Dissertation

submitted to the
Physics Faculty
of the University of Bielefeld, Germany

for the degree of
Doctor of Natural Sciences

presented by
Dipl.-Chem. Hannes Barsch
born in Kiel, Germany

Oral examination: 30th January 2008

Referees: Prof. Dr. Markus Sauer
Prof. Dr. Andreas Hütten

Printed on non-aging paper according to DIN ISO 9706.

I herewith declare that the present thesis has been written by myself without the use of unallowed aids.

Bielefeld, 7th January 2008

Hannes Barsch

Summary

This thesis presents new insights into the structure of poly-L-proline and single-stranded polythymine polymers in aqueous solution.

The combination of time-correlated single-photon counting (TCSPC) and fluorescence correlation spectroscopy (FCS) of fluorescently labeled polyproline derivatives, F-(Pro)$_N$-Trp (with N = 2 - 10 and F = MR113, MR121, R6G), suggest that subpopulations of polyprolines with a reduced end-to-end distance exist. End-to-end distances have also been predicted for different isomers of polyprolines by means of simplified molecular mechanics simulations. Incorporating interspersed *cis*-bonds into the polyproline shows that *trans*-*cis* isomerization leads to significantly shorter end-to-end distances. This fits the experimental results of subpopulations with increased photoinduced electron transfer quenching observed in fluorescent correlation spectroscopy experiments.

Fluorescently labeled single-stranded polythymines, MR121-(dT)$_N$ (with N = 2 - 100), have been investigated through FCS and molecular dynamics simulations. The spectroscopic measurements resulted in length-dependent hydrodynamic radii of the polythymine probes, which scale according to a power law with an exponent of 0.5 - 0.7 dependent on ionic strength, I, of the solvent. The persistence length of MR121-(dT)$_{100}$, L_p, have been evaluated which shows a dependence of $L_\mathrm{p} \approx I^m$ with $m = -0.22 \pm 0.01$. The comparison to molecular dynamics simulations of various labeled and unlabeled polythymine derivatives revealed a semiflexible polymer behavior with neglectable influence of the fluorescent label for polymers with $N > 30$ and that electrostatic interactions are completely shielded at sodium chloride concentrations of 100 mM. Furthermore, simulation results showed polythymine’s static flexibility to be limited by steric and geometric constraints which can be expressed by an intrinsic persistence length of 1.7 nm.

Zusammenfassung

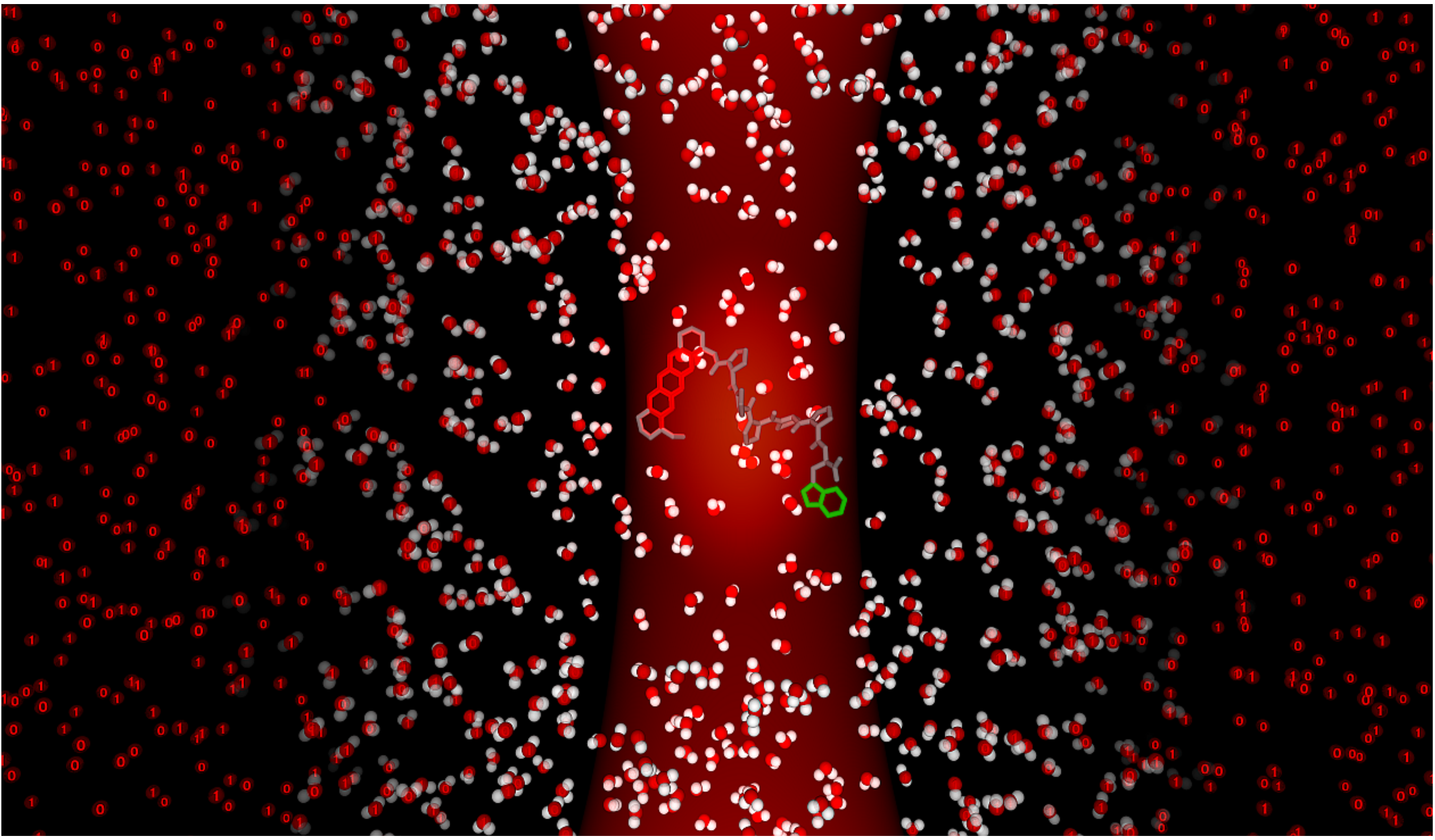
Neue Erkenntnisse bezüglich der Polymerstruktur von Poly-L-prolinen und einzelsträngigen Polythyminen in wässriger Lösung warden in dieser Arbeit präsentiert.

Die Kombination von zeitkorrelierter Einzelphotonenregistrierung (TCSPC) und Fluoreszenzkorrelationsspektroskopie (FCS) von mit Fluoreszenzfarstoffen markierten Polyprolinderivaten, F-(Pro)$_N$-Trp (mit $N = 2 - 10$ und F = MR113, MR121, R6G), legen nahe, dass Sub-Populationen der Polyproline mit verringertem Endabstand existieren. Diese Endabstände wurden ebenfalls mittels vereinfachten Molekülmechanik-Simulationen für verschiedene Isomere der Polyproline vorausgesagt. Durch Einbringen von vereinzelten *cis*-Bindungen in die Prolinpolymere kann gezeigt werden, dass *trans*-*cis*-Isomerisierung die Endabstände eines Polymers deutlich verringern kann. Diese Endabstandsverkürzung entspricht den experimentell beobachteten Sub-Populationen, welche bei FCS-Messungen erhöhte Löschung durch photoinduzierten Elektronentransfer ermittelt wurden.

Mit einem Fluoreszenzfarbstoff markierte, einzelsträngige Polythymine, MR121-(dT)$_N$ (mit $N = 2 - 100$), wurden durch FCS und moleküldynamische Simulationen untersucht. Die spektroskopischen Messungen liefern längenabhängige hydrodynamische Radien der Polythyminproben, deren Abhängigkeit durch ein Potenzgesetz mit einem Exponenten von 0,5 - 0,7 von der Ionenstärke, I, des Lösungsmittels abhängen. Die Persistenzlänge von MR121-(dT)$_{100}$, L_p, wurde berechnet und zeigt die Abhängigkeit $L_\mathrm{p} \approx I^m$ mit $m = -0{,}22 \pm 0{,}01$. Der Vergleich mit molekúdynamischen Simulationen verschiedener markierter und unmarkierter Polythyminderivate zeigt, dass sich Polythymine wie semiflexible Polmere verhalten, dass für Kettenmoleküle mit einer Länge von $N > 30$ der Einfluss der Farbstoffmarkierung vernachlässigbar ist, und dass elektrostatische Wechselwirkungen bei einer Natriumchlorid-Konzentration von 100 mM vollständig abgeschirmt sind. Weiterhin zeigten die Simulationen, dass die statische Flexibilität der Polythymine durch sterische und geometrische Einschränkungen limitiert ist, was durch eine intrinsische Persistenzlänge von 1,7 nm ausgedrückt werden kann.

Table of Contents

Investigation of Biopolymer Properties with Combined Fluorescence Correlation Spectroscopic and Computational Methods



Referees: Prof. Dr. Markus Sauer
Prof. Dr. Andreas Hütten

1. Introduction

Life is biopolymers. The role of biopolymers in nature is existential: in every living organism desoxyribonucleic acid polymers (DNA) store genetic information that holds a blueprint of the life form. This incredible amount of data, necessary to ensure the secure operation of highly complex “machinery” as a human being with all its facets is condensed to a few macromolecules consisting of four different monomers, present within each cell’s nucleus. From these data storage molecules fractional copies responsible for specific tasks are created in the form of ribonucleic acid polymers (RNA). These polymers are transported out of the cell nucleus to be translated into proteins by the ribosomes. Proteins are chain molecules consisting of 20 different amino acids. Despite the limited number of building blocks, proteins can adopt an enormous amount of different shapes, structures and functions. Proteins are nature’s work horses – they are responsible for most of the active processes involved in life.

A deeper understanding of life on the molecular level is necessary with respect to biological, medical and philosophical questions. In biological terms knowledge of molecular mechanisms within the cell leads to a detailed understanding of similarities or differences between different species. Medical science becomes more and more the research for modern drugs which act on a molecular basis – they interact with DNA, RNA or proteins by enhancement, prevention or alteration of their function to compensate malfunctions within the human body’s cells caused by diseases. Last but not least understanding life in detail is a philosophical concern: The question what life is and how it began has disturbed mankind ever since.

The detailed description of biopolymers is an impossible task as these macromolecules imply a numberless amount of degrees of freedom on the molecular level. Therefore, simplified models that predict the polymers’ properties have been developed. Dynamic and static flexibility of polynucleotides and polypeptides have been topic of research using a variety of experimental techniques as spectroscopic methods, atomic force microscopy, video imaging or dynamic light scattering. From these experimental results a simplified picture of the polymer behavior has been drawn and polymer representations like semiflexible chains and wormlike chain models have been concluded. Biopolymers that are rigid on length scales above the dimensions of the individual monomers contributing to the molecule, show polymer characteristics that are independent of the monomers' atomic structures. Double-stranded DNA, for example, can be described as polyelectrolyte, i.e. a charged polymer chain, displaying structural and dynamic properties very well predicted by the model of an ideal polymer.

The order of length, at which a biopolymer becomes flexible can be described by the persistence length, which was found to be ~ 50 nm for double-stranded DNA [1-3], ~ 15 µm for actin filaments [4, 5], and ~ 5 mm for microtubules [6, 7]. Unstructured single-stranded polynucleotides like single-stranded DNA (ss-DNA) or RNA [1, 8-11] and unstructured polypeptides [12-14] show persistence lengths on a much smaller length scale in the range of a few nanometers or less.

The great differences in flexibility of different biopolymers arise from the diverse structures of the constituting monomers and their interaction with each other. The monomer structure defines the degrees of freedom within the polymer backbone due to internal rotation around atomic bonds. Voluminous side groups of the monomers do not severely influence the backbone geometry directly, but still they further restrict the possible conformations of the polymer by attractive or repulsive interaction with other polymer moieties. These side group interaction can be short-ranged in the form of VAN DER WAALS interaction as well as long-ranged in terms of electrostatic forces. In summary side group interactions can lead to a strong influence on the persistence length if they significantly increase excluded-volume effects due to impenetrability upon folding of the polymer on itself.

Particularly, dilute polyelectrolytes show deviations from simple polymer behavior as electrostatic interactions between charged polymer moieties are long-ranged but can be effectively shielded by counter ions. Therefore, polymer characteristics become dependent on the ionic strength of the solvent, i.e. they are dependent on the solvent's dielectricity constant and its salt concentration. This influence has been originally described by the ODIJK-SKOLNICK-FIXMAN theory through an apparent persistence length. It consist of an intrinsic component and an electrostatic term which expresses the persistence length's function of ionic strength [15, 16]. This function has been predicted to scale with the inverse [15-17] or with the inverse square root [18, 19] of ionic strength, depending on specific values for intrinsic flexibility, strength of electrostatic interactions, and charge density along the polymer. Nevertheless, no global theory that covers all polymer-solvent situations has been developed, yet. Especially, no satisfying valid description of ss-DNA in aqueous solution has been given so far [20, 21].

In contrast to ss-DNA the long-ranged interaction in polypeptides tend to be more inhomogeneous. This is because of the diversity of amino acid side groups: some side groups are charged and others form hydrogen or disulfide bonds to far away moieties of the polymer. Therefore, polypeptides or proteins show a far more complex polymer behavior. Exceptionally, polypeptides only consisting of the amino acid proline, are well known to adopt a polyproline type II helix in aqueous solution. Arising from the unique feature, that proline is the only naturally occurring amino acid in which the side chain is cyclized to the backbone, its conformational space is considerably restricted [22], leading to the strong favor for the regular PPII helix formation [23, 24] in which all amide bonds form *trans* conformations. Furthermore, a second form of polyprolines, the polyproline type I helix, is known which can be established in anhydrous polar solvents under the influence of lithium chloride. In the PPI helix all amide bonds form *cis* conformation [25].

This structural regularity of PPII polyproline helices qualifies them to be used as spectroscopic ruler reference systems: Probes can be calibrated against polyprolines of defined length in spectroscopic measurements such as FÖRSTER resonance energy transfer (FRET) [26]. The end-to-end distance of polyprolines scales directly proportionally to the number of monomers which in turn defines the intensity of non-radiative dipole-dipole interaction between two fluorophores labeled to the ends of the polyproline. Such well defined molecular systems are needed for validation of FÖRSTER fluorophore pairs which again are used to derive distances and distance changes within biological samples from energy transfer efficiencies.

However, length measurements on polyprolines show significant deviations from the proposed end-to-end distance. These differences were found through ensemble FRET measurements [27], single-molecule FRET measurements [28-30], and nuclear magnetic resonance experiments [31, 32]. The question has continuously been discussed whether observed reduced polyproline end-to-end distances were due to constant bending of the polymer as a result of a wormlike chain behavior or if interspersed *cis* isomers caused the rigid polymer to kink.

In this thesis studies on two different biopolymers are presented: polythymines as representatives of single-stranded DNA and polypeptides exclusively formed of proline. By means of fluorescence correlation spectroscopy diffusion constants of fluorophore labeled polythymines with lengths of 2 - 100 monomers have been identified. These constants provide insight into the shape of the polymers in aqueous solution. Molecular dynamics simulations have been performed to further investigate the influence of the fluorophore labeled to the probes and the effect of polythymine's nucleobases on structural properties of the polymers. A comparison of experimental data, simulation results, and chain model predictions reveales that polythymine polymers can be interpreted as polyelectrolytes with limited static flexibility.

The second part of the thesis deals with the question of polyproline structure in aqueous solution. With spectroscopic measurements site-specifically fluorophore labeled polyprolyl-tryptophans consisting of 0 - 10 proline residues have been analyzed in terms of the contact formation of fluorophore and tryptophan. Subpopulations with reduced end-to-end distances of the proline backbone have been found and can be rationalized by computational modeling of the polymers.

2. Theory

In this work molecular modeling results are compared to results from fluorescence spectroscopy experiments. Both techniques are strongly connected to the quantum mechanical nature of molecules. Therefore, a short summary of the quantum mechanical aspects of matter and light are presented, followed by the resulting theories on fluorescence and molecular modeling. The fourth part deals with mathematical approaches to describe the behavior of polymers.

2.1. Quantum Mechanical Basics

Quantum mechanics (QM) are the fundamental theory of any physical observable. A golden rule in quantum mechanics states that the smaller a system gets, the larger the quantum mechanical effect is. When dealing with small systems like molecules quantum mechanics are fundamental: classical physics cannot describe molecular properties appropriately.

2.1.1. LIGHT

For three centuries light was commonly regarded as a wave phenomenon, first described by HUYGENS in the 17th century. In the early 1800s this theory was enforced by YOUNG and FRESNEL's interference experiments and at the end of the 19th century MAXWELL explained light as the propagation of electromagnetic waves with the MAXWELL equations. At the beginning of the 20th century fundamental postulations on the quantized nature of light were given by EINSTEIN's description of the photoelectric effect. Finally, DE BROGLIE developed his hypothesis that any particle shows wave-like nature.

Photoelectric effect

When light hits a metal surface electrons can be emitted by the metal. This process only depends on the wave length of the light: up to a material-specific wavelength no electrons are emitted regardless of the light's intensity. EINSTEIN interpreted this phenomenon with the quantum electro dynamical nature of light [33]: Photons, the light quanta, carry energy corresponding to their wave length and the quantum's energy is responsible for the interaction with the metal. In contrast to PLANCK's theory on the quantization of thermal radiation being absorbed or emitted by a black body [34], EINSTEIN first postulated this quantization as an intrinsic property of the electromagnetic field, independent on the interaction with matter.

The photoelectric effect, that the maximum kinetic energy, $E_{k,max}$, of emitted electrons varied with the frequency of incident radiation, could then be explained as the difference of the photon's energy, E_P, and a material specific work function, W_0:

$$E_{K,max} = E_P - W_0 \tag{2.1}$$

Based on PLANCK's theory, the photon's energy equals the product of its frequency, ν, and the PLANCK constant, h:

$$E_P = h \cdot \nu \tag{2.2}$$

The corresponding photon's wavelength, λ, is the quotient of the frequency and the speed of light, c:

$$\lambda = \frac{c}{\nu} \tag{2.3}$$

Combining equations (2.1) to (2.3) results in a description of the photoelectric effect with quantized light:

$$E_{K,max} = \frac{hc}{\lambda} - W_0 \tag{2.4}$$

2.1.2. MATTER

History

In contrast to light, matter was not attributed a wave character until the 20th century. With the rise of the concept of quantum mechanics, DE BROGLIE was the first to postulate that small particles should show wave phenomena as well. In 1925 he defined the wave length, λ, of any particle as the PLANCK constant over its momentum [35]:

$$\lambda = \frac{h}{p} \tag{2.5}$$

Two years later DAVISSON & GERMER proved DE BROGLIE's theory scattering slow electrons on nickel crystals [36]. This discovery enforced the development of quantum mechanics as it visibly confirmed that the wave-particle dualism was not a question of either or but a physical coexistence.

The smallest unit of matter in chemical respects is an atom built of a heavy, positively charged nucleus and negatively charged electrons. Following the SCHRÖDINGER equation (2.6), the electrons are forced to specific energetic states which can be visually described as orbitals. Any change of the electrons' states implies a change of the atom's energy. Despite this electronic energy of an atom the only other way to gain or loose energy is the atom's movement through space.

Molecules behave similar but in contrast to atoms molecules consist of two to thousands of atoms. In addition to electronic and transitional degrees of freedom molecules can store energy in internal vibrations and rotations. Moieties of the molecule can change their relative position in a way that is quantum mechanically driven, too.

Electronic States

Electrons of a molecule occupy molecular orbitals. These orbitals are defined by the solutions of the SCHRÖDINGER equation (2.6).

$$\hat{\mathrm{H}} \cdot \Psi = E \cdot \Psi \tag{2.6}$$

$\hat{\mathrm{H}}$ denotes the HAMILTONian operator which is the second derivative in all three spatial directions multiplied with the negative of the PLANCK constant divided by twice the electron mass. $\hat{V}$ is the operator describing the potential felt by the electron and E the energy of the orbital.

$$\left\{ -\frac{\hbar}{2m} \cdot \nabla^2 - \hat{V} \right\} \cdot \Psi = E \cdot \Psi \tag{2.7}$$

The wave function Ψ multiplied with its complementary complex, Ψ^*, results in the electron density, P_{e^-}, and yields the structure of the orbital.

$$P_{e^-} = \Psi^* \cdot \Psi \tag{2.8}$$

As molecules are many-particle systems there is no exact solution to the SCHRÖDINGER equation. The SCHRÖDINGER equation can only be solved exactly for a two-particle system as a hydrogen atom or a helium cation that can be described as a three-particle system where the nucleus consisting of two protons and two neutrons can be dealt with as one particle, leaving two effective particles. For any heavier atoms or even molecules the SCHRÖDINGER equation can only be approximated. One first step to reduce a molecule's complexity is the BORN-OPPENHEIMER approximation: the nuclei of the involved atoms with masses some thousand times greater than the electron mass can be regarded individually from the electrons. This yields in a separation of the nucleic and electronic wave functions:

$$\Psi(\text{nuclei,electrons}) = \Psi(\text{nuclei}) \cdot \Psi(\text{electrons}) \tag{2.9}$$

This separation approach halves the number of independently treated particles but still leaves a many-particle system that cannot be solved. In any polyelectronic system the electrons' charges manipulate the electrostatic potential of the other electrons. Therefore, solutions for the electronic SCHRÖDINGER equation of molecules can only be approximated.

2.1.3. INTERACTION OF LIGHT AND MATTER

Energy which is transferred to a molecule can be stored in three types of internal degrees of freedom: nuclei movement accounts for rotational and vibrational degrees of freedom and bound electrons can change the energetic state called electronic degrees of freedom. The energy differences of such changes of states are in the range of 1 - 5 eV for electronic transitions, about 0.1 eV for vibrational transitions and rotational excitation shows energy gaps of some 0.001 eV.

Apart from intermolecular collisions, changes in the state of a molecule can be achieved by interaction of a molecule with electromagnetic radiation – absorption of a photon results in an increase of the molecule's energy by the energy of the photon. Emission of a photon is the reverse of the absorption process, leaving the molecule in an energetic state where its energy after emitting a photon equals the difference of its prior energy and the photon's energy.

Due to quantum mechanics any system shows discrete energetic states only. As a system decreases in size, the amount of states decreases which leads to an increase in energy difference between two states. Large systems (as rigid bodies) are able to interact with photons of a wide energy range whereas a single molecule shows a different behavior: only photons of discrete energies are interacted with.

The BOHR frequency condition states that photons can interact with molecules only if their energy equals the energy gap in between two states of the molecule, ΔE :

$$\begin{aligned} \Delta E &= h \cdot \nu \\ &= h \cdot c \cdot \lambda \end{aligned} \tag{2.10}$$

The above mentioned energy difference for internal degrees of freedom can thus be described as photons of certain wavelengths that can interact with a molecule. Electronic, vibrational and rotational transitions can be achieved by interaction with photons of approximately 250 - 1000 nm, 10 µm and 1 mm, respectively.

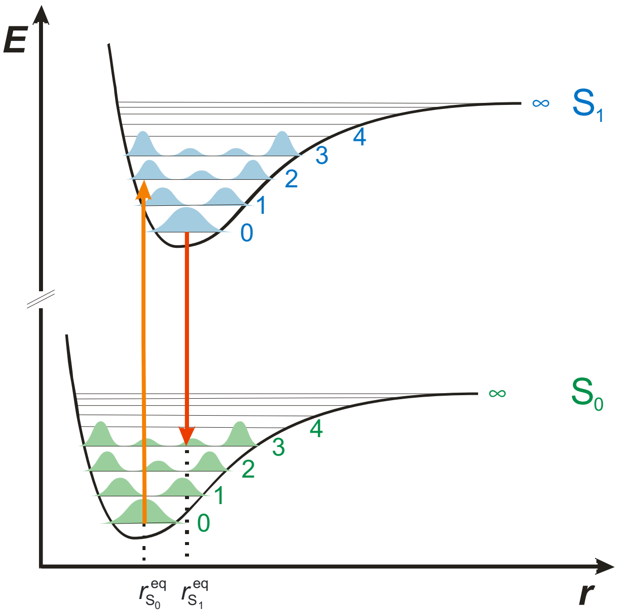


Figure 1. Electronic and vibrational states of a molecule. Due to the the high mass of the nuclei, absorption takes place vertically according to the Frank-Condon principle. After internal relaxation, i.e. vibrational relaxation to the groundstate within the eletronically excited state, a fluorescence photon is emitted as the molecule relaxes to the vibrationally excited electronic ground state.

If a molecule is electronically excited by the absorption of a photon this takes place in a fast process in the range of a femtosecond. As stated by the Born-Oppenheimer approximation the heavy atom cores in the molecule cannot react to the change in electron distribution due to their high moment of inertia (Figure 1). Therefore, the geometry of the molecule does not change directly with excitation and the electronic transition is influenced by the location probabilities of the excited electron in the two states described by the square of the electrons' wave functions, Ψ. A transition will occur with the probability $P_{\mathrm{gr}\rightarrow\mathrm{ex}}$

$$P_{\mathrm{gr}\rightarrow\mathrm{ex}} = \left\langle \Psi_{\mathrm{ex}} \middle| \mu \middle| \Psi_{\mathrm{gr}} \right\rangle = \int \Psi^{*}_{\mathrm{ex}}\left(r_1,r_2,\ldots r_n\right)\cdot \mu \cdot \Psi_{\mathrm{gr}}\left(r_1,r_2,\ldots r_n\right)\cdot d\tau \qquad (2.11)$$

Accounting for the Born-Oppenheimer approximation, the wave function is a composition of nucleus ($\Psi_{\mathrm{ex},N}$), electron ($\Psi_{\mathrm{ex,E}}$) and spin ($\Psi_{\mathrm{ex,S}}$) wave functions, and equation (2.11) expands to:

$$
\begin{aligned}
P_{\mathrm{gr}\rightarrow \mathrm{ex}} &= \left\langle \Psi_{\mathrm{ex,N}} \Psi_{\mathrm{ex,E}} \Psi_{\mathrm{ex,S}} \middle| (\mu_{\mathrm{E}} + \mu_{\mathrm{N}}) \middle| \Psi_{\mathrm{gr,N}} \Psi_{\mathrm{gr,E}} \Psi_{\mathrm{gr,S}} \right\rangle \\
&= \int \Psi^{*}_{\mathrm{ex,N}} \Psi_{\mathrm{gr,N}}\, d\tau_{\mathrm{N}} \cdot \int \Psi^{*}_{\mathrm{ex,E}}\, \mu_{\mathrm{E}} \Psi_{\mathrm{gr,E}}\, d\tau_{\mathrm{E}} \cdot \int \Psi^{*}_{\mathrm{ex,S}} \Psi_{\mathrm{gr,S}}\, d\tau_{\mathrm{S}} \\
&\quad + \int \Psi^{*}_{\mathrm{ex,E}} \Psi_{\mathrm{gr,E}}\, d\tau_{\mathrm{E}} \cdot \int \Psi^{*}_{\mathrm{ex,N}}\, \mu_{\mathrm{N}} \Psi_{\mathrm{gr,N}}\, d\tau_{\mathrm{N}} \cdot \int \Psi^{*}_{\mathrm{ex,S}} \Psi_{\mathrm{gr,S}}\, d\tau_{\mathrm{S}}
\end{aligned}
\tag{2.12}
$$

The first integral of the first summand in equation (2.12) is the FRANCK-CONDON factor, the second describes the orbital selection rules, the third the spin selection rules. The second summand’s first integral and therefore the whole summand equal zero.

In molecules electronic and spin wave functions of the ground and the excited state are not perpendicular. Therefore, the selection rules do not apply strictly. Transitions depend mainly on the conservation of spin followed by the conservation of orbital angular momentum.

spin	**orbital**	**probability**
allowed	allowed	10^{3} - 10^{5}
allowed	forbidden	10^{0} - 10^{3}
forbidden	allowed	10^{-5} - 10^{0}

Table 2-1. Transition probabilities.

Apart from the frequency of the light wave its orientation also has to fulfill certain requirements to allow interaction with matter. Electromagnetic waves are a combination of electric and magnetic fields oscillating in space and time domain with a common propagation direction but orthogonal amplitudes. The interaction of the radiation’s electronic wave vector and the molecule’s electronic transition moment is proportional to the scalar product of these two vectors (Figure 2). Therefore, parallel orientation yields in a maximal interaction.

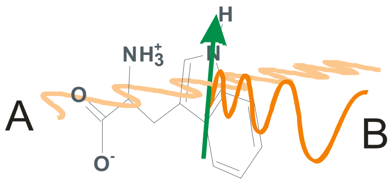


Figure 2. Interaction of photons with tryptophan. Photon A cannot interact with tryptophan as its electronic wave vector (bright orange) is orientated perpendicular to tryptophans dipole moment (green arrow). Photon B is absorbed due to coplanarity of wave vector (orange) and dipole moment.

2.2. *Principles of Photon Absorption and Fluorescence*

2.2.1. EXCITATION AND DEEXCITATION PATHWAYS

The above described quantum mechanical processes apply to absorption and emission of electromagnetic radiation by matter. At reasonable temperatures molecules reside in the electronic ground state making excitation to higher electronic, vibrational and rotational states the only pathway for light-matter interactions. Once the molecule is excited to a higher electronic state other transitions show up.

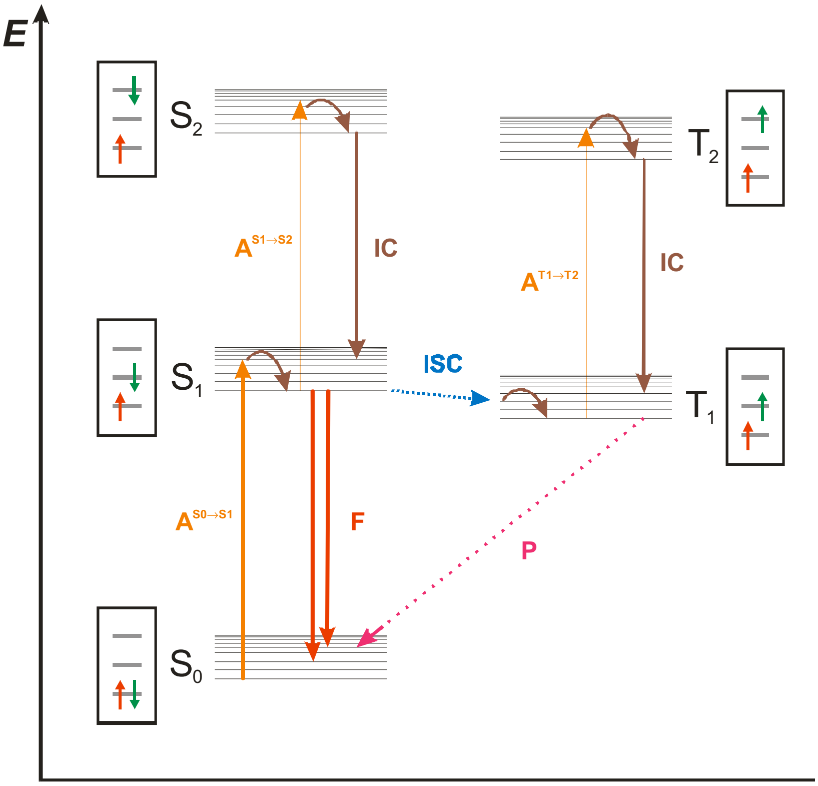


Figure 3. Jablonski diagram with singulet states at the left and triplet states at the right. For each electronical state the valence electron configuration is given (boxes). See text for a detailed describtion.

The JABLONSKI diagram (Figure 3) clarifies the possible routes of a molecule excited by radiation. Absorption leads to an electronic excited state, S_N, which is vibrationally excited but relaxes radiationless to the ground state within 10^{-12} s (VR). If the molecule is excited to an electronic state higher than S_1 internal conversion (IC) to the first excited state takes place in a time range of 10^{-13} to 10^{-11} s. From this state the molecule can either return to the electronic ground state by fluorescence (emitting a photon) or radiationless through internal conversion. Fluorescence is like photon absorption a fast process elapsing in 10^{-15} s. Thus, the FRANCK-CONDON principle applies to this process, too, and the transition proceeds to a vibrationally excited electronic ground state.

transition	**type**
$S_1 \rightarrow S_0 + \Delta G$	radiationless deactivation (IC)
$S_1 \rightarrow S_0 + h\nu_f$	fluorescence (F)
$S_1 \rightarrow T_1$	intersystem crossing (ISC)
$T_1 \rightarrow S_0 + \Delta G$	intersystem crossing (ISC)
$T_1 \rightarrow S_0 + h\nu_p$	phosphorescence (P)

Table 2-2. Transitions in excited molecules.

The electronically excited molecule can also convert to a triplet state with a multiplicity change from 1 to 3. This inter system crossing (ISC) is spin forbidden which implies a low transition probability being mainly dependent on the chemical environment of the molecule. Due to the change in spin the ISC is relatively slow with characteristic time spans of 10^{-10} to 10^{-8} s. The triplet state can be depopulated by phosphorescence where a photon is emitted and the molecule returns to a vibrationally excited electronic ground state. As this process implies a multiplicity change again, the lifetime of the triplet state ranges from 10^{-3} to 10^{2} s. On a similar time scale internal conversion can also occur without photon emission.

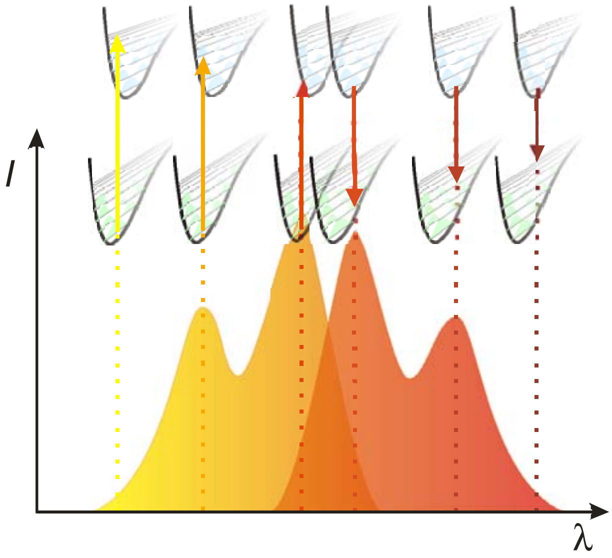


Figure 4. Stokes shift. Light of higher energy is absorbed compared to the light emitted. This is due to the dissipation of energy by vibrational relaxation after the electronic transitions.

As the molecule dissipates excessive energy in higher excited states quickly, fluorescence is commonly observed as the transition from the first (vibronic and electronic) excited state. This phenomenon is known as KASHA's rule. In combination with the fact that absorption and emission are FRANCK-CONDON processes the energy emitted by fluorescence is always lower than the energy absorbed, previously. The loss of energy through vibrational relaxation leads to the bathochrome STOKES shift (Figure 4) when comparing the two spectra. If the geometries of ground and excited state do not differ significantly, the absorption and emission spectra behave like mirror images as the energy differences of the vibrational states are comparable.

2.2.2. FLUORESCENCE QUENCHING

While a molecule stays in the first electronic excited state, S_1, the fluorescence intensity and lifetime can be altered by interactions with the surrounding. Either energy or charge is exchanged with neighborhood molecules or internal processes affect the spectral properties of the fluorophor. Several mechanisms can act on the electronic characteristics of a fluorophor, stated in Table 2-3.

type	mechanism of quenching
inter system crossing	$F^*_{S.} \rightarrow F^*_{T.}$
energy transfer	$F^* + Q \rightarrow F + Q^*$
electron transfer	$F^* + Q \rightarrow F^{+/-} + Q^{-/+}$
exciplex formation	$F^* + Q \rightarrow (FQ)^*$
proton transfer	$F^* + QH \rightarrow FH^+ + Q^-$

Table 2-3. Quenching mechanisms.

