FRET restrained high-precision structural modeling of RNA junctions

Inaugural dissertation

for the attainment of the title of doctor
in the Faculty of Mathematics and Natural Sciences
at the Heinrich Heine University Düsseldorf

presented by

Hayk Vardanyan

from Yerevan, Armenia

Düsseldorf, July 2014

from the institute for Physical Chemistry II
at the Heinrich Heine University Düsseldorf

Published by permission of the
Faculty of Mathematics and Natural Sciences at
Heinrich Heine University Düsseldorf

Supervisor: Prof. Dr. Claus A. M. Seidel
Co-supervisor: Prof. Dr. Gerhard Steger

Date of the oral examination: 15.10.2014

Summary

Hardware photon correlator

For studies of biomolecular dynamics by FCS on nanoseconds to seconds timescale, a low cost hardware correlator, as an alternative to the current commercial devices, was developed. It is implemented in a field-programmable gate array (FPGA) combined with a compact confocal fluorescence setup. Test measurements showed that the sensitivity of the developed setup is identical or even superior to the current commercial devices. It has two independent units with a time resolution of 4 ns while utilizing less than 15 % of a low-end FPGA. Increasing number of FCS application for life and material sciences [1, 2], demonstrates that flexible and low cost instrumentation with high performance will facilitate the use of FCS for even more applications in education, applied sciences and basic research.

RNA studies

It is well known that RNA molecules play an important role in living organisms [3, 4] and the functionality of these molecules is strongly related to their tertiary structures. Hence, the structural determination of these molecules is highly important. Naturally occurring RNA systems mostly exhibit a complex structure, containing several structural elements like junctions, loops and bulges at a time [5, 6]. Hence, investigation of helical junctions and the influence of the separate structural elements in the folding procedure is a key step towards understanding the functionality of the molecules.

In this work structural modelling of large RNA four- and three-way junction molecules (35 – 45 kD) based on FRET experiments, molecular dynamics and coarse-grained simulations are presented. Förster Resonance Energy Transfer (FRET) is one of the techniques used to study structure and dynamics of biomolecules on single-molecule level. FRET is a highly distance dependent mechanism describing energy transfer from one chromophore to another [7]. FRET is used as a “spectroscopic ruler” [8], as it is particularly suitable for the distance measurements between fluorescent dyes typically in the range of 30 – 80 Å. Thus, structural heterogeneities and dynamic transitions in biomolecules are directly measurable under

physiological conditions with high time resolution which is a large advantage over most other techniques in structural biology.

In total 173 independent single-molecule FRET (smFRET) measurements were performed in order to obtain the structures of RNA four-way junction and three RNA three-way junction molecules (1. fully paired RNA three-way junction, 2. RNA three-way junction with additional bulge with two cytosine (C) bases in the junction region, 3. RNA three-way junction with additional bulge with 5 cytosine (C) bases in the junction region). One major and two minor coexisting conformers were resolved for four-way junction molecule. For three-way junctions structural determination of only major populations was performed as the minor populations were shown to be due to incompletely hybridized molecules or acceptor photophysics. For each dataset distances and corresponding uncertainties were extracted. In the case of four-way junction three distances and corresponding uncertainties were extracted and successfully assigned to the three conformers of the RNA four-way junction using their distinct Mg^{2+} – affinities. For three-way junctions distance assignment to the conformers was straightforward.

In current project the FRET positioning and screening toolkit (FPS) was applied, which combines experimental and computational techniques in a hybrid approach, to derive all-atom structural models with high accuracy and precision as well as to assess the confidence levels of the obtained models [9]. Assuming the helices of the RNA molecules to be a rigid double stranded RNAs, rigid body models of the observed conformers were obtained. The models were then refined by MD simulations and coarse grained RNA folding, resulting in meaningful all-atom structural models for the investigated molecules. Confidence levels of the obtained models were estimated from cluster analysis. Quality assessment was done via rigorous error estimation. Resulting precisions are significantly better than the uncertainty of dye position with respect to macromolecule. Furthermore, this is the first time that coexisting transient minor conformers of an RNA four-way junction (J(abcd)) molecule were structurally solved.

Comparison of the 3D structures of the studied four-and three-way junctions shows that after omitting one helix from the four-way junction, structures become almost planar, deviating from complete planarity not more than 10° (planarity is described by the sum of

the mutual angles between each pair of helices. 360° corresponds to complete planarity). The resulting fully paired three-way junction has a Y shape, and only after addition of the bulges coaxial stacking occurs. This can probably be explained with the lower tension in the junction region due to the bulge. Interestingly, different pairs of helices are stacked coaxially for two bulged RNAs, which is probably the result of tertiary contacts between the relatively large bulge and one of the helices. The exact reasons for this rearrangement may be rather complex and cannot be elucidated by smFRET experiments. Noteworthy, the deviation from the planarity of the studied three-way junction molecules is very small irrespective whether there is bulge or not in the junction. Another significant fact is that two helices are coaxially stacked for bulged molecules, which is in a good agreement with solid phase data of natural RNA three-way junction systems from literature [5, 10]. Obtained results for fully paired RNAs are novel and there are no other studies to be compared with.

Finally, conformational space studies were performed for RNA three-way junction models. Entire accessible conformational space and its confinement for the FRET preferred structures is discussed.

CONTENTS

1 Introduction

It is well known that RNA molecules play an important role in living organisms [3, 4, 11-13] and the functionality of these molecules is strongly related to their tertiary structures. Hence, the understanding of how structure and function are related in biomolecules is in current research often pursued by determining structures of naturally occurring functional molecules. This approach leads to the important knowledge on how a specific molecule is able to accomplish its functionality. However, quite frequently this knowledge is not easily transferable to related compounds. In the case of RNA molecules, naturally occurring systems mostly exhibit a complex structure, containing several structural elements like junctions, loops and bulges at a time [5, 6].

Furthermore, folding often appears to be influenced by tertiary interactions between different sites. Determining structures of such complex systems does in most cases not provide detailed information about the influence of each single structural element. Although this knowledge is highly relevant to structural biology and biophysics, systematic structural studies applying modern techniques with high spatial resolution are rare. Recently, NMR spectroscopy visualized transient low-populated structures of a small RNA [14, 15]. However, structure determination of large, dynamically interconverting and structurally heterogenous RNAs was not possible by classical methods.

Helical junctions are the main architectural building blocks of RNA tertiary structures and have a major influence on RNA dynamics and function [16]. In ribozymes, junctions accelerate the folding into biologically active tertiary structure by up to three orders of magnitude [17, 18]. Thus, in order to understand the structural heterogeneity of large RNAs and its influence on folding, regulation, and functional complexity, advancing necessery knowledge of helical junctions is a key step, which also requires progress in techniques for structure determination.

For example, no structures are available for fully-paired RNA four-and three-way junctions without internal loops in the junction region among the structures listed presently in the PDB. Hence, this work is dedicated to the determination of 3D structures of RNA branched molecules (section 5) employing FRET technique.

Since the first observations of single-molecules with optical detection methods [19, 20] fluorescence spectroscopy has become an important tool in the study of the dynamic and conformational properties of biomolecules [21]. In experiments, in fact, ensemble averaging is avoided and the information on the heterogeneities and dynamic properties of a system is directly accessible [22].

Förster Resonance Energy Transfer (FRET) is one of the techniques used to study structure and dynamics of biomolecules on single-molecule level. FRET is a highly distance dependent mechanism describing energy transfer from one chromophore to another [7]. FRET is used as a "spectroscopic ruler" [8], as it is particularly suitable for the distance measurements between fluorescent dyes typically in the range of 30 – 80 Å.

In FRET, energy is transferred non-radiatively from an excited donor (D) to an acceptor (A) chromophore, provided they are in close proximity. Advantages of FRET are: high distance sensitivity ($1/R^6$ – dependence); fluorescent dyes can be site-specifically covalently bound to nucleic acids, proteins and lipids; FRET labeled molecules can be measured while immobilized on surfaces as well as freely diffusing in solution; due to the availability of fluorophores with high quantum yield and photostability [23] and highly sensitive detection systems [20, 24, 25] with high time resolution, measurements on the level are possible [26, 27] (smFRET). Hence, it is possible to obtain information about structural heterogeneity and dynamic transition in biomolecules under physiological conditions (in vitro [26, 28-30] and even in vivo [31]) with ~ nanosecond time resolution, determined by fluorescence lifetime of the dyes. Additionally, there are practically no limitations on the size of the biomolecule of interest. In summary, the fact that most biomolecules are dynamic and undergo intrinsic motions [32, 33] makes FRET favorable in comparison to X - ray crystallography (no heterogeneities and dynamics, only molecules that crystalize), NMR spectroscopy (strong limitations in size) and cryo-EM (no dynamics, not under physiological conditions) the three most well established techniques in structural biology.

Fluorescence Correlation Spectroscopy (FCS) is another powerful technique used to study processes that induce a fluctuation of the fluorescence signal [34, 35]. The characteristic relaxation times that describe the kinetic properties of each process are obtained in FCS by fitting the correlation function of the fluorescence signal. FCS is extremely useful in studying wide range of processes happening over several order of correlation times, such as

translational and rotational diffusion [36], chemical reactions [37, 38] and conformational changes [39]. In studies performed on single-molecule level FCS is used to separate the different correlation condtributions and to obtain the correct molecular fractions when multiple species are present in the measurement solution.

Main goals of my studies were the improvement of the multi-parameter fluorescence detection (smMFD) technique, used for single-molecule FRET (smFRET) experiments, and its application to the structural studies of RNA branched molecules. Hence, my thesis consists of the following topics:

1. presentation of fast hardware photon correlator implemented in a field-programmable gate array (FPGA) combined with a compact confocal fluorescence setup,
2. demonstration of the accurate smFRET-based structural models of RNA four-and three-way junctions, which proves FRET as a quantitative tool in the field of structural biology.

In section 4 fast hardware photon correlator implemented in a field-programmable gate array (FPGA) combined with a compact confocal fluorescence setup is presented. The correlator has two independent units with a time resolution of 4 ns while utilizing less than 15 % of a low-end FPGA. The device directly accepts transistor-transistor logic (TTL) signals from two photon counting detectors and calculates two auto- or cross-correlation curves in real time. Test measurements demonstrate that the performance of the correlator is comparable with the current generation of commercial devices. The sensitivity of the optical setup is identical or even superior to current commercial devices. The FPGA design and the optical setup both allow straightforward extension to multi-color applications.

In section 5 structural models of a large RNA four-and three-way junction molecules (≈ 45 kD) based on FRET experiments, molecular dynamics (MD) and SimRNA simulations are presented. Furthermore, the structural changes in three-way junction induced by the insertion of two bulges with different lengths in the junction region are investigated. In order to fully understand the structural behavior of three-way junction molecules, incomplete

variants, which resulted from the omission of the third strand from fully complementary three-way junctions, were also investigated.

Additionally, sterically allowed and FRET preferred conformational space for three-way junction models was investigated. Regions with higher and lower enthalpy of the structures (potential energy of the “bonds” and clashes in the junction region) as well as the role of the stacking interactions in the junction region are discussed. The changes in occupied conformational space induced by the bulges at the junction region are demonstrated.

2 Fluorescence

The emission of light from any substance which occurs from electronically excited states is called luminescence. Depending of the nature of excited state, two categories of luminescence are known: fluorescence and phosphorescence. A typical fluorescence lifetime (τ_0) of fluorophores (general name of light absorbing groups) or more commonly fluorescent dyes is about 10 ns. Fluorescence typically occurs from aromatic molecules (Figure 2-1) as they have a singlet ground state (Figure 2-2).

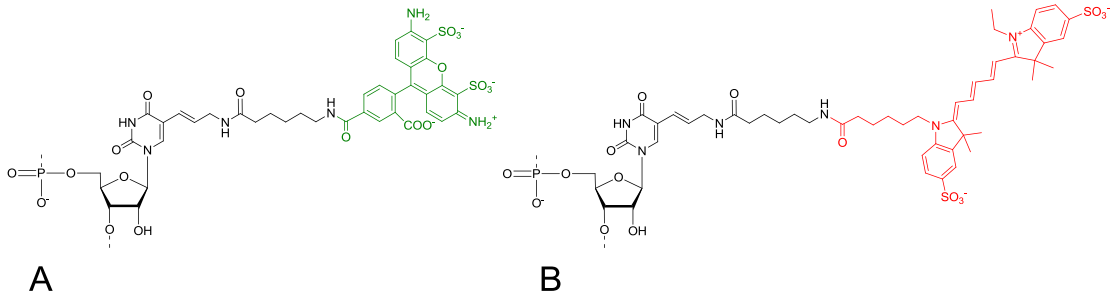


Figure 2-1 (A) Example structures of modified Uracil and the hexamethylen (C6) linkers with the Alexa 488 (A) and Cy5 (B) dyes used for RNAs. Donor and acceptor dyes are depicted in green and red, respectively.

Schematic representation of processes which occur between absorption and emission are explained by the Jablonski diagram [40] (Figure 2-2).

In solution, several processes usually occur following light absorption. The fluorophore is usually excited to higher vibrational level of either S1 or S2. Afterwards, nonradiative rapid vibrational relaxation occurs to the corresponding vibrational level. Since this process is very fast, between 10^{-14} s and 10^{-11} s; it is extremely likely to occur immediately following absorbtion. This relaxation occurs between vibrational levels, so generally electrons will not change from one electronic level to another through this method. However if vibrational energy levels strongly ovelap electronic energy levels, excited electron can transit from vibrational level in one electronic state to another vibrational level in a lower electronic state. This process is called internal conversion IC and mechanistically is identical to vibrational relaxation. IC occurs in the similar time range as vibrational relaxation, therefore, is also likely to happen following absorbtion. Yet another path a molecule may take in the dissipation of energy is called intersystem crossing (ISC), where the electron changes spin multiplicity from an excited singlet state to an excited triplet state. ISC is several orders of

magnitude slower than fluorescence (typically near 10^{-8} s), hence, fluorescence is observed only from the S_1 state and emission is independent of the excitation wavelength, Kasha's rule [41].

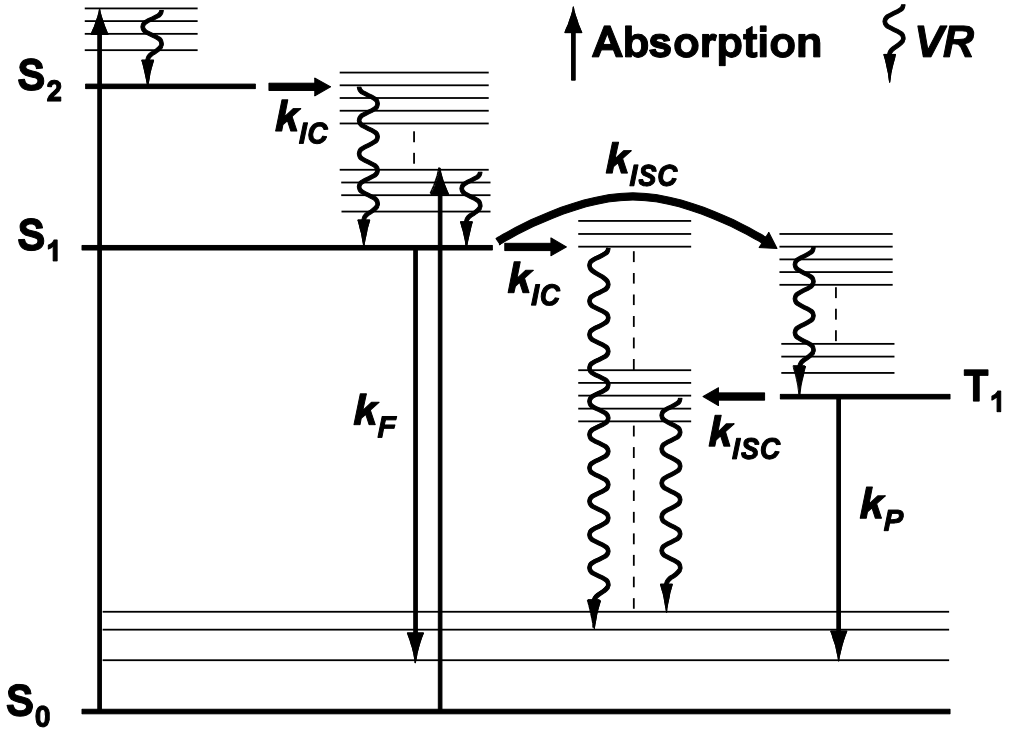


Figure 2-2 Jablonski diagram, S_0, S_1, S_2 and T_1 refer to the singlet and triplet states of the dye, respectively. k_F, k_{IC}, k_{ISC}, k_P, are the rate constants of fluorescence, internal conversion, inter-system crossing and phosphorescence, respectively. Vibrational relaxaton, *VR*, is indicated by wavy arrows. *IC, ISC* and *VR* are all non-radiative processes.

In this case fluorescence lifetime can be calculated with following differential equation,

$$\frac{dF(t)}{dt} = -(k_f + k_{IC} + k_{ISC})F(t) \qquad \textbf{Eq. 2.1-1}$$

where k_f, k_{IC} and k_{ISC} are the fluorescence, internal conversion and inter system crossing rate constants, respectively. For simplicity $k_f + k_{IC} + k_{ISC} = k_0$. The solution of Eq. 2.1-1 will be a first order exponential decay,

$$F(t) = F(0)\exp[-k_0 t] = F(0)\exp\left[-\frac{t}{\tau_0}\right] \qquad \textbf{Eq. 2.1-2}$$

$$\text{with } \tau_0 = \frac{1}{k_0}$$

The efficiency of fluorescence is quantified by the fluorescence quantum yield, $\Phi_{F(0)}$, which is defined by equation Eq. 2.1-3,

$$\Phi_{F(0)} = \frac{Emitted\ \ photons}{Absorbed\ \ photons} = \frac{k_f}{k_0} \qquad \textbf{Eq. 2.1-3}$$

Fluorescence lifetime and quantum yield are distinctive parameters of a fluorophore as the rate constants of each relaxation process, shown in the Jablonski diagram, are different for different molecules.

2.1 Fluorescence anisotropy

When excited with the polarized light, fluorophores tend to absorb components of the light with polarization parallel to their absorption dipole moment. Emission in its turn occurs with the polarization parallel to fluorophore's emission dipole moment. The extent of polarization of the emission is described in terms of fluorescence anisotropy, *r*.
In general case the fluorescence anisotropy is calculated as follows,

$$r = \frac{I_{//} - I_{\perp}}{I_{//} + 2I_{\perp}} \qquad \textbf{Eq. 2.1-1}$$

where $I_{//}$ and $I_{\perp}$ are observed intensities parallel and perpendicular to the linearly polarized excitation light respectively.

Fundamental anisotropy, r_0, is described with Eq. 2.1-2,

$$r_0 = \frac{3\cos^2(\beta) - 1}{5} \qquad \textbf{Eq. 2.1-2}$$

where β is the angle between excitation and emission transition moments. Fundamental anisotropy is a function of β and therefore varies from fluorophore to fluorophore. Anisotropy is given by Perrin equation [40],

$$r = \frac{r_0}{1 + \frac{\tau}{\rho}} \qquad \textbf{Eq. 2.1-3}$$

where r_0 is a fundamental anisotropy, τ is the fluorescence lifetime and ρ is the rotational correlation time.

