J. Biochem. **138**, 363–373 (2005)
doi:10.1093/jb/mvi149

Single-Strand Conformation Polymorphism (SSCP) of Oligodeoxyribonucleotides: An Insight into Solution Structural Dynamics of DNAs Provided by Gel Electrophoresis and Molecular Dynamics Simulations

Manish Biyani[1,2] and Koichi Nishigaki[1,2,*]

[1]Department of Functional Materials Science, Saitama University, 255 Shimo-Okubo, Sakura-ku, Saitama-shi, Saitama 338-8570 and [2]Rational Evolutionary Design of Advanced Biomolecules, Saitama Small Enterprise Promotion Corporation, SKIP City, 3-12-18 Kamiaoki, Kawaguchi, Saitama 333-0844

Received April 6, 2005; accepted July 3, 2005

Studies on the solution structure dynamics of RNA/DNA are becoming crucially important. The phenomena of SSCP (single-strand conformation polymorphism), small RNA dynamics in a cell, and others can be related to the conformational changes of single-stranded (ss) RNAs/DNAs in solution. However, little is known about those dynamics. Only the intra-structural transition of ssDNAs in solution has been reported based on Watson-Crick (W-C) base-pairing. Here, we found a general feature of the SSCP phenomenon by studying the simpler molecules of ss-oligodeoxyribonucleotides. A single base substitution or a positional exchange of nucleotide in a highly homologous series of ss-dodecanucleotides led to a change in the mobility-in-gel. This was unexpected, since most of these nucleotides [such as d(A_{11}G) or d(A_{11}C)] have no possibility of forming W-C base-pairing. MD (molecular dynamics) experiments revealed differences in shape and size between the dynamic structures of these molecules which could affect their mobility-in-gel. In addition, a high correlation was observed between the electrophoretic mobility and the size-related parameters such as end-to-end distance obtained from MD simulations. Because the simulation was considerably shorter (nanosecond) than the experimental time-scale (second), the result must be considered conservatively; but it is nevertheless encouraging for utilizing MD simulation for structural analysis of oligonucleotides.

Key words: gel electrophoresis, molecular dynamics, oligonucleotides, single-strand conformation polymorphism, solution structure dynamics.

A single nucleotide difference in single-stranded DNAs (ssDNAs) can affect the ssDNA folding and, hence, can be detected as a difference in electrophoretic mobility performed under appropriate conditions (*1*, *2*). This phenomenon is termed single-strand conformation polymorphism (SSCP) (*2*). It has been extensively used in thousands of works to analyze point mutation and SNP (single nucleotide polymorphism). However, studies on this phenomenon are remarkably few (*3*–*6*), because conventional technologies such as X-ray diffraction and NMR are currently useless in studying the dynamically unstable structures involved. In contrast, solution structures of double-stranded DNA (dsDNA) have been studied intensively, leading to the development of the reptation model (*7*), the static model (*8*) and others (*9*, *10*). These approaches that deal with dsDNAs do not need to consider sequence effects, because the structure composed of Watson-Crick (W-C) base pairing is entirely stable. This fact enables us to describe the behavior of dsDNA in terms of a small number of parameters such as the size and the persistent length. Dealing with ssDNA, however, is not so simple. The relevant facts elucidated in this field are as follows. First, the unstable structure of a four-base paired double-stranded sequence (5′GGCC/3′CCGG), the canonical structure for restriction enzyme *Hae*III recognition, was formed for brief periods and was thus cleaved in solution structure dynamics as clarified by the investigation using ssDNA (*11*). Second, single-strand conformation dynamics for ssDNAs of more than 100 nucleotides (nts) were reported, which showed that these relatively unstable structures were responsible for the difference in the SSCP phenomenon (*3*). Third, sequence-specific mobilities of oligodeoxyribonucleotides in gel were observed for quite different sequences of oligodeoxyribonucleotides (*4*), indicating the conformational effect on mobility.

Here, we report that even highly homologous oligonucleotides such as d(TTTTTTATTTTT) and d(TTTTTTTT-TTTA) exhibit the SSCP phenomenon, *i.e.*, sequence-dependent differences in electrophoretic mobility under nondenaturing conditions. To explain the principle nature of this phenomenon, we systematically studied the effects of a single base substitution in a series of highly homologous dodecanucleotides by gel electrophoresis and molecular dynamics (MD) simulation and examined the

*To whom correspondence should be addressed at: Department of Functional Materials Science, Saitama University, 255 Shimo-Okubo, Sakura-ku, Saitama-shi, Saitama 338-8570. Tel/Fax: +81-48-858-3533, E-mail: koichi@fms.saitama-u.ac.jp

relative differences in the structural dynamics among them. Discrete changes in electrophoretic mobility of dodecanucleotides were observed and interpreted in terms of the divergent dynamic structures in solution, as governed by intra-molecular interactions between different parts of a single-strand, which were further confirmed by direct structural comparison of statistical averages obtained from nanosecond MD structures of the oligonucleotides studied. Furthermore, we made an unexpected finding that nanosecond MD may semi-quantitatively predict the mobility-in-gel of an oligonucleotide. Based on these observations, the SSCP phenomenon could be more generally interpreted, applying the results for very simple oligonucleotides to much larger ssDNAs. The knowledge and approach established here with regard to solution structure dynamics of ssDNA appear to be beneficial for not only technological purposes of separating DNAs but also understanding various phenomena related to RNAs, such as systematic translational regulations performed in a cell by small non-coding microRNAs (*12*), which are believed to form not only a particular structure but also various secondary structures in solution.

MATERIALS AND METHODS

Oligodeoxyribonucleotides—All synthetic oligodeoxyribonucleotides used here (listed in Table 1) were commercially obtained from ESPEC Oligo Service Corp. (Ibaraki, Japan) in an HPLC-purified form. The concentrations were determined by measuring absorption at 260 nm and using specific extinction coefficients for A, T, G and C. Prior to electrophoresis, all were dissolved in nuclease-free water to concentrations ranging from 1 to 10 μM. Oligodeoxyribonucleotides labeled at their 5′-positions with Cy3 were also used for confirmation experiments.

Gel Electrophoresis—Slab gel electrophoresis was carried out using different concentrations of polyacrylamide gel (5% to 30% at a constant ratio of acrylamide:bis = 19:1) in 90 mM TBE (Tris-borate-EDTA) buffer with or without

Table 1. **Properties of G-, C-, A-, and T-series single-stranded dodecanucleotides obtained by experiment or simulation.**