Inter system crossing

Inter system crossing was already mentioned in 2.2.1. A spin forbidden transition from the S_1 to the long-lived T_1 state results in a temporal loss of fluorescence. Heavy atoms or triplet oxygen can facilitate this process by spin exchange with the fluorescent molecule.

Electron transfer

Electron transfer can be divided into two different mechanisms: DEXTER electron transfer and photoinduced electron transfer (PET). The DEXTER mechanism is an electron exchange between the fluorophore and the quencher whereas the photoinduced electron transfer creates charged molecules. In the DEXTER process (Figure 5) the excited electron is transferred to the quencher molecule and either in a concerted exchange or stepwise an electron of the highest occupied molecular orbital (HOMO) of the quencher is transmitted to the fluorophore. As a result the fluorophore is converted to the ground state leaving the quencher in an electronic excited state.

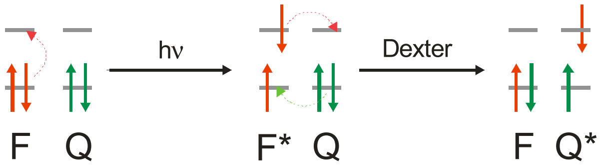


Figure 5. Dexter type quenching. After excitation fluorophore (F) and quenching molecule (Q) exchange electrons leaving F in ground state and excited Q.

The photoinduced electron transfer can proceed in two different ways: Depending on the redox potential of fluorophore and quencher the fluorescing moiety is either oxidized to a positively charged or reduced to an anionic fragment. After the fluorophore absorbs a photon the excited electron can relax to the energetically lower lowest unoccupied molecular orbital (LUMO) of the quencher. This reflects oxidizing PET occurring when the frontier orbitals (FOs) of the quenching molecule are energetically lower than the fluorophore's (Figure 6, top).

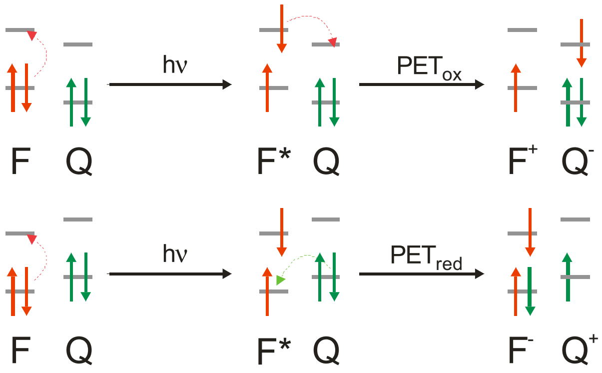


Figure 6. Photoinduced electron transfer. Depending on the redox potential of fluorophore (F) and quencher (Q) oxidative or reductive PET occur. In oxidative PET the quencher acts as an electron acceptor, in reductive PET as an electron donor.

If the quencher's FOs' energy lies above the fluorophore's, the electron hole arisen from one of the fluorophore's HOMO electron propagation to the S_1 state is compensated by the quencher's HOMO electron with adequate spin. This process reflects a reduction of the excited fluorophore, yielding it's radical anion and a positively charged quenching molecule (Figure 6, bottom).

The thermodynamics of an oxidizing photoinduced electron transfer can be expressed with the Rehm-Weller equation:

$$\Delta G = E^{F/F^+} - E^{Q/Q^-} - \Delta G_{00} - \frac{e}{\varepsilon \cdot d} \qquad (2.13)$$

The energy difference between the initial system with the excited fluorophore and ground state quencher, ΔG, determines whether the PET causes a stabilization of the fluorophore-quencher system. E^{F/F^+} and E^{Q/Q^-} are the redox potentials of the two molecules, ΔG_{00} is the transition energy between S_0 in the vibrational ground state ($\nu = 1$) and the equivalent rotational state of S_1. The last term of the Rehm-Weller equation describes the energy gained through solvation of the complex if the evolved charges are separated by a distance, d, in a solvent with the dielectric permittivity, ε. The factors incorporated are relatively easy to determine but strongly solvent dependent.

Exciplex Formation

Instead of transfer of an electron fluorophore and quencher can also build a stable excited complex (exciplex) without charge transfer (Figure 7). Exciplex formation proceeds if upon combination of the wave functions of the two molecules the LUMOs split up with a higher energy difference than the HOMOs.

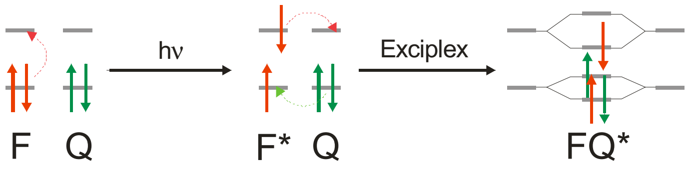


Figure 7. Exciplex formation. The orbitals of fluorophore and quencher overlap and interact. They split up and form new molecular orbitals.

Exciplexes show fluorescence preferably in non-polar solvents. The emission wavelength is red-shifted due to the smaller energy gap between the two singly occupied molecular orbitals formed upon complex formation. In polar solvents exciplexes are usually quenched by solvent molecules. The dissipation to the exciplex ground state and following dissociation to fluorophore and quencher are radiationless.

Proton Transfer

The changed electron configuration evoked by the absorption of a photon can change the acidity of a fluorophore in the excited state. For example phenols tend to show greater pK_A-values in the S_1 state as the oxygen electrons are shifted into the aromatic system, weakening the OH-bond. In such cases proton transfer reactions can stabilize the excited fluorophore and alter its emission frequency.

Resonance Energy Transfer

In contrast to the mechanisms elucidated above resonance energy transfer influences the properties of an excited fluorophore without charge transfer. In the FÖRSTER Resonance Energy Transfer (FRET) [37] the excited fluorophore's (donor's) excess energy can excite a second fluorophore (acceptor) in an non-radiative process, which requires an overlap of the emission spectrum of the donor and the absorption spectrum of the acceptor to ensure resonance between the two systems. As the energy transfer is a dipole-dipole interaction its intensity scales with the distance between the chromophores to the power of 6. Furthermore, the relative orientation of the two interacting dipoles is required to have a parallel component. FRET occurs at inter-fluorophore distances of 1 - 9 nm [26, 27] and is extensively used in bioanalytical applications as a "spectroscopic ruler" [26, 29, 38].

2.2.3. FLUORESCENCE QUANTUM YIELD AND LIFETIME

Fluorescence quantum yield and fluorescence lifetime are fundamental optical parameters in spectroscopy. Both parameters are strongly dependent on the chemical environment of a fluorophore. The fluorescence quantum yield is defined as the ratio of absorbed to emitted photons and therefore provides a measure for the proportion of radiative and radiationless relaxation processes of the excited chromophore. As indicated in Figure 3 three main pathways lead an excited molecule back to the ground state: radiationless internal conversion, inter system crossing, and the radiative process of fluorescence. The rate of the non-radiative processes, k_{NR}, is thus:

$$k_{\mathrm{NR}} = k_{\mathrm{IC}} + k_{\mathrm{ISC}} \tag{2.14}$$

Hence, if other non-significant processes are neglected, the fluorescence quantum yield is:

$$\Phi_{\mathrm{F}} = \frac{k_{\mathrm{R}}}{k_{\mathrm{R}} + k_{\mathrm{NR}}} = \frac{k_{\mathrm{F}}}{k_{\mathrm{F}} + k_{\mathrm{IC}} + k_{\mathrm{ISC}}} \tag{2.15}$$

Quantum yields close to one represent fluorophores that generally relaxes to the ground state by fluorescence.

The kinetics of the depopulation of a fluorophore’s excited state can be described with the rate constants of fluorescence, internal conversion and inter system crossing. Considering *N* excited molecules yields the differential equation

$$\frac{dN(t)}{dt} = -\left(k_{\mathrm{F}} + k_{\mathrm{IC}} + k_{\mathrm{ISC}}\right) \cdot N(t) \tag{2.16}$$

Evaluation of equation (2.16) results in an exponential decay of the molecules to the ground state.

$$N(t) = N_0 \cdot e^{-\frac{t}{k_{\mathrm{F}} + k_{\mathrm{IC}} + k_{\mathrm{ISC}}}} \tag{2.17}$$

The quotient of the exponent is the inverse of the excited state lifetime, τ_{F}, and describes the time, fluorescence intensity takes to fall off to 1 / e.

$$\tau_{\mathrm{F}} = \frac{1}{k_{\mathrm{F}} + k_{\mathrm{IC}} + k_{\mathrm{ISC}}} \tag{2.18}$$

Fluorescence lifetimes are in the nanosecond range for common organic fluorophores and can be investigated with time-resolved fluorescence measurements.

2.2.4. DYNAMIC AND STATIC FLUORESCENCE QUENCHING

Quenching can occur either if the fluorophore collides with a quenching molecule during its excite state lifetime or if the fluorophore forms a stable complex with a quencher that is long-lived compared to the excitation-fluorescence cycle. These two classes of quenching processes are called dynamic and static quenching.

Dynamic Quenching

In the dynamic collision model the excited fluorophore is hit by a diffusing quencher. In the encounter complex formed the excitation energy is non-radiatively transferred to the quencher and dissipated to the solvent environment (Figure 8).

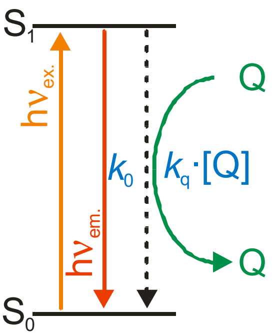


Figure 8. Dynamic quenching. The radiationless interaction of the excited fluorophore (F) and a quenching molecule (Q) during the fluorophore's excited state life time returns the fluorophore to the ground state. The rate of the collisional quenching scales with the concentration of the quencher.

Dynamic quenching influences fluorescence lifetime as well as fluorescence intensity. If the quencher concentration, [Q], exceeds the fluorophore concentration, [F], dramatically, it can be regarded as constant. The temporal decay of the fluorescence intensity, I, can then be described as:

$$I(t) = I_0(t) \cdot e^{-\left(k_0 + k_q \cdot [\mathrm{Q}]\right)t} \tag{2.19}$$

The fluorescence intensity depends on the rate constant k_0 which is the sum of the rate constants of all internal de-excitation processes (fluorescence, internal conversion, and inter system crossing). The dynamic quenching is a reaction of pseudo first order and is incorporated in the time law as the product of the rate constant k_q and the quencher concentration. The lifetime of the fluorophore's excited state is reduced due to dynamic quenching:

$$\tau = \frac{1}{k_0 + k_q[\mathrm{Q}]} \tag{2.20}$$

Encounter complex formation forces fluorophore and quencher to approach to a certain minimum distance. If the quenching process is fast compared to diffusion, it is diffusion limited. The maximum quenching constant, k_d, can then be derived from the SMOLUCHOWSKI equation:

$$k_d = 4\pi \cdot (r_F + r_Q) \cdot (D_F + D_Q) \cdot N_A \tag{2.21}$$

The quenching constant depends on the sum of the radii, r_F and r_Q, and the sum of the diffusion coefficients, D_F and D_Q. The diffusion coefficient, D, can be calculated with the STOKES-EINSTEIN equation (2.22) with the BOLTZMANN constant, k, temperature, T, the medium's viscosity, η, and the solute's hydrodynamic radius, R_h.

$$D = \frac{kT}{6\pi \cdot \eta \cdot R_h} \tag{2.22}$$

If the solute molecule is regarded as a sphere, the hydrodynamic radius can be approximated by the sphere's volume which must equal the solutes mass (as the molar mass, *M*, over AVOGADRO's number, N_A) divided by its density:

$$R_h = \sqrt[3]{\frac{3M}{4\pi \cdot \rho \cdot N_A}} \tag{2.23}$$

Dynamic fluorescence quenching can be analyzed applying the STERN-VOLLMER equation where K_D is the STERN-VOLLMER constant. To determine the dynamic quenching rate constant, k_q, fluorescence intensity or life time are measured at different quencher concentrations. For purely dynamic quenching I_0 / I or t_0 / t plotted versus the quenching molecule concentration yields a straight line with slope K_D.

Static Quenching

In contrast to dynamic quenching static quenching is marked by the formation of a complex (FQ) of the fluorophore and the quencher (Figure 9). This complexation takes place in the ground state as well as in the excited place. The excited state complex converts to the ground state mainly in a non-radiative process. Therefore, static quenching reduces the overall fluorescence intensity due to a reduction of the fluorophore concentration.

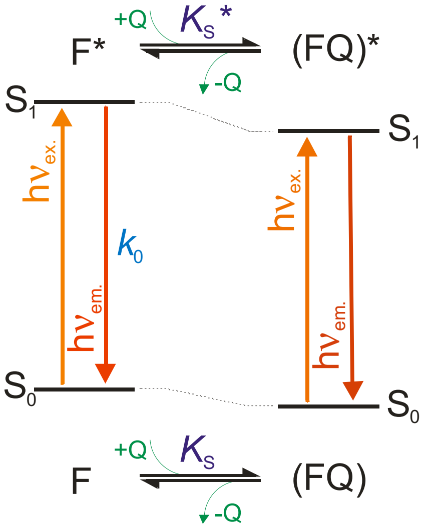


Figure 9. Static quenching. The fluorophore (F) can form a complex with a quencher molecule (Q). The complexation constants for the ground state and excited state complex formations, K_S and K_S*, determine the equilibrium of the complexation. Complex formations reduce the effective fluorophore concentration.

The fluorescence lifetime is not influenced by static quenching as the fluorescence process remains unchanged. In a STERN-VOLLMER plot K_S corresponds to the thermodynamic complexation constant for ground state complexation:

$$K_S = \frac{[FQ]}{[F] \cdot [Q]} \tag{2.24}$$

$$\frac{I}{I_0} = \frac{1}{1 + K_S \cdot [Q]} \tag{2.25}$$

A combination of steady-state and time-resolved fluorescence measurements reveals the influence of dynamic and static quenching.

2.3. Computational Chemistry

Nowadays, the demand for new application-specific chemical compounds is steadily increasing. At the same time production and research costs have to be minimized for benefit maximization. During the last decades computational chemistry became more and more established at universities as well as in chemical industry. Computational chemistry aims at the prediction of physical and chemical properties and in contrast to experimental chemistry, simulations can predict these properties for unknown substances.

In general, two main branches are distinguished between: ab initio methods that are purely based on quantum mechanics and molecular mechanics which use empirical parameters to describe a chemical system.

2.3.1. ATOMIC UNITS

If dealing with electronic calculations, a representation adapted to the properties of elementary particles simplifies notation. Charge is expressed in electron charges, mass in units of the electron mass, length as multiples of the radius of the first orbit in BOHR's treatment of the hydrogen atom (Bohr) and energy in the unit of Hartrees which correspond to the energy of two electrons separated by 1 Bohr:

$$|e| = 1.60219 \times 10^{-19}\,\mathrm{C} \tag{2.26}$$

$$m^{\mathrm{e}} = 9.10593 \times 10^{-31}\,\mathrm{kg} \tag{2.27}$$

$$1\,\mathrm{Bohr} = a_0 = \frac{\hbar^2}{4\pi \cdot m^{\mathrm{e}} \cdot e^2} = 5.29177 \times 10^{-11}\ \mathrm{m} \tag{2.28}$$

$$1\,\mathrm{Hartree} = E_{\mathrm{a}} = \frac{e^2}{4\pi \cdot \varepsilon_0 \cdot a_0} = 4.35981 \times 10^{-18}\ \mathrm{J} \tag{2.29}$$

This representation simplifies the total energy of the 1s electron in the hydrogen atom to -0.5 Hartree and the PLANCK constant, h, to 2π which turns $\hbar \equiv 1$.

2.3.2. AB INITIO METHODS

Ab initio calculations are based on quantum mechanics. A solution of the SCHRÖDINGER equation is approximated for a given molecule. Quantum mechanical calculations are usually carried out time-independently – the time-independent SCHRÖDINGER equation (2.7) is used.

As stated in 2.1.2 for a many particle system like a molecule the SCHRÖDINGER equation cannot be solved. To reduce the complexity of a molecule with N_{n} nuclei and N_{e} electrons the BORN-OPPENHEIMER approximation and the HARTREE product are used.

According to the BORN-OPPENHEIMER approximation the molecule's wave function can be split up to nucleic and electronic wave functions. The SCHRÖDINGER equation can then be written in atomic units as:

$$\left\{-\frac{1}{2}\cdot\nabla_i^{\,2}-\sum_{i=1}^{N^{\mathrm{n}}}\sum_{j=1}^{N^{\mathrm{e}}}\frac{q_i^{\mathrm{n}}}{r_j^{\mathrm{e}}-r_i^{\mathrm{n}}}+\sum_{i=1}^{N^{\mathrm{n}}}\sum_{j=i}^{N^{\mathrm{n}}}\frac{q_j^{\mathrm{n}}\cdot q_i^{\mathrm{n}}}{r_j^{\mathrm{n}}-r_i^{\mathrm{n}}}+\sum_{i=1}^{N^{\mathrm{e}}}\sum_{j=i}^{N^{\mathrm{e}}}\frac{1}{r_j^{\mathrm{e}}-r_i^{\mathrm{e}}}\right\}\Psi(s,R)=E_{\mathrm{el.}}\cdot\Psi(s,R) \tag{2.30}$$

The operator within the curly brackets is the molecular HAMILTONian. It compromises four terms: the kinetic energy of the electrons, the electron-nucleus attraction, the nucleus-nucleus repulsion and electron-electron interaction. The wave function, Ψ, depends on the spin and spatial degree of freedom, indicated by s and R, respectively.

The HARTREE product neglects electron-electron interaction. Hence, the wave function can be expanded to the product of electron-independent wave functions

$$\Psi\left(r_1, r_2, \ldots r_{N^{\mathrm{e}}}\right) = \prod_{i=1}^{N^{\mathrm{e}}} \Psi_i\left(r_i\right) \tag{2.31}$$

This functional form implies one major shortcoming because it fails to satisfy the antisymmetry principle. Fermion wave functions are antisymmetric with respect to the interchange of any set of spin-space coordinates. A way to suit this principle is to separate the space dependent part of the wave function, $\phi(R)$, from the spin dependent part, $\alpha(s)$. A spin orbital, $\chi(s,R)$ is introduced:

$$\chi(s,R) = \alpha(s) \cdot \phi(R) \tag{2.32}$$

The SLATER determinant of spin orbitals can now represent the wave function, satisfying the antisymmetry principle and assuring indistinguishability of all electrons:

$$\Psi = \frac{1}{\sqrt{N^{\mathrm{e}}!}} \begin{vmatrix} \chi_1(s_1,R_1) & \chi_2(s_1,R_1) & \cdots & \chi_{N^{\mathrm{e}}}(s_1,R_1) \\ \chi_1(s_2,R_2) & \chi_2(s_2,R_2) & \cdots & \chi_{N^{\mathrm{e}}}(s_2,R_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(s_{N^{\mathrm{e}}},R_{N^{\mathrm{e}}}) & \chi_2(s_{N^{\mathrm{e}}},R_{N^{\mathrm{e}}}) & \cdots & \chi_{N^{\mathrm{e}}}(s_{N^{\mathrm{e}}},R_{N^{\mathrm{e}}}) \end{vmatrix} \tag{2.33}$$

The assumption that the electrons can be described by the antisymmetrized product above is mathematically equivalent to assuming that each electron moves independently of all others and feels only the average COULOMB repulsion due to the positions of all other electrons and the exchange interaction due to antisymmetrization. Therefore HARTREE-FOCK is a mean field theory.

The functional form of an estimated wave function, Ψ_{Est}, can be used to calculate an energy estimation, E_{Est}, with the electronic HAMILTONian operator, $\hat{H}_{\mathrm{el}}$.

$$E_{\mathrm{Est}} = \left\langle \Psi_{\mathrm{Est}} \middle| \hat{H}_{\mathrm{el}} \middle| \Psi_{\mathrm{Est}} \right\rangle \tag{2.34}$$

Making use of the variational theorem which states that any wave function's estimate will yield a higher energy than the true energy calculated from the correct wave function, Ψ. By varying parameters within the given functional space, better approximate wave functions with energies closer to the true energy can be calculated. Thus, the correct molecular orbitals are those which minimize the approximated electronic energy best. The usual approach to obtain molecular orbitals is a linear combination of a set of given basis functions, so called atomic orbitals. These are most commonly combinations of atom-centered GAUSSian type functions as these are easily integrated over.

Besides the energy of a molecule, other molecule's physical properties, P, can be derived from the wave function with the corresponding operator, $\hat{\mathrm{P}}$:

$$P = \left\langle \Psi \middle| \hat{\mathrm{P}} \middle| \Psi \right\rangle \tag{2.35}$$

With this approach electric multipole moments or electron density distributions can be determined without the use of any empirically derived parameters The attainable accuracy is very good and errors due to various approximations made can be efficiently compensated with existing methods. The greatest drawback of ab initio calculations is their enormous demand in computational power in terms of CPU time as well as memory requirements. The practical limit for quantum mechanical calculations with modern computers lies in the order of some hundred atoms, far away from the possibility to imply the method for the investigation of biopolymers.

2.3.3. MOLECULAR MECHANICS AND DYNAMICS

In contrast to the ab initio method which is suitable for small molecules only, molecular mechanics and molecular dynamics simulations aim at the calculation of larger systems. Empiric parameters are used to describe the behavior of molecules ranging from small compounds in vacuo up to large macromolecules in a shell of solvent. Molecular dynamics simulations are one of the most commonly used tools to investigate biological molecules. They allow detailed description of fluctuations and conformational changes of proteins and nucleic acid polymers. Molecular dynamics simulations are also used for the refinement of molecular structures derived from X-ray crystallography or nuclear magnetic resonance measurements.

In molecular dynamics NEWTON's law of motion is integrated over time. The time evolution of a molecular system through phase space results a trajectory of the positions and velocities of the individual atoms constituting the system.

Molecular dynamics simulations started with a study of the interaction of 32 hard spheres performed by ALDER and WAINWRIGHT in the late 1950's [39]. This simulation delivered many important insights concerning the behavior of simple liquids. In 1964 followed the next major advance, when RAHMAN carried out the first simulation using a realistic potential for liquid argon [40]. Ten years later RAHMAN and STILLINGER performed the first molecular dynamics simulation of liquid water [41]. This work made use of the first force field describing water properties and led to a deeper understanding of the physics of glass forming liquids using the theory of inherent structures [42]. The first protein simulation appeared in 1977. MCCAMMON performed an in vacuo molecular dynamics simulation of the bovine pancreatic enzyme inhibitor [43], which lead to a new view of proteins substituting the assumption that proteins were essentially rigid structures.

During the past 30 years computationally inexpensive water models, a more realistic treatment of long-range electrostatic forces, more sophisticated force fields and last but not least the dramatic increase in computational power have made molecular dynamics simulations of biomolecules a common task. Today simulations of solvated proteins, protein-DNA complexes and membrane-lipid systems are quite common. Molecular dynamics simulation issues range from the understanding of basic thermodynamic and kinetic properties to detailed treatment of biological processes like ligand binding, small protein folding and enzymatic mechanisms.

Force Fields

Although molecules' physics are a quantum mechanical phenomena, molecular dynamics simulations are based on classical descriptions of the potentials acting on the atoms constituting the chemical system. This simplification is an approximation but due to the fact that the empirical parameters defining the potentials are derived from experimental data, they represent average properties, in sum adding up to a feasible model.

The potential energy, $V(\mathbf{r})$, can be described as the sum over the bonded and non-bonded energy terms of all atoms in the system:

$$V(\mathbf{r}) = V_{\text{bonded}} + V_{\text{non-bonded}} \tag{2.36}$$

The potential energy depends on the relative atom positions which are represented by the matrix $\mathbf{r}$.

Bonded Interaction

Bonded interactions are interactions caused by chemically bonded atoms. Most commonly used force fields include simple harmonic terms for bond length and bond angle variations and trigonometric terms for rotations about the bond between two atoms, known as torsion angle deviations.

$$V_{\text{bonded}} = V_{\text{bonds}} + V_{\text{angle}} + V_{\text{torsions}} \tag{2.37}$$

BOND LENGTH

At infinite distance in between the atoms, their potential energy is zero and as they approach each other, the energy drops to a minimum at the reference bond length. Below this distance the potential energy rises again due to the repulsion of the atoms electrons overweighting the bonding. At the sum of the atoms’ radii the potential equals the potential at infinitive distance and a further approach leads to a positive potential. A typical approximation for the potential of two encountering, attracting atoms with respect to the distance, *r*, is the MORSE potential (Figure 10, black curve).

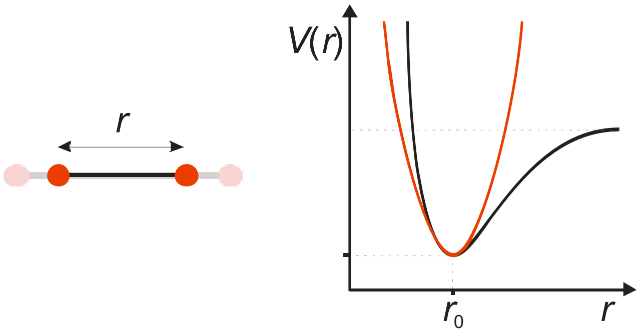


Figure 10. The potential of the bond length is approximated with a harmonic potential (red). The more exact Morse curve is shown in black.

$$V_{\text{Morse}}(r) = D^{e}\left(1 - e^{-\omega\sqrt{\frac{\mu}{2D^{e}}}\cdot\left(r - r^{0}\right)}\right)^{2} \tag{2.38}$$

The potential is defined by its depth, D^e, the reduced mass of the contributing atoms, μ, the mean bond distance, r^0, and the frequency of the bond vibration, ω. The MORSE potential is rarely used in molecular dynamics as it is inconvenient for efficient computations. Generally, a harmonic potential (Figure 10, red curve) as a simpler description of the bond length is used:

$$V(r) = \frac{k_b}{2}\left(r - r_0\right)^2 \tag{2.39}$$

The harmonic potential describes the Morse potential very well for small bond length deviations. Molecular dynamics simulations are usually performed at reasonable temperatures where bond stretching is moderate. Therefore, the harmonic potential fulfils the requirements as a force field parameter.

Both, the reference bond length, r^0, and the force constant, k^b, are specific for a pair of atoms and depend on the chemical nature of the involved atoms. Their values can be obtained from infrared stretching frequencies, microwave spectroscopy or theoretically from quantum mechanical calculations.

BOND ANGLES

The change of bond angles, θ, relative to their reference angle, θ^0, is also represented by HOOKE's law (Figure 11):

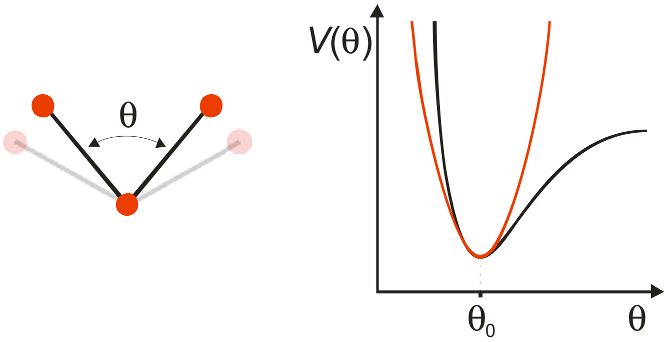


Figure 11. Bond angle variation is described by a harmonic potential (red).

$$V(\theta)=\frac{k^{\phi}}{2}\left(\theta-\theta^{0}\right)^{2} \tag{2.40}$$

Compared to bond vibrations, the angle deformation occurs at lower frequencies as the force constant is lower. Less energy is required to distort a bond angle from its reference value than to stretch or compress a bond.

TORSION ANGLES

Steric interaction, conjugation or hyperconjugation restrict substituents bond to two bonded atoms to certain relative positions. The angle between the normal vectors of the planes spanned by A, B and C and B, C, and D in the molecular fragment A-B-C-D is defined as the torsion angle. A torsion angle of 0° represents co-planarity of the four atoms with A and D heading to the same side of the plane, the torsion angle is 180° if A and D head to opposite directions in the plane.

As the rotation around a bond yields a repeating conformation every 360°, the potential of torsions is described by trigonometric functions.

$$V(\phi)=K^{\phi}\cdot\left(1-\cos\left(n\cdot\phi-\phi^{0}\right)\right) \tag{2.41}$$

The reference torsion angle is denoted by K_{ϕ}, n is the multiplicity, i.e. the number of minima and maxima per rotation, and ϕ^0 is the angular offset of the potential. A set of cosine functions is used to model the influence of multiple moieties influencing each other. Combining different potential functions with different multiplicities and amplitudes results in complex shapes for the rotational energy barrier. Figure 12 shows the addition of two potentials (black curves) to form a more accurate function.

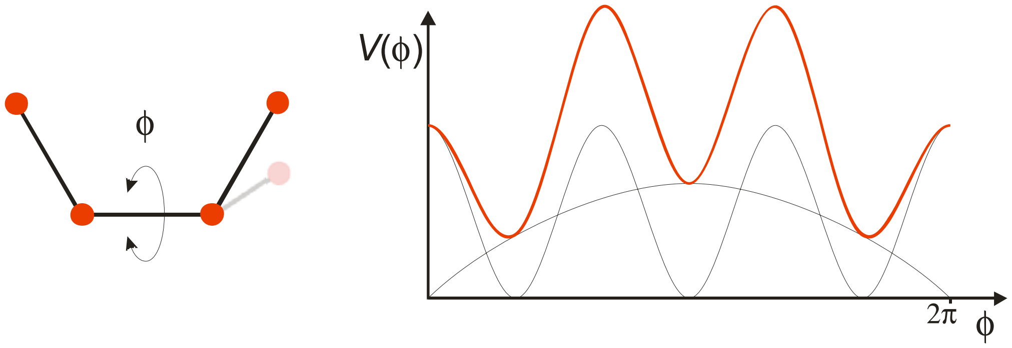


Figure 12. Dihedral angles potentials (red) are represented by a combination of trigonometric functions (black).

Non-bonded Interaction

The influence of atoms separated from each other by more than three bonds is called non-bounded interaction. The potential energy term representing its contribution has basically two components: the van der Waals and the electrostatic interaction energy:

$$V_{\text{non-bonded}} = V_{\text{vdW}} + V_{\text{el}} \tag{2.42}$$

In molecular dynamics the calculation of non-bonded interactions is the most time consuming part because basically the energy terms have to be calculated for every atom combination. The simultaneous interaction of three or more atoms is not calculated as this constitutes a many body problem. Furthermore certain polarization effects are not accounted for in all force fields.

Van der Waals interaction

Like bonded atoms, non-bonded atoms interact when approaching each other. Attractive forces are induced by dipole-dipole interactions scaling with the distance, *r*, to the power of six. At proximity repulsive forces arising from the penetration of the electron clouds act on the atoms. This potential can be approximated as the inverse of the distance to the power of twelve. A Lennard-Jones 6-12 potential is therefore used to model van der Waals interactions:

$$V(r) = 4\varepsilon \cdot \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^{6} \right] \tag{2.43}$$

The LENNARD-JONES potential (Figure 13, black curve) contains two parameters: the collision diameter, σ, which describes the distance of the atoms where attractive and repulsive forces neutralize each other, and the well depth, ε. There is plenty of evidence that the approximation of the repulsive term is a poor representation and an exponential form, e^{-r/r_0}, would represent the nature of electron-electron interactions much better. This exponential function would account for the exponential decay of the radial part of the atom's wave function and hence describe the electron distribution of the atoms in a more concise fashion.

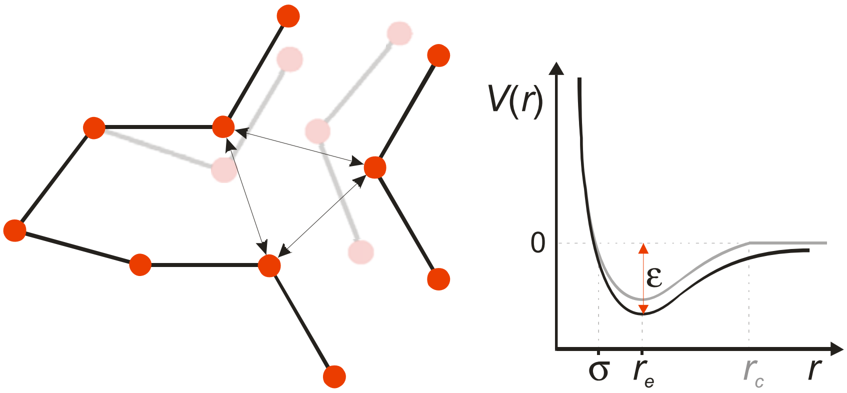


Figure 13. VAN DER WAALS interaction is represented as a LENNARD-JONES potential. A shifted function can be used to realize a cut-off criterion: beyond the cut-off distance, r_c, the potential is zero.

As stated above, VAN DER WAALS potentials have to be calculated for every atom pair within a modeled system. As the potential converges to zero rapidly ($V \propto r^{-6}$) a cut-off distance, r_c, is used in most molecular dynamics simulations. From the cut-off distance the VAN DER WAALS interaction is omitted. This can be visualized by a shift of the function towards positive potential (Figure 13, grey curve).

ELECTROSTATIC INTERACTION

The accurate reproduction of the electrostatic properties of a molecular system is of fundamental importance in force field development. Electrostatic forces are magnitudes of order greater than VAN DER WAALS interactions and therefore they have significant influence on the behavior of the system.

The electron distribution within a molecule is defined by its wave functions. As wave functions or electron density distributions are unfashionable to imply in molecular dynamics simulations a point charge model is commonly used. The charge distribution is resembled by atom centered point charges and the problem of electrostatic interaction is reduced to the calculation of COULOMB forces between atom pairs:

$$V(r) = \frac{1}{4\pi\varepsilon} \cdot \frac{q_1 \cdot q_2}{r} \tag{2.44}$$

The potential depends on the charge of the atoms, q_1 and q_2, the electric permittivity of the medium, ε, and the atoms' distance, r. For alike charges the potential is repulsive and therefore positive (Figure 14, red curve), opposite charges attract and induce a negative potential (Figure 14, black curve).

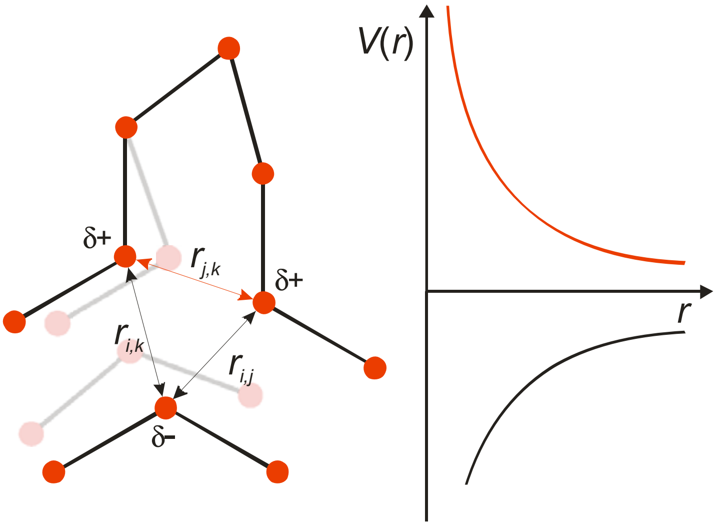


Figure 14. Electrostatic interactions can be attractive for opposite charges (black) or repulsive for alike charges (red).

The electrostatic interactions decay with the inverse of the distance. Therefore, the use of cut-off methods is not appropriate in most case and is significantly error prone.