2.2 Fluorescence Resonance Energy Transfer (FRET)

Förster Resonance Energy Transfer (FRET) is a distance dependent non radiative energy transfer from excited states of donor (D) dye to an acceptor (A) dye. FRET efficiency accurately measures distances between donor and acceptor dyes within the range of 30 – 80 Å and therefore it is often called molecular ruler. Understanding of the theory behind this technique is very important as it is the basis of all structural and dynamical investigations of the current work. The derivation of the equation presented in this section can be found in following works: see ref. [7, 42-45]).

Alexa488 and Cy5 were used in my studies as a D and A dyes, respectively (Figure 2-1). The rate of quenching of excited D by FRET is given by Eq. 2.2-1 [46],

$$k_{FRET} = \frac{1}{\tau_{D(0)}} \left(\frac{R_0}{R_{DA}} \right)^6 \qquad \textbf{Eq. 2.2-1}$$

In this equation $\tau_{D(0)}$ is the donor lifetime in the absence of A, R_{DA} is the distance between the two dyes and R_0 is the Förster radius, which equals the distance between DA at which FRET efficiency, E (Eq. 2.2-4), equals to 0.5 (Figure 2-3). The Förster radius is given by the following Eq. 2.2-2 [46],

$$R_0 = \left[\frac{9(\ln 10)}{128\pi^5 \cdot N_A} \cdot \frac{J(\lambda) \cdot \kappa^2 \cdot \Phi_{FD(0)}}{n^4} \right]^{\frac{1}{6}} \qquad \textbf{Eq. 2.2-2}$$

where N_A is the Avogadro constant, n is the refractive index of the medium between the dyes, $\phi_{D(0)}$ is the fluorescence quantum yield of D in absence of FRET, κ^2 is the dye orientation factor (Eq. 2.2-5) and $J(\lambda)$ (Figure 2-3) is the overlap integral between the absorption spectrum of A, $\varepsilon_A(\lambda)$, and the normalized emission spectrum of D, $F_D(\lambda)$. It can be written as:

$$J = \int_0^\infty F_D(\lambda)\varepsilon_A(\lambda)\lambda^4 d\lambda \qquad \textbf{Eq. 2.2-3}$$

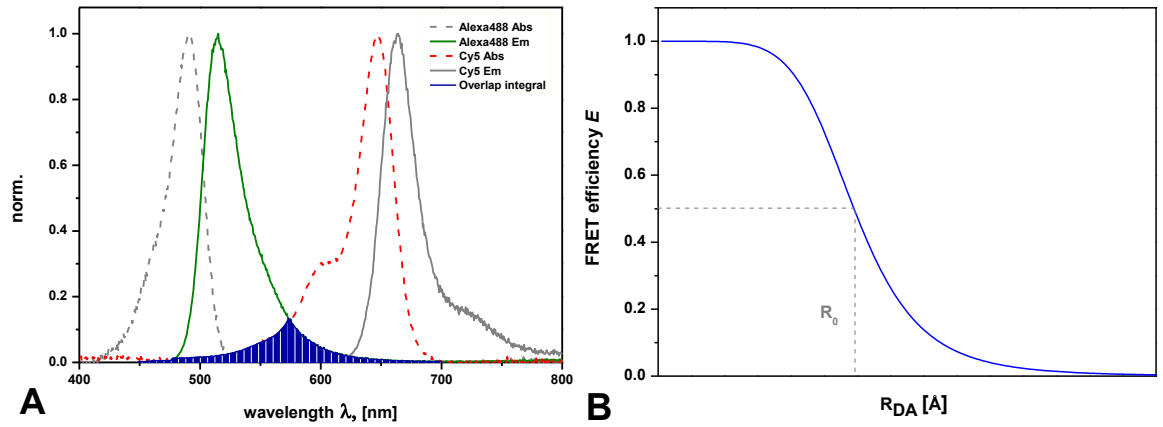


Figure 2-3 (A) Overlap integral $J(\lambda)$ (blue) between normalized fluorescence and absorbance spectra of Alexa488 (green) and Cy5 (red dashed), respectively. The absorption of Alexa488 and the emission of Cy5 are shown as gray dashed and gray solid lines, respectively. (B) The FRET efficiency E is strongly dependent on R_{DA} (see Eq. 2.2-4). E is most sensible to changes in R_{DA} in the region of R_0 ($E(R_0)$ = 0.5).

One of the most useful expressions for FRET efficiency calculation is shown in the next equation (Eq. 2.2-4) [46],

$$E = 1/(1 + R_{DA}^6 / R_0^6) \qquad \textbf{Eq. 2.2-4}$$

This means that E can be determined by measuring the fluorescence of donor (F_D) and the acceptor (F_A) as well as by measuring lifetimes of excited D in presence and absence of acceptor.

The orientation factor κ^2 depends on mutual orientation of the transition dipole moments of D and A dyes. κ^2 can be calculated with the help of the following equation (Eq. 2.2-5),

$$\kappa^2 = \left[\hat{\mu}_A \cdot \hat{\mu}_D - 3\left(\hat{\mu}_A \cdot \hat{R}_{DA}\right)\left(\hat{\mu}_D \cdot \hat{R}_{DA}\right)\right]^2 = (\sin\theta_D \sin\theta_A \cos\varphi - 2\cos\theta_D \cos\theta_A)^2 \qquad \textbf{Eq. 2.2-5}$$

One can see from equation (Eq. 2.2-5) that κ^2 can range between 0 and 4, causing big errors in FRET data analysis. Nevertheless if the rotational diffusion of the dyes is unrestricted and much faster than k_{FRET}, κ^2 can be assumed to be 2/3.

relative orientation of dipoles	κ^2
	1
	4
	0
dynamic averaging (rotational diffusion)	2/3

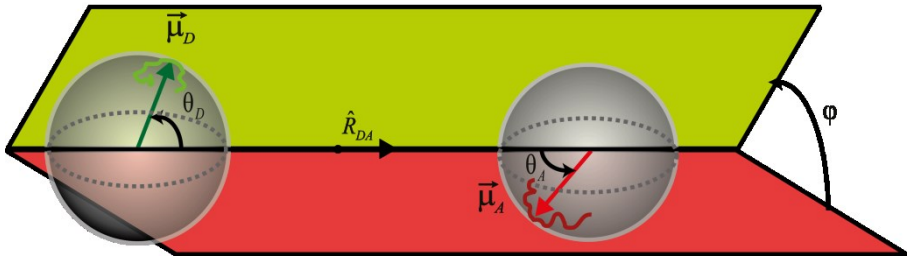


Figure 2-4 Sketch showing angles which define the orientation factor κ^2 in Eq. 2.2-5: θ_D and θ_A are the angles between the transition dipole moments of the donor and acceptor dyes (μ_D and μ_A) and the distance unit vector $\hat{R}_{DA}$, respectively, and φ is the angle between μ_D and μ_A.

2.3 Fluorescence Correlation Spectroscopy (FCS)

Fluorescence correlation spectroscopy (FCS) is a powerful statistical method to study diffusion and dynamics of biomolecules [47-50] by detecting systematic fluctuations of a fluorescent signal. As discussed before (section 2.2), FRET depends strongly on the DA distance. Hence, the donor and acceptor signal fluctuates when the distance between the dyes is changed.

FCS measurements were performed on confocal setup using multi-parameter fluorescence detection (smMFD) (section 3.2.1) for the detection of the fluorescence signal of molecules diffusing through the detection volume.

Correlation function $G_{ij}(t_C)$ of the fluorescence signal $F(t)$ at time t, allows to estimate the characteristic time of each process causing the fluctuations in the fluorescence intensity,

$$G_{ij}(t_C) = \frac{\langle F(t+t_C)_i F(t)_j \rangle}{\left(\langle F_i(t) \rangle^2 \langle F_j(t) \rangle^2 \right)} = 1 + \frac{\langle \delta F(t) \delta F(t+t_C) \rangle}{\langle F \rangle^2} \qquad \textbf{Eq. 2.3-1}$$

where $\delta F(t)$ is a fluctuation around the average signal $\langle F \rangle$. Subscripts i and j in correlation function $G_{ij}(t_C)$ refer to distinct detection channels and depending on whether $i = j$ or $i \neq j$, G function represents autocorrelation or cross-correlation functions.

In Figure 2-5 the correlation curve for freely diffusing DNA in an open volume is presented. The timescale ranges from nanoseconds, which corresponds to fluorescence emission (antibunching), to milliseconds, describing the translational diffusion of DNA. For the exemplary curve emission, rotational diffusion, triplet formation and translational diffusion processes are shown.

The correlation function can be written as a product of the several correlation terms normalized by the average number N of bright molecules in the detection volume (Eq. 2.3-2) [36],

$$G(t_C) = 1 + \frac{1}{N}[G_D(t_c)G_T(t_c)G_R(t_c)G_A(t_c)] \qquad \textbf{Eq. 2.3-2}$$

Here *D*, *T*, *R* and *A* represent translational diffusion, triplet, rotation and antibunching process, respectively. This expression, however, is accurate when the processes are independent or when the difference between timescales of the processes is at least one order of magnitude. If the processes occur in a closer time range, Eq. 2.3-2 can produce inaccurate results.

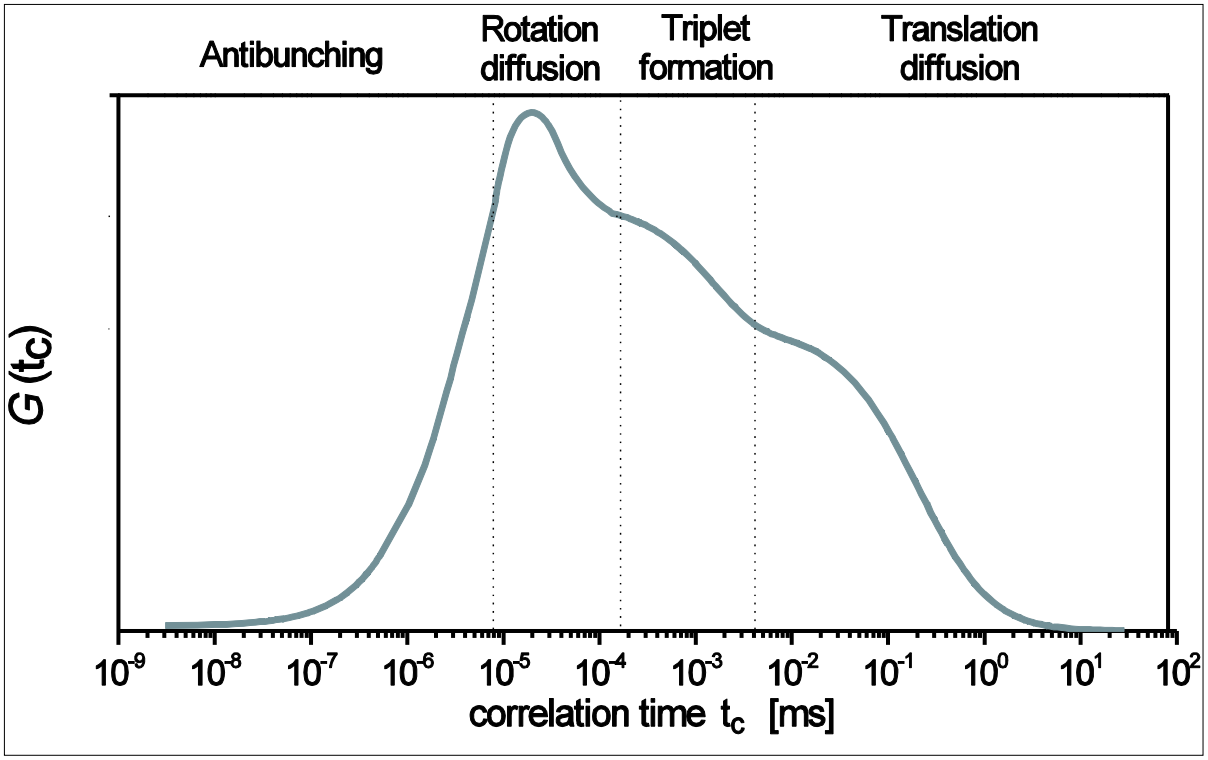


Figure 2-5. Correlation curve of the fluorescent labelled DNA molecule.

In my studies the focal volume is well approximated by a three-dimensional Gaussian with spatial distribution of the detection probabilities w,

$$w(x,y,z) = \exp\left(-2\left(x^2+y^2\right)/\omega_0^2\right)\exp\left(-2z^2/z_0^2\right) \qquad \textbf{Eq. 2.3-3}$$

Here ω_0 and z_0 are the $1/e^2$ radii in the lateral or in the axial direction, respectively. As described in [36], from Eq. 2.3-2 and Eq. 2.3-3 we have,

$$G(t_C)=1+\frac{1}{N}\cdot\left[\left(\frac{1}{1+\frac{t_C}{t_D}}\right)\times\left(\frac{1}{1+\frac{t_C}{\left(\frac{z_0}{\omega_0}\right)^2\cdot t_D}}\right)^{1/2}\times\left(1-T+T\cdot\exp\left[-\frac{t_C}{t_T}\right]\right)\times\left(1+R\cdot\exp\left[-\frac{t_C}{t_R}\right]\right)\times\left(1-\exp\left[-\frac{t_C}{t_A}\right]\right)\right] \quad \textbf{Eq. 2.3-4}$$

where t_d is the diffusion time, z_0/ω_0 is aspect ratio of the confocal volume, T is the equilibrium fraction of the triplet state, t_T is the triplet relaxation time, R and t_R are the amplitude and characteristic time of rotational diffusion and t_A is the antibunching time.

In the particular case when complex triplet and/or radical kinetics at high irradiance occur, Eq. 2.3-4 should be rewritten in the following way:

$$G(t_C)=1+\frac{1}{N}\cdot\left[\left(\frac{1}{1+\frac{t_C}{t_D}}\right)\times\left(\frac{1}{1+\frac{t_C}{\left(\frac{z_0}{\omega_0}\right)^2\cdot t_D}}\right)^{1/2}\times\left(1-T+T\cdot\exp\left[-\frac{t_C}{t_T}\right]\right)\times\left(1-K+K\cdot\exp\left[-\frac{t_C}{t_K}\right]\right)\times\left(1-\exp\left[-\frac{t_C}{t_A}\right]\right)\right] \quad \textbf{Eq. 2.3-5}$$

In Eq. 2.3-5 K and t_K are the amplitude and the characteristic time of the kinetic term, caused by the processes mentioned above. For calculation of correlation functions, hardware photon correlators are often employed [51].

3 Instrumentation

3.1 Time Correlated Single Photon Counting (TCSPC)

Ensemble time-correlated single-photon-counting (eTCSPC) measurements were performed using an IBH-5000U (IBH, Scotland) system. Data was obtained in reverse start-stop mode, which means that decay histogram was built from delay times between fluorescence photon and the following excitation pulse. TAC time (time between two consecutive laser pulses) was divided in 4096 TAC or TCSPC channels. Instrument response function, IRF, was obtained from a concentrated solution of Ludox, which is a highly scattering solution. By deconvolution of IRF, fluorescence decays were computed. Two excitation sources were

used: 470 nm diode laser (LDH-P-C 470, Picoquant, Berlin, Germany) operating at 8 MHz and 635 nm diode laser (LDH-8-1 126, Picoquant, Berlin, Germany) operating at 10 MHz for donor emission and an acceptor emission, respectively. The emission wavelength was set to 520 nm for donor emission and 665 nm for acceptor emission. The corresponding monochromator slits were set to 2 nm and 16 nm resolution for excitation and emission paths, respectively. In order to reduce the contribution of scattered light additional cut-off filters were used: > 500 nm for donor emission and > 640 nm for acceptor emission. Measurements were performed at room temperature and the concentration of the measured samples was < 1 μM.

Lifetime information from measurements was obtained by fitting fluorescence intensity decay curves using the iterative re-convolution approach [52]. The maximum number of counts was usually 50.000. The fluorescence decays $F(t)$ were modeled by double exponential decays (Eq. 3.1),

$$F(t) = x_1 \exp(-t/\tau_1) + x_2 \exp(-t/\tau_2) \qquad \textbf{Eq. 3.1}$$

where x_1, x_2 and τ_1, τ_1 are amplitudes and lifetimes of the first and the second species, respectively.

3.2 Confocal microscope setup

3.2.1 For SMD

Measurements were performed on an epi-illuminated confocal microscope (IX70, Olympus, Hamburg, Germany) with a 60x/1.2 water immersion objective (UPlanSApo 60x/1.2w, Olympus Hamburg, Germany) (Figure 3-1).

The fluorescent donor molecules (Alexa 488) are excited by a linearly polarized, active-mode-locked Argon-ion laser (Innova Saber, Coherent, Santa Clara, CA, USA, 496.5 nm, 73.5 MHz, ~ 300 ps) or by a 485 nm diode laser (LDH-D-C 485, Picoquant, Berlin, Germany) operating at 64 MHz. The laser light is focused into the dilute solution (< 50 pM) of labeled molecules by a 60x/1.2 water immersion objective. Each molecule generates a brief burst of fluorescence photons as it traverses the detection volume. This photon-train is divided

initially into its parallel and perpendicular components via a polarizing beamsplitter and then into a wavelength ranges below and above 595 nm by using a dichroic beamsplitter (595 DCXR, AHF, Tübingen, Germany). Additionally, red (HQ 720/150 nm for Cy5) and green (HQ 533/46 nm for Alexa 488 and Rh110) bandpass filters (both made by AHF, Tübingen, Germany) in front of the detectors ensure that only fluorescence photons coming from the acceptor and donor molecules are registered. An estimate of the focal geometry is acquired by determining the diffusion correlation time of 200 ± 13 µs for Rhodamine 110 and knowing its diffusion coefficient of 0.34 ± 0.03 $\mu m^2/ms$. Detection is performed using four avalanche photodiodes (SPCM-AQR-14, Laser Components, Germany or alternatively for the green channels PDM050CTC, or τ-SPAD-100, both PicoQuant, Berlin, Germany) . The signals from all detectors are guided through a passive delay unit and two routers to two synchronized time-correlated single photon counting boards (SPC 132 or SPC 832, Becker and Hickl, Berlin, Germany) connected to a PC. Bursts of fluorescence photons are distinguished from the background of 1 – 2 kHz by applying certain threshold intensity criteria [53]. Bursts during which bleaching of the acceptor occurs are excluded from further analysis by applying a criterion regarding the difference in macroscopic times, $|T_G - T_R| < 0.5$ ms, where T_G and T_R are the average macroscopic times in which all photons have been detected in the green and red channels respectively during one burst [54].

