

Biophysical Chemistry

Project:

Proving DNA Hexagon Formation with Fluorescence Correlation Spectroscopy

Johan Henriksson
Martin Niklas

Supervisors: Marcus Wilhelmsson, Michal Tokarz

Table of Contents

Aim.....	3
Theory.....	3
Nanostructures.....	3
Fluorescence Correlation Spectroscopy (FCS).....	4
Autocorrelation.....	5
Methodology.....	6
Results.....	8
Conclusions.....	8
Appendix.....	9
Code Listing for Assessing Confidence Intervals.....	9
Absorption Graphs.....	10
Correlation Plots.....	12

Aim

The aim of this mini-project was to build two different nano-sized structures of DNA by using two sets of DNA-oligomers: a circular (hexagonal) and a linear form.

To prove that these structures are different, fluorescence correlation spectroscopy (FCS) was used to monitor the different translational diffusion of these biomolecules.

This experiment was important to give evidence whether it is possible to build these circular nanoconstructs of DNA or not.

Theory

Much of the information about the Fluorescence Correlation Spectroscopy and the autocorrelation is taken from the paper "Fluorescence Correlation Spectroscopy, An Introduction to its Concepts and Applications"¹.

Nanostructures

DNA is increasingly considered as an exciting and useful building block for controllable nanotechnology. The specificity and reliability of hydrogen-bond formation between DNA bases provides an excellent basis for careful and deliberate design of sequences which can form nano-sized structures. In this project two sets of DNA-oligomers were used: one set that can hybridize with each other to form a hexagon (Figure 1) and another set that can only form a linear form (Figure 2). Both sets of oligomers are fluorescently labelled with Cy5 (Figure 3), which is one of the most commonly used DNA-fluorophores. The fluorescence of Cy5 is used to monitor the translational diffusion of the two DNA-nanoconstructs with the help of FCS.

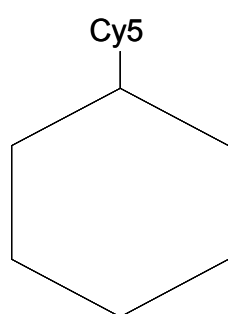


Figure 1: Hexagon

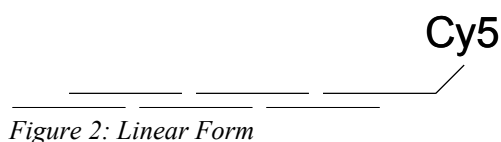


Figure 2: Linear Form

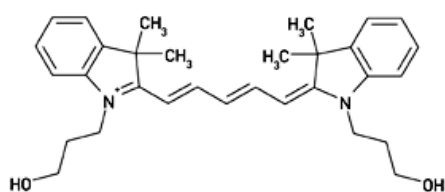


Figure 3: Fluorophore Cy5

¹"Fluorescence Correlation Spectroscopy, An Introduction to its Concepts and Applications", Petra Schwille and Elke Haustein, Experimental Biophysics Group, University of Göttingen

Figure 4, 5 and the table show the conformations of the two nano structures:

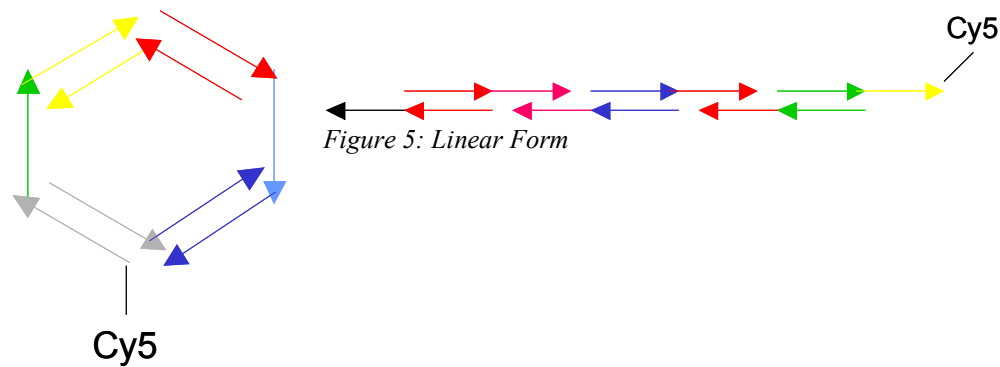


Figure 4: Hexagon

Code			
Hex6	6	5' C TGTA GAGC C TT G ATAC CATC G 3'	C'-F'
02161R	61	5' C TGTTG TGTC C TT G CTTG ATT C 3'	G'-H'
02164R	64	5' G GCTC TACA G L G ATGG AATA C L C CAAT CAAG C 3'	C-Z-H
02212R	Hex11-5	5'CY5 G AGCA CACA G L C GTTT CTCA C L C GATG GTAT C 3'	G-Y'-F
	65-5		
	6T	5' C TGTA GAGC C TT TT TTTT TTTT T 3'	
	1	5' C CATA CATA C TT C CACA GCAT C 3'	
	2	5' G GCTC TACA G TT G AGGA GGAT G 3'	
	3	5' G ATTA GCGT C TT C GATG GTAT C 3'	
	4	5' G TATG TATG G TT C ATCC TCCT C 3'	
	5	5' G ACGC TAAT C TT G ATGC TGTG G 3'	

Fluorescence Correlation Spectroscopy (FCS)

FCS is one of several standard tools used for high-resolution in spatial and temporal analysis. In contrast to many other techniques it is not trying to gain a very high emission intensity but rather spontaneous intensity fluctuations caused by the minute deviations of small systems from thermal equilibrium. These fluctuations of physical parameters are reflected by the fluorescence emission of the biomolecules. To achieve this a very low concentration (nanomolar) of the molecule of interest is required. Therefore, the basic concept of FCS is to make the number of observed molecules low enough so that each of them contributes substantially to the measured signal.

The fluctuations are quantified in their strength and duration by temporally autocorrelation which is a routine to measure the self-similarity of the recorded signal and to highlight time constants of the underlying process. In our case we

measured the lateral diffusion time of our DNA nanostructure in a specific detection volume which is very small. To make fluorescence possible fluorescent dyes, in our case Cy5, have to be attached to the molecule. During that time the particle spend in the focus chemical or photo physical reactions or conformational changes may alter the characteristics of the flourophore and give rise to additional fluctuations in the detected signal.

Figure 6 illustrates the setup of FCS. A laser beam is focused by several lenses and directed by a dicroric mirror into the objective and focused on the sample. As we used water immersion objective a drop of water had to be placed between the objective and the microscope slide. The detection volume of the specimen on which is focused is less than one femtolitre. The fluorescent light emitted by the sample is collected by the same objective and then directed to the dicroric mirror and several filters to a detector which is an avalanche diode or a photomultiplier. The filters are necessary to collect only the fluorescent light from the specimen and blocking additional light. Bandpass filters, adapted to the dye, are therefore used. Finally the recorded signal is transferred from the detector to the computer to carry out the autocorrelation and other routines.

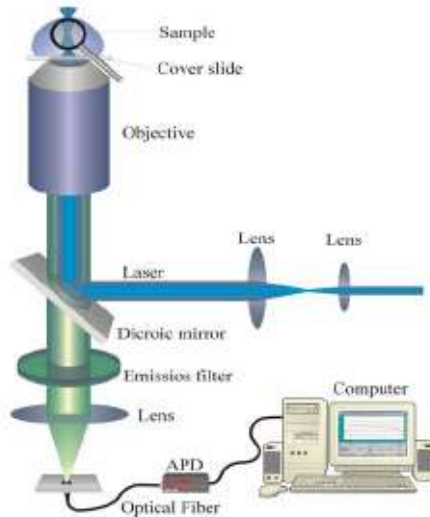


Figure 6: FCS-Setup

Autocorrelation

The non-normalized autocorrelation of the photometer output signal F is defined as:

$$(1) \quad G(\tau) = \langle \delta F(t) * \delta F(t+\tau) \rangle$$

If it is assumed that

- the concentration is constant
- fluorescence properties do not vary over time
- molecules have Brownian motion in all directions

then it is possible to derive theoretically the shape of this function:

$$(2) \quad G(\tau) = \frac{1}{\langle N \rangle \left[1 + \frac{\tau}{\tau_D} \right] \sqrt{1 + \left(\frac{r_0}{z_0} \right)^2 \frac{\tau}{\tau_D}}}$$

Earlier calibration of this equipment has shown that $\frac{r_0}{z_0} = \frac{1}{6}$.