Energy Minimization

The above described force field can be applied to a molecular system. The potential energy of the system results in:

$$\begin{aligned} V(\mathbf{R}) = {} & \sum_{i \in B} D_i^e \left(1 - e^{-\omega\sqrt{\frac{\mu_i}{2D_i^e}} \cdot \left(r - r_i^0\right)}\right)^2 + \sum_{i \in A} \frac{k_i^\phi}{2}\left(\theta_i - \theta_i^\phi\right)^2 + \sum_{i \in T} K_i^\phi \cdot \left(1 - \cos\left(n_i \cdot \phi_i - \phi_i^0\right)\right) \\ & + \sum_{i=1}^{N} \sum_{j=i+1}^{N} 4\varepsilon_{i,j} \cdot \left[\left(\frac{\sigma_{i,j}}{r_{i,j}}\right)^{12} - \left(\frac{\sigma_{i,j}}{r_{i,j}}\right)^{6}\right] + \sum_{i=1}^{N} \sum_{j=i+1}^{N} \frac{1}{4\pi\varepsilon} \cdot \frac{q_i \cdot q_j}{r_{i,j}} \end{aligned} \tag{2.45}$$

The sets B, A, and T are the sets of bonds, bond angles, and torsion angles, respectively. For all but the simplest systems the potential energy is a complicated, multidimensional function of the $3N$ Cartesian coordinates of the system. The energy surface can be searched for minima which are of special interest as these represent stable states of the molecular system.

Given a function *f* which depends on the variables $\{x_1, x_2, x_3, \ldots, x_n\}$, a minimum is defined as the point where the first derivatives of the function with respect to each of the variables is zero and the second derivatives are all positive:

$$\frac{\partial f}{dx_i} = 0; \quad \frac{\partial^2 f}{dx_i^2} > 0 \qquad \forall i = [1, 2, 3, \ldots, n] \tag{2.46}$$

A multidimensional potential function cannot be derived using standard calculus methods. Hence, in molecular mechanics an energy minimization is performed by following the energy landscape towards lower and lower energies. This algorithm implies a gradual change of the atoms' coordinates until a further change will not result in a significant energy drop. Two major issues must be considered in the development of minimization algorithms: the speed at which they converge to a minimum and the memory required for the search. Up to now no single algorithm has shown to outperform any other. Most commonly a combination of algorithms is used.

The steepest descent algorithm follows the energy landscape with the steepest slope towards the energy minimum. The step size depends on the energy difference achieved by the former step: If the energy has been lowered, the step size is increased. If it has increased, it is assumed that the step has leapt across the minimum and the step is re-performed with a fraction of the step size. This procedure is repeated until the change in energy falls below a defined threshold. The steepest descent method is a simple algorithm that works fine at relatively steeply shaped energy potentials. Near the minima it tends to leap these increasing time taken to converge.

The NEWTON-RAPHSON algorithm is a second-derivative algorithm. That is, not only the slope but also the curvature of the potential surface is taken into account when defining a minimization step. The energy potential can be written as a TAYLOR series expansion about the point $\mathbf{R}_0$:

$$V(\mathbf{R}) = V(\mathbf{R}_0) + (\mathbf{R} - \mathbf{R}_0) \cdot V'(\mathbf{R}_0) + \frac{1}{2}(\mathbf{R}_0)(\mathbf{R} - \mathbf{R}_0)^2 \cdot V''(\mathbf{R}_0) + \ldots \tag{2.47}$$

As already stated before, $\mathbf{R}$ and $\mathbf{R}_0$ are sets of *N* atomic coordinates describing the system and can be described as a 3*N*-dimensional vector. The first derivative with respect to the 3*N* variables, *V*'', is a 3*N*-dimensional vector, too. The second derivative is a 3*N* by 3*N* matrix containing the partial derivative in any pair of variables and is called the Hessian matrix. In proximity to the minimum the series can be truncated at the second order. The first derivative of the energy function describes the direction of the slope towards the minimum whereas the second derivative describes the curvature and therefore predicts where the function will change direction and what step size should approach the minimum best.

Compared to the steepest descent algorithm the NEWTON-RAPHSON method requires far more memory as the Hessian matrix has to be calculated and stored. The size of the matrix scales with N^2 whereas the first derivative scales with *N*. A commonly used method to find the minimum energy is to approach the minimum using the steepest descent algorithm for a certain number of steps and then change to the NEWTON-RAPHSON variant to encounter the minimum more effectively.

Molecular Dynamics Simulations

In molecular dynamics simulations the classical time evolution of a molecular system is calculated which seems trivial at the first glance. NEWTON's second law is integrated over time.

$$F_i = m_i \cdot a_i \qquad \forall i = [1, 2, 3, ..., n] \tag{2.48}$$

The force, F_i, exerted on an atom equals the product of its mass, m_i, and the acceleration, a_i, it experiences. Force can also be expressed as the negative gradient of the potential energy:

$$F_i = -\nabla V \qquad \forall i = [1, 2, 3, ..., n] \tag{2.49}$$

Combining (2.48) and (2.49) leads to a description of the motion caused by the potential:

$$\frac{d^2 x_i}{dt^2} = \frac{-\nabla V}{m_i} \qquad \forall i = [1, 2, 3, ..., n] \tag{2.50}$$

For “realistic” molecular systems no analytical solution of (2.50) can be derived and the equations have to be solved numerically. Computer based numerical calculations can only act discontinuously whereas the system reacts continuously in nature. This implies that the trajectory of the system can only be built by stepping through time. At each step with common time lengths of 1 - 2 fs the forces are calculated as the vector sum of the atoms' interaction with other atoms in a pair-wise additive model and the resulting accelerations are assumed to be constant for the following time step length. The atoms are then moved according to their initial velocities and the calculated accelerations. In this manner molecular dynamics simulations generate a trajectory that describes how the system' internal degrees of freedom are changed as a function of time.

Several efficient algorithms are known for integrating the equations of motion. They all use TAYLOR series expansions of the positions and dynamic properties:

$$r\left(t+\delta t\right)=r\left(t\right)+\delta t\cdot\frac{dr}{dt}+\frac{1}{2}\cdot\delta t^{2}\cdot\frac{d^{2}r}{dt^{2}}+\frac{1}{6}\cdot\delta t^{3}\cdot\frac{d^{3}r}{dt^{3}}+\ldots \quad (2.51)$$

One of the most frequently used algorithms is the VERLET algorithm [44]. It predicts the positions of the atoms by adding the future positions at t = t + δt and the past positions at t = t - δt.:

$$r\left(t+\delta t\right)+r\left(t-\delta t\right)\approx 2\cdot r\left(t\right)-r\left(t-\delta t\right)+\delta t^{2}\cdot\frac{d^{2}r}{dt^{2}} \quad (2.52)$$

The algorithm seems to cut off the TAYLOR series expansion with the acceleration term but it accounts for the third derivative of the atoms' position, too. This term is eliminated by the addition of the TAYLOR series. The velocities do not explicitly add to the VERLET integration (2.52) but they can be calculated either from the position at t = t + δt and t = t + δt as the slope of the straight through the two points of by averaging the velocities at the given times.

After having run the molecular dynamics simulation, a trajectory of the modeled system's atom coordinates versus time is achieved. From this trajectory dynamic parameters can be evaluated. Additionally, stable conformations show as heavily populated states over time. They can be analyzed and visually displayed

2.4. Biopolymers

Throughout nature any living organism depends on the function of biopolymers. DNA is used to store the information of life, it is transcribed to RNA, which is translated to proteins. Proteins are the molecular machines that transport, synthesize or communicate. In the next sections a short introduction to nucleic acids and amino acids – with a focus on proline – is given.

2.4.1. NUCLEIC ACIDS

Nucleic acids divide into two groups: deoxyribonucleic acids (DNA) and ribonucleic acids (RNA). DNA is used to store genetic information as polymers with alternating nucleobases as the carrier of information (Figure 15). DNA is unique in biochemistry as it is a quasi self-replicating molecule.

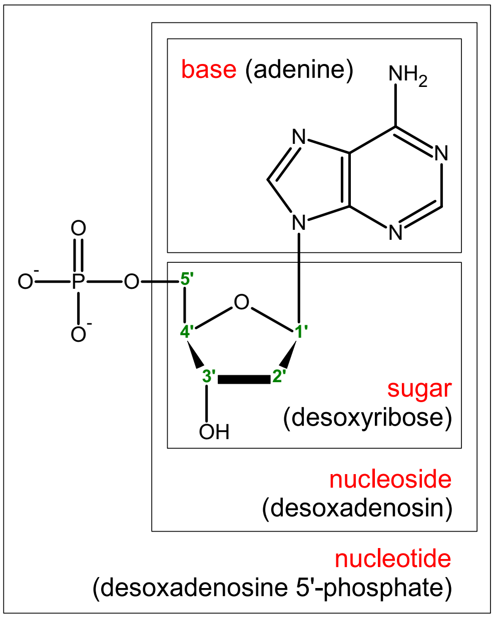


Figure 15. Chemical structure of nucleosides. One of four bases is linked to the sugar and forms the nucleoside. A phosphate at the 5' end forms the nucleotide. In DNA desoxyribose acts as the sugar backbone, whereas in RNA it is ribose with an additional hydroxyl group at $C^{2'}$.

DNA is structured in three subunits: the nucleobases are bound to the $C^{1'}$ carbon atom of the 2-deoxyribose unit. The desoxyribose is a condensed form of the sugar ribose, missing the hydroxyl group at $C^{2'}$. The $C^{5'}$ group is esterified with phosphate.

In DNA the 5’-deoxyribose phosphate forms polymers by a condensation reaction of the $C^{3'}$ hydroxyl group with the next nucleotide’s phosphate group.

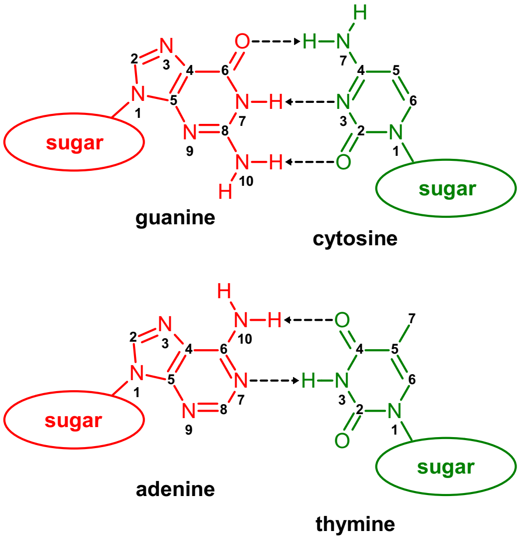


Figure 16. Chemical structure of nucleobases. Guanosine and cytosine form three hydrogen bonds, adenine and thymine two in double-stranded DNA.

Four bases are used in naturally occurring DNA: adenosine (dA), cytidine (dC), guanosine (dG), and thymidine (dT). The bases thymine and cytosine are pyrimidine derivatives, adenine and guanosine are purine derivatives.

Polymeric DNA forms dimers with matching counter-strands. In this double-stranded DNA adenine pairs with thymine and cytosine with guanosine. The first complex is formed by two hydrogen bonds, the latter by three (Figure 16).

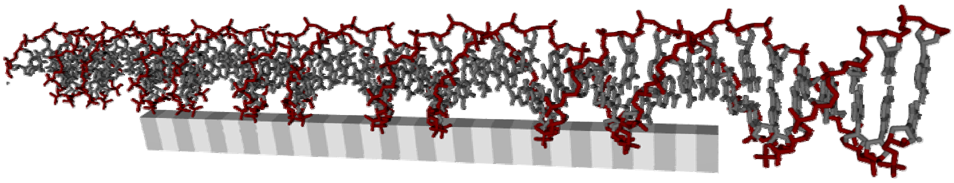


Figure 17. Double-stranded DNA. A 60-mer randomly sequenced DNA-strand and its reversed-complimentary sequence in the B double helix form. Backbone atoms are coulored red. A virtual rules with a total length of approximately 20 nm is shown next to the molecules.

Double-stranded DNA (ds-DNA) helices are right-handed and approximately 2 nm in diameter. The helical rise in the naturally occurring B-conformer is 0.34 nm per residue, the helical twist 36°. Therefore, B-DNA helices repeat every 10 residues and show a helical pitch of 3.4 nm (Figure 17). Due to the intensive hydrogen bonding between the two strands and aromatic π-π interactions between nucleobases B-DNA polymers are rigid molecules with a persistence length of ~ 50 nm (see 2.4.3 for definition).

Single stranded DNA (ss-DNA) does not form internal hydrogen bonds to a significant extend. Thus, ss-DNA is far more flexible than ds-DNA and its persistence length ranges about one order of magnitude below ds-DNA's (Figure 18).

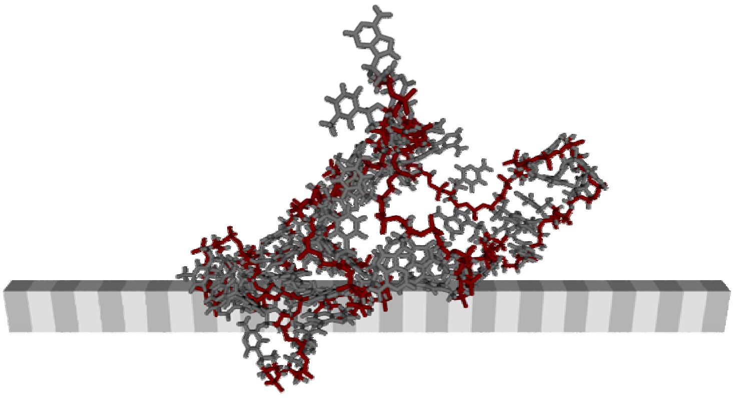


Figure 18. Single-stranded DNA. The same 60-mer randomly sequenced DNA as for the ds-DNA was modeled by simulated annealing. Next to the molecule a virtual ruler with 20 nm total length is shown.

2.4.2. PROTEINS AND PEPTIDES

Despite DNA amino acids form biopolymers as well. Whereas in nucleic acids sugar phosphates build the polymer backbone, in peptides L-α-amino acids are joined to long chains.

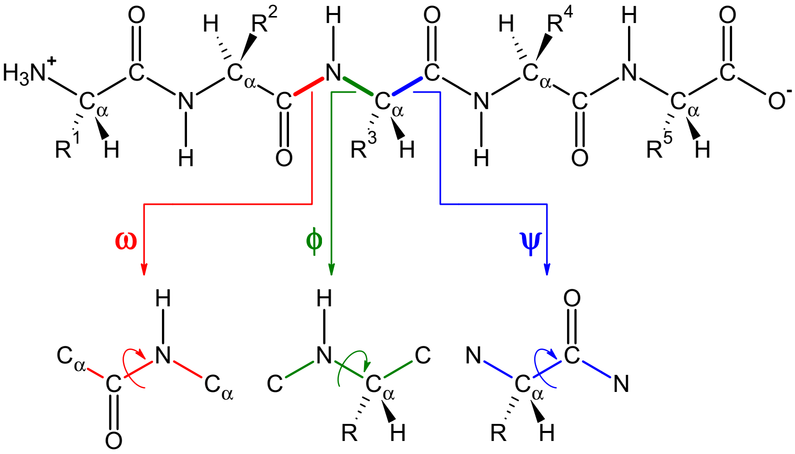


Figure 19. Dihedral angles in peptides. Rotations around the CN-bond (red), NC-bond (green), or CC-bond are specified by the dihedral angles ω, ϕ, and ψ, respectively.

This chain consists of repeating *N*-C-CO elements with side groups connected to the central carbon atom. There exist 20 naturally occurring amino acids with different side chains. All of these amino acids except proline carry a primary amino group at the central α-carbon atom which binds to the carbonyl carbon of the preceding amino acid in peptides. The peptidic amide bond and the exceptional structure of proline shall be discussed in more detail in the following sub-chapters.

The structure of the amino acid polymer backbone can be described by dihedral angles as bond lengths and angles are narrowly distributed. Three types of dihedrals exist in the peptide backbone: ω, ϕ, and ψ (Figure 19).

Peptide Bonds

Amide bonds that are formed with the polymerization of amino acids contain three atoms which can electronically be described as sp^2 hybridized. The 2s-orbital and two 2p-orbitals combine to form three degenerated orbitals with C_{3h} symmetry. These hybrid orbitals can be described by linear combinations of the specified atom orbitals and they are used to form σ-bonds to neighboring atoms. The third p-orbital can be used to form π-bonds between the three atoms (Figure 20).

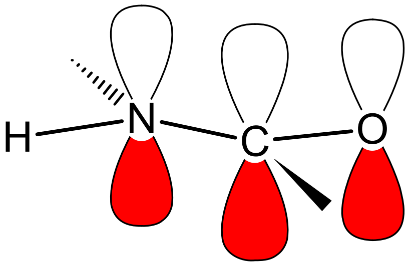


Figure 20. Peptide amide bond. The three p_Z-orbitals have to be oriented parallel for optimal overlap.

The amide group cannot be described correctly as a *N*-C=O system. The electron lone pair of the nitrogen can interact with the π*-orbital of the carbonyl group. Thus, increasing the electron density of the π*-orbital, the CO-bond is destabilized and on the other hand the CN-bond gains double bond character.

Any conformation deviating from the p-orbitals' parallel orientation will increase the system's energy significantly. Therefore, two stable conformations can be achieved when rotating the CN-bond. If the adjacent α-carbon atoms both face towards the same side the peptide bond is described as a *cis*-bond, with α-carbons in opposition peptide bonds are denoted *trans* (Figure 21).

The *trans* isomer is preferred over *cis* due to steric interaction of the polymeric chain in the *cis* conformation. The backbone atoms are spatially more demanding than hydrogen and carbonyl oxygen. For non-proline amino acids the energy difference between *cis* and *trans* lies in the range of 13 kJ·mol^{-1}. According to the BOLTZMANN distribution this implies a ratio between *cis* and *trans* of approximately 0.5 % [45].

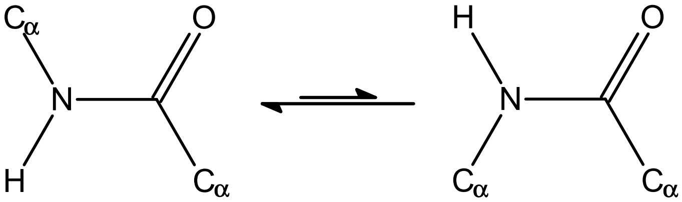


Figure 21. *cis* / *trans* Isomerisation of peptide bonds. Opposing a-carbon atoms define a *trans* amide bond, α-carbons facing the same side are denoted *cis*.

Prolines

The amino acid proline is special with respect to its structure. It is the only naturally occurring amino acid with a secondary amine group. The side chain forms a five membered ring from the a-carbon to the nitrogen (Figure 22).

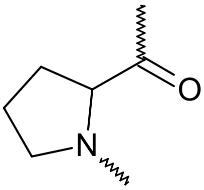


Figure 22. Proline is the only genetically encoded aminoacid with a cyclic structure.

The covalent bond between the side-chain and the nitrogen backbone atom has relevant structural consequences on the properties of the residue. In particular, (a) the cyclic nature of proline imposes severe restrictions to the conformational freedom of this residue, (b) there is a strict dependence between main- and side-chain proline conformations, (c) the *cis* state of the peptide group preceding proline residues is significantly populated, and (d) the absence of a backbone hydrogen bonding donor limits proline capability to form main-chain hydrogen bonds.

The conformational freedom of proline’s to backbone dihedral angles is restricted to $\phi \approx -75$ and $\theta \approx 145°$. The five membered ring fixes the backbone which leads to a coupling of main- and side-chain: a twist of the dihedral angle ϕ leads to a flip of the preferred ring conformation. Two ring conformations are energetically favorable: up and down puckering. Furthermore, due to the cyclic side-chain prolyl amide bonds, i.e. the amide bonds preceding a proline residue in a polymer, are less fixated to the *trans* conformation. This is because of the fact that in a proline residue the energetic difference between *cis* and *trans* is decreased as in both conformations methylene groups flank the α-carbon of the preceding residue equaling steric strains in the two isomers.

Polyprolines form regular helices: all-*trans* isomers arrange in form of the left-handed polyproline type II helix with C_3 symmetry with a helical rise of 3.1 Å per residue. All-*cis* polyprolines form a more compact, tightly wound right-handed helix, named polyproline type I helix with a helical rise of 1.9 Å per residue and a non-integer periodicity. Both helices are stabilized solely by main-chain dihedrals limitations, no hydrogen bonding contribute to their formation.

2.4.3. MATHEMATICAL DESCRIPTIONS OF BIOPOLYMERS

Due to their regularity polymers are macromolecules that can be analytically described with simpler models than on the atomic scale. A system with a large number of locally interacting identical constituents on a far bigger scale than the size of the constituents can be simplified to just a few phenomenological parameters.

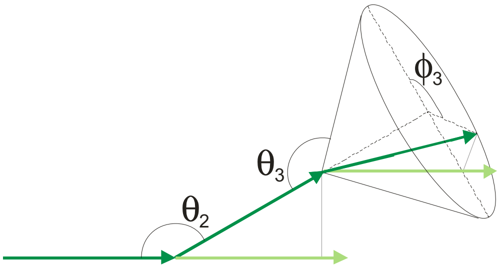


Figure 23. Spatial representation of polymer segments. The polymer segments (green vectors) are freely jointed under an angle, θ. The first two vectors form a plane, the third vector can adopt any orientation towards the circular base of a cone. The spatial orientation of the third vector is defined by the dihedral angle of the three vectors, ϕ.

A small number of effective degrees of freedom describe the system's behavior. Generally, polymers can be described as either a flexible or a semi-flexible chains.

Segments of chains can be described as vectors (Figure 23). The vector heads are represented by the pointing vectors, $\mathbf{r}_i$, the origin of the chain is given by $\mathbf{r}_0$. Thus, the segment vector, $\mathbf{s}_i$, is given as the difference of $\mathbf{r}_i$ and $\mathbf{r}_{i-1}$. The length of the segments, l_i, is then

$$l_i = |s_i| \tag{2.53}$$

For a N segments containing chain the angle between any subsequent two segments, θ, can be derived from the scalar product.

$$\cos\theta_i = \frac{s_i}{|s_i|} \cdot \frac{s_{i-1}}{|s_{i-1}|} \qquad \forall\, i \in [2, 3, ..., N] \tag{2.54}$$

The dihedral angle, ϕ, describes the twist of the third segment vector relative to the plane spanned by the prior two vectors.

$$\cos\phi_i = \left(\frac{\mathbf{s}_i}{|\mathbf{s}_i|} \times \frac{\mathbf{s}_{i-1}}{|\mathbf{s}_{i-1}|}\right) \cdot \left(\frac{\mathbf{s}_{i-1}}{|\mathbf{s}_{i-1}|} \times \frac{\mathbf{s}_{i-2}}{|\mathbf{s}_{i-2}|}\right) \qquad \forall\, i \in [3, 4, ..., N] \tag{2.55}$$

Besides the microscopic entities polymer chains also show phenomenological parameters on a larger scale: the contour length is the length of the polymer chain, i.e. the length of a segment multiplied by the number of segments:

$$L_\mathrm{c} = N \cdot l \tag{2.56}$$

The end-to-end distance is defined as the length of the vector from the origin of the chain to the head of the last segment vector:

$$L_\mathrm{e} = |r_N - r_0| \tag{2.57}$$

The geometrical center of the polymer is the average position of the segments:

$$\mathbf{r}_\mathrm{c} = \frac{\sum_{i=1}^{N} \frac{1}{2} \cdot (\mathbf{r}_i + \mathbf{r}_{i-1})}{N} \tag{2.58}$$

The center of gravity matches the geometrical center for segments with equal mass distribution. For real polymers it deviates. It is defined as

$$\mathbf{r}_m = \frac{\sum_{i=1}^{N} \frac{1}{2} \cdot (\mathbf{r}_i + \mathbf{r}_{i-1}) \cdot m_i}{\sum_{i=1}^{N} m_i} \tag{2.59}$$

The radius of gyration is calculated as the root mean square distance of the segments from the polymer's center of gravity:

$$l_\mathrm{g} = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^{N} (\mathbf{r}_i - \mathbf{r}_\mathrm{m})^2} = \sqrt{\frac{1}{2 \cdot N^2} \cdot \sum_{i=1}^{N} \sum_{j=1}^{N} (\mathbf{r}_i - \mathbf{r}_j)^2} \tag{2.60}$$

The radius of gyration can either be derived from the segments' distance to the center of gravity (first term in (2.60)) or from the differences of any two segments' positions (right term in (2.60)). The second definition only holds if all segments are of the same mass.

Finally, the persistence length of a polymer can be specified. It is a parameter for the stiffness of a polymer. Either the (average) projection of the segment vectors onto the n^{th} preceding vector is analyzed. Figure 23 shows three consecutive segment vectors (dark green) and their projections onto the first (gray lines). The function is described by an exponential decay function, C_n [46]:

$$C_n = \left\langle \cos\theta_n \right\rangle = C_0 \cdot e^{-\frac{n \cdot l}{L_\mathrm{p}}} \tag{2.61}$$

From a half-logarithmic plot of the projection lengths versus the number of separating segments the persistence length can be evaluated.

Alternatively, the persistence length can be accessed as the projection length [47].

$$L_\mathrm{p}(i,n) = \frac{\left\langle \mathbf{s}_i \cdot \left(\mathbf{r}_n - \mathbf{r}_0\right)\right\rangle}{\sqrt{\left\langle \mathbf{s}_i^2 \right\rangle}} \tag{2.62}$$

Freely Jointed Chain

A freely jointed chain can be described as a series of cylinders flexibly connected at their ends. This chain can adopt any conformation as the only restriction is the length of the segments. The angles between any two segments and the dihedral angles underlie no restrictions. Therefore the chain is penetrable and able to coil up. The shape of the chain is then controlled by statistics only (Figure 24).

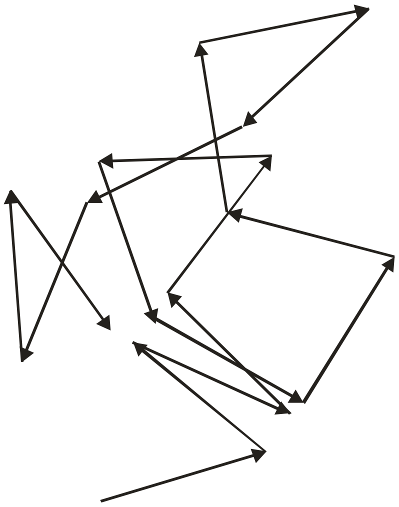


Figure 24. A possible structure for a freely jointed chain in two dimensions.

The structure of a freely jointed chain can be simulated by a random walk with constant step size.

Wormlike Chain

The wormlike chain model describes a flexible polymer chain as a continuous elastic medium. This implies that the polymer's flexibility is uniform throughout the chain (Figure 25). This obviously can only be assumed if the repeating structural unit is magnitudes of order smaller than the persistence length and the angle spanned by two consecutive segments is small. The wormlike chain model can then be seen as a special case of the freely jointed chain model.

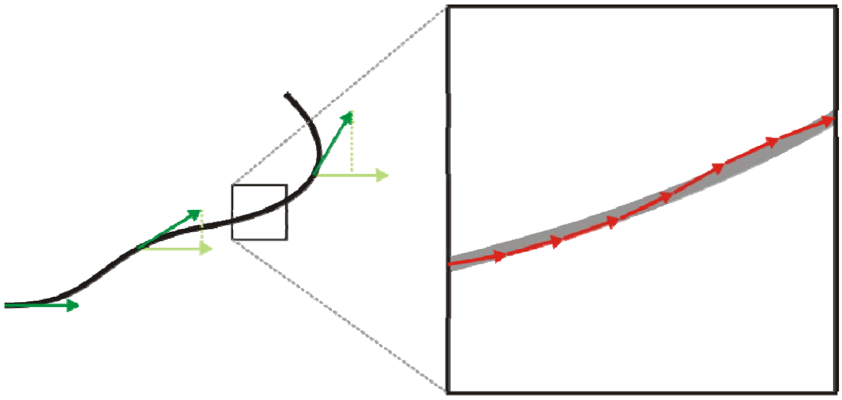


Figure 25. A polymer following the wormlike chain model. On a larger scale the chain seems to be uniformly flexible (left) as the chain segments are small compared to the persistence length and the deviation angles are minimal. A magnification shows the chain segments (right).

2.5. Fluorescence Correlation Spectroscopy

Macroscopic systems in equilibrium show no measurable fluctuations when investigated by ensemble analysis. Dynamic processes cannot be observed as with the multitude of observed molecules any individual molecule's deviation from the mean is compensated by another. Ensemble measurements can only show average properties of molecules. The collective deviation from the mean results in noise without significance. If the observation volume in a measurement is reduced, the number of observed molecules is reduced as well and the individual molecule's deviation from the mean can be resolved and interpreted. The noise in ensemble measurements becomes a parameter describing characteristic processes proceeding in a system at equilibrium. Ensemble measurements show thermodynamics whereas single-molecule measurements show kinetics.

Due to significant improvement in the field of optics during the last decades microscopes are nowadays able to detect photons emitted by single molecules. Such sensitive microscopes opened the way for fluorescence correlation spectroscopy (FCS). The basic principle of FCS is the excitation of fluorescing molecules with a laser beam followed by the detection of the emitted light. To realize the detection of only a few molecules at a time the excitation volume is reduced by focusing the laser beam and using the same optical path for excitation and emission light. Due to the Stokes shift the collected light has a higher wave length and can be separated by a dichroic filter and further be detected by highly sensitive avalanche photo diodes (APDs). With this setup nanomolar solutions of substances to be analyzed are measured for a period of time. Either by software or using a hardware correlator, data is analyzed such that any measured fluorescence is compared to any data point with a temporal pitch, yielding a distribution of likelihood to see fluorescence at a given time after a former event.

If the concentration was low enough so that only one molecule would pass the observation volume at a time, BROWNian motion would determine the average travelling time of that molecule through the detection volume. While the fluorophore resided within this volume it would emit fluorescence photons unless any internal mechanism suppresses the fluorescence.

2.5.1. LASER FOCUS

In order to achieve an excitation volume small enough such that only a limited number of molecules reside within at any time, a confocal setup can be used. In an inverted microscope the excitation laser light is collected by an objective. At the objective focal plane the laser beam is focused into a spot with dimensions according to the diffraction limit. The nature of the laser focus is determined by the point spread function, *PSF*.

In confocal microscopy the point spread function is a convolution of the effects of a number of physical components, such as the laser beam, the microscope objective, optical fibers, and the detectors. The point spread function can be approximated by a three-dimensional GAUSSian with a $1/e^2$-beam-waist-diameter of $2 \cdot \omega_0$ and a length $2 \cdot z_0$ along the optical axis (Figure 26).

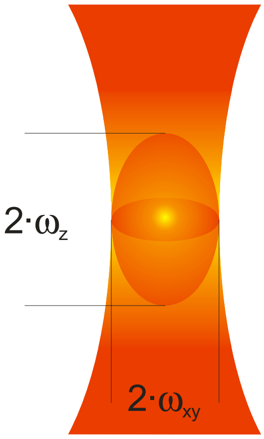


Figure 26. Focal region of a laser beam. The 1/e^2 ellipsoid and its dimensions are shown.

$$PSF(x,y,z)=e^{-\frac{2x^2+2y^2}{\omega_0^2}-\frac{2z^2}{z_0^2}} \tag{2.63}$$

$$PSF(r,\phi,z)=e^{-\frac{2r^2}{\omega_0^2}-\frac{2z^2}{z_0^2}} \tag{2.64}$$

2.5.2. AUTOCORRELATION

If a detection volume is constantly excited the measured fluorescence, $F(t)$, shows temporal fluctuations in intensity:

$$\delta F(t)=F(t)-\langle F(t)\rangle \tag{2.65}$$

$$\langle F(t)\rangle=\frac{1}{T}\int_0^T F(t)\cdot dt \tag{2.66}$$

The temporal fluorescence fluctuation function can be described as a normalized autocorrelation function, $G(\tau)$, which reflects the temporally averaged self-similarity of a signal at the moments t and $t+\tau$.

$$G(t)=\frac{\langle \delta F(t)-\delta F(t+\tau)\rangle}{\left(\delta F(t)\right)^2} \tag{2.67}$$

Due to the reduced observation volume $G(t)$ reflects the probability to re-detect a molecule detected at the point in time, t, at a later time, $t+\tau$. Therefore, $G(t)$ is the retention period distribution of single molecules within the detection volume.

Equation (2.67) can be approximated if a two-dimensional diffusion model is used. For a single species in combination with an exponential decay accounting for photophysics, the correlation function can be described as:

$$G(\tau) = \frac{1}{O} \cdot \left(1 + \frac{\tau}{\tau_\mathrm{D}}\right) \cdot \left(1 + A \cdot e^{-k\tau}\right) \tag{2.68}$$

The correlation function then depends on the number of detected molecules, *O*, the amplitude of the signal, *A* and the rate constant, *k*, describing photophysics.

2.5.3. TRANSLATIONAL DIFFUSION

One aspect of FCS measurements is to analyze the translational diffusion properties of examined species. The small detection volume used in FCS acts as a window to the microscopic behavior of molecules in BROWNian motion describing the macroscopic phenomenon of diffusion.

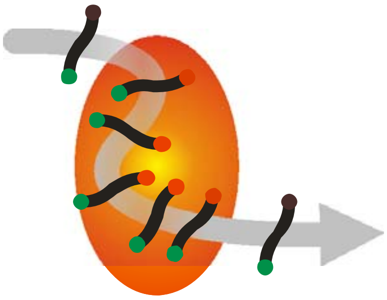


Figure 27. Schematic representation of fluorophore (red) and quencher (green) labeled molecule traversing the focal volume showing pure translational diffusion without photophysics.

As the fluorescence intensity is proportional to the number of light emitting molecules within the observation volume, its fluctuation reflects the translational diffusion of the molecule into and out of the detection volume (Figure 27). Changes in fluorescence intensity are thus due to local concentration variations of the fluorescent species. This concentration fluctuation can be expressed by the diffusion equation:

$$\frac{\partial\, \delta C(\vec{r},t)}{\partial t} = D \cdot \nabla^2 \cdot \delta C(\vec{r},t) \tag{2.69}$$

2.5.4. CONFORMATIONAL DYNAMICS

Apart from diffusing through the excitation volume, a molecule's fluorescence also reflects internal dynamics (Figure 28). As shown in 2.2.2 fluorescence depends on the local environment of an aromatic moiety. These local surroundings can be influenced by bimolecular interactions if the concentration of the fluorescing species is high enough. In typical FCS measurements the concentration of the surveyed species is so low that intermolecular interactions can be neglected. Still the fluorophore's environment can change due to intramolecular rearrangements. Electronic processes like PET or FRET can cause the quantum yield of a fluorophore to vary on a timescale several magnitudes of order faster than its translational diffusion.

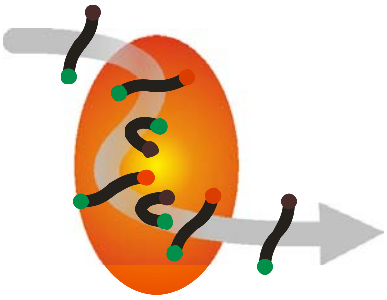


Figure 28. Schematic representation of fluorophore (red) and quencher (green) labeled molecule traversing the focal volume. Conformational changes lead to temporal fluorescence quenching.

Regarding a two-state model of a molecule undergoing a conformational change leading to a change in fluorescence quantum yield, molecules in state A will convert to state B at the rate k_{AB}, whereas molecules in state B return to state A at the rate k_{BA}.

$$\mathrm{A} \underset{k_{BA}}{\overset{k_{AB}}{\rightleftharpoons}} \mathrm{B}$$

If the conformer A shows fluorescence and conformer B is less fluorescent due to internal photophysics, their quantum yields would be Φ_A and Φ_B, respectively. Assumed that the translational diffusion time does not change due to the conformational change, hence $D_A = D_B = D$, the diffusion equation (2.69) expands to:

$$\frac{d}{dt}\begin{pmatrix} \delta C_A(\vec{r},t) \\ \delta C_B(\vec{r},t) \end{pmatrix} = D\nabla^2 \begin{pmatrix} \delta C_A(\vec{r},t) \\ \delta C_B(\vec{r},t) \end{pmatrix} + \begin{pmatrix} -k_{AB} & k_{BA} \\ k_{AB} & -k_{BA} \end{pmatrix} \begin{pmatrix} \delta C_A(\vec{r},t) \\ \delta C_B(\vec{r},t) \end{pmatrix} \tag{2.70}$$

While the first term on the right hand side determines the translational diffusion as for the one-state system in equation (2.69), the second term describes the internal kinetics.

