Studying Structural and Dynamical Properties of DNA Polymers using Scanning Fluorescence Correlation Spectroscopy

Thesis submitted in partial fulfillment
of the requirements for the degree of
"DOCTOR OF PHILOSOPHY"

By

Manish Nepal

Submitted to the Senate of Ben-Gurion University
of the Negev

June 13, 2013

Beer-Sheva

Studying Structural and Dynamical Properties of DNA Polymers using Scanning Fluorescence Correlation Spectroscopy

Thesis submitted in partial fulfillment of the requirements for the degree of "DOCTOR OF PHILOSOPHY"

by

Manish Nepal

Submitted to the Senate of Ben-Gurion University of the Negev

Approved by the advisor

Approved by the Dean of the Kreitman School of Advanced Graduate Studies

June 13, 2013

Beer-Sheva

This work was carried out under the supervision of

Prof. Oleg Krichevsky, Physics Department

and

Prof. Dina Raveh, Life Sciences Department

in Interdisciplinary Studies

Research-Student's Affidavit when Submitting the Doctoral Thesis for Judgment

I, Manish Nepal, whose signature appears below, hereby declare that

(Please mark the appropriate statements):

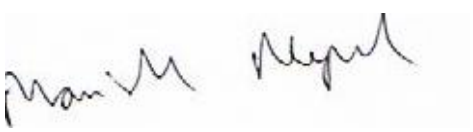
__X_ I have written this Thesis by myself, except for the help and guidance offered by my Thesis Advisors.

X_ The scientific materials included in this Thesis are products of my own research, culled from the period during which I was a research student.

_X__ This Thesis incorporates research materials produced in cooperation with others, excluding the technical help commonly received during experimental work. Therefore, I am attaching another affidavit stating the contributions made by myself and the other participants in this research, which has been approved by them and submitted with their approval.

Date: June 13th, 2013 Student's name: Manish Nepal

Signature:



Acknowledgements

First of all I would like to thank my supervisor Prof. Oleg Krichevsky who provided me the opportunity to work in his lab and for his support and guidance throughout this work.

I would also like to thank all the group members at the time in the lab for providing me support and help when needed and for all the interesting discussions I have had. I am also indebted to Tanya Gavrinyov for her procurement and logistical support.

Finally I would like to acknowledge the financial support from EC FP7 Marie Curie Fellowship program (Grant No. 215148).

Contents

Abstract

Being the carrier of genetic information, DNA is, perhaps, the most important polymer around. It is also a very interesting polymer from the Physics point of view due to the scale of its persistence length ~50 nm, that is much larger than the diameter of DNA double helix (2 nm) yet is typically much smaller than the DNA contour length (> 1μm). This defines DNA as a *semi-flexible* polymer with structural and dynamic properties distinct from those of *flexible* chains (a majority of synthetic polymers) and *stiff* polymers (such biofilaments as microtubules and actin). However, prior to our work there were no good experimental tools to measure DNA structure at the length scales above 50 nm.

In this thesis a new approach combining scanning fluorescence correlation spectroscopy (SFCS) and covalent DNA labelling is successfully implemented to study various structural and dynamical properties of DNA solutions. In particular, in dilute DNA solutions we confirm for the first time the theoretical prediction of an essentially ideal structure of DNA chains: because of low DNA flexibility the probability of collision between DNA segments is low for contour lengths up to ~ 60μm. And indeed the structure factors $S(q)$ obtained in our measurements follow the Debye expression for ideal polymers. The scaling of gyration radius R_g with DNA contour length L ($R_g \propto L^{0.52\pm0.02}$) is also consistent with the ideal worm-like chain behaviour.

With the use of specific labelling, our technique also allows us to assess DNA end-to-end distance distribution. It appears to follow Gaussian distribution in a further confirmation of the ideal nature of DNA coils.

In another application of specific labeling, we probe the structure of individual labeled DNA coils embedded in a dense mesh of other unlabeled DNA (*semi-dilute solutions*). For flexible polymers such conditions lead to shrinking of their coils with increasing concentration. However, our measurements on DNA reveal that their chain structure does not change up to the concentration of 600 $\frac{\mu g}{mL}$, about 30 times larger than the chain overlap concentration. The invariance of the individual chain structure despite the dense organisation of polymer matrix is a remarkable feature of semi-flexible polymers not previously explored experimentally. Although surprising, this experimental finding is consistent with the marginal

solution theory of semi-flexible polymers by Schaefer, Joanny and Pincus [Macromolecules 13, 1280 (1980)].

Furthermore, we develop a new method based on SFCS where scanning is performed at multiple speeds, in order to measure the segmental dynamics of DNA in semi-dilute solutions. The use of scanning FCS as compared to the standard, static FCS eliminates problems arising from photobleaching of slowly moving dye molecules. In a remarkable analogy to the invariance of DNA chain structure from dilute to semi-dilute solutions, we also find that DNA chain dynamics is only weakly affected by the dense mesh of other DNA.

Finally, we attempt to experimentally resolve one of the controversial topics in the last few decades; the effect of electrolyte strength on DNA persistence length. While overall our data are consistent with the classical Odijk-Skolnik-Fixman (OSF) theory of polyelectrolytes, we find significant deviations from OSF at very low electrolyte concentrations.

Keywords: Polymer solutions, DNA solution structure, DNA dynamics, semi-flexible polymers, Fluorescence Correlation Spectroscopy

1. Introduction

The importance of DNA as a carrier of genetic information inspired the discovery of its double helical chemical structure in the past. However, on length scale beyond a few nanometers, DNA is a polymer composed of repeating nucleotide bases and its physical properties can be expected to be explained by the theories of polymer physics [1-5]. Unfortunately there had been lack of good experimental tools to measure the large scale (>50-100 nm) structure of DNA solutions. Our lab has recently shown that fluorescence correlation spectroscopy (FCS) (6-19) technique in a scanning mode [20] (scanning FCS, or SFCS) in combination with proper DNA staining enables such measurements.

While DNA is probably the most studied polymer and a variety of methods have been used in the past to study different DNA properties [21-28] only static light scattering (SLS) approach was used to probe the large scale structure of DNA solutions. However, SLS does not work well with DNA solutions due to poor scattering by DNA as well as because of the ubiquitous presence of dust particles in the sample which add to the noise [29]. Poor scattering at low concentration forces researchers to perform measurements at high concentrations and then extrapolate these data to the low concentration range. However, the theoretical justification for the extrapolation procedure is debatable [30].

Various studies [21-30] established that DNA is a semi-flexible polymer with bending fluctuations over the length scale of ~50 nm, known as persistence length l_p. DNA width of about 2 nm is much smaller than the persistence length. DNA flexibility is only apparent above persistence length whereas rod like behavior is observed below it. A particular interest in the context of this thesis is the structure of DNA chains at length scales much larger than its persistence length and in the so called "good" solvent. Polymer conformations depend on the solvent types which are classified according to how well monomers dissolve in the solvent [2]. Quantitatively, the interaction potential $U(\vec{r})$ between monomers in the solvent determines the type of the solvent through the second virial coefficient $v = \int d\vec{r}(1 - \exp(-U(\vec{r})/k_B T))$. In 'good' solvent ($v > 0$), monomers prefer solvent molecules rather than other monomers and as a result polymer chains swell. In 'bad' solvent ($v < 0$), monomers prefer other monomers and avoid solvent molecules which results in collapse of the chain. In 'theta' solvent ($v = 0$), monomer-monomer and monomer-solvent interactions

are exactly balanced leading to ideal chain behavior. For the length scales above l_p, the statistical properties of an ideal chain conform to those of random walk, with separations between two monomers j and k being Gaussian-distributed. The structure factor of ideal chains has an exact analytical expression known as Debye function [31].

DNA, due to its stiffness occupies larger volume in space as opposed to flexible chains of same length. This results in low probability of collisions between monomers and relatively weak excluded volume interactions. This situation has been termed 'marginal' by Schaefer, Joanny and Pincus (SJP) with predictions of 'theta'-like conditions for dilute polymer solutions and invariance of chain conformation in semi-dilute solutions [5].

However, prior to our measurements SJP theory has not been well supported. E.g. the structure of DNA coils in dilute regime, where different chains are well separated in space, is a contentious issue in the literature. Theoretical calculations show ideal chain behavior with no coil expansion for DNA with the contour length up to 60 microns [4], while the experimental studies in the past reported the scaling behavior characteristic of real chains and a significant coil expansion [24-30].

In our lab, we previously used SFCS in combination with non-specific DNA labeling to obtain the information on the overall structure of DNA solution [20]. The method is based on measuring correlations in emission intensity while performing confocal scan of fluorescently stained DNA solutions with high constant speed v. The basic idea is that the scanning is so fast that the sample is essentially "frozen" while passing through the sampling volume. Then the measured temporal correlation function $G(t)$ just reflects the spatial correlations with $r = vt$ providing the conversion from time to space. More precisely we show that the measured $G(r = vt)$ is a convolution of the pair correlation function $g(r)$ of the sample with the detection profile $F(r)$: $G(r) \propto \int d\vec{r}' g(\vec{r} - \vec{r}')F(\vec{r}')$. These two can be deconvolved in Fourier domain leading to the structure factor $S(\vec{q}) \equiv g(\vec{q}) \propto G(\vec{q})/F(\vec{q})$ of the labeled DNA solution.

Previously, our lab used this method to probe the structure of DNA randomly labeled with an intercalating dye (ethidium bromide, or EtBr). With such labeling approach, only the overall structure of DNA polymer solution is obtained and only particular buffer conditions could be used due to the dependence of EtBr binding to DNA on electrolyte concentration. However, the true power of our SFCS approach lies in the possibility of combining it with specific fluorescence labeling, so that a very diverse set of questions regarding DNA polymer

structure and dynamics can be addressed. In this work, SFCS, in combination with covalent DNA labeling is used extensively to study individual DNA coil structure in both dilute and semi-dilute solutions. Having DNA as our "test" polymer, we address here such classical questions of polymer physics as coil structure in dilute solutions, their end-to-end distance distribution, structure of individual coils in semi-dilute and concentrated regimes and the segmental dynamics of polymers in semi-dilute regime.

Furthermore, we exploit our approach to probe the dependence of DNA persistence length on salt concentration. DNA behaves as an ideal chain when electrostatic interactions arising from the negatively charged phosphate backbone is totally screened i.e. at high salt concentration. Chain expansion due to electrostatic repulsion is expected at low salt concentration and this is due to, according to classical theory by Odijk (1977) and Skolnick and Fixman (1977) (OSF), increase in persistence length. According to OSF theory DNA persistence length has two components in polyelectrolyte regime, an electrostatic part and a fixed neutral part which are additive and predicts reciprocal dependence of electrostatic part of the persistence length on salt concentration. The classical OSF theory, however, has been challenged by Manning citing previous experiments. In this work we attempt to resolve this issue by performing experiments on electrolyte concentration range not explored previously.

The following chapters of the thesis are organized as follows:

In Chapter 2 I give the overview of polymer physics concepts relevant for this work. I present the goals of my research in the Chapter 3. Then Chapter 4 introduces FCS and SFCS, and provides details on the physical principles and technical aspects of these techniques. In Chapter 5 I discuss the experimental design and working methods which include details of our experimental set up followed by molecular biology techniques adapted for the preparation of appropriate DNA samples for experiments. Next three chapters are devoted to the results and discussions of the experiments performed. The concluding remarks are presented in the final chapter.

2. Polymer Physics Background

Polymers are large macromolecules made up of repetitive units called monomers. The number of monomers in a single chain, which is also known as degree of polymerization, can easily reach up to 10^4-10^5. From chemical point of view polymers are characterized by their degree of polymerization, microstructure, chemical compositions and polymer architecture. However, from physical point of view conformations adopted by a polymer chain are more interesting since they provide information on the flexibility of chain, interactions between monomers and interactions with surrounding environment that the chain is in. Polymer chain conformation is essentially a shape adopted at an instant of time by the chain. Chain flexibility is an important issue that affects its conformation. Often chain flexibility depends on scale we choose to look at. Hence a chain that appears like a rigid rod at small scales, can also be considered as flexible at much larger scales. Attractive or repulsive interactions between monomers of the same chain or with monomers of other chains or solvent molecules also dictate chain conformations that lead to rich variety of polymer solution structures. Polymer conformations often leads to the situation where volume occupied by the chain (pervaded volume) is much larger than the sum of volumes of its monomers.

Physical descriptions of polymers usually omit of much of the chemical characteristics of the chain and focus only on universal features that are valid for broad range of polymers. Often simple models such as random walk on a lattice are used to explain static polymer features.

In the following of this Chapter we start with an overview of different polymer solution regimes and then discuss the statistics of the chains in dilute solutions, first those of ideal chains and then of real chains. We proceed by characterizing the properties of semi-dilute solutions of flexible polymers and then introduce a particular "marginal" regime for semi-flexible polymers that was predicted by SJP theory [5]. We finish this Chapter with the discussion of experimental knowledge of DNA solution structure prior to this work.

2.1. Polymer Solution Types

As mentioned above, monomer-monomer interactions are quantitatively described by the second virial coefficient of monomer-monomer interaction potential $U(r)$ as:

$$v = \int (1 - \exp[-U(\vec{r})/k_B T])d\vec{r} \quad (1)$$

The potential $U(r)$ is, in general, repulsive at short distances since two monomers cannot occupy the same space, and attractive at large distances due to Van der Waals interactions between them. Depending on temperature T and relative contributions of attraction and repulsion, the second virial coefficient can be either positive, negative or zero. It is easy to see this, if one assumes that the repulsive part $U_{rep}(r)$ is described by a hard core potential:

$$U_{rep} = \begin{cases} \infty & r < d \\ 0 & r \geq d \end{cases},$$

where d is monomer size, and that the Van der Waals energy $|U_{VdW}| \ll k_B T$ for $r > d$. Then by separating the integration for $r < d$ and $r > d$ ranges, and by expanding the exponent in the latter range, one can rewrite Eq. 1 as:

$$v = \frac{4}{3}\pi d^3 - \frac{A}{k_B T} \quad (2)$$

where $A = -4\pi \int_d^\infty U_{VdW}(r) r^2 dr > 0$ (since Van der Waals energy is negative).

First of all we notice that for the pure hard core interaction (no attraction), the second virial coefficient is just equal to the excluded volume of two spherical monomers. The presence of the attractive interactions decreases v.

While $v > 0$ (i.e. when the interactions are overall repulsive), the second virial coefficient allows to map the statistical behavior of a polymer chain with complicated interaction potential $U(r)$ between its monomers, to that of a chain with monomers interacting through a simple hard core potential such that each monomer "excludes" any other monomer within volume v around it.

As follows from Eq. 2 at some temperature, called θ- temperature, v turns to 0: statistically the attractive interactions are exactly balanced out by the repulsive interactions between the monomers. Further changes in the temperature lead to the overall attractive interactions between the monomers.

Respectively, in polymer physics three types of solvents are generally defined with respect to the monomer-monomer interactions: 1) in *good* solutions monomers of the chain repel each other, i.e. monomers “like” to be dissolved in the solvent ($v > 0$), 2) in *bad* solvents monomer interactions are overall attractive ($v < 0$) and the polymers tend to

precipitate from such solutions, and 3) in "*theta*" solvents polymers behave as ideal chains with no interactions between their monomers ($v = 0$).

Furthermore, three concentration regimes are distinguished for *flexible* polymers: 1) in dilute polymer solutions the distances between polymer coils exceed coil diameters, 2) in semi-dilute solutions, polymer coils overlap and interpenetrate, although the majority of molecules in solution are those of solvent, and 3) in polymer melts, the solvent is the minority phase. In general, problems of polymer physics refers to the individual chain structure and polymer network structure in various experimental conditions of different types of solvent and different concentration regimes. The main polymer parameters of interest are polymer flexibility (characterized by the persistence or Kuhn length), polymer contour length, excluded volume interactions, monomer concentration and the solution correlation/screening length.

In dilute solutions in theta solvents, the collisions between polymer segments can be neglected and the polymers conformations essentially map random walks. Polymer coils are called "ideal" in this situation, and as we discuss further in Section 2.2, their size, defined by their gyration radius R_g or mean square end-to-end distance R scales as $L^{1/2}$ with their contour length L. Polymer end-to-end distance is Gaussian distributed, as expected for a random walk.

However, theta solvent situation is rather unique. In "good" solvents, polymer segments repel each other and the chain is "swollen" by these interactions. As derived in Section 2.3, the coil size scales according to Flory prediction of $R_g \sim L^{3/5}$ in this regime. Rather than a simple random walk, the chain configuration is similar to self-avoiding random walk. The end-to-end distribution is not Gaussian and has minimum at zero end-to-end distance (in contrast to a Gaussian distribution which has maximum at this point). In bad solvents, the coils are collapsed into dense globules and their size scales as $R_g \sim L^{1/3}$.

In semi-dilute solutions in good solvents, the collisions between segments of different chains screen the interactions between the segments of the same polymer. The screening can be assessed directly by measuring solution structure. The concentration dependent screening of monomer correlations was first proposed by Edwards [1], who considered a semi-dilute solution of ideal polymers. Since the structure of ideal coils in dilute solutions is similar to a Coulomb potential, then in similarity to Debye-Huckel screening in electrolytes, the structure of dense polymer solutions exhibit screening with the characteristic length scaling as $\xi \sim c^{-0.5}$

with monomer concentration c. Although qualitatively correct, Edward's theory gives incorrect prediction for the concentration dependence of ξ for real chains in both good and theta solvents. The main deficiency of the theory is its essential mean-field approach neglecting the correlations in segmental collisions. As described in the Section **Error! Reference source not found.**, the theoretical approach based on scaling theory and backed by experimental evidence predicts the screening length dependence on inverse concentration with exponent 3/4 and 1 in "good" and "theta" solvents respectively [2,3]. The individual polymer coils become ideal at scales larger than ξ, i.e. their size scales as $R_g \sim L^{1/2}$.

Nevertheless, SJP theory [5] and our experimental results show that Edwards' prediction is correct for semi-dilute solutions of *semi-flexible* polymers: in the so-called *marginal* regime that we discuss at the end of this Chapter.

2.2. Ideal Chains in Dilute Solutions

Linear polymer chains that do not have interactions between distant monomers along the chain are known as ideal chains. At a particular temperature, called the theta temperature, there are no net interactions among monomers in a chain (second virial coefficient $v = 0$) and the polymers behave as ideal. Ideal chain conformation is also present in polymer melts and concentrated solutions due to screening of interactions by surrounding chains. In addition, SJP theory predicts ideal chain behavior for semi-flexible polymers such as DNA. Thus we will start by introducing idealized models of chains with no long-range interactions between monomers. We first calculate their end-to-end distance and then define their gyration radii. On the example of solutions of ideal chains, we then introduce two structural characteristics of polymer solutions: the pair-correlation function and the structure factor.

2.2.1. Freely jointed chain

Let's consider flexible polymer chain of n bond vectors $\vec{r}_i$. The characteristic size of su The sum of all bond vectors gives end-to-end vector :

$$\vec{R} = \sum_{i=1}^{n} \vec{r}_i$$

The ensemble average of end-to-end vector is obviously zero since there is no preferred direction of the vector in the ensemble of ideal polymer state. However, the average mean square end-to-end distance is non-zero and can be written as:

$$\left\langle R^2 \right\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n}\left\langle \vec{r}_i \cdot \vec{r}_j \right\rangle$$

The scalar product between bond vectors can simply be written in terms of cosines of angle between them assuming constant bond vector length $b = |\vec{r}_i|$:

$$\left\langle R^2 \right\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n}\left\langle \vec{r}_i \cdot \vec{r}_j \right\rangle = b^2\sum_{i=1}^{n}\sum_{j=1}^{n}\left\langle \cos\theta_{ij} \right\rangle = nb^2 + b^2\sum_{i=1}^{n}\sum_{j\neq i}^{n}\left\langle \cos\theta_{ij} \right\rangle$$

In the freely jointed chain model, different bond vectors are totally uncorrelated $\left\langle \cos\theta_{ij} \right\rangle = 0$ for $i \neq j$. Then:

$$\left\langle R^2 \right\rangle = nb^2$$

Many other models of ideal polymers, such as those discussed in the next sections, can be mapped onto the freely jointed chain model by choosing an appropriate effective segment length replacing the actual bond vectors. This effective segment length is known as Kuhn length.