The constant that we want to find is τ_D . This is done using non-linear least squares. The problem here is that confidence intervals are tricky to find

because of the non-linearity. We are not aware of a way to compute these exactly and hence we will only compute an estimate. To do this, G has to be linearised:

$$G(\tau)_{linear} = G(\tau_0, \tau_{D0}, \langle N \rangle_0) + G'_{\langle N \rangle}(\tau_0, \tau_{D0}, \langle N \rangle_0)(\langle N \rangle - \langle N \rangle_0) + G'_{\tau_{D0}}(\tau_0, \tau_{D0}, \langle N \rangle_0)(\tau_D - \tau_{D0})$$

It turns out that the dependence on $\langle N \rangle$ is minimal compared to dependence on τ_D and hence we neglect it. This makes the parameters follow a univariate distribution (much easier statistics). Formally, the model becomes:

$$\tau_D - \tau_{D0} = \frac{G_i - G(\tau_0, \tau_{D0}, \langle N \rangle_0)}{G'_{\tau_{D0}}(\tau_0, \tau_{D0}, \langle N \rangle_0)} + \xi_i$$

where ξ_i are Normally i.i.d errors. This is a linear regression model and the significance can be found in the normal way. This means that the parameter is distributed as

$$\frac{\tau_D - \tau_{D0}}{s} \sim t_{n-2}$$

with sample variance defined as usual. In the end, the purpose is to compare the different estimates for τ_D . When comparing two randomly distributed variables, one should actually use a different approach but it turns out that this brute approximation is enough to give the wanted conclusions.

There is one error made in this analysis however; the samples are highly correlated. Once again we know no math that can be used for a correlation this big. There are two fast ways out; either have this in mind when interpreting the results or assume a different effective degrees of freedom.

The diffusion constant is related to the effective radius through

$$D = \frac{kT}{5\pi\eta_v R_h}$$

Because the linear conformation is less rigid than the hexagonal structure, it should have less effective radius and hence move faster.

Methodology

First of all we had to build our nano structures consisting of different sets of DNA-oligomers and label them with the fluorescent dye Cy5.

The oligomers solved in a buffer were already given to us but we had to dilute them because one needs concentrations of around 10 nM in order to carry out measurements with FCS.

In the first step we determined the actual concentrations of our samples. This was done by using absorption spectroscopy. The concentration can be calculated by the formula: $c = A/l * \epsilon$

- c is the concentration
- l = 0,3 cm is the length of our sample holder

- A is the absorption
- ϵ is the extinction coefficient (it can be looked up in tables)

with the help of the absorption graphs (see appendix, Figure 9- 14).

It is important to record a baseline of the buffer and subtract it from all the measured values of the DNA-oligomers.

The following table shows the actual concentrations of our samples:

Sample	Excitation spectrum [nm]	Wavelength correlated to absorption peak [nm]	Absorption maximum	concentration [μm]
61	350 - 220	260	0,151	2,66
64	350 - 220	259	0,297	3,21
65-5	350 - 220	260	0,171	1,91
65-5 (dye)	800 - 200	648	0,110	1,47
Hex6	350 - 220	260	0,186	2,96

Most DNA-oligomer have their absorption maximum at about 260 nm but we also had to scan 65-5 in the range of 800 – 200 nm because the dye has its absorption maximum at about 648 nm.

In this process we made solutions of every DNA-oligomer of 1 μ M and before we did the FCS we diluted them again to achieve nM.

After that we mixed 380 μ l each of the oligomers 61, 64, 65-5 and Hex6 to form the hexagonal nano-structure and 380 μ l each of the oligomers 61, 64, 65-5 and **S4** to form the linear structure.

After the mixture both solutions were constantly heated and cooled down to hybridise all DNA-oligomers in order to form the structures.

