Folding and Unfolding Kinetics of DNA Hairpins in Flowing Solution by Multi-parameter Fluorescence Correlation Spectroscopy

Jaemyeong Jung and Alan Van Orden*
Department of Chemistry, Colorado State UniVersity, Fort Collins, Colorado 80523

J. Phys. Chem. B ***2005****, 109, 3648-3657*

Out lines I

- **Introduction**

  Fluorescence correlation spectroscopy (FCS)

  Cross (Auto)-correlation spectroscopy

  DNA hairpin structure

- **Dynamic equilibrium between the folded and unfolded conformations of ssDNA hairpin**

- **Conclusion**

Fluorescence correlation spectroscopy (FCS)

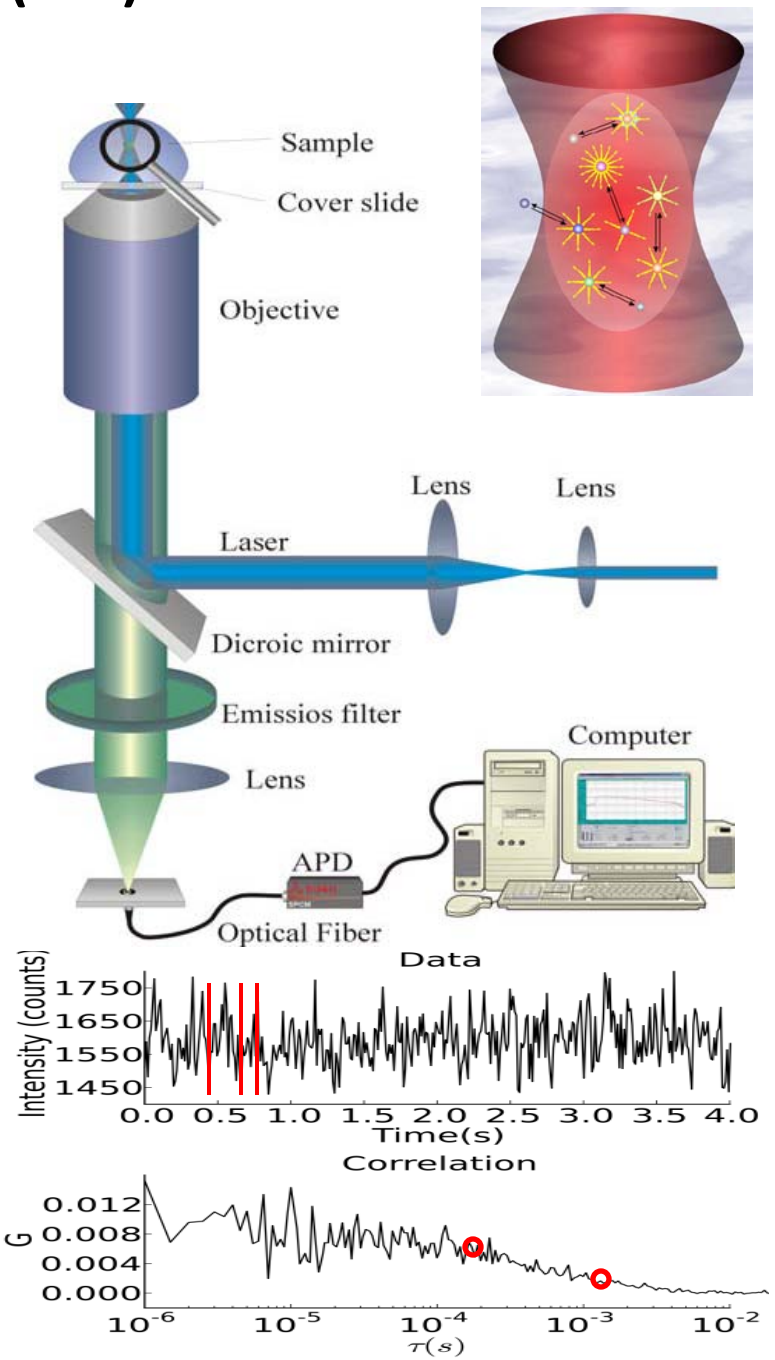
A technique for analyzing the spontaneous fluctuations in the fluorescence signal

diffusion
concentrations
kinetic chemical reaction
…

Correlation function indicates the statistical correlation between two different points in time.

$$G_A(\tau) = \int_{-\infty}^{\infty} f(t) f(t+\tau)\, dt$$

The decay of correlation function present the kinetic of the fluctuations.



Examples:

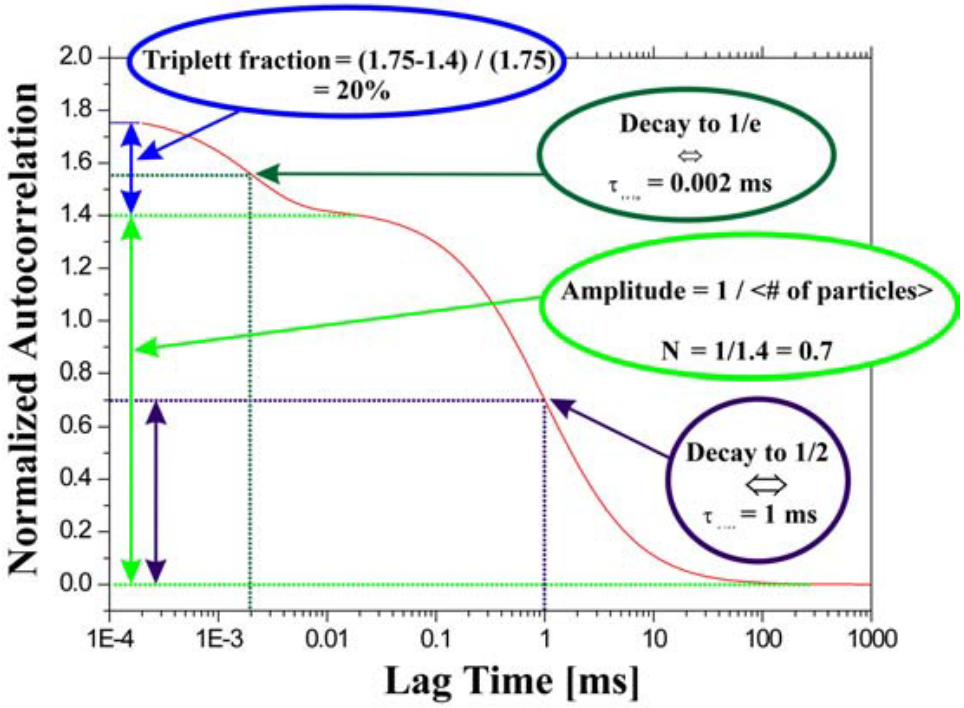
Freely diffusion and triplet fluctuation:

when the dynamic processes are in different time scale

$$G_{Total}(\tau) = G_{Diff}(\tau) G_{triplet}(\tau)$$

$$G_{Diff}(\tau) = \frac{1}{\langle N \rangle} \frac{1}{\left(1 + \frac{\tau}{\tau_D}\right)} \frac{1}{\sqrt{1 + \left(\frac{r_0}{z}\right)^2 \frac{\tau}{\tau_D}}}$$

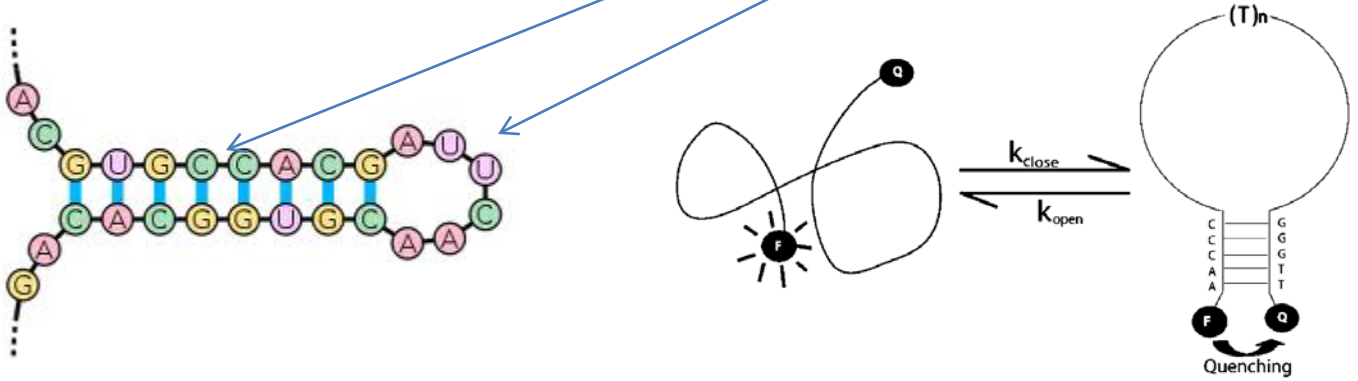
$$G_{triplet}(\tau) = 1 - T + T e^{-\frac{\tau}{\tau_{triplet}}}$$



In case that the dynamic processes are in the same time scale. FCS analysis can not distinguish

DNA hairpin structure

single-stranded DNA oligonucleotides that form a stem-loop DNA hairpin structure



Hairpin structures are not static: they fluctuate between different conformations. In a simplified description, all of the conformations can be divided into two main states: **the open state and the closed one**

Kinetics of conformational fluctuations in DNA hairpin-loops

(molecular beacons/fluorescence correlation spectroscopy/folding kinetics/polymer conformation/fluorescence energy transfer)

GRÉGOIRE BONNET, OLEG KRICHEVSKY, AND ALBERT LIBCHABER†

Center for Studies in Physics and Biology, The Rockefeller University, 1230 York Avenue, New York NY 10021

Communicated by Hans Frauenfelder, Los Alamos National Laboratory, Los Alamos, NM, May 15, 1998 (received for review April 8, 1998)

ABSTRACT The kinetics of DNA hairpin-loop fluctuations has been investigated by using a combination of fluorescence energy transfer and fluorescence correlation spectroscopy. We measure the chemical rates and the activation energies associated with the opening and the closing of the hairpin for different sizes and sequences of the loop and for various salt concentrations. The rate of unzipping of the hairpin stem is essentially independent of the characteristics of the loop, whereas the rate of closing varies greatly with the loop length and sequence. The closing rate scales with the loop length, with an exponent 2.6 ± 0.3. The closing rate is increased at higher salt concentrations. For hairpin closing, a loop of adenosine repeats leads to smaller rates and higher activation energies than a loop with thymine repeats.

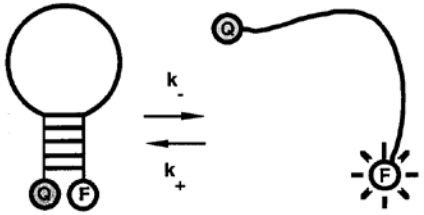


FIG. 1. Sketch of the DNA molecular beacon. The five bases at the two ends of the beacon are complementary to each other. The size of the loop and its content are varied in our experiments. The beacon flips between open and closed states with the characteristic rates k_-

Experiment setup

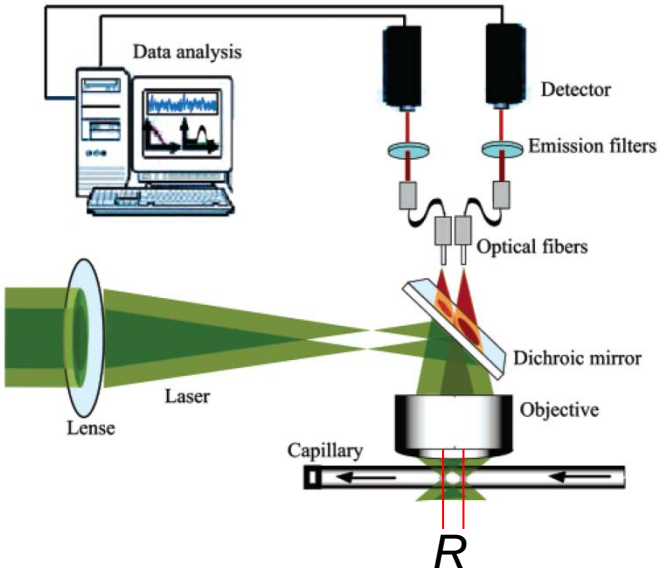
From autocorrelation analysis of a standard R6G solution

$$w = 0.306 \pm 0.092\,\mu m$$

$$k = w/z \approx 0.104 \pm 0.002$$

Combining with cross-autocorrelation analysis

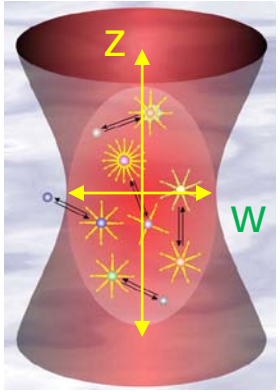
$$R = 2.26 \pm 0.07\,\mu m$$



5'-R6G-AACCC-$(T)_{21}$-GGGTT-dabcyl-3'