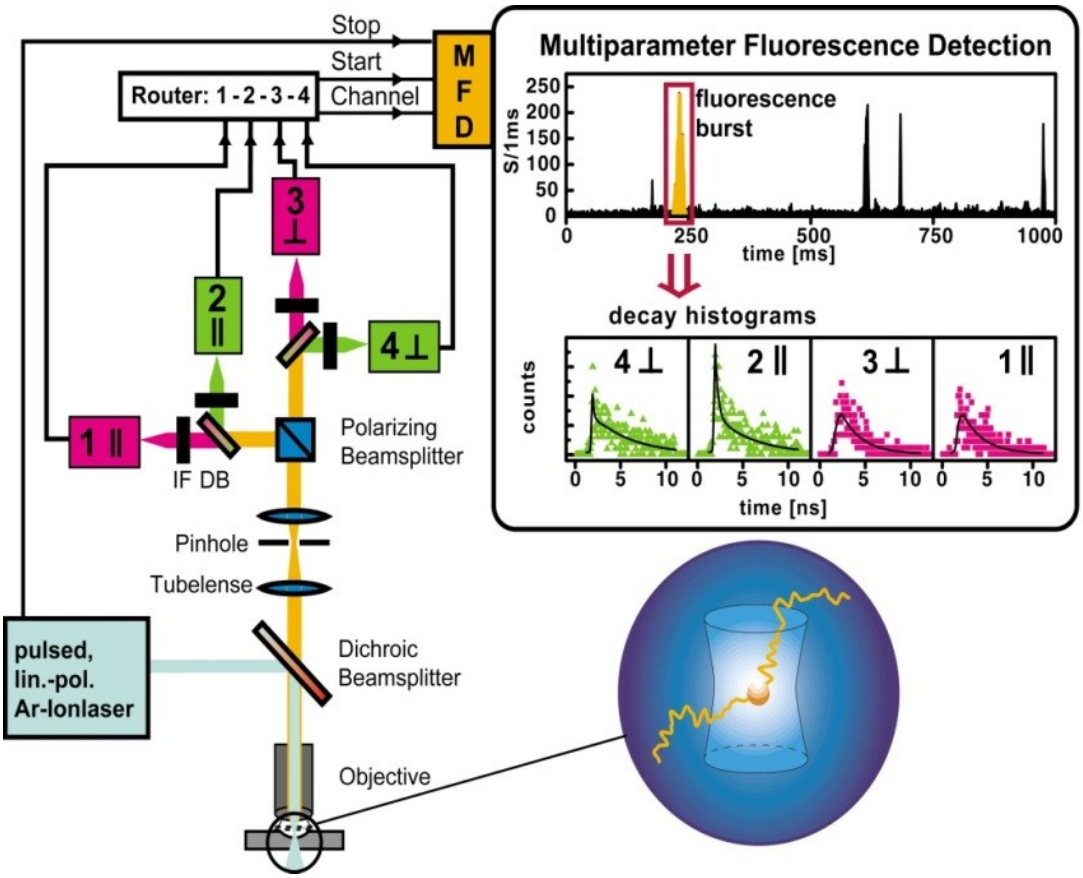


Figure 3-1 Experimental setup for smMFD measurements. For more details see [22].

In addition to detecting incomplete labeling and characterizing fluorophore quenching and mobility, which allows addressing the quality of the FRET measurements, smMFD resolves multiple FRET states [29, 55, 56], which will be used later in this work.

3.2.2 For FCS

FCS measurements of Rhodamine 110 (Rh110) diffusion and photophysics were performed using a home-built confocal setup (Figure 3-2).

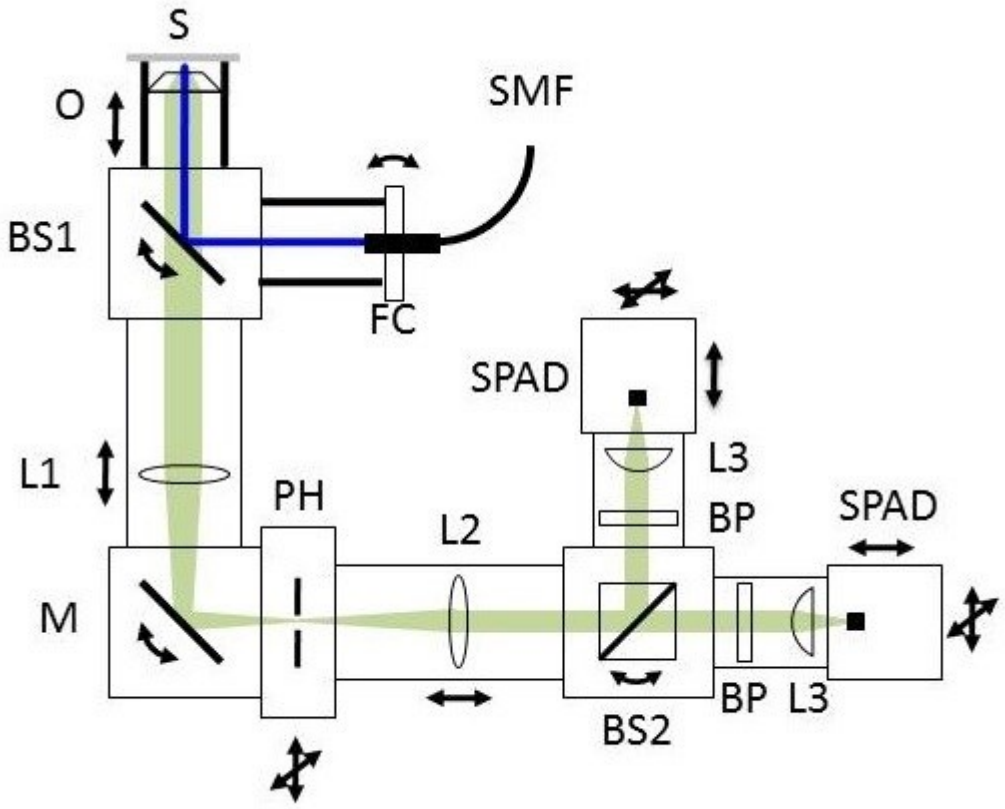


Figure 3-2 Schematic drawing of the FCS setup used to perform test measurements.BS1: dichroic beamsplitter; BS2: polarizing beamsplitter; L1: tube lens f = 160 mm; L2: collimating lens f = 100 mm; L3: lens f = 10 mm; SMF: single mode fiber; FC: fiber coupler with 2-axis tilt; O: objective; S: sample plate; M: broadband mirror; PH: pinhole on 8 position wheel; BP: band pass filter; SPAD: single photon avalanche diode.

The excitation source is a tunable Ar-ion laser (35-LAP-431-220, Melles-Griot, Bensheim, Germany) set to 496 nm and coupled via single mode fiber (Schäfter&Kirchhoff, Hamburg, Germany) to a modular system consisting of galvanized aluminum cubes connected via dove-tail adapters. The main dichroic (BS 500, AHF Tübingen, Germany) is mounted on a micrometer driven rotation and tilt manipulator inside the first cube. The objective (UPlanSApo 60x/1.2 w, Olympus Hamburg, Germany) is attached to a z-micrometer (SM1Z, Thorlabs Dachau, Germany) at the exit port. A tube lens (f = 160 mm achromatic lens, Linos, Göttingen, Germany) focuses the fluorescence light leaving the second exit port onto a pinhole (Plano Wetzlar, Germany), which is mounted on an 8-position wheel. Pinhole sizes can be varied between 25 µm and 5 mm (70 µm was used in the presented experiments). The spatially filtered light is then collimated by a second lens (f = 100 mm achromatic lens,

Linos) before being divided by a polarizing beam splitter (TSWP 633, Linos). After passing bandpass filters (HC525/39, AHF) to remove scattered laser light and limit the detected fluorescence range the two beams are finally focused by two plano-convex lenses (f = 10 mm, Linos) onto two single photon avalanche diodes (PD200A, MPD Bolzano, Italy or PerkinElmer SPCM AQR-14/AQRH-14). Detectors are attached via xy-micrometer manipulators. All optical components are broadband AR coated and can be exchanged quickly. The modular design allows for easy extension, i.e. to set up four, six or eight detection channels to record multiparameter data (polarization and various spectral ranges). The sensitivity of the whole setup was found to be at least as sensitive as commercial microscopes like the Olympus IX71 equipped with equivalent detectors.

4 Hardware photon correlator

4.1 Introduction

In order to calculate the fluorescence correlation functions, hardware correlators [51] or real-time software correlation is used [57-60]. In both cases a dedicated data acquisition or processing board is required to build a FCS setup. In this work a field-programmable gate array (FPGA)-implementation of a hardware photon correlator is introduced. Specifically, it is based on a general-purpose Xilinx SP605 evaluation board (Xilinx, USA) equipped with a value line Spartan 6 FPGA chip (XC6SLX45T). Photon detectors can be directly connected to the SP605 board requiring no additional custom-built hardware. Test measurements of fluorescence fluctuations of Rhodamine 110 show that the time resolution of 4 ns is easily achieved in practice. In this respect, design of the correlator is comparable with “fast” versions of commercial devices, such as the ALV 6010/200. Two correlator units utilize less than 15% of the FPGA’s resources, which suggests that the design should fit even into low-end FPGAs. Alternatively, more correlators can be implemented in parallel without sacrificing time resolution, which is for instance useful for FCS in combination with Förster resonance energy transfer (FRET). Additional features of the correlator include a real-time display of photon count rates, and a display of the intensity trace with millisecond time resolution.

4.2 Methods and materials

Performance of the correlator is demonstrated using a compact home-built optical setup employing only the minimum of required optical components, ensuring maximized optical throughput and stability of alignment (section 3.2.2).

4.3 Results

The FCS curves generated with current SP605-based design and commercial ALV 6010/200 correlator with a specified time resolution of 5 ns are compared in Figure 4-1.

In the ~ 100 ns – 1 s range, the curves are hardly distinguishable. The photon antibunching term [61] is slightly more pronounced in the curve computed using the correlator at shorter correlation times. Afterwards the correlation curves were fitted using Eq. 4.3-1, accounting for translational diffusion, triplet and reaction kinetics, and photon antibunching.

$$G(t_c) = 1 + \frac{1}{N}\left(\frac{1}{1+t_c/t_d}\right)\left(\frac{1}{1+t_c/\left[(z_0/\omega_0)^2 t_d\right]}\right)^{1/2}\left(1-T+T\exp(-t_c/t_T)\right) \times\left(1-K+K\exp(-t_c/t_K)\right)\left(1-\exp(-t_c/t_A)\right) \quad \textbf{Eq. 4.3-1}$$

Here N is an average number of bright molecules in the detection volume, t_d is the diffusion time, z_0/ω_0 is aspect ratio of the confocal volume, T is the equilibrium fraction of the triplet state, t_T is the triplet relaxation time, and t_A is the antibunching time.

The second kinetic term with an amplitude K and a characteristic time t_K is caused by saturation and complex triplet and/or radical kinetics at high irradiance [62, 63]. Investigation of this effect however is out of the scope of this work.

The fitted parameters for the FCS curves generated using proposed design and the ALV 6010/200 agree typically within a few percent (see Figure 4-1 caption). In particular N and t_D show almost perfect agreement (N: 0.4%; t_D: 0.2%).

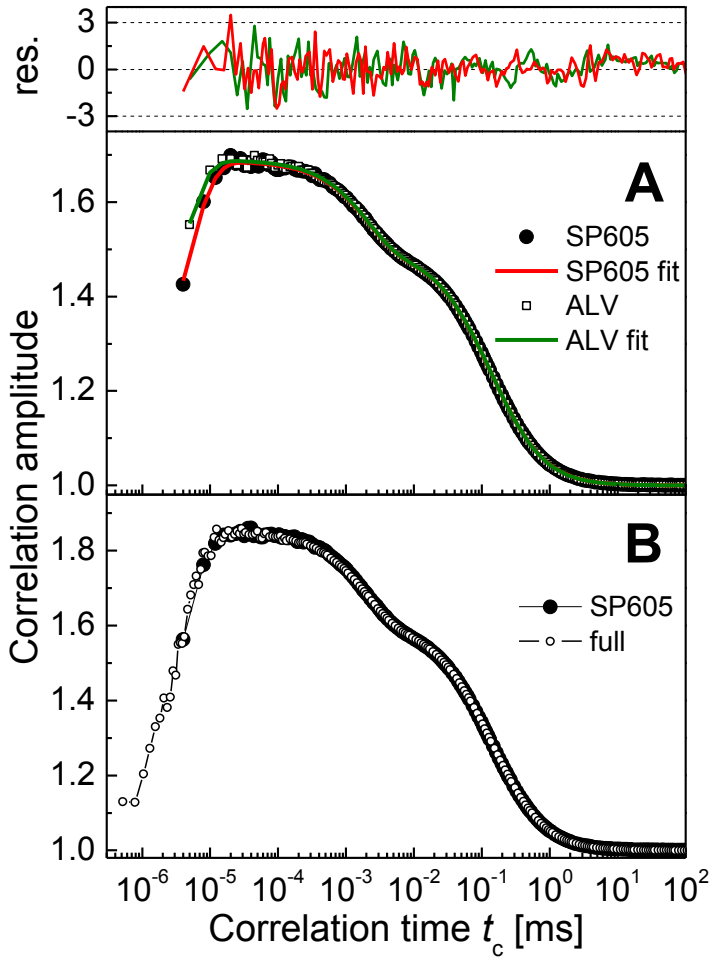


Figure 4-1 Comparison of the FCS curves of Rhodamine 110 generated by ALV 6010/200 (□) and the SP605-based correlator presented in this work (●). The curves are fitted by (Eq. 4.3-1) with the following parameters: ALV: N = 1.450; t_D = 0.144 ms; z_0/ω_0 = 4.1; T = 0.274; t_T = 2.03 µs; K = 0.217; t_K = 0.591 ms; t_A = 3.1 ns; SP605: N = 1.455; t_D = 0.144 ms; z_0/ω_0 = 4.0; T = 0.272; t_T = 2.03 µs; K = 0.186; t_K = 0.52 µs; t_A = 4.0 ns. Weighted residuals are shown above. B: Comparison with the FCS curve calculated by software correlation of photon traces recorded with TCSPC cards (○). In all cases two cross-correlation functions (detector1→detector2 and detector2→detector1) were calculated and the average was taken. In plot (B) the curves were measured with slightly different concentrations of Rh110 and were scaled to the same amplitude. The mean irradiance was ca. 50 kW/cm^2 (450 µW at the objective).

The deviations can be attributed to limited statistics of the measurements and to instability of the fit with respect to K and t_K because this process is slower than diffusion. Weighted residuals (Figure 4-1A, upper panel) show no features specific for either curve. To ensure that the observed anticorrelation in the ns range is not due to an artifact, photon traces with picosecond time resolution using two TCSPC cards (SPC832, Becker & Hickl, Germany) on the same experimental setup were recorded. Then the full software correlation of the recorded

signals [59] was performed. Figure 4-1B shows that FCS curve from the newly developed correlator agrees very well with the computed full correlation curve over the whole time range. This comparison convincingly demonstrates that the electronic time resolution of 4 ns is readily achieved in practice.

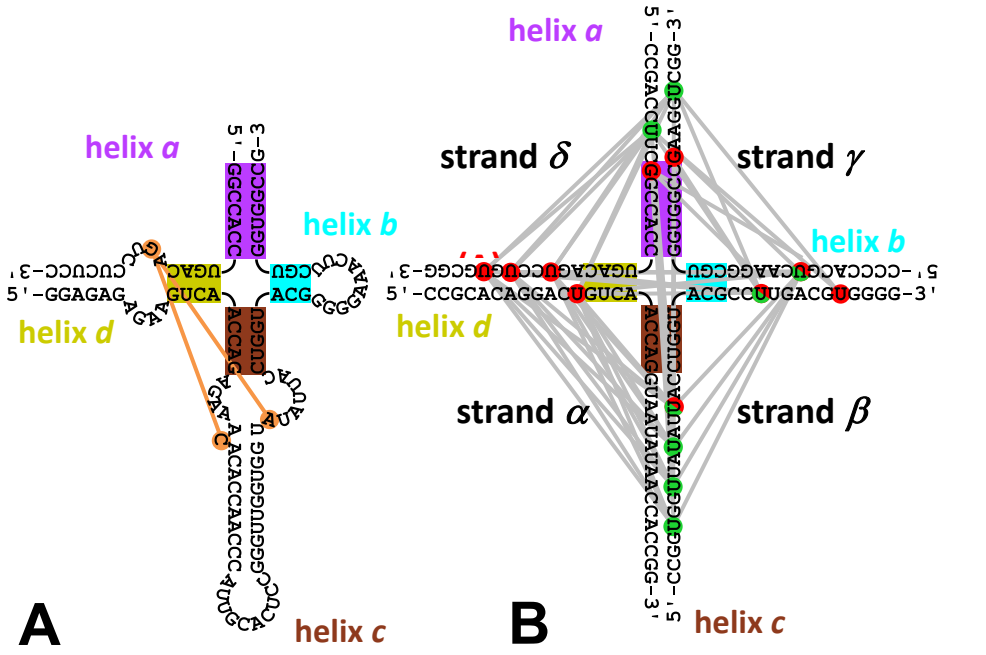
4.4 Discussion

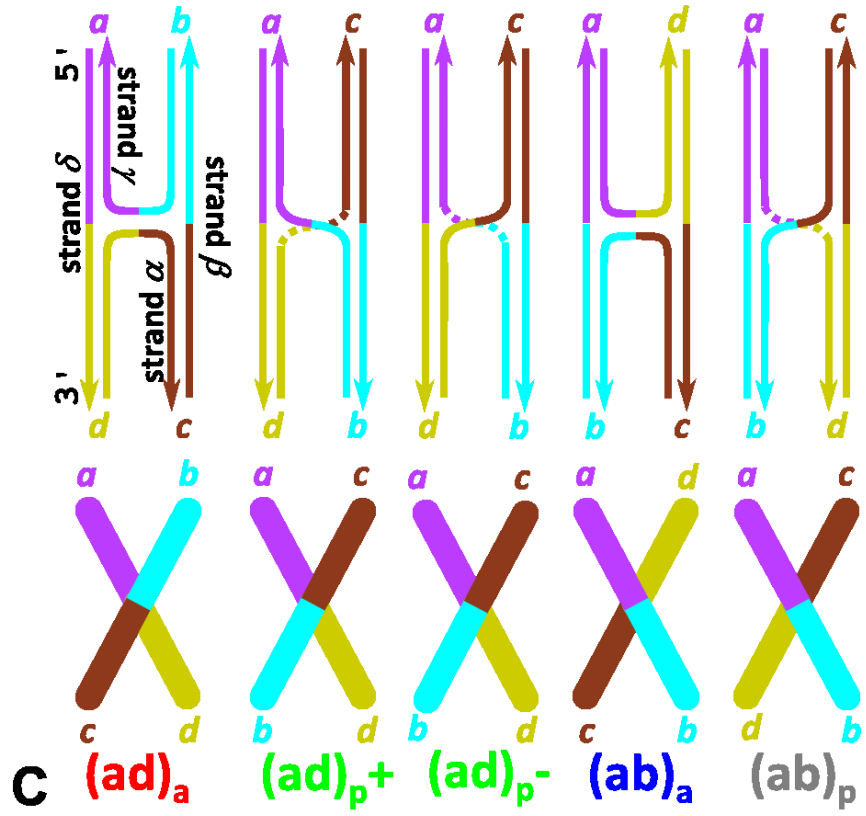
In this work a fast hardware correlator with features and performance similar to current commercial devices was presented. In order to achieve significantly better (sub-ns) time resolution high-end electronics with TCSPC capabilities should be used (for example Becker&Hickl SPC series or DPC-230; PicoQuant PicoHarp or HydraHarp modules). Due to low costs of the SP605 board and of the compact FCS setup described in section 3.2.2, as well as the fact that it can be easily extended to more than two detection channels and several parallel correlator units, this correlator should be very useful for research, general analytic applications and education.

