3420 *Biochemistry* **2010,** *49,* 3420–3426
DOI: 10.1021/bi902047x

BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Probing the Transition State for Nucleic Acid Hybridization Using Φ-Value Analysis[†]

Jandi Kim and Jong-Shik Shin*

Department of Biotechnology, Yonsei University, Shinchon-Dong 134, Seodaemun-Gu, Seoul 120-749, Korea

Received December 1, 2009; Revised Manuscript Received February 10, 2010

ABSTRACT: Genetic regulation by noncoding RNA elements such as microRNA and small interfering RNA (siRNA) involves hybridization of a short single-stranded RNA with a complementary segment in a target mRNA. The physical basis of the hybridization process between the structured nucleic acids is not well understood primarily because of the lack of information about the transition-state structure. Here we use transition-state theory, inspired by ϕ-value analysis in protein folding studies, to provide quantitative analysis of the relationship between changes in the secondary structure stability and the activation free energy. Time course monitoring of the hybridization reaction was performed under pseudo-steady-state conditions using a single fluorophore. The ϕ-value analysis indicates that the native secondary structure remains intact in the transition state. The nativelike transition state was confirmed via examination of the salt dependence of the hybridization kinetics, indicating that the number of sodium ions associated with the transition state was not substantially affected by changes in the native secondary structure. These results propose that hybridization between structured nucleic acids undergoes a transition state leading to formation of a nucleation complex and then is followed by sequential displacement of preexisting base pairings involving successive small energy barriers. The proposed mechanism might provide new insight into physical processes during small RNA-mediated gene silencing, which is essential to selection of a target mRNA segment for siRNA design.

Formation of secondary structures in a reliable manner driven by sequence-specific hybridization has made nucleic acid construction materials suitable for storage and transmission of genetic information (*1*). Recent discoveries on small silencing RNAs, including microRNA, small interfering RNA (siRNA),[1] and Piwi-interacting RNA, reveal that nucleic acids do not merely serve as a passive information carrier but actively participate in gene regulation (*2*−*5*). The RNA interference involves sequence-dependent hybridization of a short single-stranded guide RNA with a complementary region in the target mRNA of which gene expression is to be silenced (*5*−*8*). It is known that the gene silencing efficiency is greatly dependent on the local secondary structure of the target region (*9*−*15*). This suggests that hybridization kinetics combined with local thermodynamic stability in the context of mRNA secondary structures are important in determining the gene silencing efficiency. Therefore, elucidating the physical basis for target recognition and gene silencing efficacy in RNA interference requires a thorough understanding of the molecular details of hybridization between structured nucleic acids.

Despite extensive research efforts, developing a generally applicable kinetic model that enables prediction of the effect of secondary structure on hybridization kinetics remains a challenge (*16*−*19*). One of the major hurdles to developing such a model results from the fact that nucleic acids usually adopt multiple secondary structures rather than a single minimum free energy structure (*14*). Therefore, the solution structure of a given oligonucleotide should be considered as a Boltzmann ensemble of all possible structures separated by substantial energy barriers. As a result, the hybridization process would occur through multiple reaction pathways, reminiscent of a protein folding process (*20*). Another problem in deriving the kinetic model is that any structural information about the transition state for the hybridization reaction is not yet available. The transition-state structure is crucial to identification of a rate-determining step, permitting prediction of the activation free energy that determines a rate constant. For example, if the transition state retains most of the preexisting self-structures, the rate constant would be determined mainly by the nucleation step and the ensuing zipping up combined with self-structure destruction would not significantly affect the overall rate. However, there have been no prior studies that quantitatively investigate what the transition state for the hybridization reaction looks like.

It is known that hybridization between unstructured oligonucleotides (i.e., devoid of intramolecular base pairing) follows a two-state model and the nucleation step involving interstrand base pairing with a length of a few base pairs is a rate-determining step (*21*). However, the presence of secondary structures renders the hybridization mechanism much more complicated. Previous studies have shown that a zipper model (i.e., nucleation followed by zippering), instead of the two-state model, permits better interpretations of kinetic data, which indicates that the hybridization process between structured oligonucleotides involves multiple energy barriers flanking a nucleation complex (*16*, *17*, *22*). Because of the ignorance of the transition-state structure, it is not clear whether the transition state occurs prior to the nucleation complex. Recently, we showed that extending a single-stranded region

[†]This work was supported by the BK21 program of the Korean Ministry of Education, start-up research funds provided by Yonsei University, and Seoul R&BD Programs (NT080612 and KU080657).

*To whom correspondence should be addressed: Engineering Building 2 (Room 507), Yonsei University, Shinchon-Dong 134, Seodaemun-Gu, Seoul 120-749, South Korea. E-mail: enzymo@yonsei.ac.kr. Telephone: (+82) 2-2123-5884. Fax: (+82) 2-362-7265.

[1]Abbreviations: siRNA, small interfering RNA; G, guide strand; R, reference strand; V, variant strand; F, folded structure; U, unfolded structure; P, final hybridization product; NC, nucleation complex.

 Published on Web 03/08/2010

adjacent to a nucleation site directly increased the activation energy, leading us to hypothesize that nucleation is a rate-determining step for hybridization between structured oligonucleotides (*23*). However, this hypothesis should be verified by additional studies. First, the activation free energy should not be significantly affected by the thermodynamic stability of preexisting secondary structures that are to be disrupted during a zipping up process. Second, the preexisting secondary structures should be intact in the transition state because nucleation occurs preferentially at single-stranded regions.