5'-R6G-AACCC-$(T)_{30}$-GGGTT-dabcyl-3'

poly(dT)40 oligonucleotide labeled at the 5' end with R6G for comparison purposes



Theory

This study investigate the fluctuations are cause by:

-Diffusion and unidirectional flow;
-Folding and unfolding dynamics of the hairpin;
-“Triplet blinking”

Cross-correlation function

$$G_C(\tau)-1=\frac{\gamma}{N}\underbrace{\frac{1}{\left(1+\frac{\tau}{\tau_d}\right)}\frac{1}{\sqrt{1+k^2\frac{\tau}{\tau_d}}}}_{G_{Diff}(\tau)}\underbrace{\left(1-T+Te^{-\tau/\tau_T}\right)}_{G_{triplet}(\tau)}\underbrace{\left[1+Be^{-\left(\tau/\tau_R\right)^{\beta}}\right]}_{G_R(\tau)}\underbrace{\exp\left(\frac{-r^2\left(1-\tau/\tau_F\right)}{1+\tau/\tau_d}\right)}_{Flowing}$$

N : number of molecules (one focal volume)

γ : geometric correction factor (compair to ideal Gaussian function)

τ_d : transit time through a focal volume

τ_F : transit time between two focal volumes R/V_x

r : the ratio R/w T : quantum yeild τ_T : time constant of triplet state

B : amplitude factor $B=K\frac{(1-Q)^2}{(1+QK)^2}$ K: equiblirium distribution F/UnF Q: relative fluorescence intensity

τ_R : relaxation time of folding and unfolding reaction

β: strectch parameter.

In the cross-correlation function, the hydrodynamic parameters relating to translational diffusion and uniform translation are effectively decoupled

Theory

$$G_C(\tau) - 1 = \frac{\gamma}{N} A \exp\left(\frac{-r^2\left(1 - \tau/\tau_F\right)}{1 + \tau/\tau_d}\right)$$

$$\text{Flowing}$$

One can determine τ_F i.e. τ_d from the position and the dispersion of the correlation peaks

While τ_F can be controlled by flowing rate and distance between two focal spot. It is reasonable to satify the condition in which

$$\tau_F \gg \tau_T \text{ and } \tau_F \gg \tau_R$$

$$\Rightarrow G_C(\tau) - 1 = \frac{\gamma}{N}(1 - T)\frac{1}{\left(1 + \frac{\tau}{\tau_d}\right)}\frac{1}{\sqrt{1 + k^2\frac{\tau}{\tau_d}}}\exp\left(\frac{-r^2\left(1 - \tau/\tau_F\right)}{1 + \tau/\tau_d}\right)$$

$$G_{Diff}(\tau) \qquad \text{Flowing}$$

All experiments will be done with $\tau_F \approx 1 \div 2ms$ which can be controlled by flowing rate and distance between two focal spot.

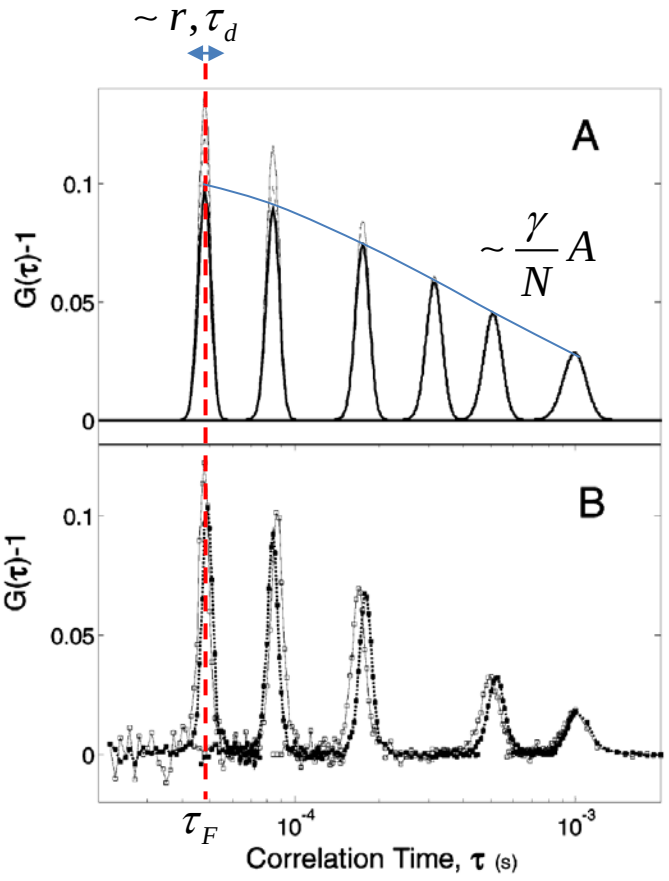


Figure 4. (A) A series of theoretical cross-correlation functions calculated using eq 2 with varying τ_F. The solid bold curves represent a DNA sample which does not exhibit conformational fluctuations. The dotted curves and the thin solid curves represent hairpin DNA samples with $\tau_R = 150\,\mu s$ and $\tau_R = 50\,\mu s$, respectively. (B) Experimental cross correlation functions measured for poly(dT)$_{40}$ (solid squares connected by dotted curves) and hp-T$_{30}$ (open squares connected by solid curves) samples.