S. No	Mnemonic	Dodecamers seq. (5′ to 3′)	MW[a]	μ[b] (10^{-4} cm^2 V^{-1} s^{-1})	$d_{e\text{-}e}$[c] (Å)	r_{ces}[d] (Å)	τ_{rces}[f] (ns)	*RMSD*[e] (Å)	γ_{RMSD}[g] (ns)
1	G_{12}	GGGGGGGGGGGG	1.141	0.459 ± 0.006	18.26	14.34	0.5	8.96	0.2
2	G_6CG_5	GGGGGGCGGGGG	1.129	0.480 ± 0.002	32.28	18.89	1.4	7.36	0.6
3	$G_{11}C$	GGGGGGGGGGGC	1.129	0.477 ± 0.004	32.02	17.24	1.9	6.44	0.5
4	G_6AG_5	GGGGGGAGGGGG	1.136	0.464 ± 0.004	19.56	16.04	0.6	9.56	0.2
5	$G_{11}A$	GGGGGGGGGGGA	1.136	0.466 ± 0.006	26.49	16.38	0.7	6.82	0.2
6	G_6TG_5	GGGGGGTGGGGG	1.134	0.468 ± 0.002	26.55	16.45	0.9	7.70	0.3
7	$G_{11}T$	GGGGGGGGGGGT	1.134	0.459 ± 0.005	21.88	15.97	1.7	7.83	0.5
8	C_6GC_5	CCCCCCGCCCCC	1.012	0.487 ± 0.002	13.71	14.49	0.3	10.49	0.1
9	$C_{11}G$	CCCCCCCCCCCG	1.012	0.479 ± 0.001	10.85	13.15	0.7	9.44	0.2
10	C_{12}	CCCCCCCCCCCC	1.000	0.492 ± 0.003	21.96	17.01	2.6	8.69	1.1
11	C_6AC_5	CCCCCCACCCCC	1.007	0.492 ± 0.003	18.73	13.83	0.4	9.59	0.1
12	$C_{11}A$	CCCCCCCCCCCA	1.007	0.488 ± 0.001	15.46	13.40	1.0	10.79	0.3
13	C_6TC_5	CCCCCCTCCCCC	1.004	0.490 ± 0.001	16.76	14.62	1.3	8.94	0.4
14	$C_{11}T$	CCCCCCCCCCCT	1.004	0.486 ± 0.002	13.27	13.90	0.5	9.23	0.1
15	A_6GA_5	AAAAAAGAAAAA	1.089	0.497 ± 0.001	15.44	13.20	0.3	9.46	0.1
16	$A_{11}G$	AAAAAAAAAAAG	1.089	0.492 ± 0.002	11.50	12.73	0.1	10.01	0.0
17	A_6CA_5	AAAAAACAAAAA	1.078	0.503 ± 0.000	21.94	15.79	2.1	8.34	0.8
18	$A_{11}C$	AAAAAAAAAAAC	1.078	0.505 ± 0.003	24.05	15.99	1.3	7.79	0.4
19	A_{12}	AAAAAAAAAAAA	1.085	0.504 ± 0.003	31.81	19.09	2.9	7.16	1.2
20	A_6TA_5	AAAAAATAAAAA	1.082	0.504 ± 0.001	34.26	19.47	4.1	5.98	5.3
21	$A_{11}T$	AAAAAAAAAAAT	1.082	0.501 ± 0.003	19.48	15.90	0.8	8.31	0.3
22	T_6GT_5	TTTTTTGTTTTT	1.060	0.425 ± 0.001	21.86	15.83	1.0	8.71	0.3
23	$T_{11}G$	TTTTTTTTTTTG	1.060	0.419 ± 0.002	18.63	14.84	0.1	9.31	0.0
24	T_6CT_5	TTTTTTCTTTTT	1.048	0.434 ± 0.001	33.29	19.06	2.3	6.45	1.0
25	$T_{11}C$	TTTTTTTTTTTC	1.048	0.436 ± 0.003	33.85	20.40	1.6	7.17	1.8
26	T_6AT_5	TTTTTTATTTTT	1.055	0.440 ± 0.000	16.21	14.87	1.3	7.83	0.4
27	$T_{11}A$	TTTTTTTTTTTA	1.055	0.433 ± 0.003	19.55	15.64	0.6	8.36	0.1
28	T_{12}	TTTTTTTTTTTT	1.053	0.430 ± 0.003	27.64	17.62	1.8	7.46	0.9
29	$T_3GT_4GT_3$	TTTGTTTTGTTT	1.067	0.423 ± 0.002	13.71	13.20	0.3	8.74	0.1
30	$T_3CT_4GT_3$	TTTCTTTTGTTT	1.056	0.434 ± 0.001	16.99	14.29	1.7	7.77	0.5
31	$T_3GT_4CT_3$	TTTGTTTTCTTT	1.056	0.435 ± 0.000	17.46	15.50	1.2	9.71	0.3
32	$T_3CT_4CT_3$	TTTCTTTTCTTT	1.044	0.442 ± 0.001	26.35	17.95	3.0	7.53	1.5
33	$T_3AT_4AT_3$	TTTATTTTATTT	1.058	0.461 ± 0.003	23.57	15.02	2.9	7.65	1.4

[a]Relative molecular weight to that of C_{12} (3,408 Dalton). [b]Absolute electrophoretic mobility with standard deviation values through a 20%T native gel, calculated from more than 3 independent experimental trials. [c,d,e]Average simulation values of end-to-end distance, radius of completely enclosing sphere, and root mean square distance, respectively obtained by MD calculation starting wih B-form structure. [f]Half transition times observed in the trajectories of r_{ces} (see text). [g]Time ratio for extended *versus* folded conformations observed during their 5 ns trajectories of *RMSD*.

8 M urea. This buffer concentration was adopted since relative mobilities of dodecanucleotides are basically not affected by TBE buffer concentrations in the range of 45–450 mM (*13*). The temperature was controlled at 20°C for non-denaturing gel or at 60°C for denaturing gel using a circulating water bath. All gels were freshly cast prior to the experiments. The amounts of polymerization initiators, TEMED and ammonium persulfate, were adjusted for each gel concentration to ensure gel polymerization within 30 min. Gels were subjected to a prerun until the electric current stabilized. After electrophoresis, oligodeoxyribonucleotides were visualized by silver staining as described elsewhere.

In cases of multiple bands in G-series, which were observed to be dependent on the concentration of migrating sample (Biyani *et al.*, GENE, in press), assigning of the monomer band was performed by changing the concentration of monomers and confirmed by analyzing the Cy3-labelled oligodeoxyribonucleotides. All gel images were captured by scanning the gels using a fluoroimager (Quantity One; Bio-Rad, USA), then processed by image-analysis software, which enabled accurate measurement of differences in relative mobility. The electrophoretic mobility (μ) was calculated as $\mu = d/Et$ (in $cm^2/V{\cdot}s$), where d is the distance of migration (cm), E is the electric field strength (V/cm), and t is the duration of electrophoresis (s).

Molecular Dynamics Simulations—All molecular dynamics (MD) simulations were performed using the *sander* module of the AMBER 7 molecular simulation program package (*14*, *15*). AMBER 94 force field (*16*) was adopted in all calculations with the GB/SA continuum solvation model (*17*), and a Debye-Huckel approximation was used to account for salt effects (incorporated into version 7 of the AMBER package). All the parameters for GB/SA calculations were taken from those established by Tsui and Case (*18*), while the salt concentration was set at 0.2 M. Direct comparisons had already shown (*18–21*) a reasonable accord between the GB (Generalized Born) solvation model and the explicit solvent model, which was confirmed in this study by showing the qualitative similarities in structural transitions calculated using the explicit solvent and GB/SA models for few oligodeoxyribonucleotides. All calculations were performed on ONYX-3400 supercomputers (SGI) at the Information Processing Center of Saitama University.

Prior to MD simulation, two independent setups were prepared for each single-stranded DNA. In one case, a starting geometry was generated using the AMBER7/*nucgen* module in a helical conformation. For this, the coordinates of the atoms for starting structures of single-stranded dodecanucleotide DNA (*e.g.*, see inset in Fig. 3) were built up from the corresponding duplex geometries simply by first constructing canonical right-handed A- or B-forms of standard nucleic acids duplexes and then deleting the atoms of the complementary strand to the desired single-stranded oligonucleotides. Thus the stability of the helix in initial starting structures for a single-stranded DNA in a A- or B-form is governed by the stacking interactions between the adjacent bases. The coordinates of the missing hydrogen atoms were added to the *nucgen* structures using the AMBER7/tleap editing program. In the other case, the initial structure used was an unstacked random-coiled (nonhelical) conformation obtained by a high temperature (900 K) dynamics simulation beginning with a single-stranded helical structure and then selected after 100 ps. Prior to MD simulation under physiological conditions, the starting conformation was subjected to two rounds of 500 steps of steepest descent energy minimization in order to remove any possible local atomic clashes. The first round was performed to minimize the potential by displacing the position of hydrogen atoms, keeping all heavy atoms in each residue constrained to their original positions ($ntr = 1$) using a harmonic potential with a force constant of $k = 5$ kcal/mol/Å^2. The second round was carried out for the whole structure without any harmonic restraints. The atomic systems were then subjected to a 10 ps equilibration with a continuously-reducing energy constraint (2 to 0.1 kcal/mol/Å^2). SHAKE (*22*) was applied to all bonds involving hydrogen atoms. The translational and rotational motion was removed at every nscm = 1,000 steps. During equilibration the temperature of the system was gradually raised from 0 to 293 K. The final structure, attained after the equilibration process, was used as the starting structure for MD at 293 K. Each system was then simulated for 5 ns and some were tested up to 50 ns.