We note that since this model is equivalent to a simple random walk, it is easy to obtain not only the variance but also the whole distribution of the end-to-end distance. E.g. the distribution of end-to-end displacement along the *X* axis is given by a Gaussian:

$$P(x) \propto \exp\left(-\frac{3x^2}{2R^2}\right) = \exp\left(-\frac{3x^2}{2nb^2}\right) \tag{3}$$

This means that upon the separation of polymer ends by a distance x, the chain loses its entropy:

$$\Delta S = \Delta \ln P = -\frac{3x^2}{2nb^2}.$$

and increases its free energy by

$$F_{ent} = -k_B T \Delta S = k_B T \frac{3x^2}{2nb^2}. \tag{4}$$

2.2.2. Freely rotating chain

The main assumptions inherent in this model of an ideal chain are fixed bond angles, bond lengths and equal likelihood of torsion angles. In order to calculate mean square end-to-end distance only the correlation along the adjacent bond vectors is taken into account, while

the component that is perpendicular averages out to zero due to free rotations of torsion angle. Therefore correlation between bond vectors $\vec{r}_i$ and $\vec{r}_{i+k}$ separated by k links can be written as:

$$\left\langle \vec{r}_i \cdot \vec{r}_{i+k} \right\rangle = b^2 (\cos\theta)^k .$$

Then

$$\left\langle R^2 \right\rangle = nb^2 + b^2 \sum_{i=1}^{n} \sum_{j \neq i}^{n} \left\langle \vec{r}_i \cdot \vec{r}_j \right\rangle = nb^2 + 2b^2 \sum_{i=1}^{n} \sum_{j>i}^{n} \left\langle \vec{r}_i \cdot \vec{r}_j \right\rangle =$$

$$= nb^2 + 2b^2 \sum_{i=1}^{n} \sum_{k=1}^{n-i} \left\langle \vec{r}_i \cdot \vec{r}_{i+k} \right\rangle = nb^2 + 2b^2 \sum_{i=1}^{n} \sum_{k=1}^{n-i} \cos^k \theta$$

The $\cos^k \theta$ term decays rapidly with increasing k, so for long enough polymers the last summation can be extended to infinite series resulting in simplification of the expression:

$$\left\langle R^2 \right\rangle = nb^2 + 2nb^2 \frac{\cos\theta}{1-\cos\theta} = nb^2 \frac{1+\cos\theta}{1-\cos\theta} = Ll_k ,$$

where we introduced Kuhn length l_k by requesting that the chain contour length $L = nb$ be divided into n' segments of of length l_k: $L = n'l_k$ so that $\left\langle R^2 \right\rangle = n'l_k^2$ as in the freely jointed chain model. Then $l_k = \left\langle R^2 \right\rangle / L$ and

$$l_k = b \frac{1+\cos\theta}{1-\cos\theta}$$

2.2.3. Worm like chain model (WLC)

The worm like chain model is a continuous version of Kratky-Porod model that takes into account the special case of small bond angle due to chain stiffness and is highly relevant for semiflexible polymers such as DNA. One can show that correlations between tangent vectors $\vec{t}(s')$ and $\vec{t}(s)$ to the polymer contour at contour points s and s' decay exponentially with their separation (as they do for freely rotating chain):

$$\left\langle \vec{t}(s)\vec{t}(s') \right\rangle = \exp\left(-\frac{|s-s'|}{l_p} \right),$$

where l_p is the persistence length of the chain that defines its flexibility. Then:

$$\left\langle R^2 \right\rangle = \left\langle \left(\int_0^L \vec{t}(s) ds \right)^2 \right\rangle = \int_0^L \int_0^L \left\langle \vec{t}(s) \vec{t}(s') \right\rangle ds' ds =$$

$$= \int_0^L \left[\int_0^L \exp\left(-\frac{|s - s'|}{l_p} \right) ds' \right] ds = 2l_p^2 \left(\frac{L}{l_p} - 1 + \exp\left(-\frac{L}{l_p} \right) \right).$$

It is obvious from the final expression for the WLC model that when we take the limit of contour length much smaller than persistence length we can expand the exponential part to third order and recover rod like limit. Similarly, for chains much longer than persistence length one can recover ideal chain limit:

$$\left\langle R^2 \right\rangle \cong 2Ll_p,$$

Thus the Kuhn length for worm like chain is $l_k = 2l_p$. The Kuhn length is simply the length scale below which polymer behaves like a rod due to flexibility constraint and above Kuhn length, the polymer retains its flexible coil like characteristics.

2.2.4. Radius of gyration

A convenient way to describe the size of polymers including those polymers that are not linear is their gyration radius R_g. This is because all physical objects can be assigned a gyration radius that describes their inertia with respect to rotation. In the case of polymers, the square of gyration radius of a polymer chain is defined as an average of squared distance between monomers R_i of that chain from the chain center of mass R_{cm}:

$$R_g^2 = \left\langle \frac{1}{n} \sum_{i=1}^{n} \left(\vec{R}_i - \vec{R}_{cm} \right)^2 \right\rangle, \tag{5}$$

where the center of mass of the polymer chain is the average of all monomer position vectors assuming same monomer mass:

$$R_{cm} = \frac{1}{n} \sum_{j=1}^{n} \vec{R}_j . \tag{6}$$

Substituting Eq. 5 into Eq. 6 we obtain a convenient expression for the gyration of radius:

$$R_g^2 = \frac{1}{n}\sum_{i=1}^{n}\left\langle\left(\vec{R}_i^2 - 2\vec{R}_i\vec{R}_{cm} + \vec{R}_{cm}^2\right)\right\rangle = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left\langle\left(\vec{R}_i^2 - 2\vec{R}_i\vec{R}_j + \vec{R}_i\vec{R}_j\right)\right\rangle = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left\langle\left(\vec{R}_i^2 - \vec{R}_i\vec{R}_j\right)\right\rangle =$$

$$= \frac{1}{2}\left[\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left\langle(\vec{R}_i^2 - \vec{R}_i\vec{R}_j)\right\rangle + \frac{1}{n^2}\sum_{j=1}^{n}\sum_{i=1}^{n}\left\langle(\vec{R}_j^2 - \vec{R}_j\vec{R}_i)\right\rangle\right] = \frac{1}{2n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left\langle\left(\vec{R}_i - \vec{R}_j\right)^2\right\rangle.$$

Finally, by counting each monomer pair only once in the double sum we obtain:

$$R_g^2 = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j>i}^{n}\left\langle\left(\vec{R}_i - \vec{R}_j\right)^2\right\rangle . \tag{7}$$

2.2.5. Radius of gyration of ideal chains

The Eq. 7 can now be used to calculate mean square radius of gyration of ideal chains. We will consider very long chains only ($L \gg l_k$) for which the resulting formula is generic. For such long chains we can change the summation over monomers into integration over contour of the chain. We also change indices into continuous coordinates. Hence we get:

$$R_g^2 = \frac{1}{n^2}\int_0^n di \int_i^n \left\langle\left(\vec{R}(i) - \vec{R}(j)\right)^2\right\rangle dj .$$

In the above expression $\vec{R}(i)$ is the monomer position vector. Assuming Gaussian distribution of monomer-monomer distances for ideal chain $\left\langle\left(\vec{R}(i) - \vec{R}(j)\right)^2\right\rangle = |i-j|b^2$, we get:

$$\left\langle R_g^2\right\rangle = \frac{b^2}{n^2}\int_0^n\int_i^n (j-i)didj = \frac{nb^2}{6} = \frac{\left\langle R^2\right\rangle}{6}$$

Thus there is a simple relation between mean square end-to-end distance and mean square gyration radius of an ideal chain.

2.2.6. Density-density correlation function

The structure of simple and complex fluids is conveniently described by the density-density correlation function $g(r) = \langle\delta c(0)\delta c(r)\rangle/\bar{c}$, where for polymer solutions $c(r)$ and $\bar{c}$ are respectively an instantaneous local monomer concentration at the position $\vec{r}$ and an average monomer concentration and $\delta c(r) = c(r) - \bar{c}$. The angular brackets denote

ensemble averaging and the solution is assumed to be homogenous and isotropic. In dilute solutions, the density-density correlation function describes the average concentration of monomers as a function of distance from any given monomer.

For dilute solutions of ideal chains one can estimate the functional form of density-density correlation function inside the coil (i.e. for $r < R$) by counting the number m of monomers within a sphere of radius r and calculating their density $\sim m/r^3$. For ideal chain statistics $r^2 \sim mb^2$ and then:

$$g(r) \approx \frac{m}{r^3} \cong \frac{1}{b^2 r} \ , \tag{8}$$

Hence we can see that pair correlation function decreases with increasing distance in a manner similar to the Coulomb potential.

2.2.7. Structure factor

In experiments, such as in different scattering methods (X-ray, visible light, neutron etc), one measures the Fourier transform of the density-density correlation function, called the structure factor $S(q)$. For the ideal chains it can be calculated exactly. It was first given by Debye and hence it is known as Debye function:

$$S(q) = \frac{2}{q^4 R_g^4}\left[\exp(-q^2 R_g^2) - 1 + q^2 R_g^2\right], \tag{9}$$

For $q > \pi/R_g$, corresponding to spatial distances within the coil ($r < R$), the Debye function gives $S(q) \propto 1/q^2$. This is expected as the Fourier transform of $1/r$ behaves as $1/q^2$.

In this thesis work the Debye function is extensively used as a fitting function for the structure factor of DNA solutions measured with Scanning Fluorescence Correlation Spectroscopy (SFCS) technique to provide experimental evidence for the ideal chain behavior of DNA coils, measure coil gyration radii, as well as to study DNA solution structure under various solvent/buffer conditions.

2.3. Real chains in dilute solutions

2.3.1. Excluded volume interactions

In good solutions ($v > 0$) the interactions between distant monomers on a chain render the polymer behavior non-ideal. In order understand the importance these interactions, we need to first estimate number of monomer-monomer contacts in a single chain. As discussed above, the second virial coefficient allows us to treat the monomers as hard spheres with the excluded volume v. In the mean field approximation, the probability of a monomer to collide with any other monomer is equal to the fraction ϕ of the chain volume occupied by monomers. For an ideal chain (e.g. in the Freely Jointed Chain model) we expect:

$$\phi \sim \frac{nv}{R^3} \sim \frac{nv}{(b\sqrt{n})^3} \sim \frac{v}{b^3\sqrt{n}}\,, \tag{10}$$

The above expression shows that the probability of a collision for any given monomer in a long chain is very small. However, because of the relatively low entropy of a polymer chain (as e.g. compared to the solution of free monomers), it appears that the relevant parameter is the total number N_{col} of monomer collisions within the chain:

$$N_{col} \sim \phi n \sim \frac{v}{b^3}\sqrt{n}\,, \tag{11}$$

and this can be a large number for a long chain. Indeed, for a flexible chain far from the θ-conditions one expects $v \sim b^3$, so that $N_{col} \sim \sqrt{n}$ and so there are numerous collisions between the monomers in rather short flexible chains.

These collisions strongly affect polymer conformations and have to be taken into account in theoretical descriptions. The first such description was given by Flory described in the next subsection.

2.3.2. Flory theory of real chains

In conditions when the net interactions between monomers are repulsive ($v > 0$) the solvent properties are called *good* and the chains are called *real*, as opposed to *ideal* chains with no interactions between the monomers ($v = 0$). The real chains are swollen due to the net repulsive interactions between monomers. However, swelling also results in the loss of chain entropy. Flory theory estimates the repulsive and entropic contributions to the free

energy, minimization of which gives the numerical value of the universal exponent related to the chain size.

The interaction energy F_{int} is estimated within the second virial coefficient expansion that gives k_BT per each collision:

$$F_{int} = k_BTN_{col} \sim k_BTn\phi \sim k_BTn\frac{nv}{R^3} = k_BT\frac{n^2v}{R^3}, \quad (12)$$

Swelling is expected to separate the ends of the chain by the distance $\sim R$. Thus the entropic term is estimated from the energy of extension of an ideal chain Eq. 4:

$$F_{ent} \approx k_BT\frac{R^2}{nb^2},$$

The total free energy given by

$$F = F_{int} + F_{ent} \approx k_BT\left(\frac{n^2v}{R^3} + \frac{R^2}{nb^2}\right), \quad (13)$$

is then minimized with respect to R giving the Flory radius:

$$R_f \approx (vb^2)^{1/5}n^{3/5}, \quad (14)$$

Thus the real chain size R_f has a stronger dependence on the polymer length n than that of the ideal chain: the exponent is 3/5 vs. 1/2 for the ideal case. The chain is swollen.

The Flory theory can be generalized to any dimension d by defining $\phi = nv/R^d$ in the Eq. 12 resulting in the general power law:

$$R_f \propto n^{\nu}, \quad (15)$$

where the exponent $\nu = 3/(d+2)$.

Remarkably, despite theory's simplicity the exponents given by Flory theory are correct up to 4 dimensions. In fact, the theory overestimating both parts of the total free energy. However, thanks to a spectacular mutual cancellation of errors, the resulting exponents are correct.

2.3.3. Excluded volume interactions in semi-flexible polymers

As we have seen in the Subsection 2.3.1, flexible chains display significant excluded volume interactions and therefore their mean end-to-end distance or gyration radius scales

with number of monomers as predicted by Flory theory. According to WLC model, semi-flexible chains with contour length L much larger than Kuhn length l_k behave like flexible polymers. Naively, one would expect that for $L \gg l_k$ the semiflexible chains would experience significant excluded volume interactions. We show here this is not quite the case: in a significant range of contour lengths, semiflexible polymers behave as ideal chains.

As we discussed in the Subsection 2.2.3 the WLC model can be mapped onto the Freely Jointed Chain model by assuming $b = l_k$ and $n = L/l_k$: the Kuhn are the effective monomers. However, these monomers do not have a roughly spherical shape as for flexible chains, but are rather similar to rigid thin cylinders of length l_k and a diameter $d \ll l_k$ representing the hard core diameter of the chain segments. The excluded volume for such cylinders can be pictured as a $l_k \times l_k \times d$ cuboid, and therefore $v \approx l_k^2 d$. We can use then Eq.11 to estimate the number of collisions within such a chain:

$$N_{col} \sim \frac{v}{l_k^3}\sqrt{n} \sim \frac{d}{l_k}\sqrt{n}\,, \tag{16}$$

The contour length L_t (or n_t in terms of number of monomers) up to which a chain behaves like an ideal chain with no excluded volume interactions can be found from the condition $N_{col} = 1$. This leads to:

$$n_t \sim \frac{l_k^2}{d^2}\,, \tag{17}$$

For flexible polymers $d \sim l_k$, which gives the trivial contour length of about one monomer $n_t \approx 1$: as expected, one cannot neglect excluded volume interactions in flexible chains in good solvents. However, for semi-flexible polymers like DNA with $l_k \gg d$ and with effective excluded volume $v = l_k^2 d$ we obtain $n_t = l_k^2/d^2 \gg 1$. For double stranded DNA $l_k = 100\ nm$ and effective diameter $d = 4\ nm$ (the diameter of DNA double helix is 2 nm and the additional ~ 2 nm come under physiological buffer conditions from the screened electrostatic repulsion between the charged backbones), we end up with very significant threshold contour length $L_t = n_t l_k \approx 60\ \mu m$ [4] . Hence DNA should behave like an ideal chain up to very large contour lengths. Overall, the probability of collisions between Kuhn segments is relatively small owing to the relatively large volume occupied by a semi-flexible chain.

The semi-flexible chains with contour lengths larger than threshold size ($L \gg L_t$) behave as Flory chains, yet at the scales below L_t they still obey the ideal chain statistics.

2.4. Semi-dilute polymer solutions

2.4.1. Semi-dilute solutions of flexible polymers

In dilute solutions the distance between individual chains is much larger than the coil size. Thus interactions between different chains are rare and the conformation of each chain is independent of the overall solution concentration. However, as polymer concentration is increased the chains approach each other and eventually their coils start to overlap and interpenetrate. The overlap concentration c^* is the monomer concentration of solution at which polymer chains start to "touch" each other.

If we consider a polymer solution at overlap concentration when chains occupy all the space between them then the concentration is simply the number of monomer in each chain divided by the chain volume which is assumed to be spherical i.e. $c^* \cong \frac{n}{R^3}$. For ideal polymers $R \propto n^{1/2}$ leading to $c^* \propto n^{-1/2}$. For a real polymers with $R \propto n^{3/5}$ we obtain $c^* \propto n^{-4/5}$. Obviously the overlap concentration decreases with polymer length in both cases.

As the concentration is further increased beyond the overlap concentration, polymer chains start to form a 3D mesh. The collisions between different chains lead to the screening of correlations in monomer positions beyond the characteristic screening length ξ (also called mesh size). For individual chains, at the length scale below mesh size there is no screening effects and chain segment retains its dilute conformation. At length scale larger than ξ the polymer loses the distinction between its own monomers and those of other chains in its vicinity and as a result the statistics are those of an ideal chain. Thus in the semi-dilute regime the coils of the flexible polymers are expected to contract with monomer concentration *c*.

The first prediction for such screening and for the dependence of ξ on monomer concentration *c* was given by Edwards [1]. Edwards considered ideal chains (at all scales) and gave a nice analogy to Debye-Huckel screening in electrolytes: since the correlations within individual chains decay as a Coulomb potential (see Eq. 8), the correlations within the

dense solutions of such chains decay as $\xi \propto c^{-1/2}$ as for electrostatic screening in electrolytes or plasma.

While the prediction of screening by Edwards is correct, the $\xi(c)$ dependence derived by him is wrong for flexible polymers: first of all ideal polymers are not supposed to collide, and then the assumption of uncorrelated collisions between different segments made in Edwars' theory is incorrect.

A proper dependence of the screening length on monomer concentration in both good and theta solvents is given by the scaling theory [2]. Assuming that there are g monomers per mesh size we have $c \sim g/\xi^3$. Since within a mesh size there are no collisions between different polymers, the statistics at this scale is that of a free polymer, so that $\xi \propto g^{\nu}$. Combining the last two expressions we obtain:

$$\xi \propto \frac{1}{c^{\nu/(3\nu-1)}}, \tag{18}$$

For flexible polymers in good solvents $\nu = 3/5$ and therefore $\xi \propto c^{-3/4}$, while for theta conditions $\nu = 1/2$ and, respectively, $\xi \propto c^{-1}$.

2.4.2. Semidilute solutions of semi-flexible polymers. Marginal regime.

As discussed in the Subsection 2.3.3, due to their large sizes coils of semi-flexible polymers behave like ideal chains with no excluded volume interactions in dilute solution. In semi-dilute solutions of such polymers an interesting scenario emerges since such solutions have chains overlapping with each other yet the collisions between different polymers may be quite rare, again because of relatively sparse organization of semi-flexible chains. The screening effects in such solutions cannot be described within the scaling theory.

This unique situation of semi-dilute solution of semi-flexible chains was addressed theoretically by Schaefer, Joanny and Pincus (SJP). They predicted a large range of concentrations where the properties of such polymer solutions are “marginal” [5]. SJP treat weak excluded volume interactions in semi-dilute solution as a perturbation of dilute solution structure $S_0(q)$. The expressions they obtained for semi-dilute solution structure $S(q)$ and screening length ξ using mean-field approximations are:

$$S(q) = \frac{S_0(q)}{1 + kT(v + wc)S_0(q)},$$

$$\xi = \frac{l_p}{(vc + wc^2)^{1/2}},$$

Where v and w are the second and third virial coefficients. In the marginal semi-dilute region when $wc \ll v$ we get the important $\xi \sim c^{-1/2}$ dependence, same as Edwards' theoretical prediction. Essentially, the Edwards' theory works in this case since both its assumptions: 1) the ideal structure of coils in dilute solutions, and 2) lack of correlations in segmental collisions - are satisfied in the case of semi-flexible polymers (both due to the sparse structure of such coils). Interestingly, although by the strength of excluded volume interactions marginal solutions lie in-between "good" and "theta" solutions, their exponent of concentration dependence of screening length is not intermediate between these two cases.

Because of the relatively weak interactions, the structure of individual semi-flexible chains is not supposed to be affected by the mesh (as opposed to the flexible coils that contract with concentration). This is somewhat counter-intuitive since if the chains keep the same structure as in dilute solutions then there should be no screening either. What this means, in fact, is that in the marginal regime relatively small changes in the structure of individual coils can lead to big changes in the screening length, i.e. in the structure of the whole polymer solution.

2.5. Experiments on DNA solution structure prior to this work

Summarizing the theoretical results on the solution structure of semi-flexible polymers, one expects: 1) up to rather large contour lengths $L_t \sim l_k^3/d^2$ coils should behave as ideal in dilute solutions, 2) the screening length dependence on monomer concentration is $\xi \propto c^{-1/2}$ in semi-dilute solutions, 3) the structure of individual coils in semi-dilute solutions is not significantly affected by the mesh.