Theory

Auto-correlation function

$$G_A(\tau)-1=\frac{\gamma}{N}(1-T)\underbrace{\frac{1}{\left(1+\frac{\tau}{\tau_d}\right)}\frac{1}{\sqrt{1+k^2\frac{\tau}{\tau_d}}}}_{G_{Diff}(\tau)}\underbrace{\left(1-T_{eq}e^{-\tau/\tau_T}\right)\left[1+Be^{-\left(\tau/\tau_R\right)^{\beta}}\right]}_{G_{triplet}(\tau)\quad G_R(\tau)}\underbrace{\exp\left[-\left(\frac{r\tau}{\tau_F}\right)^2\frac{1}{1+\tau/\tau_d}\right]}_{Flowing}$$

We have determined $\frac{\gamma}{N}(1-T);r;\tau_d;\tau_F$ from cross-correlation analysis

➡ One can define $T_{eq};\tau_T;B;\tau_R$ and β by fitting the autocorrelation data

Results

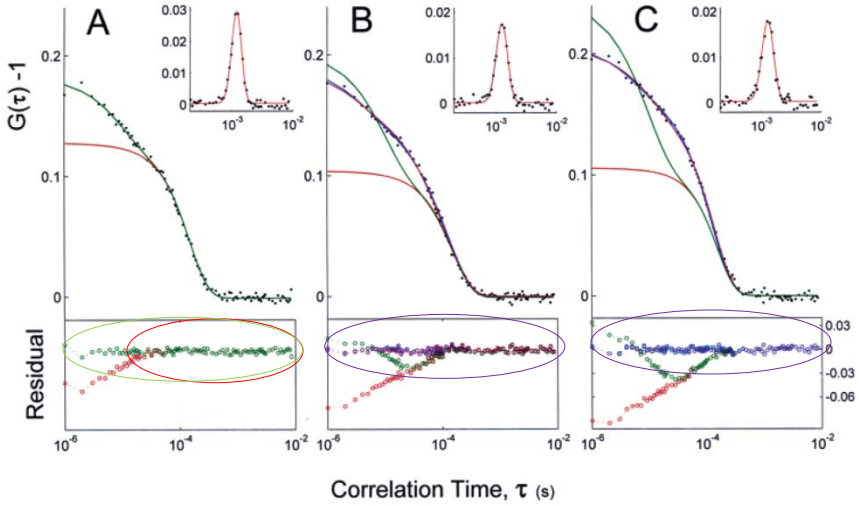


Figure 5. The experimental autocorrelation data (black dots) and fitting curves using our models for a poly(dT)$_{40}$ (A), a hp-T$_{21}$(B), and a hp-T$_{30}$(C) DNA sample. (Insets show corresponding cross correlation data and fitting curves.) Red colored curves are calculated autocorrelation functions obtained from the parameters(a, r, τ_d, and τ_F), determined from the cross-correlation analysis. Green colored curves fit the data by a 'triplet model,' which excludes the $g_R(\tau)$ contribution from the eq 5. The experimental data also were analyzed by fitting to eq 5, with a single-exponential decay model (blue), and nonexponential decay model (violet). The residuals from the least-squares fitting routines are shown at the bottom. The color scheme for the residual plots is the same as the one used for the autocorrelation functions. The fitted parameters are shown in Table 1.

TABLE 1: Parameters from FCS Analysis[a,b]

Cross Correlation Analysis

	poly(dT)$_{40}$	hp-T$_{21}$	hp-T$_{30}$
τ_F (ms)	1.311(12)	1.299(13)	1.268(13)
r	7.61(22)	7.90(31)	7.77(39)
a	0.1121(67)	0.1054(63)	0.1059(64)
τ_d (μs)	383(31)	266(16)	332(22)
standard error[c]	6.84×10^{-4}	6.92×10^{-4}	7.28×10^{-4}

Auto Correlation Analysis

		hp-T$_{21}$		hp-T$_{30}$	
	poly(dT)$_{40}$	exponential	nonexponential	exponential	nonexponential
τ_T (μs)	9.4(1.1)	3.95(69)	2.92(98)	4.94(78)	3.62(91)
T_{eq}	0.587(41)	0.293(62)	0.306(92)	0.346(46)	0.329(96)
B		0.603(18)	0.689(78)	0.754(30)	0.82(12)
τ_R (μs)		51.3(3.1)	45.5(8.7)	102.7(9.2)	121(19)
β			0.72(14)		0.67(12)
$\langle\tau_R\rangle$ (μs)			56(17)		159(43)
standard error[c]	2.54×10^{-3}	2.55×10^{-3}	2.51×10^{-3}	2.57×10^{-3}	2.44×10^{-3}

Equilibrium Measurement

		hp-T$_{21}$		hp-T$_{30}$	
	poly(dT)$_{40}$	exponential	nonexponential	exponential	nonexponential
K_{eq}		28.4(2.8)		22.5(2.0)	
Q		0.156(22)		0.153(26)	
τ_{op} (ms)		1.49(17)	1.64(52)	2.08(26)	3.73(97)
τ_{cl} (μs)		53.1(3.2)	58(16)	108.0(9.6)	166(45)

Folding and unfolding time constant

The obtained diffusion coefficients of poly (dT)40 DNA at 6.12x10^{-7} cm^2/s can be compared with work of Nkobo (2001) which is ~7.1x10^{-7} cm^2/s.

This set parameters is larger than expected by a factor of 2. Which is claimed that the R6G label is interacting with the ssDNA in a way that somehow stabilizes the triplet state of fluorophore.

For size dependence, Nkobo proposed that D scales as M$^{-0.68}$ i.e D_{T21}/D_{T30} is predicted to be 1.19 while the ratio of 1.25 was found from cross-correlation analysis.

$$G_C(\tau) - 1 = a \frac{1}{\left(1+\frac{\tau}{\tau_d}\right)} \frac{1}{\sqrt{1+k^2\frac{\tau}{\tau_d}}} \exp\left[-\left(\frac{r\tau}{\tau_F}\right)^2 \frac{1}{1+\tau/\tau_d}\right]$$

$$G_A(\tau) - 1 = a \frac{1}{\left(1+\frac{\tau}{\tau_d}\right)} \frac{1}{\sqrt{1+k^2\frac{\tau}{\tau_d}}} \left(1 - T_{eq} e^{-\tau/\tau_T}\right) \exp\left[-\left(\frac{r\tau}{\tau_F}\right)^2 \frac{1}{1+\tau/\tau_d}\right]$$

adjusting set of parameters: $T_{eq}, \tau_T, B, \tau_R$ and β

$$B = K \frac{(1-Q)^2}{(1+QK)^2} \rightarrow \tau_{op} = \langle\tau_R\rangle(K+1)$$

$$\rightarrow \tau_{cl} = \frac{\langle\tau_R\rangle(K+1)}{K}$$

Melting curve

Discussion

For a given stem sequence, the time constants for hairpin formation increase as a function of the loop size.

	hp-T_{21}		hp-T_{30}	
	exponential	nonexponential	exponential	nonexponential
τ_R (μs)	51.3(3.1)	45.5(8.7)	102.7(9.2)	121(19)

The folding time depends on the sequence of the hairpin loop. The mean relaxation time of hp-T30 is 159us while hp-A30 is 450us agreeing with results of Klenerman (2000)

DNA hairpin folding exhibits non-Arrhenius kinetice. A stretched exponential model is needed to characterize the relaxation process.