For evaluating nonbonded interactions and to reduce the time spent on computing long-range interactions, the cut-off distance with *nrespa* variables (which allows evaluation of slowly-varying terms involving derivatives with respect to the effective radii and pair interactions whose distances are greater than the cutoff) were used. In AMBER, the cutoff parameter for GB simulations controls not only the truncation of Coulombic and Lennard-Jones parameters but also the exclusion of distant pairs of atoms in computing the effective Born radii. Inclusion of *nrespa* variables allows the reciprocal part of the Ewald procedure to be carried out while evaluating short-range interactions. The use of a 10-Å cutoff with $nrespa = 2$ and $dt = 0.002$ ps is known to be effective for GB simulations. In recent studies, use of the GB model with cutoff has been shown to yield results closely mimicking those from a corresponding GB simulation with no cutoff (*19*, *23*). This applicability was further examined in advance by carrying out GB simulations of several oligodeoxyribonucleotides (G_{12}, G_6AG_5, G_6TG_5, G_6CG_5 for G-series and C_6GC_5, C_6AC_5, C_6TC_5, C_{12} for C-series) with (10 Å) or without (300 Å) cutoff values. The use of a cutoff worked well for GB simulations and reduced the computing time spent on long-range interactions by a factor of more than four. The correlation coefficient between with and without cutoffs in terms of half-ransition time for each series was 0.97.

The average and instantaneous temperatures and various energies were monitored over the course of the simulation and atomic coordinates were saved every picosecond. All other simulation properties were derived from analyses of these snapshots. The RMSD value for each point on a trajectory was calculated against the initial starting structure using the AMBER7/*Carnal* module. Snapshots from a trajectory were illustrated with a molecular visualization program, such as *VMD* (*24*) or *RasMol* (*25*), for visual inspection.

Measures Used for Quantitating Structural Dynamics—To characterize the structural dynamics of oligodeoxyribonucleotides, here we introduced and employed three

measurements: the radius of completely enclosing sphere (r_{ces}), end-to-end distance ($d_{e\text{-}e}$) and half-transition time ($\tau_{1/2}$). The overall shape of a single-stranded oligodeoxyribonucleotide was determined by the size of r_{ces}, which represents the radius of a sphere that encloses 95% of the atoms with its center located at the center of mass; and by $d_{e\text{-}e}$, which is the distance between the backbone-end atoms (non-hydrogen) of a single-stranded oligodeoxyribonucleotide. The half-transition time, $\tau_{1/2}$, is the midpoint of the collapse starting from an extended state to a folded one. The calculation of r_{ces} and $d_{e\text{-}e}$ was performed using in-house programs, while the others were calculated with AMBER7/*Carnal* module.

RESULTS

Distinctive Mobility-in-Gel of Dodecanucleotides Differing by One Base—Homododecanucleotides with a single base substitution were serially studied. Base substitution in this series resulted in different mobility during non-denaturing gel electrophoresis (Fig. 1). Remarkably, oligodeoxyribonucleotides having the same nucleotide composition but with different positional arrangements can be separated, as shown by the pairs, $G_6TG_5/G_{11}T$ $C_6GC_5/C_{11}G$, $C_6AC_5/C_{11}A$, $T_6GT_5/T_{11}G$, and $T_6AT_5/T_{11}A$. These results were surprising, and even though we had observed the SSCP phenomenon previously [in which mainly transient Watson-Crick interactions are working and thus point-substituted ssDNAs can have different mobilities due to differences in the overall dynamic structures (*4*)], it seemed unlikely that these small molecules would form significant structures. This observation suggested that, although the premise of the previous interpretation is probably correct, a greater contribution of non-W-C interactions should be considered. These non-W-C interactions must include not only well-established interactions (for example, Hoogsteen-type interactions), but also interactions yet not recognized (hydrogen bonding and base stacking). This point is reinforced by the current findings.

Figure 1 also shows that: (i) a single base replacement causes, more or less, a mobility shift during gel electrophoresis, as in X_6YX_5 and $X_{11}Y$, where X (= G, C, A or T) is the homododecanucleotide type and Y is the substituted nucleotide type other than X; (ii) double-base substitutions resulted in a greater mobility shift (Fig. 1e; *e.g.*, $T_{12} < T_6CT_5 < T_3CT_4CT_3$); (iii) a single base replacement with G in homododecanucleotides [$d(C)_{12}$, $d(A)_{12}$, and $d(T)_{12}$] caused a decrease in mobility while a single base replacement with C in homododecanucleotides [$d(G)_{12}$, $d(A)_{12}$, and $d(T)_{12}$] caused an increase in mobility. These interesting observations are novel, as none of the previous reports discussed such short and simple oligodeoxyribonucleotides (*1–4*, *26*). The closest notion previously established is that of sequence-dependent mobility profiles, which states that each oligodeoxyribonucleotide behaves specifically with regard to mobility when subjected to a temperature gradient due to dynamic W-C interactions (*11*).

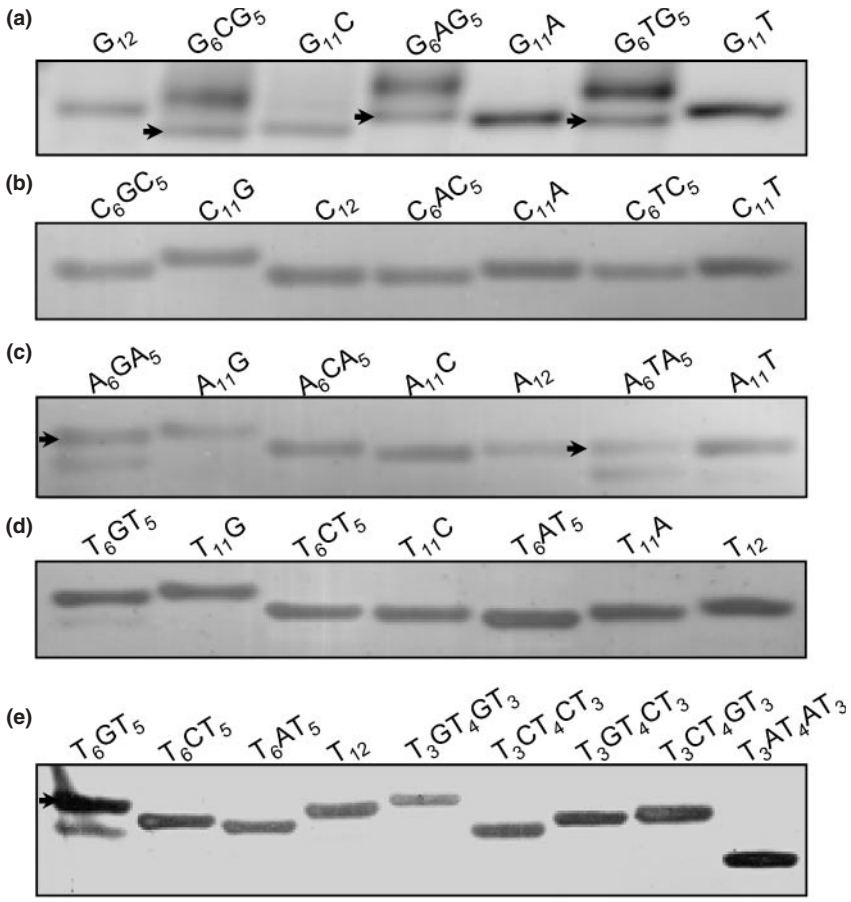


Fig. 1. **SSCP of simple single-stranded oligonucleotides.** Homo-dodecanucleotides possessing a single base substitution at the center, near the center or at the 3′-terminal were subjected to electrophoresis under non-denaturing conditions on 20% neutral polyacrylamide gel and visualized by silver staining (Ref. *45*). Separation of G series (a), C series (b), A series (c), T series (d) and double base substitution T series (e) are shown. In case of the G series with G_6XG_5 type dodecanucleotides, several intense bands (representing multi-stranded species) appeared above the monomer band, while with $G_{11}X$ type dodecanucleotides, ladder pattern bands (representing aggregated lower mobility structures (Ref. *46*), not shown here) appeared (where X denotes non-guanine base). In case of multiple bands, an arrow shows the monomeric form of the parent sequence band assigned by another iterative experiment. Reproducibility of relative migration order was tested by more than 3 independent experiments. Double base substitution T series were performed to further confirm the effect of specific-base substitution type.

Remarkably, Bowling *et al.* reported a phenomenon in which the difference in the 3′-terminal dideoxynucleotides (ddN) of ssDNA resulted in different mobilities during gel sequencing (*27*), and is thus consistent with our results that a 3′-terminal C allows the ssDNA to migrate faster, while a 3′-terminal T causes slower migration (in addition, in our experiment, a 3′-terminal G was shown to decrease mobility substantially).