2.5.5. PHOTOPHYSICS

A molecule traversing the focal volume in FCS measurements can undergo photophysical effects during its retention period within the focus. A transition of the excited molecule to a triplet state can lead to short periods on the time scale of milliseconds to seconds (Figure 29) due to the triplet state's long lifetime. These off-states of the fluorophore then causes fluctuations visible in the correlation function.

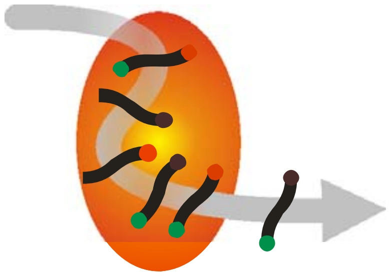


Figure 29. Schematic representation of fluorophore (red) and quencher (green) labeled molecule traversing the focal volume. Intersystem crossing to a triplet state prohibits fluorescence on a short time scale.

Fluorophores can also be irreversibly destroyed by photobleaching (Figure 30). The fluorophore can react with surrounding molecules during the excited state lifetime. Usually the fluorophore reacts after inter-system crossing to a triplet state. This state is long-lived, leaving a relatively long time span for the fluorophore to undergo chemical reactions. Photodestruction of the fluorophore is enforced by high excitation light intensities and radicals like oxygen within the solvent.

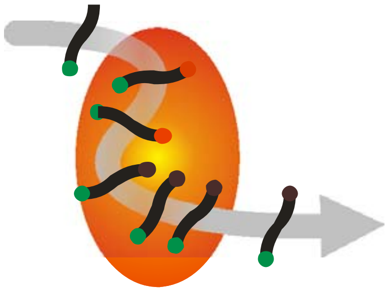


Figure 30. Schematic representation of fluorophore (red) and quencher (green) labeled molecule traversing the focal volume. The fluorophore is photobleached, fluorecence is lost permanently.

These shortened time periods, during which the molecule can be detected, pretend a faster diffusion in FCS measurements, i.e. shifting the diffusional regime towards shorter times in the correlation curve.

2.5.6. FLUORESCENCE CORRELATION SPECTROSCOPY LIMITS

The drawback of FCS is that it works only within a very limited concentration range: If the concentration of fluorescing molecules becomes too large (typically $> 10^{-8}$ M), the contribution from correlated photons from individual molecules, scaling with the number N of molecules within the detection volume, becomes small compared with the contribution of uncorrelated photons from different molecules, scaling with N^2 [48].

Another limiting factor in FCS measurements is that kinetic data can only be measured during diffusion time, i.e. the average time a molecule stays in the detection volume. Diffusion through the laser focus takes place at time scales of a few milliseconds, depending on the size of the diffusing particle. Furthermore, diffusion events and processes of similar timescales can complicate data analysis.

3. Materials and Methods

3.1. Computer system

All calculations contributing to this thesis were performed on the linux cluster BiCEPs. The cluster was designed, assembled and installed as part of this work. It contains a front end and twelve compute nodes.

3.1.1. FRONTEND

The frontend is the head of the cluster and the only computer connected to the university intranet. It is used to manage the whole cluster by the super user and as the login computer for ordinary users. Jobs are started from the frontend, which distributes them to the compute nodes.

frontend hardware	
mainboard	Tyan Thunder K8W
central processing units	2 × AMD Opteron 248
random access memory	8 × Infenion DIMM 512 MB DDR CL 2,5
graphics card	Gainward FX PowerPack Ultra/960
Myrinet card	Myrinet M3F-PCIXD-2
hard disk drives	4 × Maxtor 6B200M0
	1 × Maxtor 6Y080P0

Table 3-1. Frontend hardware.

3.1.2. COMPUTE NODES

Compute nodes on BiCEPs are mainly designed as the frontend, lacking half the memory and a powerful graphics adapter. They are set up in a 19” rack mount enclosure.

compute node hardware	
mainboard	Thunder K8S Pro S2882GNR
central processing units	2 × AMD Opteron 248
random access memory	4 × Infenion DIMM 512 MB DDR CL 2,5
Myrinet card	Myrinet M3F-PCIXD-2
hard disk drives	Samsung SV0802N

Table 3-2. Compute node hardware.

3.1.3. NETWORKS

All nodes are connected via Gigabit Ethernet (GE) through a 3Com SuperStack 3 switch with 24 ports. This network is used for non time-critical data transfer between compute nodes and the frontend

As a second network the high-bandwidth low-latency Myrinet network serves as communication layer for parallel processes in molecular dynamics simulations. Myrinet is an optical network with a bandwidth of 2 $\text{Gbit}\cdot\text{s}^{-1}$ and a typical latency of 2.6 µs. Compared to GE, Myrinet responds 100-fold faster. As molecular dynamics simulate the trajectory of a set of molecules by integrating Newton’s law in steps of femtoseconds, the only way to split up the task is to break down the modeled volume. After each time step the whole volume has to be rebuilt and potentials acting on the atoms have to be refreshed. Therefore, the simulation data is split up on the frontend, distributed to the nodes, calculated on the nodes, and returned to the frontend for every molecular dynamics step, which explains the strong dependency of molecular dynamics speed on the network’s bandwidth and latency.

All nodes and the frontend link to two Myrinet M3-SW16-8F Line Cards assembled in a Myrinet M3-E32 Switch Enclosure via 5 m and 25 m fiber cables, respectively.

3.1.4. OPERATING SYSTEM

The cluster runs an installation of Rocks 4.2. This is a freely distributed cluster management enabled operating system based on CentOS. CentOS is an open source version of the RedHat Enterprise Server linux system.

Rocks manages the cluster using a MySQL database. The software and services configuration are set at the frontend which automatically guides the installation of the compute nodes. Manually compiled software can be integrated as RPM packages.

3.1.5. SOFTWARE

The Rocks installation includes Myrinet support for GE, MPICH parallel libraries, PBS batch system, and Ganglia cluster monitoring software In addition computational chemistry software and its prerequisites were compiled and installed.

- MPICH for Myrinet (version 1.2.7)
- OctaveForge (version 2004.11.16)
- OpenBabel (version 2.0.2)

- Gamess (version 2006-09-07)
- AMBER (version 8)
- Visual Molecular Dynamics (version 1.8.5)
- PovRay (version 3.6.1)

For compilation of the software the Portland Group compiler suite (version 5.20) was used.

3.2. Parameterization

Parameterization of fragments used in AMBER [49] molecular dynamics simulations were carried out using R.E.D III [50]. This tool provides aid in the process of charge generation. It creates input files for ab initio geometry optimization and energy calculations with GAMESS [51], provides a tool for molecule reorientations and calculates atom centered charges with RESP [52].

As the general procedure in molecular dynamics simulations is to build large molecules from small units called residues, these molecular fragments must be parameterized in the first place. AMBER comes with parameter sets for a wide range of biomolecular residues like ribonucleic acids, amino acids and sugars as well as common solvents like water. It lacks any special modifications like the fluorescent dyes or modified desoxyribonucleic acid used in this work.

To circumvent this, an AMBER force field library for common fluorescent dyes, aliphatic linkers, and modified bases and amino acids was established for the molecular fragments listed in appendix 9.A.

Fluorescent dyes, needed to detect biopolymers in spectroscopic measurements, are commonly labeled to the biomolecule through amide bonds. In peptides the marker moiety is directly connected to a *N*-terminal amino acid or to a lysine’s side chain amino group, or to the C-terminus through an aliphatic 1,*n*-diamine. In DNA or RNA an 1-aminoalkane-*n*-ol of variable length can be used to convert a terminal phosphate group into an amine being capable of binding to the activated fluorescent dye. Light active moieties can also be introduced to the central region of ribonucleic acids by the means of modified bases, such as a 5-modified thymine, dM, (E)-3-(2,4-dioxo-1H-pyrimidin-5-yl)prop-2-enoic acid, or Stable Abasic Site, dS, 1,2-Dideoxyribose-5-phosphate. Figure 31 shows this general approach to introduce a fluorescent marker to the above named types of biomolecules.

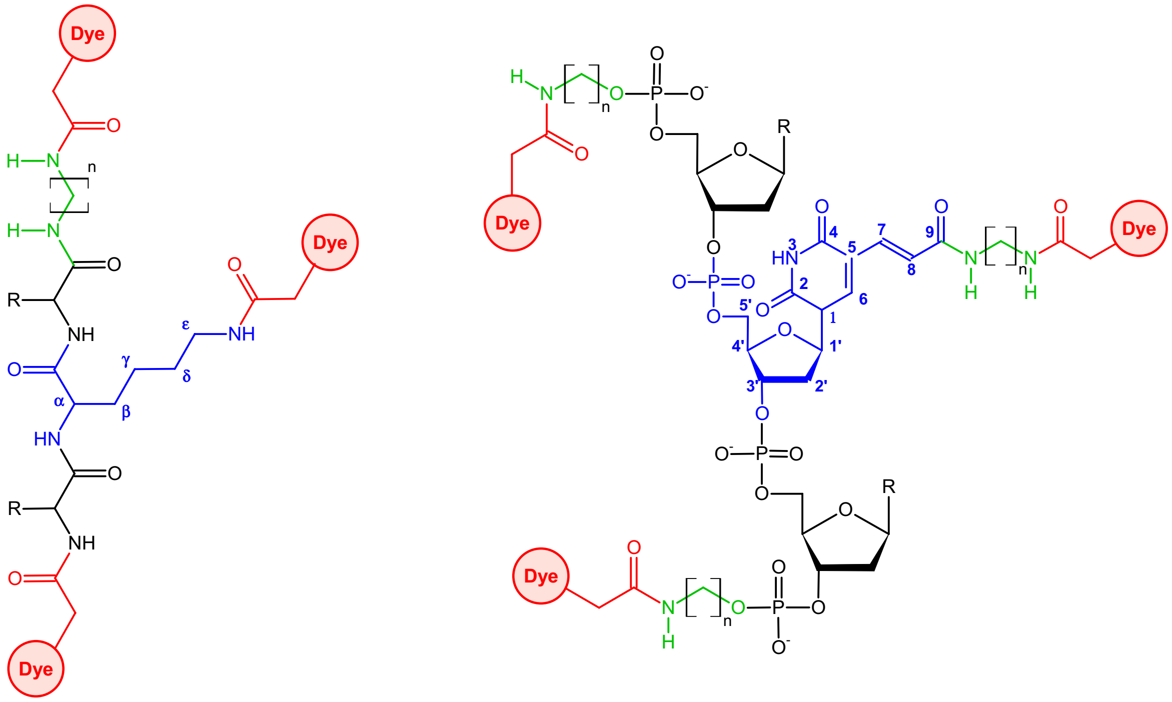


Figure 31. Biomolecule labeling. Peptides (left) can be fluorescent dye (red) labeled C-terminally (top), N-terminally (bottom) or internally at a lysine's (blue) ε-amine nitrogen. Nucleic acids (right) are labeled at the 5' end (top), the 3' end (bottom) or internally through an introduced modified thymine base (blue). Aliphatic linkers of variable length are used to convert non-conpliant binding sites to primary amines (green).

3.2.1. DYES

To generate a force field for fluorescent dyes, a model molecule was used. As the dyes are bound to biomolecules through amide bonds, the models for these were the *N*-methyl amides of the dyes. An arbitrary three dimensional starting structure of the model compound was created with ACD/ChemSketch [53]. The two dimensional structure was converted to 3D within ChemSketch and checked visually for chemical reasonability to ensure closeness to the energetic global minimum. The exported MDL mol file was converted to PDB format using OpenBabel [54].

With the aid of the tool AnteRED which is part of the R.E.D III package, an input file for the geometry optimization was generated. To ensure compatibility with the CORNELL et al. force field [55] the ab initio calculation was performed as a restricted HARTREE-FOCK calculation with the 6.31G* basis set. The convergence criterion for this geometry optimization was set to 10^{-6} Hartree · Bohr^{-1}, the SCF convergence value was 10^{-8} Hartree · Bohr^{-1}, and an analytical solution of the initial hessian matrix, i.e. the energy second derivative tensor, was chosen. The hessian matrix was updated every 10th iteration.

orientation	atom 1	atom 2	atom 3
1	A	B	C
2	B	D	C
3	C	E	D
4	D	A	E

Table 3-3: Molecule re-orientation scheme for ESP calculations.

The geometry optimized structure was used to calculate the electrostatic potential of the molecule. The most error prone step in point charge generation is the fit of atom centered charges to the ESP. Depending on the molecular orientation in respect to the coordinate system, point charges tend to differ. To reduce this deviation the ab initio potential calculation was performed up to four times with different orientations of the molecule. The reorientation scheme for the model structures was to pick five atoms from the molecule and use groups of three to define an orientation (Table 3-3). The molecules were re-orientated such that atom 1 is centered in the origin, atom 2 is placed on the x-axis and atom 3 lies on the xy-plane (equations (3.1) to (3.4)).

$$\mathbf{O} = \mathbf{r}_1 \tag{3.1}$$

$$\mathbf{x} = \frac{\mathbf{r}_2 - \mathbf{r}_1}{\left|\mathbf{r}_2 - \mathbf{r}_1\right|} \tag{3.2}$$

$$\mathbf{y} = \mathbf{z} \times \mathbf{x} \tag{3.3}$$

$$\mathbf{z} = \frac{\mathbf{r}_3 - \mathbf{r}_1}{\left|\mathbf{r}_3 - \mathbf{r}_1\right|} \times \frac{\mathbf{r}_2 - \mathbf{r}_1}{\left|\mathbf{r}_2 - \mathbf{r}_1\right|} \tag{3.4}$$

ESP calculations were also performed at the 6.31G* level. The SCF was converged to 10^{-6} Hartree · Bohr^{-1}. For the potential calculation a set of points on several fused sphere VAN DER WAALS surfaces, with points selected using an algorithm described by KOLLMAN and SINGH [56] was used. This method resulted in electrostatic potential data consisting of 739 to 2167 data points.

RESP was then used to fit these data to a set of atom centered charges. In a first step all atom charges besides the charges of the protecting group were fitted independently to a convergence level of $5 \cdot 10^{-5}$ C. The protecting group was set to a group charge equivalent to its chemical charge. In a second step all methyl or methylene groups were refitted with the constraint of hydrogen atoms being equally charged. This fit was run to a convergence criterion of a change in charge lower than 10^{-4} C.

To convert the model structure used for ab initio calculations to a molecule fragment compatible to AMBER residue libraries, the protecting group' collective atom charge was restrained to 0 for the two applied fits.

The RED III Perl script created a `mol2` file with atom charges which were used with AMBER's leap module to create an AMBER residue library file of the molecule fragment. This was done by specifying atom types for each atom of the fragment. Atom types were taken from AMBER's CORNELL et al. force field.

3.2.2. NUCLEIC ACIDS

The procedure of the nucleic acid parameterization was closely related to the above described method applied to the fluorescent dyes. As nucleic acids form polymers, they can adopt four chemically similar types within the macromolecule: At the 5' end of the polymer the ribose backbone is capped by a hydroxyl group at $C^{5'}$ (dN^5), at the polymer's 3' end the terminal $C^{3'}$ carries a hydroxyl group (dN^3), within a polymer the residue carries a phosphate group at $C^{5'}$ (dN), and finally a neutral monomer form contains the nucleic acid flanked by two hydroxyl groups at $C^{5'}$ and $C^{3'}$ (dN^n).

Furthermore, the ribose backbone is usually treated equivalently in molecular dynamics simulations. This implies that the atom charges of the ribose moieties are equal in all nucleic acid residues, except for the $C^{1'}$ carbon atoms and their hydrogen atoms as these compensate for the differing charges of the bases.

The approach to incorporate these restraints a many structure model was used in the parameterization procedure. The different nucleic acids (dA, dC, dG, dT, dS, and dM) were geometry optimized as neutral monomers (dN^n) and dimethyl phosphate (DMP) was used in its *cis*-*cis* conformation to reflect the bridging phosphate groups in the later polymer. In the charge fitting process the ribose backbone atoms were restrained to equal charges (Figure 32, green atoms). The methyl groups of the DMP and the 5' and 3' hydroxyl groups were restraint to a neutral summed atom charge, respectively (Figure 32, red boxes).

After the atom centered charge fit the nucleic acids and DMP are assembled to form dN^5, dN^3, and dN, respectively, neglecting phosphate methyl and ribose hydroxyl groups where necessary.

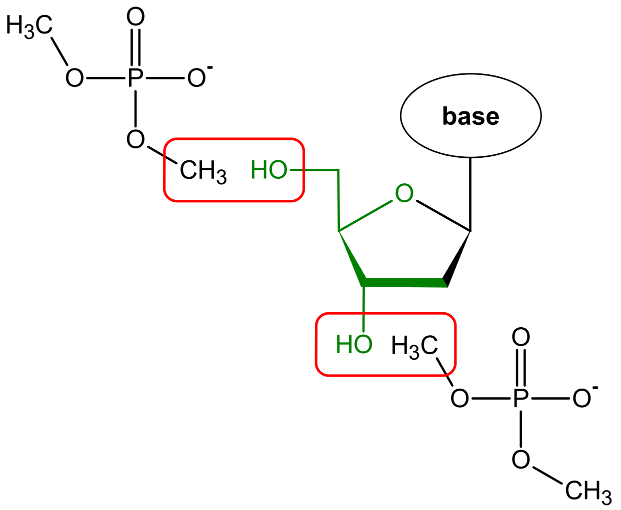


Figure 32. Schematic representation of the parametrization method used for nucleic acids. Atoms in green color are restrained to equal charges in all nucleic acids. Red boxes show the bimolecular charge constraint between the ribose hydroxyl groups and the DMP methyl groups.

Despite of the multi-molecule approach all calculations and fits are performed as for the fluorescent dyes described above.

3.2.3. LINKING FRAGMENTS

Fluorescent dyes carry a carbonylic acid group that is used to label them to biomolecules. Whereas peptides can be marked directly at the *N*-terminus, aliphatic linkers are used for the C-terminal and for nucleic acid polymer labeling (green fragments in Figure 31).

These linker molecules can be parameterized in the same fashion as dyes and nucleic acids. Linkers used in nucleic acid polymers must be restrained to the same net charge as the terminal nucleic residues as the linker-dye construct replaces the terminal residue.

3.3. Polyprolines

3.3.1. SPECTROSCOPY EXPERIMENTS

Sample Preparation

Synthetic peptides with the sequence $(\text{Pro})_N$-Trp ($N = 0 - 10$) (Thermo, Ulm, Germany) were modified with the amino-reactive oxazine fluorophores MR121-NHS, MR113-NHS, and the rhodamine derivative R6G-NHS (Invitrogen, Karlsruhe, Germany) at the *N*-terminus by using classical *N*-hydroxysuccinimidylester (NHS-ester) chemistry. For chemical structures of the used fluorescent dyes see appendix 9.A. 50 µg of reactive fluorophore were dissolved in 5 µl of acetonitrile. Approximately 500 µg of peptide were dissolved in 50 µl of aqueous sodium bicarbonate (50 mM, pH 8.3) containing 30 % (vol / vol) acetonitrile. Fluorophore and peptide solutions were mixed and incubated for 1 h at room temperature. Labeled peptide conjugates were purified by reversed-phase (Hypersil-ODS C18 column) high performance liquid chromatography (Agilent Technologies, Waldbronn, Germany) by using a linear gradient of 0 - 75 % acetonitrile in 0.1 M aqueous triethylammonium acetate.

Ensemble Spectroscopy

Absorption and steady-state emission measurements were carried out on a UV-visible (Lambda 25, PerkinElmer, Waltham, MA) or fluorescence spectrometer (Cary Eclipse, Varian, Darmstadt, Germany) at 20°C. Fluorescence lifetimes were measured by time-correlated single-photon counting at 20°C under the magic angle (54.7°) with a standard TCSPC spectrometer (model 5000MC, IBH, Glasgow, U.K.) equipped with a pulsed laser diode (635 nm). In all cases a bi-exponential model (3.5) the measured decay could adequately be described. The pre-exponential factors, a_i, describe the ratio of the de-excitation pathways and τ_i denote the lifetimes of the excited state before de-excitation.

$$I = I_0 \cdot e^{-\left(\frac{\Delta t}{a_1 \cdot \tau_1 + a_2 \cdot \tau_2}\right)} \tag{3.5}$$

Because of the limited time resolution, strongly quenched populations with decay times shorter than ~ 50 ps were not revealed. Molecular aggregation, glass absorption of peptides or fluorophores, and re-absorption of emitted photons were excluded in all ensemble experiments by concentration-independent results up to ~ 1 µM. The steady-state ($\Phi_{ss} = I / I_0$) and dynamic ($\Phi_{dyn} = \tau_2 / \tau_0$) quantum yield were calculated from fluorescence intensity, I, and lifetime, τ_2, where τ_0 and I_0 are the fluorescence lifetime and intensity of a non-quenched reference sample. The steady-state quantum yield is the product of static and dynamic quantum yield ($\Phi_{\mathrm{ss}} = \Phi_{\mathrm{stat}} \cdot \Phi_{\mathrm{dyn}}$), where Φ_{dyn} reflects collisional quenching and Φ_{stat} reflects formation of non-fluorescent complexes [57].

Fluorescence Correlation Spectroscopy

FCS experiments were performed on a confocal fluorescence microscope that essentially consists of a standard inverse fluorescence microscope equipped with a HeNe laser, emitting at 632.8 nm, as excitation source. The collimated laser beam was coupled into an oil-immersion objective (63, numerical aperture of 1.4, Zeiss) by a dichroic beam splitter (645DLRP, Omega Optical, Brattleboro, VT). The fluorescence signal was collected by the same objective, filtered by a band-pass filter (700RDF75, Omega Optical), and imaged onto the active area of two single-photon avalanche photodiodes (APDs) (AQR-14, EG & G, Vaudreuil, QC, Canada), sharing the fluorescence signal by a cubic nonpolarizing beamsplitter (Linos, Göttingen, Germany). The signals of the APDs were recorded in the cross-correlation mode by using a digital real-time correlator device (ALV-6010, ALV, Langen, Germany). Fluorescently modified peptides were diluted to a final concentration of ~ 0.5 nM in phosphate-buffered saline (PBS buffer, pH 7.4) containing 0.3 $mg \cdot ml^{-1}$ bovine serum albumin (BSA) and 0.05 % Tween-20 to suppress glass surface interactions. Excitation intensities were adjusted to 1 mW (at the objective’s back aperture) for measuring sub-microsecond kinetics and reduced to 50 μW, avoiding possible artifacts due to photobleaching and fluorescence saturation of the fluorophore, for determination of sample concentrations.

Fluctuations in the fluorescence signal $I(t)$ were analyzed *via* the second order autocorrelation function. For simple diffusion the correlation function can be described by two parameters, the occupancy, *O*, of the observation volume and the characteristic diffusion time, τ_D [58, 59]. For each FCS measurement a relative occupancy, O_0, was calculated by normalizing *O* to the absolute molecule concentration (as determined from the optical density at the fluorophore’s absorption maximum). Dividing the average fluorescence signal *I* by the average number of molecules within the observation volume, *V,* yielded the average photon emission rate per individual molecule, *B*.

3.3.2. COMPUTATIONAL METHODS

Separation Approach

To investigate contact formations possibilities of quencher and dye in *N*-terminally dye labeled polyprolines carrying C-terminal tryptophan (F-Pro$_N$-Trp, F = MR113, MR121, R6G), the system is split up into three parts: The conformational space of the dye is simulated through a spatial sampling of possible geometries by permutating the rotatable dihedral angles of the dye's linker. This results in a point cloud of the conformers' aromatic systems' centers that is regarded as independent of the polyproline's *cis* / *trans* conformation. The tryptophan is treated in the same way.

To investigate the influence of *cis* peptide bonds within polyprolines on the quenching possibility of tryptophan conformations of all *cis* / *trans* permutations for polyprolines up to a length of ten monomers are calculated. Above this length only all-*trans* and mono-*cis* conformations are taken into account, as the number of conformers rises with two to the power of the number of peptide bonds, yielding 512 structures for a 10-mer and doubling with every additional monomer.

MODEL STRUCTURES

Model structures for spatial distribution computation are achieved with AMBER's Leap using the CORNELL et al. force field in combination with previously developed libraries for MR113, MR121 and Rhodamine 6G.

ENERGY CALCULATION

To gain insight into the relative free energy of the built model structure an energy minimization run is performed with just one step length to prevent the structure from relaxing into a lower energetic state. As there is no positional change of the atoms during minimization and all model structures are identical in respect to the proline chain the difference in energy between two structures should roughly resemble the change in energy due to dihedral rotation of the linker bonds.

Energy calculations are performed with implicit solvent model and a cutoff distance of 300 Å.

VAN DER WAALS CONTACTS

To examine a generated structure for VAN DER WAALS contacts between the non-linker fragment of the terminal residue and the polyproline a distance matrix is generated with `ptraj`. It is a $N \times N$ square matrix where N is number of atoms contributing to the molecule and contains the distances between any pairs of atoms.

The distance matrix is compared with a matrix of VAN DER WAALS distances between atom pairs within the molecule. This violation matrix, **V**, is generated by summing up $r_{\mathrm{vdW},i}$ and $r_{\mathrm{vdW},j}$ for *i* and *j* being element of all atom indices.

$$\mathbf{V}(i,j)=\frac{1}{2}\cdot\left(\left(\left(\frac{r_{i,j}^{\mathrm{vdW}}}{r_{i,j}}\right)^{12}-\left(\frac{r_{i,j}^{\mathrm{vdW}}}{r_{i,j}}\right)^{6}\right)+\mathrm{abs}\left(\left(\frac{r_{i,j}^{\mathrm{vdW}}}{r_{i,j}}\right)^{12}-\left(\frac{r_{i,j}^{\mathrm{vdW}}}{r_{i,j}}\right)^{6}\right)\right) \tag{3.6}$$

The violation matrix turns zero wherever the distance between two atoms is greater than the VAN DER WAALS distance and yields a factor proportional to the repulsive VAN DER WAALS force if the distance is shorter.

A submatrix using only the first rows of the violation matrix (representing non-linker dye atoms) and the end columns (representing the polyproline atoms) easily shows VAN DER WAALS contacts through a positive determinant. VAN DER WAALS radii were taken from [60].

MODEL STRUCTURE NORMALIZATION

All model structures are translated and rotated such that the peptide bond connecting the terminal residue to the polyproline is centered in the origin. Peptidic nitrogen is placed on positive x-axis, carbonyl oxygen on the xy-plane in the second quadrant.

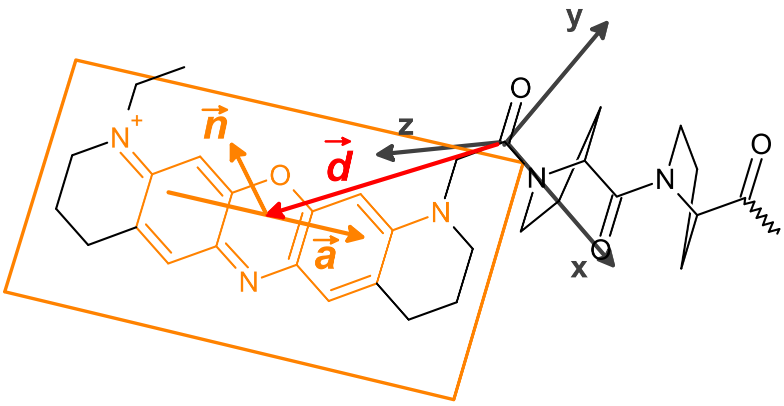


Figure 33. Vectorial description of MR113 fluorescent label. The cartesian coordianate system is aligned with the peptide bond connecting dye and first proline residue. The fluorescent moiety's orientation is described by the pointing vector, **d**, to the dye's center of mass, its normal vector, **n**, orthogonal to the plane of the dye's aromatic system, and the axis vector, **a**, parallel to the aromatic system.

Normalization is performed in octave. The PDB file is read in with the `creadpdb` function of the octave-forge package. In octave a rotation matrix is generated and the coordinates are transformed. Then a plane is fitted through all aromatic heavy atoms of the terminal residue. The normal vector of the plane resembles the normal vector of the aromatic system. To determine the aromatic axis vector of the aromatic system the mean of two groups' atom positions are calculated and subtracted as indicated in figure.

Using these three parameters centre vector, **d**, normal vector, **n,** and axis vector, **a**, the position and the orientation of the dye is fully described and can easily be adopted to other peptidic polymers (Figure 33).

POINT CLOUD DATA CREATION

Permutating dihedral angles results in an exponential rise of simulated structures, *N*, depending on the number of rotatable bonds, *k*, and the dihedral step size dϕ:

$$N = \left(\frac{360^\circ}{\Delta\phi} \right)^k \tag{3.7}$$

Step sizes of 5°, 20°, 2° and 10° for MR113, MR121, Rhodamine 6G and tryptophan yield point cloud data sets of 5184, 104976, 180 and 46656 entries, respectively. As for each proline conformation any data point of the dye's point cloud has to be compared to any point of tryptophan's the amount of distance calculations equals the product of the number of data sets of the dye point cloud and tryptophan point cloud, ranging from 8.4 million for Rhodamine 6G and 4.9 billion calculations resulting in segmentation faults when being calculated with origin. Therefore, point cloud data sets are reduced by summing up data points within angular proximity to each other.

Point Cloud Data Reduction

The surface of a sphere is divided into surface bins with similar areas. In spherical coordinates data points are grouped depending on their azimuthal angle θ and their polar angle ϕ (Figure 34).

With a binning frequency *n* giving the number of bins per hemi circle the number of azimutal bins is

$$n_\theta = n \tag{3.8}$$

Azimuthal bins are defined as

$$B_\theta(k): \left(\frac{k}{n} \cdot \pi \right) \le \theta < \left(\frac{k+1}{n} \cdot \pi \right) \qquad \forall \;\; 0 \le k \le (n-1) \tag{3.9}$$

with a mean angle θ_k of

$$\theta_k = \frac{2 \cdot k + 1}{2 \cdot n} \cdot \pi \tag{3.10}$$

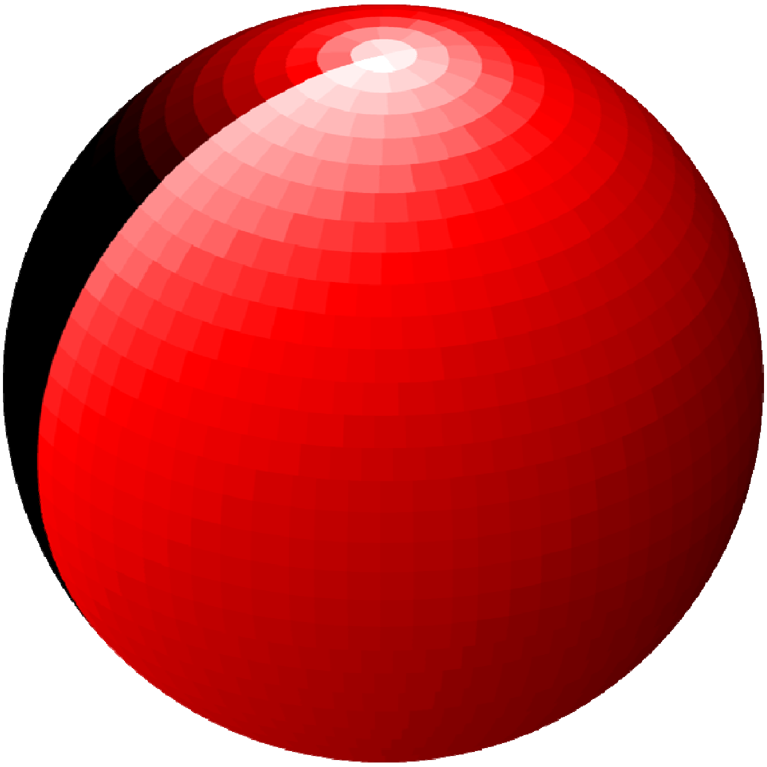


Figure 34. Spherical binning concept. A sphere's surface is divided into equally sized areas. At the poles fewer longitudinal bins are used than at the equator where the number of equatorial bins reaches the number of azimuthal bins.

The radial bin size in polar direction depends on the mean azimuthal angle of the bin. It is proportional to the radius of the intersection circle at θ_k. The number of polar bins for a given azimuthal bin is

$$n_\phi = \lfloor \sin\theta_k \cdot n \rfloor + 1 \tag{3.11}$$

where the truncated brackets indicate the floor function. Therefore, polar bins are given by

$$B_\phi(k,l): \frac{\pi \cdot l}{\lfloor \sin\theta_k \cdot n \rfloor + 1} - \pi \le \phi < \frac{\pi \cdot (l+1)}{\lfloor \sin\theta_k \cdot n \rfloor + 1} - \pi \qquad \forall\ 0 \le l \le \left(2 \cdot \lfloor \sin\theta_k \cdot n \rfloor + 1\right) \tag{3.12}$$

with a mean angle of

$$\phi_l = \frac{2 \cdot l - 2 \cdot \lfloor \sin\theta_k \cdot n \rfloor - 1}{2 \cdot \lfloor \sin\theta_k \cdot n \rfloor + 2} \cdot \pi \tag{3.13}$$

Close to the sphere's poles fewer polar bins per azimuthal bin are used than in equatorial regions where n_ϕ converges to 2 *n*.

$$k = \left\lfloor \frac{n \cdot \theta}{\pi} \right\rfloor \tag{3.14}$$

$$l = \left\lfloor \frac{\phi + \pi}{\pi} \cdot \left(\left\lfloor n \cdot \sin\left(\frac{2 \cdot \left\lfloor \frac{n \cdot \theta}{\pi} \right\rfloor + 1}{2n} \cdot \pi \right) \right\rfloor + 1 \right) \right\rfloor \tag{3.15}$$

Dye-Tryptophan Distance Distribution Calculation

Two coordinate systems are used for distance calculations: The first amide group in the model structure used for dye point cloud calculations, connecting dye and first proline residue, defines the first Cartesian coordinate system, $\Re^3$: The carbonyl carbon sets the origin, amide nitrogen defines positive x and carbonyl oxygen resides in the second quadrant of the xy-plane. The same coordinate system is used in proline model structures where the flanking glycine residue replaces the fluorescent dye and in tryptophan point cloud generation where the amide group between the last proline and terminal tryptophan are centered to the origin. The second coordinate system, $\Re^{3'}$, represents the orientation of the amide group between the last proline and terminal glycine residues in proline model structures. The equivalent atoms as used for $\Re^3$-definition span x', y' and z'.

The distance, **d**, between the fluorescent dye and tryptophan in dependency of the dihedral angles of the two moieties is calculated as

$$d = \left| \mathbf{p} + \mathbf{t'} - \mathbf{d} \right| \tag{3.16}$$

where the vectors **d**, **p** represent the fluorescent dye's center of mass and the position of the terminal amide bond's carbonyl carbon. Tryptophan's center of mass is described by **t'**, the vector from carbonyl carbon to the aromatic system's center of mass and is calculated as the original vector, **t**, multiplied by a rotation matrix to compensate differing spatial orientation of the terminal amide bond (Figure 35).