In dilute solutions, DNA might be well a unique semi-flexible polymer for which a wide range of a predicted ideal behaviour can be checked: synthetic semi-flexible polymers are not stiff enough to produce a significant range ideal behaviour, while other biopolymers such as actin and tubulin filaments are too rigid, and so bend only slightly and do not form coils. DNA on the other hand has rather large Kuhn length, so that for physiological buffer conditions the effective aspect ratio of its Kuhn segments $l_k/d \sim 25$ is a large number. Yet DNA contour length can reach tens of microns and more so that $L \gg l_k$, i.e. DNA of such lengths appears as a flexible chain. We note that the situation with DNA is somewhat

different from polymers in θ -solvents, where the second virial coefficient of interactions between monomers is zero. This is not the case for DNA: the second virial coefficient is non-zero, but the number of collisions within a chain is small due to the polymer stiffness.

As discussed in the Subsection 2.3.3 the estimated crossover contour length for DNA in physiological conditions is about $60\mu m$ or 200 kbp [4,5]. Hence for shorter chains one does not expect departure from Gaussian chain statistics. This means no significant coil expansion due to repulsive interactions between monomers within the coil of size up to 200Kbp.

Prior to our measurements the experimental evidence for this claim was lacking and quite a few indirect experimental results actually contradicted the ideal chain behavior for DNA. The experimental studies demonstrating strong excluded volume interactions with the Flory scaling $R_g \propto L^{0.57-0.59}$ include measurement of diffusion coefficient of DNA lengths $2.6 < L < 130\mu m$ [25], the segment distribution of substrate tethered DNA of lengths $15 < L < 60\mu m$ [26], and end-to-end distance distribution measurements by Atomic Force Microscope (AFM) of DNA length $0.5 < L < 15\mu m$ [27]. Moreover, static light scattering (SLS) technique that is capable of directly measuring polymer coil structure in the solution also shows non Gaussian attribute of DNA coils of size $9 < L < 60\mu m$ with the scaling $R_g \propto L^{0.54-0.55}$ and significant coil expansion of ~30% [28-30]. The use SLS approach to measure of DNA coil structure in dilute solution is not straightforward because of the poor scattering by DNA that requires performing the experiments at higher concentrations and then extrapolation their results to $c = 0$ [30]. Such extrapolation by itself requires a theory [28]. Another problem with SLS is its sensitivity to scattering by dust particles which are often close in size to DNA molecules [29] and difficult to filter out.

Thus, prior to our experiments there has been a significant discrepancy between the theoretical predictions and measured data in dilute DNA solutions.

For semi-dilute solutions of semi-flexible polymers, the experimental attempts to test the existence of the *marginal* regime using synthetic polymers have been inconclusive [5, 36-39]. As well as for dilute solutions, here the main problem is relatively short Kuhn length of synthetic polymers even those that are considered semi-flexible. DNA molecules can serve as an excellent material to study the marginal regime. Indeed, our lab developed a new method based on SFCS to measure the structure of DNA solutions and published the first evidence of the marginal regime – the concentration dependence of the screening length of $\xi \propto c^{-1/2}$ [20]. However, another prediction of SJP theory – the lack of changes in individual coil

structure in semidilute regime has not been tested. In this these I show that our SFCS approach can assess this feature as well.

3. Research Objectives

The main thrust of this work is in understanding the structure of DNA solutions both in dilute and semi-dilute regime. We use SFCS method in combination with non-specific labeling of DNA by intercalating dye, as well as by covalently bound dye, to show that the structure of DNA coils in dilute solutions obeys the Debye expression for ideal chains.

We further develop our method to measure the end-to-end distance distributions for DNA chains using the SFCS in combination with specific labeling of DNA ends. We show that the end-to-end distances are normally distributed, again as expected for ideal chains.

Based on the above two measurements we check the end-to-end distance and gyration radius dependences on DNA contour length and find them consistent with the ideal chain behavior.

Then having small fraction of DNA chains labeled covalently and embedded into the mesh of non-labeled DNA, we test the predictions of SJP theory for the structure of individual coils in the semi-dilute regime.

The FCS formalism needed for all of our measurements is discussed in the next chapter, the instrumentation and sample preparation methods are given in the Chapter 5, and the results are described in the Chapter 6.

The experiments in the Chapter 6 are done at physiological salt (~150 mM) concentrations, in conditions when DNA charge is screened within ~ 1nm distances from the double helix and DNA behaves essentially as a neutral polymer. At lower ionic strengths, the increase in electrostatic interactions leads to the stiffening of DNA chain. While Odijk-Skolnik-Fixman (OSF) theory of this stiffening is widely accepted, there are some others (e.g. by Manning) that give different predictions. I give the background material on this topic and our experimental results at low salt concentrations in the Chapter 7.

While most of my theses deal with the structure of DNA solutions, in the Chapter 8 I present proof-of-the-concept results for an SFCS based method to measure the segmental dynamics of DNA in semi-dilute solutions. Generally, DNA dynamics can be monitored with the standard, static FCS. However, in dense solutions the dynamics is so slow that fluorophores photobleach before they diffuse out of the confocal volume. This totally distorts the characteristics of the sampling volume and leads to unreliable results. We show that by incorporating scanning into FCS, we avoid this type of problems.

4. Fluorescence Correlation Spectroscopy

Molecular system at equilibrium undergoes thermal fluctuations revealing its kinetic and dynamic properties which can be investigated using Fluorescence Correlation Spectroscopy (FCS). Originally the technique was developed by Magde and co-workers [6-8] to measure diffusion and binding kinetics of Ethidium bromide onto double stranded DNA. It was later significantly improved by Rigler's group through the incorporation of confocal illumination/detection scheme [9]. FCS, as the name suggests, is essentially an experimental tool to analyze correlation of fluorescence intensity fluctuations due to events at the molecular level. Molecular systems have fluctuating properties which are related to their response to external perturbations according to the fluctuation-dissipation theorem [10]. Several techniques including FCS can be used to exploit this feature to study the equilibrium statistics of the molecular system.

Equilibrium state of the molecular systems at molecular scale is highly dynamic due to naturally present thermal noise and this leads to large fluctuations in a small observed system. FCS is capable of exploiting this naturally present phenomenon of deviations from equilibrium at the molecular scale. In FCS either molecules are fluorescent or tagged with fluorescent dye. The amplitude of fluorescence intensity fluctuations depends on the number of fluorescent molecules in the sampling volume, and hence can be used to determine concentrations. The kinetics of the emission fluctuations depends on the kinetics of thermal motion of the labelled molecules and can be used to measure e.g. molecular diffusion coefficients. In general, any dynamic feature, such as binding kinetics or protein/polymer/DNA conformational dynamics, can be measured as long as it is coupled to changes in fluorescence.

4.1. FCS History

Magde, Elson and Webb in 1972 were the first to introduce the technique of Fluorescence Correlation Spectroscopy [6]. The first application of this technique by them was for the study of binding of ethidium bromide (EtBr) onto double stranded DNA. EtBr is an intercalating fluorescent dye whose fluorescence quantum yield increases drastically when bound to DNA. The fluctuation in emitted intensity when the dye is excited by appropriate laser light is caused by both binding/de-binding and diffusion of molecules in and out of

observation volume. In the first FCS experimental set-up the laser used was 514 nm at 6 kW/cm^2 and photomultiplier tube was used to image the observation volume which was aided by parabolic mirrors to collect the fluorescence light. The scattered excitation light was filtered by the solution of K2Cr2O7. The observation volume was much larger than what is possible in today's FCS set-up. Nevertheless this simple technology allowed Magde, Elson and Webb to obtain the diffusion coefficient of labelled DNA.

Several technological improvements in the initial FCS scheme were incorporated later such as dichroic mirrors and filters to separate emission from excitation light, improved and stable lasers, avalanche photo-diode (APD) instead of photomultiplier tubes etc. The major improvement came in the 90's with the incorporation of confocal illumination-detection scheme which resulted in drastic decrease in the observation volume to a size of about 0.2-1 fl. This scheme resulted in much better signal to noise ratio and improvement in spatial resolution only limited by diffraction. Small observation volume, although reducing total emission intensity, results in increased amplitude of intensity fluctuations hence its importance for single molecule applications. Since the initial development FCS has been used for various single molecule studies such as conformational changes, screening for drugs, diffusion inside the cells and on membrane, adsorption and enzyme kinetics [11].

4.2. FCS Instrumentation

In this section technical aspect of FCS experiments are discussed following the references [9], [11] and [12]. An outline of FCS set-up utilising confocal illumination-detection scheme is shown in Fig. 4.1. A microscope is often a part of the set-up. Laser beam of appropriate wavelength for a given application is usually expanded to appropriate size before directing towards the microscope using reflecting mirrors (not shown). The dichroic mirror deflects the beam into the high power (usually water immersion) objective which focuses the beam into the sample of fluorescent molecules. The molecules are excited by the laser light which results in fluorescence emission and this emission is collected by the same objective and passed through the dichroic mirror and a band pass filter to avoid unwanted light. The emitted light is then focused with a lens onto the pinhole located in the image plane of the optical system. The pinhole rejects out of focus light very efficiently and also improves the size of the sampling volume laterally. An avalanche photodiode with high detection efficiency and

time resolution is used to detect fluorescent emission. The detected signal is then subjected to real-time processing by hardware correlator and displayed on the computer screen.

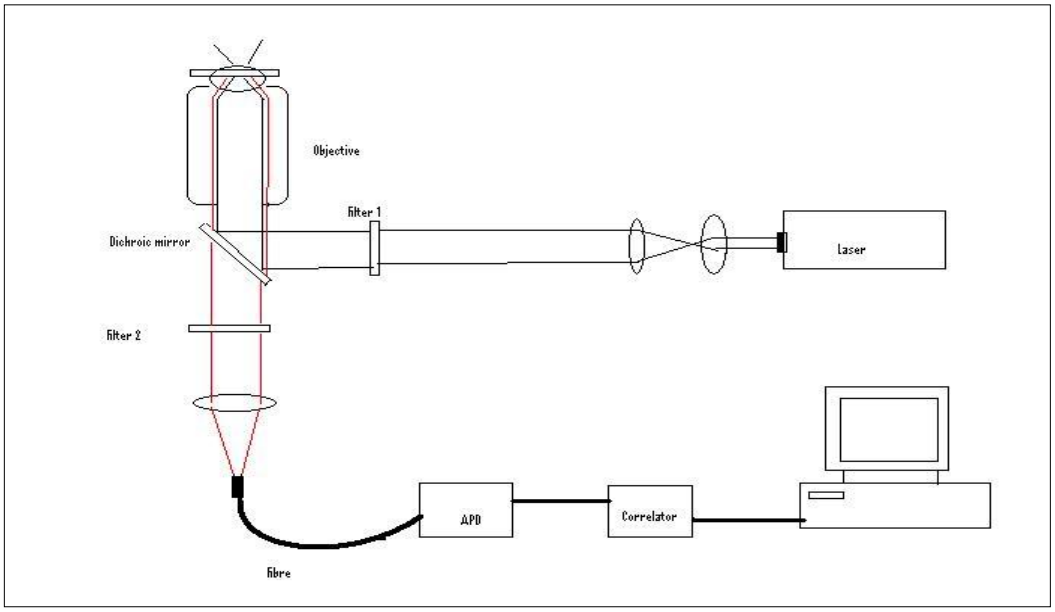


Fig. 4.1. The schematics of an FCS setup.

4.3. FCS Formalism

In FCS experiment a time-averaged correlation function of the emission fluctuations is measured. However, assuming that the system is ergodic, it is more convenient to perform derivations using ensemble averaging instead of time averaging. Here we first derive general FCS formalism and then apply it to the particular cases of single labelled object diffusion and to the measurements of the solution structure factor.

We define the autocorrelation function of fluctuations $\delta I_{em}(t) = I_{em}(t) - < I_{em} >$ in emitted intensity $I_{em}(t)$ for a time lag t as:

$$G(t) = \frac{\langle \delta I_{em}(0) \delta I_{em}(t) \rangle}{\langle I_{em} \rangle}. \tag{19}$$

The normalization here is somewhat different from the standard one (in which the denominator is $\langle I_{em} \rangle^2$), but here we find our definition more convenient since $G(t)$ defined this way does not depend explicitly on the molecular concentration. The fluctuations in emission result from the concentration fluctuations in the observation volume:

$$\delta I_{em}(t) = Q \int \delta c(\vec{r}, t)\, I_{ex}(\vec{r}) d\vec{r}, \tag{20}$$

where $\delta c(\vec{r},t) = c(\vec{r},t) - \bar{c}$ is the deviation of local concentration field $c(\vec{r},t)$ from the average concentration $\bar{c}$, Q is solely a function of the fluorophore brightness and the setup efficiency in collecting photons, and $I(\vec{r})$ is the excitation intensity profile. Then the numerator in Eq. 19 can be written as:

$$\langle \delta I_{em}(t)\delta I_{em}(0)\rangle = Q^2 \iint d\vec{r}\, d\vec{r}' I(\vec{r}) I(\vec{r}') \langle \delta c(\vec{r},0)\delta c(\vec{r}',t)\rangle \,.$$

Assuming translational symmetry of the sample and moving to Fourier space, we further obtain:

$$\langle \delta I_{em}(t)\delta I_{em}(0)\rangle = \\ = Q^2(2\pi)^3\Omega^{-1}\int d\vec{q}\, |I(\vec{q})|^2 \langle \delta c^*(\vec{q},0)\delta c(\vec{q},t)\rangle \,, \tag{21}$$

where Ω is the total sample volume. Now, using the representation $c(\vec{r},t) = \sum_j^N \delta(\vec{r} - \vec{r}_j(t))$, where $\vec{r}_j(t)$ is a position of a fluorescent molecule and N is the total number of molecules, we have:

$$\langle \delta I_{em}(t)\delta I_{em}(0)\rangle = Q^2\bar{c}\int d\vec{q}\, |I(\vec{q})|^2 \left[\frac{1}{N}\sum_{j,k=1}^{N} \langle e^{i\vec{q}(\vec{r}_j(t)-\vec{r}_k(0))}\rangle - \bar{c}\delta(q)\right]. \tag{22}$$

For confocal FCS setups the rear aperture of the objective lens is usually underfilled and the excitation profile (actually, an effective point spread function, PSF) is well approximated by an axisymmetric Gaussian:

$$I(\vec{r}) = I_0 \exp\left(-\frac{2(x^2+y^2)}{w_{xy}^2} - \frac{2z^2}{w_z^2}\right), \tag{23}$$

whose Fourier transform (also called mode transfer function, MTF) is another Gaussian:

$$I(\vec{q}) = I_0 \frac{w_{xy}^2 w_z}{8} exp\left(-\frac{w_{xy}^2}{8}\left(q_x^2+q_y^2\right) - \frac{w_z^2}{8}q_z^2\right), \tag{24}$$

where w_{xy} and w_z characterize the extent of excitation-illumination profile in the radial and axial directions respectively. The aspect ratio $\omega = w_z/w_{xy}$ of the profile is typically ~5, so that $\omega^2 \gg 1$.

The average emission then is:

$$\langle I_{em}\rangle = Q\bar{c}\int I(\vec{r})d\vec{r} = \left(\frac{\pi}{2}\right)^{3/2} QI_0\bar{c}w_{xy}^2 w_z\,. \tag{25}$$

We can now proceed to consider particular cases of the FCS application.

4.3.1. Static FCS to measure dynamics of single labelled molecules

By far the most frequent use of FCS is in measuring the diffusion of small molecules. Here we will derive a more general formula that can be used to measure the dynamics of single-labeled large molecules provided that the labels displacements are normally distributed. With the generalized expression one can study e.g. the internal dynamics of DNA molecules [50].

For independent point sources of fluorescence, all of the terms in the sum in the Eq. 22 for $j \neq k$ vanish except for $q = 0$. Their contribution for $q = 0$ cancels out with that of $\bar{c}\delta(q)$. The terms with $j = k$ are all similar so that the Eq. 22 can be rewritten as:

$$\langle \delta I_{em}(t)\delta I_{em}(0)\rangle = Q^2\bar{c}\int d\vec{q}\,|I(\vec{q})|^2\langle e^{i\vec{q}\Delta\vec{r}(t)}\rangle\,, \tag{26}$$

where $\Delta\vec{r}(t) = \vec{r}_j(t) - \vec{r}_j(0)$ is the fluorescence source displacement over time *t*. For normally distributed displacements $\langle e^{i\vec{q}\Delta\vec{r}(t)}\rangle = e^{-q^2\langle\Delta r^2(t)\rangle/6}$ holds with $\langle\Delta r^2(\mathrm{t})\rangle$ being the mean square displacement of a label at time *t*. Substituting $q^2 = q_x^2 + q_y^2 + q_z^2$ and Eq. 24 into Eq. 26 and integrating over each component of $\vec{q}$, we get:

$$\langle \delta I_{em}(t)\delta I_{em}(0)\rangle = \frac{\pi^{3/2}Q^2I_0^2w_{xy}^2w_z}{8\left(1+\frac{2\langle\Delta r^2(\mathrm{t})\rangle}{3w_{xy}^2}\right)\sqrt{1+\frac{2\langle\Delta r^2(\mathrm{t})\rangle}{3w_z^2}}}\,. \tag{27}$$

Finally, introducing the aspect ratio $\omega = w_z/w_{xy}$ and substituting the above expression and Eq. 25 into Eq. 19, we obtain:

$$G = \frac{QI_0}{2\sqrt{2}}\left(1+\frac{2\langle\Delta r^2(\mathrm{t})\rangle}{3w_{xy}^2}\right)^{-1}\left(1+\frac{2\langle\Delta r^2(\mathrm{t})\rangle}{3\omega^2w_{xy}^2}\right)^{-1/2}\,. \tag{28}$$

For the diffusion of small molecules, such as simple fluorophores we have $\langle\Delta r^2(\mathrm{t})\rangle = 6Dt$ where *D* is the diffusion coefficient of the molecules. Then Eq. 28 converts into the standard FCS expression:

$$G = \frac{QI_0}{2\sqrt{2}}\left(1+\frac{t}{\tau}\right)^{-1}\left(1+\frac{t}{\omega^2\tau}\right)^{-1/2}, \quad (29)$$

where $\tau = w_{xy}^2/4D$ is the characteristic diffusion time of the molecule across the confocal volume. Respectively, it defines the characteristic decay time of the correlation function. Clearly, under the normalization used the amplitude of correlation function in Eqs. 28&29 does not depend on fluorophore concentration but only on its molecular brightness.

4.3.2. SFCS to measure the structure of multiple labelled sample

The standard FCS measures the kinetics of diffusing fluorescent molecules through temporal correlation of emission intensity fluctuations. The ability to measure spatial correlation in the observation volume opens up the possibility of structural measurements of sample which can be realised through scanning mode. The idea of incorporating either sample or beam scanning motion in FCS to enhance diffusion of molecules dates back to 70's [13,14]. This and closely related Image Correlation Spectroscopy (ICS) approach was implemented to study the dynamics of biological molecules and their aggregates [15-19]. Only recently SFCS was implemented to measure the structure rather than the dynamics of solutions [20]. Much of this thesis work involves use of this tool developed in our lab.

In combination with an appropriate formalism and a particular scanning pattern, SFCF can be adapted to measure spatial correlation $g(R)$ in fluorophore positions or its Fourier transform the structure factor $S(q)$ [20]. The basic idea is to provide enough drift velocity V to the sample of fluorescent molecules, such as uniformly labelled DNA, through confocal observation volume so that the Brownian motion is negligible while the coils move through the sampling volume. This results in fluctuations in emission intensity as a result of inhomogeneous spatial distribution of dye. Scanning the sample with high drift velocity essentially converts spatial correlation function $g(R)$ in fluorophore positions to temporal correlation function $G(t) = \langle \delta I_{em}(0)\delta I_{em}(t)\rangle$ with $R = Vt$. More precisely, the measured correlation function $G(R = Vt)$ is a convolution of a characteristic excitation-detection profile of the instrument and spatial correlation function $g(R)$ of dye positions. De-convolution which leads to separation of spatial correlation from instrumental characteristics can be carried out in Fourier space which results in Fourier transform of $g(R)$. In the case when fluorophores label DNA densely (i.e. the distances between them are much smaller than

light wave length) $g(R)$ reflects the monomer density-density correlation function and its Fourier transform gives the solution structure factor $S(q)$.