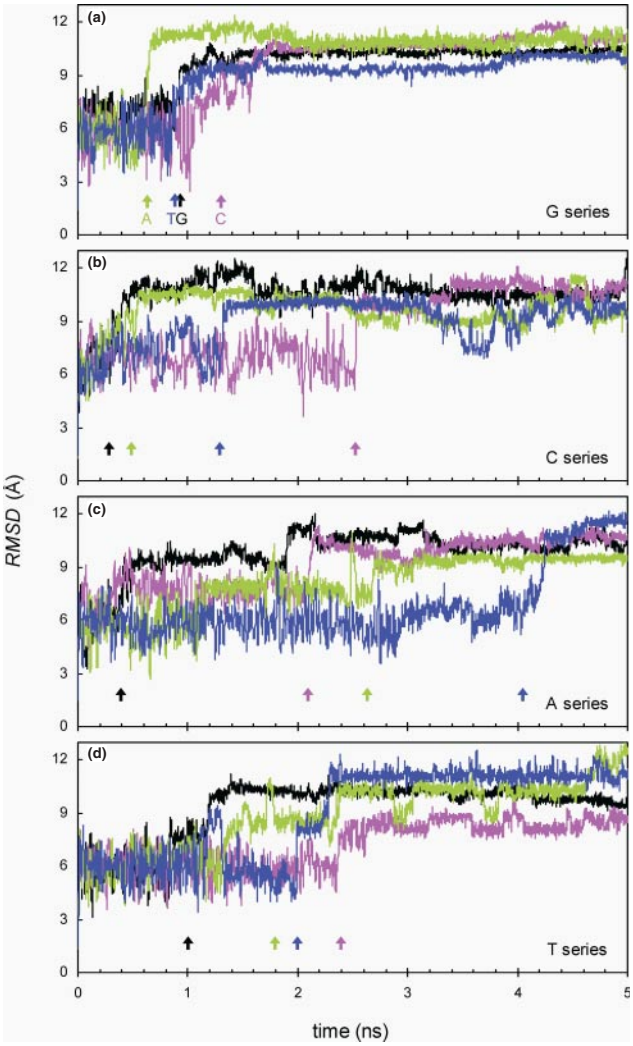
These results clearly indicate that a conformational factor is responsible for this phenomenon. In addition, this conformation must be less stable, because there are no known stable structures that would allow separation of T_6GT_5 and $T_{11}G$ (both dodecanucleotides composed of 11Ts and one G). Therefore, for such simple molecules, the only rational interpretation is the assumption that there are intra-molecular 'conformational dynamics' which lead to effectively different (time-averaged) molecular shapes, since it is evident that mainly intra-molecular interactions without W-C base pairings are working in these homo-dodecanucleotide series, as there are 4, 3, 3, and 2 distinct possibilities for forming at least two hydrogen bonds between two homo base pairs of G, A, T, and C, respectively (*28*). This intra-molecule hypothesis is supported by the observation here (unpublished data) that the denaturant negated this phenomenon by reducing the relative frequency of intramolecular interactions (*29*). If this is the case, MD simulation may be able to demonstrate this phenomenon, although seconds-to-minutes simulations are not presently available. Recently, the modeling calculation on single-stranded DNA has shown very encouraging results to understand to what extent the DNA sequence, strand length, backbone modification and ion-strand interaction influences its structural and dynamics conformation (*30*, *31*). We therefore performed realistic calculations for nanosecond MD, expecting to find clues to interpret the experimental results.

Simulation Measurements for Structural Dynamics—All dodecanucleotides (33 species) used in electrophoresis were then analyzed by MD calculation (see Table 1). The simulation results were analyzed in terms of root-mean-squared deviation (RMSD), end-to-end distance ($d_{e\text{-}e}$), radius of completely enclosing sphere (r_{ces}), and snapshots, as described in "MATERIALS AND METHODS."

Root-Mean-Square Deviation—The root-mean-squared deviation (RMSD), which describes the "distance" between two conformations, may able to indicate the flexibility and range of conformational space during dynamics simulations. We have compared the overall structural features by looking at the time evolution of RMSD values with respect to each starting B-form of all the studied dodecanucleotides [Fig. 2, only the X_6YX_5 series is shown, where X (=G, C, A or T) is the homododecanucleotide type and Y is the substituted nucleotide type]. Initial fluctuations are large in magnitude, associated with short or long-term drift, which later vanishes in all cases, and then remain constant for the remaining duration of the simulation (this continued up to 50 ns for nucleotides tested). So, in general, each dodecanucleotide underwent passage from a helically stacked starting conformation to a folded conformation through unstacking and/or restacking processes and tended to shift to conformations of finite RMSD value within period of 4 nanoseconds of simulation. Interestingly, this parameter (rate of transition between starting B-form to folded conformation) is observed to be dependent on the specific type of base substitution, *e.g.*, a single base substitution with C at the center of the all four homonucleotides results in more fluctuation (both band-width and transition delay time) than one substitution with G, illustrating the concept of "single base-induced difference in conformation dynamics." If these observations have any relationship with macroscopic phenomena, a quantitative study of the base substitution effect is desired.

Essentially the same results were obtained for different initial starting structures (*i.e.*, B-form, A-form and stretched conformation) (as partly shown for G-series in Supplementary Figure S1 (see: http://gp.fms.saitama-u.ac.jp/jb-suppli.html/), *i.e.*, the order of half-transition times ($G_{12} < G_6AG_5 < G_6TG_5 < G_6CG_5$) remains the same for all three cases). This means that the trajectory of conformation occurring during simulation is most likely to be induced by a specific single-base substitution and independent of the initial starting structure. This also suggests that there is a local stability minimum (trough) in the conformation space, which all these simple dodecanucleotides are able to attain in nanosecond trials (5 ns) with this potential. Although it is unknown whether this is a global minimum or a case of entrapment in a local minimum, it appears to be sufficiently stable on a nanosecond time scale.

Recognizing Intramolecular Interactions—Direct visualization of snapshot images of simulated oligodeoxyribonucleotides after their collapse into globular structures was used to analyze all possible intramolecular hydrogen bonding (Hoogsteen type) and base-stacking interactions. A relaxed definition of the hydrogen bond is used in the present study. The definition is that the maximum distance allowed between the atoms bearing the donor and acceptor atoms is 2.5 Å. The averaged snapshot at the later stages of the nanosecond trajectory revealed that both ends of the dG_{12} strand were sufficiently close to form a semi-stable hairpin-like folded structure (with a type of base-stacking and possible intramolecular hydrogen bonding) (Fig. 3; time-averaged images for G-series are shown). Visual inspection confirmed that the hairpin-like structure dominates. The appearance of this structure depended on the type and position of the single base substitution. Substitution with a C at the center (*i.e.*, G_6CG_5) or at the 3′-terminal (*i.e.*, $G_{11}C$) of dG_{12} resulted in a less self-folded and more extended conformation due to the lower degree of stacking of this conformation, whereas A- or T-substitution at the center or at the 3′-terminal of dG_{12} enhanced the tendency for self-folding (Fig. 3). If the stretched conformer migrates faster than the folded one (Supposition A), then (although this observation is in the order of only nanoseconds) the above observations are consistent with the second-to-minute behavior of gel experimental results (Table 1; *e.g.*, serial # 1–7, relative differences between electrophoretic mobility and simulated values for almost all sequences in G-series and Ref. *27*). This consistency appears to be coincidental, because there is no theoretical basis for accordance between nanosecond structures and second-to-minute phenomena. In other words, we first need to know the entire landscape of conformational space of oligodeoxyribonucleotides, which can be used to determine whether a particular stable structure is in the global minimum. Nevertheless, we can conservatively



state that this brief, nanosecond simulation may have attained the global minimum for the oligodeoxyribonucleotides used (Supposition B) because they have such small and simple structures. Therefore, the current observations regarding nanosecond structures of oligodeoxyribonucleotides may be applicable to second-to-minute behavior, which is further supported by the following observations.