$0 \le \beta \le 1$ $\beta = 0$ Corresponding to a completely random folding mechanism

$\beta \approx 1$ Corresponding to two state, single barrier folding

$\beta_{polyA} \approx 0.5$ Was introduced by Klenerman and explained by Ansari that because of base-stacking of polyadenine chain that incease the roughness of the folding free energy surface.

$\beta_{polyT} \approx 0.7$ Was found in this experiment proves that theses intramolecular interaction are largely absent for polythymine chain

Folding and unfolding time constant are in qulitative agreement with conventional FCS analysis results doned by Libchaber group (2000) which used single exponential model.

$\tau_{op} \sim L^{\alpha^{op}}$ and $\tau_{cl} \sim L^{\alpha^{cl}}$

Jung		Libchaber		Ansari	
α^{op}	α^{cl}	α^{op}	α^{cl}	α^{op}	α^{cl}
2.30	2.96		2.6		2.2

Conclusion I

It is possible to extract relevant information about the hydrodynamic and the reaction rate properties of DNA hairpin molecules by analyzing fluorescence fluctuations observed at two spatially offset detection volumes.

Cross-correlation analysis of the two detection volumes revealed independently information about the hydrodynamic properties

Autocorrelation functions could then be constrained to allow determination of the intramolecular dynamics properties.

This research present evidence for heterogeneity in folding mechanism of large polythymine containing hairpin loops coming from the observation of nonexponential relaxation kinetics

The correlation analysis aonle is not sufficient to determine the equilibrium distribution of folded and unfolded hairpins. The photon counting histogram can be used to resolve these parameters

JACS
ARTICLES
Published on Web 01/05/2006

A Three-State Mechanism for DNA Hairpin Folding Characterized by Multiparameter Fluorescence Fluctuation Spectroscopy

Jaemyeong Jung and Alan Van Orden*

Contribution from the Department of Chemistry, Colorado State University, Fort Collins, Colorado 80523

Received September 2, 2005; E-mail: vanorden@lamar.colostate.edu

Trying with two state model

$$G_A(\tau)-1=\frac{1}{\bar{N}_{CC}}\frac{1}{\left(1+\frac{\tau}{\tau_d}\right)}\frac{1}{\sqrt{1+k^2\frac{\tau}{\tau_d}}}\left(1-T_{eq}e^{-\tau/\tau_T}\right)\left[1+Be^{-\left(\tau/\tau_R\right)^{\beta=1}}\right]\exp\left[-\left(\frac{r\tau}{\tau_F}\right)^2\frac{1}{1+\tau/\tau_d}\right]$$

$\bar{N}_{CC}=\bar{N}_1+\bar{N}_2$ Is the average number of hp DNA in both open and close conformation

FCS analysis

cross-correlation		autocorrelation	
r	11.8(06)	T_{eq}	0.198(24)
$\bar{N}_{cc}$	4.26(05)	B_{eq}	0.987(97)
τ_F (ms)	1.47(15)	τ_T (μs)	4.12(71)
τ_D (μs)	214(13)	τ_R (μs)	62.7(5.1)
		$[G(0)-1]^{-1}$	1.65(01)

But they want to go further

K, B, τ_{op} and τ_{cl}

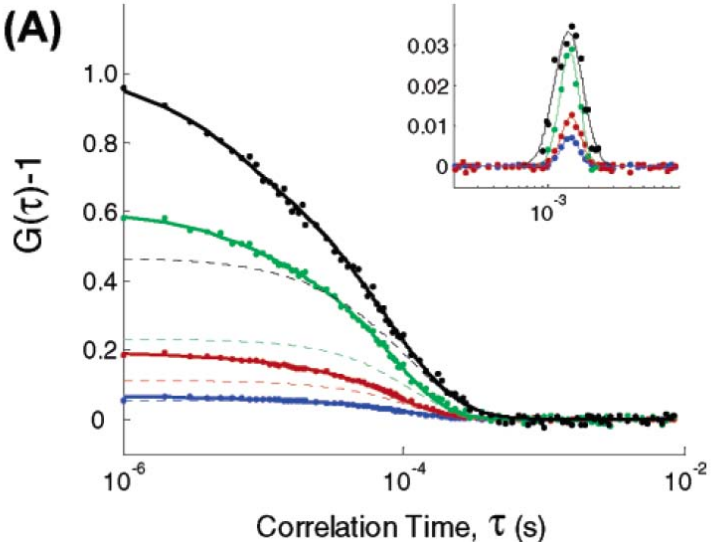


Figure 2. Experiment data (dots) and corresponding fitting curves (solid lines) from simultaneous FCS and PCH measurements vs NaCl concentration (blue, 0; green, 25; red, 100; and black, 500 mM NaCl). (A) The observed autocorrelation functions, with the cross-correlation functions shown in the inset. The dashed curves are the predicted autocorrelation functions assuming pure diffusion and flow. (B) The photon counting histograms obtained from the same samples using 9 μs sampling intervals. The DNA concentration was held constant at ~25 nM for all the samples.