It is easier to start the formal treatment from the Eq. 21. Assuming that the forced sample motion through the excitation beam is the only reason for the changes in the spatial distribution of the fluorophores, we have $\delta c(\vec{r},t) = \delta c(\vec{r} - \vec{V}t, 0)$ and as a result in the Fourier domain $\delta\tilde{c}(\vec{q},t) = \delta\tilde{c}(\vec{q},0)e^{i\vec{q}\vec{V}t}$. We substitute this into Eq. 21 and change the variables $\vec{R} = \vec{V}t$ to get:

$$G(\vec{R}) \propto \langle \delta I_{em}(t)\delta I_{em}(0)\rangle \propto \int d\vec{q}\, |I(\vec{q})|^2 \langle |\delta c(\vec{q})|^2\rangle e^{i\vec{q}\vec{R}}\,, \tag{30}$$

We skip normalization here for simplicity and since we are going to normalize the structure factor to 1 at the end, consistently with the Debye formula Eq. 9. The above expression obviously means that:

$$G(\vec{q}) \propto \langle |\delta c(\vec{q})|^2\rangle |I(\vec{q})|^2 \propto S(\vec{q})|I(\vec{q})|^2\,, \tag{31}$$

where we use the fact that $\langle |\delta c(\vec{q})|^2\rangle$ is a Fourier transform of the density-density correlation function defined in the Subsection 2.2.6 and therefore up to a numeric coefficient is equal to the structure factor for strongly labelled DNA [22]. Hence the above expression allows us to extract the structure factor $S(q)$ of DNA solution from the Fourier transform of the measured SFCS functions $G(R)$ provided that we calibrate instrumental MTF characteristics $I(\vec{q})$.

4.3.3. SFCS to measure end-to-end distribution in DNA

The true power of our approach as compared to the static light scattering is in the specific fluorescence labelling capability. Only the labelled places of interested will produce light and the resulting correlations. One of the examples of such approach that we consider here is in the specific labelling of DNA polymer ends and the resulting ability to study the end-to-end distributions.

We start again with Eq. 22 and we assume high speed of scanning, so that there are no internal motions at the time scales of passage through the beam. Then $\vec{r}_j(t) = \vec{r}_j(0) + \vec{V}t$. Similar to the derivation in 4.3.1 the majority of terms in the sum of Eq. 22 are trivial as they contain the positions of independent labels. However, the labels attached to the ends of the same DNA molecule are correlated. Let e.g. labels #1 and #2 belong to the same molecule. Then we can rewrite Eq. 22 as:

$$\langle \delta I_{em}(t)\delta I_{em}(0)\rangle = Q^2\bar{c}\int d\vec{q}\,|I(\vec{q})|^2 e^{i\vec{q}\vec{V}t}\left[1+\langle e^{i\vec{q}\vec{r}_{12}}\rangle\right], \quad (32)$$

where $\vec{r}_{12}=\vec{r}_2-\vec{r}_1$ is the end-to-end distance of a DNA molecule. The first term in the brackets comes from the autocorrelation of a label with itself, while the second term reflects cross-correlation in positions between the two labels belonging to the same molecule. The term $\langle e^{i\vec{q}\vec{r}_{12}}\rangle$ is the characteristic function/Fourier transform of the end-to-end probability density function $P_{ee}(\vec{r}_{12})$. Indeed by definition:

$$\langle e^{i\vec{q}\vec{r}_{12}}\rangle = \int d\vec{r}_{12}\,P_{ee}(\vec{r}_{12})e^{i\vec{q}\vec{r}_{12}} = P_{ee}(\vec{q})\,.$$

So that if $P_{ee}(\vec{r}_{12})$ is normally distributed as expected for ideal chains, then $P_{ee}(\vec{q})$ should also be Gaussian. Substituting this and Eqs. 32&25 into Eq. 19 we have for the double labelled molecules:

$$G(\vec{R}) = \frac{Q}{\int I(\vec{r})\vec{r}}\left[\int d\vec{q}\,|I(\vec{q})|^2 e^{i\vec{q}\vec{R}} + \int d\vec{q}\,|I(\vec{q})|^2 P_{ee}(\vec{q})e^{i\vec{q}\vec{R}}\right], \quad (33)$$

where we again changed the variables $\vec{R}=\vec{V}t$. The first contribution $G_1(\overrightarrow{R}) = \frac{Q}{\int I(\vec{r})\vec{r}}\int d\vec{q}\,|I(\vec{q})|^2 e^{i\vec{q}\vec{R}}$ can be measured on single labelled objects and subtracted from the total correlation function. The difference $G_{12}(\vec{R}) = G(\vec{R}) - G_1(\vec{R})$ describes the cross-correlation between the DNA ends. Moreover, as can be seen from the Eq. 33 $G_{12}(\vec{R})$ is related to the characteristic function of the end-to-end distribution in a manner that the correlation function of multiple labelled sample is related to the sample's structure factor (Eq. 31):

$$G_{12}(\vec{q}) = P_{ee}(\vec{q})|I(\vec{q})|^2\,. \quad (34)$$

Notice also that under the normalization choosen $G(\vec{R})$ and $G_1(\vec{R})$ do not depend on concentration, so one can just measure them and subtract them directly.

4.3.4. SFCS to measure dynamics of single labelled molecules

As I show in the last chapter, the static FCS approach discussed in the subsection 4.3.1 does not work very well on the dense samples where the dynamics is very slow, and the labels might photobleach before they leave the confocal volume. In this cases, scanning FCS might

be useful. In addition, as I describe in the next chapter, we use scanning FCS to calibrate the characteristics of our sampling volume. Here we derive the formalism for such an approach.

As in the subsection 4.3.1 we consider the thermal dynamics of independent (not necessarily small) molecules labelled at a single position. Now, however, we will assume that the sample is moved with constant speed $\vec{V}$. In contrast to the approach of the subsections 4.3.2 and 4.3.3, we do not assume that the speed is high.

The counterpart of the Eq. 26 for the current case is:

$$\langle \delta I_{em}(t)\delta I_{em}(0)\rangle = Q^2\bar{c}\int d\vec{q}\,|I(\vec{q})|^2 e^{i\vec{q}\vec{V}t}\langle e^{i\vec{q}\Delta\vec{r}(t)}\rangle\,, \tag{35}$$

where $\Delta\vec{r}(t)$ is the label displacement sorely due to the thermal motion. We will again assume that these displacements are normally distributed so that $\langle e^{i\vec{q}\Delta\vec{r}(t)}\rangle = e^{-q^2\langle\Delta r^2(t)\rangle/6}$ and consider the experimentally relevant case of lateral scanning direction, e.g. in the *X* direction $V_x = V; V_y = V_z = 0$. Once we substitute Eq. 24 integrals over different directions separate. Straightforward integration over dq_x gives:

$$\int dq_x \exp\left[-\frac{w_{xy}^2}{4}q_x^2 - \frac{\langle\Delta r^2(t)\rangle}{6}q_x^2 + iq_xVt\right]$$
$$= 2\sqrt{\frac{\pi}{w_{xy}^2\left(1+\frac{2\langle\Delta r^2(\mathrm{t})\rangle}{3w_{xy}^2}\right)}}\exp\left[-\frac{(Vt)^2}{w_{xy}^2\left(1+\frac{2\langle\Delta r^2(\mathrm{t})\rangle}{3w_{xy}^2}\right)}\right].$$

The integrations over other directions give the same result as in the subsection 4.3.1. The overall result for the correlation function G_V measured at speed *V* is:

$$G_V = G_0\exp\left[-\frac{(Vt)^2}{w_{xy}^2\left(1+\frac{2\langle\Delta r^2(\mathrm{t})\rangle}{3w_{xy}^2}\right)}\right], \tag{36}$$

where G_0 is the static correlation function given by the Eq. 28.

In the particular case of the diffusion of small molecules $\langle\Delta r^2(\mathrm{t})\rangle = 6Dt$ and

$$G_V = G_0\exp\left[-\frac{(Vt)^2}{w_{xy}^2(1+t/\tau)}\right], \tag{37}$$

where $\tau = w_{xy}^2/(4D)$ as before and G_0 is given by the Eq. 29. This result was previously derived by Magde *et al* [14].

5. Materials and Methods

5.1. SFCS Instrumentation and calibration

5.1.1. SFCS setup

In the previous section we explained the basic idea and theoretical formalism behind the SFCS technique to measure the structure of polymer solution. In this section we discuss our experimental realisation of SFCS approach. This is the primary experimental tool employed in this thesis work in combination with various molecular biology techniques for specific and non-specific fluorescent labelling of DNA molecules.

The measurements are carried out in a home built confocal SFCS setup. It consists of an Ar-ion laser (Advantage 163D Spectra-Physics) providing duel wavelength options (488 nm and 514 nm) as an excitation light source for fluorophore. The two linear polarizers are used to adjust the beam power which is typically kept between 2-12 μW, measured before deflecting mirrors of the setup. The deflecting mirrors, which are also used for alignment purpose, direct the beam into the microscope unit along the optical axis. Inside the microscope a dichroic beamsplitter (Q525 Chroma) is used to deflect the laser line into a high power objective lens (UPLAPO 60X1.2W, Olympus). The same objective also collects the emitted light from the sample which is passed through a bandpass filter (HQ565/80, Chroma) into a multimode optical fibre (50 μm core). A fiber-coupled photon counting avalanche photodiode (SPCM-AQR-15-FC PerkinElmer) receives the emission from the output of the multimode optical fibre which is then fed into digital correlator (Flex2k-12Dx2, correlator.com). For the scanning of sample a flexure XYZ piezo-stage (Trotor 101, PiezoSystem Jena) is used. The piezo-stage is equipped with capacitive sensors making precise monitoring of stage position and speed possible. The stage motion and monitoring of positions is controlled by analog voltages supplied through DAQ board. Furthermore the output of the position sensors is also measured by the DAQ board. Also it was possible to keep the speed of the stage constant to within 1-2% standard deviation by feeding specifically optimized analog signal pattern. The sample speed for structure measurements (i.e. experiments described in the Chapters 6 and 7) were kept at around 4mm/sec which is the maximum speed attained in the setup. Also, it was noted that results are independent of speed even up to the speed as low as 1 mm/sec.

In principle the SFCS formalism strictly requires constant velocity scan of the sample along the straight line. However, this requirement can be relaxed to include also a constant speed scan along any trajectory if the radius of curvature of that trajectory at any given point is much larger than the confocal volume size of $\sim 0.25 \mu m$. In practice the sample is scanned in lateral plane along a trajectory of an imperfect circle of $\sim 40 \mu m$ diameter whose centre orbits along another circular trajectory of $\sim 30 \mu m$ diameter (Fig. 5.1). This pattern of scanning is chosen to cover a large surface thereby alleviating the problem of frequent revisiting of same positions which is necessary to avoid photobleaching and correlated noise.

The typical sample volume in our experiment is about 1.5 μl. The solution is sandwiched between two clean glass cover slips which fit onto the sample holder. The sample is then aligned in the middle of the laser beam and the water immersed objective is lowered to focus inside the sample.

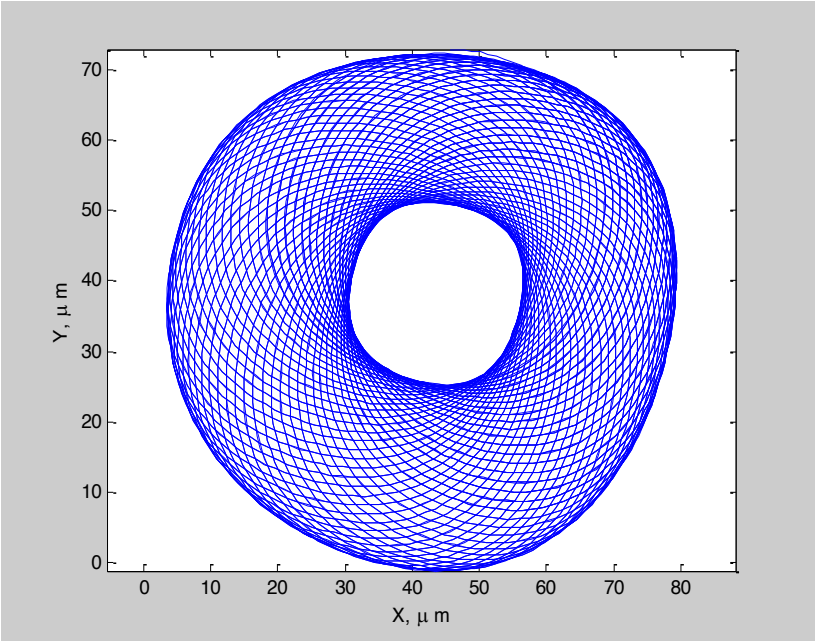


Fig. 5.1. Typical sample trajectory in the scanning mode.

5.1.2. PSF Measurement using SFCS

Calibration of instrumental mode transfer function $I(\vec{q})$ is necessary for the structure measurements. The direct measurement of PSF carried out through the imaging of small

fluorescent beads (50 nm in size) shows the field consistent with Gaussian assumption (Fig. 5.2). However, the image is too noisy to determine the Fourier transform of the PSF.

Thus we assume Gaussian MTF (Eq. 24) and need then to determine only the width of the confocal volume w_{xy} in XY plane and its aspect ratio $\omega = w_z/w_{xy}$. To determine ω, we fit the correlation functions obtained with the expression (cf. Eq. 29):

$$G_0(t) = A\left(1+\frac{t}{\tau}\right)^{-1}\left(1+\frac{t}{\omega^2\tau}\right)^{-1/2}$$

where A is the amplitude of correlation function and τ is the characteristic diffusion time across the observation volume. Typically we obtain $\omega^2 \approx 30$.

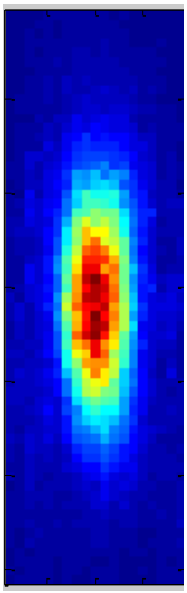


Fig. 5.2. Instrumental PSF as measured by XZ scans of small (50 nm) fluorescent beads. The horizontal axis is in the *X* direction and the vertical axis is in *Z* direction. The pixel size is ~ 50nm.

Then ω_{xy} is obtained from the comparison of correlation curves obtained from diffusing Rh6G molecules in static and scanning FCS measurements. Combining the expression for static FCS correlation function G_0 with that of scanning FCS Eq. 37 we see that the following holds:

$$\left(\frac{G_V(t)}{G_0(t)}\right)^{A/G_0(t)} = \exp\left[-\frac{(Vt)^2}{w_{xy}^2}\right], \tag{38}$$

as long as $t \ll \omega^2\tau$, the condition that is easily satisfied since $\omega^2 \approx 30 \gg 1$. Thus we just fit the left hand side of the Eq. 38 vs. Vt with a Gaussian to determine w_{xy}. Typically, we have $w_{xy} = 0.25 \pm 0.01\ \mu m$.

5.1.3. Fourier transform of the correlation function

For the determination of the structure factor or of the characteristic function of the end-to-end distribution through the Eqs. 31 and 34 respectively, we need to determine the Fourier transform of the correlation function $G(\vec{q})$ or $G_{12}(\vec{q})$. In general, this requires scanning in all directions in 3D and performing a Fourier transform of the 3D function $G(\vec{R})$ or $G_{12}(\vec{R})$. However, since the underlying density-density correlation function $g(r)$ and the end-to-end distribution $P_{ee}(r)$ are spherically symmetric, there is enough information in the lateral scans in order to obtain the necessary Fourier transforms.

In practice we do only lateral scans in the *XY* plane and we make use of the fact that $\omega^2 \approx 30 \gg 1$. This allows us to first assume no axial dependence of the PSF, i.e. $\omega^2 \to \infty$, resulting in $|I(\vec{q})|^2$ and $G(\vec{q})$ being non-zero only for $q_z = 0$. For the isotropic system the Fourier-Bessel transform (2D Fourier transform for isotropic function) of measured $G(R)$ gives $G(\vec{q})$ and $S(\vec{q}) \propto G(\vec{q})\exp(w_{xy}^2 q^2/4)$. From here on an iterative procedure starts which takes into account the finite ω^2. Each iteration involves the calculation of the expected $G(R)$ from $I(\vec{r})$ and the previously obtained $S(q)$. The difference between the expected and measured $G(\vec{r})$ is the correction for the $S(q)$. However due to the high accuracy of the initial approximation, one step iteration is usually sufficient for the convergence.

The same procedure is applied to G_{12} (Eq. 34) in order to extract the characteristic function $P_{ee}(q)$ of the end-to-end distance

5.2. Molecular biology techniques

Several existing molecular biology techniques were adapted for the preparation of covalently labelled DNA molecules in order to study various fundamental physical aspects of DNA coils such as end-to-end distance distribution, conformational changes, monomer dynamics, and concentration dependent effects. Furthermore in order to prepare clean and homogeneous DNA samples we implemented adaptive and novel extraction and purification techniques. In this chapter we discuss the molecular biology techniques employed in this work. Also we discuss the experimental difficulties and challenges of appropriate sample

preparation for our studies and our adaptive and often novel approach to deal with those challenges.

5.2.1. Polymerase Chain Reaction (PCR)

Polymerase chain Reaction is an indispensable tool in molecular biology in order to amplify a few copies of DNA chain into several orders of magnitude. Kary Mullis is credited for its discovery in 1983 which has now become the most widely used technique in molecular biology and bio-chemistry for various applications such as DNA cloning for sequencing, disease diagnosis (hereditary and infectious), genetic fingerprinting, functional analysis of genes etc.

The key to the method of PCR is thermal cycling which is essentially a repeated heating and cooling of the reaction mixture, containing all the necessary ingredients, for melting (separation of DNA strands) and enzymatic replication of DNA. The thermal cycling of reaction mixture to a series of temperature steps ensures first the physical separation of template DNA strands then attaching of short pieces of oligonucleotides also known as primers to the complementary regions of template and then finally synthesis and amplification of DNA by polymerase enzyme. The polymerase enzyme assembles or synthesizes a new DNA strand from the nucleotides, the four building blocks of DNA. However the initiation of synthesis by polymerase requires DNA primers that are complementary to the targeted region of amplification in the template which also makes the procedure selective. All PCR techniques and its variants require a heat-stable DNA polymerase enzyme, the most common one being Taq polymerase which was originally isolated from the bacterium *Thermus aquaticus*. One of the most important features of PCR based technique is the ability to exponentially amplify a DNA segment which is possible since the DNA amplified is itself used as a template as the reaction proceeds. In addition to primers and enzyme other ingredients required in PCR reaction are template, buffer, nucleotides and magnesium ion in a typical reaction volume of $50 - 200\mu l$. Furthermore typical thermal cycle or temperature change include initialization step at 98°C for 1-5 min, melting step at 94-98℃ for 20-30 sec, annealing step at 50-60℃ for 20-40 sec, extension or elongation step at 65-72℃ at the rate of 500-1000bp/min, and a single final elongation step for the same duration as that of the elongation step.

5.2.2. PCR and Covalent DNA Labelling

In the context of the research objectives of this thesis covalent attachment of dye to DNA is of crucial importance. Non-specific labelling of DNA such as with intercalating dye EtBr does not allow us to study specific polymeric properties such as end-to-end distribution or the structure of individual chain in the matrix of other chains. Also, we cannot change buffer or electrolyte concentration of non-specifically stained DNA solution since the binding of such dyes to DNA depends on the buffer conditions. Furthermore, the intercalating dyes like EtBr are weakly fluorescent molecules requiring high labeling densities which affect contour length [23] and may also affect persistence length. Covalent binding of strongly fluorescent dye (carboxyrhodamine 6G) to the DNA molecules was achieved by the substitution of one of the nucleotide bases with a modified nucleotide or an analog (amino-allyl-dUTP, AA-dUTP) during PCR synthesis of DNA and subsequent reaction of such DNA molecules with the fluorescent dyes.