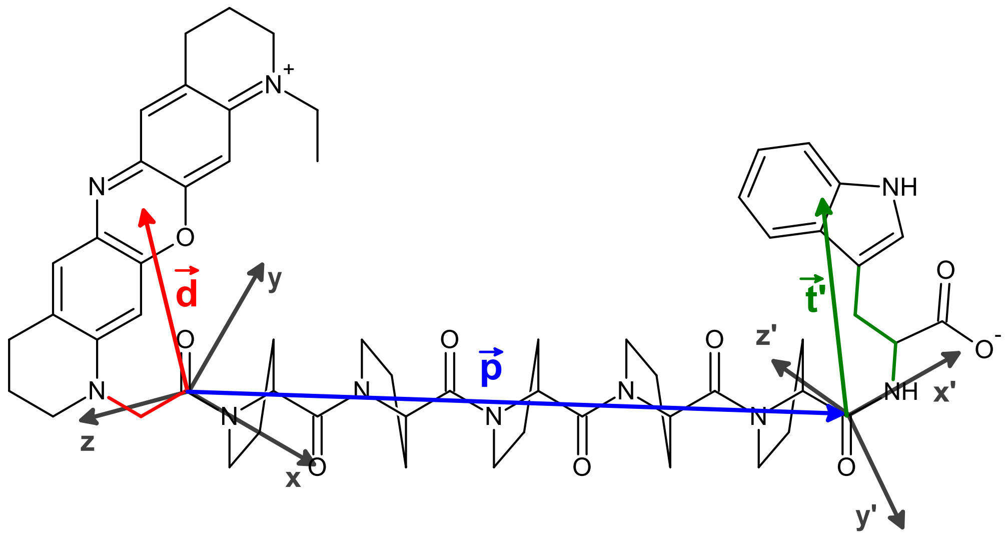


Figure 35. The proline chain is described by the vector p (blue), pointing from the carbon of the first amide bond to the carbon of the last. Both amide groups span the cartesian coordinate systems [**x**, **y**, **z**] and [**x'**, **y'**, **z'**], respectively. Fluorescent dye and tryptophan are described as vectors **d** (red) and **t'** (green) pointing from the first coordinate system's origin to the dye's center of mass and from the second coordinate system's origin to the tryptophan's center of mass, respectively.

Dye and tryptophan point clouds can be represented by two matrices **D** and **T**, containing *d* and *t* vectors to the aromatic centre of the moieties, respectively.

$$\mathbf{D}=\begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ \vdots & \vdots & \vdots \\ x_d & y_d & z_d \end{pmatrix} \qquad \mathbf{T}=\begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ \vdots & \vdots & \vdots \\ x_t & y_t & z_t \end{pmatrix} \tag{3.17}$$

For a given proline length and conformation the polyproline can be represented by a vector, **p**, pointing from first to last amide bond carbonyl carbon. Furthermore, a 3×3 transformation matrix, **P**, is specified converting amide carbon, oxygen and nitrogen of the first amide bond to the equivalent atoms within the last.

Dye tryptophan distance distributions can be calculated as

$$\begin{aligned} \mathbf{r}_{i,j} &= \left| \mathbf{p}+\mathbf{t}_j{}' - \mathbf{d}_i \right| \\ &= \left| \mathbf{p}+\mathbf{t}_j \cdot \mathbf{P} - \mathbf{d}_i \right| \qquad \forall \quad \mathbf{d}_i \in \mathbf{D}, \quad \mathbf{t}_j \in \mathbf{T} \end{aligned} \tag{3.18}$$

In matrix notation **T** is multiplied by **P** to yield a matrix **T***, holding rotated vectors accounting for the proline-tryptophan amide group’s orientation. If **D’**, **T’** and **P’** are single column matrices of *d*, *t* and *d·t* rows in size containing ones, a distance vector matrix, **R**, is generated using the Kronecker product:

$$
\begin{aligned}
\mathbf{R} &= \left(\mathbf{T} \bullet \mathbf{P}\right) \otimes \mathbf{D}' - \mathbf{T}' \otimes \mathbf{D} + \mathbf{P'} \otimes \vec{\mathbf{p}} \\
&= \quad \mathbf{T}^{*} \otimes \mathbf{D}' - \mathbf{T}' \otimes \mathbf{D} + \mathbf{P'} \otimes \vec{\mathbf{p}} \\
&= \begin{pmatrix}
x_1^{t^*} & y_1^{t^*} & z_1^{t^*} \\
\vdots & \vdots & \vdots \\
x_1^{t^*} & y_1^{t^*} & z_1^{t^*} \\
x_2^{t^*} & y_2^{t^*} & z_2^{t^*} \\
\vdots & \vdots & \vdots \\
x_2^{t^*} & y_2^{t^*} & z_2^{t^*} \\
 & \vdots & \\
 & \vdots & \\
x_t^{t^*} & y_t^{t^*} & z_t^{t^*} \\
\vdots & \vdots & \vdots \\
x_t^{t^*} & y_t^{t^*} & z_t^{t^*}
\end{pmatrix}
-
\begin{pmatrix}
x_1^{d} & y_1^{d} & z_1^{d} \\
\vdots & \vdots & \vdots \\
x_d^{d} & y_d^{d} & z_d^{d} \\
x_1^{d} & y_1^{d} & z_1^{d} \\
\vdots & \vdots & \vdots \\
x_d^{d} & y_d^{d} & z_d^{d} \\
 & \vdots & \\
 & \vdots & \\
x_1^{d} & y_1^{d} & z_1^{d} \\
\vdots & \vdots & \vdots \\
x_d^{d} & y_d^{d} & z_d^{d}
\end{pmatrix}
+
\begin{pmatrix}
x^{p} & y^{p} & z^{p} \\
 & \vdots & \\
 & \vdots & \\
 & \vdots & \\
 & \vdots & \\
 & \vdots & \\
 & \vdots & \\
 & \vdots & \\
 & \vdots & \\
 & \vdots & \\
x^{p} & y^{p} & z^{p}
\end{pmatrix}
\end{aligned}
\tag{3.19}
$$

R is a ($d{\cdot}t$)×3 matrix containing point-to-point vectors between any two point of the clouds. Distances, **r**, are the square root of the sum of the rows components' squares:

$$
\mathbf{r} = \left[\left[\begin{pmatrix}
x_1 & y_1 & z_1 \\
x_2 & y_2 & z_2 \\
\vdots & \vdots & \vdots \\
x_i & y_i & z_i
\end{pmatrix} \bullet \begin{pmatrix}
x_1 & x_2 & \cdots & x_i \\
y_1 & y_2 & \cdots & y_i \\
z_1 & z_2 & \cdots & z_i
\end{pmatrix}\right]\right]
\tag{3.20}
$$

The structure with smallest distance between dye and tryptophan was extracted from the matrix **r**. The smallest value of its rows indicates the geometry of the structure. From the row number, $i_{\mathbf{r},\min}$, the dihedral angle set could be retrieved:

$$
i_{\mathrm{trp}} = \left\lfloor \frac{i_{\mathbf{r},\min}}{t} \right\rfloor + 1
\tag{3.21}
$$

$$
i_{dye} = i_{\mathbf{r},\min} \bmod t
\tag{3.22}
$$

3.4. Polythymines

3.4.1. FLUORESCENCE CORRELATION SPECTROSCOPY

All FCS measurements were performed on a home-build confocal microscope as sketched in Figure 36.

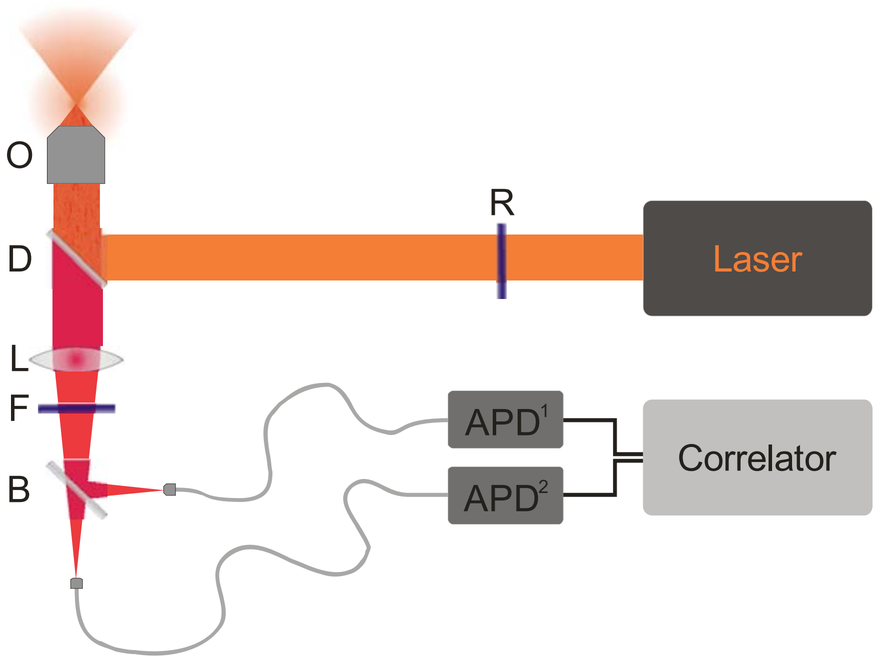


Figure 36. FCS setup. The laser light (orange) passes a reducer to regulate the excitation power, is reflected by the dichroic mirror, D, runs through the objective, O, and is focused into solution. Emitted light (red) is collected by the same objective and passes the dichroic mirror and a concave lens, L. It is split into two rays by a beam splitter and focused into two optical fibres. These channel the emitted light onto two avalanche photo diodes which are connected to a hardware correlator.

MR121 was excited at 632.8 nm (50 µW measured at the back aperture of the objective) using a HeNe laser. Photobleaching was not detectable at the chosen excitation power. The collimated laser beam was coupled into an oil-immersion objective (63x, NA 1.4, Zeiss, Jena, Germany) by a dichroic beam splitter (645DLRP, Omega Optics, Brattleboro, VT, USA). The fluorescence signal was collected by the same objective, filtered by a band-pass filter (700DF75; Omega Optics, Brattleboro, VT, USA), separated into two beams using a cubic non-polarizing beam splitter (Linos, Göttingen, Germany), coupled into multi-mode optical fibers with a diameter of ~100 µm, and imaged onto the active area of two single-photon avalanche photodiodes (AQR-14, EG&G, Quebec, Canada). The signals of the APDs were cross-correlated (15 min for each measurement) using a digital real time multi-tau correlator device (ALV-6010, ALV-GmbH, Langen, Germany) with a time resolution down to 6.25 ns.

Sample Preparation

The fluorescent oxazine derivative MR121 (chemical structure in appendix 9.A.2) was provided by Prof. K.-H. Drexhage (Universität-Gesamthochschule Siegen, Germany). The synthesis of amino-reactive MR121-NHS (MR121-*N*-hydroxysuccinimidylester) was performed as described by Neuweiler [61].

DNA probes carrying a C_6-amino linker at the 5'-end (IBA, Göttingen, Germany) were fluorescently labeled using standard *N*-hydroxysuccinimidylester (NHS-ester) chemistry. All other chemicals for conjugation chemistry and measurements were purchased from Sigma-Aldrich, Munich, Germany. Fluorescently modified probes were purified by reversed-phase (Hypersil-ODS C-18 particles, 5 µm diameter, 120 Å pore size) high performance liquid chromatography (Agilent Technologies, Böblingen, Germany) using a linear gradient of 0 - 75% acetonitrile in 0.1 M aqueous triethylammonium acetate.

The amino-terminal conjugates were diluted to a final concentration of ~ 1 nM in 10 mM sodium phosphate buffer, pH 7.0, containing 0 - 1 M NaCl, 0.1 mM ethylene diamine tetraacetate (EDTA), and 0.3 $mg{\cdot}ml^{-1}$ bovine serum albumin (BSA) to suppress glass surface adsorption. DNA samples were transferred onto a microscope slide and covered by a cover slip. Sample temperature was controlled by a custom-built objective heater to be constant at 20 °C.

3.4.2. MOLECULAR DYNAMICS SIMULATIONS

All molecular calculations were carried out with the AMBER molecular simulation program package. Trajectories were simulated for polythymine $(dT)_N$, with $N = (2, 4, 8, 16, 32, 64, 128)$, and MR121-labeled polythymines, MR121-$(dT)_N$, of equal length. Further simulations were performed for stable abasic sites $(dS)_N$, representing the pure backbone of ss-DNA, and for penetrable polythymine $(dT_{pen})_N$, a hypothetical ss-DNA that can penetrate itself, since all VAN DER WAALS interactions were switched off. Starting structures were generated from AMBER's residue library `all-nucleic94.lib` with LEAP and energy-minimized with SANDER for a maximum of 10,000 steps using the steepest-descent algorithm for four steps and conjugate gradient thereafter. MD simulations were run with SANDER using the AMBER force field FF99, an integration time of 1 fs, and a non-bonded cutoff of 12 Å and a simulation temperature of 800 K. Attractive non-bonded interactions were deactivated using SANDER's NMR refinement options (HB, ELEC, ATTRACT). For the simulation of $(dT_{pen})_N$, VAN DER WAALS interactions were deactivated using SANDER's NMR refinement options (NB). Snapshots were taken from the trajectories every 1 ps starting after the first 250 ps of simulation time. The offset was set to exclude system equilibration (heating up from 0 K to 800 K) and to assure relaxation from the unlikely starting geometry of an extended polynucleotide as it was generated from the residue library. Simulations were carried out over 20,250,000 time steps of 1 fs each, yielding a usable trajectory of 20-ns length with $n_{\mathrm{Sim}} = 20{,}000{,}000$ snapshots. Due to the elongated structure derived from the Leap libraries collapsing to a more realistic conformation, AMBER's internal virtual box was sporadically busted, leading to abortion of the simulation. This was overcome by restarting the simulations, with an increased box size. Atom coordinates and velocities were taken from the prior simulation. Concatenating all produced trajectories resulted in the full size trajectory used for further analysis.

Simulation Temperature

With rising temperature the energy of a molecule increases and energy barriers between local minima on a molecules potential hyper surface can be overcome more easily. On the other hand rising simulation temperature also raises the energy accumulated in bond length, angle, and dihedral angle distortion. As molecular dynamics force fields use simplified energy potentials which are accurate for small deviations from the standard value, simulations at high temperature could yield falsified results. To ensure sufficient sampling of allowed conformations combined with model consistency trial simulations of $(dT)_{16}$ were run at temperatures ranging from 200 to 1200 K resulting in a chosen simulation temperature of 800 K.

Polymer Property Evaluation

As pure and dye labeled polynucleotides were investigated, the 5'-oxygen atom was used as the polymer's origin. The polymer segments, $\mathbf{s}_i$, were defined from one phosphor atom to the subsequent (dark green arrow in Figure 37). Due to the missing phosphate groups at the 5'- and 3'-terminal ends, the polythymines consist of $N - 2$ ordinary segments, flanked by two capping segments, $\mathbf{s}_{5'}$ and $\mathbf{s}_{3'}$ (light green arrows in Figure 37), pointing from the $O^{5'}$-atom of the first thymidine to the first phosphor atom of the polymer and from the last phosphor atom to the terminal hydrogen, H^{3T}.

The end-to-end distance (red line in Figure 37) was defined as the distance between first thymidine's $O^{5'}$-atom and terminal hydrogen, fulfilling the definition in 2.4.3 (equation (2.57)) with the restriction of two capping segments slightly deviating from normal with regard to segment length. The fluorescent dye MR121 linked to the

5’-end of the polymer (grey moiety in Figure 37) was neglected in all calculations of polymer properties except calculation of radius of gyration.

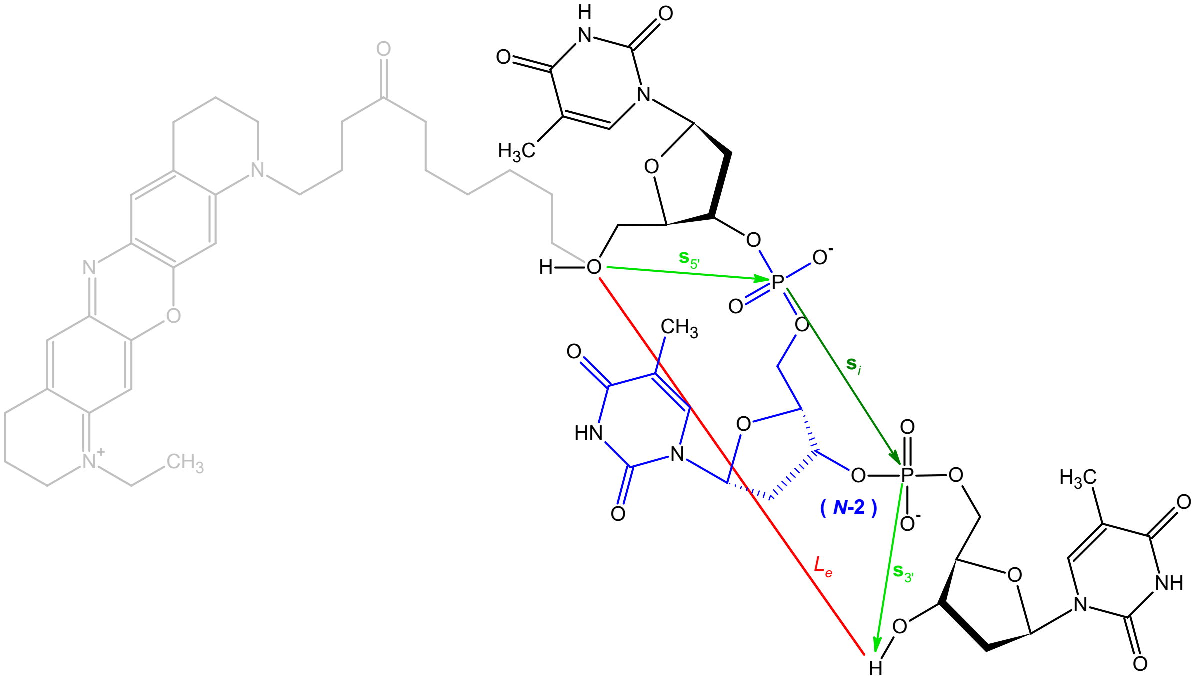


Figure 37. Schematic representation of the examiced polythimines. The nucleic acid chain is shown with segment vectors indicated by green arrows, end-to-end distance by a red line. The center thymidine residue which repeats *N* - 2 times is colored blue. The dye MR121 is shown in grey and represents the 5'-terminal hydrogen substitution in dye labeled probes.

The radius of gyration was calculated with VMD according to equation (2.60). It was averaged over the whole simulation length.

Time averages of the segment vectors’ orientational lengths, $C_{i,j}(t)$, were calculated as:

$$C_{i,j}(t) = \left\langle \mathbf{s}_i(t) \cdot \frac{\mathbf{s}_j(t)}{\left|\mathbf{s}_j(t)\right|} \right\rangle \qquad i \in \left[1, 2, 3, \ldots, (N-2)\right]; \qquad j \in \left[i, (i+1), \ldots, (N-2)\right] \tag{3.23}$$

These projections of the segment vectors onto a preceding normalized segment vector shows the directional correlation between the two vectors. While two neighbored segments are aligned in similar directions segments far apart from each other show less correlation in terms of orientation. The data sets for each vector pair was temporally averaged

$$C_n(i,j) = \sum_{t=1}^{N_{\text{Sim}}} C_{i,j}(t) \qquad i \in \left[1, 2, 3, \ldots, (N-2)\right]; \quad j \in \left[i, (i+1), \ldots, (N-2)\right] \tag{3.24}$$

From the orientation lengths the persistence length, $L_{\text{p,OC}}$, was evaluated by fitting the orientational correlation function for two segment vectors separated by *i* segments, C_i:

$$C_k = \sum_{i=1}^{(N-k-2)} \frac{C_n\left(i,(i+k)\right)}{N-k-2} \tag{3.25}$$

This function representing the orientational correlation in dependence on the number of separating segments. It can be fitted as discussed in 2.4.3 to evaluate the persistence length $\text{L}_{\text{p,OC}}$ using a mono-exponential decay.

$$C_i = \langle l \rangle \cdot \langle \cos\theta_i \rangle = C_0 \cdot \mathrm{e}^{\frac{-i\cdot\langle l \rangle}{L_{p,oc}}} \tag{3.26}$$

Alternatively the persistence length was evaluated as the projection length, $L_{\text{p,Pro}}$. This is the end-to-end distance vector, projected onto the first segment. The sum of all segment projections onto the first segment was used as end-to-end distance her, neglecting the two capping segments.

$$L_{p,\text{Pro}} = \sum_{i=1}^{N-2} \mathbf{s}_i \cdot \frac{\mathbf{s}_1}{\left|\mathbf{s}_1\right|} \tag{3.27}$$

The persistence length calculations were performed with different scripts: First, phosphor atom coordinates were extracted from the trajectory with AMBER’s ptraj. A custom Perl script was used to generate the orientational length data files which were histogramized by an octave script.

4. Results and Discussion

4.1. Polyprolines

The conformation of polyprolines has been heavily discussed during the past decades. Synthesized polyprolines are used as spectroscopic rulers. As these peptides are believed to adopt a rigid structure, the end-to-end distance should be directly proportional to the number of proline residues contributing to the polymer. Length calibrations on the nanometer length scale have been performed to experimentally determine FÖRSTER radii of fluorescent dye pairs which are then used in distance measurements in biological probes. Therefore, distance mapping in biology using FRET is based on the assumption that polyprolines form regular, rigid polyproline type II helices in aqueous solution.

Proline residues also induce rigidity in peptides that have proven to be responsible for folding kinetics in naturally occurring proteins due to the slow *cis-trans* isomerization of prolyl amide bonds. A deeper knowledge of prolyl isomerization in biologically relevant surrounding can thus help to understand protein folding processes better.

Polyproline derivatives with fluorescent dyes connected to the *N*-terminus and C-terminal tryptophan have been investigated through fluorescence spectroscopy and computational methods to verify the PPII helix formation of dissolved polyprolines. All employed fluorescent dyes can be quenched by tryptophan on VAN DER WAALS contact, i.e. at proximities in the order of 1 nm, and are therefore a reporter in terms of end-to-end contact formation between dye and tryptophan. The question whether *cis-trans* isomerization influences the polyproline structure in aqueous solution towards shorter end-to-end distances could thus be analyzed.

4.1.1. EXPERIMENTAL DATA

Steady-state fluorescence quenching experiments

Polyproline conjugates F-$(\text{Pro})_N$-Trp with three different fluorophores, F (MR121, MR113, and R6G), and $0 < N < 10$ were investigated using steady-state fluorescence spectroscopy, time-resolved spectroscopy, and fluorescent correlation spectroscopy. Steady-state fluorescence measurements revealed a strong dependence of the steady-state quantum yield, Φ_{ss}, on the number of proline residues. The steady-state quantum yield is calculated from the fluorescence emission, I, of the quenched molecule relative to I_0 of a non-quenched control sample. From investigating bimolecular interactions between tryptophan and various fluorophores, it is known that Φ_{ss} is reduced through static and dynamic quenching processes based on photoinduced electron transfer between fluorophore and tryptophan. Static quenching describes formation of efficiently quenched complexes, in which fluorophore and tryptophan favor a stacked configuration that is stabilized by hydrophobic interactions [62]. Dynamic quenching takes place during transient diffusional encounters of fluorophore and tryptophan [57, 63].

A bimolecular complex exhibits a stability (i.e. a binding energy between fluorophore and tryptophan) in the order of 10 kJ · mol^{-1} with an enthalphic contribution between 20 and 30 kJ · mol^{-1} [57]. These quenching mechanisms are only effective upon interaction between fluorophore and tryptophan, as seen from fluorescence emission of the control sample F-(Pro)$_6$, in which the quenching moiety, tryptophan, was missing. It exhibits a Φ_{ss} similar to that of the fluorophore itself.

Time-resolved fluorescence measurements of F-(Pro)$_N$-Trp conjugates revealed bi-exponential decays with a main nanosecond component, τ_2, and a weak picosecond component, τ_1. From the nanosecond lifetime $\Phi_{dyn} = \tau_2 / \tau_0$ was calculated. The dynamic quenching mechanism was found to be significant for conjugates with $N < 4$, to decrease monotonically with decreasing N, and to reach a minimum on the order of 30 % when fluorophore and tryptophan residue are labeled directly to each other. The weak picosecond component can be attributed to residual fluorescence of inefficiently quenched fluorophore-tryptophan complexes due to deviations from co-planarity (e.g. stacked, side-to-face or side-to-side interactions) [62].

From steady-state and dynamic quantum yield the static quantum yield $\Phi_{stat} = \Phi_{ss} / \Phi_{dyn}$ was calculated, which reflects the formation of essentially non-fluorescent complexes that are quenched on time scales below the ~ 50 ps temporal resolution of our time-correlated measurements and thus do not effectively contribute to the presented TCSPC measurements. Φ_{stat} as a function of the number of proline residues was found to follow a transition from ~ 0.1 to ~ 1 where the transition width depends on the fluorophore linker. An oxazine fluorophore connected through a C_2 carbon linker (MR113) showed a transition with a midpoint at four proline residues and a width of less than one residue when fitted to a sigmoidal function (Figure 38).

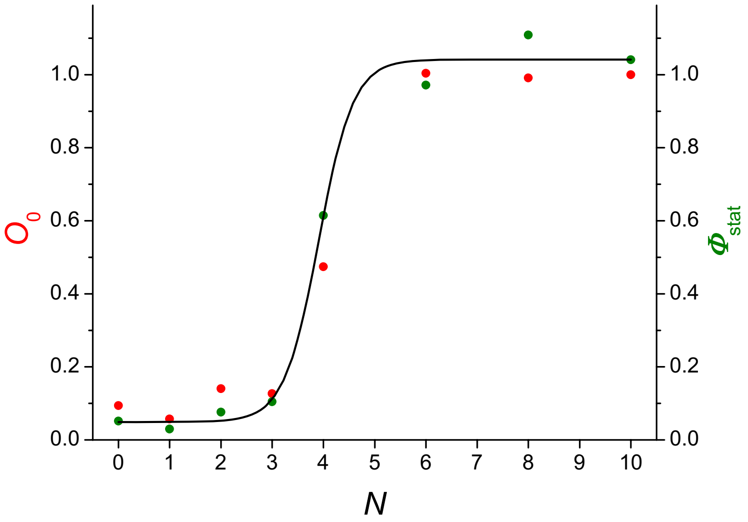


Figure 38. Occupancy and static quenching quantum yield versus proline length, *N*, for MR113-(Pro)$_N$-Trp.

The same oxazine fluorophore connected through a C_4 carbon linker (MR121) had a midpoint of 5 and a width of 2 proline residues (Figure 39). The broader transition results from the greater flexibility and length of the linker.

For R6G, which has an extremely rigid linker with only a single rotational degree of freedom, the transition was shifted to shorter peptides (Figure 40). These results demonstrate that an interaction between fluorophore and tryptophan resulting in efficient quenching and the width of the transition is determined by the length of the molecular linker and the total degrees of freedom for reorientation of fluorophore and the indole moiety of the tryptophan.

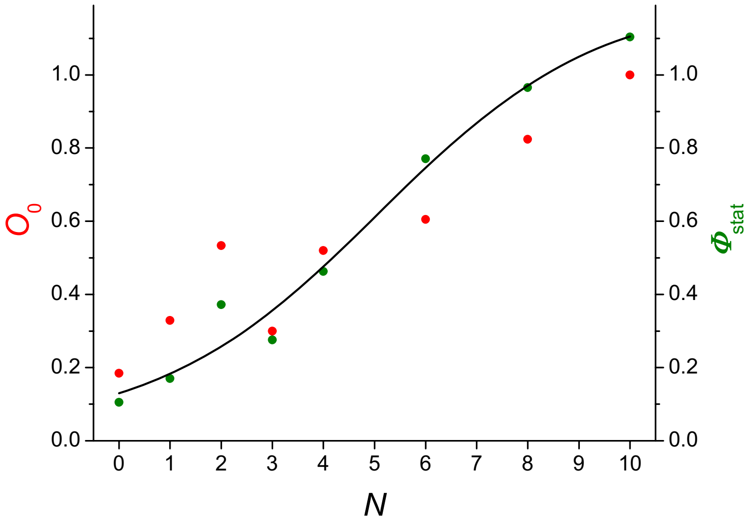


Figure 39. Occupancy and Static quenching quantum yield versus proline length, *N*, for MR121-(Pro)$_N$-Trp.

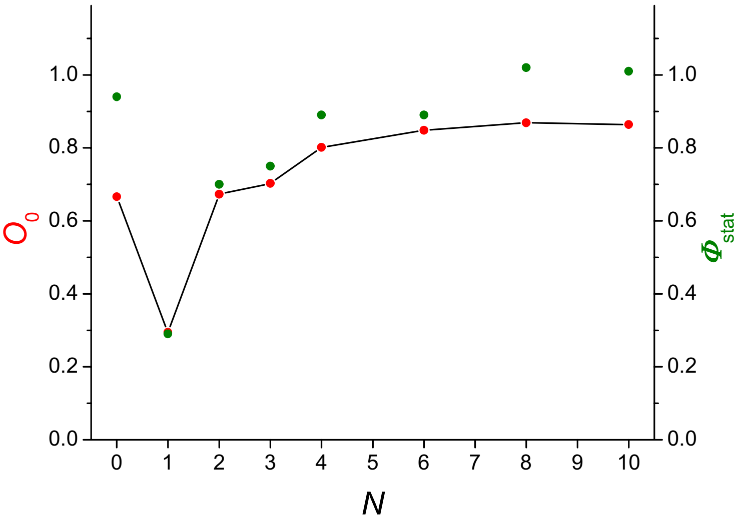


Figure 40. Occupancy and static quenching quantum yield versus proline length, N, for R6G-$(Pro)_N$-Trp.

Although *through-bond* photoinduced electron transfer has been observed in polyprolines separating strong organic donors and acceptors [64-67], any significant contribution in these experiments can be ruled out by the observation of ~ 100 % quantum yield for the R6G-tryptophan conjugate. This molecule prevents complex formation by the Rhodamine 6G’s short and stiff linker. Additionally, the use of saturated alkyl chains as fluorophore linker, and the steep sigmoidal distance dependence observed for MR113-tryptophan conjugates proof that no through-bond electron transfer occurs.

Fluorescence Correlation Spectroscopy

Having established the fact that in short polyprolines up to 90 % of all tryptophan-fluorophore pairs have formed a non-fluorescent complex, FCS was used to further investigate the underlying kinetics of complex formation and hence to probe conformational dynamics of the polypeptides. An FCS measurement, as presented in Figure 41 yields characteristic parameters for all independent processes which result in fluctuations of the detected fluorescence signal. One such process is translational diffusion of fluorescent molecules through the observation volume. The characteristic time constant, τ_D, for this process is on the order of 500 µs. Any additional independent stochastic process that results in fluctuations on the ns to ms time scale would appear as additive part as previously shown for various quenching interactions [63, 68-70].

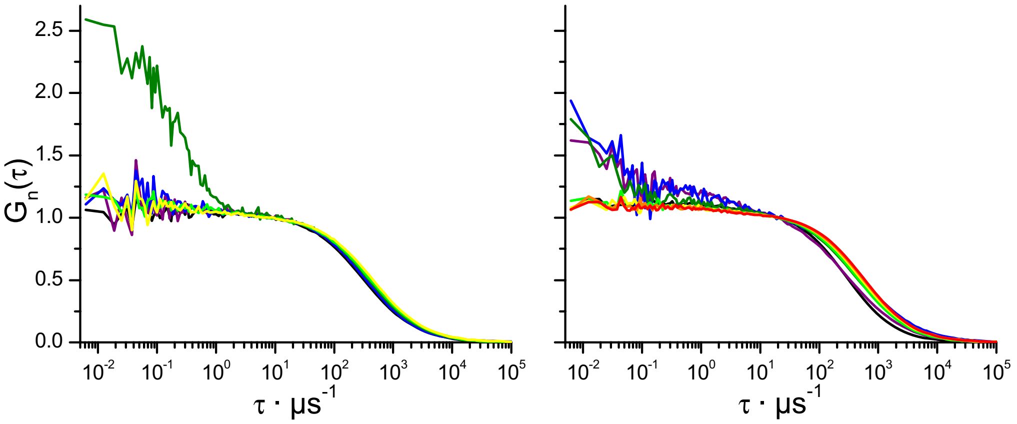


Figure 41. Correlation functions of MR113-$(Pro)_N$-Trp and MR121-$(Pro)_N$-Trp. The diffusion time is represented by the transition point of the decaying curves. See text for detailed description.

Fast processes with kinetics below 1 µs are observed for the MR113-$(Pro)_2$-Trp and MR121-labeled constructs with 0 - 2 proline residues. These fast fluctuations are caused by a subpopulation of the named samples; the major fraction (50 - 90 %) of these molecules show no such fluctuations.

Figure 38 - Figure 40 also show FCS results for F-$(Pro)_N$-Trp samples with F being MR121, MR113, or R6G, and $N = (0 - 4, 6, 8, 10)$ corresponding to the above discussed ensemble measurements. First the relative occupancy, O_0, was estimated, i.e. the average number of fluorescent molecules inside the observation volume normalized to the absolute sample concentration (estimated from the optical density measured at the fluorophore's absorption maximum). A strong decrease of O_0 with decreasing number of proline residues could be observed that is in excellent agreement with the amount of static quenching as estimated in ensemble measurements. The measurements reveal that static quenching is due to formation of non-fluorescent complexes which are stable for more than ~1 ms, since no association / dissociation kinetics contribute to the correlation curve and increase the overall amplitude.

Collisional quenching, on the other hand, decreases only the dynamic quantum yield without causing molecules to be switched off for extended periods of time (no influence on O_0) but resulting in a reduction of the brightness per molecule B. Figure 42 indeed shows very good agreement of Φ_{dyn} from time-resolved fluorescence measurements and B from FCS experiments for all samples.

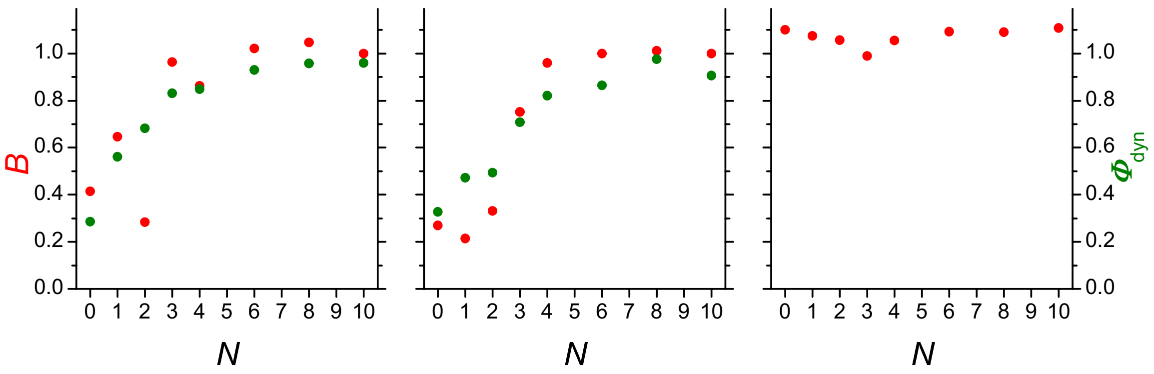


Figure 42. Brightness and dynamic quenching quantum yield versus proline length for F-(Pro)$_N$-Trp with fluorophores MR113 (left), MR121 (middle), and R6G (right), respectively.

The most striking result from FCS measurements is the absence of any fluctuation on time scales between 6 ns (limited by the temporal resolution of the FCS setup) and ~1 ms (limited by the molecules diffusion time through the observation volume) for most samples. In particular no fluctuations for nearly all MR113 samples were found with the exception of MR113-(Pro)$_2$-Trp, and for MR121-Pro-Trp and MR121-(Pro)$_5$-Trp. Correlation data for polyprolines labeled with R6G is not presented since R6G exhibits strong intersystem crossing resulting in a correlation decay on the microsecond time scale that could not be distinguished from complex formation kinetics.