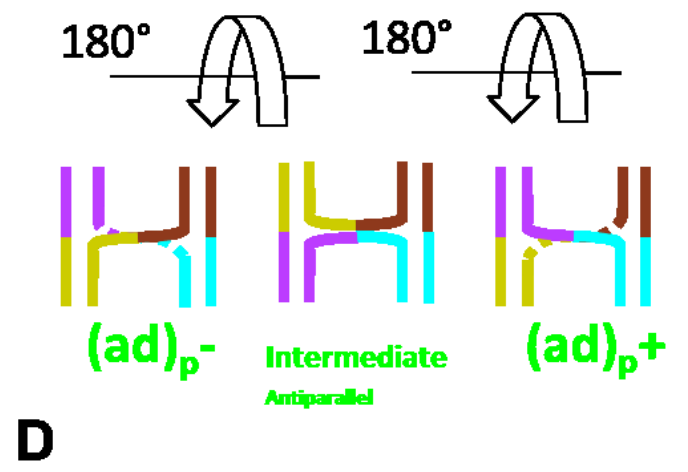
5 RNA branched molecules

5.1 Introduction

The RNA four-way junction sequence was designed to introduce perfect Watson-Crick base pairing throughout the molecule, and is related to the hairpin ribozyme [64] (PDB-ID: 1M5K, Figure 5-1A,B), which allows to rationalize the fundamentals of ribozyme architecture. The RNA three-way junction molecule resulted from the omitting helix *d* from the four-way junction molecule. For J(abc(C2)) and J(abc(C5)) two and five unpaired nucleotides were added in junction region, respectively (Figure 5-3).


helix a
helix b
helix c
helix d
A
strand δ
strand γ
strand α
strand β
B


strand δ
strand γ
strand α
strand β
5'
3'
(ad)a
(ad)p+
(ad)p-
(ab)a
(ab)p
C


180°
180°
(ad)p-
Intermediate
Antiparallel
(ad)p+
D

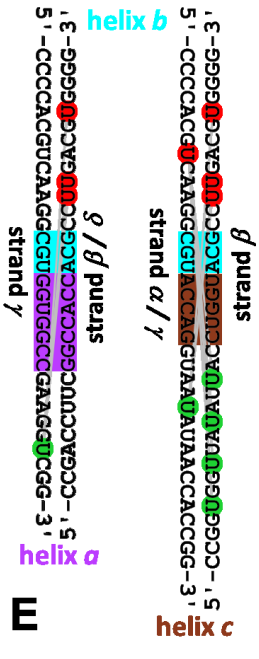

helix a
helix b
helix c
strand γ
strand β/δ
strand α/γ
strand β
E

Figure 5-1 Secondary structures of (A) the hairpin ribozyme (1M5K [64]) and (B) the RNA four-way juntion used in this work. Regions with matching sequence are highlighted. Interactions between bases of different helices are indicated in orange. Donor (D) and acceptor (A) labeling positions are shown in green and red, respectively. They are named according to dye type, single strand (α, β, γ or δ), and number of the base starting from the 5'-end and the junction arm, e.g. (A)δ28d (Figure 5-3). (C) Four possible conformers with stacked helix pairs of the J(abcd). (D) Detailed description of the difference between $(ad)_{p-}$ and $(ad)_{p+}$ conformers. (E) Double-stranded RNA with primary structures corresponding to the ones of helices *a* and *b* (left) or *b* and *c* (right) of the J(abcd) to verify the A-RNA structure of the arms.

In current project FRET positioning and screening toolkit (FPS) was applied, which combines experimental and computational techniques in a hybrid approach, to derive all-atom structural models with high accuracy and precision as well as to assess the models' confidence levels [9].

Different positions on the studied molecules were chosen for internal labeling for FRET with Alexa488 as a donor (D) and Cy5 as an acceptor (A) dyes. In total, 51 combinations of DA pairs for J(abcd), 44 for J(abc), 39 for J(abc(C2)) and 39 for J(abc(C2)) molecules were sequentially measured and analyzed (see section 5.2.3).

Additionally, 10 distances were measured along two double-stranded (ds) RNA molecules with sequences corresponding to two of the different arms of the RNA four-way junction, respectively (Figure 5-1E). This control experiment demonstrated that an A-RNA structure can be assumed for the helices of the studied molecules (section 5.5.2).

5.2 Materials

5.2.1 Oligonucleotides

Unlabeled and ultrapure labeled with hexamethylen linkers (PAGE Grade) RNA oligonucleotides were purchased from Purimex (Grebestein, Germany).

See Figure 5-2 for chemical structures of modified U and dG nucleotides with linker and dye and Figure 5-3 for the sequences and labeling positions of the RNA strands.

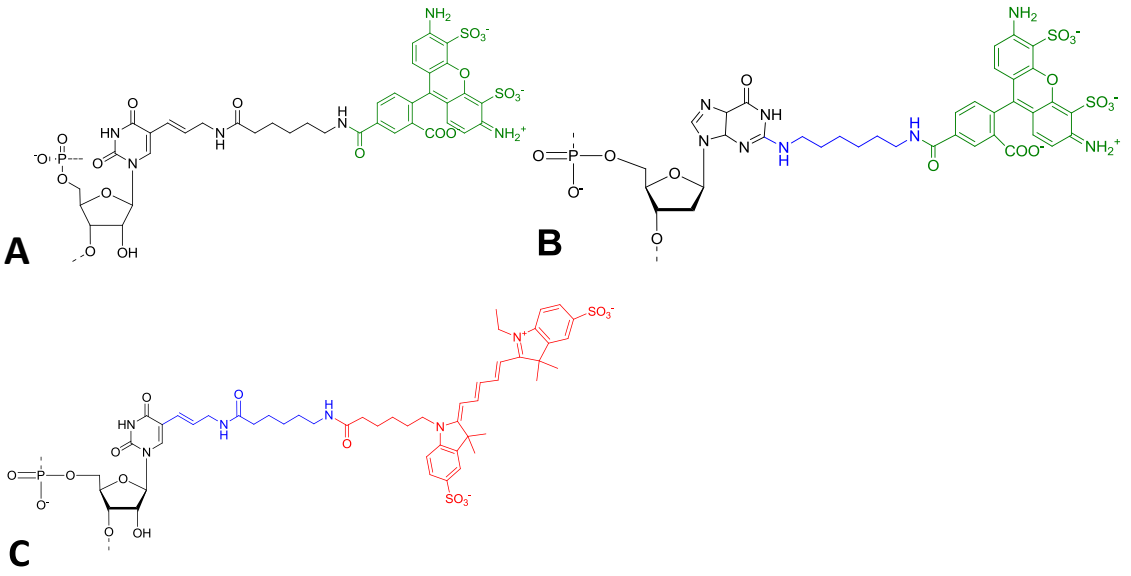


Figure 5-2 Structures of modified U (A,C) and dG (B) nucleotide residues with hexamethylen (C6) linkers (blue) and fluorescent dye Alexa488 (green) (A,B) and Cy5 (red) (C).

5.2.2 DNA sequences

Sequences and labeling positions (green for Alexa488 and red for Cy5) for the dsDNA used for calibration of the detection efficiency ratio (see section 5.4.5).

5'-d(GCA ATA CTT GGA CTA GTC TAG GCG AAC GTT TAA GGC GAT CTC TGT TTA CAA CTC CGA AAT AGG CCG)-3'

5'-d(CGG CCT ATT TCG GAG TTG TAA ACA GAG ATC GCC TTA AAC GTT CGC CTA GAC TAG TCC AAG TAT TGC)-3'

5.2.3 **Nomenclature**

Descriptive nomenclature of four- and three-way junction molecules and labeling positions was chosen to make the comparison of these molecules possible. Junction was designated by the capital letter J. Helices of the particular molecule were listed one after another starting with helix *a*: *abcd* for four-way junction and *abc* for three-way junction molecules. In case of 2- and 5-nucleotides bulges in the junction region, "C2" and "C5" were added to the name of the molecule resulting in J(abc(C2)) and J(abc(C5)), respectively. Here C stands for cytosine, 2/5 is the number of the unpaired cytosine in the bulge and the position of C2/C5 means that the bulge is located between helices *c* and *a.* Hence, J(abcd), J(abc), J(abc(C2)) and J(abc(C5)) describe RNA four-way junction, RNA three-way junction, RNA three-way junction with 2C bulge and RNA three-way junction with 5C bulge, respectively (Figure 5-3).
The following names were given to the single strands: α, β, γ, δ, $\delta\alpha-\Delta d$, $\delta\alpha$(C2)$-\Delta d$ or $\delta\alpha$(C5)$-\Delta d$. In case of three-way junctions, the common strand of helices *a* and *c* consisted of the first half of the strand δ and the second part of the strand α. Therefore the name of this common strand is $\delta\alpha-\Delta d$, where Δd means that helix *d* is omitted. For J(abc(C2)) and J(abc(C5)) the same strand would be $\delta\alpha$(C2)$-\Delta d$ or $\delta\alpha$(C5)$-\Delta d$, respectively. Labeling positions were named according to dye type (D, A), single strand, number of base starting from the 5'-end and the junction arm, e.g. (A)$\delta\alpha$(C2)$-\Delta$d30c. In case of (D)γ29a and (A)β33b labeling positions, the names of the samples of each molecule will be: (D)γ29a/(A)β33b, (D)γ29a/(A)β33b/$\delta\alpha-\Delta d$, (D)γ29a/(A)β33b/$\delta\alpha$(C2)$-\Delta d$ and (D)γ29a/(A)β33b/$\delta\alpha$(C5)$-\Delta d$ for RNA four-way junction, RNA three-way junction, RNA three-way junction with 2C bulge and RNA three-way junction with 5C bulge, respectively (Figure 5-3).

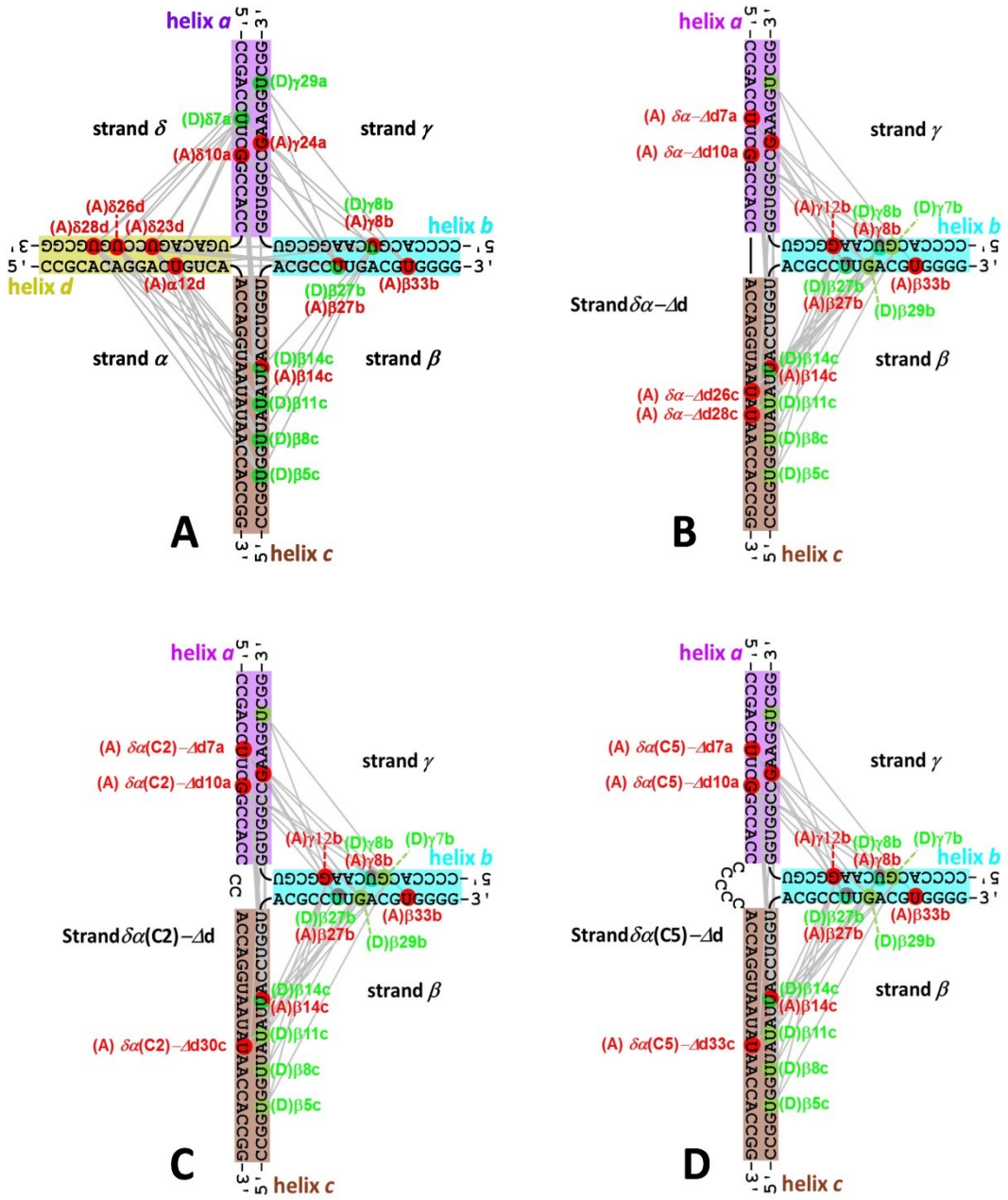


Figure 5-3 Secondary structures of RNA molecules used in this work: (A) RNA four-way junction J(abcd), (B) RNA three-way junction J(abc), (C) RNA three-way junction with additional 2C bulge in the junction region J(abc(C2)) and (D) RNA three-way junction with a 5C bulge in the junction region J(abc(C5)). Identical helices are highlighted with the same color in all molecules. Alexa488 and Cy5 were used as donor (shown as green spheres) and acceptor (shown as red circles) dyes, respectively (see chemical structures of fluorophores in Figure 5-2). Resulting DA pairs are indicated with gray lines.

5.3 Procedures

5.3.1 Hybridization procedure

Hybridization of molecules was done by mixing single donor, acceptor labeled and two unlabeled strands (for four-way junction) and one unlabeled strand (for three-way junctions) with the ratio of 1:3:4:4 and 1:3:4 respectively. The concentration of Donor labeled (acceptor labeled) RNA single strands in solution was 0.5-2μM. For Donor or Acceptor only molecules (DO and AO) unlabeled strands were taken four times more than labeled ones: the ratio 1:4:4:4. For incomplete molecules donor and acceptor labelled strands with the ratio 1:3 and no unlabeled strands were used. Hybridization buffer contained 20 mM KH_2PO_4/K_2HPO_4, 100 mM KCl, and 20 mM $MgCl_2$, pH 6.5. The solution was heated up until 85°C inside a thermo-cycler (primus 96 advanced, peqLab, Erlangen, Germany) with 0.1°C/s, and then cooled down up to 25°C with 2°C/h. Afterwards it was quickly cooled down until 4°C.

5.3.2 Measurement buffer for fluorescence measurements

The measurement buffer contained 20 mM KH_2PO_4/K_2HPO_4, 100 mM KCl and 20 mM $MgCl_2$, pH 6.5. Additionally, approximately 0.5 mM of Trolox [65] was added to decrease the bleaching of Cy5 in the measurements.

5.3.3 Sample preparation for SAXS-measurements

For the preparation of the four-way junction a ratio of the four single strands of 1:1:1:1 was used. Four RNA strands were dissolved in hybridization buffer (20 mM KH_2PO_4/K_2HPO_4, 100 mM KCl, 20 mM $MgCl_2$, pH 6.5) to a final concentration of 20 μM. The sample was heated up to a temperature of 90°C for 15 minutes in a thermo-cycler (Biometra T3000) with a rate of

0.1°C/s. The sample was cooled down to 28°C with a rate of 2°C/h., followed by immediate cooling to 4°C.

For purification, the sample was subjected onto 8% PAA-gel. The gel was run at 80 V for 1.5 h at 9°C in TB-buffer (90 mM Tris/Borate, 1mM $MgCl_2$, pH 8.3). Elution of the desired product was carried out according to the protocol of Lipfert et al. [66]. Briefly, desired bands were cut out and 5 ml TEN-buffer (10 mM Tris, 1 mM EDTA, 100 mM NaCl, pH 8) was added. Samples were frozen and warmed up to room temperature (3 cycles) followed by shaking over night at 4°C. After centrifugation and filtration of the supernatant, RNA was precipitated from ethanol.

5.4 Methods

5.4.1 Calculation of dye distributions via the AV approach

Dye distributions were modeled using the accessible volume (AV) approach [67, 68] according to the methodology described in [69] and in [9]. In brief, dyes are approximated by a sphere with an empirical radius of R_{dye}, where the central atom of the fluorophore is connected by a flexible linkage of a certain effective length L_{link} and width w_{link} to the nucleobase. Overall length of the linkage is given by the actual length of the linker and the internal chemical structure of the dye. A geometric search algorithm finds all dye positions within the linkage length from the attachment point, which do not cause steric clashes with the macromolecular surface. All allowed positions are considered as equally probable, which allows one to define an accessible volume for the dye (AV). To take the three quite different dimensions of a fluorophore into account, the real physical dimensions for each calculation of position distribution is used and three independent AV simulations with three different radii $R_{dye(i)}$ are performed and superimposed (Figure 5-4). Thus, the obtained position distribution represents an average weighted by the number of allowed positions. The attachment atoms are C5s and C2s for labeled Us and dGs, respectively (see Figure 5-2).