Two step PCR was adapted for non-specific amplification of Lambda DNA segments of various lengths. We successfully synthesized modified nucleotide incorporated DNA fragments of 4.2, 8.5 and 14.2 kbp using a single forward primer CCGTTCTTCTTCGTCATAAC and different reverse primers GCA CTC TTTCTCGTAGGTACT, CGCTTTATTACCATCCTCAG and CAC GCA GGG GAA ATA TCT TT respectively. Modified to native nucleotide ratio of 1:1 in the reaction mixture was used. Denaturation of template was done at 98℃ for 10sec followed by annealing at 3℃ above lower annealing temperature recommended by primer manufacturer (Midland Certified) for the given primer pair. Extension time was kept at 45 sec/Kbp for the first step which included 15 cycles whereas in the second step it was increased to 2min/Kbp for 20 cycles. Final concentrations of DNTPs, Mg, and Primers in reaction volume of 50uL, were 400uM, 5mM, and 5uM respectively. The amount of template and polymerase (Vent exo-, NEB) in the reaction mixture were 5ngm and 0.5uL respectively. After PCR, DNA purification kit (GE Healthcare) was used to remove salts, unused primers and dNTPs. The analog AA-dUTP has a chemically reactive amino group which reacts and binds to the succinimidyl ester moiety of the dye Rh6G. Following the overnight reaction of base substituted DNA molecules with Rh6G dye in 0.1M NaHCO3 buffer, gel electrophoresis was performed to isolate the DNA molecules of correct size from non-specific PCR products and again extracted using the purification kit following the recommended protocol from the

manufacturer. Extracted DNA samples were in Tris pH 7.4 buffer with 100mM NaCl. Typical dye to base pair ratio of ~ 1:200 was obtained which is small enough not to perturb the DNA coil structure. Furthermore, since the modified nucleotide AA-dUTP is also a triphosphate and similar in size to its native counterpart in DNA, neither the persistence length nor the contour length of such labeled DNA is changed significantly. This allows us to measure DNA coil structure in almost native and unperturbed state.

To synthesize DNA labeled specifically at its ends (for measuring end-to-end distributions) we first prepared short DNA fragments using PCR with one of the bases substituted by an analog (AA-dUTP). Fragments were then reacted with Rh6G dye and cut with restriction enzyme and then ligated to longer DNA pieces with the same restriction overhangs. We designed primers with restriction site in order to facilitate enzymatic cleavage and ligation to longer pieces. Gel electrophoresis technique was used to separate longer labeled DNA from unreacted fragments and unwanted concatamers. Both single and double labeled DNA samples were prepared in order to subtract autocorrelation from measured total correlation of double labeled DNA so as to measure end-to-end correlation function directly.

5.2.3. Gel Electrophoresis

Gel Electrophoresis is a common method of separation of macromolecules in bio-chemistry and molecular biology. The macromolecules typically separated by gel electrophoresis, based on size and charge, are proteins, DNA and RNA. The purpose of gel electrophoretic separation of macromolecules is often analytical such as post amplification of DNA by PCR. In this method electric field is applied along the gel matrix which forces charged molecules to move, the speed of motion being dependent on the size of the molecules. Shorter molecules move faster through the pores and travel further than the longer and heavier ones for a given period of time which is the basis of separation in gel electrophoresis. Gel is usually made of either agar or polyacrylamide of which the former is the only one used in the context of this thesis. Agarose gel is composed of long unbranched carbohydrate chains without cross-links, unlike polyacrylamide, that forms the pores large enough for the separation of large macromolecules. For most applications agarose gel concentration is between 0.5% to 2% w/v in the same buffer as the running buffer (Tris/Acetate/EDTA, TAE). The optimum and the most frequently used concentration, however, is 1% w/v. The concentration dictates the pore size and strength of the gel.

In practice the sample containing molecules being separated are dispensed into small wells at one end of the gel slab sitting in an electrophoresis chamber containing running buffer. The gel slab containing the sample is immersed in the buffer which provides ions for the current and to maintain the pH at a relatively stable value. The electrophoresis chamber is connected to the power source providing the appropriate voltage for the necessary current. The voltage supplied forces the positively charged species to move towards the cathode and the negatively charged ones to move toward the anode. Electrophoresis run time depends on several factors such as voltage supplied, gel size and concentration, size of the macromolecules and separation or resolution required. After the completion of gel electrophoresis macromolecules such as DNA in the gel is stained with a dye and illuminated with an ultraviolet lamp for visualization. The most common dye for visualization of DNA is Ethidium Bromide which intercalates with DNA and fluoresces reddish orange when illuminated with UV. The gel can also be photographed for the analysis or the desired band of DNA can be cut out of the gel for purification.

5.2.4. Field Inversion Gel electrophoresis (FIGE)

Conventional gel electrophoresis is not able to separate DNA molecules beyond the size of 20Kb. Molecules larger than 20Kb have size-independent mobilities through gel matrix hence cannot be seen as distinct bands in the conventional gel electrophoresis. In order to circumvent this limitation a technique that involves periodically inverting the electric field was developed [24]. Successful separation of DNA molecules much larger than 20Kb (up to 200Kb) was accomplished by this simple technique. The key to this technique is periodic inversion of electric field which involves the repetition of switching cycle of certain forward and reverse durations. The switching cycle is chosen based on the size of the DNA fragments to be separated. The larger the molecules lower the switching frequency and vice versa. Since FIGE requires longer time than the conventional electrophoresis heat exchanger is needed to keep the running buffer and gel at a fixed temperature.

We adapted field inversion gel electrophoresis technique to separate lambda DNA concatomers for the purpose of SFCS structure measurements. We used home built polarity changer and commercial function generator (Agilent) for the switching cycle and connected them to the conventional gel electrophoresis unit.

The conditions we used for the field inversion gel electrophoresis were 0.2x TAE buffer, 0.75% agarose gel concentration, 250 mHz switching frequency with 3 sec forward and 1 sec

reverse pulse, electric field gradient of 8V/cm and total run time of 12 and half hours including the initial half an hour straight run.

We managed to get good separation of longer DNA as in ref. 24 without the need for complicated recirculation of buffer through heat exchanger by simply diluting the standard TAE buffer fivefold which lowered the current and minimised gel heating. We keep the switching cycle and total run time as in ref. 24 depending on the size of the DNA fragments. We checked the validity of our approach by separating the DNA fragments of standard marker and restriction enzyme cleaved lambda DNA.

5.2.5. Extraction and Purification

DNA embedded in gel after the electrophoresis can be extracted and purified using several types of commercially available purification kits. For the most part of this work we also used such kit (DNA purification kit, GE Healthcare). However for the DNA size larger than 50Kb (Lambda concatamers) such extraction kits are ineffective and also not recommended by the manufacturers.

We noticed an interesting phenomenon in which large DNA molecules embedded in gel pieces could be forced out to a certain extent. This simple technique involves freezing gel pieces containing DNA in a clean tube at -20℃ for about half an hour and then thawing at room temperature. This results in separation of buffer fluid containing DNA from the gel matrix. This procedure can be repeated for probably slightly better yield. We did not notice significant difference between shorter or longer freezing time in terms of yield. We estimated the yield to be about 5-15% depending on the length of DNA which is significantly better in comparison to the commercial kits tried although they were not recommended by the manufacturers for such large DNA molecules. For our purpose since the volume as well as the buffer was not optimum we used commercially available concentrator/buffer exchanger (Ultratech, Milipore) for the final sample preparation for the SFCS measurements. The main advantage of this technique is its simplicity since it requires no special equipments or chemicals and also sample handling is very gentle (no significant amount of pippeting) which is important for large DNA molecules. We verified the validity of our technique by repeating field inversion gel electrophoresis of extracted long DNA sample and checking it against the commercially available lambda concatamer marker.

5.2.6. DNA preparation for experiments with nonspecific labelling

Most of our measurements in dilute solutions were performed on DNA labeled non-specifically with EtBr. DNA fragments of length shorter than that of lambda DNA (48502 bp) were obtained by digesting lambda DNA as well as pUC18 (2686 bp) and pBR322 (4361 bp) with commercially available restriction enzymes.

The longest DNA fragments used in our measurements were lambda DNA and its dimer. Lambda DNA was annealed (65℃ for 10 min) first to open their overhangs and then ligated to each other in order to prepare the dimer. Since ligation reaction proceeds beyond dimerization, we used field inversion gel electrophoresis to separate dimer from other longer fragments.

The experiments were performed in phosphate buffered saline (PBS) in physiological electrolyte condtions (10 mM phosphate pH 7.4 buffer with 137 mM NaCl, 2.7 mM KCl). In the case of non-specific labeling the EtBr dye concentration was about $10\mu M$. At this concentration the dye to base pair ratio of about 1:5 is obtained. EtBr labeling of DNA extends the contour length by a factor of ~ 1.4 without affecting it persistence length [23,32]. Prior to each measurement the samples of lambda and its dimer were annealed at 65℃ for 10 min to make sure the overhangs did not close the chains.

For the measurements in the dilute regime, the concentration of lambda dimer was less than $\sim 1\mu g/mL$ and that of lambda DNA and shorter molecules was less than $4\mu g/mL$ whereas the overlap concentration of lambda DNA is $10 - 15\mu g/mL$ [20].

6. Static DNA structure studies

In this Chapter I first present our measurements of the structure factor of DNA molecules in dilute solutions, then of their end-to-end distributions and, finally, of the structure of individual DNA molecules in semi-dilute solutions.

6.1. Dilute solutions

In dilute solutions individual chains are on average much further from each other than their size measured in either gyration radius or end-to-end distance. This is true for any kind of polymers including semi-flexible polymers like DNA. Furthermore, in dilute solution individual chain conformation is independent of concentration of the solution and only factors that dictate chain conformations are excluded volume interactions and solvent quality. As discussed in the introductory chapters, our main motivation for the study of the structure of DNA coils in dilute solution stems from the fact that there is a significant discrepancy between the theoretical predictions and experimental findings on such solutions.

In this work we demonstrate for the first time the measurement of the structure of DNA coils ranging in size from $\sim 0.6\mu m$ to $\sim 46\mu m$ in dilute solution. We performed SFCS measurements on both non-specifically labeled (EtBr) as well as covalently labeled (Rh6G).

The examples of measured SFCS correlation functions for various DNA sizes in dilute solutions are presented in Fig. 6.1. The measured correlation functions are plotted vs. $r^2 = (Vt)^2$. As seen in the plot the spatial correlations between DNA segments extend to larger distances for longer DNA coils. For all the measured correlation functions we obtained good fits with the expression Eq. 30 where we used Debye function (Eq. 9) for the structure factor of an ideal chain with the gyration radius R_g as a fitting parameter and $I(\vec{q})$ as given by the Eq. 24 with parameters w_{xy} and $w_z = \omega w_{xy}$ measured as described in the Subsection 5.1.2. An example of w_{xy} measurement according to the Eq. 38 is represented by the dashed line the Fig. 6.1.

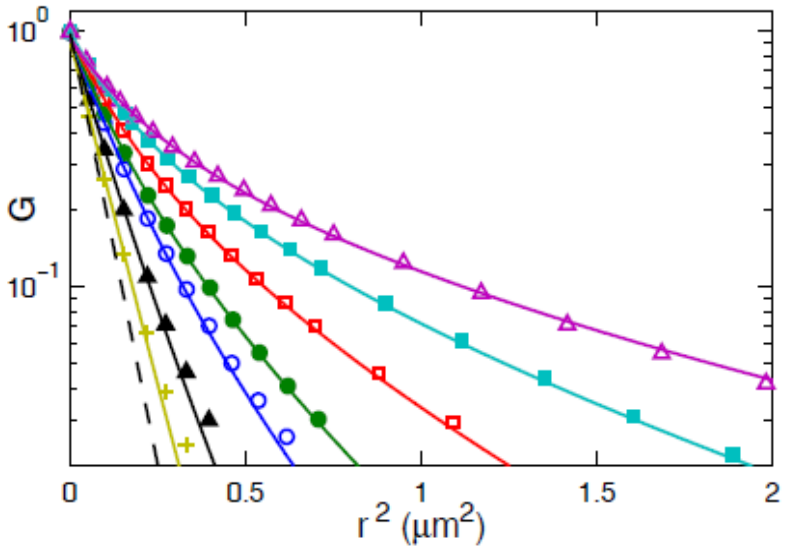


Fig. 6.1: Examples of measured SFCS correlation functions of dilute labelled DNA solutions normalized to unity. Symbols representing chain length are 0.64 (crosses), 2.1 (full triangles), 4.5 (open circles), 8.0 (full circles), 11.5 (open squares), 23.1 (full squares) and 46.2 μm (open triangles). Lines represent fits to the experimental data with Debye structure factor expression for ideal chain. The dashed line is SFCS correlation function of Rh6G as independent point sources for the calibration of the PSF of the optical set up according to Eq. 38 .

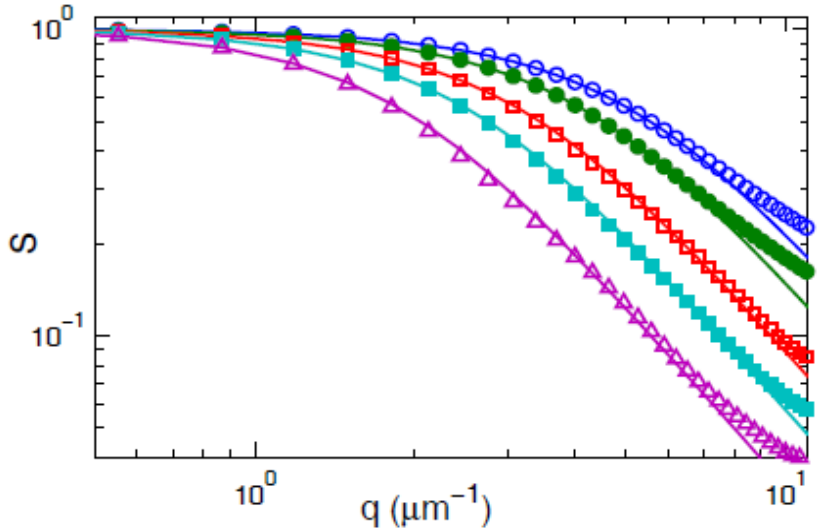


Fig. 6.2: Symbols representing measured structure factor $S(q)$ of dilute solutions of various DNA lengths: 4.5 (open circles), 8.0 (full circles), 11.5 (open squares), 23.1 (full squares), and 46.2 μm (open triangles). Lines are fits of measured data with Debye expression of structure factor for ideal chains Eq. 9.

Instead of fitting the correlation functions $G(r)$ directly, we can extract the solution structure factor $S(q)$ using the Eq. 31 and the procedure outlined in the Subsection 5.1.3. For DNA shorter than 4 μm this method does not give sufficient dynamic range of reliable measurements as the size of DNA coil is smaller or too close to the size of the sampling volume. The examples of structure factors $S(q)$ of DNA longer than 4 μm are presented in the Fig. 6.2. As expected, the dynamic range is larger for longer DNA chains. The extracted structure factors can be fit well with the Debye expression Eq. 9 for the ideal chain (Fig. 6.2).

In the Fig. 6.2. we can notice the upward trend in the extracted structure factors for $q > 9\mu m^{-1}$. This seems to be due to an artefact caused by the small deviations of the optical field from the Gaussian as measured by imaging of small beads.

Next, we aimed at confirming our findings of EtBr labelled DNA coil structure measurements with covalently labelled DNA. It was important since that intercalating dyes like EtBr are weakly fluorescent requiring high labelling density. This affects DNA contour length and may affect its persistence length. Besides, with covalently labelled DNA we can change buffer conditions at will, while the binding of EtBr strongly depends on electrolyte concentration.

We thus covalently tagged DNA molecules uniformly along their contour with strongly fluorescent Rh6G using PCR (Subsection 5.2.2). Three types of DNA were synthesized with lengths of 4.2, 8.5, 14.2 Kbp. We present our measurements on covalently labelled DNA molecules of three different lengths in Fig. 6.3. As well as our data on EtBr labelled DNA they can be fit well with Debye structure factors for ideal chains.

In the Fig. 6.4 we show the contour length L dependence of the gyration radii R_g of DNA coils as determined from the fits with Debye function to measured $S(q)$. The extension of DNA contour length of EtBr labelled samples was taken into account according to Ref. 23. There is an excellent agreement between the two data sets; one obtained from DNA labelled with intercalating dye EtBr and another from covalently attached Rh6G. The power law obtained on the dependence of gyration radius on contour length $R_g \propto L^{0.52\pm0.02}$ is close to the ideal chain exponent of ½ indicating no significant coil expansion. Furthermore we fit the data of $R_g(L)$ dependence with the expression for the ideal worm-like chain [34]:

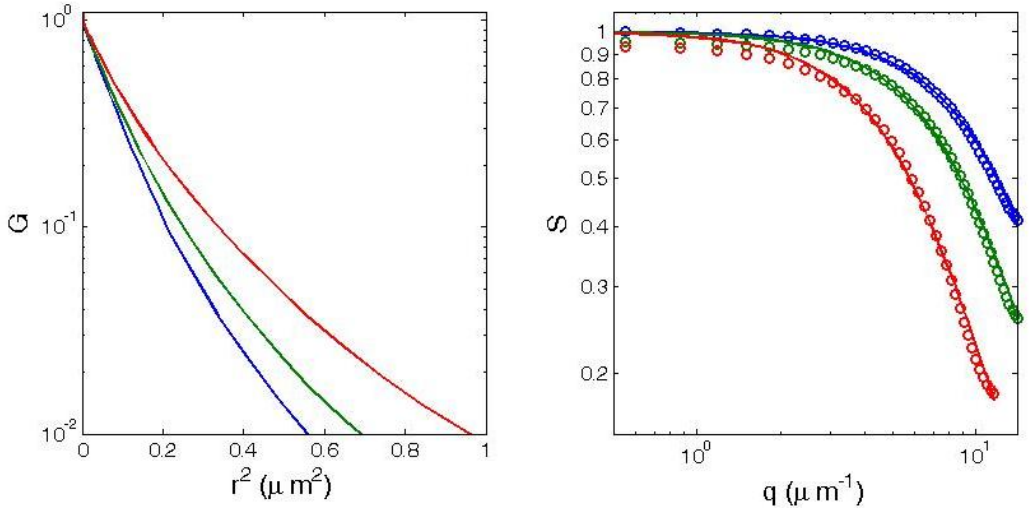


Fig. 6.3: SFCS measurements of correlation functions of covalently labelled DNA plotted against displacements (left panel). DNA lengths from left to right on the left panel: 4.2 Kbp, 8.5 Kbp and 14.2 Kbp. Extracted structure factors (circles) from SFCS measurements and corresponding fits with Debye expression (lines) on the right panel.

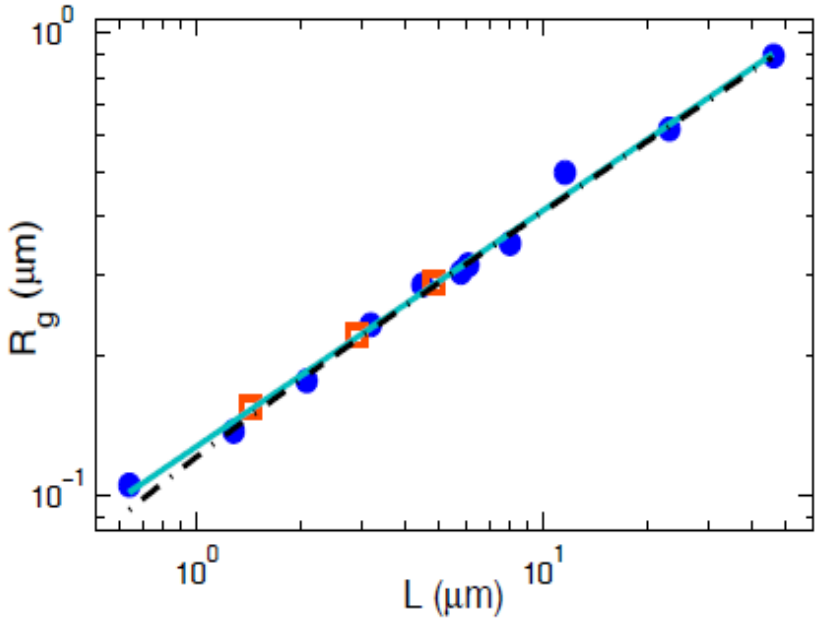


Fig. 6.4. The gyration radius R_g dependence on DNA contour length L. Data in circles and squares represent DNA labelled with EtBr and covalently attached Rh6G, respectively. Gyration radius values were extracted from the direct fits to SFCS correlation functions. Error bars are similar in size to symbols. The best power law fit $R_g \propto L^{0.52\pm0.02}$ is represented by solid line. Dashed line is fit with worm like chain expression giving $l_p = 51 \pm 1\ nm$. [33]

$$R_g^2(L) = \frac{Ll_p}{3} - l_p^2 + \frac{2l_p^3}{L} - \frac{2l_p^4}{L^2}\left(1 - \exp\left(-\frac{L}{l_p}\right)\right) \quad (39)$$

The fit is rather good (Fig. 6.4) and give the DNA persistence length of 51 ± 1 nm which agrees well with the known value of ~50 nm.