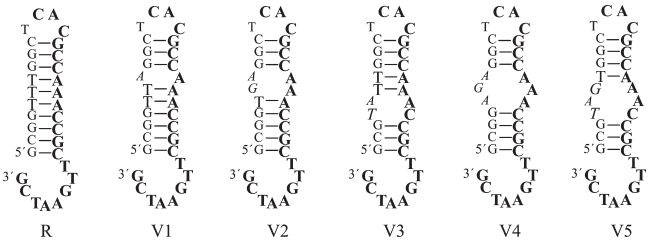
Addressing such issues requires a systematic analysis of thermodynamic and kinetic data obtained with structurally defined oligonucleotides. We noticed that the hybridization process resembles protein folding in that the approach to the final state is driven by gradual energy gain obtained from structural changes leading to the most stable structure. The transition state for protein folding has been successfully studied by ϕ-value analysis (*20*, *24*−*26*), which led us to use the same approach to characterize the transition state for nucleic acid hybridization. In the ϕ-value analysis, stability changes caused by amino acid substitutions were measured both in a transition state and in a folded state via observation of unfolding and refolding kinetics. If the region around the mutation site is already folded in the transition state, the transition state would exhibit a stability change similar to that of the folded state. By applying the ϕ-value analysis to nucleic acid hybridization, we found that the preexisting self-structure is completely retained in the transition state. This result clearly shows that formation of a nucleation complex is a rate-determining step, which provides insight into effective siRNA design.

MATERIALS AND METHODS

DNA Oligonucleotides. All DNA oligonucleotides (HPLC-purified grade) were purchased from IDT, Inc. (Coralville, IA). Labeling of the G strand with Oregon Green 488 was performed by the supplier. DNA stock solutions (100 μM) were prepared in TE buffer (pH 8.0) (10 mM Tris and 1 mM EDTA).

Folding Calculation of DNA Oligonucleotides. DNA secondary structures and thermodynamic stability were determined by mfold (*27*) which uses free energy data from SantaLucia and Hicks (*28*). Unless otherwise specified, the folding temperature and NaCl concentration were set to 25 °C and 300 mM, respectively. The default setting was used for other folding parameters, including the percent suboptimality number (5), the upper bound on the number of computed folding (50), the window parameter (0 for G and 1 for R and V), and the maximum distance between paired bases (no limit).

Kinetic Measurements and Analysis. Fluorescence measurements were performed as described previously with a fluorometer equipped with a Peltier temperature controller and a photomultiplier detector (PTI Co.) (*23*). Unless otherwise specified, the hybridization reactions (working volume of 100 μL) were performed in a sodium phosphate buffer (50 mM, pH 6.5) supplemented with 250 mM NaCl. In the experiments that aimed to examine the salt dependence of the hybridization reactions, 50−250 mM NaCl was a supplement for the same sodium phosphate buffer. Typical reaction conditions were as follows: 5 nM Oregon Green 488-labeled G strand and 50 nM R or V strand. Excitation and emission wavelengths were set to 496 and 517 nm, respectively. Changes in the fluorescence emission were recorded for 100 s.



G: 5′-CGATTCAAGCGGTTTGGCGTG-Oregon green 488

FIGURE 1: DNA sequences of the oligonucleotides used in this study. Secondary structures for R and V strands were predicted by mfold.

To obtain hybridization rate constants (k_{hyb}), kinetic traces from the hybridization reaction under pseudo-steady-state conditions (i.e., $[V]_0 \gg [G]_0$) were subjected to curve fitting. Under the reaction conditions, the reaction follows pseudo-first-order kinetics.

$$d[G]/dt = -k_{hyb}[V]_0[G] \quad (1)$$

Integration of eq 1 yields a single-exponential function for [G] whose exponent is linearly proportional to $[V]_0$.

$$[G] = [G]_0 e^{-k_{hyb}[V]_0 t} \quad (2)$$

Because both the G strand and the final hybridization product (denoted by P) contribute to fluorescence emission, the total fluorescence intensity with respect to time is

$$f(t) = \varepsilon_P[P] + \varepsilon_G[G], \quad (3)$$

where ε is the molar fluorescence intensity. Because [P] equals $[G]_0 - [G]$, eq 3 becomes

$$f(t) = \varepsilon_P[G]_0 + (\varepsilon_G - \varepsilon_P)[G]_0 e^{-k_{hyb}[V]_0 t} \quad (4)$$

Equation 4 was used for curve fitting to obtain k_{hyb} values.

RESULTS AND DISCUSSION

Oligonucleotide Design. In the ϕ-value analysis in the study of the transition state for protein folding, stability changes in the native structure caused by amino acid substitutions were compared with those in the transition state (*20*, *24*−*26*). Protein folding is a unimolecular process, whereas the hybridization reaction is a bimolecular one involving two oligonucleotides. Therefore, to apply the ϕ-value analysis to the hybridization reaction, we introduced structural perturbation into only one oligonucleotide. To this end, we designed a series of structurally defined oligonucleotides (the reference and variant strands, denoted by R and V, respectively) consisting of a hairpin region serving as a structural framework and an eight-base dangling end serving as a nucleation site (Figure 1). The R and V strands share an identical segment (shown in bold type) perfectly complementary to a 21-base strand (denoted by G), so the hybridization with the G strand results in formation of a 21 bp duplex with an 11-base overhang. Base changes introduced into V strands (shown in italic) induce a perturbation in secondary structure stability, leading to generation of an internal loop. The number of the base changes was limited to three, so the structural perturbation in V strands is not high enough to impair the overall hairpin structure. The mfold calculations (*27*), in which the maximal energy difference allowed was set to 30%, predicted only one structure for R and V strands as shown in Figure 1.