These results demonstrate that quenched complexes form, und thus proline conformations exist, which are stable for time periods longer than the diffusion time (> 1 ms). It is important to remember that dynamic quenching is too fast to be reflected in FCS data (< 6 ns) and that fast fluctuations, as e.g. detected for MR113-(Pro)$_2$-Trp, originate only from those molecules which are represented by the overall FCS amplitude. For MR113-(Pro)$_2$-Trp the occupancy, O_0, amounts to about 10 % of the total sample concentration meaning that the additional fast correlation component originates from only this subpopulation and likely represents sterically unfavorable encounters between fluorophore and tryptophan. In other words, even though some samples show fast correlation components representing fluctuations of the quenching efficiency on the microsecond time scale, the majority of complexes, once they are formed, are stable for an extended period of time that is longer than our observation time (~ 1 ms).

This observation is in contrast to complex formation as observed in flexible gylcine-serine-peptides [63] and a polyproline peptide with two glycine residues incorporated (MR121-$(Pro)_3$-$(Gly)_2$-$(Pro)_3$-Trp). As glycine is known to introduce considerable flexibility in a polypeptide chain [71-73], both molecules are highly dynamic and reveal a strong nanosecond decay representing complex formation and dissociation as mediated by conformational dynamics of the peptide chain [63].

4.1.2. COMPUTATIONAL SIMULATIONS

In order to elucidate the influence of prolyl isomerization on end-to-end distances in polyprolines, molecular structures of the experimentally investigated constructs were set up using molecular mechanics force fields and energetically allowed structures with various distributions of *trans*/*cis*-bonds were modeled. A strong influence of the amount and position of individual *cis*-bonds on end-to-end distances was found. To mimic the experimental situation, the minimal distances d_{min} between fluorophore (MR113, MR121, and R6G; attached at the *N*-terminus) and tryptophan (last residue of the polypeptide) was determined in conformations that are sterically allowed for each given polyproline structure.

Polyproline structures were determined from an ideal all-*trans* PPII helix, as predicted from crystal structure [74] with individual prolyl-bonds forced into a *cis*-conformation. Following energy minimization yielded an energetically acceptable polyproline conformer incorporating the PPII distortion by a *cis* amide bond. Minimal distances between fluorophore and tryptophan residue, which predict if contact-induced quenching interactions are feasible, were calculated by analyzing the conformational space of fluorophore / linker and tryptophan residue. Sampling the rotational degrees of freedom for all bonds linking the planar fluorophore to the peptide backbone (and correspondingly those linking the planar indole moiety to the peptide backbone), and excluding those conformations that lead to steric clashes within the whole structure, point clouds representing allowed center of mass positions for fluorophore and indole moieties were built.

The conformational space that can be accessed by the fluorophore MR113 and tryptophan, respectively, is shown in (Figure 43). Sterically allowed conformations are shown as points representing the center of mass of the moiety's aromatic system. A polyproline in a PPII helical conformation with five residues in length connects the fluorescent dye and the tryptophan.

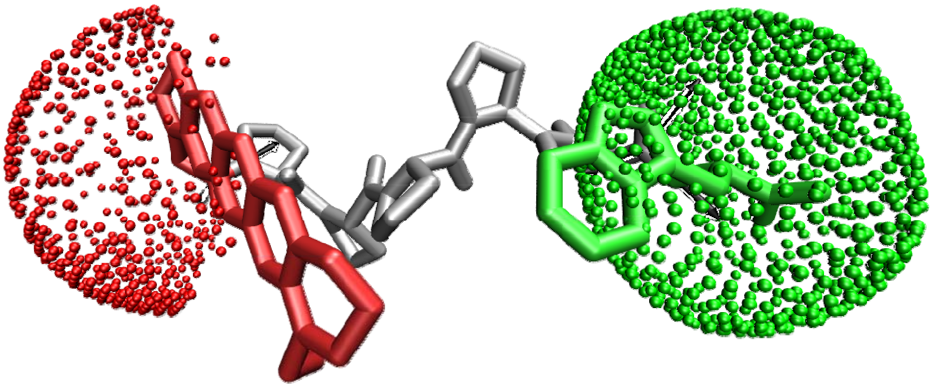


Figure 43. MR113-(Pro)$_5$-Trp in all-*trans* conformation. Point clouds of sterically allowed terminal residues are shown for the MR113 moiety (red) and tryptophan (green).

A conformational change of the proline backbone can induce significant changes to the fluorophore-tryptophan distance. If the prolyl-bond between the 3rd and 4th proline residues converts to *cis*-geometry, the fluorophore and the quenching tryptophan can approach to a distance at which photoinduced electron transfer becomes feasible (Figure 44).

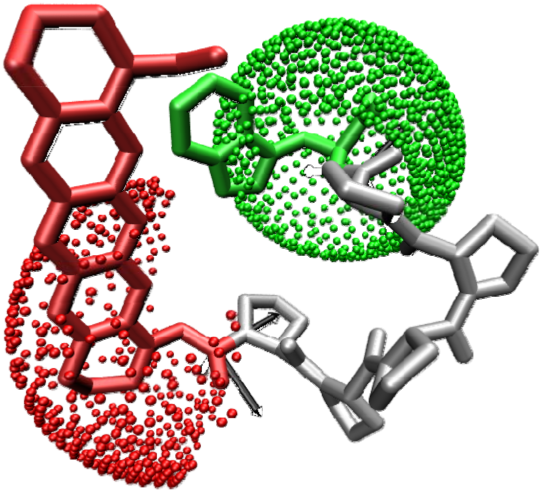


Figure 44. MR113-(Pro)$_5$-Trp in ttct-conformation. Point clouds of sterically allowed terminal residues are shown for the MR113 moiety (red) and tryptophan (green).

Obviously, MR121 with 4 rotational degrees of freedom in the C$_4$-linker samples a larger conformational space than MR113 with 2 rotational degrees of freedom

(Figure 45). Due to this grater flexibility the point cloud of allowed geometries extends further in angular terms. According to the longer linker length, the point cloud's diameter is increased compared to MR113.

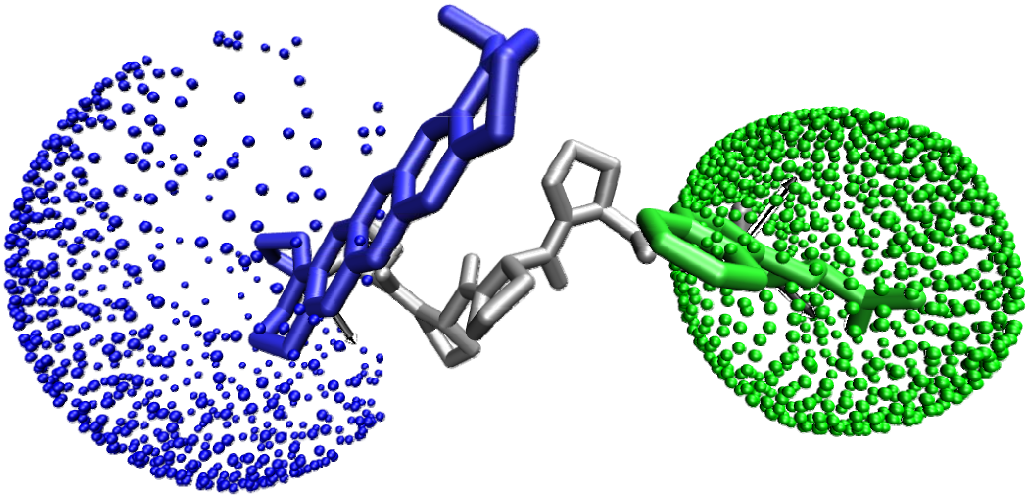


Figure 45. MR121-(Pro)$_5$-Trp in all-*trans* conformation. Point clouds of sterically allowed terminal residues are shown for the MR121 moiety (blue) and tryptophan (green).

Rhodamine 6G consists of two aromatic systems: the chromophore and a phenyl moiety perpendicular to the first. The phenyl group carries the carbonyl functional group in *para*-position. The only rotatable bonds, not involved in aromatic rings, are the bond between the carbonyl carbon and the phenyl ring and the bond between the two aromatic systems. Both bonds are co-axial and restricted in rotation as the both connect moieties with delocalized π-bonds. Therefore, Rhodamine 6G is extremely rigid and the point cloud degenerates to a small cylindrical distribution (Figure 46).

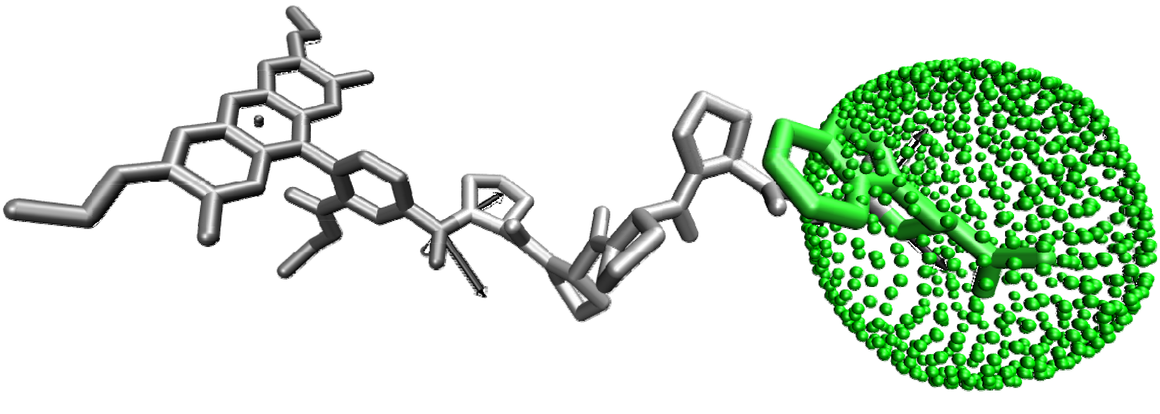


Figure 46. R6G-(Pro)$_5$-Trp in all-*trans* conformation. Point clouds of sterically allowed terminal residues are shown for the R6G moiety (grey, within the aromatic system's center) and tryptophan (green).

The minimal distance between the fluorophore's and tryptophan's center of mass is defined as d_{min} for any two given point clouds. Determination of d_{min} for point clouds separated by the various polyproline structures allows to assess in a conservative manner whether fluorescence quenching is, in principle, possible or not. In (Figure 47) d_{min} is plotted for a series of polyprolines MR113-(Pro)$_N$-Trp (with $2 < N < 20$) with single *cis*-bonds at the indicated position in the proline chain.

Compared to MR113-labeled constructs, for a given number of spacing proline residues, the distance, d_{min}, between fluorescent dye and tryptophan is 0.2 - 0.3 nm longer for MR121-labeled samples in a PPII helix. This difference agrees with the length dependence of the experimental fluorescence quantum yield showing a transition for MR113 that is shifted by about one proline residue (helical rise in PPII helix is 0.3 nm) compared to MR121. (Figure 48) shows the minimal dye-to-tryptophan distances for F-(Pro)$_N$-Trp for $N = 2 - 6$, 8, 10, 12, 15, 20 and F = MR113, MR121, and R6G, respectively. The all-*trans* geometry and any mono-*cis* structures are accounted for. The end-to-end distance of a proline chain in PPII-geometry (all-*trans*) is shown as a black line.

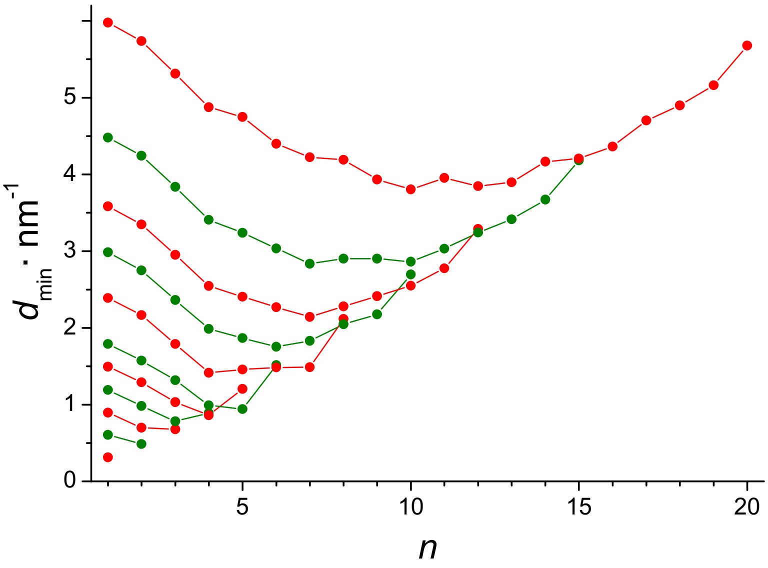


Figure 47. Dye-to-tryptophan distances of *n*-*cis*-MR113-(Pro)$_N$-Trp with N = 2 - 6, 8, 10, 12, 15, 20 (from bottom to top). All-*trans* geometries are added at $n = 1$ and $n = N$.

Apparently, d_{min} changes dramatically upon introduction of a single *cis*-isomer, in particular when positioned at the center of the proline chain. A photoinduced electron transfer quenching interaction becomes feasible when fluorophore and tryptophan are at VAN DER WAALS contact which has been estimated to occur for d_{min} between 0.5 nm and 1.0 nm [62] (indicated by the dotted lines in Figure 48).

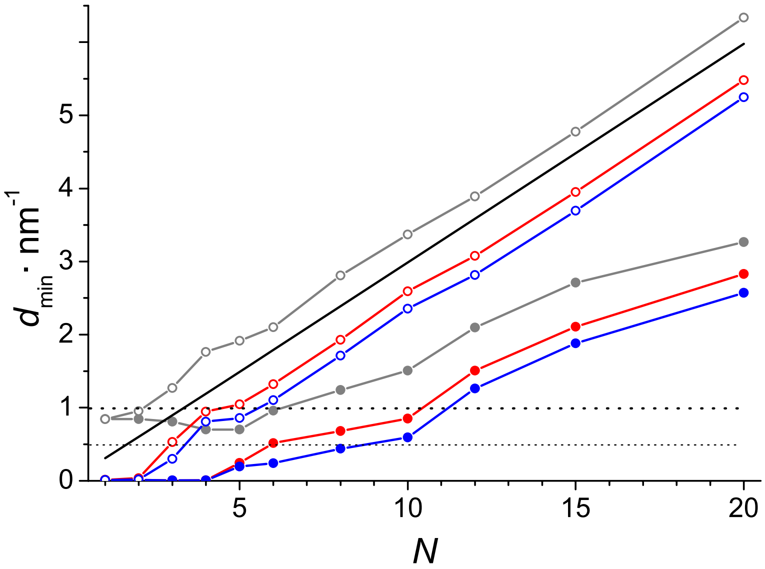


Figure 48. Dye-to-tryptophan distances of *n-cis*-Dye-(Pro)$_N$-Trp (full circles) and all-*trans*-Dye-(Pro)$_N$-Trp (open circles) with Dye = MR113 (red), MR121 (blue), and R6G (gray), respectively.

If the isomerization of more than one prolyl amide bond to *cis* configuration is taken into account, the minimal possible distances between the fluorophore and tryptophan decrease even more drastically. Figure 49 shows these distances versus the polymer length.

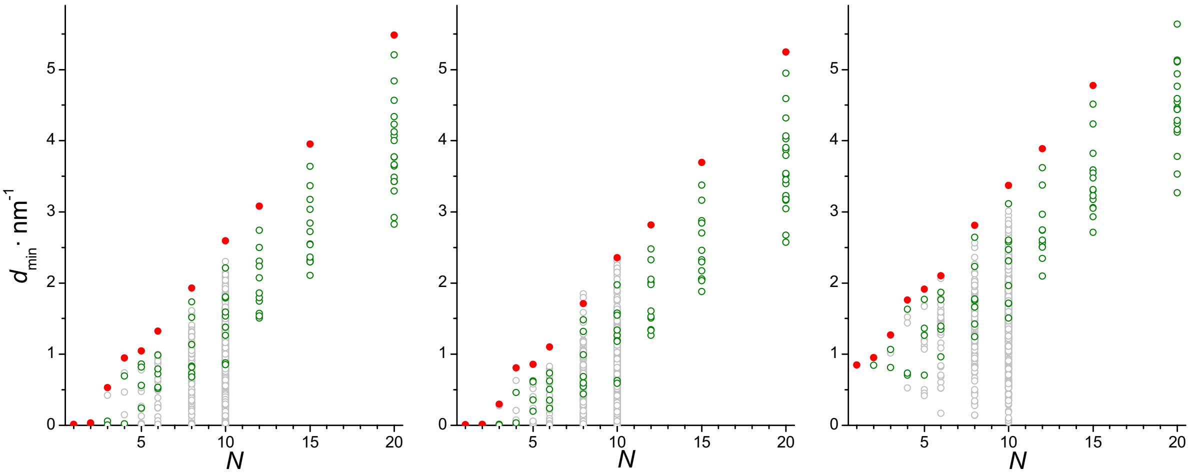


Figure 49. Dye-to-tryptophan distances of F-(Pro)$_N$-Trp with F = MR113 (left), MR121 (middle), and R6G (right), respectively. All-*trans* conformations are colored red, mono-*cis* green, and poly-*cis* gray.

For polyproline lengths above 10 monomers poly-*cis* isomers were not analyzed due to the enormous number of possible conformers scaling with $2^{(n-1)}$. However, introduction of more than one amide bond in *cis* conformation can reduce the end-to-end distance of polyprolines arbitrarily although in aqueous solution this would be very unlikely to happen.

The computational results demonstrate that prolyl *trans*/*cis* isomerization in all-*trans* polyproline helices allows contact formation between fluorophore and tryptophan to occur even in longer constructs where both moieties are well separated in the all-*trans* PPII configuration. This further explains why photoinduced electron transfer quenching is sensitive to prolyl-isomerization as shown in the above discussed experiments and suggests that, given the non-zero probability of prolyl *cis* isomers, average end-to-end distances in a polyproline ensemble should be significantly smaller than expected for a pure PPII structure. As the rate of prolyl isomerization was estimated to be on a minute time scale, the populations of various isomers form an equilibrium distribution that does not change significantly during the above discussed ensemble and FCS experiments in aqueous solution.

The data in Figure 48 can not prove the occurrence of PET quenching in structures with d_{min} below a certain threshold. Nevertheless, it can predict structures in which PET quenching is strictly prohibited. As experimental quantum yield's clearly indicate PET quenching interactions in F-(Pro)$_N$-Trp with $N < 8$ (for MR121), $N < 5$ (for MR113), and $N < 4$ (for R6G), these calculations suggest that the experimental result is not consistent with a polyproline PPII helix but can be explained by interspersed *cis*-prolyl isomers.

4.2. Polymer Properties of Polythymines

Polythymines show typical polymer properties: They form linear chains that can interact with the surrounding solvent, with existent ions and with itself. These interactions and the fact that the polythymine sugar backbone is restricted in terms of conformational freedom result in chain geometries that can be summarized by standard polymer models.

The structure of a solvated chain molecule defines how it interacts with the solvent and therefore characterizes the diffusion of the polymer due to BROWNian motion.

Therefore FCS measurements which can monitor and quantify diffusion can be used to derive polymer properties. Spectroscopic methods rely in turn on the fluorescence of the probe or an artificially introduced moiety of the probe. As autofluorescence of DNA shows poor intensity, investigated probes need to be labeled with organic fluorophores. These obviously alter the probe and can influence its structural properties so that diffusion constants determined through FCS measurements cannot be transferred to unlabeled samples. To number the label's weight on polymer properties of the studied polythymines, spectroscopic results have been compared to molecular dynamic simulations of dye labeled and pure polythymines.

The combination of FCS and molecular dynamics simulations can thus reveal the influence of fluorescent dye labeling on the polymer properties.

4.2.1. EXPERIMENTAL DATA

Using FCS, translational diffusion of polydesoxythymine, MR121-$(dT)_N$, can be monitored in aqueous solution. The nucleic acid polymer has to be labeled site-specifically with the oxazine derivative MR121 at the 5'-end.

Thymine bases are known to exhibit negligible base-pairing interactions, and the weakest base-stacking interactions compared to all other nucleobases. Polythymine is known to behave as an unstructured polymer that randomly samples a conformational space depending on its contour length, i.e., the total length along the polymer backbone. MR121-$(dT)_N$ with various contour lengths by changing N from 5 to 100 were investigated.

As a fluorescent molecule diffuses through the confocal observation volume, it will emit photons in the red spectral region (emission maximum of MR121 is 673 nm) upon excitation with 633 nm laser light. For a fixed observation volume, which is determined by the laser excitation profile and the detection optics, the observed fluorescence will fluctuate with a characteristic time constant that depends on the diffusion constant of the molecule. Analyzing such fluctuations by calculating the autocorrelation curve of the fluorescence signal yields a time constant, τ_D, that is inversely proportional to the diffusion constant, D, and directly proportional to an apparent hydrodynamic radius as defined by the Stokes-Einstein equation (2.22). All observed correlation functions $G(\tau)$ could be well fitted by the appropriate analytical expression (2.68), independent of polymer size.

As the contour length of polythymine was increased, FCS curves were shifted toward longer diffusion times, revealing an increase of the apparent hydrodynamic radius. The accuracy of our diffusion measurements using FCS on a custom-made confocal microscope was determined by estimating the relative standard deviation of repeated measurements (measurement time 100 s, excitation power 50 mW) and found to be ~ 5% for most of the presented experiments (the accuracy can be improved to < 1% using longer measurement times and increased excitation power). All FCS measurements were performed at a concentration of ~ 1 nM, making intermolecular interactions negligible. A small contribution from photophysical processes was found to be dependent on excitation power (with an amplitude K well below 0.05 at an excitation power of 50 mW [75]), but independent of the investigated sample. Diffusion times were independent of sample concentration and excitation power for excitation powers well below the saturation limit. Assuming a diffusion constant of $D = 4.26 \times 10^{-6}$ cm^2·s^1 for MR121 (as measured for the structurally similar fluorophore Atto655 [48]), diffusion constants were calculated for polythymine to range from ~ 2.26×10^{-6} cm^2·s^1 to ~ 4×10^{-5} cm^2·s^1.

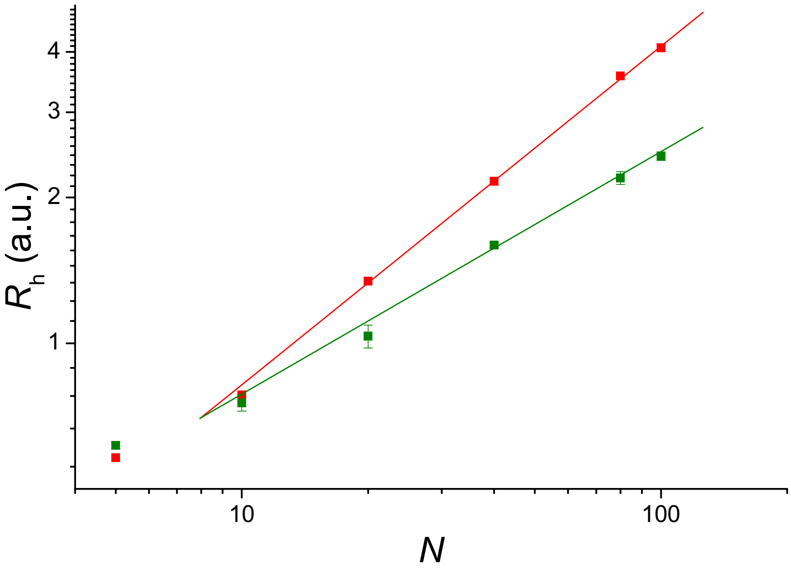


Figure 50. Hydrodynamic radii of MR121-(dT)$_N$ from FCS experiments at NaCl concentrations of 10 mM (red) and 1 M (green).

Inverse diffusion constants and, thus, hydrodynamic radii R_H were observed to scale with contour length (expressed through the number of dT residues, N) as $R_h \propto N^{\nu}$ for polynucleotides with more than ten residues (Figure 50). Below ten bases, the influence of the fluorophore on the diffusion time disguises polymer properties. For

MR121-(dT)$_N$ with 10 < N < 100 v varied between 0.5 (1 M NaCl) and 0.7 (10 mM NaCl) as a function of the ionic strength of the solvent.

In Figure 51, the dependence of R_h on ionic strength is shown in detail for MR121-(dT)$_{100}$. Ionic strength was changed by varying the concentration of sodium chloride from 0 to 3 M in bi-distilled water. Identical results were found in 10 mM phosphate buffer for NaCl concentrations > 10 mM. The hydrodynamic radius scales with ionic strength I as $R_\mathrm{h} \propto I^{\alpha} R_\mathrm{h}$ with α = (-0.11 ± 0.01) for NaCl concentrations < 10 mM. For NaCl concentrations below 5 mM, R_h saturates at a constant level.

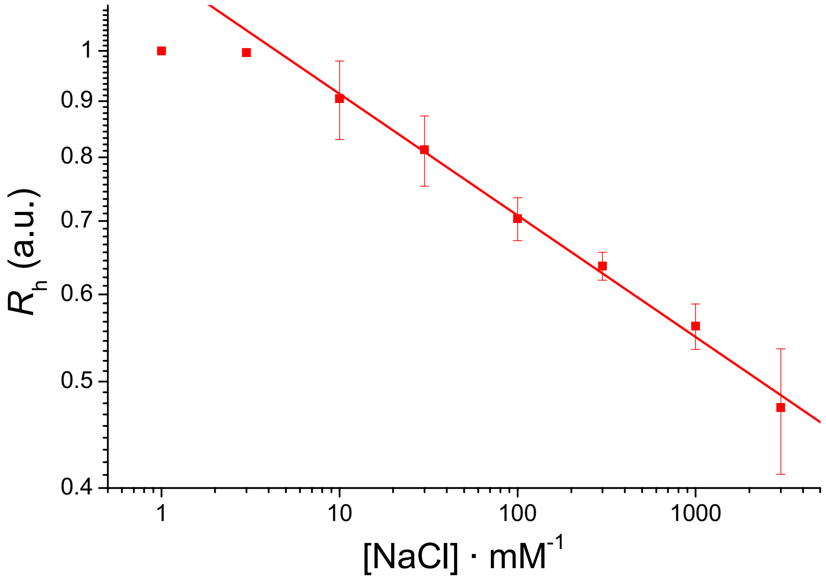


Figure 51. Salt dependency of the hydrodynamic radius of MR121-(dT)$_{100}$. For monovalent salt concentrations below ~ 5 mM, R_h saturates at a constant level whereas higher concentrations induce a contraction of the polymer.

4.2.2. MOLECULAR DYNAMICS SIMULATIONS

The conformational space of polythymine derivatives was sampled using an Amber force field that was limited to short-range (VAN DER WAALS) interactions in vacuum. The approach constitutes a hard-sphere model of ss-DNA in which excluded volume is set by VAN DER WAALS radii. By this approach, all constraints on conformational dynamics given by bond length, bond angles, and bond rotations, as well as excluded volume due to steric interactions could be explored.

Four different polythymine derivatives were investigated: (dT)$_N$, MR121-(dT)$_N$, (dS)$_N$, and (dT$_\mathrm{pen}$)$_N$. Simulations of (dT)$_N$ and MR121-(dT)$_N$ were performed to probe the differences between MR121-labeled and non-labeled polythymine. A polynucleotide made of abasic sites $(dS)_N$, i.e., a sugar-phosphate backbone without attached base moieties, was investigated to test whether the attached bases have any influence on polymer conformations. A hypothetical polynucleotide, $(dT_{pen})_N$, that is self-penetrable but consists of the same bond configurations and residues as polythymine, was simulated to identify any excluded-volume effect. The simulation was realized by switching off all VAN DER WAALS interactions.

To ensure sufficient sampling, simulations at temperatures ranging from 200 K to 1000 K were performed for $(dT)_{16}$ (Figure 52) . At low simulation temperatures the polymer is not able to overcome rotational barriers within a reasonable simulation time. This results in shifted end-to-end distance and radius of gyration distributions towards higher values. At a nominal temperature of 800 K, sufficient sampling of the geometrically allowed conformational space without severe stabilization of individual states due to local energy minima was observed. The temperature was still low enough to allow for 1 fs integration steps without producing instabilities during numerical integration of NEWTON's second law.

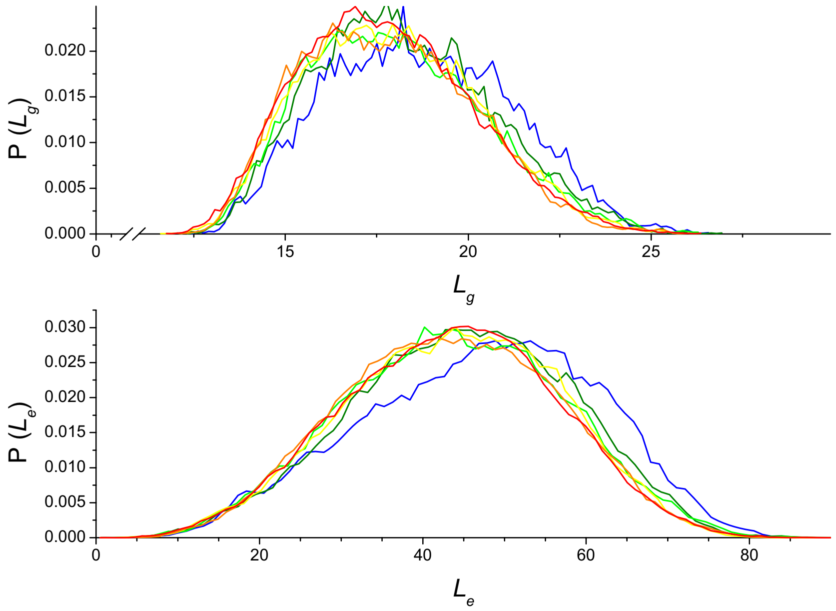


Figure 52. Radius of gyration (top) and end-to-end distance distributions (bottom) for $(dT)_{16}$ at temperatures between 200 K (blue) and 1200 K (red).

For each molecule, with contour length fixed by the specified number of residues, three parameters to characterize the polymer were extracted:

a) end-to-end distances, L_e
b) radius of gyration, L_g
c) persistence length, L_p

Average values of the square end-to-end distances and the radius of gyration were calculated directly from the corresponding distribution functions.

First, average values of squared end-to-end distances $\left\langle L_e^2 \right\rangle$ and radii of gyration $\left\langle L_e^2 \right\rangle$ were investigated. (Figure 53) shows $\left\langle L_e^2 \right\rangle$ as a function of the polymer's contour length, i.e. the number of residues, N for $2 \leq N \leq 128$. The polymers in which electrostatic forces and repulsive van der Waals interaction are included show a dependency in the contour length of $\left\langle L_e^2 \right\rangle \propto N^{\mu}$ with an exponent $\mu > 1$. An ideal polymer would scale linear with the contour length.

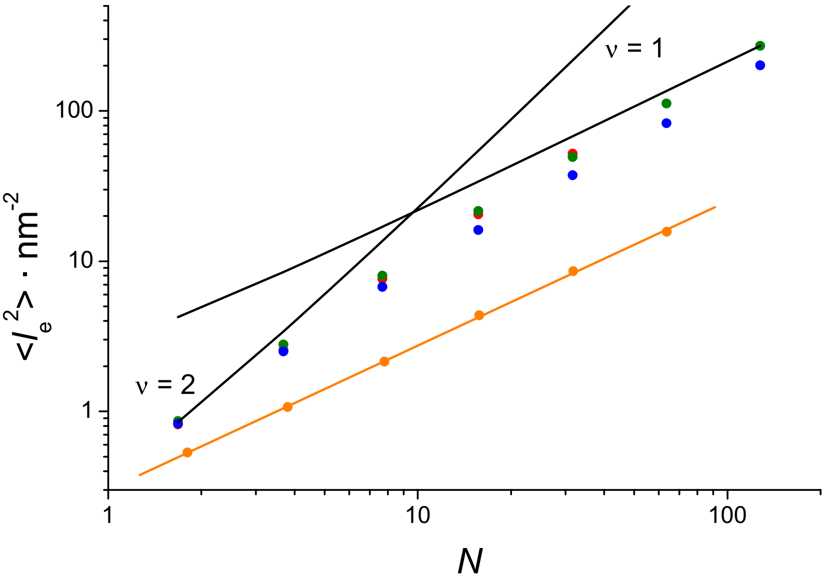


Figure 53. Average square end-to-end distances of $(dT)_N$ (green), MR121-$(dT)_N$ (red), $(dS)_N$ (blue) and $(dT_{pen})_N$ (orange), Black lines indicate power law dependencies with an exponent of μ for ideal polymers ($\mu = 1$) and rigid rods ($\mu = 2$).

Therefore, van der Waals and electrostatic interaction cause the polymer to swell. This is confirmed by an end-to-end distance power-law fit of the $(dT)_N$ polymer data from simulations without the influence of any van der Waals or electrostatic forces: $(dT_{pen})_N$ yields an exponent $\mu = 0.96 \pm 0.01$ and thus validates the absence of excluded-volume effects. Two power-law functions are included in Figure 53 for comparison: a function with $\mu = 1$ which represents an ideal freely jointed chain polymer behavior and a function with $\mu = 2$ reflecting a rigid-rod like polymer. The penetrable polythymine behaves as a ideal freely jointed chain. In contrast the polymers being simulated with repulsive interactions behave like rigid rods for very short chain lengths, i.e. on length scales where the contour length is much smaller than the persistence length. At greater contour lengths these simulations show the swollen polymer properties of nucleic acid chains.

Nucleobases do not influence the behavior of the ss-DNA significantly. At short contour length scales stable abasic site chains are indistinguishable from polythymine with respect to end-to-end distance. Higher numbers of residues show abasic site polymers in a slightly more compressed average conformation compared to $(dT)_N$. Therefore the assumption that base stacking plays a minor role in polythymine structure formation is enforced.

The fluorophore MR121 does not significantly influence the structure of polythymine, either. The plots of end-to-end distances versus polymer length for pure $(dT)_N$ and MR121-$(dT)_N$ overlay perfectly. This emphasizes the assumption that experimental data gained from measurements of structure-relevant properties can be performed with spectroscopic means. Polymer structure seems to be independent of dye labeling in ss-DNA.

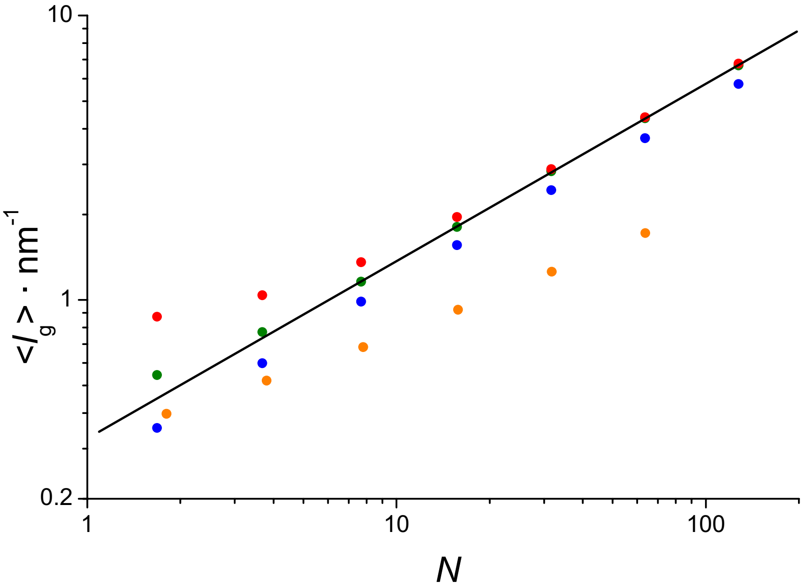


Figure 54. Average radii of gyration of $(dT)_N$ (green), MR121-$(dT)_N$ (red), $(dS)_N$ (blue) and $(dT_{pen})_N$ (orange). A power law fit to $(dT)_N$ with $N > 5$ is shown in black.