Thus our measurements are consistent with the theoretical prediction of ideal chain behavior for the DNA coils. The measured structure factor $S(q)$ shows good agreement with Debye expression for the ideal chains [31]. Furthermore the gyration radius R_g, obtained by fitting the measured data with the Debye expression, scales with the DNA contour length L in agreement with an ideal worm like chain formula.

6.2. DNA end-to-end distance distribution

With excluded volume interactions "turned on" the overall polymer structure undergoes expansion and change in the scaling laws. These are significant changes. However, the nature of the polymer end-to-end distribution changes yet more drastically. In an ideal chain the end-to-end distribution is Gaussian as in a random walk, and has its maximal probability density at zero end-to-end separation. In a *real* chain, the excluded volume repulsion between polymer ends through the clouds of monomers connected to them results in zero probability of finding the two ends of the chain close to each other. This is essentially a phase transition in the shape of the end-to-end distribution (pictured schematically in the Fig. 6.5)

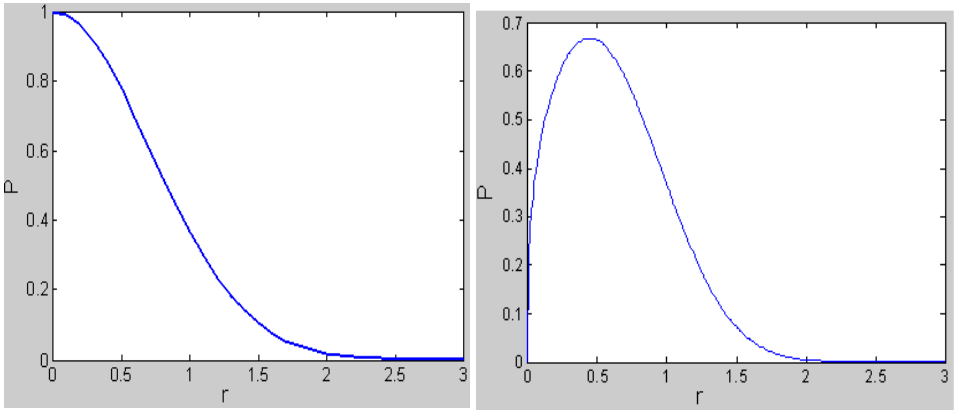


Fig. 6.5. Schematic representation of end-to-end distance distribution $P_{ee}(r)$ of an ideal (left) and a real (right) chain.

Therefore we wondered whether the end-to-end distribution $P_{ee}(t)$ is more sensitive to the excluded volume interactions than the structure factor $S(q)$, so that we might be able to

observe nonideality in DNA behaviour based on the measurements of $P_{ee}(t)$. We exploited our covalent DNA labelling technique and SFCS to measure DNA end-to-end distribution.

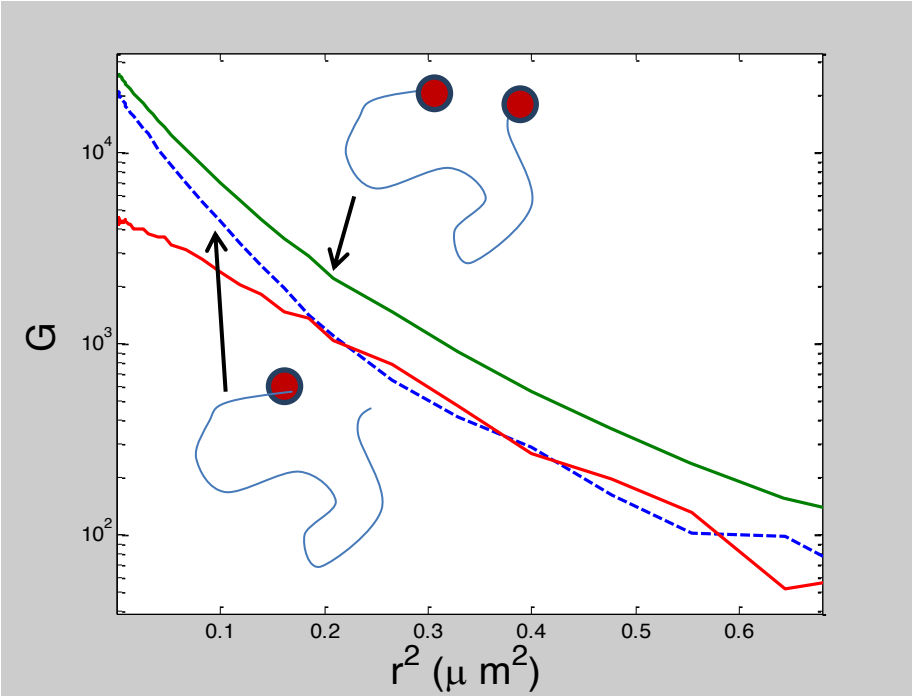


Fig. 6.6. An example of SFCS measurement of DNA end-to-end distribution (see Subsection 4.3.3). The green line is the measurement of SFCS correlation function $G(r)$ on DNA with both ends labeled. $G(r)$ contains contributions from cross correlation between the two labeled ends, but also auto correlation of each of the ends with itself. The pure autocorrelation contribution $G_1(r)$ can be measured on single labeled DNA samples (blue line). Then the difference $G_{12}(r) = G(r) - G_1(r)$ reflects solely the cross-correlation in the positions of the ends (red line). These results are for DNA of 2.8 Kbp.

The idea of the measurement as outlined in the Subsection 4.3.3 is to determine the cross correlation function $G_{12}(r)$ of the two ends by taking the difference of the correlation functions collected from double- and single- end labelled DNA. The Fourier transform of $G_{12}(r)$ then allows to determine the characteristic function of P_{ee} (see Eq. 34)

We performed SFCS measurements on both single- and double- end labeled DNA samples of several lengths. Fig. 6.6 shows an example of such a measurement for 2.8Kbp DNA. The resulting characteristic function $P_{ee}(q)$ is presented in the Fig. 6.7.

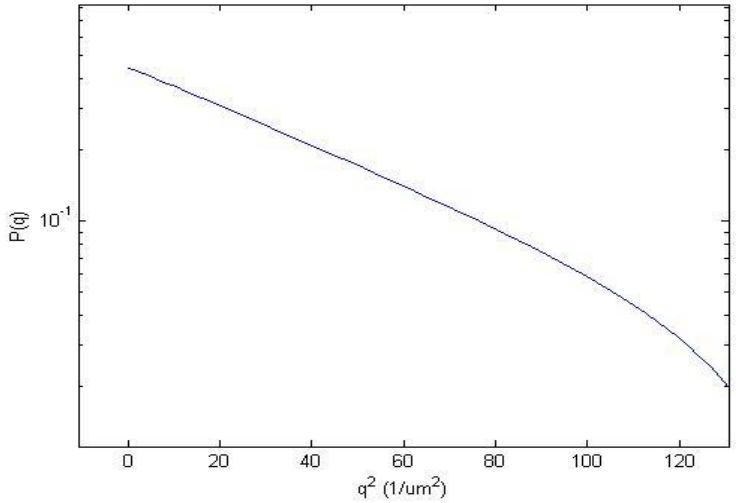


Fig. 6.7 The characteristic function $P_{ee}(q)$ of the end-to-end distribution of 2.8 kbp DNA molecule (red line in the Fig. 6.6).

In the Fig. 6.6 we can see that the two measured functions $G(r)$ and $G_1(r)$ are not Gaussian but their difference, the cross-correlation function $G_{12}(r)$, is Gaussian as evident from the linear nature of the function in the presented scale. Since the excitation-detection profile of our set up is close to Gaussian this means that the underlying end-to-end distribution is also Gaussian. Respectively, $P_{ee}(q)$ in the Fig. 6.7 is also Gaussian up to the very high wave vectors, where the measurement is limited by the optical resolution. This result is consistent with the theoretical prediction for ideal chains and also complements our findings from structure factor measurements using uniformly labeled DNA.

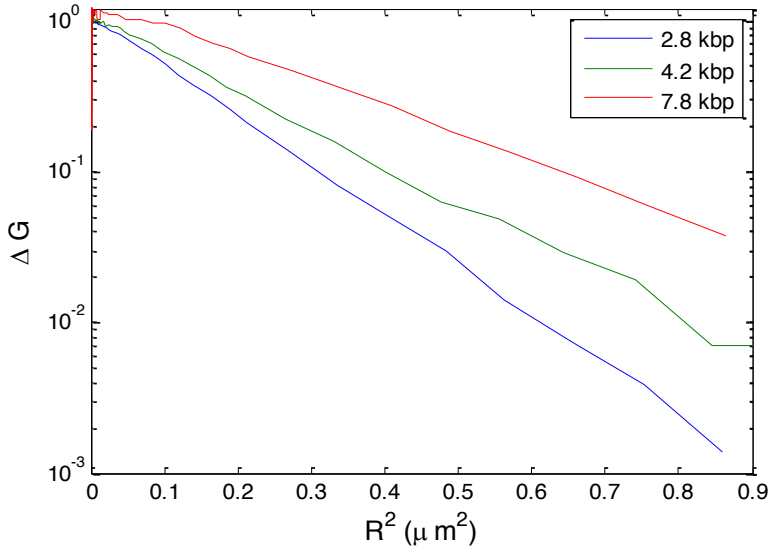


Fig. 6.8 Cross-correlations functions $G_{12}(r)$ in end positions for DNA of different lengths. DNA lengths, top to bottom: 7.8kbp (red), 4.2kbp (green), 2.8 Kbp (blue). The dependences are close to Gaussian reflecting the underlying Gaussian distribution of end-to-end distances.

Our measurements of cross-correlations in end positions on longer DNA molecules also gave similar results (Fig. 6.8): the obtained $G_{12}(r)$ are close to Gaussians, consistent with the ideal polymer behaviour of DNA. We can measure mean square end-to-end distance $\langle R^2 \rangle$ between DNA ends from the slope of the linear part of Fourier transform of the cross-correlations $P(q)$ since $P(q) \sim e^{-q^2 R^2/6}$. The measured $\langle R^2 \rangle$ values of 0.114, 0.15, and 0.27 μm^2 compare well to the theoretical estimate of 0.11, 0.14, and 0.28 μm^2 for 2.8, 4.2, 7.8 Kbp DNA respectively with DNA persistence length of ~50 nm.

6.3. Semi-dilute Solutions

In semi-dilute polymer solutions chains overlap and interpenetrate gradually as the concentration is increased. The interesting implications of such solutions from the theoretical perspective is mainly two fold; how the screening length depends on concentration and what happens to the structure of individual coil as the concentration is raised. The answer differs according to the nature of the polymer. As discussed in the Subsection 2.4.2 for the case of semi-flexible polymers like DNA the theory predicts a "marginal" regime with the screening length scaling with monomer concentration $\xi \sim c^{-1/2}$. Our lab previously reported such scaling using SFCS approach on DNA solution [20]. Similarly, the theory also predicts that the structure of each individual chain is largely unaffected by the presence of other chains [5, 35] unlike in the case of flexible polymers which contract in semi-dilute solutions.

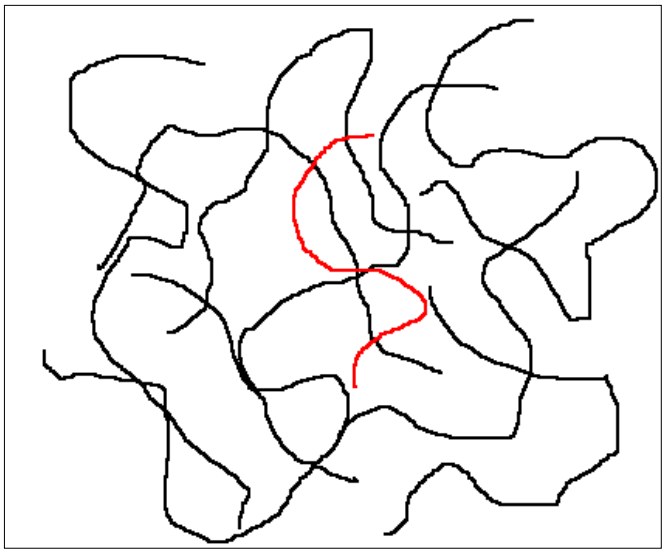


Fig. 6.9. Schematic representation of a labeled DNA chain (red) in the matrix of unlabeled chains (black).

Here, in combination with SFCS, we exploit our covalent DNA labelling technique to probe the structure of individual coil in semi-dilute solutions. We mix small amount of covalently labeled DNA chains with a large amount of unlabeled chains (Fig. 6.9). We keep the concentration of labeled chain fixed and we vary the concentration of unlabeled chain. There is one point in dilute regime where the concentration of both types of chains is low. The size of the labeled DNA is $4.8\mu m$ (14.2Kbp) and its concentration is kept low at $\sim 1-2\mu g/mL$ in order to avoid direct interactions among these chains. The size of the unlabeled DNA is $16.5\mu m$ (48.5 Kbp, Lambda DNA) and we vary its concentration from $\sim 1\mu g/mL$ (dilute regime) up to $1040\mu g/mL$. The semi-dilute regime for Lambda DNA begins at $c^*\sim 10\mu g/mL$,whereas c^* is about twice larger for the labeled chain of $4.8\mu m$. From the technical point of view our novel approach is similar to the method of neutron scattering applied to the flexible polymer system in which a small fraction of molecules is selectively deuterated. However, we would like to stress here that in the case of semi-flexible polymers it is only our technique that makes it possible to study structure of individual chains embedded in dense solutions.

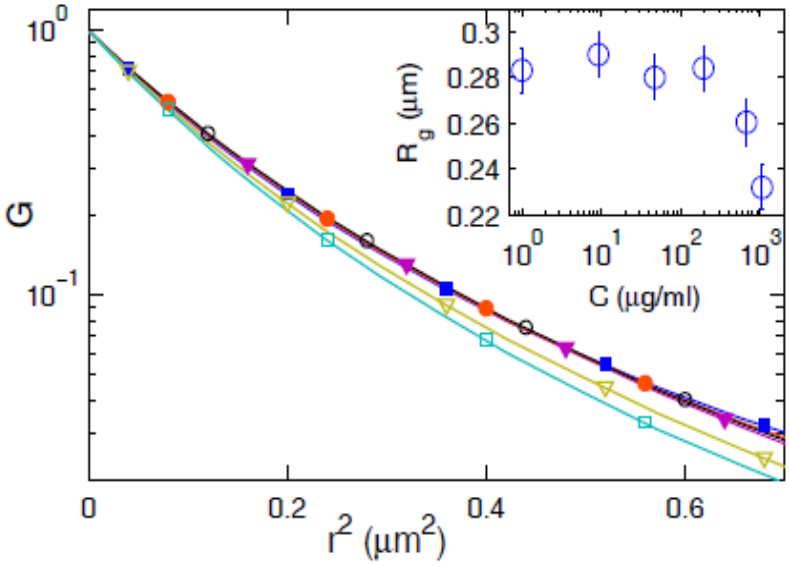


Fig. 6.10 Measured SFCS correlation functions normalized to unity of individual $4.8\mu m$ covalently labeled chains embedded in the matrix of unlabeled $16.5\mu m$ chains at different concentrations. Total DNA concentrations marked by different symbols are: 1 (full triangles), 9 (full squares), 47 (full circles), 190 (open circles), 670 (open triangles), and 1040 $\mu g/mL$ (open squares). To avoid clutter every data points are not symbolised. Inset shows plot of extracted gyration radii against concentration (33)

In Fig. 6.10, we present measured SFCS correlation functions of covalently labeled 4.8 μm DNA mixed with various concentrations of unlabeled lambda DNA ($16.5\mu m$). The

correlation functions represent the structure of only the labeled chains as the unlabeled chains are "invisible". We do not observe any changes in the chain conformation up to the concentration of $\sim 600 \mu g/mL$ consistent with theoretical predictions for marginal solutions [5, 40]. This is a significant finding since it is deep inside semi-dilute regime where it has been previously shown that the screening length changes by an order of magnitude from $\sim 1 \mu m$ to 100 nm [20]. Our measurements reveal an amazing feature of semi-dilute semi-flexible polymer solutions where small rearrangements in the structure of individual coils can bring about drastic changes in the collective behaviour i.e. screening of spatial correlations. This is opposite to what is observed for flexible chains in semi-dilute solutions which significantly change conformations and contract.

There is a weak contraction of DNA chain with increasing DNA concentrations revealed by the SFCS correlation functions at very high concentration. Quantitatively we noticed $\sim 18\%$ contraction over 5.5 fold increase in DNA concentration from 190 to $1040 \mu g/mL$. In this concentration range the solution enters a new 'concentrated' regime in which DNA persistence length and screening length are comparable and the marginal theory neither predicts change in screening length nor the gyration radius [5, 40]. This means the observed coil contraction at the highest probed concentrations is at odds with the theoretical predictions. Further experiments are necessary to understand the observed discrepancy between the theoretical predictions and experiment at the crossover to the 'concentrated' regime.

7. Structure of DNA polyelectrolyte solutions

DNA, which is strongly charged along its phosphate backbone, is in fact an example of a polyelectrolyte (polymer containing ionizable monomers which dissociate in solvent giving polyelectrolyte plus counter-ions). The implication of long range Coulomb interactions, due to charged monomers, is to affect coil conformation and effective monomer size. However the electrostatic interactions between monomers can be almost completely screened with high salt concentration allowing one to treat such solutions as polymer solution instead of polyelectrolyte solution. In fact all our measurements in the previous sections were of DNA polymer solutions where high salt concentration in the buffer medium screens the electrostatic interactions completely along the DNA chain. Our ability to covalently attach fluorophore to DNA molecules allows us to vary salt concentration in the buffer solvent, something that could not be done with intercalating dyes since binding of such dyes to DNA depends on salt concentration in the solvent. In this section we present our work on the structure of DNA polyelectrolyte solutions attempting to address the controversial issue of electrostatic persistence length dependence on salt concentration.

7.1. Background and Motivation

In the context of this experimental work we focus on one particular aspect of stiff polyelectrolyte chains i.e. electrostatic persistence length and its dependence on ionic strength. Apart from bulk properties such as viscosity and osmotic pressure polyelectrolytes in general are not as well understood [41] as neutral polymer solutions [2]. The main difficulties in proper understanding of polyelectrolytes are due to the presence of charges on the chain and accompanying counter-ions. The long range nature of Coulomb interactions due to the charges along the chain provides a new length scale (Debye length, $r_d = \kappa^{-1}$) to the chain. The implications of such interactions are two-fold; locally the chain stiffness increases whereas globally there is an effect similar to an excluded volume between chain segments.

7.1.1. Counterion condensation

A peculiar feature of polyelectrolyte solution system is the phenomenon of counterion (or Manning) condensation [42, 43]. Condensation of counterions on polyelectrolyte is brought about by the electrostatic attractions between polyelectrolyte chains and counterions in the

solution. However, such attractions and resultant condensation is at the expense of the loss of translational entropy by counterions due to their localization near the polyelectrolyte chains.

We can analyze a simple charged rod to arrive at the condition necessary for the counterion condensation (Fig. 7.1). We simply need to find mean equilibrium distance between counterions and the charged rod. For this we need to know the free energy change as we attempt to move from a certain distance r_1 from the rod to r_2 where $r_2 > r_1$ in a cylindrical geometry. The entropic gain of the translational motion from $r_1$to r_2 can be written as:

$$\Delta F_1 \sim k_B T \Delta S \sim k_B T ln(V_2/V_1) \sim k_B T ln(r_2/r_1)$$

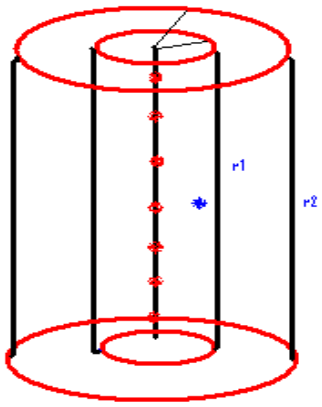


Fig. 7.1 Schematic representation of a charged rod and a counterion at distance r_1.

The resultant decrease in attraction energy as we move from $r_1$to r_2:

$$\Delta F_2 \sim -e\Delta\phi \sim -e\frac{\rho}{\varepsilon} ln(r_2/r_1) \sim -\frac{e^2}{\varepsilon a} ln(r_2/r_1),$$

where a is distance between charges along the cylinder axis and $\rho = e/a$ is the linear charge density.