Now, all three samples (65-5, hexagonal, linear) were diluted to a concentration of 10nM. This was to make sure no more than one molecule at a time was moving in the lighted area. The equipment was tested first to make sure reasonable data was sampled. Our test results are not shown here. The real run was sampled at 500kHz during 60 seconds. After the autocorrelations were obtained, the parameters were obtained using non-linear least squares. The fitting was done using the program "ISS VistaFCS for Leica" (internally, it uses a kind of steepest descent for optimizing). Evaluation of confidence intervals was done in Matlab (see appendix).

Results

The fitted values for τ_D are given in the table below for the different samples.

	65-5	Hexagonal	Linear
τ_D (μs)	229.84	293	265.538
99.5% interval (df=60)	+ -0.53	+ -0.93	+ -0.82
99.5% interval (df=5)	+ -0.67	+ -1.1832	+ -1.04

The graphs showing the fitted curves are attached in the appendix (Figure 15, 16).

Conclusions

Ignoring the correlation between the samples, statistics shows that the fitted values (τ_D for 65-5, linear and hexagonal form) are distinct even with 5 degrees of freedom. Since this is a very little, we consider this to be too large a difference to be of coincidence and hence conclude that the values are different. By assumption, the structures we formed by DNA-oligomers should be different. The linear conformation moves also faster which supports this conclusion.

Appendix

Code Listing for Assessing Confidence Intervals

```
%%%%%%%%% Linear
a=load('lin3.csv');
taud=265.381e-6;
meann=4.18837;

%%%%%%%%%% Hex
%a=load('hex3.csv');
%taud=293e-6;
%meann=4;

%%%%%%%%%% 65
%a=load('653.csv');
%taud=229.841e-6;
%meann=4.177198;

tau=a(:,1);
coval=a(:,2);

%figure(1)
%semilogx(tau,coval,'o')

r0z0=1/6;

covalest=gfun(tau, taud, meann, r0z0);

figure(2)
semilogx(tau,coval,'o', tau, gfun(tau, taud, meann, r0z0));

%Find G'
gnprim=[gfun(tau, taud, meann+0.01, r0z0) - gfun(tau, taud, meann, r0z0)]/0.01;
gtaudprim=[gfun(tau, taud+0.01, meann, r0z0) - gfun(tau, taud, meann, r0z0)]/0.01;

t995=2.660;
t99 =2.390;
t975=2.000;
tdf5=3.365;

%Normalize error
err=(coval-covalest)./(gtaudprim);

%Find interval
s=sqrt(var(err));
taud
inter=s*t995
inter2=s*tdf5

%Compare relative importance of differentials
plot(tau,gnprim,'-',tau,gtaudprim,'-')

function y=gfun(tau, taud, meann, r0z0)
y=1./((meann.*(1+tau./taud).*sqrt(1+r0z0.^2 * tau./taud )));
```

Absorption Graphs

The x-axis of the following six graphs shows the absorption and the y-axis shows the wavelength of the excitation light.