The equilibrium distribution of folded and unfolded molecules

In two state regime the melting curves can be analyzed to determine the equilibrium constant $K_{melt} = \bar{N}_2 / \bar{N}_1$

Following by the equation

$$K_{melt}(T) = \frac{I(85°C) - I(T)}{I(T) - I(5°C)}$$

It was expected that NaCl stabilizes the folded form of the hairpin. Hence, K_{melt} rises with added NaCl

The goal is to simply characterize this behavior directly from FCS experiment. Such idea comes from the relation

$$B = K\frac{(1-Q)^2}{(1+QK)^2} = \frac{\bar{N}_1\bar{N}_2(\varepsilon_1 - \varepsilon_2)^2}{\bar{N}_1\varepsilon_1 + \bar{N}_2\varepsilon_2}$$

$\varepsilon_1, \varepsilon_2$ are the specific brightnesses of hpDNA in two state

$Q = \varepsilon_2 / \varepsilon_1$

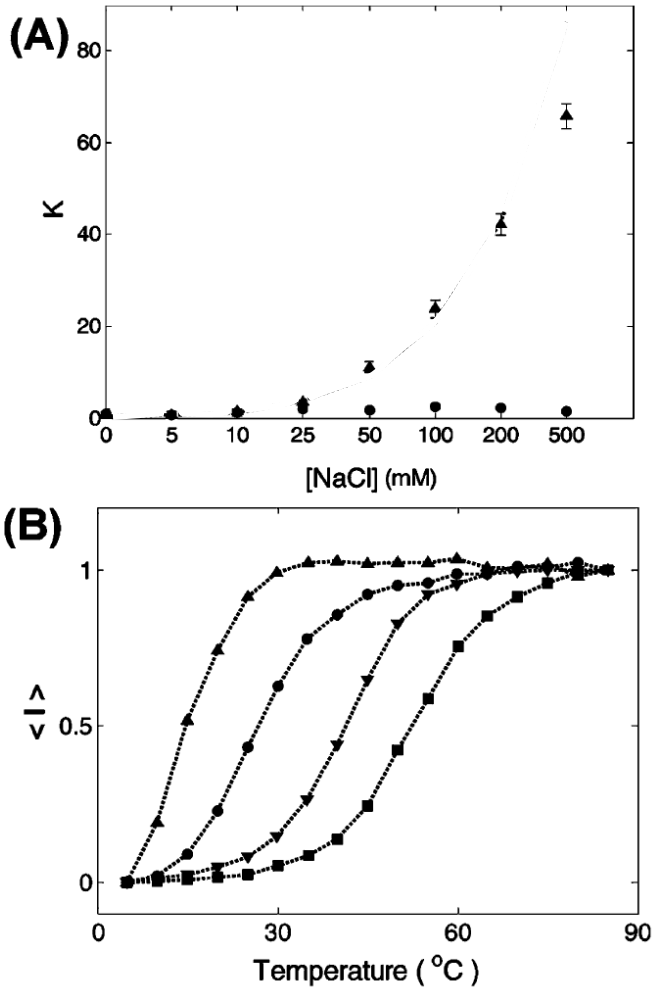
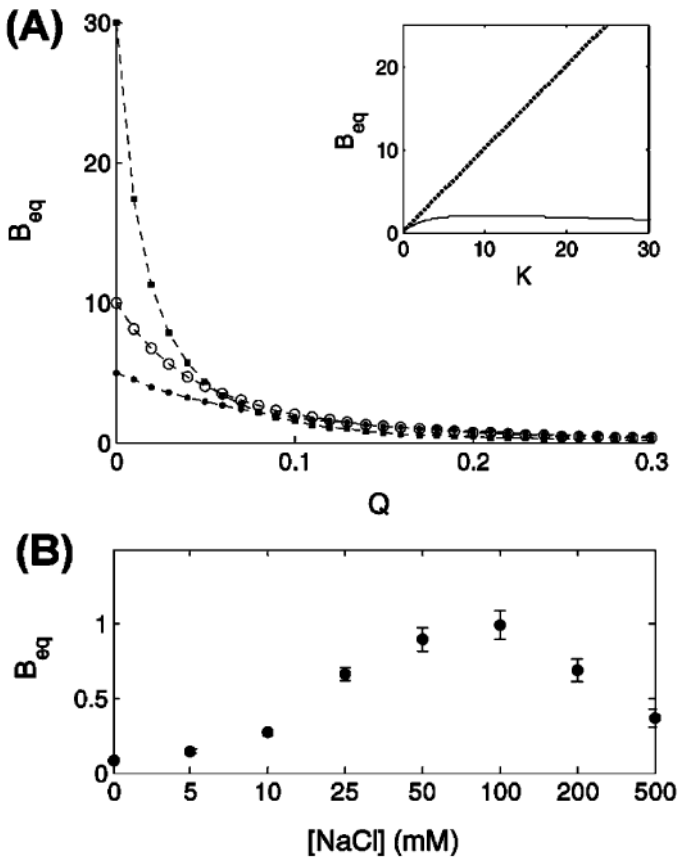


Figure 3. (A) Equilibrium constants of DNA hairpin samples vs NaCl concentration and (B) corresponding melting profiles [data sets with NaCl concentrations of 0 (▲), 25 (●), 100 (▼), and 500 mM (■) are shown]. In panel A, K_{melt} (▲) represents the equilibrium constants evaluated from the melting curves according to eq 12. K_{FFS} (●) represents the equilibrium constants determined from our FCS and PCH analysis. The dotted line in panel A is $K_{melt,3S}$ calculated according to eq 25.

The equilibrium distribution of folded and unfolded molecules

$$B = K\frac{(1-Q)^2}{(1+QK)^2} = \frac{\bar{N}_1\bar{N}_2(\varepsilon_1-\varepsilon_2)^2}{\bar{N}_1\varepsilon_1+\bar{N}_2\varepsilon_2}$$



In case of B<1; Q~0.1 it is not enough of using FCS data alone to interpret the equilibrium distribution

Combining FCS with PHC analysis

A photon counting histogram analysis investigates the amplitude of the fluctuations

The number of detected photons from a constant intensity light source is governed by Poisson statistics

$$p\left(k,\langle k\rangle\right)=\frac{\left(\langle k\rangle\right)^{k}e^{-\eta_E E}}{k!}\equiv Poi\left(k,\langle k\rangle\right)$$

k is the number of detected photons

$\langle k\rangle=\frac{\varepsilon}{V_0}\int_{V_0}\overline{\mathrm{PSF}}\left(\vec{r}\right)d\vec{r}$ is the average number of detected photons