We can define now a dimensionless counterion condensation parameter as $u \equiv \frac{e^2}{\varepsilon a k_B T}$. Since both forms of free energy are proportional to ln (r_2/r_1), we have two possibilities: u is either less than or greater than 1. Obviously if $u < 1$ then, $\Delta F_1 > |\Delta F_2|$ and entropy dominates and there will be no counterion condensation. However, if $u > 1$ then $\Delta F_1 < |\Delta F_2|$ and there will be counterion condensation. Furthermore , the term $\frac{e^2}{\varepsilon k_B T} = l_B$ is Bjerrum length which is 7.14 $A°$ in water at 25℃. For DNA, since the distance between charges a is 1.7 °, and counterion condensation is expected as $u = 4.2$. The counterions will bind DNA

till the effective linear charge is equivalent to one electron charge per Bjerrum lengths, i.e. $u = 1$.

7.1.2. Debye-Huckel screening

Counterion condensation phenomenon is associated with electrostatic screening, which is also known as Debye-Huckel screening, of interactions between charges in the polyelectrolyte. The Debye length scale ($r_d = \kappa^{-1}$) is introduced in the polyelectrolyte system through the screening effect. Here we briefly describe the phenomenon.

The screening phenomenon starts with the Poisson equation for the electrostatic potential φ in a neutral ionic solution ($\sum_i q_i n_i = 0$) with small ionic concentrations n_i and charges $q_i = z_i e$:

$$\nabla^2 \varphi = -\frac{\rho}{\epsilon_r \epsilon_0}$$

Ions are in thermal equilibrium and free to move so they obey Boltzmann statistics:

$$n_i = n_i^0 \exp\left(-\frac{z_i e \varphi}{k_B T}\right)$$

Since charge density $\rho = \sum_i z_i e n_i$, we can combine the above two equations to get non-linear Poisson-Boltzmann (PB) equation for the electrostatic potential:

$$\nabla^2 \varphi = -\sum_i \frac{z_i e n_i^0}{\epsilon_r \epsilon_0} e^{-\frac{z_i e \varphi}{k_B T}}$$

The non-linear PB equation can be solved only numerically. However, for weak electrostatic potential $\left(\frac{z_i e \varphi}{k_B T} < 1\right)$ the equation can be linearized by expanding the exponents. After expansion the first term is reduced to zero due to charge neutrality and we are left with a linear equation for φ:

$$\nabla^2 \varphi = \frac{1}{r_d^2} \varphi$$

where $r_d = \sqrt{\frac{\epsilon_r \epsilon_0 k_b T}{\Sigma z_i^2 e^2 n_i^0}}$ is the Debye length.

In e.g spherically symmetric case the solution of the linearized PB equation above is:

$$\varphi(r) = \frac{A}{r} \exp(-r/r_d) \quad .$$

In a polyelectrolyte system with finite salt concentration the electrostatic interactions between charged monomers are screened by the ions and the interaction strength declines exponentially with the distance. At distances much smaller than Debye length, there are still interactions among charges through unscreened coulomb potential. If the polyelectrolyte chain size is much smaller than the Debye length then there is no effect of salt on the chain. On the other hand when salt concentration is high or Debye length is much smaller than chain size then the polyelectrolyte chain behaves like a neutral polymer in good solvent. The main implication for chains like DNA at intermediate salt concentration is the additional chain stiffening induced by electrostatic interactions or salt concentration dependent electrostatic persistence length l_e.

7.1.3. Electrostatic persistence length

In the context of this thesis we limit our discussion to the conformational changes and consequent persistence length change in DNA as we reduce salt concentration. The double helical DNA chain is locally stiff but also has long range flexibility and is well described by worm like chain (WLC) model [34]. The model predicts size of the chain based on just two parameters, the contour length and the length over which correlation between segment directions persists. This local stiffness or the persistence length of DNA is generally considered as useful concept in terms of understanding the rigidity of the molecule and its various biological implications such as nucleosomal organization, DNA packaging in capsids, transcription and condensation. It is worthwhile to mention that persistence length is independent of contour length and chains smaller than persistence length are considered rigid rods. In high salt limit (>0.1M NaCl) when electrostatic interactions are almost totally screened, DNA persistence length l_p~50 nm has been reported by various experiments [44]. The dependence of DNA persistence length on electrolyte concentration is a controversial issue in the literature. The classical theory by Odijk [45] and Skolnick and Fixman [46] (OSF) predict reciprocal dependence on molar concentration of salt for the electrostatic part of the persistence length l_e to which a fixed non-electrostatic or "neutral" part l_0 is additive and is independent of salt concentration. For polyelectrolyte in buffers containing 1:1 salt such as NaCl, OSF prediction reads:

$$l_p = l_0 + l_e = l_0 + \frac{u^2}{4\kappa^2 l_b} = l_0 + \frac{3.24 \cdot 10^{-2}}{I}\ nm \tag{40}$$

For strongly charged polyelectrolyte like DNA for which $u > 1$, the counterion condensation on the chain reduce the charge density and make it $u = 1$ [43], so that

$$l_p = l_0 + l_e = l_0 + \frac{1}{4\kappa^2 l_B} = l_0 + \frac{3.24 \cdot 10^{-2}}{I}\ nm. \quad (41)$$

One of the main predictions of the OSF theory is that $l_e \sim \kappa^{-2} \sim I^{-1}$. Experimental investigation verifying OSF prediction of quadratic dependence of electrostatic persistence length of DNA on Debye length has been reported e.g. in [47]. Other theories predict $l_e \sim \kappa^{-1}$ dependence [48].

In this part of the thesis, however, we will only concern a challenge of OSF by Manning [49] who concluded that the persistence length is affected only by tension force imposed along the DNA backbone by the electrostatic repulsion between charges, that is not accounted for in the OSF theories. According to this theory the electrostatic and neutral contributions to the persistence length are not additive but instead multiplicative and the overall prediction for the dependence of DNA persistence length on electrolyte concentration is given as:

$$l_p = A\left[\left(\frac{2l_B}{a} - 1\right)\frac{\kappa a e^{-\kappa a}}{1 - e^{-\kappa a}} - 1 - ln(1 - e^{-\kappa a})\right], \quad (42)$$

where the parameter A depends on the rigidity of the "neutral" backbone and $a \sim 0.17$ nm is the distance between charges on DNA. For monovalent electrolyte concentrations below 100mM , the above equation can be linearized since $\kappa a \ll 1$. In this electrolyte concentration range, the dominant concentration dependence of persistence length in the Manning theory stems from $\log(\kappa a)$ term, i.e. the dependence of the l_p on salt concentration is logarithmic. Manning [49] cites a number of experimental results described in the literature as an evidence to support his theory and criticizes the results of Ref. 47, obtained in single molecule DNA stretching experiments, that support OSF theory.

We thus tried to resolve this controversy using our SFCS method to measure DNA persistence length for different electrolyte concentrations.

7.2. Experimental findings and discussions

We make use of our SFCS approach and of covalent DNA labeling to address the issue of persistence length dependence on salt concentration. We performed measurements of the structure of 4.2 Kbp DNA in dilute solutions at various NaCl concentrations. These

molecules are much larger than the persistence length l_p and hence can be considered as polymers while still short enough for the excluded volume interactions between their segments to be insignificant.

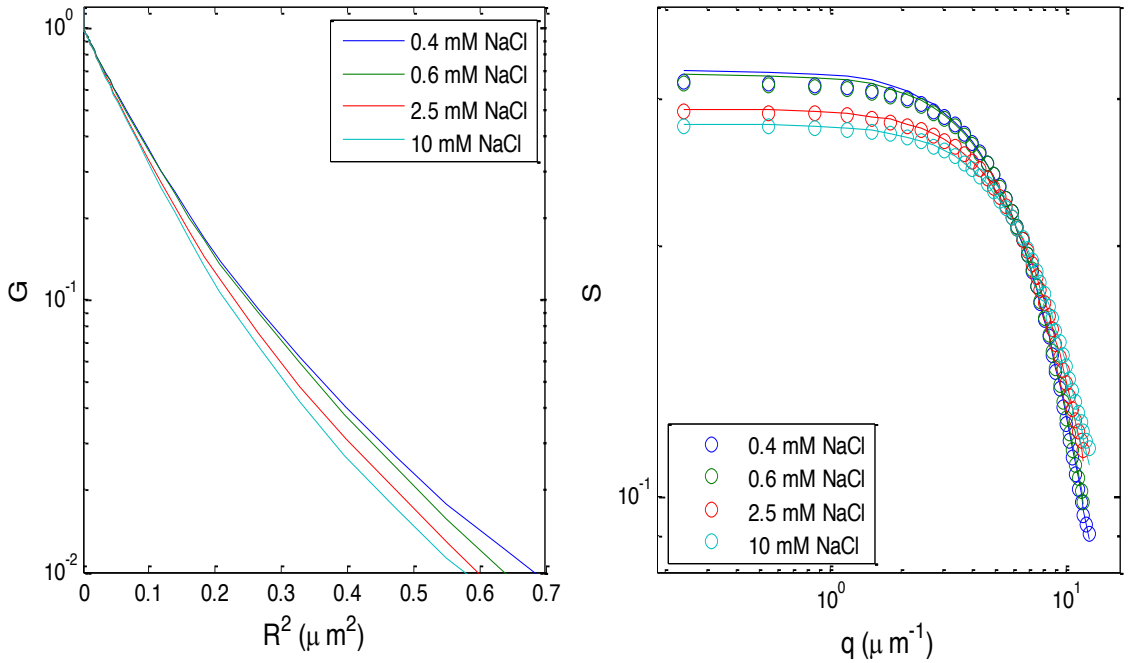


Fig. 7.2 SFCS measurements of structure of covalently labeled 4.2 Kbp DNA solutions at various salt concentrations. In the left panel, SFCS correlation functions are shown vs. displacement at various salt concentrations (left to right): 10, 2.5, 0.6 and 0.4 mM. Extracted structure factors (circles) from SFCS measurements and corresponding fits with Debye function (lines) are shown in the right panel.

Since Manning theory predicts rather weak log (c) dependence of l_p in a wide range of concentrations and also for OSF theory l_p converges to a constant l_0 at C > 0.01M, it is very difficult to distinguish between the two theories in the concentration range of C > 0.01M. Thus the measurements were performed in the low concentration C < 0.01M range which has only been done in a few experimental studies before.

Measured SFCS correlation functions G and structure factors extracted from those measurements at different electrolyte concentrations are presented in Fig. 7.2. As salt concentration decreases, the chain stiffens and the coils expand resulting in slower decay of the correlation functions. The structure factors are fitted with the Debye expression for the structure of ideal chain with chain gyration radius R_g as an only parameter. The worm-like chain expression Eq. 39 is then solved for l_p using the R_g from Debye fit and known contour length of 4.2 Kbp DNA in order to extract the persistence length. Good fits by the Debye function support the assumption that the chains behave as ideal even in low salt conditions.

In Fig. 7.3 we plot the extracted persistence length against the electrolyte concentration together with fits to OSF (Eq. 41) and Manning's (Eq. 42) theoretical predictions. In both fits a single and same parameter is used i.e. electrolyte independent or "neutral" persistence length l_0 in OSF theory and in the case of Manning's theory the amplitude A which is determined by the "neutral" persistence length. It seems from the Fig. 7.3 that Manning's prediction of $\log(c)$ dependence of persistence length is supported by our data. However, Manning's theory lacks enough parameters to fit the data numerically. On the other hand our data can be fit reasonably well to the OSF theory with a single fitting parameter $l_0 = 50$ nm.

We mention here that in Ref. 47, the fit of their data to classical OSF theory is erroneous as pointed out by Manning [49]. The OSF line drawn in Ref. 47 through the data is off by factor of ~ 2.5 compared to the real OSF theory.

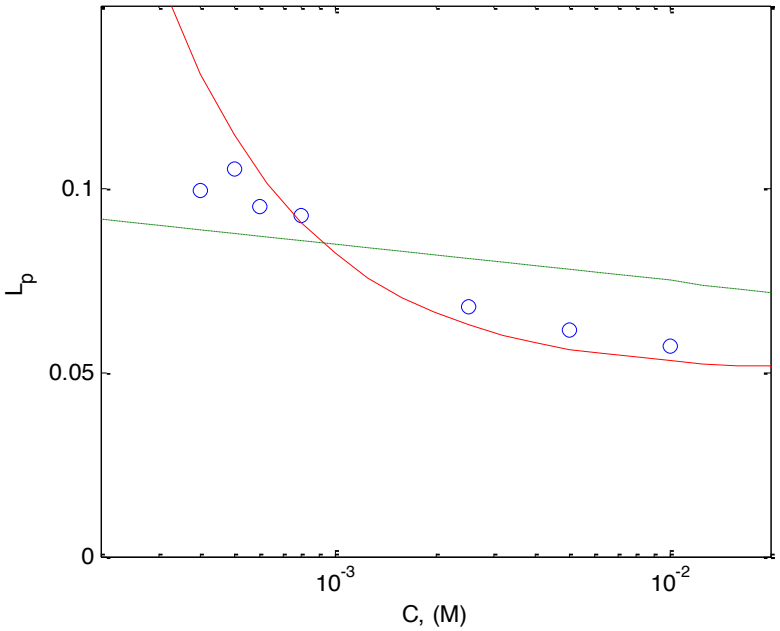


Fig. 7.3 DNA persistence length dependence on salt (NaCl) concentration as measured by SFCS. Circles are the experimental data whereas solid line and dashed line are fit with OSF and Manning's theory respectively.

Although we obtained data consistent with OSF prediction in the concentration range depicted in Fig. 7.3, we also noticed significant discrepancy between data and theory at electrolyte concentrations below 1 mM. Initially we attributed this to uncertainty about the residual salt concentration left in the sample after its preparation. However, the measurements performed on the different dilutions of the sample with pure water showed that residual salts do not pose problems. We note also that there are always free ions in water coming from

dissolved CO_2 (which is a weak acid). However, our conductivity measurements on similarly prepared samples and controls showed the background electrolyte concentrations of $\sim 70\mu M$. This is too small concentration to explain the discrepancy between theory and experiment at low salt concentrations. In any case, we add this value to the concentrations in our plots.

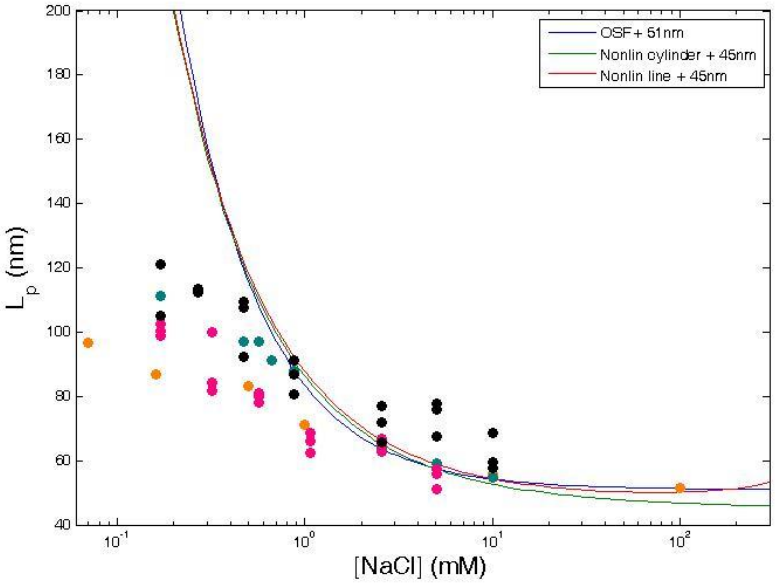


Fig. 7.4. DNA Persistence length dependence on salt concentration measured by SFCS. Different colors represent samples prepared and measured on different days. Lines drawn represent OSF (blue), non-linear with cylindrical charge distribution (green) and non-linear line charge distribution (red). The theoretical curves are courtesy of Prof. Yoram Burak (HUJI).

We performed several further measurements at salt concentration extending well below 1 mM to check if the deviation from OSF theory in our measurements was more notable at yet low concentrations. We are not aware of any data of electrostatic persistence length measurements below 2 mM . Manning [49] compiles measurements of persistence length by several groups; however, none of the measurements are for below 2 mM. We prepared our samples in pure water and added salts without any buffers since the presence of the buffer would not allow us to go below 1 mM. We checked if sample prepared in buffer would be any different from sample prepared in pure water at a given concentration of salt and found no difference in terms of SFCS structure factor measurements. We present a compilation of all our data sets which includes measurements below 1 mM in the Fig. 7.4. Different sets of samples prepared and measured on different days are represented by different colours. While the spread in the data for the same concentration values is significant, it is similar or better than measured by other methods [47, 49]. Besides, within each set of our data the consistency is rather good.

We compared our data sets also to the theoretical estimates of the persistence length dependence on salt arising from the solutions of the non-linear Poisson-Boltzmann kindly provided to us by Prof. Yoram Burak. In Burak's calculations two forms of linearized Debye-Huckel equations are used. In one form, the effective charge is assumed spread around a cylindrical rim around polyelectrolyte chain without salt penetrating the rim whereas in the other form line charge distribution along the polyelectrolyte axis replaces the cylinder and salt is assumed present everywhere in the vicinity. As can be seen from the Fig. 7.4, at the lower end of the salt concentration the measured data deviates significantly from all three theories.

It might be argued that the deviations of our data from OSF prediction at low salt concentration could be due to DNA strand separation as a result of unscreened electrostatic interactions. So we measured the DNA melting curves using UV absorption by DNA at 260nm in pure water (i.e. $\sim 70\mu M$ of electrolyte due to CO_2). DNA in a conventional buffer and physiological electrolyte strength melts at $\sim$ 98℃ and in spectrophotometric measurements the optical density of single stranded DNA is ~ 1.5 times than that of double stranded DNA.

Our measurements (Fig. 7.5) show that melting occurs around 60℃ when no salts are present in solution, so no melting or strand separation should occur near the room temperature where our SFCS measurements were performed at $\sim$ 18℃.

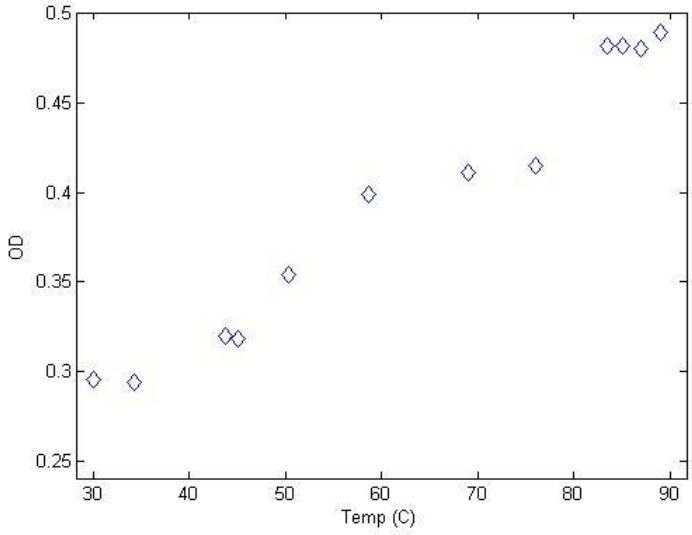


Fig. 7.5 Spectrophotometer measurement of optical density of 4.3 Kbp DNA sample at low electrolyte strength.

It is not clear at this point what causes the discrepancy between OSF theory and our results at the lowest salt concentrations. Possibly, small DNA bubbles locally open at low salts to form stranded DNA bubbles, decreasing the overall persistence length.

8. Polymer Dynamics in semi-dilute solutions

SFCS can also be exploited to measure kinetics of motion of labeled DNA segment within large polymer chain. In the past FCS has been utilised in our lab to measure dynamics of both single and double stranded DNA chains in dilute regime [50]. In this study, freely diffusing long DNA molecules with single fluorescent label attached to a single end base were monitored with FCS and the time dependence of mean square displacement (MSD) of monomers were extracted using the Eq. 28.

However, this simple FCS approach does not work well for slowly diffusing fluorescent molecules such as DNA in semi-dilute or concentrated solutions. When the characteristic diffusion time through the sampling volume approaches $1s$, the fluorophores photobleach with high probability while inside the sampling volume. This totally changes the character of the correlation function so that even the data at the short time scale cannot be trusted: the photobleaching mostly occurs in the center of the sampling volume and this changes its characteristics, leading to higher relative contribution to the signal of the molecules at the periphery.