Overall Size of Oligodeoxyribonucleotides—The same series was further analyzed in terms of r_{ces} (radius of 95% enclosing sphere) (for definition, see "MATERIALS AND METHODS" section) for those oligodeoxyribonucleotides which appeared in Fig. 1. Due to the molecular simplicity of the studied oligodeoxyribonucleotides, which had almost identical molecular weights, we opted to use the r_{ces} term for analysis instead of the R_g (radius of gyration) term. In almost all cases, we observed a single sharp transition from a swollen (B-form) to a compact (folded) structure in terms of r_{ces} within 5 ns in the simulation (see Supplementary Figure S1: http://gp.fms.saitama-u.ac.jp/jb-suppli.html). The majority finished the main transition before the 4-ns step, and then remained constant with small ripples (tested up to 50 ns for a few sequences). A general feature of the homo-dodecanucleotide series generated by a single-base substitution near the center of sequence was observed, *i.e.*, the order of half-transition times from a swollen (initial) to a compact structure. For the homo-G, homo-T and homo-C series it is $X_6GX_5 < X_6AX_5 < X_6TX_5 < X_6CX_5$,

Fig. 2. **Time-evolution of the root-mean-square deviation (RMSD) of single-stranded homo-oligodeoxyribonucleotides**. The four series of homo-dodecanucleotides having a single base substitution near the center of the sequence [$d(G_6YG_5)$, $d(C_6YC_5)$, $d(A_6YA_5)$, and $d(T_6YT_5)$ (top to bottom) where Y = G, C, A or T] are depicted. The type of substitution is shown by color [G (black), C (pink), A (green), and T (blue)]. The arrows indicate the half-transition time. Plots were made for 1 ps running average.

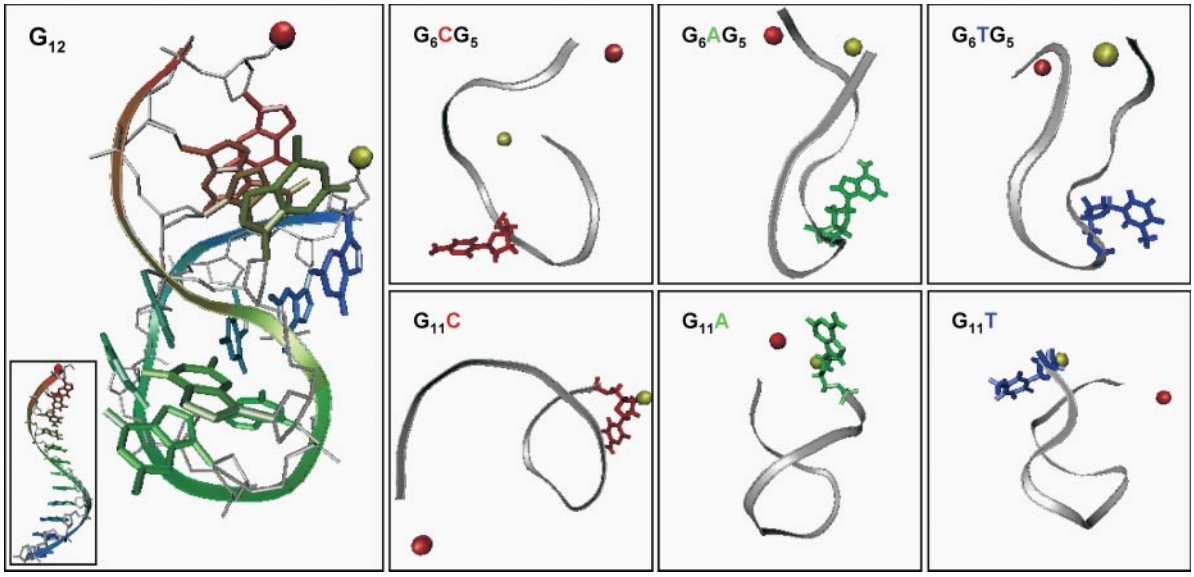


Fig. 3. **Snapshots of G-series single-stranded dodecanucleotides**. These averaged snapshots were taken as the final structure from the last half nanosecond of 5 ns simulations. DNA backbone is shown in ribbon with 5′- and 3′- termini drawn in red and yellow spheres, respectively. The substituted base is sterically drawn. Inset: helically-stacked single-stranded conformation of the B-form used as starting structure is shown.

where X = G, C, or T, while a minor difference was observed in the homo-A series (*i.e.*, A_6GA_5 < A_6CA_5 < A_{12} < A_6TA_5). In order to test the reproducibility of the final values in the 5-ns simulation, two different and independent setups (starting with different conformations than B-form) were prepared and subjected to MD analysis. For the homo-G and -A series, all four X_6YX_5 type oligodeoxyribonucleotides starting from different initial conformations (*i.e.*, B-form, A-form, and stretched conformation) show basically similar results in terms of half-transition time (see Table 2), *i.e.*, order of transitions, indicating that the long-time-elapsed shape of the oligodeoxyribonucleotides is less dependent on their initial structures. The homo-dodecanucleotide series with a single-base substitution at the 3′-terminal was also examined and analyzed data are listed in Table 1. The most notable observation in this series was that G-substitution resulted in a rapid formation of intramolecular interactions and thus more bent (or folded) structures, which probably corresponds to the low mobility during gel electrophoresis.

End-to-End Distance Fluctuations—All oligodeoxyribonucleotides possessing a single nucleotide difference were well separated on polyacrylamide gels, suggesting that they differ in shape depending on the type of base substitution. To characterize each conformation, the distance between the two ends of each single strand was measured based on the computed trajectories, and these data are shown in the form of a spectrum in Fig. 4. This enables us to recognize differences in the conformation dynamics between highly homologous oligodeoxyribonucleotides. Each homododecanucleotide type (X) has a specific distribution of end-to-end distances depending on the substituted nucleotide type (Y). In general, during electrophoresis, molecules migrate through a matrix of pores which are large enough for the molecule. The DNA-matrix interactions modulated by the structure of DNA influence the absolute mobility (*32*). Therefore, more and/or stronger contacts reduce the electrophoretic mobility of DNA. Consistently, the slowly migrating sequences appear to have more bent structures, which are associated with a retardation effect. As shown in Fig. 4, the distribution of end-to-end distances for the slowly migrating sequences (*i.e.*, a single base substitution by G in each homododecanucleotide type) shifts further toward shorter distances than those of the faster migrating sequences (*i.e.*, a single base substitution by C in each homododecanucleotide type). The possibility of non–W-C intra-molecular base interactions supports this interpretation, since substitution by a G base results in a higher possibility of Hoogsteen-type hydrogen bonding and base-stacking interactions with neighboring nucleotides in a strand than substitution by a non-G base (such as C). This means that substitution by one G in homododecanucleotide series can cause a higher rate of exchange *via* intra-molecular interactions and shift the end-to-end distribution towards shorter distances, *i.e.*, towards compact conformation, while in contrast substitution by one C leads to a lower rate of exchange *via* intra-molecular interactions and, therefore, a higher value of end-to-end distribution. Numerically, the average end-to-end distances for the faster migrating oligodeoxyribonucleotides are larger than those for the more slowly migrating oligodeoxyribonucleotides (see Table 1). These findings support the significant correlation between the experimental results and those of the simulation.

Table 2. **Conserved orders of structural transitions beginning with different initial conformations among oligonunucleotides.**

Mnemonic	Half-transition time (ns)[a]		
	B-form	A-form	Stretched
G-series			
G_{12}	0.5 (1)	0.3 (1)	0.1 (1)
G_6AG_5	0.6 (2)	0.5 (2)	0.3 (2)
G_6TG_5	0.9 (3)	0.6 (3)	0.7 (3)
G_6CG_5	1.4 (4)	1.5 (4)	1.9 (4)
A-series			
A_6GA_5	0.3 (1)	0.3 (1)	0.4 (2)
A_6CA_5	2.1 (2)	1.2 (2)	0.5 (3)
A_{12}	2.9 (3)	1.7 (3)	0.3 (1)
A_6TA_5	4.1 (4)	2.2 (4)	0.6 (4)

[a]Half-transition times observed in the trajectories of r_{ces}. Ordering of transitions are shown in parentheses.

Correspondence of Simulation and Gel Electrophoresis Experiments—Because the simulated values can be supposed to have a relationship with mobility-in-gel during electrophoresis, the average values of end-to-end distance and time ratio for extended versus folded conformations from RMSD trajectories over the entire simulation were calculated and compared with experimental values, *i.e.*, mobility-in-gel (μ) for all oligodeoxyribonucleotides tested here, as shown in Fig. 5a. A parallel relationship was found to exist for almost all values between the experimental values (mobility) and the two types of simulation-derived values ($d_{e\text{-}e}$ and γ_{RMSD}). This is more clearly shown in Fig. 5b by drawing two curves with maximum overlapping, and is extremely unlikely to have happened by chance [The correlation coefficients between the experimental (mobility) and simulated ($d_{e\text{-}e}$) values is 0.93 for the G-series and 0.91 for the C-series]. The correspondence between the snapshot images (shapes) and the mobility-in-gel, as discussed above, suggests that nanosecond MD may be able to depict the structural dynamics of oligodeoxyribonucleotides, and may be used to quantitatively obtain average molecular sizes and shapes. A possible conclusion is that the phenomena (*i.e.*, structural dynamics) observed in the nanosecond simulation-time range may be applicable to estimate the phenomena in the experimental time range (second-to-minute). This may be particularly true for molecules with simple, less stable structures like homooligodeoxyribonucleotides. It should be easy to survey the entire conformational range of these molecules, which is rather diminished due to their relative simplicity.