In contrast to the R and V strands, the G strand was designed not to form any stable secondary structure. The mfold calculation

showed that the free energy of the predicted secondary structure is lower than that of the unfolded structure by 0.3 kcal/mol, which is smaller than the thermal energy RT (i.e., 0.6 kcal/mol) (*23*). Therefore, we envisaged that alteration of the hybridization kinetics of the G strand with different V strands results only from the structural perturbation introduced into the hairpin region of V strands.

Kinetic and Thermodynamic Analysis. We used a single fluorophore as a probe to monitor the hybridization reaction as described previously (*23*). It is known that fluorescence emission of a dye covalently attached to oligonucleotides is influenced by chemical environments, including base composition and secondary structures in the vicinity of the dye (*29*−*32*). In this study, we labeled the 3′ end of the G strand, which is an opposite end to the predesigned nucleation site, with Oregon Green 488 to detect the final hybridization product selectively over the nucleation complex. The fluorescence emission from the G strand decreased by ~25% at λ_{max} (i.e., 517 nm) upon hybridization with the R strand (Figure 2).

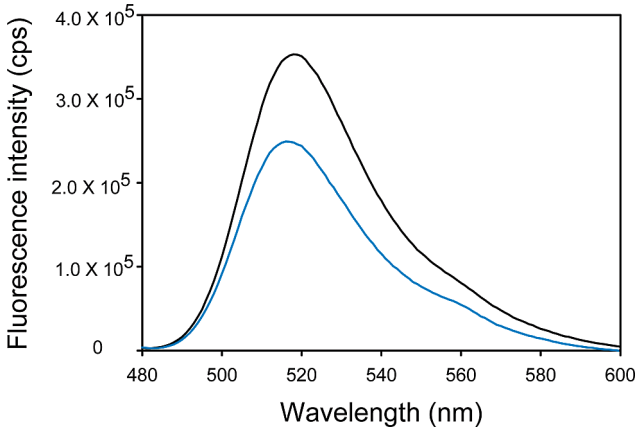


FIGURE 2: Change in the fluorescence emission of the G strand (black) upon hybridization with the R strand (blue).

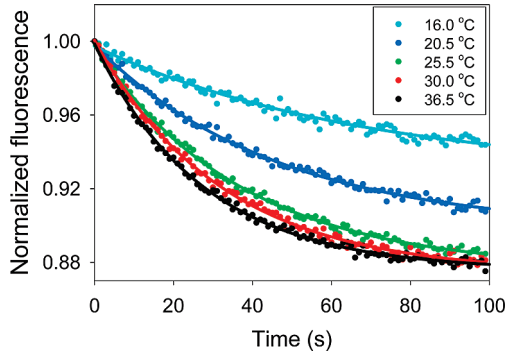


FIGURE 3: Kinetic traces for the hybridization reaction between R (50 nM) and G (5 nM) strands at different temperatures. Solid lines represent fitting curves.

Time course monitoring of the hybridization reaction under pseudo-steady-state conditions yielded exponentially decreasing kinetic traces that fit well to a single-exponential function (Figure 3). We measured hybridization rate constants (k_{hyb}) at different temperatures (16−37 °C) as listed in Table 1, so activation enthalpy ($\Delta H^{\ddagger}$) and activation entropy ($\Delta S^{\ddagger}$) were obtained from Eyring plots (Figure 4) using

$$\ln \frac{k_{hyb}}{T} = \ln \frac{\kappa k_B}{h} + \frac{\Delta S^{\ddagger}}{R} - \frac{\Delta H^{\ddagger}}{R}\frac{1}{T} \quad (5)$$

where T is the temperature, κ is the transmission coefficient, k_B is Boltzmann's constant, h is Planck's constant, and R is the gas constant (*24*). As shown in Figure 4, the linear regression analysis showed good correlations between $1/T$ and $\ln(k_{hyb}/T)$, although a slight curvature occurred for R ($r^2 = 0.91$), V1 ($r^2 = 0.93$), and V5 ($r^2 = 0.88$) strands. Unlike the activation enthalpy, the activation entropy could not be determined directly from the Eyring plot because the value of κ is unknown. In the ϕ-value analysis for protein folding, the unknown κ cancels out in the calculation of changes in activation free energy (*24*). We used the same strategy in this study to deal with the unknown κ value.

Table 2 shows the thermodynamic parameters [$\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}/R + \ln(\kappa k_B/h)$] obtained from the kinetic measurements and the secondary structure stability of R and V strands (ΔG^F, in which superscript F represents a folded structure) calculated by mfold. The ΔG^F values represent the free energy differences between the folded state adopting a secondary structure and the unfolded state devoid of intramolecular base pairing. The ΔG^F values showed correlations with neither $\Delta H^{\ddagger}$ nor $\Delta S^{\ddagger}/R + \ln(\kappa k_B/h)$. Because $\Delta S^{\ddagger}$ values could not be obtained from the kinetic

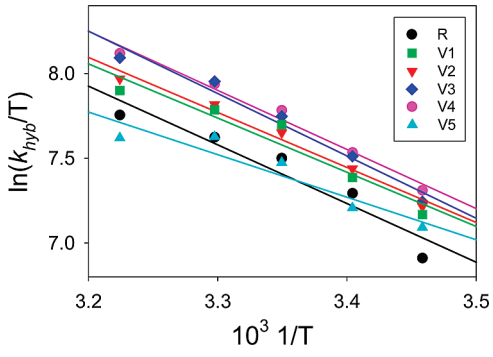


FIGURE 4: Eyring plots for determining the activation enthalpy and entropy. The unit of temperature is kelvin.