Thus we devised a variation of the scanning FCS to measure molecular dynamics. Within this approach we measure the FCS correlation function while the sample is scanned at multiple different speeds, so we call this method multiple speed scanning FCS and abbreviate as mSFCS. As the molecules in scanning mode spend little time in the sampling volume as they are actively moved from there by the setup, we can avoid photobleaching problem even in the samples with very slow dynamics. In this section we present our successful application of SFCS in measuring end dynamics of long DNA polymer in dilute, semi-dilute and concentrated solutions.

In our approach of single end labeled long DNA sample preparation, small DNA fragments ($<$ Kuhn length, $\sim$ 280 bp) with incorporated modified nucleotides are synthesized first with PCR and then reacted with fluorescent dyes (Rh6G). These labeled DNA fragments are then ligated to long DNA chains, which is followed by gel electrophoresis to separate unlabeled and unwanted DNA fragments from the desired sample. In order to make sure single end labelling, we create matching restriction overhang on only one end of the long chains. The samples were measured in Tris buffer and 100 mM NaCl electrolyte strength. For

the measurements on semi-dilute and concentrated regimes we mix samples with unlabeled Lambda DNA at various concentrations.

The idea of the mSFCS method is that we scan the sample sequentially at different speeds $\{V_i\}$ from very slow $\sim 1\mu m/s$ to very high $\sim 4000\,\mu m/s$, typically with increase in the speed of 10% in between different runs. As a result, from each sample we obtain a set of correlation curves (Fig. 8.1, left panel), with a shift of the curve to the shorter times with higher speeds. Focusing on any particular time point, taking the values from the curves scanned at different speeds for the same delay time (inset in the left panel Fig. 8.1), we obtain the set of correlation values $G_v(r_i, t)$, where $r_i = V_i t$. All these values relate to the same time lag, but to different displacements, i.e. they show how far a concentration fluctuation propagated in space over this particular time *t*. Quantitatively, in the Subsection 4.3.4 we derived the Eq. 36 that holds for the measured correlation function in these conditions:

$$G_V = G_0(t)\exp\left[-\frac{(Vt)^2}{w_{xy}^2\left(1+\frac{2\langle\Delta r^2(\mathrm{t})\rangle}{3w_{xy}^2}\right)}\right],$$

Thus fitting $G_V(r,t)$ with a Gaussian as a function of displacement $r = Vt$ for any particular time *t* gives $w_{xy}^2\left(1+\frac{2\langle\Delta r^2(\mathrm{t})\rangle}{3w_{xy}^2}\right)$. Provided that w_{xy} is calibrated, we can calculate from here the segmental MSD $\langle\Delta r^2(\mathrm{t})\rangle$ for this particular time. Then performing this procedure for every available time point, we recover the whole temporal dependence $\langle\Delta r^2(\mathrm{t})\rangle$.

In dilute solutions, this method produces results consistent with our previous approach of extracting the monomer MSD from static FCS data (Fig. 8.1, blue symbols and line in the right panel). Its real strength, however, is in dense and crowded solutions, where the dynamics are so slow that in the usual FCS approach the sample bleaches under the spot and the resulting data are not reliable. Such an application in the dense regime is also shown in Fig. 8.1. For the highest DNA concentration tried ($3500\,\mu g/mL \sim 150c^*$) the difference between the results of static FCS and mSFCS is large (green circles and a line). Notice that the static FCS approach does not give reasonable results here, exhibiting almost the same kinetics as a dilute sample. However, mSFCS shows significant slowing down of the kinetics, as expected.

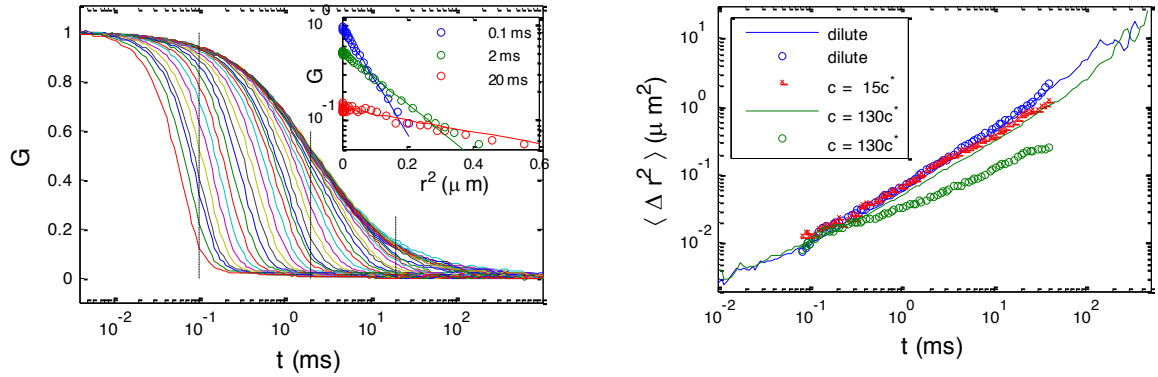


Fig. 8.1. Example of the application of mSFCS technique to the dynamics of single labelled 9kbp DNA. Left panel: the results of the multiple scans at different speeds are analyzed at different points, inset: extraction of MSD for three time points. Right panel: the comparison of static FCS (blue line) and mSFCS data (blue circles) in dilute regime and deep in the semidilute regime (green line and data, c = 130c*). Notice the discrepancy between mSFCS and FCS results in the semidilute regime. Red crosses: mSFCS data at 15 overlap concentration do not exhibit any obvious slowing down.

Remarkably, for moderate DNA concentrations of $\sim 415 \mu g/mL$ that are still deep inside semi-dilute regime ($\sim 15 c^*$), we do not see any significant change in the dynamical properties of DNA ends. Compared to the dilute solution ($\sim 2 \mu g/mL$) measurements, the dynamics at a concentration of $\sim 415 \mu g/mL$ is not even distinguishable except at long time lags ($t > 5$ ms). This means that even though chains overlap and interpenetrate significantly, they are not entangled, apparently, for the same reason that excluded volume effects in DNA are weak: because of high DNA stiffness there is ample free volume within each chain, so that the segmental dynamics is essentially uninhibited.

9. Conclusions

A new approach combining covalent DNA labelling technique and Scanning Fluorescence Correlation Spectroscopy (SFCS) has been developed to measure various statistical and dynamical properties of DNA molecules, such as solution structure factor, DNA end-to-end distance distribution and the temporal dependence of segmental MSD. New protocols, necessary for the appropriate sample preparation, such as DNA labelling, electrophoretic separation of long DNA and subsequent extraction from the agarose gel were implemented.

Measurement of the structure of individual DNA chains over two decades in length (0.6 to 46 μm) show that DNA follows ideal chain behaviour with no internal segmental collisions. This was corroborated by both measurements of DNA structure factor an of DNA end-to-end distance distribution.

Measurements of the structure of individual DNA chains in semi-dilute regime reveal no effect of the dense matrix on the conformations of individual chains. Interestingly, at moderate concentrations in semi-dilute regime, even dynamical properties of DNA segments are unaffected by the presence of dense matrix of DNA chains. Our data in dilute and in semidilute solutions support the findings of Schaefer, Joanny and Pincus theory [5].

Furthermore the effect of ionic strength on DNA persistence length was investigated. Overall we get a good agreement with Odijk-Skolnik-Fixman theory [45, 46]. At very low electrolyte concentrations (< 1mM NaCl), the experimental results deviate from the theoretical predictions. Further research is needed to understand this discrepancy.

We have published part of the presented data in [33].

Finally, we note that many of our findings on DNA structure are supported by a recent work on DNA brushes by Bar-Ziv group at the Weizmann institute [53].

10. Bibliography

1) Edwards SF. "Theory of Polymer Solutions at Intermediate Concentration". Proceedings of the Physical Society of London (1966), 88:265.
2) deGennes PG. Scaling Concepts in Polymer Physics: Cornell University Press, 1979.
3) Daoud M, Cotton JP, Farnoux B, Jannink G, Sarma G, Benoit H, Duplessix R,Picot C, Gennes PGD. "Solutions of Flexible Polymers - Neutron Experiments and Interpretation". Macromolecules (1975), 8:804-818.
4) J. F. Marko and E. D. Siggia, Phys. Rev. E 52, 2912 (1995).
5) D. W. Schaefer, J. F. Joanny, and P. Pincus, Macromolecules 13, 1280 (1980).
6) D. Magde, E. L. Elson, and W.W Webb. Thermodynamic fluctuations in a reacting system - measurement by fluorescence correlation spectroscopy.Phys. Rev. Lett., 29(11):705-708, 1972.
7) E. L. Elson and D. Magde. Fluorescence correlation spectroscopy. i. conceptual basis and theory. Biopolymers, 13(1):1-27, 1974.
8) D. Magde, E. L. Elson, and W. W. Webb. Fluorescence correlation spectroscopy.i. an experimental realization. Biopolymers, 13(1):29-61, 1974.
9) R. Rigler, U. Mets, J. Widengren, and P. Kask. Fluorescence correlation spectroscopy with high count rate and low background: analysis of translational diffusion. Europ. J. Biophys., 22:169-175, 1993.
10) R. Kubo. Fluctuation-dissipation theorem. Reports on Progress in Physics, 29:255–284, 1966.
11) N. L. Thompson. Fluorescence correlation spectroscopy. In J. R. Lakowicz, editor, Topics in Fluorescence Spectroscopy, volume 1. Plenum Press, New York, 1991.
12) O. Krichevsky and G. Bonnet. Fluorescence correlation spectroscopy: the technique and its applications. Rep. Prog. Phys., 65(2):251297, 2002.

13) M. Weissman, H. Schindler, and G. Feher, Proc. Natl. Acad. Sci. USA 73, 2776, 1976.
14) D. Magde, W. W. Webb, and E. L. Elson, Biopolymers 17, 361, 1978.
15) N. O. Petersen, D. C. Johnson, and M. J. Schlesinger, Biophys. J. 49, 817, 1986.
16) D. E. Koppel, F. Morgan, A. E. Cowan, and J. H. Carson, Biophys. J. 66, 502, 1994.
17) M. A. Digman, P. Sengupta, P. W. Wiseman, C. M. Brown, A. R. Horwitz, and E. Gratton, Biophys. J. 88, L33, 2005.
18) Z. Petrasek and P. Schwille, Biophys. J. 94, 1437, 2008.
19) J. P. Skinner, Y. Chen, and J. D. Mäuller, Biophys. J. 89,1288 ,2005.
20) E. Shafran, A. Yaniv and O. Krichevsky, *Phys*. *Rev*. *Lett*. 104, 128101, 2010.
21) B. J. Bern and R. Pecora, *Dynamic Light Scattering* (Wiley, New York, 1976).
22) T. Romantsov, I. Fishov, and O. Krichevsky, Biophys. J. 92, 2875, 2007.
23) M. S. Rocha, M. C. Ferreira, and O. N. Mesquita, J. Chem. Phys. 127, 105108, 2007.
24) G. F. Carle, M. frank, M. V. Olson, *Science* 232, 65, 1986.
25) R. M. Robertson, S. Laib, and D. E. Smith, Proc. Natl. Acad. Sci. U S A 103, 7310, 2006.
26) R. Lehner, J. Koota, G. Maret, and T. Gisler, *Phys*. *Rev*. *Lett*. 96, 107801, 2006.
27) F. Valle, M. Favre, P. De Los Rios, A. Rosa, and G. Di-etler, *Phys*. *Rev*. *Lett*. 95, 158105, 2005.
28) P. Sharp and V. A. Bloomfield, *Biopolymers* 6, 1201, 1968.
29) J. Harpst and J. Dawson, *Biophysical J*. 55, 1237, 1989.
30) J. A. Harpst, *Biophys Chem* 11, 295, 1980.
31) P. Debye, J. Phys. Colloid. Chem. 51, 18, 1947.
32) S. B. Smith, Y. Cui, and C. Bustamante, Science 271, 795, 1996.
33) M. Nepal, A. Yaniv, E. Shafran, O. Krichevsky, Phys. Rev. Lett. 110, 058102, 2013.
34) H. Benoit and P. Doty, J. Phys. Chem. 57, 958, 1953.

35) R. Verma, J. C. Crocker, T. C. Lubensky, and A. G. Yodh, Macromol. 33, 177, 2000.

36) W. Brown and T. Nicolai, Colloid & Polymer Science 268, 977, 1990.

37) P. Wiltzius, H. R. Haller, D. S. Cannell, and D. W. Schaefer, Phys. Rev. Lett. 51, 1183, 1983.

38) A. Bennett, P. Daivis, R. Shanks, and R. Knott, Polymer 45, 8531, 2004.

39) T. Uematsu, C. Svanberg, M. Nyden, and P. Jacobsson, Phys. Rev. E 68, 051803, 2003.

40) R. Verma, J. C. Crocker, T. C. Lubensky, and A. G. Yodh, Phys. Rev. Lett. 81, 4004, 1998.

41) L. Wang and V. A. Bloomfield, Macromolecules 23, 804, 1990.

42) Oosawa F. Polyelectrolytes. New York: Marcel Dekker; 1971.

43) GS Manning. Limiting laws and counterion condensation in polyelectrolyte solutions 1. Colligative properties. J. Chem. Phys. 51:924, 1969.

44) P. J. Hagerman *Annu*. *Re*V. *Biophys*. *Biophys*. *Chem*. *17*, 265, 1988.

45) T. Odijk, *J*. *Polym*. *Sci*. *B Polym*. *Phys*. *15*, 477, 1977.

46) J. Skolnick, M. Fixman, *Macromol*. *10*, 944, 1977.

47) C. G. Baumann, S. B. Smith, V. A. Bloomfield, C. Bustamante, *Proc*. *Natl*. *Acad*. *Sci*., *94*, 6185, 1997.

48) A. V. Dobrynin, *Macromol*., *38*, 9304, 2005.

49) G. S. Manning, *Biophysical J*., 91, 3607, 2006.

50) R. Shusterman, S. Alon, T. Gavrinyov and O. Krichevsky, Phys. Rev. lett. 92, 048303, 2004.

51) P. E. J. Rouse, J. Chem. Phys. 21, 1272, 1953.

52) B. H. Zimm, J. Chem. Phys. 24, 269, 1956.

53) Bracha D, Karzbrun E, Shemer G, Pincus PA, Bar-Ziv RH.
Proc Natl Acad Sci U S A., 4534-8, 2013.

תקציר

בהיותו הנשא של מטען גנטי, DNA מהווה , כנראה, את הפולימר החשוב ביותר למדע. בנוסף, מולקולת הDNA מושכת את עניינם של פיזיקאים היות ואורך ההתמדה שלה הוא כ50 נאנומטר. אורך זה הוא גדול בהרבה מקוטרה של מולקולת DNA (כ-2 נאנומטר) ומצד שני יכול להיות קטן בהרבה מאורך המולקולה (< 1 מיקרון). תכונות אלה מגדירות את ה-DNA כפולימר גמיש-למחצה בעל תכונות מבניות ודינמיות השונות מאלה של פולימרים גמישים (המהווים את רוב הפולימרים המלאכותיים המוכרים) ושל פולימרים קשיחים (כדוגמת סיבים ביולוגיים כמו מיקרוטובולים או אקטין). באופן מפתיע, עד פרסום עבודתינו, לא היו כלים ניסיוניים מוצלחים שאיפשרו מדידה של מבנה ה-DNA בקנה מידה שמעל 50 נאנומטר.

בתיזה זו נציג גישה חדשה של שילוב בין ספקטרוסקופיית מתאם פלורוסנטי בסריקה (SFCS) לבין סימון קוולנטי של מולקולות DNA בכדי למדוד תכונות מבניות ודינמיות של תמיסות DNA. באופן פרטני, עבור תמיסות דלילות של DNA איששנו לראשונה את התחזית התיאורטית בדבר התכונות הלמעשה אידיאליות של מבנה שרשראות ה-DNA: כתוצאה מגמישותה המוגבלת של המולקולה, ההסתברות להתנגשויות בין מקטעי DNA באורכים שעד כ-60 מיקרון היא נמוכה. ואכן, גורמי המבנה S(q)שמדדנו תואמים לתחזית של דבאי לגבי פולימרים אידאלים.

בעזרת סימון קוונלטי ספציפי, הטכניקה שלנו מאפשרת הערכה של התפלגות אורכי קצה-אל-קצה של DNA. התפלגות זו נראית כהתפלגות גאוסיאנית, עובדה שמחזקת את הטענה לגבי הטבע האידאלי של סלילי DNA.

שימוש נוסף של סימון ספציפי הוא באפשרות למדוד את המבנה של מולקולות DNA מסומנות בודדות המוטמעות בתוך רשת צפופה של מולקולות לא מסומנות (תמיסות דלילות-למחצה). עבור פולימרים גמישים, תנאים אלה מובילים להצטמקות של המולקולה עם עליית הריכוז. אולם, מדידות שערכנו על DNA חושפות שהמבנה המרחבי שלו נשאר זהה אפילו עד ריכוזים של כ-$\frac{\mu g}{ml}$ 600, פי 30 גבוה יותר מריכוז החפיפה של DNA באורכים שמדדנו. אי התלות של המבנה המרחבי בין המבנה המרחבי של מולקולה בודדת לבין צפיפות תמיסת הפולימרים היא תכונה מעניינת של פולימרים גמישים-למחצה לה אנו מביאים הוכחות ניסואיות לראשונה. למרות שתכונה זו נראית מפתיעה, ממצא ניסויי זה עומד בכפיפה אחת עם תורת התמיסות הגבוליות של פולימרים גמישים-למחצה של שייפר, ג'ואני ופינקוס [Macromolecules 13, 1280 (1980)].

בנוסף, פיתחנו שיטה חדשה המבוססת על SFCS, בה הסריקה מבוצעת במספר מהירויות, המאפשרת מדידת דינמיקה של מקטעי DNA בתמיסות דלילות-למחצה. באנלוגיה מדהימה לאי התלות של מבנה המולקולה בצפיפות התמיסה, אנו רואים השפעה מועטה ביותר של המולקולות הסמוכות על הדינמיקה של מקטע DNA כלשהו.

לבסוף, ניסינו לפתור באופן ניסויי את אחת הסוגיות שהיו שנויות במחלוקת בעשורים האחרונים השפעת ריכוז ועצמת האלקטרוליטים בתמיסה על אורך ההתמדה של DNA. למרות שבאופן כללי, תוצאותינו תואמות לתורה הקלאסית של אודיק, שקולניק ופיקסמן (OSF) על פוליאלקטרוליטים, אנו מוצאים סטיות משמעותיות מתורת OSF עבור ריכוזי אלקטרוליטים נמוכים מאוד.

הצהרת תלמיד המחקר עם הגשת עבודת הדוקטור לשיפוט

אני החתום מטה מצהיר/ה בזאת: (אנא סמן):

X__ חיברתי את חיבורי בעצמי, להוציא עזרת ההדרכה שקיבלתי מאת מנחה/ים.

X__ החומר המדעי הנכלל בעבודה זו הינו פרי מחקרי מתקופת היותי תלמיד/ת מחקר.

X__ בעבודה נכלל חומר מחקרי שהוא פרי שיתוף עם אחרים, למעט עזרה טכנית

הנהוגה בעבודה ניסיונית. לפי כך מצורפת בזאת הצהרה על תרומתי ותרומת שותפי למחקר, שאושרה על ידם ומוגשת בהסכמתם.

תאריך _13.06.2013__ שם התלמיד/ה _מניש נפל___ חתימה _

העבודה נעשתה בהדרכת

פרופ' אולג קריצ'בסי, מחלקה לפיסיקה

פרופ' דינה רווה , מחלקה למדעי חיים

בלימודים בין תחומים

קביעת תכונות מבניות ודינמיות של פולימרי DNA בעזרת ספקטרוסקופיית מתאם פלורוסנטי בסריקה

מחקר לשם מילוי חלקי של הדרישות לקבלת תואר "דוקטור לפילוסופיה"

מאת

נפל מניש

הוגש לסינאט אוניברסיטת בן גוריון בנגב

אישור המנחה _____ _____

אישור דיקן בית הספר ללימודי מחקר מתקדמים ע"ש קרייטמן _____

5 תמוז, 5773 **13.06.2013**

באר שבע

קביעת תכונות מבניות ודינמיות של פולימרי DNA בעזרת ספקטרוסקופיית מתאם פלורוסנטי בסריקה.

מחקר לשם מילוי חלקי של הדרישות לקבלת תואר "דוקטור לפילוסופיה"

מאת

נפל מניש

הוגש לסינאט אוניברסיטת בן גוריון בנגב

5 תמוז, 5773 **13.06.2013**

באר שבע