DISCUSSION

We have developed a novel approach to examine the relationship between the structure and the dynamics of ssDNA in solution by means of gel electrophoresis and molecular dynamics simulation. Based on these data, we can construct the following general model for solution structure dynamics of ssDNA (possibly RNA also). If we define a structure as a molecular state that is held for a fixed

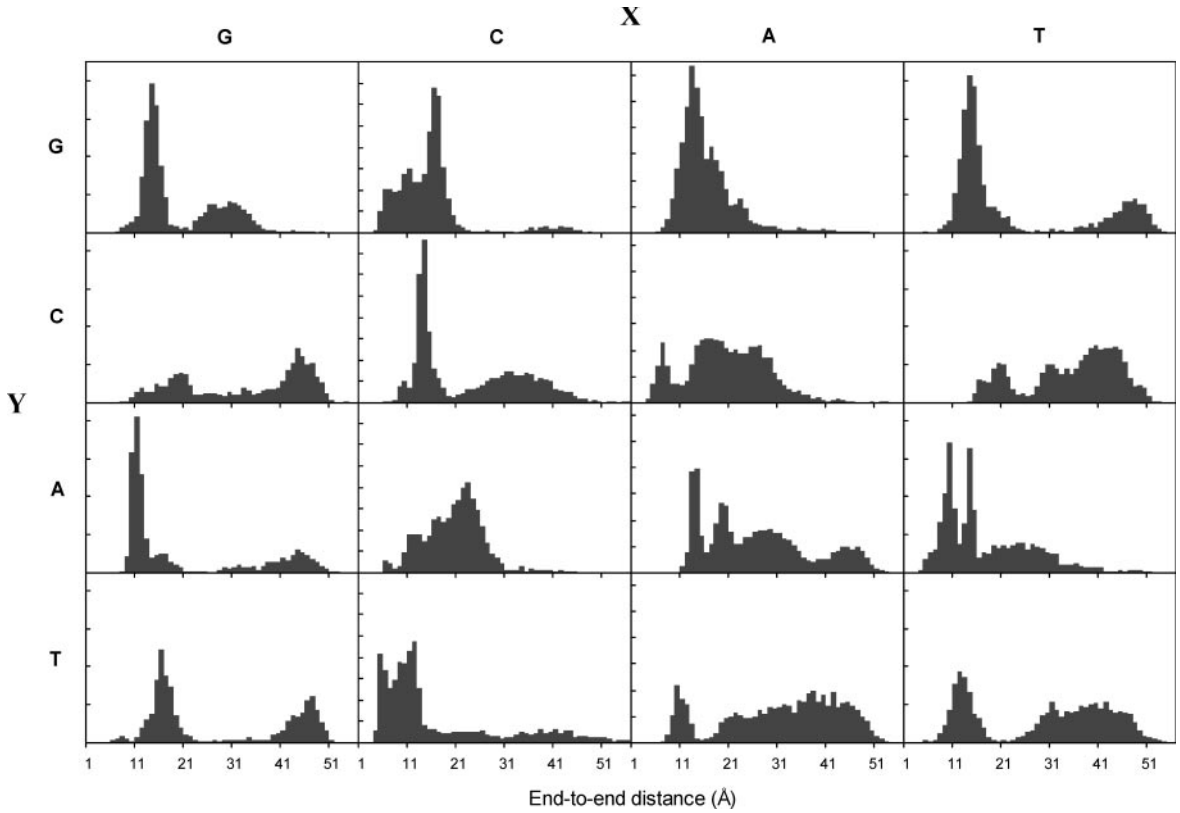


Fig. 4. **Distribution of end-to-end distance of homododecanucleotides possessing a single-base substitution at the center of the sequence.** X denotes the homododecanucleotide type, while Y denotes the base substitution type. Each distribution is calculated from MD simulation of 2 ns (G-series), 5 ns (C- and A-series) or 3 ns (T-series), which corresponds to half-transition period.

amount of time (minimum life time; τ_{min}, though arbitrary), we can count the number of structures a particular molecule of ssDNA will form in a finite period of time (we must adopt an approximation of the identity of a structure. In other words, we have to assign a range of similar structures to a single structure). Some structures are held for longer periods due to the stability of the canonical double-stranded (ds) structure and the hairpin structure; some are held for less time due to shortened or mismatch-containing ds structures (*11*); while others are formed to break via W-C pairings (*4*) or non-W-C pairings (which include both hydrogen bonding and base stacking) as shown in this paper. If a structure S_i of a life-time τ_i appears n_i times within the duration of T_o, then;

$$N_m = \sum_{i=1}^{i_{max}} n_i \quad (1)$$

$$\bar{\tau} = \sum_{i=1}^{i_{max}} \tau_i \cdot n_i / N_m \quad (2)$$

$$\sum_{i=1}^{i_{max}} \tau_i \cdot n_i = \bar{\tau} \cdot N_m \leqslant T_0 \quad (3)$$

where N_m and $\bar{\tau}$ are the total number of structures and the average life time of the structures, respectively, regarding molecule m.

This situation is schematically illustrated in Fig. 6. The distribution of dynamic structures depends on the sequence of the ssDNA and the definition of τ_{min}. Therefore, the novel SSCP phenomenon reported in this paper shows the significance of negligibly weak, previously neglected, intra-molecular interactions between adjacent nucleotides of a strand, which can and do contribute to the distribution of dynamic structures by generating structures having short life times (near τ_{min}). In other words, models that consider only stable structures of ssDNA are not able to reach correct conclusions and often overlook such phenomena as that reported here. [Our preliminary experiments performed by changing gel concentration and operation temperature support this discussion (data not shown): raising temperature and gel concentration (leading to smaller pore sizes) resulted in less distinct mobilities among oligodeoxyribonucleotides, probably meaning that they reduced the contribution of intra-molecular binding interactions to the whole ensemble of structural dynamics. Note that gels of smaller pore sizes lead to higher frequency of inter-molecular interactions between DNA–DNA and DNA–gel (*32*).]

Based on this structural dynamic model of ssDNA (possibly RNA also) in solution, we will now be able to interpret more quantitatively the rates and efficiencies of hybridization-related reactions, such as general PCR (*33*, *34*), RNA interference (*35*), RNA splicing (*36*), ribozyme/deoxyribozyme (*37*), antisense DNA/RNA (*12*, *38*), and sequencing-related phenomena (*27*). For this purpose, the effectiveness of the present ns-MD is very encouraging, because ns-MD may substantiate the entire image of dynamic structures. This rather optimistic prospect is

