the 9-hydroxy group forms multiple hydrogen bonds with a flanking guanine [18]. However, there may exist variable binding sites composed of non-Pyu DNA segments, *i.e.*, GAA repeat [19, 20]; (GAA)_n•(TTC)_n, polyadenylate site; AATAAA•TTTATT, but these sequences may work in different ways to regulate gene transcription.

The binding of monovalent cations in the minor groove of DNA A-tracts, runs of 4-6 contiguous adenine (or thymine) residues, is a topic of continuing interest [21, 22]. The localization of cations in the A-tract minor groove would lead to asymmetric neutralization of charged phosphate residues [23, 24], giving directionality to the bending fluctuations of the helix backbone [25-28]. The molecular dynamic modeling of A-tract oligonucleotides disclosed that the A-tract appeared to act as positioning elements that make the helix phasing more precise, while Py-Pu steps in the open hinge state serve as curvature elements [22, 29]. The stable-curvature of the helix axis requires the co-operative binding of monovalent cation(s) to two or more closely spaced A-tracts [30]. However, it was known that the extended A-tract had a bifurcated hy-

drogen bonding in the A-tract major groove [31, 32]. Molecular dynamic simulation study also disclosed that A-tract was essentially straight (wedge angles of <1) and more rigid than generic B-form DNA with slight base-pair inclination, high propeller twist and a minor groove narrowing 5' to 3' [33]. We thought the curvature of A-tract DNA is caused by the increased electrostatic potential of Py-Pu step and the bifurcated hydrogen bonds between dA:dT pair. The former may indicate that the Py-Pu step is identical to the central point of Pyu DNA segment at which the total negative charge of Pys gathered and the cations are preferentially bound, and the latter may cause the conformational change of A-tract by the increase of base pair polarity in dA:dT pair.

Taken together, the Pyu DNA segments hypothesized in this study showed stronger hybridization, depending on the size and number of the DNA segment than the Puy DNA segment. The Pyu DNA segment is characterized by the DNA base pair polarity eliciting electrostatic charge from hydrogen bonds between pyrimidine and purine bases, which can be transferred to the phosphate groups of nucleotides, enable the intramolecular electrostatic interaction between both strands of a Pyu DNA segment without electron transfer. Therefore, we presume that the Pyu DNA segment producing unique hybridization stable enough to maintain its conformation can be a basic unit of genetic code. Based on the DNA base pair polarity the potential energy of DNA hybridization can be calculated and subsequently it is able to be utilized not only in the prediction of the efficiency of probe hybridization but also in the gene analysis for multiple purposes, *i.e.*, the mutation analysis and gene regulation.

Acknowledgements

We deeply thank to Dr. In Sung Choi who kindly encourage and advise us. This work was supported by research grants from Korean Science and Engineering Foundation (R11-2002-001-03003-0 and R01-2003-000-10891-0) and Korean Research Foundation (KRF-2007-013).