Although internal structure seems to be untouched by the attachment of a MR121 moiety, the fluorophore influences short polymers regarding total polymer dimensions: The radius of gyration of MR121 labeled polythymines deviates from unlabeled at short polymer lengths (Figure 54). Up to ~ 20 residues the dye increases l_g compared to the pure nucleic acid chain.

The radius of gyration scales perfectly well for unlabeled polythymine over all sizes with $2 < N < 128$ and follows a power law with exponent 0.62 ± 0.01. This result is in perfect agreement with the experimental observation that the hydrodynamic radius, being proportional to $<L_g>$, scales as a power law for more than 10 - 20 residues and deviates for short MR121-(dT)$_N$ in a similar way as $<l_g>$ does (Figure 55). Overall, the length dependence of $<L_g>$ for labeled polythymine is in excellent agreement with experimental R_h of polythymine in aqueous buffer with 100 mM NaCl as Figure 55 shows. The measured power law exponent of 0.58 ± 0.02 is furthermore in excellent agreement with the theoretical exponent for a polymer in which excluded volume effects contribute, numerically estimated to 0.588 [76].

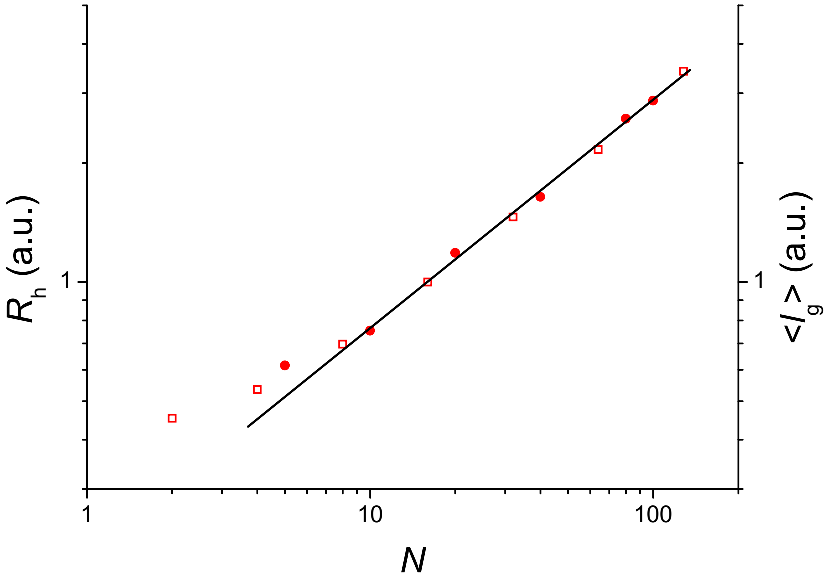


Figure 55. Hydrodynamic radii from FCS experiments (squares) and average radii of gyration from simulations (circles) of MR121-(dT)$_N$. Data are scaled to arbitrary units such that data points for MR121-(dT)$_{10}$ overlap.

The length dependence of the polymer dimension can be visualized by FLORY's characteristic ratio

$$F_N = \frac{\langle L_e{}^2 \rangle}{N \cdot l^2} \tag{4.1}$$

The quotient in equation *(4.1)* is the average square end-to-end distance of an equivalent freely jointed chain. An equivalent chain with a virtual segment length $l = 0.66$ nm can be assumed. This is the average distance between two phosphor atoms in the nucleic acid polymers. The number of segments N is equal to the number of residues. The contour length yields in $L_c = N \cdot l$. FLORY's ratio follows a strong increase over a characteristic length of ~ 20 residues asymptotically approaching the limit of C_∞ = 5 - 6 for $(dT)_N$ and C_∞ = 3 - 4 for $(dS)_N$. The observation that C_∞ is length independent and close to 1 for $(dT_{pen})_N$ shows that steric interactions of side-groups influence polymer flexibility to a larger degree than constraint bond angles do.

The ratio of average square end-to-end distance and average square radius of gyration is a scale referred to as the structure factor, G:

$$G = \frac{\langle L_e^{\,2} \rangle}{\langle L_g^{\,2} \rangle} \tag{4.2}$$

The structure factor of a freely jointed chain is G_{FJC} = 6, whereas G_{RR} = 12 for a rigid rod [76, 77]. With increasing contour length it increases for all polymers except $(dS)_N$ and asymptotically approaches a limit of 6 to 7 for ~ 20 residues. The asymptotic value is close to the structure factor of an ideal freely jointed chain and again reflects the fact that ss-DNA in the limit of long chains is well described as a polymer with large although not unrestricted flexibility. Structure values are reduced for shorter polymers. This is due to the fact that base residues and MR121 contribute more to L_g than they do to L_e.

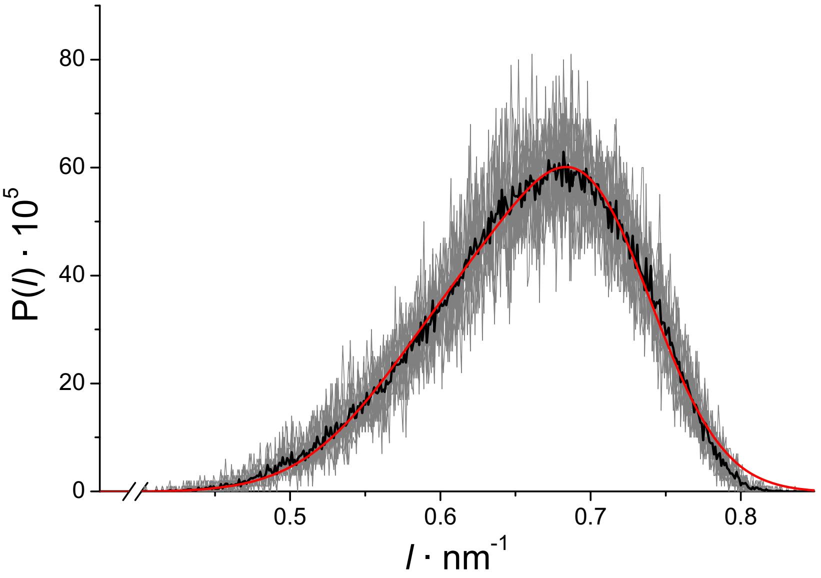


Figure 56. Segment length distribution of $(dT)_{16}$. Histograms were created for each P-P segment's length over simulation time. An overlay of all P-P segment length histograms is shown in grey, the average over all l in black and a double GAUSSian fit as a red line.

Radial distributions of the nucleic acid polymers can be compared to the wormlike chain model. To judge polymers of different lengths against each other the contour length is expressed in units of persistence length:

$$l_{\mathrm{e}} = \frac{L_{\mathrm{e}}}{L_{\mathrm{p}}} \tag{4.3}$$

Polynucleotides with N = (64, 32, 16, 8, 4) and wormlike chains with contour lengths according to the nucleic acid chains were compared. The contour length was estimated by extracting average distances between successive phosphorus atoms in the DNA backbone (Figure 56) . A mean P-P distance for each polymer was extracted that is independent of the bond position in a single polynucleotide, and independent of the number of residues N.

All P-P distances were distributed with a standard deviation of ~ 20% of the mean value and could be fitted by a double GAUSSian distribution, f:

$$\mathrm{f} = P_0\left(l\right) + \frac{a_0}{\omega_0 \cdot \sqrt{\pi/2}} \cdot e^{-2\cdot\left(\frac{x - x_{c0}}{\omega_0}\right)^2} + \frac{a_1}{\omega_1 \cdot \sqrt{\pi/2}} \cdot e^{-2\cdot\left(\frac{x - x_{c1}}{\omega_1}\right)^2} \tag{4.4}$$

The good fit using two GAUSSian distribution implies that two conformers of the sugar backbone are present: a conformation with a mean P-P distance of 0.63 ± 0.07 nm and a more extended component with 0.70 ± 0.01 nm. These P-P distances fit the distances stated by OLSON [8, 78] within ~ 10 % and resemble different sugar puckering. The $C^{3'}$-exo conformer of the desoxyribose ring extends the segment length compared to the $C^{2'}$-exo conformation. According to the double GAUSSian fit these two ribose forms in average contribute to the distribution with a ratio of 60 : 40. With respect to the BOLTZMANN distribution this reflects an energy difference between the two conformers ΔE = 2.65 kJ · mol^{-1}.

A further partitioning into smaller virtual bonds, e.g., partitioning into the two virtual bonds P-C5’ and C5’-P, showed no advantage. While the length deviations from the mean for P-C5’ were low as the two atoms are separated by two bonds only, the relative length fluctuations for the virtual bond C5’-P were comparable to those of P-P vectors. These results are consistent with earlier calculations by OLSON and FLORY, who showed that ss-DNA can be described as a freely rotating chain with virtual bonds of constant length representing individual P-P vectors.

The average length of P-P segments, l, was 0.66 ± 0.07 nm (std. dev.), resulting in a contour length of $L_c = l \cdot N$ between 1.3 nm for $N = 2$ and 84.5 nm for $N = 128$. To show in which size range and to what degree the simulated distribution function resembles that of a wormlike chain, simulations with corresponding wormlike chain distribution functions were overlaid. Correspondence was based on matching contour length and average square end-to-end distance L_e. Wormlike chain distribution functions were calculated using the analytical expression given by THIRUMALAI and HA [19]. The plots (Figure 57) demonstrate that G(r) for polythymine and G(r) for the corresponding wormlike chain show better agreement with increasing contour length, becoming nearly equal for ~ 30 residues. For $(dS)_N$ and $(dT_{pen})_N$, simulated and ideal wormlike chain distributions show good agreement for slightly shorter contour lengths of ~ 10 residues, again reflecting the influence of excluded volume.

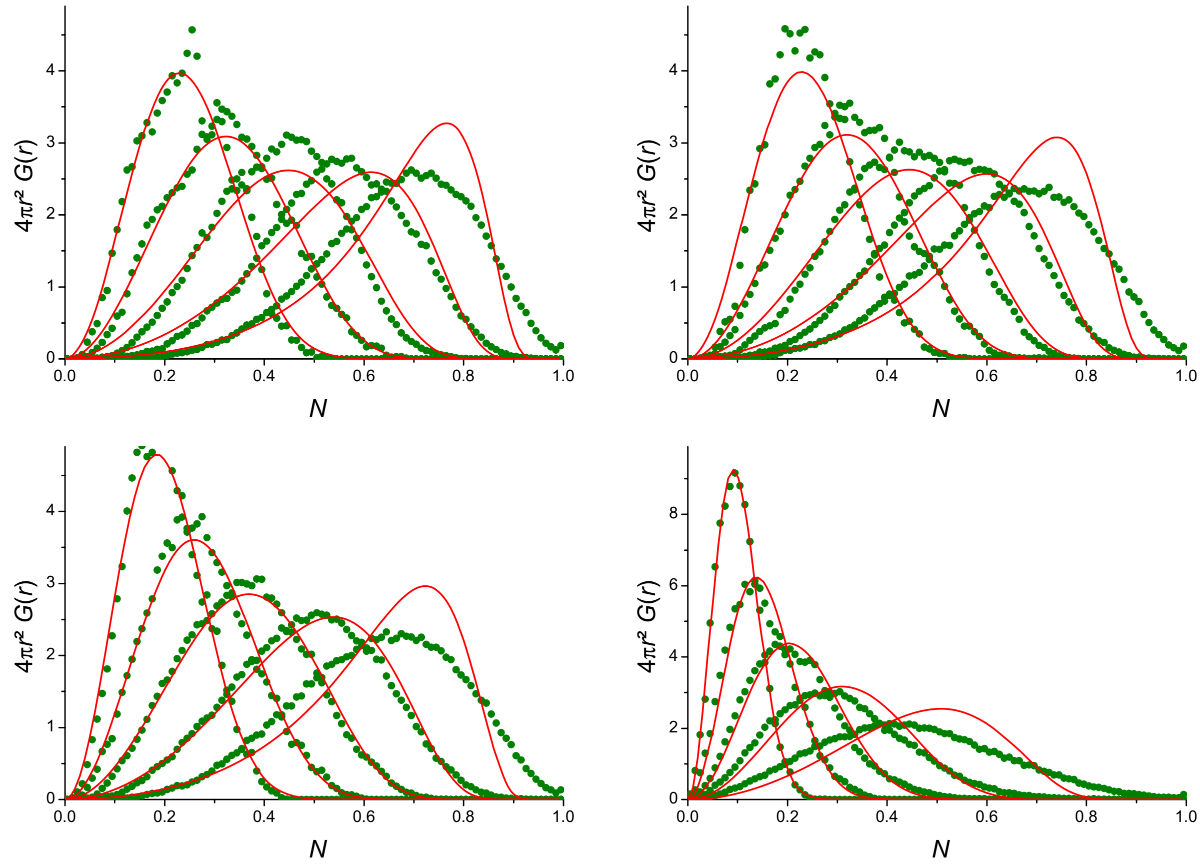


Figure 57. Radial end-to-end distance distribution functions of $(dT)_N$, MR121-$(dT)_N$, $d(S)_N$, and $(dT_{pen})_N$ (from top left to bottom right). End-to-end distances from molecular dynamics simulations are shown as green circles, wormlike chain distributions as red lines. All sets of curves show $(dN)_N$ with N = (64, 32, 16, 8, 4), respectively, from left to right, contour length is given in units of persistence length.

Attachment of the fluorescent label MR121 at the 5'-end of the ss-DNA has only a weak effect on G(r). (Figure 58) shows the simulated end-to-end distances for polythymine with and without MR121. End-to-end distances are calculated from the first nucleic residue's $O^{5'}$ to the last residue's H^{3T}, neglecting the fluorophore if present. As stated before, the fluorophore does not influence the polymer backbone structure and therefore does not contribute to nucleic acid chain properties.

When plotting distribution functions $P(L_g)$, where the fluorophore is now included in the calculation of L_g, a difference between labeled and unlabeled molecules appears as illustrated in Figure 58. Obviously the difference is more pronounced for shorter polythymines, as it is visible in the $P(L_g)$ overlay for MR121-$(dT)_N$ and $(dT)_N$. Whereas $P(L_g)$ for short polythymines is considerably broadened and shifted toward larger values upon MR121 attachment, the fluorophore's influence on $P(L_g)$ decreases with increasing length and becomes insignificant for ~ 60 thymine residues. Distribution functions are similar in shape for all molecules with four or more residues and resemble a GAUSSian distribution with a slight asymmetry.

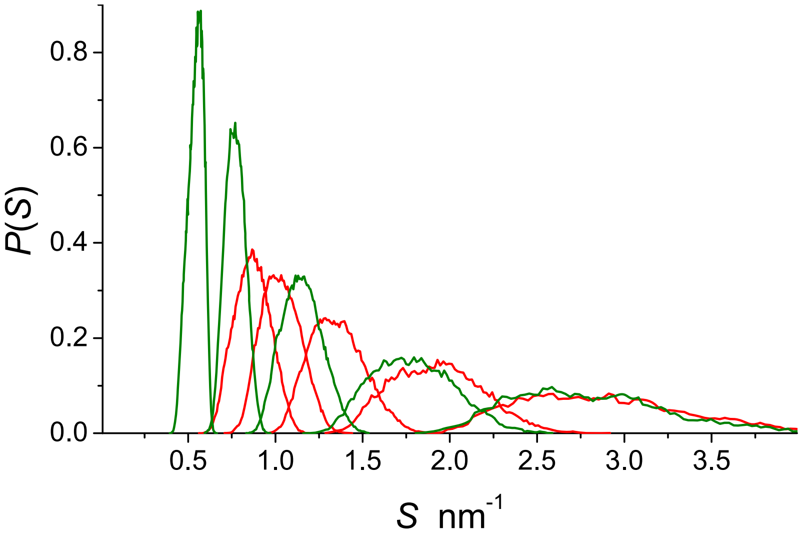


Figure 58. Radius of gyration distribution estimated by molecular dynamics simulations. Distribution functions for polythymine are shown in green, for MR121-labeled polymer in red. Plots represent polymers of 2, 4, 8, 16 and 32 residues in length, respectively (from left to right). Functions are normalized to equal integrals.

Last, the persistence length from simulated data was estimated and the experimental observation of a salt dependent hydrodynamic radius interpreted by an electrostatic persistence length that changes with ionic strength. The persistence length can be estimated in multiple ways from simulated data [46]: First, the orientational correlation length $L_{p,oc}$ and projection length $L_{p,pro}$ were calculated. Projection length values around 1.5 nm are weakly dependent on the contour length and slightly smaller for abasic-site polymers. A fit of orientational correlation functions revealed non-exponential decays and thus did not yield reliable $L_{\mathrm{p,oc}}$ values. Second, the persistence length $L_{\mathrm{p,WLC}}$ for a wormlike chain model was calculated with equal $\langle L_{\mathrm{e}}^2 \rangle$ and corresponding contour length compared to simulated data. The relation between $\langle L_{\mathrm{e}}^2 \rangle$, contour length L_{c}, and persistence length $L_{\mathrm{p,WLC}}$ is determined by the following equation [77]:

$$\langle L_{\mathrm{e}}^2 \rangle = 2L_{p,WLC} \cdot \left(e^{-\frac{L_c}{L_{p,WLC}}} - 1 + \frac{L_c}{L_{p,WLC}} \right) \tag{4.5}$$

It is important to remember that an increasing deviation between WLC and simulated end-to-end distance distributions was found for chains with less then ~ 30 residues.

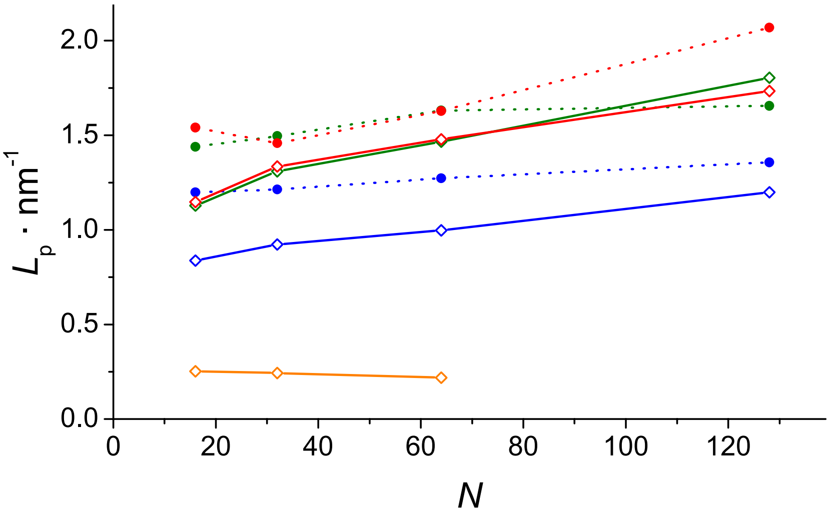


Figure 59. Persistence lengths of $(\mathrm{dN})_N$ for N > 16. Solid lines show L_p determined from the corresponding wormlike chain model distribution, dashed lines L_p derived from the projection length calculated from molecular dynamics simulation data. The persistence length of $(\mathrm{dT_{pen}})_N$ was too small to be estimated by projection length.

This is again reflected in the plot of $L_{\mathrm{p,WLC}}$ as a function of contour length (Figure 59), where a slight length dependence appears. Overall, a persistence length of ~ 1.7 nm for polythymine $(\mathrm{dT})_{128}$, ~ 1.2 nm for abasic sites $(\mathrm{dS})_{128}$, and ~ 0.2 nm for penetrable polythymine $(\mathrm{dT_{pen}})_{128}$ were determined. The conclusion that polythymine with 30 - 100 residues can be well described as a wormlike chain with respect to geometric and steric constraints could therefore be drawn.

5. Conclusion and Outlook

Spectroscopic measurements in combination with computational structure prediction methods reveal new insights into biopolymer behavior. Time resolved spectroscopy in the form of time-correlated single photon counting and FCS give a detailed view on molecular behavior on timescales ranging from nanoseconds to microseconds. Computational simulations can show intramolecular structure variations either as possible static states of a molecule or as dynamic time trajectories on timescales of nanosecond within reasonable computation time. The combination of theoretically acquired data with experimental results can therefore help to better understand molecular mechanisms involved in biopolymers.

The first part of this thesis presents new insights in the structure of poly-L-proline polymers. During the past decades deviations of calculated end-to-end distances of polyprolines and experimental ones derived from FRET measurements have been intensively discussed [27-29]. Whereas shorter experimentally observed end-to-end distances were explained by continuous bending of the polymer resembling a wormlike chain behavior on the one hand [29], interspersed *cis* isomers within the favored *trans* conformation have previously been suggested on the other hand [28, 79, 80]. However, experimental evidence for strong heterogeneities in polyprolines have never been reported, yet. Average FRET efficiencies from polyproline samples on a length scale of 1 - 24 amino acids have been analyzed by comparison to a wormlike chain model and persistence lengths between 2 and 4 nm have been reported [27-29].

The combination of time-correlated single photon counting and fluorescent correlation spectroscopy of polyprolines proposes that subpopulations of polyprolines with a reduced end-to-end distance exist. These more compact conformers of polyproline show no fluctuations within the observation time in fluorescent correlation spectroscopy experiments (~ 1 ms). Considering the nanosecond timescale that have been reported for end-to-end contact formation rates of unstructured, highly flexible poly-(Gly-Ser) peptides [81, 82] these subpopulations therefore cannot be due to homogenously chain bending as such distortion would fluctuate on timescales ranging from nanoseconds to a microsecond, detectable by fluorescent correlation spectroscopy.

Polyprolines were also investigated by means of computational structure prediction. Due to the exponentially rising number of possible *cis*-*trans* isomers with growing chain size, a reduced model of fluorescently labeled polyprolyl-tryptophans was developed. This model treats polyproline moieties using molecular mechanics, whereas terminal fluorescent label and tryptophan were examined in a spatial sampling approach. End-to-end distances for different isomers of polyprolines, incorporating interspersed *cis*-bonds were derived from simulations and show that *trans*-*cis* isomerization leads to significantly shorter end-to-end distances which fits the experimental results of subpopulations with increased photoinduced electron transfer quenching observed in fluorescent correlation spectroscopy experiments for MR121-(Pro)$_N$-Trp with $N = 2 - 10$.

The results of the performed studies on poly-L-prolines strongly indicate that these polymers deviate from the regular polyproline type II helix structure in aqueous solution by the formation of *cis* isomers. This effects a decrease of the proposed polyproline end-to-end distance and has to be accounted for, especially if the polymer is used for length calibrations on a molecular scale.

The second part of this work deals with the polymer properties of single-stranded polythymine that have been investigated through FCS in combination with molecular mechanics simulations. The spectroscopic measurements resulted in length-dependent diffusion constants of fluorescently labeled polythymine probes at low concentrations preventing influence by intermolecular interactions. The diffusion constants show relationships to contour lengths which can be approximated as the number of monomers, N, multiplied with the average contour length of a monomer. Monomer dimensions were calculated as the distance between two adjacent phosphorus atoms. The inverse of the diffusion constants being proportional to the hydrodynamic radii scale with the contour lengths of the polymers as N^{ν}. with ν between 0.5 and 0.7.

Molecular dynamics were performed in vacuo with a force field restricted to short-range repulsive VAN DER WAALS interactions. Any electrostatic or attractive VAN DER WAALS forces were neglected. From the temporal trajectories of these simulations radii of gyration were calculated, which scale with the polymers' contour length in perfect agreement with the semiflexible polymer model taking excluded volume into account where $\nu = 0.588$.

A comparison between simulated radii of gyration and experimental hydrodynamic radii revealed that electrostatic interaction within polythymine is completely shielded at a monovalent salt concentration of 100 mM. At this salt concentration solvent cations compensate for the negative charges aligned at the polymer backbone and the polymer behaves equivalent to a non-charged chain molecule.

Furthermore, the comparison of fluorescent dye labeled single-stranded polythymine with non-labeled showed that the fluorophore moiety does not influence the intrinsic polymer structure. Polymer properties like the end-to-end distance and persistence length are independent of the fluorescent label. The radius of gyration is not significantly manipulated by dye labeling for polythymine polymers of ~ 30 monomers and longer. This justifies the use of fluorescent dye labeled probes in spectroscopic measurements to analyze single stranded polythymines of adequate length with respect to diffusion properties and conformational investigations and transfer results to unlabeled polythymines.

Polythymine polymers with lengths of 30 nucleotides or more can also be described by the wormlike chain model with respect to end-to-end distance distributions. The crude mathematical model reflects the polymer behavior of polythymine adequately. The structure factor of 6 - 7, the characteristic ratio according to FLORY of 5 - 6, and the persistence length of 1.7 nm determined from simulation data reveal limited static flexibility imposed by geometrical constraints and steric interactions.

The comparison of polythymine and polymers consisting of stable abasic sites (pure desoxyribose phosphate backbone lacking any nucleobases) revealed that base stacking or steric hindrance contributes to static flexibility of the polymer only to a small extend. Polymer stiffness arises from conformational restrictions within the backbone of the chain molecule.

These results yield an experimental confirmation that polythymine of a few tens of nucleotides in length resembles an unstructured polyelectrolyte that can be described as a semiflexible polymer. The evaluated persistence length and its scaling with the number of residues provide important experimental results for the verification of polyelectrolyte theories and can thus help to develop a complete and accurate description of single-stranded DNA in general.

The combination of spectroscopic measurement and computational simulation has proven to be a fruitful method in terms of biopolymer analysis. Spectroscopic results can better be interpreted by comparison to simulations on the atomic level. As biopolymers are large-scale molecules with numberless degrees of freedom, classical computational simulations like molecular dynamics of the solvated system are a extraordinarily time-consuming task. The simplified modeling methods shown in this thesis offer a trade-off between computational feasibility and simulation detail. Yet, they provide novel descriptions of biopolymer properties.

With ever increasing computational power the model detail feasible will drastically rise in future. The ultimate task in computational terms would be nano- to microsecond simulations of the probes used in this work as a solvated system, treating the interaction between the fluorescent dye and the quenching tryptophan quantum mechanically. Taking MOORE's law [83] of doubling computational power every two years into account, this challenge seems not to be possible during the next decade.

6. References

1 Smith, S. B., Cui, Y. and Bustamante, C. (1996),"Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules", *Science* **271,** 795-799.

2 Hagerman, P. J. (1988),"Flexibility of DNA", *Annual Review of Biophysics and Biophysical Chemistry* **17,** 265-286.

3 Hagerman, P. J. (1997),"Flexibility of RNA", *Annual Review of Biophysics and Biomolecular Structure* **26,** 139-156.

4 Yanagida, T., Nakase, M., Nishiyama, K. and Oosawa, F. (1984),"Direct observation of motion of single F-actin filaments in the presence of myosin", *Nature* **307,** 58-60.

5 Le Goff, L., Hallatschek, O., Frey, E. and Amblard, F. (2002),"Tracer Studies on F-Actin Fluctuations", *Physical Review Letters* **89,** 258101.

6 Pampaloni, F., Lattanzi, G., Jonas, A., Surrey, T., Frey, E. and Florin, E.-L. (2006),"Thermal fluctuations of grafted microtubules provide evidence of a length-dependent persistence length", *Proceedings of the National Academy of Science* **103,** 10248-10253.

7 Gittes, F., Mickey, B., Nettleton, J. and Howard, J. (1993),"Flexural rigidity of microtubules and actin filaments measured from thermal fluctuations in shape", *Journal of Cell Biology* **120,** 923-934.

8 Olson, W. K. (1975),"Configurational Statistics of Polynucleotide Chains. A Single Virtual Bond Treatment", *Macromolecules* **8,** 272-275.

9 Olson, W. K. and Flory, P. J. (1972),"Spatial configurations of polynucleotide chains. III. Polydeoxyribonucleotides", *Biopolymers* **11,** 57-66.

10 Mills, J. B., Vacano, E. and Hagerman, P. J. (1999),"Flexibility of single-stranded DNA: use of gapped duplex helices to determine the persistence lengths of Poly(dT) and Poly(dA)", *Journal of Molecular Biology* **285,** 245-257.

11 Murphy, M. C., Rasnik, I., Cheng, W., Lohman, T. M. and Ha, T. (2004),"Probing Single-Stranded DNA Conformational Flexibility Using Fluorescence Spectroscopy", *Biophysical Journal* **86,** 2530-2537.

12 Klimov, D. K. and Thirumalai, D. (2002),"Stiffness of the distal loop restricts the structural heterogeneity of the transition state ensemble in SH3 domains", *Journal of Molecular Biology* **317,** 721-737.

13 Lapidus, L. J., Steinbach, P. J., Eaton, W. A., Szabo, A. and Hofrichter, J. (2002),"Effects of Chain Stiffness on the Dynamics of Loop Formation in Polypeptides. Appendix: Testing a 1-Dimensional Diffusion Model for Peptide Dynamics", *Journal of Physical Chemistry B* **106,** 11628-11640.

14 Zhou, H. X. (2001),"Loops in Proteins Can Be Modeled as Worm-Like Chains", *Journal of Physical Chemistry B* **105,** 6763-6766.

15 Odijk, T. (1977),"Polyelectrolytes near the rod limit", *Journal of Polymer Science: Polymer Physics* **15,** 477-483.

16 Skolnick, J. and Fixman, M. (1977),"Electrostatic Persistence Length of a Wormlike Polyelectrolyte", *Macromolecules* **10,** 944-948.

17 Everaers, R., Milchev, A. and Yamakov, V. (2002),"The electrostatic persistence length of polymers beyond the OSF limit", *The European Physical Journal E* **8,** 3-14.

18 Barrat, J. L. and Joanny, J. F. (1993),"Persistence length of polyelectrolyte chains", *Europhysics Letters* **24,** 333-338.

19 Ha, B. Y. and Thirumalai, D. (1995),"Electrostatic Persistence Length of a Polyelectrolyte Chain", *Macromolecules* **28,** 577-581.

20 Zandi, R., Rudnick, J. and Golestanian, R. (2002),"Radial distribution function of rod-like polyelectrolytes", *The European Physical Journal E - Soft Matter* **9,** 41-46.

21 Laurence, T. A., Kong, X., Jäger, M. and Weiss, S. (2005),"Probing structural heterogeneities and fluctuations of nucleic acids and denatured proteins", *Proceedings of the National Academy of Science* **102,** 17348-17353.

22 Schimmel, P. R. and Flory, P. J. (1967),"Conformational Energy and Configurational Statistics of Poly-L-Proline", *Proceedings of the National Academy of Science* **58,** 52-59.

23 Cowan, P. M. and McGavin, S. (1955),"Structure of Poly-L-Proline", *Nature* **176,** 501-503.

24 Steinberg, I. Z., Harrington, W. F., Berger, A., Sela, M. and Katchalski, E. (1960),"The Configurational Changes of Poly-L-proline in Solution", *Journal of the American Chemical Society* **82,** 5263-5279.

25 Traub, W. and Shmueli, U. (1963),"Structure of Poly-L-Proline I", *Nature* **198,** 1165 - 1166.

26 Stryer, L. and Haugland, R. P. (1967),"Energy Transfer: A Spectroscopic Ruler", *Proceedings of the National Academy of Sciences* **58,** 719-726.

27 Sahoo, H., Roccatano, D., Hennig, A. and Nau, W. M. (2007),"A 10-Å Spectroscopic Ruler Applied to Short Polyprolines", *Journal of the American Chemical Society* **129,** 9762-9772.

28 Watkins, L. P., Chang, H. and Yang, H. (2006),"Quantitative Single-Molecule Conformational Distributions: A Case Study with Poly-(L-proline)", *Journal of Physical Chemistry A* **110,** 5191-5203.

29 Schuler, B., Lipman, E. A., Steinbach, P. J., Kumke, M. and Eaton, W. A. (2005),"Polyproline and the "spectroscopic ruler" revisited with single-molecule fluorescence", *Proceedings of the National Academy of Science* **102,** 2754-2759.

30 Schuler, B., Lipman, E. A. and Eaton, W. A. (2002),"Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy", *Nature* **419,** 743-747.

31 Wu, C. C., Komoroski, R. A. and Mandelkern, L. (1975),"The Observation of Cis Residues in Poly(L-proline) in Aqueous Solution", *Macromolecules* **8,** 635-637.

32 Jacob, J., Baker, B., Bryant, R. G. and Cafiso, D. S. (1999),"Distance estimates from paramagnetic enhancements of nuclear relaxation in linear and flexible model peptides", *Biophysical Journal* **77,** 1086-1092.

33 Einstein, A. (1905),"Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt", *Annalen der Physik* **322,** 132-148.

34 Planck, M. (1901),"Ueber das Gesetz der Energieverteilung im Normalspectrum", *Annalen der Physik* **309,** 553-563.

35 de Broglie, L. (1925),"Sur la fréquence propre de l'électron", *Comptes Rendus* **180,** 498-500.

36 Davisson, C. and Germer, L. H. (1927),"Diffraction of Electrons by a Crystal of Nickel", *Physical Review* **30,** 705-740.

37 Förster, T. (1946),"Energieumwanderung und Fluoreszenz", *Zeitschrift für Naturwissenschaften* **33,** 166-175.

38 Selvin, P. R. (2000),"The renaissance of fluorescence resonance energy transfer", *Nat Structural Molecular Biology* **7,** 730-734.

39 Alder, B. J. and Wainwright, T. E. (1959),"Studies in Molecular Dynamics. I. General Method", *The Journal of Chemical Physics* **31,** 459-466.

40 Rahman, A. (1964),"Correlations in the Motion of Atoms in Liquid Argon", *Physical Review* **136,** 405-411.

41 Rahman, A. and Stillinger, F. H. (1974),"Propagation of sound in water. A molecular-dynamics study", *Physical Review A* **10,** 368-378.

42 Stillinger, F. H. and Weber, T. A. (1984),"Inherent pair correlation in simple liquids", *The Journal of Chemical Physics* **80,** 4434-4437.

43 McCammon, J. A., Gelin, B. R. and Karplus, M. (1977),"Dynamics of folded proteins", *Nature* **267,** 585-590.

44 Verlet, L. (1967),"Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules", *Physical Review* **159,** 98-103.

45 Editor: Dugave, C. (2006) "*cis-trans Isomerization in Biochemistry*" (Wiley-VCH Verlag GmbH & Co. KGaA,, Wiesbaden, Germany).

46 Ullner, M. and Woodward, C. E. (2002),"Orientational Correlation Function and Persistence Lengths of Flexible Polyelectrolytes", *Macromolecules* **35,** 1437-1445.

47 Yamakawa, H. (1971), "*Modern theory of polymer solutions",* (Harper & Row, New York, NY, USA).

48 Dertinger, T., Pacheco, V., von der Hocht, I., Hartmann, R., Gregor, I. and Enderlein, J. (2007),"Two-Focus Fluorescence Correlation Spectroscopy: A New Tool for Accurate and Absolute Diffusion Measurements", *ChemPhysChem* **8,** 433-443.

49 Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. S., Cheatham III, T. E., DeBolt, S., Ferguson, D., Seibel, G. and Kollman, P. (1995),"AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules", *Computer Physics Communications* **91,** 1-41.

50 Pigache, A., Cieplak, P. and Dupradeau, F. Y. (2004), *"Automatic and highly reproducible RESP and ESP charge derivation: application to the development of programs RED and X RED"*, 227th ACS National Meeting, Anaheim, CA, USA.

51 Schmidt, M. W., Baldridge, K. K., Boatz, J. A., Elbert, S. T., Gordon, M. S., Jensen, J. H., Koseki, S., Matsunaga, N., Nguyen, K. A., Su, S., Windus, T. L., Dupuis, M. and Montgomery Jr, J. A. (1993),"General atomic and molecular electronic structure system", *Journal of Computational Chemistry* **14,** 1347-1363.

52 Bayly, C. I., Cieplak, P., Cornell, W. and Kollman, P. A. (1993),"A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model", *Journal of Physical Chemistry* **97,** 10269-10280.