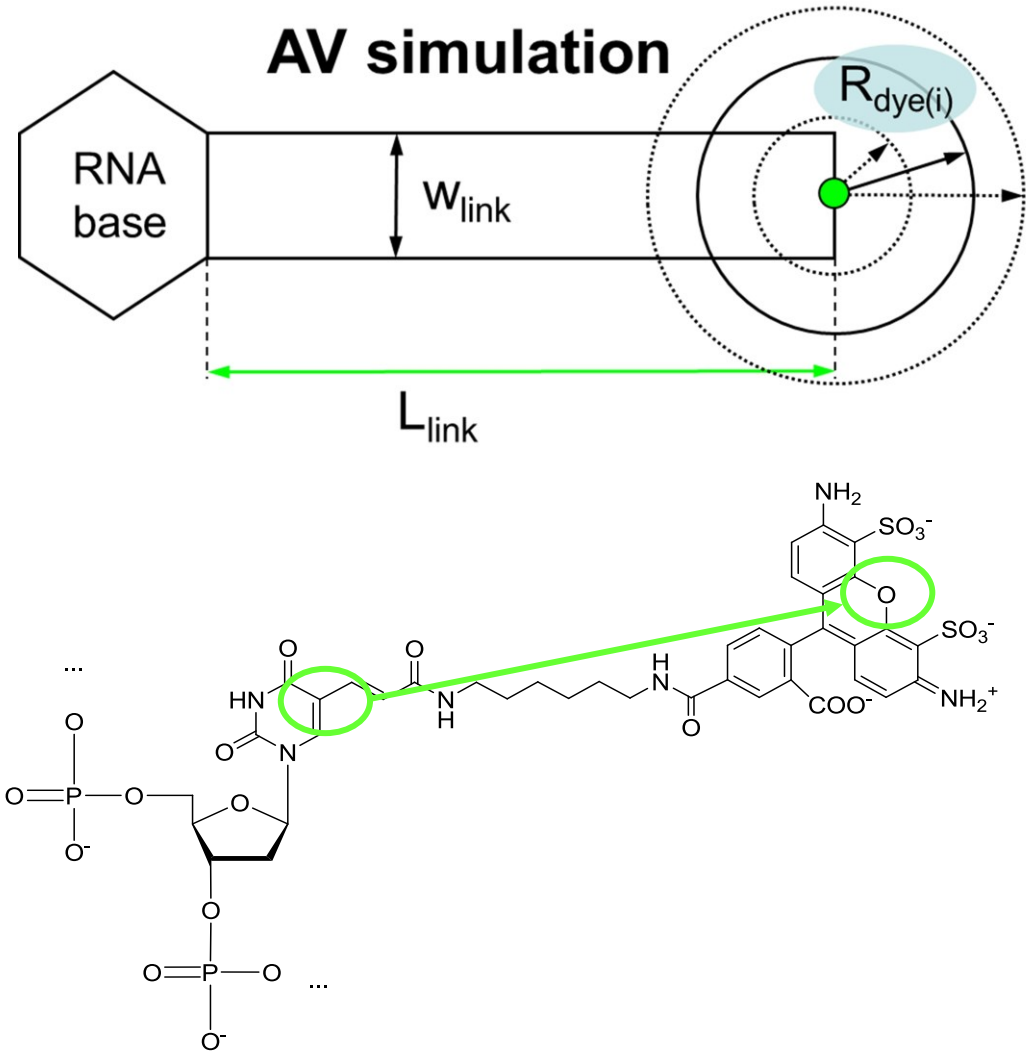


Figure 5-4 Schematic comparison of AV and MD simulations of dye positions. $R_{dye(i)}$, L_{link} and w_{link} are shown as arrows. Following parameters were used for AV simulations: $R_{dye(1)}$ = 5 Å, $R_{dye(2)}$ = 4.5 Å and $R_{dye(3)}$ = 1.5 Å for Alexa488 (D) and $R_{dye(1)}$ = 11 Å, $R_{dye(2)}$ = 3 Å and $R_{dye(3)}$ = 1.5 Å for Cy5 (A); L_{link} = 20 Å for Alexa488 and L_{link} = 22 Å for Cy5; w_{link} = 4.5 Å.

5.4.2 AVs for U and dG labeling positions

In Figure 5-5 accessible volumes calculated for dyes attached to modified U and dG bases are compared. In the case of U and dG labeling, the distances of mean dye position from the helical axis are ~ 10 Å and ~ 17 Å, respectively. Combined use of these two labeling approaches helps to unambiguously distinguish between "competing" structures, as without FRET distances involving dG labeling positions, more than one structure with good agreement with FRET data were obtained.

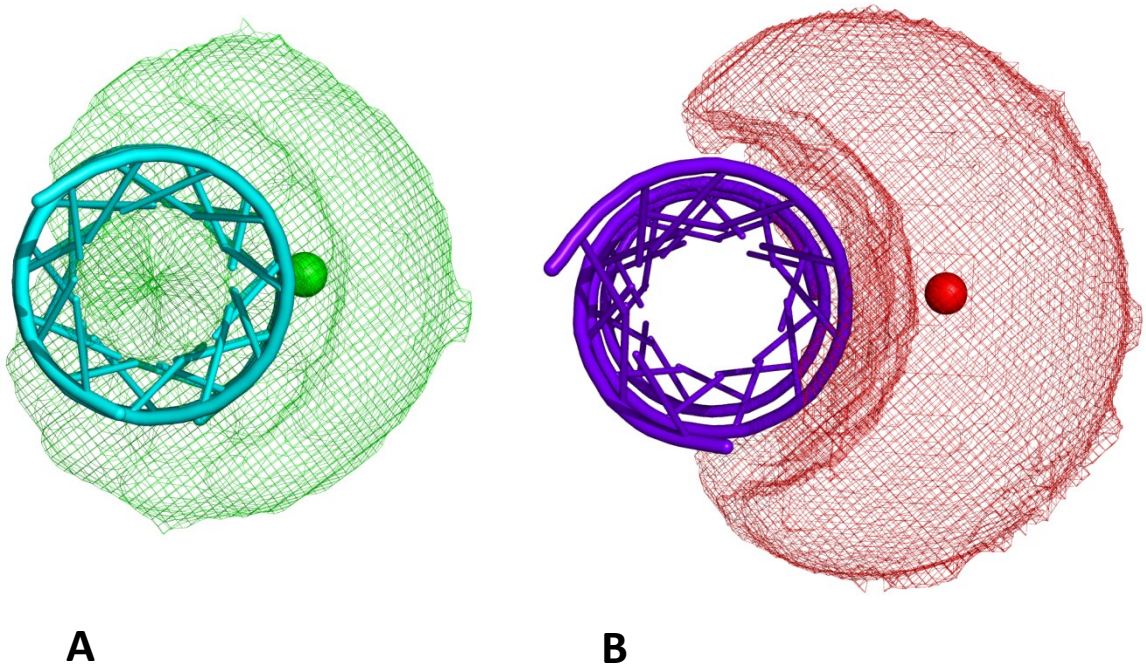


Figure 5-5 Accessible volumes of Alexa488 (A, green) and Cy5 (B, red) attached at positions (D)β8c (A, modified U) and (A)δα–ΔD10a (B, modified G), respectively. Mean positions of the dyes are displayed as green and red spheres. Following parameters were used for the AV simulation: Alexa488: L_{linker} = 20 Å; w_{linker} = 4.5 Å; $R_{dye(1)}$ = 5 Å; $R_{dye(2)}$ = 4.5 Å; $R_{dye(3)}$ = 1.5 Å; Cy5: L_{linker} = 22 Å; w_{linker} = 4.5 Å; $R_{dye(1)}$ = 11 Å; $R_{dye(2)}$ = 3 Å; $R_{dye(3)}$ = 1.5 Å.

5.4.3 Distance determination via Photon Distribution Analysis (PDA)

Multi-parameter fluorescence detection (smMFD) simultaneously acquires all fluorescence parameters [56], which allows for calculating the FRET efficiency based on both fluorescence intensity and fluorescence lifetime. DA distances (R_{DA}) from the fluorescence intensities of D and A (F_D and F_A, respectively) are calculated via,

$$R_{DA} = \left(\Phi_{FA} \frac{F_D}{F_A} \right)^{\frac{1}{6}} R_{0r} \qquad \textbf{Eq. 5.4-1}$$

where Φ_{FA} is the fluorescence quantum yield of A (Table 3). In Eq. 5.4-1 the reduced Förster radius R_{0r} [70] is used, which, in contrast to the Förster radius R_0 [56], does not depend on the quantum yield of D in absence of FRET $\Phi_{FD(0)}$. Throughout this work R_{0r} = 53.97 Å is used (using R_{0r} = 53.97 Å is equivalent to using R_0 = 52 Å and $\Phi_{FD(0)}$ = 0.8). F_D and F_A can be calculated from the signals measured in the green and red detection channels S_G and S_R, respectively, via Eq. 5.4-2 and Eq. 5.4-3,

$$F_D = \frac{F_G}{g_G} = \frac{S_G - \langle B_G \rangle}{g_G} \qquad \textbf{Eq. 5.4-2}$$

$$F_A = \frac{F_R}{g_R} = \frac{S_R - \alpha F_G - \langle B_R \rangle}{g_R} \qquad \textbf{Eq. 5.4-3}$$

where F_G and F_R are the fluorescence signals in the green and the red signal channels, respectively, α is the crosstalk factor which is determined as the ratio between donor photons detected in the red channels and those detected in the green channels ($\alpha = F_{R(D)} / F_{G(D)}$) for the D only labeled sample, g_G and g_R are the detection efficiencies in the green and red channels, respectively (see section 5.4.5 for the determination of g_G/g_R), and $\langle B_G \rangle$ and $\langle B_R \rangle$ are the mean background intensities in the green and red channels, respectively.

To accurately predict the shape of S_G/S_R (or equivalently F_D/F_A) histograms in the presence of FRET PDA is used, which explicitly takes into account shot noise, background contributions and additional broadening due to complex acceptor photophysics [46, 71-73]. Additionally, it provides a meaningful reduced chi-squared value (χ_r^2) directly derived from photon statistics. PDA calculates the probability of observing a certain combination of photon counts $P(S_G, S_R)$

$$P(S_G,S_R) = \sum_{F_G+B_G=S_G;F_R+B_R=S_R} P(F)P(F_G,F_R|F)P(B_G)P(B_R) \qquad \textbf{Eq. 5.4-4}$$

The intensity distribution of the fluorescence only contribution to the signal, $P(F)$, is obtained from the total measured signal intensity distribution $P(S)$ by deconvolution assuming that the background signals B_G and B_R obey Poisson distributions, $P(B_G)$ and $P(B_R)$, with known mean intensities $\langle B_G \rangle$ and $\langle B_R \rangle$. $P(F_G, F_R \mid F)$ represents the conditional probability of observing a particular combination of green and red fluorescence photons, F_G and F_R, provided the total number of registered fluorescence photons is F, and can be expressed as a binomial distribution [73]. Subsequently, $P(S_G, S_R)$ may be further manipulated to generate a theoretical histogram of any FRET-related parameter as discussed elsewhere [72].

In this work a model accounting for up to three FRET states (Gaussian distributed distances) and a D-only contribution was used. Additional broadening of FRET states was accounted for by a global parameter σ_{app} as justified in [46]. In some cases an impurity (mostly 1 – 3%) with an apparent distance of typically 70 – 90 Å had to be taken into account. This state was in most of cases present in respective D-only samples with amplitude of a few percent. Thus, for n FRET states $2n + 1$ to $2n + 3$ fit parameters were required depending on whether the impurity state was considered. The fit quality was judged by the reduced chi-squared (χ^2_r) parameter and by visually inspecting weighted residuals plots.

5.4.4 Static FRET line and distribution of possible σ_{DA}-values

The static FRET line represents the expected dependence between FRET indicators derived from intensities (e.g. F_D/F_A) and the fluorescence lifetime of the donor. In the most simple case it is given by the well-known equation $E = 1-\tau_{DA}/\tau_{D(0)}$. In reality this relationship does not hold because the distributions of donor-acceptor distances due to flexible dye linkers are not accounted for. In addition, non-exponential fluorescence decay of the donor dye itself must be considered (see Table 3 for the fluorescence decay fits for all D positions). These effects can be corrected for as described in [9, 74]. As there is no analytical expression for the $E(\tau_{DA})$ dependence that considers the above effects, a polynomial approximation is used. In this work the following approximation was used,

$$F_D/F_A = \frac{\Phi_{FD(0)}}{\Phi_{FA}} \Bigg/ \left(\frac{\left\langle \tau_{D(0)} \right\rangle_x}{c_3\left\langle \tau_{D(A)} \right\rangle_f^3 + c_2\left\langle \tau_{D(A)} \right\rangle_f^2 + c_1\left\langle \tau_{D(A)} \right\rangle_f + c_0} - 1 \right) \qquad \textbf{Eq. 5.4-5}$$

where $\langle \tau_i \rangle_x$ and $\langle \tau_i \rangle_f$ are species and fluorescence averaged mean lifetimes, respectively. The polynomial coefficients $c_{(i)}$ are calculated assuming 6 Å half-width of the DA distance distribution (σ_{DA}) [69] (see Table 4 for the resulting coefficients). For highly asymmetric AVs (see section 5.4.2) σ_{DA} depends also on the mutual orientation of D and A clouds, which implies that individual σ_{DA}-values should be used for different samples, labeling positions, and/or even FRET states. Considering various possible orientations of calculated dyes' AVs, it was estimated that σ_{DA} can vary between ca. 5.5 and 12 Å. To fit static FRET lines to the observed FRET states, values of σ_{DA} between 6 and 9 Å were required, which is within the expected range.

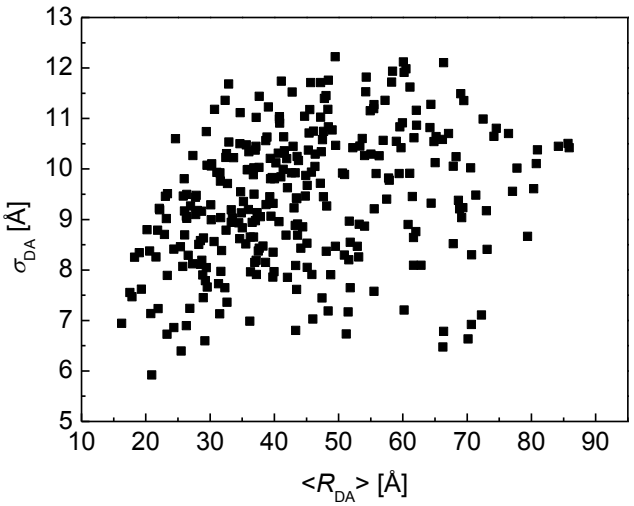


Figure 5-6 Possible values of σ_{DA} calculated for randomly positioned and oriented AVs of Alexa488 and Cy5 initially calculated for positions (D)β8c and (A)δ23d, respectively.

5.4.5 Determination of detection efficiency ratio g_G/g_R

To be able to convert model distances into probabilities of observing green photons, the detection efficiency ratio g_G/g_R is needed (g_G and g_R stand for the detection efficiencies of "green" and "red" channels, respectively). These values are calculated for each measurement session by requiring that the linker-corrected static FRET line [74] (see Eq. 5.4-5) goes through the observed FRET population in a 2D histogram of F_D/F_A vs $\tau_{D(A)}$ for a measurement of a FRET-labeled dsDNA (see section 5.2.2). For the FRET line it is assumed that σ_{DA} = 6 Å, $\Phi_{FD(0)}$ = 0.8 and $\Phi_{D(0)}$ = 4.1 ns (mono-exponential decay).

5.4.6 Confidence intervals for fit parameters in PDA

To estimate the errors of fitted parameters due to photon statistics, ($\Delta R_{DA}(E)$), the parameter space for sets of variables providing acceptable fits is explored. All free fit parameters are varied simultaneously in a random manner. The χ^2_r – values are calculated at 100000 random points yielding 1000 - 60000 points with χ^2_r – values below $\chi^2_{r,min}$ + $(2/N_{bins})^{1/2}$ (here N_{bins} is the number of histogram bins, and $\chi^2_{r,min}$ is the reduced chi-squared

of the best fit) (Figure 5-7 red points). The range where such fits are possible is assigned as 1σ confidence interval. While one could calculate χ^2_r thresholds more strictly from the χ^2 distribution [75], in practice $\chi^2_{r,min}$ is often affected by experimental imperfections and can be considerably larger than one. For this reason, the simple test mentioned above is preferred which relates χ^2_r values to that of the best fit. The overall error ΔR_{DA}, resulting from photon statistics ($\Delta R_{DA}(E)$) and from the uncertainty of the mutual orientation of the donor and acceptor dyes (κ^2 errors, $\Delta R_{DA}(\kappa^2)$), is then calculated according to error propagation rules (Eq. 5.4-4):

$$\Delta R_{DA}^2 = \Delta R_{DA}^2(E) + \Delta R_{DA}^2(\kappa^2) \qquad \textbf{Eq. 5.4-6}$$

For the major peaks $\Delta R_{DA}(E)$ contributes only weakly to the overall error (between 0.5 and 2 %). For the minor peaks $\Delta R_{DA}(E)$ can be considerably higher, especially if they are overlapped by other peaks (mostly between 3 and 10 %, up to ~ 20 % for a few cases (see sections 5.5.4.2 and 5.5.4.3 for J(abcd) and sections 5.5.5.2 and 5.5.5.3 for three-way junction molecules).

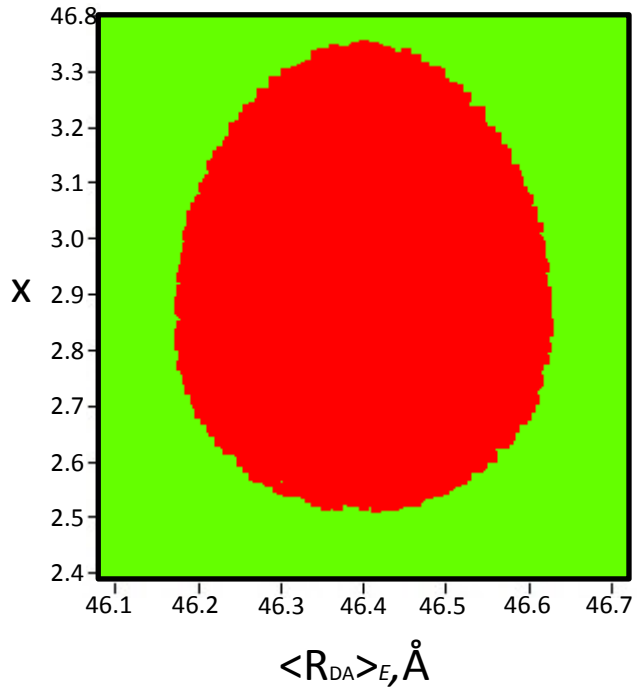


Figure 5-7 A typical example of the estimation procedure of the errors of fitting parameters for molecule (A)δα_Δd10a/(D)β27b/γ. Red dots correspond to the structures

with the χ^2_r values smaller than $\chi^2_{r,min} + (2/N_{bins})^{1/2}$. $<R_{DA}>_E$ and x are measured distance and the relative amplitude of the major state.