Table 1: Hybridization Rate Constants (k_{hyb}) Determined by Kinetic Analysis[a]

	k_{hyb} ($\times 10^5$ M^{-1} s^{-1})[b]					
T (°C)	R	V1	V2	V3	V4	V5
16.0	2.90 ± 0.18	3.75 ± 0.16	3.93 ± 0.21	4.04 ± 0.22	4.33 ± 0.24	3.48 ± 0.18
20.5	4.32 ± 0.28	4.74 ± 0.33	4.99 ± 0.30	5.37 ± 0.28	5.49 ± 0.22	3.96 ± 0.29
25.5	5.39 ± 0.25	6.59 ± 0.41	6.28 ± 0.24	6.91 ± 0.43	7.16 ± 0.33	5.26 ± 0.25
30.0	6.20 ± 0.59	7.29 ± 0.37	7.52 ± 0.32	8.62 ± 0.37	8.49 ± 0.46	6.22 ± 0.43
36.5	7.23 ± 0.61	8.35 ± 0.45	8.94 ± 0.31	10.12 ± 0.46	10.40 ± 0.35	6.31 ± 0.47

[a]The concentrations of G and V (or R) strands used for kinetic measurements were 5 and 50 nM, respectively. [b]The k_{hyb} values are means of three independent measurements.

Table 2: Secondary Structure Stability and Thermodynamic Parameters for the Hybridization Reactions

	R	V1	V2	V3	V4	V5
ΔG^{F} (kcal/mol)[a]	−14.49	−10.27	−9.01	−8.02	−6.54	−5.16
$\Delta H^{‡}$ (kcal/mol)[b]	7.03 ± 1.21	6.50 ± 1.03	6.58 ± 0.53	7.45 ± 0.75	7.08 ± 0.54	5.10 ± 1.05
$\Delta S^{‡}/R + \ln(\kappa k_B/h)$[b]	19.24 ± 2.04	18.53 ± 1.73	18.69 ± 0.90	20.26 ± 1.26	19.63 ± 0.91	15.99 ± 1.77

[a]The ΔG^{F} values (i.e., secondary structure stability) were calculated by mfold at 25 °C and 300 mM NaCl. [b]The values were determined by the Eyring equation.

Table 3: Differences in the Thermodynamic Parameters for the Hybridization Reaction

	V1	V2	V3	V4	V5
$\Delta\Delta G^{F}_{V-R}$ (kcal/mol)[a]	4.22	5.48	6.47	7.95	9.33
$\Delta\Delta H^{‡}_{VG-RG}$ (kcal/mol)	−0.53	−0.45	0.42	0.05	−1.93
$\Delta\Delta S^{‡}_{VG-RG}$ (cal mol^{-1} K^{-1})[b]	−1.41	−1.09	2.03	0.77	−6.47
$\Delta\Delta G^{‡}_{VG-RG}$ (kcal/mol)[c]	−0.11	−0.13	−0.19	−0.18	0.00

[a]The $\Delta\Delta G^{F}_{V-R}$ values (i.e., differences in the secondary structure stability of V strands in reference to the R strand) correspond to $\Delta G^{F}_{V} - \Delta G^{F}_{R}$. [b]The $\Delta\Delta S^{‡}_{VG-RG}$ values were calculated using the $\Delta S^{‡}/R + \ln(\kappa k_B/h)$ values listed in Table 2. [c]The values were calculated with the equation $\Delta\Delta G^{‡}_{VG-RG} = \Delta\Delta H^{‡}_{VG-RG} - T\Delta\Delta S^{‡}_{VG-RG}$.

measurements because of the uncertainty of the κ value, we could not determine the activation free energy ($\Delta G^{‡}$). We did not use the $\kappa \approx 1$ approximation to obtain $\Delta S^{‡}$ values because there is no theoretical or experimental estimate for the κ value in the nucleic acid hybridization. Instead, we calculated differences in the activation entropy ($\Delta\Delta S^{‡}$) using the R strand as a reference to cancel out the unknown κ value, so the activation free energy ($\Delta\Delta G^{‡}$) could be obtained. This calculation is based on the assumption of a constant κ value by considering the structural similarity of V strands to the R strand. We presumed that the internal loop in V strands does not significantly alter the transition-state structure. Table 3 shows the differences in the thermodynamic parameters. For example, $\Delta\Delta H^{‡}_{VG-RG}$ represents $\Delta H^{‡}_{VG} - \Delta H^{‡}_{RG}$, in which VG and RG represent the hybridization between V and G and the hybridization between R and G, respectively.