53 Advanced Chemistry Development, Inc. (2006), *"ACD/ChemSketch* 10.02", http://www.acdlabs.com.

54 GNU General Public License (2007), *"Open Babel* 2.0.2", http://openbabel.sourceforge.net.

55 Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W. and Kollman, P. A. (1995),"A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules", *Journal of the American Chemical Society* **117,** 5179-5197.

56 Singh, U. C. and Kollman, P. A. (1984),"An approach to computing electrostatic charges for molecules", *Journal of Computational Chemistry* **5,** 129-145.

57 Doose, S., Neuweiler, H. and Sauer, M. (2005),"A close look at fluorescence quenching of organic dyes by tryptophan", *ChemPhysChem* **6,** 2277-2285.

58 Magde, D., Elson, E. L. and Webb, W. W. (1974),"Fluorescence correlation spectroscopy. II. An experimental realization", *Biopolymers* **13,** 29-61.

59 Hess, S. T., Huang, S., Heikal, A. A. and Webb, W. W. (2002),"Biological and chemical applications of fluorescence correlation spectroscopy: a review", *Biochemistry* **41,** 697-705.

60 Latscha, H. P., Schilling, G. and Klein, H. A. (1990), "*Chemie-Datensammlung",* (Springer-Verlag, Berlin, Germany).

61 Neuweiler, H. (2002),*"Synthesis, Spectroscopic Characterization, and Molecular Dynamics Simulation of Fluorescently Labeled Peptides for Single-Molecule Detection of p53 Antibodies)"*, Ph. D. Thesis, Combined Faculties for Natural Sciences and Mathematics, Ruprecht-Karls-University, Heidelberg, Germany

62 Vaiana, A. C., Neuweiler, H., Schulz, A., Wolfrum, J., Sauer, M. and Smith, J. C. (2003),"Fluorescence quenching of dyes by tryptophan: interactions at atomic detail from combination of experiment and computer simulation", *J Am Chem Soc* **125,** 14564-14572.

63 Neuweiler, H., Lollmann, M., Doose, S. and Sauer, M. (2006),"Dynamics of Unfolded Polypeptide Chains in Crowded Environment Studied by Fluorescence Correlation Spectroscopy", *J Mol Biol* **365,** 856-869.

64 Mishra, A. K., Chandrasekar, R., Faraggi, M. and Klapper, M. H. (1994),"Long-Range Electron-Transfer in Peptides - Tyrosine Reduction of the Indolyl Radical - Reaction-Mechanism, Modulation of Reaction-Rate, and Physiological Considerations", *Journal of the American Chemical Society* **116,** 1414-1422.

65 Defelippis, M. R., Faraggi, M. and Klapper, M. H. (1990),"Evidence for through-Bond Long-Range Electron-Transfer in Peptides", *Journal of the American Chemical Society* **112,** 5640-5642.

66 Isied, S. S. and Vassilian, A. (1984),"Electron-Transfer across Polypeptides .3. Oligoproline Bridging Ligands", *Journal of the American Chemical Society* **106,** 1732-1736.

67 Gray, H. B. and Winkler, J. R. (2005),"Long-range electron transfer", *Proc Natl Acad Sci U S A* **102,** 3534-3539.

68 Kim, J., Doose, S., Neuweiler, H. and Sauer, M. (2006),"The initial step of DNA hairpin folding: a kinetic analysis using fluorescence correlation spectroscopy", *Nucl. Acids Res.* **34,** 2516-2527.

69 Neuweiler, H., Doose, S. and Sauer, M. (2005),"A microscopic view of miniprotein folding: Enhanced folding efficiency through formation of an intermediate", *Proc. Natl. Acad. Sci. USA* **102,** 16650-16655.

70 Chattopadhyay, K., Elson, E. L. and Frieden, C. (2005),"The kinetics of conformational fluctuations in an unfolded protein measured by fluorescence methods", *Proc Natl Acad Sci U S A* **102,** 2385-2389.

71 Huang, F. and Nau, W. M. (2003),"A conformational flexibility scale for amino acids in peptides", *Angew Chem Int Ed Engl* **42,** 2269-2272.

72 Krieger, F., Moglich, A. and Kiefhaber, T. (2005),"Effect of proline and glycine residues on dynamics and barriers of loop formation in polypeptide chains", *J Am Chem Soc* **127,** 3346-3352.

73 Brant, D. A. and Flory, P. J. (1965),"The Configuration of Random Polypeptide Chains. II. Theory", *Journal of the American Chemical Society* **87,** 2791-2800.

74 Cowan, P. M. and McGavin, S. (1955),"Structure of poly-L-proline", *Nature* **176,** 501-503.

75 Kim, J., Doose, S., Neuweiler, H. and Sauer, M. (2006),"The initial step of DNA hairpin folding: a kinetic analysis using fluorescence correlation spectroscopy", *Nucleic Acids Research* **34,** 2516-2527.

76 Flory, P. J. (1988), "*Statistical Mechanics of Chain Molecules",* (Carl Hanser Verlag, Munich, Germany).

77 Benoit, H. and Doty, P. (1953),"Light Scattering from Non-Gaussian Chains", *Journal of Physical Chemistry* **57,** 958-963.

78 Olson, W. K. and Flory, P. J. (1972),"Spatial configurations of polynucleotide chains. I. Steric interactions in polyribonucleotides: A virtual bond model", *Biopolymers* **11,** 1-23.

79 Mandelkern, L. and Mattice, W. L. (1971),"Conformational properties of poly-L-proline from II in dilute solution", *Journal of the American Chemical Society* **93,** 1769-1777.

80 Tanaka, S. and Scheraga, H. A. (1975),"Calculation of the Characteristic Ratio of Randomly Coiled Poly(L-proline)", *Macromolecules* **8,** 623-631.

81 Neuweiler, H., Lollmann, M., Doose, S. and Sauer, M. (2007),"Dynamics of Unfolded Polypeptide Chains in Crowded Environment Studied by Fluorescence Correlation Spectroscopy", *Journal of Molecular Biology* **365,** 856-869.

82 Krieger, F., Moglich, A. and Kiefhaber, T. (2005),"Effect of Proline and Glycine Residues on Dynamics and Barriers of Loop Formation in Polypeptide Chains", *Journal of the American Chemical Society* **127,** 3346-3352.

83 Schaller, R. R. (1997),"Moore's law: past, present and future", *Spectrum, IEEE* **34,** 52-59.

7. Publication List

1 Piestert, O., Barsch, H., Buschmann, V., Heinlein, T., Knemeyer, J. P., Weston, K. D. and Sauer, M. (2003),"A Single-Molecule Sensitive DNA Hairpin System Based on Intramolecular Electron Transfer", *Nano Letters* **3**, 979-982.

2 Doose, S., Barsch, H. and Sauer, M. (2007),"Polymer Properties of Polythymine as Revealed by Translational Diffusion", *Biophysical Journal* **93**, 1224-1234.

3 Doose, S., Neuweiler, H., Barsch, H. and Sauer, M. (2007),"Probing polyproline structure and dynamics by photoinduced electron transfer provides evidence for deviations from a regular polyproline type II helix", *Procedings of the National Academy of Science* **104**, 17400-17405.

8. Acknowledgements

This thesis would not have been possible without several people who constantly encouraged me to keep proceeding even when things did not work out as should have.

At first place I would like to thank Prof. Dr. Markus Sauer for giving me the possibility to work in his group. Although my wish to do computational chemistry did not fit the experimentally orientated work of the other group members, you believed in my skills to put up a linux cluster and achieve results from simulations worth publishing. Thanks for all the freedom you gave me throughout my work and thank you for helping me to keep my roots in Heidelberg – I guess I will miss cruising along Lake Eder with you.

All through the work on polyprolines and polythymines Dr. Sören Doose was the one who supervised my work, who constantly shared thoughts and ideas and who had always read that one more paper necessary for success. With your calm and patient mind you were not only the perfect colleague to share the office with for a long time but also a great help in terms of research!

Working in the Applied Laser Physics and Laser Spectroscopy group has always been a pleasure. Any question or need for discussions has fallen on fertile ground at all times. Thanks, D3alle!

Whereas computers were my passion, I felt like an alien in the lab. Thank you, Dr. Sören Doose and Dr. Hannes Neuweiler, for your support with FCS measurements!

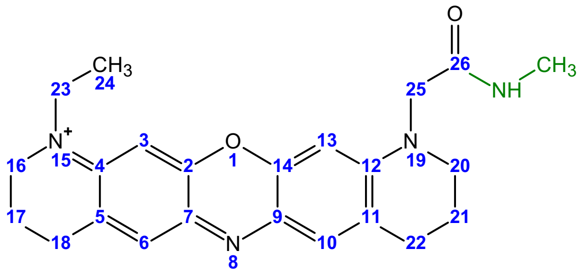
Apart from research my life in Bielefeld would have been boring and lonesome without friends. It was a gift to meet Robert Kasper and Sabine Esch who always lent an ear when work or private life put a damper on me. Thanks for your time, for your advice and for your sympathy!

Finally, I want to thank Heike A. Müller. You had to endure our time being separated much longer than supposed but you kept your faith in me. You could not help me with research but during the past years you patiently encouraged me to move forward personally which is worth more than any title.

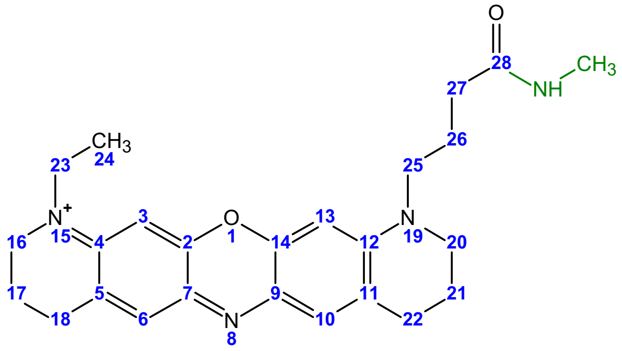
9. Appendices

A Molecular Structures of Parametrized Residues

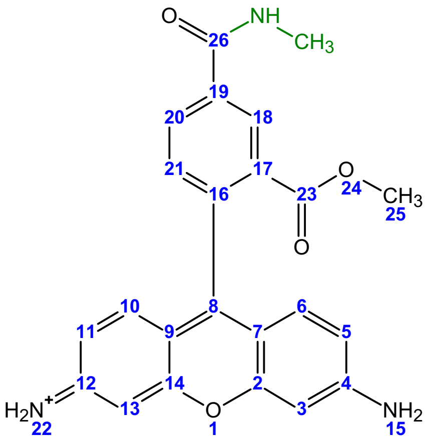
A.1. MR113



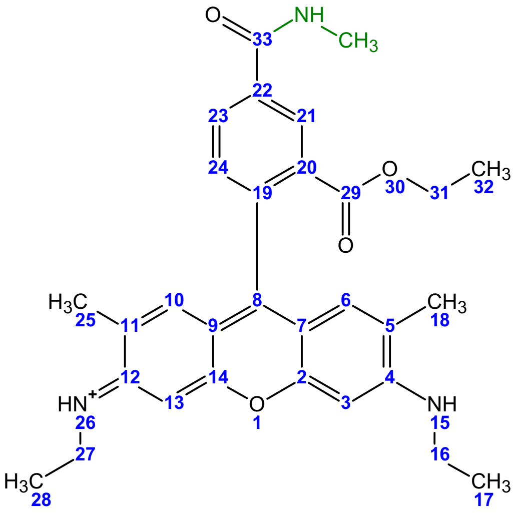
A.2. MR121 (M21)



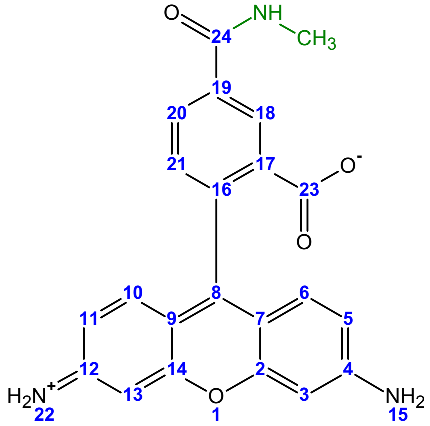
A.3. Rhodamine 123 (R23)



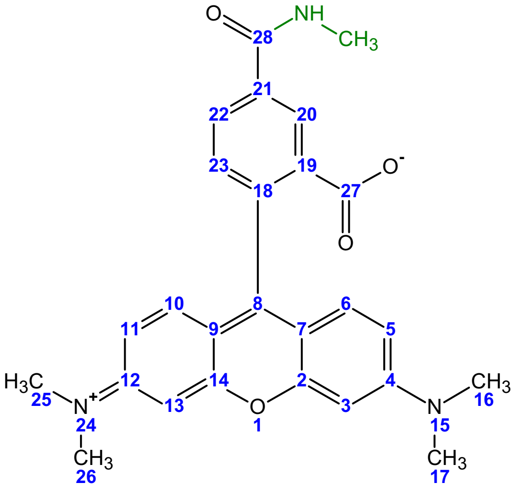
A.4. Rhodamine 6G (R6G)



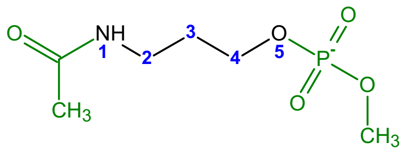
A.5. Rhodamine Green (RGR)



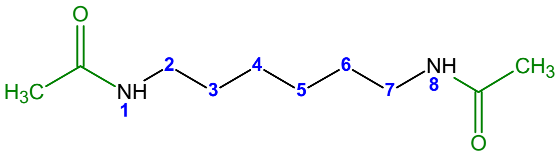
A.6. Tetramethyl-Rhodamine (TMR)



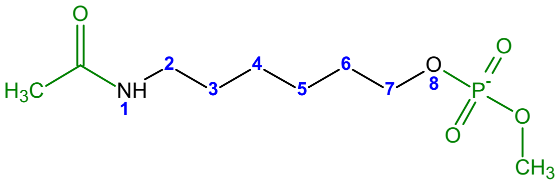
A.7. C$_3$-Linker-NO (3NO)



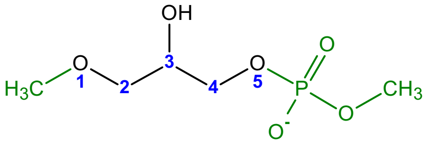
A.8. C$_6$-Linker-NN (6NN)



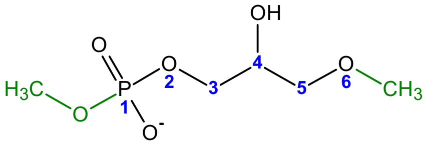
A.9. C$_6$-Linker-NO (6NO)



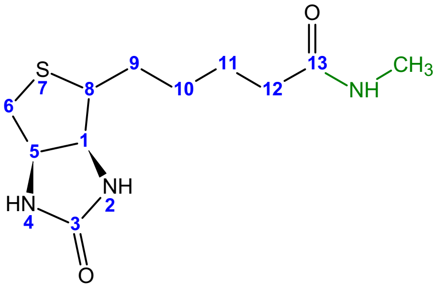
A.10. Glycerol (GL5)



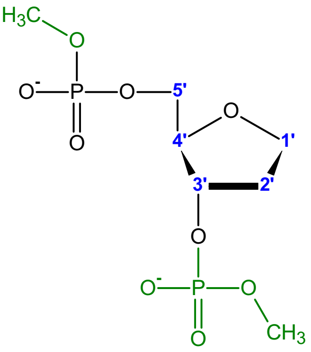
A.11. Glycerol Phosphate (GL3)



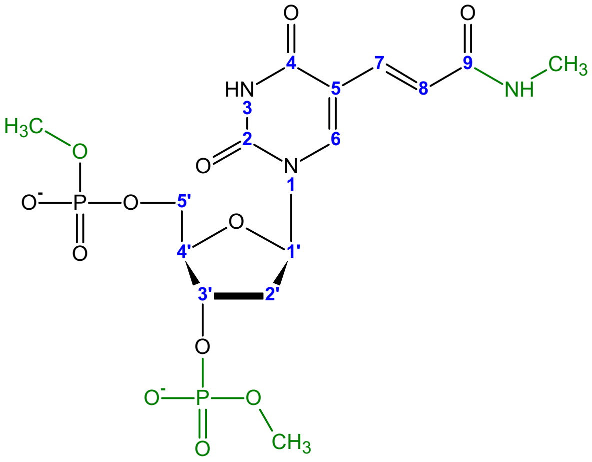
A.12. Biotin (BTN)



A.13. Stable Abasic Site (dS)



A.14. 5-modified Thymine (dM)



B Atoms Used for Reorientation

Name	A	B	C	D	E
MR113	O^1	C^{20}	C^{10}	C^6	C^{16}
MR121	O^1	C^{16}	C^6	C^{10}	C^{20}
Rhodamine 123	O^1	C^{10}	C^{20}	C^{18}	C^6
Rhodamine 6G	O^1	C^{11}	C^{23}	C^{21}	C^6
Rhodamine Green	O^1	C^{11}	C^{20}	C^{18}	C^5
Tetramethyl Rhodamine	O^1	C^{11}	C^{22}	C^{20}	C^5
dA, dC, dG, dT, dS, dTm	$C^{5'}$	$C^{4'}$	$C^{1'}$	$C^{2'}$	$C^{3'}$
C_3-Linker-NO	N^1	C^2	C^3	C^4	O^5
C_6-Linker-NN	N^1	C^3	C^5	C^7	N^8
C_6-Linker-NO	N^1	C^3	C^5	C^7	O^8
C_7-Linker-ON	P^1	C^3	C^6	C^8	N^{10}
Glycerol	O^1	C^2	C^3	C^4	O^5
Glycerol Phosphate	O^1	C^2	C^3	C^4	O^5
Biotin	N^2	N^1	C^6	C^2	C^{11}

C Atom Centered Charges

C.1. MR113 (M13)

O^{1}	OS	-0.1622	N^{15}	N2	-0.0638	C^{22}	CT	-0.0288
C^{2}	CA	0.2644	C^{16}	CT	-0.0747	H^{221}	HC	0.0501
C^{3}	CA	-0.4102	H^{161}	HC	0.0940	H^{222}	HC	0.0501
H^{3}	HA	0.1932	H^{162}	HC	0.0940	C^{23}	CT	-0.0293
C^{4}	CA	0.1305	C^{17}	CT	-0.0500	H^{231}	H1	0.0719
C^{5}	CA	0.0467	H^{171}	HC	0.0516	H^{232}	H1	0.0719
C^{6}	CA	-0.2992	H^{172}	HC	0.0516	C^{24}	CT	-0.0340
H^{6}	HA	0.1913	C^{18}	CT	-0.0427	H^{241}	HC	0.0305
C^{7}	CA	0.3464	H^{181}	HC	0.0615	H^{242}	HC	0.0305
N^{8}	NC	-0.4373	H^{182}	HC	0.0615	H^{243}	HC	0.0305
C^{9}	CA	0.3171	N^{19}	N2	-0.0289	C^{25}	CT	-0.1482
C^{10}	CA	-0.2297	C^{20}	CT	-0.0135	H^{251}	H1	0.0791
H^{10}	HA	0.1786	H^{201}	H1	0.0619	H^{252}	H1	0.0791
C^{11}	CA	0.0242	H^{202}	H1	0.0619	C^{26}	C	0.8143
C^{12}	CA	0.0141	C^{21}	CT	-0.0174	O^{26}	O	-0.5804
C^{13}	CA	-0.1600	H^{211}	HC	0.0518			
C^{14}	CA	0.1542	H^{212}	HC	0.0518			

C.2. MR121 (M21)

O^{1}	OS	-0.2005	C^{16}	CT	-0.0754	C^{23}	CT	-0.0326
C^{2}	CA	0.2744	H^{161}	HC	0.0918	H^{231}	H1	0.0716
C^{3}	CA	-0.3739	H^{162}	HC	0.0918	H^{232}	H1	0.0716
H^{3}	HA	0.1882	C^{17}	CT	-0.0462	C^{24}	CT	-0.0340
C^{4}	CA	0.1056	H^{171}	HC	0.0500	H^{241}	HC	0.0290
C^{5}	CA	0.0323	H^{172}	HC	0.0500	H^{242}	HC	0.0290
C^{6}	CA	-0.2565	C^{18}	CT	-0.0416	H^{243}	HC	0.0290
H^{6}	HA	0.1806	H^{181}	HC	0.0583	C^{25}	CT	-0.0412
C^{7}	CA	0.2810	H^{182}	HC	0.0583	H^{251}	H1	0.0699
N^{8}	NC	-0.4125	N^{19}	N2	-0.0472	H^{252}	H1	0.0699
C^{9}	CA	0.2820	C^{20}	CT	-0.0264	C^{26}	CT	-0.0038
C^{10}	CA	-0.2497	H^{201}	H1	0.0684	H^{261}	HC	0.0301
H^{10}	HA	0.1810	H^{202}	H1	0.0684	H^{262}	HC	0.0301
C^{11}	CA	0.0259	C^{21}	CT	-0.0225	C^{27}	CT	-0.0031
C^{12}	CA	0.0588	H^{211}	HC	0.0442	H^{271}	HC	0.0141
C^{13}	CA	-0.2190	H^{212}	HC	0.0442	H^{272}	HC	0.0141

H^{13}	H	0.0975	C^{22}	CT	-0.0433	C^{28}	C	0.5489
C^{14}	CA	0.2501	H^{221}	HC	0.0582	O^{28}	O	-0.5128
N^{15}	N2	-0.0640	H^{222}	HC	0.0582			

C.3. Rhodamine 123 (R23)

O^{1}	OS	-0.2160	C^{13}	CA	-0.3626	H^{222}	HN	0.4277
C^{2}	CA	0.2653	H^{13}	HA	0.1927	C^{23}	C	0.6956
C^{3}	CA	-0.3673	C^{14}	CA	0.2715	O^{23}	OS	-0.5140
H^{3}	HA	0.1942	N^{15}	*N**	-0.8695	O^{24}	CT	-0.3113
C^{4}	CA	0.4522	H^{151}	HN	0.4251	C^{25}	HC	-0.1416
C^{5}	CA	-0.2378	H^{152}	HN	0.4251	H^{251}	HC	0.1219
H^{5}	HA	0.1782	C^{16}	CD	0.0058	H^{252}	HC	0.1219
C^{6}	CA	-0.2043	C^{17}	CA	-0.0525	H^{253}	C	0.1219
H^{6}	HA	0.1979	C^{18}	CA	-0.2034	C^{26}	O	0.5457
C^{7}	CA	0.0482	H^{18}	HA	0.1297	O^{26}	O	-0.5771
C^{8}	CD	0.0644	C^{19}	CA	0.1164	$N^{2}A$	*N*	-0.4530
C^{9}	CA	0.0438	C^{20}	CA	-0.1385	$H^{2}A$	H	0.3263
C^{10}	CA	-0.2046	H^{20}	HA	0.1881	$C^{1}A$	C	-0.0540
H^{10}	HA	0.2059	C^{21}	CA	-0.1785	$H^{1}A^{1}$	H	0.0714
C^{11}	CA	-0.2306	H^{21}	HA	0.1590	$H^{1}A^{2}$	H	0.0714
H^{11}	HA	0.1774	N^{22}	*N**	-0.8713	$H^{1}A^{3}$	H	0.0714
C^{12}	CA	0.4443	H^{221}	HN	0.4277			

C.4. Rhodamine 6G (R6G)

O^{1}	OS	-0.1890	C^{17}	CT	-0.1454	H^{26}	H	0.3317
C^{2}	CA	0.2546	H^{171}	HC	0.0645	C^{27}	CT	-0.0215
C^{3}	CA	-0.2509	H^{172}	HC	0.0645	H^{271}	H1	0.0801
H^{3}	HA	0.1292	H^{173}	HC	0.0645	H^{272}	H1	0.0801
C^{4}	CA	0.0789	C^{18}	CT	-0.1635	C^{28}	CT	-0.0126
C^{5}	CA	0.0558	H^{181}	HC	0.0681	H^{281}	HC	0.0252
C^{6}	CA	-0.2345	H^{182}	HC	0.0681	H^{282}	HC	0.0252
C^{7}	CA	0.0161	H^{183}	HC	0.0681	H^{283}	HC	0.0252
C^{8}	CD	0.0625	C^{19}	CD	0.0135	C^{29}	C	0.6306
C^{9}	CA	0.0323	C^{20}	CA	-0.0187	O^{29}	O	-0.4803
C^{10}	CA	-0.2538	C^{21}	CA	-0.1959	O^{30}	OS	-0.3458
H^{10}	HA	0.1989	H^{21}	HA	0.1335	C^{31}	CT	0.0819
C^{11}	CA	0.0492	C^{22}	CA	0.0795	H^{311}	H1	0.0730
C^{12}	CA	0.1300	C^{23}	CA	-0.1217	H^{312}	H1	0.0730

C^{13}	CA	-0.2360	H^{23}	HA	0.1818	C^{32}	CT	-0.0639
H^{13}	HA	0.1614	C^{24}	CA	-0.1993	H^{321}	HC	0.0361
C^{14}	CA	0.2098	H^{24}	HA	0.1667	H^{322}	HC	0.0361
N^{15}	N2	-0.2616	C^{25}	CT	-0.1488	H^{323}	HC	0.0361
H^{15}	H	0.2956	H^{251}	HC	0.0656	C^{33}	C	0.5966
C^{16}	CT	-0.0952	H^{252}	HC	0.0656	O^{33}	O	-0.5886
H^{161}	H1	0.1169	H^{253}	HC	0.0656	H^{6}	HA	0.1875
H^{162}	H1	0.1169	N^{26}	N2	-0.3687			

C.5. Rhodamine Green (RGR)

O^{1}	OS	-0.2639	C^{11}	CA	-0.2179	C^{20}	CA	-0.1044
C^{2}	CA	0.1794	H^{11}	HA	0.1627	H^{20}	HA	0.1666
C^{3}	CA	-0.2883	C^{12}	CA	0.3339	C^{21}	CA	-0.2373
H^{3}	HA	0.1673	C^{13}	CA	-0.2936	H^{21}	HA	0.1721
C^{4}	CA	0.3315	H^{13}	HA	0.1695	N^{22}	*N**	-0.8982
C^{5}	CA	-0.2199	C^{14}	CA	0.1729	H^{221}	HN	0.3883
H^{5}	HA	0.1629	N^{15}	*N**	-0.9004	H^{222}	HN	0.3883
C^{6}	CA	-0.2908	H^{151}	HN	0.3890	C^{23}	C	0.6956
H^{6}	HA	0.2054	H^{152}	HN	0.3890	O^{231}	O2	-0.4443
C^{7}	CA	0.0546	C^{16}	CD	0.0140	O^{232}	O2	-0.5443
C^{8}	CD	0.2527	C^{17}	CA	-0.0545	C^{24}	C	0.5433
C^{9}	CA	0.0616	C^{18}	CA	-0.2056	O^{24}	O	-0.5880
C^{10}	CA	-0.2911	H^{18}	HA	0.1625			
H^{10}	HA	0.2111	C^{19}	CA	0.0685			

C.6. Tetramethyl Rhodamine (TMR)

O^{1}	OS	-0.2952	H^{13}	HA	0.1452	C^{23}	CA	-0.2055
C^{2}	CA	0.2864	C^{14}	CA	0.1930	H^{23}	HA	0.1509
C^{3}	CA	-0.3346	N^{15}	*N**	-0.2926	N^{24}	*N**	-0.1194
H^{3}	HA	0.1563	C^{16}	CT	-0.1939	C^{25}	CT	-0.1704
C^{4}	CA	0.1970	H^{161}	HC	0.0917	H^{251}	HC	0.0907
C^{5}	CA	-0.2483	H^{162}	HC	0.0917	H^{252}	HC	0.0907
H^{5}	HA	0.1530	H^{163}	HC	0.0917	H^{253}	HC	0.0907
C^{6}	CA	0.0689	C^{17}	CT	-0.0539	C^{26}	CT	-0.1621
H^{6}	HA	0.0799	H^{171}	HC	0.0703	H^{261}	HC	0.0872
C^{7}	CA	0.0073	H^{172}	HC	0.0703	H^{262}	HC	0.0872
C^{8}	CD	-0.0012	H^{173}	HC	0.0703	H^{263}	HC	0.0872
C^{9}	CA	0.0321	C^{18}	CD	-0.0115	C^{27}	C	0.7416

C^{10}	CA	-0.2122	C^{19}	CA	0.0245	O^{271}	O2	-0.5871
H^{10}	HA	0.1838	C^{20}	CA	-0.2469	O^{272}	O2	-0.3827
C^{11}	CA	-0.2073	H^{20}	HA	0.1826	C^{28}	C	0.5552
H^{11}	HA	0.1489	C^{21}	CA	0.0598	O^{28}	O	-0.5895
C^{12}	CA	0.0921	C^{22}	CA	-0.1130			
C^{13}	CA	-0.2164	H^{22}	HA	0.1657			

C.7. C_3-Linker-NO (3NO)

N^1	*N*	-0.4703	H^{22}	H1	0.0742	C^4	CT	0.1308
H^1	H	0.2678	C^3	CT	0.0025	H^{41}	H1	0.0457
C^2	CT	-0.0568	H^{31}	HC	0.0357	H^{42}	H1	0.0457
H^{21}	H1	0.0742	H^{32}	HC	0.0357	O^5	OS	-0.4930

C.8. C_6-Linker-NN (6NN)

N^1	*N*	-0.5041	C^4	CT	-0.0094	H^{62}	HC	0.0227
H^1	H	0.3207	H^{41}	HC	0.0033	C^7	CT	-0.0165
C^2	CT	-0.0090	H^{42}	HC	0.0033	H^{71}	H1	0.0714
H^{21}	H1	0.0690	C^5	CT	-0.0053	H^{72}	H1	0.0714
H^{22}	H1	0.0690	H^{51}	HC	0.0031	N^8	*N*	-0.5071
C^3	CT	0.0123	H^{52}	HC	0.0031	H^8	H	0.3231
H^{31}	HC	0.0217	C^6	CT	0.0125			
H^{32}	HC	0.0217	H^{61}	HC	0.0227			

C.9. C_6-Linker-NO (6NO)

N^1	*N*	-0.4829	H^{32}	HC	0.0073	C^6	CT	0.0330
H^1	H	0.2755	C^4	CT	0.0082	H^{61}	HC	0.0169
C^2	CT	-0.0010	H^{41}	HC	-0.0025	H^{62}	HC	0.0169
H^{21}	H1	0.0693	H^{42}	HC	-0.0025	C^7	CT	0.1521
H^{22}	H1	0.0693	C^5	CT	0.0180	H^{71}	H1	0.0145
C^3	CT	0.0106	H^{51}	HC	-0.0201	H^{72}	H1	0.0145
H^{31}	HC	0.0073	H^{52}	HC	-0.0201	O^8	OS	-0.4922

C.10. C_7-Linker-ON (7ON)

P^1	P	1.2193	H^{42}	HC	0.0236	H^{72}	HC	-0.0023
O^{11}	O2	-0.7829	C^5	CT	0.0177	C^8	CT	0.0044
O^{12}	O2	-0.8084	H^{51}	HC	-0.0210	H^{81}	HC	0.0144

O^{2}	OS	-0.4919	H^{52}	HC	-0.0210	H^{82}	HC	0.0144
C^{3}	CT	0.0916	C^{6}	CT	0.0237	C^{9}	CT	-0.0413
H^{31}	H1	0.0331	H^{61}	HC	-0.0117	H^{91}	H1	0.0753
H^{32}	H1	0.0331	H^{62}	HC	-0.0117	H^{92}	H1	0.0753
C^{4}	CT	0.0275	C^{7}	CT	0.0122	N^{10}	*N*	-0.4579
H^{41}	HC	0.0236	H^{71}	HC	-0.0023	H^{10}	H	0.2711

C.11. Glycerol (GL5)

C^{2}	OS	-0.0478	C^{3}	CT	0.2386	C^{4}	CT	0.0047
H^{21}	CT	0.0769	H^{31}	H1	0.1252	H^{41}	H1	0.0457
H^{22}	H1	0.0769	O^{3}	OH	-0.7741	H^{42}	H1	0.0457
O^{1}	H1	-0.1913	H^{32}	HO	0.5604	O^{5}	OS	-0.1608

C.12. Glycerol Phosphate (GL3)

C^{2}	OS	-0.0521	H^{31}	H1	0.0874	H^{42}	H1	0.0387
H^{21}	CT	0.0712	O^{3}	OH	-0.6435	O^{5}	OS	-0.5044
H^{22}	H1	0.0712	H^{32}	HO	0.4034	P	P	12.681
O^{1}	H1	-0.2015	C^{4}	CT	0.0934	O^{P1}	O2	-0.7955
C^{3}	CT	0.2412	H^{41}	H1	0.0387	O^{P2}	O2	-0.8083

C.13. Biotin (BTN)

C^{1}	CT	0.0288	C^{6}	CT	0.0294	H^{101}	HC	0.0342
H^{1}	H1	0.1090	H^{61}	H1	0.0758	H^{102}	HC	0.0342
N^{2}	*N*	-0.5787	H^{62}	H1	0.0758	C^{11}	CT	0.0144
H^{2}	H	0.3525	S^{7}	S	-0.2723	H^{111}	HC	0.0146
C^{3}	C	0.7726	C^{8}	CT	-0.2368	H^{112}	HC	0.0146
O^{3}	O	-0.6298	H^{8}	H1	0.2060	C^{12}	CT	-0.1473
N^{1}	*N*	-0.5543	C^{9}	CT	-0.0603	H^{121}	HC	0.0482
H^{1}	H	0.3411	H^{91}	HC	0.0870	H^{122}	HC	0.0482
C^{5}	CT	0.0251	H^{92}	HC	0.0870	C^{13}	C	0.5500
H^{5}	H1	0.1052	C^{10}	CT	-0.0286	O^{13}	O	-0.5456

C.14. Stable Abasic Site (dS)

P	P	1.2168	$C^{4'}$	CT	0.2079	$H^{2'2}$	HC	0.0414
O^{P1}	O2	-0.7868	$H^{4'}$	H1	0.0878	$C^{1'}$	CT	0.0033
O^{P2}	O2	-0.7968	$O^{4'}$	OS	-0.4299	$H^{1'1}$	HC	0.0930

$O^{5'}$	OS	-0.5589	$C^{3'}$	CT	0.1535	$H^{1'2}$	HC	0.0930
$C^{5'}$	CT	0.0065	$H^{3'}$	H1	0.0657	$O^{3'}$	OS	-0.5013
$H^{5'1}$	H1	0.0670	$C^{2'}$	CT	-0.0706			
$H^{5'2}$	H1	0.0670	$H^{2'1}$	HC	0.0414			

C.15. 5-modified Thymine (dT)

P	P	1.2168	$C^{2'}$	CT	-0.0706	H^{8}	HA	0.2066
O^{P1}	O2	-0.7868	$H^{2'1}$	HC	0.0414	C^{9}	C	0.6481
O^{P2}	O2	-0.7968	$H^{2'2}$	HC	0.0414	O^{9}	O	-0.5620
$O^{5'}$	OS	-0.5589	$C^{1'}$	CT	0.2677	C^{4}	C	0.5103
$C^{5'}$	CT	0.0065	$H^{1'}$	H2	0.1163	O^{4}	O	-0.5235
$H^{5'1}$	H1	0.0670	N^{1}	N'	-0.0791	N^{3}	NA	-0.3644
$H^{5'2}$	H1	0.0670	C^{6}	CM	-0.2115	H^{3}	H	0.3146
$C^{4'}$	CT	0.2079	H^{6}	H4	0.2880	C^{2}	C	0.5312
$H^{4'}$	H1	0.0878	C^{5}	CM	-0.0039	O^{2}	O	-0.5669
$O^{4'}$	OS	-0.4299	C^{7}	CE	-0.0809	$O^{3'}$	OS	-0.5013
$C^{3'}$	CT	0.1535	H^{7}	HA	0.0349			
$H^{3'}$	H1	0.0657	C^{8}	CE	-0.3364			