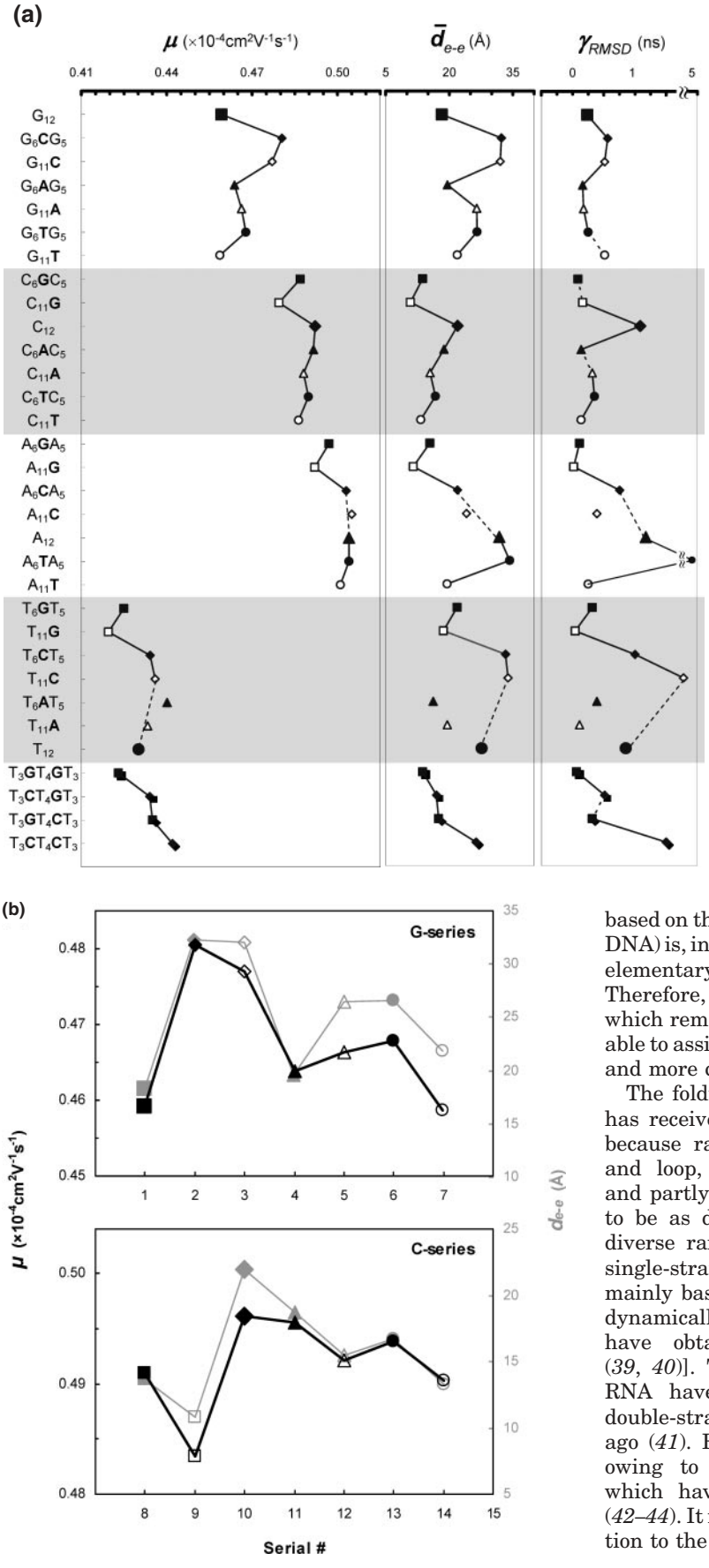


Fig. 5. **Comparison of experimental and simulation results.** (a) Mobility in gel (μ), average end-to-end distance ($d_{e\text{-}e}$) and time ratio for extended and folded conformations from RMSD trajectories are plotted for each dodecanucleotide. All the dodecanucleotide series are aligned in a fixed order: center G (X_6GX_5; filled square), 3′-G ($X_{11}G$; hollow square), center C (X_6CX_5; filled diamond), 3′-C ($X_{11}C$; hollow diamond), center A (X_6AX_5; filled triangle), 3′-A ($X_{11}A$; hollow triangle), center T (X_6TX_5; filled circle), 3′-T ($X_{11}A$; hollow circle), while four homododecanucleotides are shown by enlarged symbols. For μ, the error bars are omitted for clarity (see Table 1). (b) Emphasized presentation of the correspondence. The graph for G and C series is re-drawn with proportionally enlarged and adjusted scale of μ so as to get a maximum fit with the scale of $d_{e\text{-}e}$.

based on the fact that the nature of a large biopolymer (*e.g.*, DNA) is, in principle, similar to the linear combination of its elementary segments (*e.g.*, oligodeoxyribonucleotides). Therefore, detailed analysis of oligodeoxyribonucleotides, which remains possible by this type of simulation, may be able to assist in elucidation of structural dynamics of larger and more complex DNAs/RNAs.

The folding problem of single-stranded DNA and RNA has received less attention than that of proteins, partly because rather primitive structures, such as the stem and loop, seem to be satisfactory for such problems and partly because the roles of DNA/RNA do not appear to be as diverse as those of proteins, which require a diverse range of structures. The current notion is that single-stranded DNA (RNA) forms solution structures mainly based on W-C base pairing, which is the thermodynamically stable conformation [computer algorithms have obtained secondary structures for DNA/RNA (*39*, *40*)]. Therefore, to date, stable structures of DNA/RNA have been investigated since the discovery of double-stranded DNA by Watson & Crick about 50 years ago (*41*). However, the situation is drastically changing owing to the recent discovery of non-coding RNAs, which have been widely found in genome sequences (*42–44*). It is hoped that this paper may help to direct attention to the solution structure dynamics of single-stranded

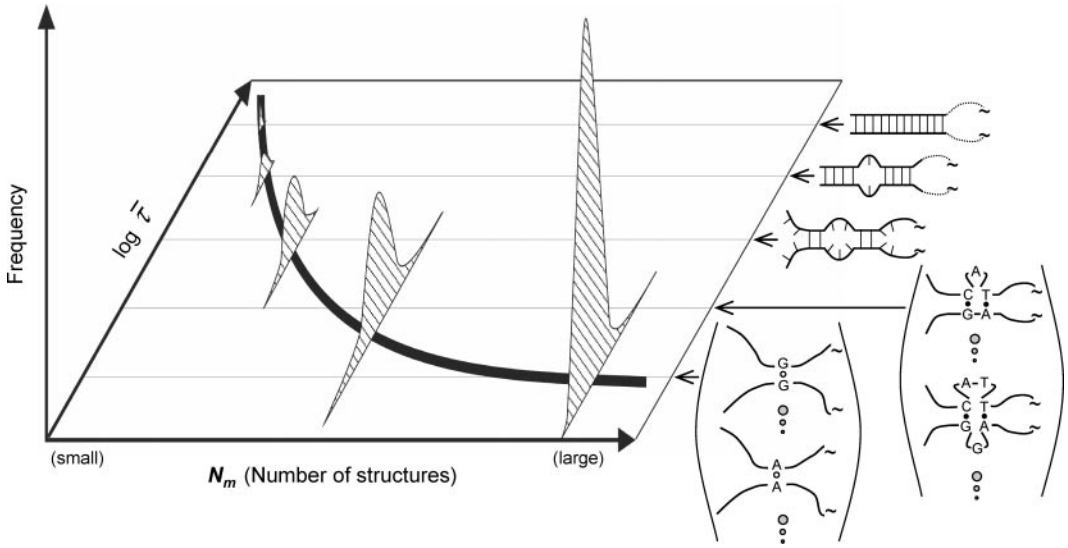


Fig. 6. **Schematic drawing of hypothetical solution structures dynamics of ssDNA (RNA)**. The symbols N_m and $\bar{\tau}$ represent the number of structures of a single-stranded molecule and their average life-time, respectively. The solid hyperbolic curve stands for Eq. 3 (in the text), while the hatched bell-shapes denote the distribution of structures having the average stability of $\bar{\tau}$. The representative structure at each level of $\bar{\tau}$ is shown schematically at right. Conceptually, each structure can be formed either intra- or intermolecularly. For clarity, only W-C and non-W-C base-to-base interactions are included, and other types of possible interactions (such as hydrophobic, stacking or metal ion interaction) are excluded.

DNA (RNA), which must contribute to understanding systematic translational regulations performed by non-coding RNAs and to establishing the way to control gene networking.

CONCLUSIONS

First, we reported the novel phenomenon of non-identical mobility of similar oligodeoxyribonucleotides. Next, we clarified a possible correlation between the conformational properties of oligodeoxyribonucleotides obtained by molecular dynamics simulations and their mobilities obtained by gel electrophoresis. Our results showed that the diversity of molecular conformation of ssDNA is determined by its sequence and is influenced by even a single base substitution. The structural dynamics of ssDNA in solution can be precisely described only by considering its unstable structures, which may be obtained using nanosecond molecular dynamics.

The authors are grateful for the financial support of the joint research project by Aishin, Co. and Taitec, Co. This work was partially supported as a part of the Rational Evolutionary Design of Advanced Biomolecules (REDS) Project. The authors appreciate the preliminary experiments performed by Satoshi Urata and Takashi Aoyama.