5.4.7 Rigid body docking

Rigid body docking was performed as described in [9]. For four-way junction the helices were kept as whole, and were connected to each other only with covalent bonds, which also exist in the natural structures. For three-way junctions, however, more sophisticated model was used in order to provide additional flexibility in the junction region: two nucleotides on each helix close to the junction were cut and the rest of helices were assumed to be rigid bodies (see section 5.5.5.6 for more details). For J(abc) and J(abc(C2)) molecules cut nucleotides were modeled as separate bodies. For J(abc(C5)) two cut nucleotides on each strand were modeled as one body. Nucleotides in bulges for J(abc(C2)) and J(abc(C5)) were modeled as separate bodies. All separate bodies were connected with neighboring bodies and rigid bodies with covalent, hydrogen and stabilizing artificial bonds as is shown in Figure **5-8** and in Figure 5-30

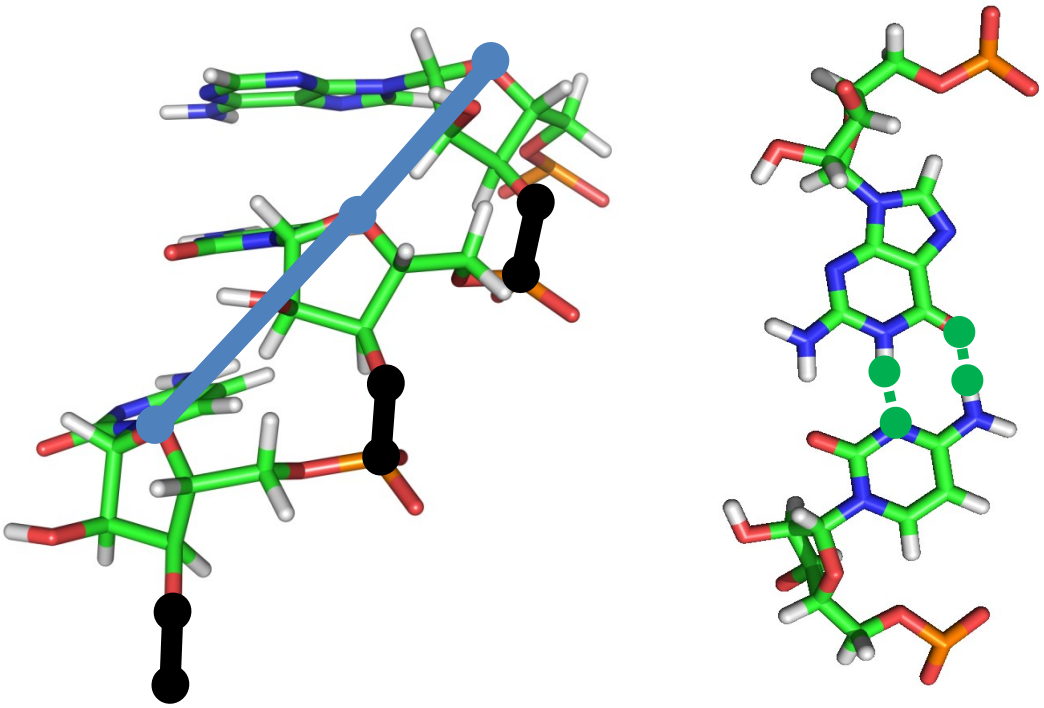


Figure 5-8 Atoms used for covalent, hydrogen and artificial bonds. Covalent bond connects 2' oxygen atom of the sugar of one nucleotide with phosphorus atom of the phosphate group of the next nucleotide in the same strand (blue). Hydrogen bonds were connecting O6 and H1 atoms on Guanine with H4 and N3 atoms on cytosine, respectively, and H3 and O4 atoms on Uracil and N1 and H6 atoms on Adenine, respectively (green). Artificial bonds connect two sugar oxygen atoms in two consequent nucleotides (black).

We explicitly model the accessible volumes (AVs, see section 5.4.1) of fluorophores [9]. This makes it possible to estimate mean dye positions with respect to RNA and average over distributions of DA distances (for further details see sections 5.4.9 and [9]). Mean dye positions were rigidly fixed to the labeled RNA helix and connected with "springs" with relaxed lengths given by the corresponding values of R_{mp} (distance between the mean positions of the dyes). The strengths of the "springs" were derived from experimental errors ΔR_{DA} (see section 5.4.6). Examples of input files for FPS for J(abc) molecule are shown in Table 35 and Table 36. Viscosity, clash tolerance and other parameters used in FPS are presented in Table 37. The weighted data-model deviation for a set of distances, given by Eq. 5.4-7 was then relaxed to the nearest minimum of its "energy", which corresponds to a local χ^2_r minimum.

$$\chi_E^2 = \sum_{i=1}^{n} \frac{(R_{\mathrm{DA}(i)} - R_{\mathrm{model}(i)})^2}{\Delta R_{\mathrm{DA}(i)}^2} \qquad \textbf{Eq. 5.4-7}$$

The procedure was repeated many times for random initial conditions, which ensures exhaustive sampling of the configuration space. In a second run ("refinement"), the AVs are re-modeled for all solutions found in the first run. This takes into account possible steric clashes of fluorophores with molecule's arms they are *not* attached to. All structures are then optimized using the new AVs. The solutions were sorted by χ^2_r and clustered as shown in the sections 5.5.4.7 and 5.5.5.6.1.2 for four-and three-way junctions, respectively.

5.4.8 Metropolis Monte Carlo sampling

Metropolis Monte Carlo (MMC) [76] sampling was performed to explore the allowed conformational space.

In a nutshell: the MMC method is used to generate a set of N configurations of the system $\xi_1, \xi_2, \xi_3 \ldots \xi_N$ in such way that

$$\lim_{n\to\infty} \frac{N_\xi}{N} = \mathrm{P}(\xi) \qquad \textbf{Eq. 5.4-8}$$

Here P(ξ) is the probability Boltzmann distribution and N_ξ is the number of configurations ξ.
MMC consists of 3 steps:

First step: Pick a configuration ξ_n.
The initial structure in my case is the structure generated by RBD approach (section 5.4.7).

Second step: Pick a trial configuration ξ_m (a configuration similar to ξ_n) and compute the probability ratio R = P(ξ_m)/ P(ξ_n). Do a check for a random number t = [0:1] as following: ξ_{n+1} = ξ_m if $t \leq$ R, otherwise ξ_{n+1} = ξ_n (see Eq. 5.4-9 for the acceptance critera).

$$\exp\left[-(E_{new} - E_{old})/kT\right] \geq random(0,1) \qquad \textbf{Eq. 5.4-9}$$

Here E_{new} and E_{old} are the energies of the new and old configurations, respectively. kT is Boltzmann's constant (k) multiplied by temperature (T), and random(0,1) is a random number between 0 and 1.
In my case χ^2 of the structures were used as E in Eq. 5.4-9. Trial structure was generated by rotating a random helix for 0.1° around random axis and translating it with the 0.05 Å step. Then, the overall χ^2 of the new structure was calculated and was compared with the χ^2 of the previous structure. If the new structure was accepted, it was considered as the next step in sampling, otherwise the previous structure was taken for the next step

Third step: Do the step 2, replacing ξ_{n+1} by ξ_n.

Generated structures were saved at given intervals for further analysis. Step 3 is repeated N times, where N is sufficiently large number (~ $15*10^6$ in my case). In this work MMC was performed at different "temperatures" (kT = 0.01, 0.1, 1, 10) to create the landscape of the enthalpy of the generated structures (section 5.5.5.11).

5.4.9 R_{mp} - $\langle R_{DA}\rangle_E$ conversion function

Although the optimization problem (Eq. 5.4-5) can be defined for $\langle R_{DA}\rangle_E$ values, this would require re-calculation of the dyes' AVs after each iteration, which is a time-consuming procedure. To avoid this, a conversion function between R_{mp} and $\langle R_{DA}\rangle_E$. was generated. This was done by generating a large number of random orientations and positions of dyes' AVs. For each pair of dye clouds, R_{mp} and $\langle R_{DA}\rangle_E$ were calculated. The resulting $R_{mp}(\langle R_{DA}\rangle_E)$ dependence was approximated with a 3rd order polynomial and used as a conversion function. The RMS deviation between the polynomial approximation and the $R_{mp}(\langle R_{DA}\rangle_E)$ dependence was typically 0.5 – 0.7 Å.

5.4.10 Model discrimination of docking results

Model discrimination is done according to [9]. Solutions are considered ambiguous if the respective χ_r^2 values do not differ significantly. A threshold given by Eq. 5.4-10 is typically applied,

$$\chi_r^2 < \chi_{r,\min}^2 + \left[2/(n-p)\right]^{1/2} = \chi_{r,\max}^2 \qquad \textbf{Eq. 5.4-10}$$

where n is the number of distance restraints (51, 44, 39, 39 for J(abcd), J(abc), J(abc(C2), J(abc(C5), respectively; the restraints between the helix ends at the junction are not considered here as they cause a reduced number of degrees of freedom) and p is the number of degrees of freedom (9 and 6 for four- and three- way junctions, respectively. Translational degrees of freedom are not taken into account due to the strong bonds between helices; rotation of two helices against the third one gives 3 × 3 (for four-way junction) and 2 × 3 (for three-way junction) rotational degrees of freedom. Eq. 5.4-10 roughly corresponds to the variance of the chi-squared distribution of 2 × (degrees of freedom) [75] (magenta lines in Figure 5-14, Figure 5-35, Figure 5-36 and Figure 5-37). The fact that $\chi_{r,\min}^2$ is often larger than one is attributed to systematic experimental errors and

to possible violations of the AV and/or rigid body models. Other criteria defining different levels of significance can be applied here in a straightforward way.

5.4.11 **Precision estimation for rigid body models**

The precision of rigid body models was estimated by bootstrapping [77] as described in [9]. Briefly, all model distances found for the optimized structures were simultaneously perturbed by adding normally distributed random numbers with a mean of zero and the standard deviation given by the experimental errors $\{\Delta R_{DA}\}$. Afterwards the structures were re-optimized using the perturbed distances. This procedure was repeated 100 times yielding a set of structures representing the distribution of possible positions of the helices. For this set, RMSD values were calculated for each phosphorous atom with respect to the original model. The average RMSD over all P atoms of a structure is used to characterize the overall precision of the rigid body model.

5.4.12 **Molecular Dynamics simulations**

Molecular dynamics (MD) simulations were performed by the group of Prof. Gohlke from HHU in Duesseldorf. The Amber 11 suite of programs were used[78], together with the force field as described by Hornak et al. [79], using modifications suggested by Pérez et al. [80], Banás et al. [81], and Joung et al. [82].

The starting structure for each conformer, which was obtained from rigid body docking, was placed in an octahedral periodic box of TIP3P water molecules [83]. The distance between the edges of the water box and the closest atom of the RNA was at least 11 Å, resulting in a system of ~ 250,000 atoms. Mg^{2+} and K^{+} ions were added to reproduce a concentration of 20 mM and 100 mM, respectively. Cl^{-} ions were added to neutralize the simulation system.

The system was minimized by 50 steps of steepest descent minimization followed by 450 steps of conjugate gradient minimization. The particle mesh Ewald method [84] was used to treat long range electrostatic interactions, and bond lengths involving bonds to hydrogen atoms were constrained using SHAKE [85]. The time step for all MD simulations was 2 fs,

with a direct-space nonbonded cutoff of 8 Å. Applying harmonic restraints with force constants of 5 kcal mol^{-1} $Å^{-2}$ to all solute atoms, canonical ensemble (NVT)-MD for 50 ps was carried out, during which the system was heated from 100 to 300 K. Subsequent isothermal isobaric ensemble (NPT)-MD was used for 150 ps to adjust the solvent density. Finally, the force constants of the harmonic restraints on solute atom positions were gradually reduced to 1 kcal mol^{-1} $Å^{-2}$ during 50 ps of NVT-MD.

From the following 40 ns of NVT-MD at 300 K, conformations were extracted for every 20 ps. In case of RNA four-way junction harmonic restraints were applied on phosphorus atoms more than six bases away from the junction region, using force constants that have been chosen such that the positional uncertainties of the phosphorous atoms as calculated by bootstrapping for the rigid body docking models are reproduced.

5.4.12.1 **Clustering of MD simulations for J(abcd)**

Clustering of the MD trajectories was performed with the hierarchical clustering method using the average linkage algorithm as implemented in the AmberTools 13 package [86]. The first 1000 snapshots of each trajectory, corresponding to 20 ns of simulation time, were omitted in the cluster analysis in order to account for equilibration of the MD simulation. Clusters were formed according to pair-wise RMSD of the phosphorus atoms of the RNA, calculated between the snapshots of the trajectory. From the resulting 100 clusters, cluster representatives, which are the conformations that have the smallest RMSD to all other conformations in their cluster, were chosen for the further analyses.

5.4.13 **SimRNA**

SimRNA is a de novo RNA folding method. Its uses a coarse-grained representation of RNA, samples the conformational space using the Monte Carlo approach and evaluates the energy of conformations using a statistical potential derived from analysis of experimentally solved RNA structures. A detailed description of the early version of this method is presented in [87]. The current version uses five pseudo-atoms per nucleotide residue (P, C4', N1, C2, and

C4 for pyrimidines and P, C4', N9, C2, C6 for purines) rather than only three (P, C4', N1/N9) as in the early version. As a result of coarse-graining, SimRNA is faster by a factor of ~1000 compared to all-atom MD simulations, and therefore can be used to fold RNA without the need of pre-formed starting structures, and to sample a very large space of possible conformations. SimRNA can simulate the folding of molecules comprising any number of RNA chains, and can use any combination of restraints on atomic positions, on interatomic distances or on the desired secondary structure. Models resulting from SimRNA simulations can be easily converted from coarse-grained to all-atom representation, which is done using a built-in tool. The resulting full-atom models can be subjected to local optimization (using. e.g. energy minimization or Molecular Dynamics) to idealize interatomic distances, angles, and steric interactions.

5.4.13.1 **Restraints for simulations with SimRNA**

Three types of restraints have been used in the SimRNA Monte-Carlo simulations: restraints on base–pairing (RNA secondary structure), on distances between individual residues according to FRET data, and on maintaining the A-helix conformation.

1) Restraints on the formation of base pairs corresponding to the predicted secondary structure were imposed using a built in mechanism of SimRNA. Additional penalty was introduced for specified pairs of nucleotide residues, unless they formed a canonical cis–Watson-Crick base pair. In order not to influence the internal dynamics of the junction, this type of restraints were only imposed on base pairs located more than three nucleotide residues away from the junction. Hence, base pairs in the region of the junction could form in the course of the simulation only due to the attractive energies in the SimRNA force field, and not owing to restraints.

2) FRET-based distance restraints were imposed on distances between P atoms in pairs of specified nucleotide residues belonging to each of the D-A pairs for which FRET data was available. A penalty was introduced when the distance between a specified pair of atoms at a given step of the simulation exceeded either the minimal or the maximal allowed value specified for that pair, and the value of the penalty was linearly proportional to the violation observed. The minimal and maximal values of distance restraints were taken from the

positions of P atoms in rigid-body models described in section 5.4.7. Values for each conformer (rounded to 3 decimal figures) are shown in Table 1.

3) Additional distance restraints were used for idealization of the backbone geometry in helices, to maintain the A-helix structure and avoid distortions resulting from e.g. “pulling” forces of FRET-based restraints as well as to decrease the “wobbling” of helices observed in high temperature simulations with restrains on base-pairing alone. The distances used were generated by creating an ideal RNA A-helix using the X3DNA program [88] and computing the distances between P atoms of residues separated by 6, 7, 8, 9, 10 positions in the RNA chain. This procedure provided 218 distance restraints for the entire structure of the RNA four-way junction molecule (see Table 2).

The above-mentioned spatial restraints were used as additional components of the SimRNA energy function. In the first strategy, corresponding to folding without experimental restraints, only restraints on base–pairing were used, with the weight of the secondary structure set to 1.0 (arbitrary internal SimRNA units). The lowest energy structures from the folding simulations underwent refinement using SimRNA and both secondary structure and FRET distance restraints.

In the second strategy, experimental FRET restraints were used in the initial folding simulations, together with those rigidifying the helices, with weights equal to 0.1. Additional refinement simulations were made with the weight of FRET restraints increased to 0.2, to 0.5, and to 1.0.

Table 1 Sets of FRET restraints used in the simulations.

Pairs of residues	distance between P-atoms [Å]					
	$(ad)_a$		$(ad)_p$		$(ab)_a$	
	minimal	maximal	minimal	maximal	minimal	maximal
α12d_β5c	49.036	52.366	56.951	61.071	58.071	63.191
α12d_β8c	51.662	55.082	54.261	57.991	50.718	55.748
α12d_β11c	43.303	46.483	42.937	47.017	43.604	48.364
α12d_β14c	28.634	31.824	30.841	34.621	33.933	38.633
α12d_γ8b	46.922	50.602	33.338	37.108	35.231	40.371
α12d_γ29a	50.319	54.559	46.773	51.853	34.523	40.903
α12d_δ7a	34.797	38.227	32.044	36.054	18.823	23.873
β14c_γ8b	40.893	43.623	42.170	45.280	33.737	37.697
β14c_γ29a	53.501	56.791	34.838	39.258	69.033	74.233
β14c_δ7a	42.764	45.244	22.265	25.615	53.994	57.864

β27b_γ29a	24.353	27.983	50.977	55.587	56.161	62.271
β27b_δ7a	20.936	23.756	42.686	46.226	44.917	49.697
β33b_γ29a	46.949	51.229	55.250	60.150	73.998	81.538
β33b_δ7a	44.178	47.648	53.502	57.332	60.486	66.696
γ8b_β5c	65.946	68.816	65.767	69.217	53.649	58.029
γ8b_β8c	56.138	59.098	54.976	58.036	42.397	46.687
γ8b_β11c	44.487	47.207	45.139	48.549	29.915	33.935
γ8b_β14c	40.893	43.623	42.170	45.280	33.737	37.697
γ8b_δ7a	31.597	34.567	37.679	41.019	44.324	48.634
δ23d_β5c	42.079	45.009	61.989	65.809	65.020	69.450
δ23d_β8c	47.550	50.570	60.523	63.953	61.166	65.506
δ23d_β11c	38.693	41.473	52.626	56.406	53.681	57.751
δ23d_β14c	27.394	30.184	38.427	41.907	40.714	44.724
δ23d_β27b	58.636	61.766	28.637	32.307	53.847	58.767
δ23d_γ8b	54.246	57.526	41.169	44.639	48.300	52.750
δ23d_γ29a	65.272	69.112	61.850	66.630	42.862	48.552
δ26d_β5c	57.582	60.332	70.055	74.285	76.560	81.760
δ26d_β11c	51.987	54.587	61.319	65.509	61.640	66.480
δ26d_β14c	41.122	43.732	47.088	50.978	50.804	55.584
δ26d_β27b	62.732	65.682	39.408	43.488	54.855	60.545
δ26d_γ8b	58.131	61.231	51.855	55.735	47.945	53.165
δ26d_γ29a	70.229	73.889	67.641	72.831	35.727	42.187
δ28d_β5c	61.042	64.412	69.336	73.936	80.103	85.323
δ28d_β8c	64.602	68.062	72.216	76.426	75.068	80.198
δ28d_β11c	52.421	55.641	64.116	68.676	64.269	69.129
δ28d_β14c	43.910	47.140	49.246	53.506	53.640	58.440
δ28d_β27b	63.120	66.690	48.880	53.330	58.379	64.089
δ28d_γ8b	56.298	60.018	60.923	65.173	49.024	54.264
δ10a_β5c	51.956	54.426	54.330	57.870	66.643	71.343
δ10a_β8c	46.493	49.053	47.673	50.823	55.125	59.735
δ10a_β11c	40.662	42.982	32.770	36.270	50.407	54.747
δ10a_β14c	28.715	31.045	29.988	33.188	46.582	50.862
δ10a_β27b	24.615	27.285	30.538	33.928	32.317	37.507
γ24a_β5c	61.938	64.608	57.656	61.376	70.942	75.532
γ24a_β8c	58.992	61.752	53.179	56.509	61.762	66.262
γ24a_β11c	55.622	58.142	37.394	41.074	59.279	63.509
γ24a_β14c	42.159	44.689	38.043	41.423	52.006	56.176

Table 2 Helix rigidifying restraints used in simulations.