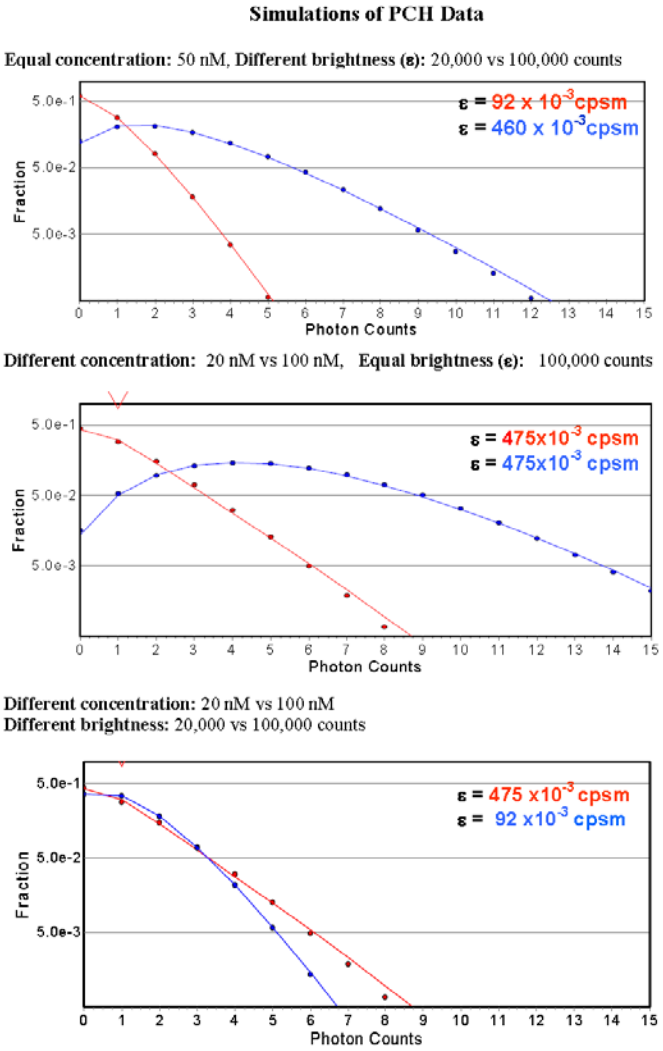
ε the molecular brightness

V_0 total sample volume

In two component system

$$P\left(k;\bar{N}_1,\varepsilon_1,\bar{N}_2,\varepsilon_2\right)=P\left(k;\bar{N}_1,\varepsilon_1\right)\otimes P\left(k;\bar{N}_2,\varepsilon_2\right)$$

By fitting the PCH data, one can obtain these parameters i.e. the equilibrium distribution



Failure of two state reaction mechanism

If the sub millisecond chemical relaxation process observed by FCS represents the complete DNA hairpin folding reaction then K_{melt} and $K_{FCS+PCH}$ should be identical

This assumption works at low concentration of NaCl in which the open form of hpDHA is favored

At higher concentration the deviation becomes quite dramatic

It is the first obvious clue showing that the DNA hairpin folding reaction may be more complicated than previously recognized

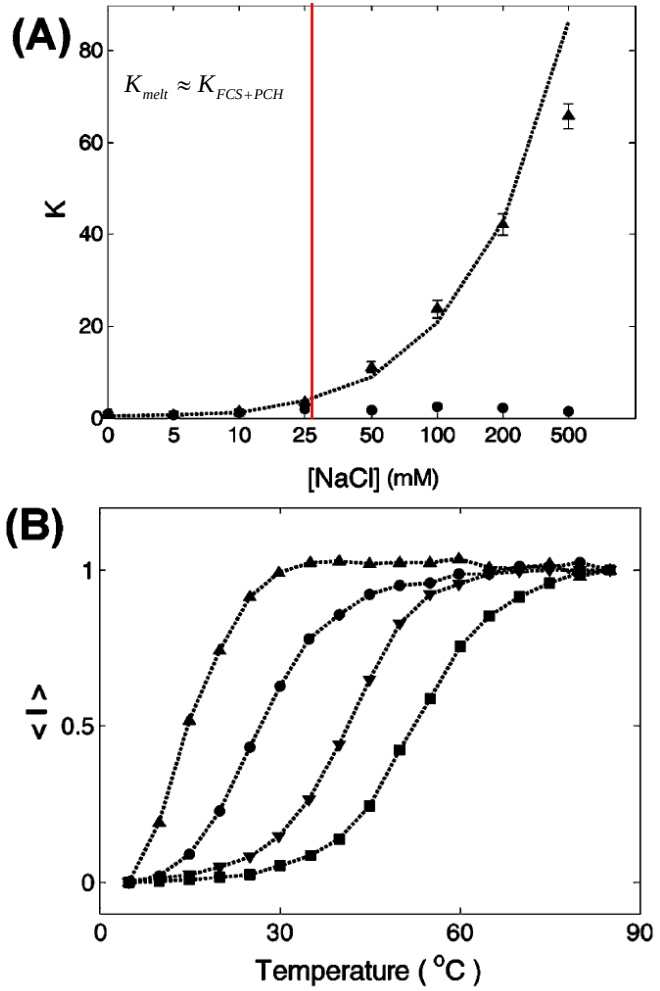


Figure 3. (A) Equilibrium constants of DNA hairpin samples vs NaCl concentration and (B) corresponding melting profiles [data sets with NaCl concentrations of 0 (▲), 25 (●), 100 (▼), and 500 mM (■) are shown]. In panel A, K_{melt} (▲) represents the equilibrium constants evaluated from the melting curves according to eq 12. K_{FFS} (●) represents the equilibrium constants determined from our FCS and PCH analysis. The dotted line in panel A is $K_{melt,3S}$ calculated according to eq 25.