Theoretical Considerations for φ-Value Analysis. Using the differences in the free energies listed in Table 3 (i.e., $\Delta\Delta G^{F}_{V-R}$ and $\Delta\Delta G^{‡}_{VG-RG}$), we conducted a ϕ-value analysis to evaluate how much native secondary structure of the R strand remains in the transition state. This analysis has been extensively used in the protein folding study to probe the degree of nativeness of the transition state (*20*, *24*−*26*). To determine the ϕ value, we considered the base changes introduced into V strands as destabilizing mutations affecting free energy levels of both a folded structure and a transition state. As used in the ϕ-value analysis for protein folding, we reasoned that the ratio between changes in the energy levels of the folded structures and the transition states might indicate the degree of nativeness in the transition state around the mutation site (*20*, *24*−*26*). For example, a ϕ value close to the unity would indicate a nativelike transition state around the mutation site.

A free energy diagram for the hybridization reaction is shown in Figure 5. Starting at folded structures (denoted by F), the hybridization proceeds through a transition state (denoted by ‡) and leads to a final hybridization product (denoted by P). To clarify the definition of secondary structure stability (ΔG^{F}), the

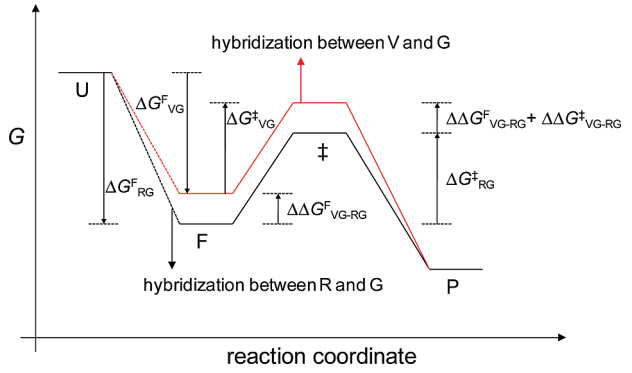


FIGURE 5: Free energy diagram for the hybridization reaction. Arrows designate the direction of structural change relevant to the free energy change.

energy level of an unfolded structure (denoted by U) is included in Figure 5.

Unlike ‡ and P that are bimolecular complexes, F and U consist of two physically independent oligonucleotides. Therefore, the energy levels of F and U represent summation of individual free energies as follows.

$$G^{F}_{RG} = G^{F}_{R} + G^{F}_{G} \text{ and } G^{F}_{VG} = G^{F}_{V} + G^{F}_{G} \quad (6)$$

$$G^{U}_{RG} = G^{U}_{R} + G^{U}_{G} \text{ and } G^{U}_{VG} = G^{U}_{V} + G^{U}_{G} \quad (7)$$

Because of the destabilizing mutations in V strands, energy levels of the folded structure and the transition state for the hybridization between V and G strands (depicted by the red line) are higher than the corresponding energy levels for the hybridization between R and G strands (depicted by the black line). In contrast, the destabilizing mutations are thought not to substantially affect the energy levels of U and P because the mutation sites are located in the single-stranded region in both structures. We ignored the free energy changes caused by base substitutions in a single-stranded region, which is a generally accepted assumption for nucleic acid folding calculations (*27*). The assumption that the energy level of U is not affected by a destabilizing mutation has also been used in the ϕ-value analysis of protein folding (*24*−*26*).

We defined the ϕ value as how much of the stability perturbation introduced into the folded structure is retained in the transition state.

$$\phi = (G^{‡}_{VG} - G^{‡}_{RG})/(G^{F}_{VG} - G^{F}_{RG}) \quad (8)$$

In the definition of the ϕ value, the hybridization between R and G was used as a reference. It is crucial in the calculation of the ϕ value to relate the energy differences in eq 8 to the $\Delta\Delta G^{F}_{V-R}$ and $\Delta\Delta G^{‡}_{VG-RG}$ values listed in Table 3. The denominator of eq 8 can be related to $\Delta\Delta G^{F}_{V-R}$. First, $G^{F}_{VG} - G^{F}_{RG}$ is expressed by individual free energies of V and R

strands using eq 6

$$G^{F}_{VG} - G^{F}_{RG} = G^{F}_{V} - G^{F}_{R} \quad (9)$$

Second, as shown in Figure 5, $\Delta\Delta G^{F}_{VG-RG}$ is equal to $\Delta G^{F}_{VG} - \Delta G^{F}_{RG}$ which can be expressed by free energies of unfolded and folded structures

$$\Delta\Delta G^{F}_{VG-RG} = (G^{F}_{VG} - G^{U}_{VG}) - (G^{F}_{RG} - G^{U}_{RG}) \quad (10)$$

By using eqs 6 and 7, eq 10 becomes

$$\Delta\Delta G^{F}_{VG-RG} = (G^{F}_{V} - G^{F}_{R}) - (G^{U}_{V} - G^{U}_{R}) = G^{F}_{V} - G^{F}_{R} \quad (11)$$

because G^{U}_{V} is equal to G^{U}_{R} as discussed before.

Rearranging eq 11 leads to

$$\Delta\Delta G^{F}_{VG-RG} = (G^{F}_{V} - G^{U}_{V}) - (G^{F}_{R} - G^{U}_{R}) = \Delta\Delta G^{F}_{V-R} \quad (12)$$

Taken together with eqs 9 and 11, eq 12 indicates that the denominator in eq 8 is equal to $\Delta\Delta G^{F}_{V-R}$.