Pairs of residues	distance between P-atoms [Å]	
	minimal	maximal
α1d_δ25d, α2d_δ24d, α3d_δ23d, α4d_δ22d, α5d_δ21d, α6d_δ20d, α7d_δ19d, β1c_α30c, β2c_α29c, β3c_α28c, β4c_α27c, β5c_α26c, β6c_α25c, β7c_α24c, β8c_α23c, β9c_α22c, β10c_α21c, β11c_α20c, β12c_α19c, γ1b_β30b, γ2b_β29b, γ3b_β28b, γ4b_β27b, γ5b_β26b, γ6b_β25b, γ7b_β24b, δ1a_γ25a, δ2a_γ24a, δ3a_γ23a, δ4a_γ22a, δ5a_γ21a, δ6a_γ20a, δ7a_γ19a	9.125	10.125
α2d_δ25d, α3d_δ24d, α4d_δ23d, α5d_δ22d, α6d_δ21d, α7d_δ20d, β2c_α30c, β3c_α29c, β4c_α28c, β5c_α27c, β6c_α26c, β7c_α25c, β8c_α24c, β9c_α23c, β10c_α22c, β11c_α21c, β12c_α20c, γ2b_β30b, γ3b_β29b, γ4b_β28b, γ5b_β27b, γ6b_β26b, γ7b_β25b, δ2a_γ25a, δ3a_γ24a, δ4a_γ23a, δ5a_γ22a, δ6a_γ21a, δ7a_γ20a	7.482	8.482
α1d_δ24d, α2d_δ23d, α3d_δ22d, α4d_δ21d, α5d_δ20d, α6d_δ19d, β1c_α29c, β2c_α28c, β3c_α27c, β4c_α26c, β5c_α25c, β6c_α24c, β7c_α23c, β8c_α22c, β9c_α21c, β10c_α20c, β11c_α19c, γ1b_β29b, γ2b_β28b, γ3b_β27b, γ4b_β26b, γ5b_β25b, γ6b_β24b, δ1a_γ24a, δ2a_γ23a, δ3a_γ22a, δ4a_γ21a, δ5a_γ20a, δ6a_γ19a	12.952	13.952
α37c_β5c, α36c_β6c, α35c_β7c, α34c_β8c, α33c_β9c, α32c_β10c, α31c_β11c, α30c_β12c, α29c_β13c, α28c_β14c, α27c_β15c, α26c_β16c, α25c_β17c, α24c_β18c, α23c_β19c, β37b_γ5b, β36b_γ6b, β35b_γ7b, β34b_γ8b, β33b_γ9b, β32b_γ10b, β31b_γ11b, β30b_γ12b, β29b_γ13b, β28b_γ14b, γ32a_δ5a, γ31a_δ6a, γ30a_δ7a, γ29a_δ8a, γ28a_δ9a, γ27a_δ10a, γ26a_δ11a, γ25a_δ12a, γ24a_δ13a, γ23a_δ14a, δ32d_α5d, δ31d_α6d, δ30d_α7d, δ29d_α8d, δ28d_α9d, δ27d_α10d, δ26d_α11d, δ25d_α12d, δ24d_α13d, δ23d_α14d	17.995	17.995

α36c_β5c, α35c_β6c, α34c_β7c, α33c_β8c, α32c_β9c, α31c_β10c, α30c_β11c, α29c_β12c, α28c_β13c, α27c_β14c, α26c_β15c, α25c_β16c, α24c_β17c, α23c_β18c, β36b_γ5b, β35b_γ6b, β34b_γ7b, β33b_γ8b, β32b_γ9b, β31b_γ10b, β30b_γ11b, β29b_γ12b, β28b_γ13b, γ31a_δ5a, γ30a_δ6a, γ29a_δ7a, γ28a_δ8a, γ27a_δ9a, γ26a_δ10a, γ25a_δ11a, γ24a_δ12a, γ23a_δ13a, δ31d_α5d, δ30d_α6d, δ29d_α7d, δ28d_α8d, δ27d_α9d, δ26d_α10d, δ25d_α11d, δ24d_α12d, δ23d_α13d	16.469	16.469
α37c_β6c, α36c_β7c, α35c_β8c, α34c_β9c, α33c_β10c, α32c_β11c, α31c_β12c, α30c_β13c, α29c_β14c, α28c_β15c, α27c_β16c, α26c_β17c, α25c_β18c, α24c_β19c, β37b_γ6b, β36b_γ7b, β35b_γ8b, β34b_γ9b, β33b_γ10b, β32b_γ11b, β31b_γ12b, β30b_γ13b, β29b_γ14b, γ32a_δ6a, γ31a_δ7a, γ30a_δ8a, γ29a_δ9a, γ28a_δ10a, γ27a_δ11a, γ26a_δ12a, γ25a_δ13a, γ24a_δ14a, δ32d_α6d, δ31d_α7d, δ30d_α8d, δ29d_α9d, δ28d_α10d, δ27d_α11d, δ26d_α12d, δ25d_α13d, δ24d_α14d	20.843	20.843

5.4.13.2 SimRNA folding simulations and refinement of models

Simulations of the J(abcd) structure formation performed with SimRNA were divided into the following stages: i) folding with multiple independent Replica Exchange Monte-Carlo (REMC) simulations without using any experimental restraints, ii) selection of low-energy conformations with the correct secondary structure, iii) refinement of selected structural models by SimRNA REMC simulations using FRET distance restraints, and iv) conversion of coarse-grained models into full-atom versions and their local optimization.

All folding simulations were initiated with separated RNA chains as a starting configuration. The REMC sampling scheme was used with nine replicas, with temperature values uniformly distributed between 0.55 and 0.90 (SimRNA internal dimensionless units). Each simulation comprised 16 000 000 iterations, in which conformational changes were introduced and energies of the resulting conformations were monitored. In every simulation after each period of 16000 iterations (i.e. 1000 times per simulation) the lowest energy conformation

from each replica was recorded (i.e. 9000 conformations and corresponding structural models per simulation).

For folding without experimental FRET restraints (i.e. with restraints on base-pairing only) 140 independent simulations were carried out. Each simulation run was screened for conformations with correct secondary structure, and if more than one such conformation was found, then one with the lowest SimRNA energy was selected for further refinement and analysis.

Refinement was carried out with SimRNA, with the temperature range from 0.40 to 0.60 and all other parameters set the same as for the unrestrained simulations (nine replicas in each simulation, temperatures uniformly distributed in a given range, 16 000 000 iterations, lowest energy conformations recorded for each period of 16000 iterations), and single lowest energy conformation from each simulation was selected. All selected coarse-grained models generated by SimRNA were finally converted to all-atom representation using a built-in SimRNA tool and were locally optimized using the AMBER ff10 force-field [89]. The local structure optimization comprised two stages (2000 iterations of steepest descent minimization each). During the first stage, positions of all atoms present in the coarse-grained representation of SimRNA were fixed, and only the positions of newly added atoms were allowed to relax. During the second stage of minimization, positions of all atoms other than P and all atoms in the junction region (up to three base pairs away from the junction) were allowed to move.

As a second computational strategy folding with FRET-restraints (and with helix-rigidifying restraints) was used. Here, 27 independent simulations were carried out, 9 per each set of restraints, corresponding to the three different conformers observed experimentally. For each of these three groups of simulation runs, structural models were recorded and pooled together and 1 % of lowest-energy structures were retained for clustering. Representatives of five largest clusters, closest in terms of RMSD to centroids of the respective clusters, were selected for further processing. The refinement of these selected structures was carried out with SimRNA, with the temperature range from 0.40 to 0.80 (i.e. lower than in the initial folding simulations) and all other parameters set the same as in the folding stage (nine replicas in each simulation, temperatures uniformly distributed in a given range, 16 000 000 iterations, lowest energy structures recorded for each period of 16000 iterations), and a single lowest-energy structure from each simulation was selected for further processing.

Each of the selected structures underwent all-atom reconstruction and local refinement with the AMBER ff10 force-field as described above.

5.4.13.3 Clustering of SimRNA simulations for J(abcd)

For each of the 3 conformers, 1 % lowest-energy structures from all SimRNA simulations have been clustered based on pairwise RMSD comparisons. The minimal radius of clusters to be recorded has been limited to 3 Å. 100 largest clusters were selected and centroid of each cluster was chosen for the χ^2 estimation.

5.4.14 Precision estimation for MD models

Cluster centroid (see section 5.4.12.1) with the best agreement to FRET determines a $\chi^2_{r,max}$ value (see section 5.4.10). All trajectory members with χ^2_r values below this ($\chi^2_{r,max}$) represent the distribution of possible conformations. For this set, RMSD values were calculated for each phosphorous atom with respect to the cluster centroid with the best agreement to FRET. The average RMSD over all P atoms of a structure is used to characterize the overall precision of the MD model.

5.4.15 Precision estimation for SimRNA models

The cluster centroid (see section 5.4.13.3) with the best agreement to FRET determines a $\chi^2_{r,max}$ value (see section 5.4.10). All cluster centroids with χ^2_r values below this ($\chi^2_{r,max}$) represent the distribution of possible conformations. For this set, RMSD values were calculated for each phosphorous atom with respect to the cluster centroid with the best agreement to FRET. The average RMSD over all P atoms of a structure is used to characterize the overall precision of SimRNA body model.

5.4.16 **SAXS measurements**

All data were taken on beam line 4-2 at the Stanford Synchrotron Radiation Lightsource (SSRL), CA, USA, employing a sample-detector distance of 2.5 m and a Mar 325 CCD detector. The data were collected using a capillary flow cell at a temperature of 20 °C and an X-ray energy of 9 keV. Details of the beamline are described elsewhere [90]. SAXS data of the RNA four-way junction were collected at RNA concentrations ranging from ~4 to 60 µM. Buffer blanks were collected before and after each of the concentration series and subtracted for background correction.

For each condition, ten exposures of 1 s each were taken. Data were image corrected, normalized by incident flux, circularly averaged, and profiles calculated using SASTool [90]. The scattering profiles showed no signs of radiation damage, as tested by comparing scattering profiles of subsequent exposures of the same sample (data not shown). The magnitude of the moment transfer vector is defined as $q = 4\pi \cdot \sin(\theta)/\lambda$, where 2θ is the total scattering angle and λ is the X-ray wavelength. The usable q-range, after removing parasitic scattering at the lowest q values, was $0.02 < q < 0.33$ Å^{-1}.

Scattering profiles at different RNA concentrations were superimposable after scaling by forward scattering intensity (Figure 5-9A) and showed reasonable linearity in the Guinier region [91, 92](i.e. in a plot of *ln(I)* vs. q^2 for the lowest q values; (Figure 5-9B), indicating the absence of aggregation or interparticle interference[93]. In addition, no significant differences between scattering data obtained in the presence of 10 or 20 mM $MgCl_2$ (Figure 5-9) was observed.

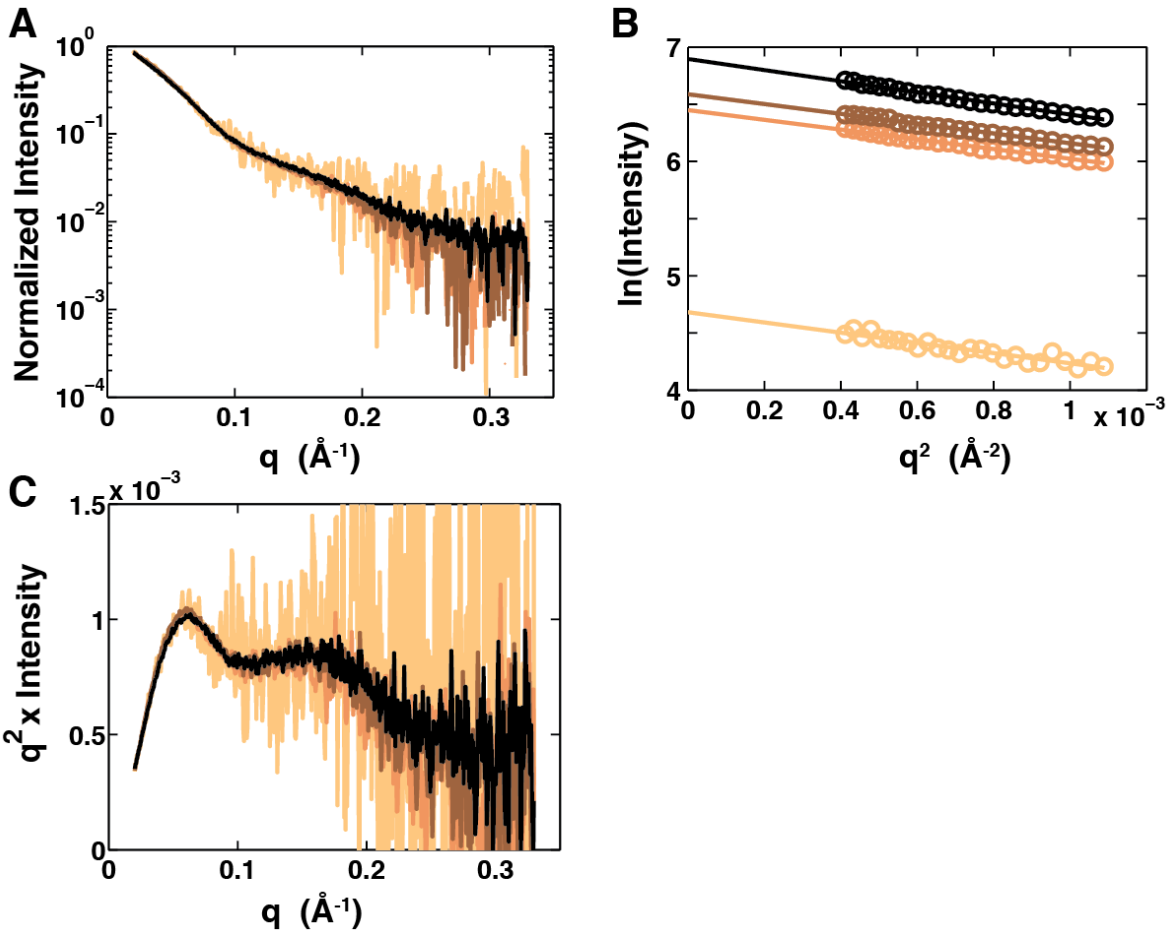


Figure 5-9 Small-angle X-ray scattering data for the J(abcd). (A) Scattering profiles for the J(abcd) at RNA concentrations of approximately 4, 30, 40, and 60 µM (light to dark). The profiles at RNA concentrations of 4 and 30 µM were measured in the presence of 20 mM $MgCl_2$, the profiles at 40 and 60 µM in the presence of 10 mM $MgCl_2$. Scattering profiles were normalized by forward scattering intensity obtained from the Guinier fits (see next panel). (B) Guinier plots of *ln(I)* vs. q^2 for small *q* values for the same profiles as in panel A. Data are shown as circles and linear fits are shown as straight lines. (C) Kratky plots of $q^2 \cdot I(q)$ vs. *q* for the same data as in panel A.

5.4.17 **SAXS data analysis**

We used the crysol program [94] to predict scattering profiles for structural models of the RNA four-way junction construct. Control calculations predicting scattering profiles with the program FoXS [95] gave very similar results (data not shown). Neither crysol nor FoXS take into account the small contribution of the ion atmosphere to the overall scattering intensity; however, the contribution from the ion scattering is typically small and does not influence

the overall shape of the predicted scattering significantly. Notably, both crysol and FoXS have been successfully applied to RNA modeling (see e.g. [96, 97]). Control calculations, where explicit ions from the MD simulations into the structural models were included, were performed to predict scattering profiles. Including explicit ions increased the predicted radius of gyration by 1 – 2 Å, without significantly altering the overall shape of the predicted scattering profiles or the quality of the fits (data not shown). To compare the predicted scattering profiles to experimental SAXS data, a χ^2 criterion was used (Eq. 5.4-11),

$$\chi^2 = \Sigma_i \, (I_{exp,i} - I_{theory,i})^2 / \sigma_i^2 \qquad \textbf{Eq. 5.4-11}$$

where $I_{exp,i}$ and $I_{theory,i}$ are the measured and predicted scattering profiles at momentum transfer value q_i, respectively, and σ_i is the experimental error. $I_{exp,i}$ and σ_i were obtained as the mean and standard deviation from multiple experimental measurements. The χ^2 criterion applied to SAXS data does not have the usual statistical interpretation via the χ^2-distribution, as the intensities at different q_i are not independent normal variables. Nonetheless, profiles with lower χ^2 values provide a better fit to the experimental data, so it provides a valid criterion to compare fits for different models. In addition to plotting scattering intensity vs. q, the Kratky representation of $q^2 \cdot I(q)$ vs. q was employed to visually compare scattering profiles (see Figure 5-9). The latter is more sensitive to differences in conformation of the scattering macromolecule [66, 93].

5.5 Results

Our RNA samples were measured in free diffusion by multi-parameter fluorescence detection (smMFD, section 3.2.1) to simultaneously acquire all fluorescence parameters [56]. A high Mg^{2+}-concentration (20 mM $MgCl_2$) was chosen to avoid structural transitions within the observation time (~ 1 ms) [98]. FRET-averaged mean DA distances $\langle R_{DA} \rangle_E$ are btained from the FRET indicator F_D/F_A (ratio of D and A fluorescence) [56] (section 5.4.3).

Figure 5-10A shows two dimensional MFD-histograms of F_D/F_A and the donor anisotropy r_D versus the donor fluorescence lifetime $\tau_{D(A)}$ for the sample *(D)β11c_(A)δ23d*. The three FRET and the D-only species follow the theoretically expected dependence between $\tau_{D(A)}$ and F_D/F_A for a static molecule (section 5.4.4). The low donor anisotropies prove fast rotational diffusion, minimizing dye orientation problems in the FRET analysis.

5.5.1 Measurement of D only fluorescence decays and quantum yields Φ_{FA}

According to Eq. 5.4-1, for the determination of R_{DA} the acceptor quantum yield Φ_{FA} needs to be known (see also section 5.4.3). Furthermore, to determine the static FRET line the fluorescence decay of the donor in absence of FRET needs to be determined (see section 5.4.4). It has been shown that quenching of Alexa488 and Cy5 attached to dsRNAs using long C6 dye linkers is mostly dynamic [69]. Thus, to determine Φ_{FA} and the D fluorescence decays for each A and D labeling position, respectively, eTCSPC measurements for the single labeled molecules were performed (see section 3.1). Values for Φ_{FA} were assumed to be proportional to the species averaged lifetime $\langle\tau\rangle_x = x_1\tau_1 + x_2\tau_2$ and calculated by calibrating against a sample labeled with Cy5 with a known fluorescence quantum yield. Single labeled dsDNA with a C6-hexamethylen linker and Cy5 with $\langle\tau\rangle_x$ = 1.16 ns and Φ_{FA} = 0.32 due to the presence of ~ 20% cis-trans isomerization under conditions (instead of Φ_{FA} = 0.4 expected for ensemble measurements) [69, 99] were used for calibration. Additionally Φ_{FA} = 0.32 was used for the static FRET line to calibrate g_G/g_R with the dsDNA (see section 5.4.5). Therefore, even if the determined values for Φ_{FA} are wrong, errors due to wrong calibration cancel out and Eq. 5.4-5 will yield the correct distance R_{DA}. For D samples, it was assumed that free Alexa488 has $\tau_{D(0)}$ = 4.1 ns and $\Phi_{FD(0)}$ = 0.8 [69, 99]. The results of the fits of the fluorescence decays and resulting values for the fluorescence and the species averaged lifetimes ($\langle\tau\rangle_x$ and $\langle\tau\rangle_f$, respectively) and for Φ_{FA} are compiled in Table 3. The fitted fluorescence decays of the D-only molecules were used to calculate static FRET lines

according to section 5.4.4 while assuming σ_{DA} = 6 Å. The resulting polynomial coefficients are listed in Table 4.

Table 3 Fluorescence lifetimes (τ) and quantum yields (Φ_{FA}) of single labeled RNA four- and three-way junction samples. See section 7.6 and AppendixIV in [100] for all data and fit plots.