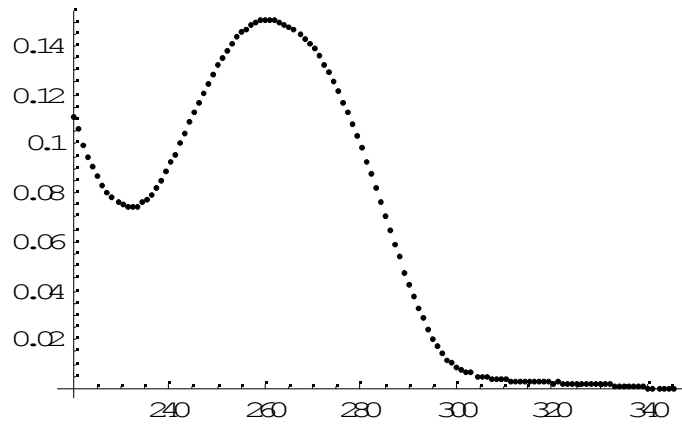


Figure 9: DNA 61

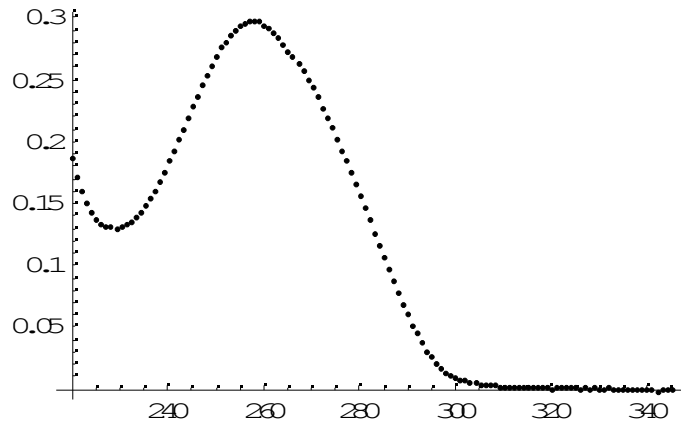


Figure 10: DNA 64

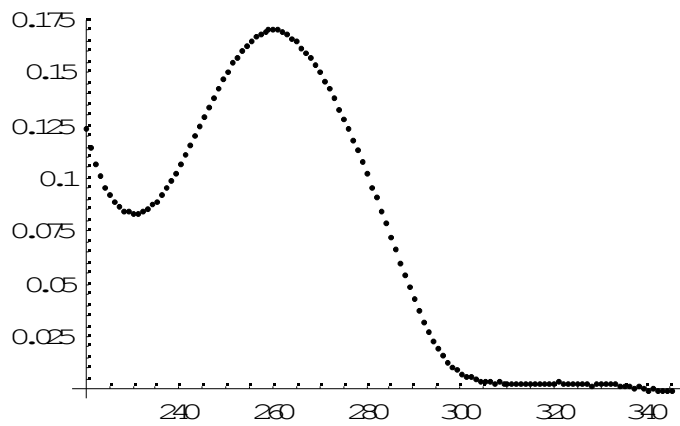


Figure 11: DNA 65-5

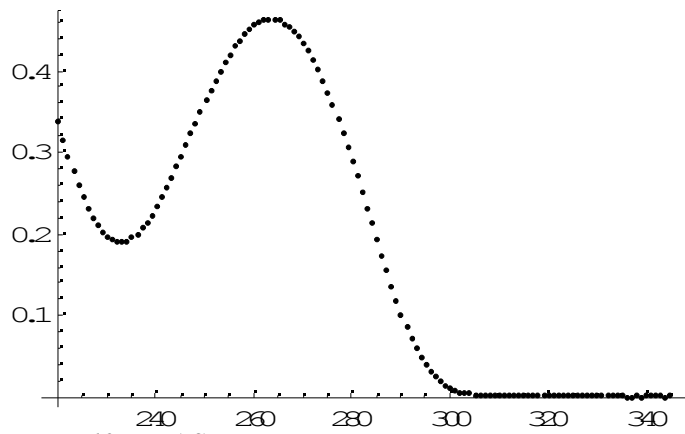


Figure 12: DNA S4

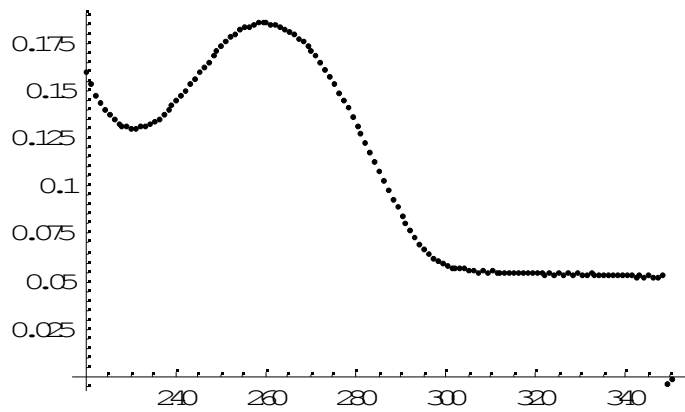


Figure 13: Hex 6

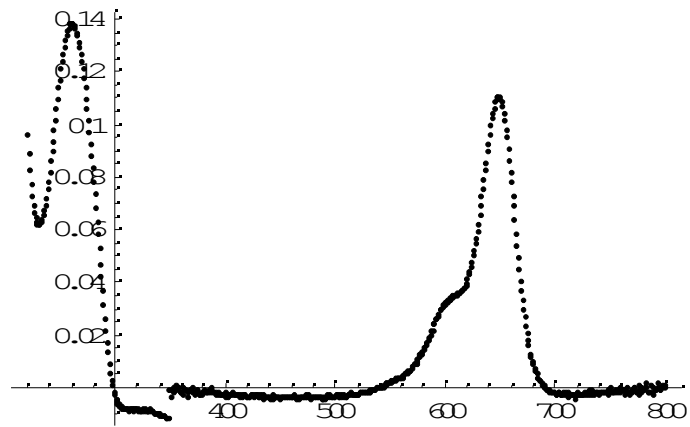


Figure 14: DNA 65-5 (dye)