Likewise with eq 10, $\Delta\Delta G^{‡}_{VG-RG}$ is expressed by

$$\Delta\Delta G^{‡}_{VG-RG} = \Delta G^{‡}_{VG} - \Delta G^{‡}_{RG} = (G^{‡}_{VG} - G^{F}_{VG}) - (G^{‡}_{RG} - G^{F}_{RG}) \quad (13)$$

Substitution of G^{F}_{VG} and G^{F}_{RG} with eq 6 yields

$$\Delta\Delta G^{‡}_{VG-RG} = (G^{‡}_{VG} - G^{‡}_{RG}) - (G^{F}_{V} - G^{F}_{R}) \quad (14)$$

Rearranging eq 14 gives

$$G^{‡}_{VG} - G^{‡}_{RG} = \Delta\Delta G^{F}_{V-R} + \Delta\Delta G^{‡}_{VG-RG} \quad (15)$$

Therefore, the ϕ value can be expressed by the free energy values listed in Table 3.

$$\phi = (\Delta\Delta G^{F}_{V-R} + \Delta\Delta G^{‡}_{VG-RG})/\Delta\Delta G^{F}_{V-R} \quad (16)$$

ϕ-Value Analysis of Nucleic Acid Hybridization. As shown in Figure 6, the Brønsted plot shows a linear correlation between $\Delta\Delta G^{F}_{V-R}$ and $\Delta\Delta G^{F}_{V-R} + \Delta\Delta G^{‡}_{VG-RG}$, supporting the idea that the constant κ was a reasonable assumption. The slope obtained from the plot (i.e., the ϕ value) was 1.02 (r^2 = 0.99), suggesting that the native structure remains intact in the transition state around the mutation sites. Because the mutations were introduced into the stem region of the R strand, this result indicates that the preexisting base pairings remain intact in the transition state even though the stem region should be disrupted by strand displacement for completion of the hybridization reaction. This finding clearly supports the general notion that formation of a nucleation complex is a rate-determining step (*16*, *17*, *21*, *22*).

To confirm the nativelike transition state as indicated by the ϕ-value analysis, we examined the effect of salt concentration on the hybridization kinetics. As shown in Figure 7, ln k_{hyb} exhibited linear correlations with ln[Na^+]. It is known that the slope from the plot corresponds to the change in the number of bound sodium ions during a rate-determining step (*33*). The ln k_{hyb}/ln[Na^+] values determined from Figure 7 were 1.22 (for R), 1.21 (for V1), 1.10 (for V2), 1.13 (for V3), 1.18 (for V4) ,and 1.16 (for V5), yielding a mean value of 1.17 ± 0.05. The result indicates that, compared to the native structure, approximately 1.2 more sodium ions are required to neutralize negative charges that

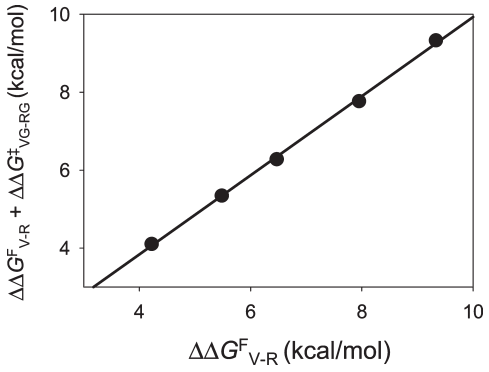


FIGURE 6: Brønsted plot to obtain a ϕ value for the hybridization reaction.

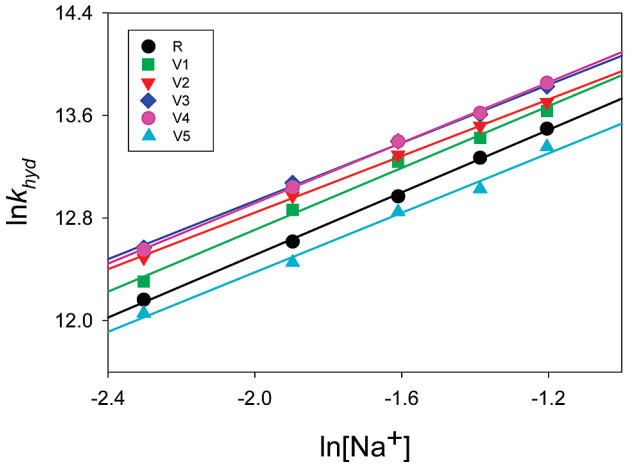


FIGURE 7: Effect of sodium ion concentration on the hybridization kinetics at 37 °C.

accumulate in the transition state. This number is close to 1.4 which was previously measured for formation of a duplex of d(GGAATTCC) (*33*), suggesting that the numbers of newly formed base pairs in the transition states for this study are similar with that for duplex formation. It is noteworthy that the length of the nucleation site in this study is also eight nucleotides. The similar ln k_{hyb}/ln[Na^+] values in this study and for formation of the duplex of d(GGAATTCC) are in agreement with the recent transition path simulation study which revealed that the nucleation event involves molecular sites with approximately 4 bp (*34*). Altogether, it is obvious that the transition states do not involve structural changes in the hairpin region but the similar degree of structural change at the nucleation site.