Failure of two state reaction mechanism

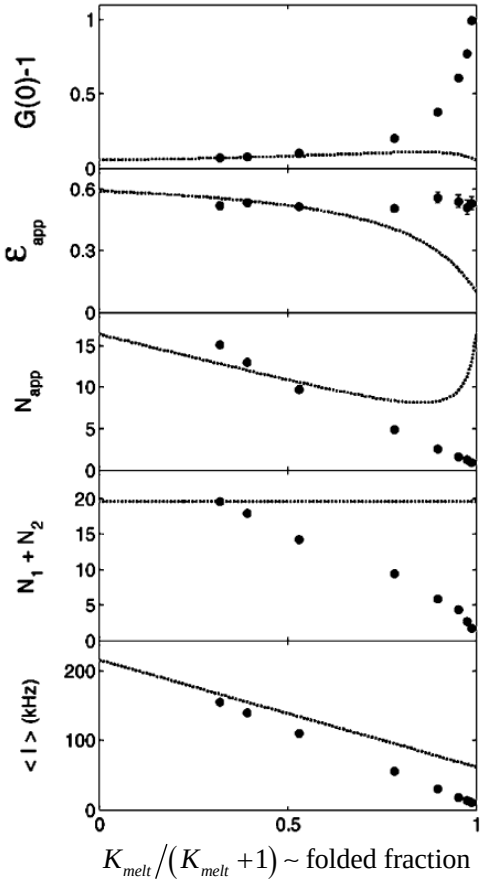
The disagreements are clear on the other parameters
Is there something wrong with the experiment?

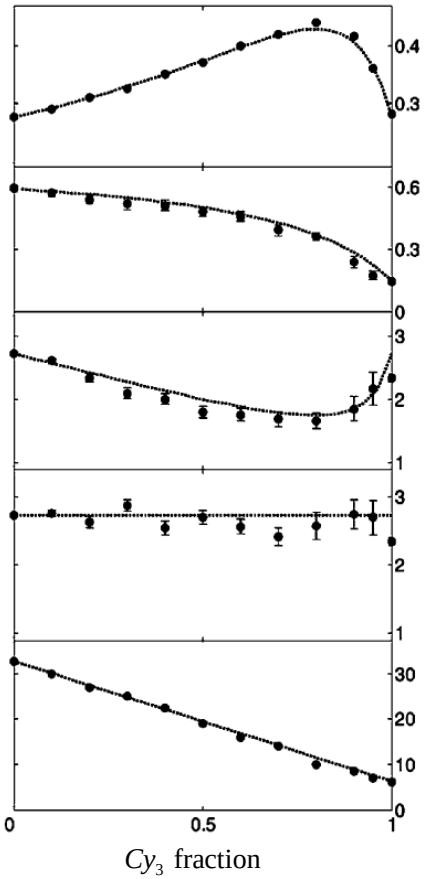
Verify with the mixture of two different brightness dyes (Cy3 +Rh6G)

Apparent brightness

Apparent number of molecule

Average photon count rate





The experiment is capable to characterizing two state behavior but hairpin folding mechanism

Three state reaction mechanism

The assumptions

-Some of the DNA being converted into a non fluorescent dark state

-Such dark state is being formed on a longer time scale (~1ms) than the FCS correlation time (sub millisecond). Once it is formed, it is stable

$$\bar{N}_1 \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \bar{N}_2 \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} \bar{N}_3 \qquad K_1 = \frac{\bar{N}_2}{\bar{N}_1}, K_2 = \frac{\bar{N}_3}{\bar{N}_2}$$

$\bar{N}_1$: refers to the unfolded DNA conformation for which the R6G fluorescence is unquenched

$\bar{N}_2$: refers to a reaction intermediate that is stable on the sub-millisecond time scale

$\bar{N}_3$: is fully folded DNA hairpin

$$G_{2S}(\tau) - 1 \Rightarrow G_{3S}(\tau) - 1 = \frac{1}{\bar{N}_{total}}\left(1 + B_1 e^{\lambda_1 \tau} + B_2 e^{\lambda_2 \tau}\right) g_D(\tau)\, g_T(\tau)\, g_F(\tau)$$

$$\bar{N}_{total} = \bar{N}_1 + \bar{N}_2 + \bar{N}_3$$

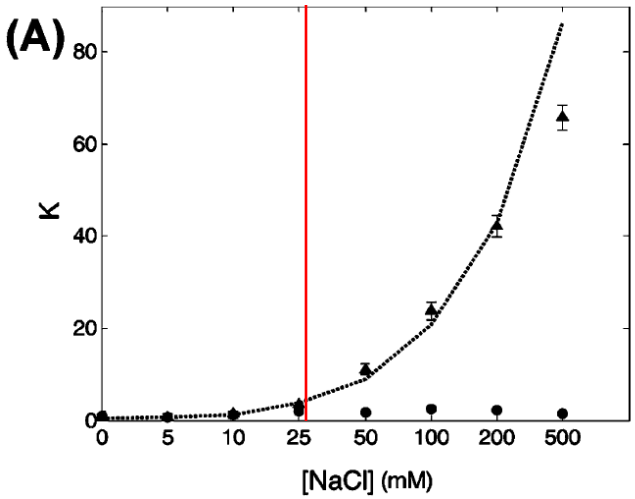
$B_1, B_2, \lambda_1, \lambda_2$: are functions of the rate and equilibrium constants

Qualitative discussions:

The equilibrium constant in 2 state regime which is not describe completely melting curve is now only the K_1 in 3 states model

At low concentration of NaCl, the $\bar{N}_1 \rightleftharpoons \bar{N}_2$ reaction dominates, so K and K_{melt} are essentially equivalent.

As NaCl is added, the $\bar{N}_2 \rightleftharpoons \bar{N}_3$ reaction becomes dominant, such that K_{melt} rises while K change more slowly because NaCl does not affect the $\bar{N}_1 \rightleftharpoons \bar{N}_2$ reaction as strongly as it does on complete folding reaction



The FCS and PCH measurements are only sensitive to $\bar{N}_1$ and $\bar{N}_2$. At low NaCl concentration, represent the dominant species so the PCH and FCS parameters behave as expected for two state reaction.

Quantitative discussions:

$$K_{melt,3S,i} = \frac{\langle I \rangle_{\max} - \langle I \rangle_i}{\langle I \rangle_i - \langle I \rangle_{\min}}$$

$\langle I \rangle_i$: is the average fluorescence count rate for a given NaCl concentration

$\langle I \rangle_{\max}$ and $\langle I \rangle_{\max}$: are the fluorescence count rate when all DNA are in the close form or the open form

Conclusion II:

Extra PCH analysis is simultaneously used to investigate the equilibrium constant by the same experiment which are in remarkably good agreement with what introduced by the melt curve.

A three state reaction mechanism has been proposed and proved by experiment data at very high concentration of NaCl in which, the folded DNA are favored. This regime consists of a rapid equilibrium between open and intermediate forms of the DNA and a fully folded form that is stable on the time scale of FCS experiment .

Thank you for your attention