References

1. **Zuckerkindl E** [1976] Gene control in eukaryotes and the c-value paradox "excess" DNA as an impediment to transcription of coding sequences. *J Mol Evol* 9: 73-104.
2. **Peters H, Wilm B, Sakai N, Imai K, Maas R, Balling R** [1999] Pax1 and Pax9 synergistically regulate vertebral column development. *Development* 126: 5399-5408.
3. **Aramini JM, Germann MW** [1998] NMR studies of DNA duplexes containing alpha-anomeric nucleotides and polarity reversals. *Biochem Cell Biol* 76: 403-410.
4. **Weisz K, Leitner D, Krafft C, Welfle H** [2000] Structural heterogeneity in intramolecular DNA triple helices. *Biol Chem* 381: 275-283.
5. **van Dongen MJ, Mooren MM, Willems EF, van der Marel GA, van Boom JH, Wijmenga SS, Hilbers CW** [1997] Structural features of the DNA hairpin d(ATCCTA-GTTA-TAGGAT): formation of a G-A base pair in the loop. *Nucleic Acids Res* 25: 1537-1547.
6. **Bjornson KP, Hsieh J, Amaratunga M, Lohman TM** [1998] Kinetic mechanism for the sequential binding of two single-stranded oligodeoxynucleotides to the *Escherichia coli* Rep helicase dimer. *Biochemistry* 37: 891-899.
7. **Cox JM, Hayward MM, Sanchez JF, Gegnas LD, van der Zee S, Dennis JH, Sigler PB, Schepartz A** [1997] Bidirectional binding of the TATA box binding protein to the TATA box. *Proc Natl Acad Sci USA* 94: 13475-13480.
8. **Grove A, Galeone A, Yu E, Mayol L, Geiduschek EP** [1998] Affinity, stability and polarity of binding of the TATA binding protein governed by flexure at the TATA Box. *J Mol Biol* 282: 731-739.
9. **Perals K, Cornet F, Merlet Y, Delon I, Louarn JM** [2000] Functional polarization of the *Escherichia coli* chromosome terminus: the dif site acts in chromosome dimer resolution only when located between long stretches of opposite polarity. *Mol Microbiol* 36: 33-43.
10. **Ho PS** [1994] The non-B-DNA structure of d(CA/TG)_n does not differ from that of Z-DNA. *Proc Natl Acad Sci USA* 91: 9549-9553.
11. **Bichara M, Schumacher S, Fuchs RP** [1995] Genetic instability within monotonous runs of CpG sequences in *Escherichia coli*. *Genetics* 140: 897-907.
12. **Toyota M, Shen L, Ohe-Toyota M, Hamilton SR, Sinicrope FA, Issa JP** [2000] Aberrant methylation of the Cyclooxygenase 2 CpG island in colorectal tumors. *Cancer Res* 60: 4044-4048.
13. **Loukola A, Vilkkki S, Singh J, Launonen L V, Aaltonen A** [2000] Germline and somatic mutation analysis of MLH3 in MSI-positive colorectal cancer. *Am J Pathol* 157: 347-352.
14. **Young J, Biden KG, Simms LA, Huggard P, Karamatic R, Eyre HJ, Sutherland GR, Herath N, et al.** [2001] HPP1: A transmembrane protein-encoding gene commonly methylated in colorectal polyps and cancers. *Proc Natl Acad Sci USA* 98: 265-270.
15. **Jin P, Warren ST** [2000] Understanding the molecular basis of fragile X syndrome. *Hum Mol Genet* 9: 901-908.
16. **Liquori CL, Ricker K, Moseley ML, Jacobsen JF, Kress W, Naylor SL, Day JW, Ranum LP** [2001] Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9. *Science* 293: 864-867.
17. **Wang YH** [2007] Chromatin structure of repeating CTG/CAG and CGG/CCG sequences in human disease. *Front Biosci* 12: 4731-4741.
18. **Chaires JB, Fox KR, Herrera JE, Britt M, Waring MJ** [1987] Site and sequence specificity of the daunomycin-DNA interaction. *Biochemistry* 26: 8227-8236.
19. **Fischer KM** [1998] Expanded (CAG)_n, (CGG)_n and (GAA)_n trinucleotide repeat microsatellites, and mutant purine synthesis and pigmentation genes cause schizophrenia and autism. *Med Hypotheses* 51: 223-233.
20. **Bidichandani SI, Purandare SM, Taylor EE, Gumin G, Machkhas H, Harati Y, Gibbs RA, Ashizawa T, et al.** [1999] Somatic sequence variation at the Friedreich ataxia locus includes complete contraction of the expanded GAA triplet repeat, significant length variation in serially passaged lymphoblasts and enhanced mutagenesis in the flanking sequence. *Hum Mol Genet* 8: 2425-2436.
21. **Olson WK, Zhurkin VB** [1996] Twenty years of DNA bending. Adenine Press, Schenectady.
22. **Beveridge DL, Dixit SB, Barreiro G, Thayer KM** [2004] Molecular dynamics simulations of DNA curvature and flexibility: Helix phasing and premelting. *Biopolymers* 73: 380-403.
23. **Strauss JK, Maher LJ 3rd** [1994] DNA bending by asymmetric phosphate neutralization. *Science* 266: 1829-1834.
24. **Strauss JK, Roberts C, Nelson MG, Switzer C, Maher LJ 3rd** [1996] DNA bending by hexamethylene-tethered ammonium ions. *Proc Natl Acad Sci U S A* 93: 9515-9520.
25. **Hud NV, Sklenar V, Feigon J** [1999] Localization of ammonium ions in the minor groove of DNA duplexes in solution and the origin of DNA A-tract bending. *J Mol Biol* 286: 651-660.
26. **Range K, Mayaan E, Maher LJ 3rd, York DM** [2005] The contribution of phosphate-phosphate repulsions to the free energy of DNA bending. *Nucleic Acids Res* 33: 1257-1268.
27. **Manning GS** [2006] The contribution of transient counterion imbalances to DNA bending fluctuations. *Biophys J* 90: 3208-3215.
28. **McDonald RJ, Dragan AI, Kirk WR, Neff KL, Privalov PL, Maher LJ 3rd** [2007] DNA bending by charged peptides: electrophoretic and spectroscopic analyses. *Biochemistry* 46: 2306-2316.
29. **Olson WK, Gorin AA, Lu XJ, Hock LM, Zhurkin VB** [1998] DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. *Proc Natl Acad Sci USA* 95: 11163-11168.
30. **Lu Y, Stellwagen NC** [2007] Monovalent cation binding by curved DNA molecules containing variable numbers of A-tracts. *Biophys J* 94: 1719-1725.
31. **Nelson HC, Finch JT, Luisi BF, Klug A** [1987] The structure of an oligo(dA).oligo(dT) tract and its biological implications. *Nature* 330: 221-226.
32. **Coll M, Frederick CA, Wang AH, Rich A** [1987] A bifurcated hydrogen-bonded conformation in the d(A.T) base pairs of the DNA dodecamer

- d(CGCAAATTTGCG) and its complex with distamycin. Proc Natl Acad Sci USA 84: 8385-8389.
33. **McConnell KJ, Beveridge DL** [2001] Molecular dynamics simulations of B'-DNA: sequence effects on A-tract-induced bending and flexibility. J Mol Biol 314: 23-40.
 34. **Ma C, Sun L, Bloomfield VA** [1995] Condensation of plasmids enhanced by Z-DNA conformation of d(CG)_n inserts. Biochemistry 34: 3521-3528.
 35. **Patel DJ, Canuel LL** [1976] Ethidium bromide-(dC-dG-dC-dG)₂ complex in solution: Intercalation and sequence specificity of drug binding at the tetranucleotide duplex level. Proc Natl Acad Sci USA 73: 3343-3347.