REFERENCES

1. Nishigaki, K., Husimi, Y., and Tsubota, M. (1986) Detection of differences in higher order structure between highly homologous single-stranded DNAs by low-temperature denaturant gradient gel electrophoresis. *J. Biochem*. **99**, 663–671
2. Orita, M., Iwahana, H., Kanazawa, H., Hayashi, K., and Sekiya, T. (1989) Detection of polymorphism of human DNA by gel electrophoresis as single strand conformation polymorphism. *Proc. Natl. Acad. Sci. USA* **86**, 2766–2770
3. Nakabayashi, Y. and Nishigaki, K. (1996) Single-strand conformation polymorphism (SSCP) can be explained by semistable conformation dynamics of single-stranded DNA. *J. Biochem.* **120**, 320–325
4. Nishigaki, K., Miura, T., Tsubota, M., Sutoh, A., Amano, N., and Husimi, Y. (1992) Structural analysis of nucleic acids by precise denaturing gradient gel electrophoresis: II. Applications to the analysis of subtle and drastic mobility changes of oligo- and polynucleotides. *J. Biochem.* **111**, 151–156
5. Zhang, Y., Zhou, H., and Ou-Yang, Z.C. (2001) Stretching single-stranded DNA: interplay of electrostatic, base-pairing, and base-pair stacking interactions. *Biophys. J.* **81**, 133–143
6. Dessinges, M.N., Maier, B., Zhang, Y., Peliti, M., Bensimon, D., and Croquette, V. (2002) Stretching single stranded DNA, a model polyelectrolyte. *Phys. Rev. Lett.* **89**, 248102
7. Viovy, J.L. (2000) Electrophoresis of DNA and other polyelectrolytes: Physical mechanisms. *Rev. Modern Phys.* **72**, 813–872
8. Calladine, C.R. and Drew, H.R. (1996) A useful role for "static" models in elucidating the behaviour of DNA in solution. *J. Mol. Biol.* **257**, 479–485
9. Slater, G.W. and Song-Yan, Wu. (1995) Reptation, entropic trapping, percolation, and rouse dynamics of polymers in "random" environments. *Phys. Rev. Lett.* **75**, 164–167
10. Zhou, H., Zhang, Y., and Ou-Yang, Z. (2000) Elastic property of single double-stranded DNA molecules: theoretical study and comparison with experiments. *Phys. Rev. E* **62**, 1045–1058
11. Nishigaki, K., Kaneko, Y., Wakuda, H., Husimi, Y., and Tanaka, T. (1985) Type II restriction endonucleases cleave single-stranded DNAs in general. *Nucleic Acids Res.* **13**, 5747–5760
12. Carrington, J.C. and Ambros, V. (2003) Role of microRNAs in plant and animal development. *Science* **301**, 336–338
13. Biyani, M. and Nishigaki, K. (2003) Sequence-specific and non-specific mobilities of single-stranded oligonucleotides observed by changing the borate buffer concentration. *Electrophoresis* **24**, 628–633
14. Weiner, P.K. and Kollman, P.A. (1981) AMBER: assisted model building with energy refinement. A general program for

modeling molecules and their interactions. *J. Comput. Chem.* **2**, 287–303
15. Case, D.A. *et al*. (2002) *AMBER7 Manual*, University of California, San Francisco, CA
16. Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, K.M., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W., and Kollman, P.A. (1995) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **117**, 5179–5197
17. Still, W.C., Tempczyk, A., Hawley, R.C., and Hendrickson, T. (1990) Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* **112**, 6127–6129
18. Tsui, V. and Case, D.A. (2000) Molecular dynamics simulations of nucleic acids with a generalized Born solvation model. *J. Am. Chem. Soc.* **122**, 2489–2498
19. Dominy, B. and Brooks, C. III (1999) Development of a generalized Born model parameterization for proteins and nucleic acids. *J. Phys. Chem. B* **103**, 3765–3773
20. Calimat, N., Schaefer, M., and Simonson, T. (2001) Protein molecular dynamics with the generalized Born/ACE solvent model. *Proteins* **45**, 144–158
21. Xia, B., Tsui, V., Case, D.A., Dyson, H.J., and Wright, P.E. (2002) Comparison of protein solution structures refined by molecular dynamics simulation in vacuum, with a generalized Born model, and with explicit water. *J. Biomol. NMR* **22**, 317–331
22. Ryckaert, J.P., Ciccotti, G., and Bereendsen, H.J.C. (1977) Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**, 327–341
23. Tsui, V. and Case, D.A. (2001) Theory and applications of the Generalized Born solvation model in macromolecular simulations. *Biopolymers* **56**, 275–291
24. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: Visual molecular dynamics. *J. Mol. Graphics* **14**, 33–38
25. Sayle, R.A. and Milner-White, E.J. (1995) RASMOL: biomolecular graphics for all. *Trends Biochem. Sci.* **20**, 374
26. Hirao, I., Nishimura, Y., Naraoka, T., Watanabe, K., Arata, Y., and Miura, K. (1989) Extraordinary stable structure of short single-stranded DNA fragments containing a specific base sequence: d(GCGAAAGC). *Nucleic Acids Res.* **17**, 2223–2231
27. Bowling, J.M., Bruner, K.L., Cmarik, J.L., and Tibbetts, C. (1991) Neighboring nucleotide interactions during DNA sequencing gel electrophoresis. *Nucleic Acids Res.* **19**, 3089–3097
28. Sanger, F. (1981) Determination of nucleotide sequences in DNA. *Science* **214**, 1205–1210
29. Baba, Y., Tsuhako, M., Sawa, T., and Akashi, M. (1993) Effect of urea concentration on the base-specific separation of oligodeoxynucleotides in capillary affinity gel electrophoresis. *J. Chromatogr. A* **652**, 93–99
30. Sen, S. and Nilsson, L. (2001) MD simulations of homomorphous PNA, DNA, and RNA single strands: characterization and comparison of conformations and dynamics. *J. Am. Chem. Soc.* **123**, 7414–7422
31. Martinez, J.M., Elmroth, S.K., and Kloo, L. (2001) Influence of sodium ions on the dynamics and structure of single-stranded DNA oligomers: a molecular dynamics study. *J. Am. Chem. Soc.* **123**, 12279–12289
32. Niederweis, M., Lederer, T., and Hillen, W. (1994) Matrix effects suggest an important influence of DNA-polyacrylamide interactions on the electrophoretic mobility of DNA. *J. Biol. Chem.* **269**, 10156–10162
33. Sakuma, Y. and Nishigaki, K. (1994) Computer prediction of general PCR products based on dynamical solution structures of DNA. *J. Biochem.* **116**, 736–741
34. Nishigaki, K., Saito, A., Hasegawa, T., and Naimuddin, M. (2000) Whole genome sequence-enabled prediction of sequences performed for random PCR products of Escherichia coli. *Nucleic Acids Res.* **28**, 1879–1884
35. Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998) Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **391**, 806–811
36. Sharp, P.A. (1994) Split genes and RNA splicing. *Cell* **77**, 805–815
37. Cech, T.R. (1993) Structure and mechanism of the large catalytic RNAs: Group I and Group II introns and ribonuclease P in *The RNA World* (Gesteland, R.F. and Atkins, J.F., eds.) pp. 239–269, Cold Spring Harbor Press
38. Stein, C.A. and Cheng, Y.C. (1993) Antisense oligonucleotides as therapeutic agents—is the bullet really magical? *Science* **261**, 1004–1012
39. Zuker, M. and Stiegler, P. (1981) Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucleic Acids Res.* **9**, 133–148
40. Nussinov, R. and Jacobason, A.B. (1980) Fast algorithm for predicting the secondary structure of single-stranded RNA. *Proc. Natl. Acad. Sci. USA* **77**, 6903–6913
41. Watson, J.D. and Crick, F.H.C. (1953) A structure for Deoxyribose Nucleic Acid. *Nature* **171**, 737–738
42. Eddy, S.R. (2001) Non-coding RNA genes and the modern RNA world. *Nat. Rev. Genet.* **2**, 919–929
43. Rivas, E., Klein, R.J., Jones, T.A., and Eddy, S.R. (2001) Computational identification of noncoding RNAs in E. coli by comparative genomics. *Curr. Biol.* **11**, 1369–1373
44. McCutcheon, J.P. and Eddy, S.R. (2003) Computational identification of non-coding RNAs in Saccharomyces cerevisiae by comparative genomics. *Nucleic Acids Res.* **31**, 4119–4128
45. Bassam, B.J., Caetano-Anolles, G., and Gresshoff P.M. (1991) Fast and sensitive silver staining of DNA in polyacrylamide gels. *Anal. Biochem.* **196**, 80–83
46. Biyani, M. and Nishigaki, K. (2005) Structural characterization of ultra-stable higher-ordered aggregates generated by novel guanine-rich DNA sequences. *Gene* (in press)