Proposed Kinetic Model. The result of the high ϕ value close to unity gives a clue about a kinetic mechanism of the hybridization between structured nucleic acids (Figure 8). The hybridization process usually involves destruction of preexisting secondary structures. There are two possible mechanisms for strand displacement after a nucleation step (*35*): a dissociative pathway (simultaneous base pairings after complete melting of local secondary structures) and a sequential displacement pathway (zipping up from a nucleation site). The dissociative pathway should involve an additional high-energy transition state flanked by a nucleation complex and a final hybridization product. This putative transition state is likely to cost even more activation free energy than the nucleation step, because activation energies for dehybridization reactions are known to be much higher than those for hybridization reactions (*21*, *33*). This would lead to a

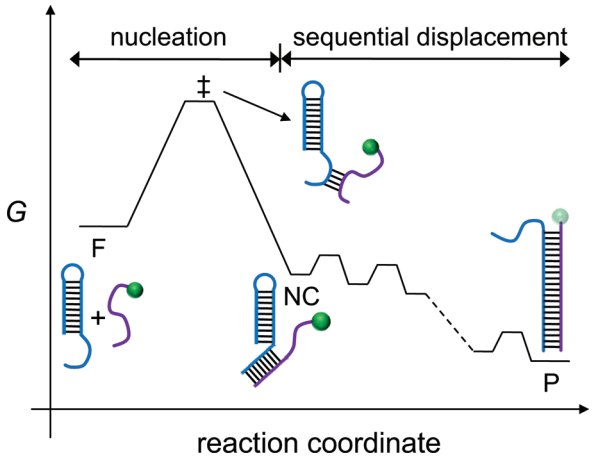


FIGURE 8: Proposed free energy profile for the hybridization reaction between R and G. NC represents a nucleation complex. The green sphere represents the dye that becomes faint in the final hybridization product.

shift of a rate-determining step from nucleation to strand displacement, which is contradictory to the high ϕ value. Therefore, our result is in accordance with the sequential displacement pathway. Zhang and Winfree recently showed that a three-step model considering nucleation, branch migration, and displacement is successful in predicting the kinetics of toehold exchange processes, supportive of the sequential displacement pathway (*36*). We propose that hybridization between structured nucleic acids proceeds through a transition state which retains all preexisting secondary structures, followed by sequential displacement of preexisting base pairings involving successive small energy barriers.

CONCLUSIONS

In this study, ϕ-value analysis has been applied to nucleic acid hybridization for the first time. We found that the preexisting secondary structure remains intact in the transition state, clearly providing compelling support for the previous observation that nucleation is a rate-determining step for hybridization of structured nucleic acids. Formation of the nucleation complex is followed by a zipping up process that involves disruption of a self-structure adjacent to the nucleation site. The proposed transition-state structure may shed light on the understanding of hybridization processes involved in small RNA-mediated gene silencing (*2*−*5*), which is essential to rational selection of a target mRNA segment for siRNA design. Currently, software packages for selecting efficient siRNA for a given mRNA are available (*14*, *37*−*39*). We expect that our proposed kinetic model can be implemented in the siRNA design tools to improve the rate of success of highly efficient siRNAs. To the best of our knowledge, this is the first direct demonstration of a transition-state determination for nucleic acid hybridization.