Correlation Plots

The x-axis represents τ_D and the y-axis the correlation.

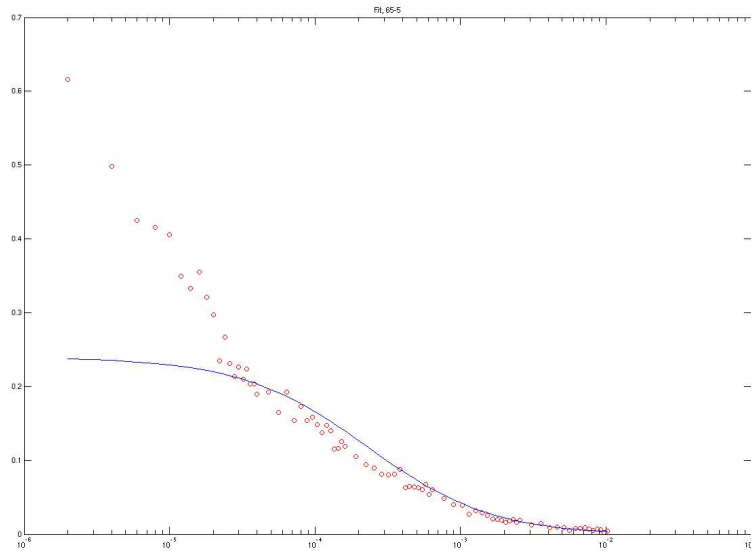


Figure 15: Fitted curve, DNA 65-5

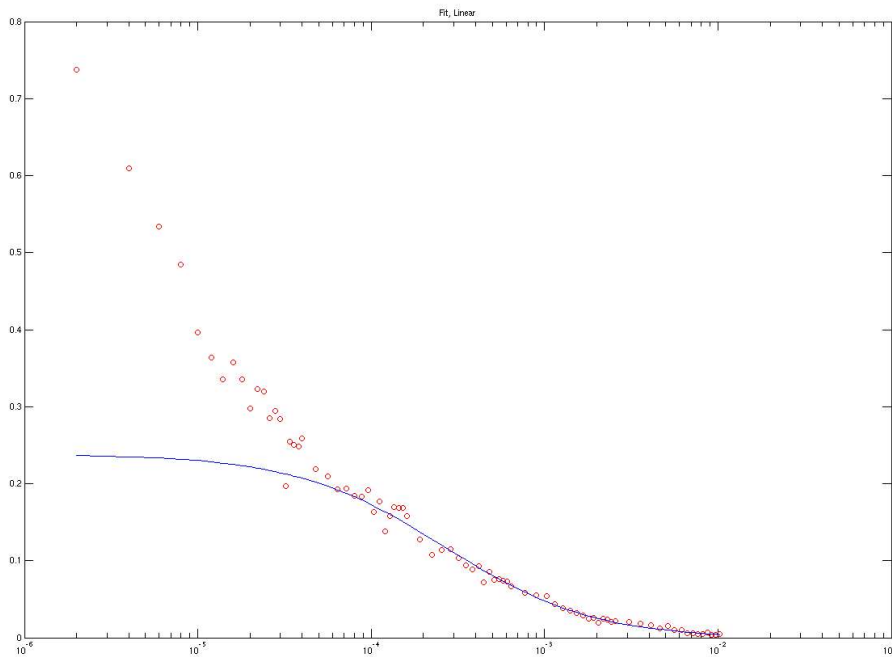


Figure 16: Fitted curve, Linear

In both plots the red dots are the sampled values (correlation function) and the blue line is the fitted curve.