REFERENCES

1. Varshavsky, A. (2006) Discovering the RNA double helix and hybridization. *Cell 127*, 1295–1297.
2. Zamore, P. D., and Haley, B. (2005) Ribo-gnome: The big world of small RNAs. *Science 309*, 1519–1524.
3. Meister, G., and Tuschl, T. (2004) Mechanisms of gene silencing by double-stranded RNA. *Nature 431*, 343–349.
4. Ellington, A. D. (2007) What's so great about RNA? *ACS Chem. Biol. 2*, 445–448.
5. Ghildiyal, M., and Zamore, P. D. (2009) Small silencing RNAs: An expanding universe. *Nat. Rev. Genet. 10*, 94–108.
6. Bartel, D. P. (2009) MicroRNAs: Target recognition and regulatory functions. *Cell 136*, 215–233.
7. Brodersen, P., and Voinnet, O. (2009) Revisiting the principles of microRNA target recognition and mode of action. *Nat. Rev. Cell Biol. 10*, 141–148.
8. Filipowicz, W., Bhattacharyya, S. N., and Sonenberg, N. (2008) Mechanisms of post-transcriptional regulation by microRNAs: Are the answers in sight? *Nat. Rev. Genet. 9*, 102–114.
9. Luo, K. Q., and Chang, D. C. (2004) The gene-silencing efficiency of siRNA is strongly dependent on the local structure of mRNA at the targeted region. *Biochem. Biophys. Res. Commun. 318*, 303–310.
10. Schubert, S., Grünweller, A., Erdmann, V. A., and Kurreck, J. (2005) Local RNA target structure influences siRNA efficacy: Systematic analysis of intentionally designed binding regions. *J. Mol. Biol. 348*, 883–893.
11. Gredell, J. A., Berger, A. K., and Walton, S. P. (2008) Impact of target mRNA structure on siRNA silencing efficiency: A large-scale study. *Biotechnol. Bioeng. 100*, 744–755.
12. Köberle, C., Kaufmann, S. H. E., and Patzel, V. (2006) Selecting effective siRNAs based on guide RNA structure. *Nat. Protoc. 1*, 1832–1839.
13. Kertesz, M., Iovino, N., Unnerstall, U., Gaul, U., and Segal, E. (2007) The role of site accessibility in microRNA target recognition. *Nat. Genet. 39*, 1278–1284.
14. Tafer, H., Ameres, S. L., Obernosterer, G., Gebeshuber, C. A., Schroeder, R., Martinez, J., and Hofacker, I. L. (2008) The impact of target site accessibility on the design of effective siRNAs. *Nat. Biotechnol. 26*, 578–583.
15. Rudnick, S. I., Swaminathan, J., Sumaroka, M., Liebhaber, S., and Gewirtz, A. M. (2008) Effects of local mRNA structure on posttranscriptional gene silencing. *Proc. Natl. Acad. Sci. U.S.A. 105*, 13787–13792.
16. Chen, X., Zhou, Y., Qu, P., and Xin, S. Z. (2008) Base-by-base dynamics in DNA hybridization probed by fluorescence correlation spectroscopy. *J. Am. Chem. Soc. 130*, 16947–16952.
17. Gao, Y., Wolf, L. K., and Georgiadis, R. M. (2006) Secondary structure effects on DNA hybridization kinetics: A solution versus surface comparison. *Nucleic Acids Res. 34*, 3370–3377.
18. Wang, J.-Y., and Drlica, K. (2003) Modeling hybridization kinetics. *Math. Biosci. 183*, 37–47.
19. Erickson, D., Li, D., and Krull, U. J. (2003) Modeling of DNA hybridization kinetics for spatially resolved biochips. *Anal. Biochem. 317*, 186–200.
20. Dobson, C. M. (2003) Protein folding and misfolding. *Nature 426*, 884–890.
21. Bloomfield, V. A., Crothers, D. M., and Tinoco, I. (2000) Nucleic Acids: Structure, Properties and Functions, Chapter 8, University Science Books, Sausalito, CA.
22. Chen, C., Wang, W., Wang, Z., Wei, F., and Zhao, X. S. (2007) Influence of secondary structure on kinetics and reaction mechanism of DNA hybridization. *Nucleic Acids Res. 35*, 2875–2884.
23. Jo, J.-J., Kim, M.-J., Son, J.-T., Kim, J., and Shin, J.-S. (2009) Single-fluorophore monitoring of DNA hybridization for investigating the effect of secondary structure on the nucleation step. *Biochem. Biophys. Res. Commun. 385*, 88–93.
24. Fersht, A. (1998) Structure and mechanism in protein science: A guide to enzyme catalysis and protein folding, Chapter 18, Freeman, New York.
25. Fersht, A. R., and Daggett, V. (2002) Protein folding and unfolding at atomic resolution. *Cell 108*, 573–582.
26. Fersht, A. R., and Sato, S. (2004) Φ-Value analysis and the nature of protein-folding transition states. *Proc. Natl. Acad. Sci. U.S.A. 101*, 7976–7981.
27. Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res. 31*, 3406–3415.
28. SantaLucia, J., Jr., and Hicks, D. (2004) The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct. 33*, 415–440.
29. Torimura, M., Kurata, S., Yamada, K., Yokomaku, T., Kamagata, Y., Kanagawa, T., and Kurane, R. (2001) Fluorescence-quenching phenomenon by photoinduced electron transfer between a fluorescent dye and a nucleotide base. *Anal. Sci. 17*, 155–160.
30. Nazarenko, I., Pires, R., Lowe, B., Obaidy, M., and Rashtchian, A. (2002) Effect of primary and secondary structure of oligodeoxyribonucleotides on the fluorescent properties of conjugated dyes. *Nucleic Acids Res. 30*, 2089–2095.
31. Knemeyer, J.-P., Marme, N., and Sauer, M. (2000) Probes for detection of specific DNA sequences at the single-molecule level. *Anal. Chem. 72*, 3717–3724.

32. Noble, J. E., Wang, L., Cole, K. D., and Gaigalas, A. K. (2005) The effect of overhanging nucleotides on fluorescence properties of hybridising oligonucleotides labelled with Alexa-488 and FAM fluorophores. *Biophys. Chem. 113*, 255–263.
33. Braunlin, W. H., and Bloomfield, V. A. (1991) ^{1}H NMR study of the base-pairing reactions of d(GGAATTCC): Salt effects on the equilibria and kinetics of strand association. *Biochemistry 30*, 754–758.
34. Sambriski, E. J., Schwartz, D. C., and de Pablo, J. J. (2009) Uncovering pathways in DNA oligonucleotide hybridization via transition state analysis. *Proc. Natl. Acad. Sci. U.S.A. 106*, 18125–18130.
35. Reynaldo, L. P., Vologodskii, A. V., Neri, B. P., and Lyamichev, V. I. (2000) The kinetics of oligonucleotide replacements. *J. Mol. Biol. 297*, 511–520.
36. Zhang, D. Y., and Winfree, E. (2009) Control of DNA strand displacement kinetics using toehold exchange. *J. Am. Chem. Soc. 131*, 17303–17314.
37. Lu, Z. J., and Mathews, D. H. (2008) Efficient siRNA selection using hybridization thermodynamics. *Nucleic Acids Res. 36*, 640–647.
38. Owczarzy, R., Tataurov, A. V., Wu, Y., Manthey, J. A., McQuisten, K. A., Almabrazi, H. G., Pedersen, K. F., Lin, Y., Garretson, J., McEntaggart, N. O., Sailor, C. A., Dawson, R. B., and Peek, A. S. (2008) IDT SciTools: A suite for analysis and design of nucleic acid oligomers. *Nucleic Acids Res. 36*, W163–W169.
39. Lu, Z. J., and Mathews, D. H. (2008) OligoWalk: An online siRNA design tool utilizing hybridization thermodynamics. *Nucleic Acids Res. 36*, W104–W108.