D-only					
	τ_1, ns (x_1)	τ_2, ns (x_2)	$\langle\tau\rangle_x$, ns	$\langle\tau\rangle_f$, ns	
(D)β5c	4.10(92%)	1.34(8%)	3.87	4.02	
(D)β8c	4.03(90%)	1.23(11%)	3.74	3.93	
(D)β11c	4.03(95%)	2.28(5%)	3.94	3.98	
(D)β14c	4.12(93%)	0.61(7%)	3.86	4.08	
(D)β27b	3.87(79%)	0.40(21%)	3.15	3.77	
(D)β29b	3.96(92%)	0.74(8%)	3.7	3.91	
(D)γ8b	3.85(81%)	1.18(19%)	3.34	3.67	
(D)γ29a	3.62(75%)	0.79(25%)	2.92	3.43	
(D)γ7b	3.92(88%)	1.22(12%)	3.6	3.81	
(D)δ7a	3.99(82%)	0.51(18%)	3.35	3.89	
A-only					
	τ_1, ns (x_1)	τ_2, ns (x_2)	$\langle\tau\rangle_x$, ns	$\langle\tau\rangle_f$, ns	Φ_{FA}
(A)α12d	1.90(27%)	1.03(74%)	1.26	1.38	0.35
(A)β14c	1.82(23%)	1.03(77%)	1.21	1.3	0.34
(A)β27b	1.63(27%)	0.99(73%)	1.16	1.23	0.32
(A)β33b	1.91(72%)	1.14(28%)	1.69	1.76	0.47
(A)γ8b	1.81(25%)	1.03(75%)	1.23	1.32	0.34
(A)γ24a	1.53(41%)	0.93(59%)	1.17	1.25	0.32
(A)δα_Δd7a	1.33(55%)	0.82(45%)	1.11	1.16	0.302
(A)δα_Δd10a	1.68(39%	0.99(61%)	1.26	1.35	0.35
(A)δα-Δd26c	1.6(25%)	0.84(75%)	1.02	1.13	0.281
(A)δα_Δd28c	1.38(46%)	0.87(54%)	1.1	1.16	0.301
(A)γ12b	1.36(47%)	0.77(53%)	1.05	1.13	0.286
(A)δ10a	1.68(39%)	0.99(61%)	1.26	1.35	0.35
(A)δ23d	1.65(22%)	1.00(78%)	1.15	1.21	0.32
(A)δ26d	1.76(62%)	1.20(38%)	1.54	1.59	0.43
(A)δ28d	1.90(79%)	1.13(21%)	1.73	1.79	0.48

Table 4 Polynomial coefficients ci used for the calculation of the static FRET-lines (see section 5.4.4).

D position	c_0	c_1	c_2	c_3
(D)β5c	-0.0501	0.5383	0.2748	-0.0391
(D)β8c	-0.0516	0.5528	0.2795	-0.0413
(D)β11c	-0.0449	0.5003	0.2837	-0.0391
(D)β14c	-0.0518	0.5563	0.2638	-0.0373
(D)β27b	-0.0672	0.7212	0.2455	-0.0443
(D)γ8b	-0.058	0.6107	0.2937	-0.05
(D)γ29a	-0.0675	0.7118	0.2845	-0.0569
(D)γ7b	-0.0516	0.5632	0.2878	-0.0446
(D)β29b	-0.0507	0.5584	0.2766	-0.0411
(D)δ7a	-0.0644	0.6786	0.2494	-0.0418

5.5.2 smFRET measurements of double stranded RNAs (dsRNA)

To prove the assumption that the helices in RBD have perfect A-RNA structures, smMFD measurements (see section 3.2.1) of the dsRNAs shown in Figure 5-1E were performed (half of them published in Sindbert et al. [69]). Additionally, using AV modeling (see section 5.4.1) and assuming perfect A-RNA structure, expected distances $\langle R_{DA} \rangle_E$ were calculated. Comparing the measured and the calculated distances (see Table 5, RMSD = 2.8 Å), shows that the assumption of A-RNA form is justified. The differences between the distances measured for dsRNA molecule bc and those for corresponding FRET pairs and the major conformer $(ad)_a$ (will be discussed in section 5.5.4.1) of the J(abcd) (see Table 5, for $(ad)_a$ helices *b* and *c* should be stacked) are significantly larger (RMSD = 4.6 Å). This indicates that the helices in the RNA four-way junction are not perfectly coaxially stacked.

Table 5 Measured and calculated distances $\langle R_{DA}\rangle_E$ for dsRNAs and those for the corresponding FRET pairs in the J(abcd).

dsRNA molecule	FRET pair	$\langle R_{DA}\rangle_E$, measured, [Å]	$\langle R_{DA}\rangle_E$, calculated, [Å]	Corresponding FRET pair in J(abcd)	$\langle R_{DA}\rangle_E$, for $(ad)_a$, [Å]
ab	(D)γ29a_(A)β/γ28b	76.4	78.1	(D)γ29a_(A)β33b	49.5
ab	(D)γ29a_(A)β/γ24b	61.1	63.0		
ab	(D)γ29a_(A)β/γ23b	60.5	60.6	(D)γ29a_(A)β27b	42.1
bc	(D)α/γ26c_(A)β33	76.4	68.9		
bc	(D)α/γ26c_(A)β28	57.5	57.0		
bc	(D)α/γ26c_(A)β27	57.6	55.6		
bc	(D)β14c_(A)α/γ8b	41.4	40.9	(D)β14c_(A)γ8b	46.1
bc	(D)β11c_(A)α/γ8b	48.8	49.8	(D)β11c_(A)γ8b	50.7
bc	(D)β8c_(A)α/γ8b	55.2	56.3	(D)β8c_(A)γ8b	59.3
bc	(D)β5c_(A)α/γ8b	57.7	61.0	(D)β5c_(A)γ8b	64.1

5.5.3 Distance and errors from PDA

To extract values of $\langle R_{DA}\rangle_E$ from noisy single-molecule data photon distribution analysis (PDA, section 5.4.3) [71, 73] was used.

show an exemplary PDA together with the resulting distances, relative amplitudes, and overall uncertainties for the fitted FRET species for four- and three- way junctions (see sections 5.5.4.2 and 5.5.5.2 for all resulting distances, amplitudes, and errors). The broadening beyond shot noise due to multiple acceptor states was accounted for by a global parameter for the half widths of the FRET peaks σ_{app}. In Figure 5-10D (upper) the dependence of σ_{app} on the distance ⟨RDA⟩E for the same acceptor is presented for RNA four- and three-way junctions investigated in the current work. Additionally, the same dependence for another two sets of acceptors for three-way junction molecules is shown (Figure 5-10D, lower). The expected linear dependence for all acceptors is observed [46], which proves the consistency of the FRET data for all molecules.

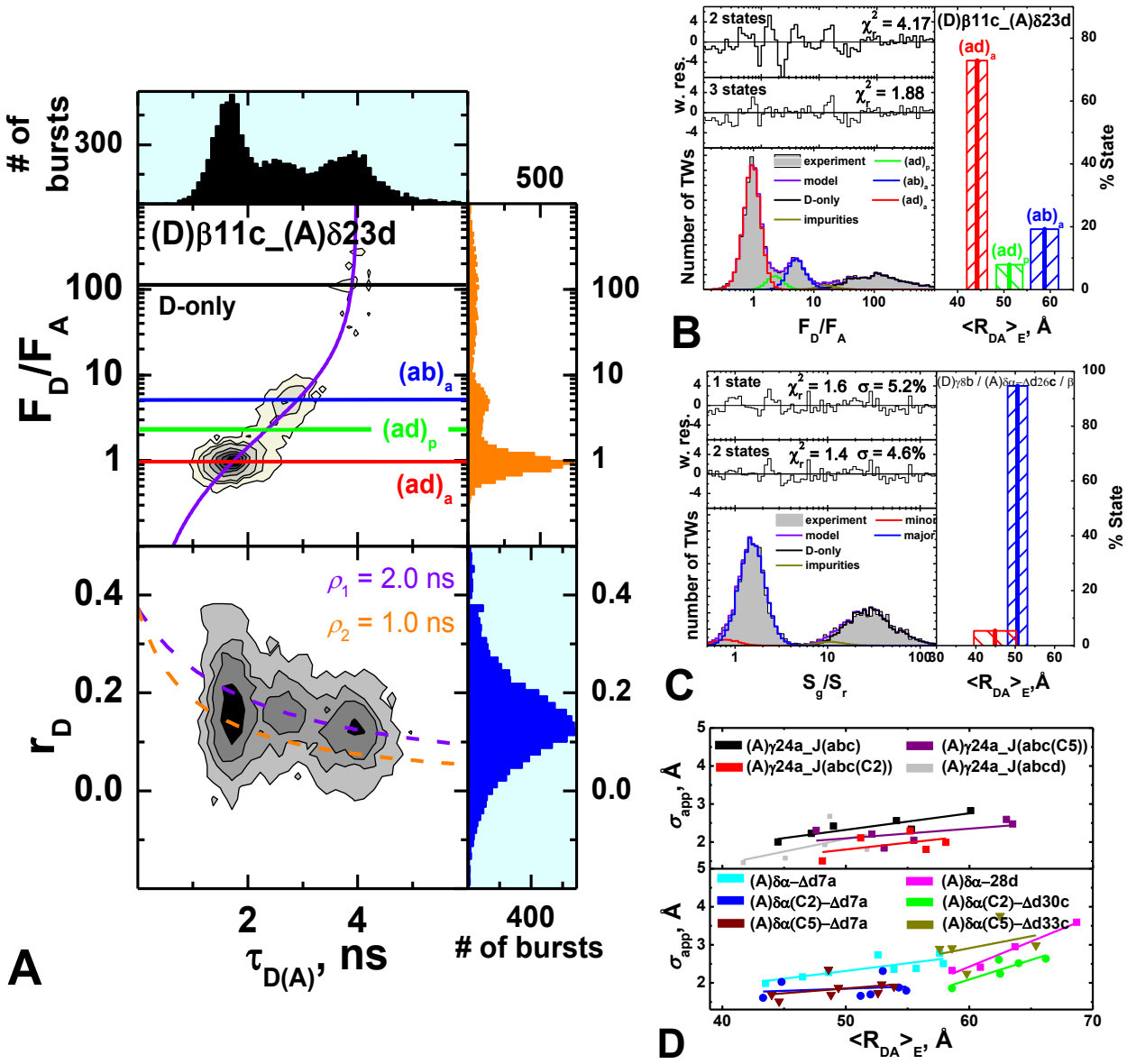


Figure 5-10 (A) 2D burst frequency histograms of F_D/F_A (upper panel) and r_D (lower panel) versus $\tau_{D(A)}$, respectively, for sample *(D)β11c_(A)δ23d*. The number of molecules (fluorescence bursts) in each bin is gray scaled from white (lowest) to black (highest). 1D-histograms are shown as projections. The static FRET line (purple, section 5.4.4) and horizontal lines indicating the D-only and the FRET states are overlaid in the F_D/F_A vs $\tau_{D(A)}$ plot. The dashed lines in the r_D-$\tau_{D(A)}$ diagram are given by the Perrin equation $r_D = r_0/(1 + \tau_{D(A)}/\rho)$, with indicated rotational correlation times and $r_0 = 0.374$. (B,C) PDA (section 5.4.3) for the samples *(D)β11c_(A)δ23d and* (D)γ8b / (A)δα–Δd26c / β . The experimental F_D/F_A histogram is fitted using three FRET states ($\langle R_{DA}\rangle_{E(1)}$, $\langle R_{DA}\rangle_{E(2)}$ and $\langle R_{DA}\rangle_{E(3)}$) (two FRET states on (C)) with a global relative apparent width σ_{app}, one D-only and one state accounting for impurities (also present in D-only samples; see sections 5.5.5.2 and 5.5.5.3 for all parameters). Weighted residuals for three- and two FRET states fits (two and one FRET states on (C)) are shown in the middle and the upper plots, respectively. The right panel shows the distances, relative amplitudes, and the confidence intervals (striped boxes) for

the three FRET states (two FRET states on (C)). (D) For the acceptors γ24a (upper) and δα–Δd7a and δα–Δd28c (lower) the distances $\langle R_{DA} \rangle_E$ versus the apparent peak widths σ_{app}, both determined by PDA, are plotted for four- and three-way junctions. Linear fit was used to fit data points with the same acceptors.

DA-pair	Equation type	Color	Equation number	Equation
(A)γ12b / (D)β14c	Static FRET line corrected for linker movement	Blue	Eq. 5.4-5	F_D/F_A = (0.7693/0.3200)/((3.9794/((-0.0766*x^3)+(0.5926*x^2)+-0.1547*x+0.0343))-1)
(A)γ12b / (D)β14c	Perrin equation	Dashed blue, dashed orange	Eq. 2.1-2	r_D = 0.3740/(1+x/2) r_D = 0.3740/(1+x/1.0)

5.5.4 RNA four-way junction

5.5.4.1 Fitting the data

For 22 of the 51 datasets, three FRET states were necessary to reach a satisfactory fit quality in PDA (χ_r^2 and weighted residuals; Figure 5-10B): one major state with an average relative amplitude ~ 70 % (± 9 % standard deviation) and two minor states with average amplitudes of ~ 15 % (± 6 % standard deviation), respectively. This confirms the presence of at least three quasi-static conformers in equilibrium [98]. Two states (one major and one minor) were sufficient for the remaining datasets (Figure 5-11).

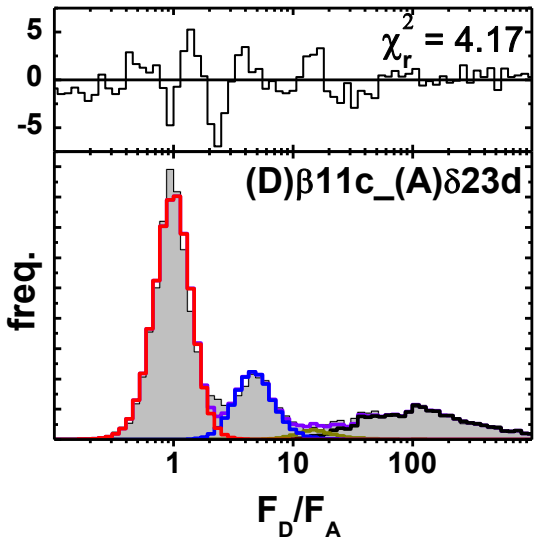


Figure 5-11 PDA for sample *(D)β11c_(A)δ23d* (selected bursts). F_D/F_A histogram of experimental data (gray area) is fitted (purple solid line) using the following parameters: 47.5 % of $\langle R_{DA}\rangle_{E(1)}$ = 44.5 Å (red); 13.9 % of $\langle R_{DA}\rangle_{E(2)}$ = 57.8 Å (blue); σ_{app} = 4.5 % of $\langle R_{DA}\rangle_E$; 36.5 % of D-only; 2.0 % of impurities with apparent R_{DA} = 70.5 Å; χ_r^2 = 4.17. Weighted residuals are shown in the upper plot.

We note that for certain dye positions, very similar DA distances are expected for different conformers, which leads to the observed overlap of FRET peaks (e.g. labeling of helices *a* and *d* for conformers $(ad)_a$ and $(ad)_p$, see Figure 5-1C).

For quantitative FRET analysis FPS requires the calculation of specific overall errors for the individual distances $\{\Delta R_{DA}\}$ considering error propagation rules. $\{\Delta R_{DA}\}$ includes uncertainties due to photon statistics, $\{\Delta R_{DA}(E)\}$ as determined by PDA (between 0.3 % and 21.9 %, see section 5.4.6), and the mutual orientation of the dyes described by the orientation factor κ^2 ($\{\Delta R_{DA}(\kappa^2)\}$, 5.0% [69]), such that $\{\Delta R_{DA}\}$ ranges between 5.0 % and 22.5 %.

5.5.4.2 **Distances and errors**

Table 6 Values for measured distances $\langle R_{DA}\rangle_E$ and their relative amplitudes *x*, measurement errors ΔR_{DA} (resulting from $\Delta R_{DA}(\kappa^2)$ and $\Delta R_{DA}(E)$, see section 5.4.6) and model distances R_{model} resulting from rigid body docking. See AppendixIV in [100] for all data and fit plots.

			$(ad)_a$			$(ad)_p$			$(ab)_a$		
#	DA-pair	σ_{app}	$<R_{DA}>_E$, Å (x_1)	ΔR_{DA}	R_{model}, Å	$<R_{DA}>_E$, Å (x_2)	ΔR_{DA}	R_{model}, Å	$<R_{DA}>_E$, Å (x_3)	ΔR_{DA}	R_{model}, Å
1	(D)β5c_(A)α12d	4.10%	53.4(78.9%)	5.10%	51.4	53.4(10.2%)	15.50%	52.5	60.2(10.9%)	13.60%	55.5
2	(D)β8c_(A)α12d	3.90%	58.7(73.8%)	5.00%	52.7	58.7(11.8%)	10.90%	53.1	48.3(14.5%)	5.50%	50.2
3	(D)β11c_(A)α12d	3.80%	56.4(85.6%)	5.00%	50.4	46.6(9.7%)	5.70%	47.6	41(4.7%)	7.00%	43.5
4	(D)β14c_(A)α12d	5.90%	44.2(78.3%)	5.00%	38.7	44.2(10.8%)	10.40%	36.3	44.2(10.8%)	10.40%	35.8
5	(D)Υ8b_(A)α12d	4.90%	46.8(82.2%)	5.00%	43.8	36.9(8.9%)	11.60%	40.9	36.9(8.9%)	11.60%	38
6	(D)Υ29a_(A)α12d	3.60%	55.1(85.7%)	5.20%	52.1	47.7(4%)	9.70%	50.9	55.1(10.3%)	11.60%	50.7
7	(D)δ7a_(A)α12d	4.60%	39.8(70.5%)	5.10%	36.4	49.1(14.8%)	10.60%	36.6	49.1(14.8%)	10.60%	34.2
8	(D)Υ8b_(A)β14c	4.30%	44.3(80.8%)	5.20%	43.4	50.5(7.5%)	7.20%	45	44.3(11.7%)	13.50%	40.6
9	(D)Υ29a_(A)β14c	4.50%	52.2(83.8%)	5.10%	54.9	42.2(3.1%)	6.70%	41.1	56.1(13.1%)	10.70%	70.1
10	(D)δ7a_(A)β14c	4.30%	43.9(88%)	5.00%	48.5	38.9(6.2%)	6.70%	31.2	51(5.8%)	6.30%	52.8
11	(D)Υ29a_(A)β27b	3.70%	42.1(69.5%)	5.00%	40.9	50.8(7.6%)	7.90%	50.4	60.5(22.9%)	6.70%	59.2
12	(D)δ7a_(A)β27b	3.70%	39.3(66.4%)	5.10%	37.1	46.3(13.5%)	7.80%	49.5	51.8(20.1%)	5.80%	51
13	(D)Υ29a_(A)β33b	4.30%	49.5(80.7%)	5.00%	51.8	59.8(8.1%)	6.50%	60	75.1(11.2%)	10.40%	80.1
14	(D)δ7a_(A)β33b	4.40%	52.7(67.9%)	5.00%	54.8	59.7(12.6%)	5.70%	63.5	73.4(19.4%)	5.10%	66.4
15	(D)β11c_(A)Υ24a	5.50%	48.7(75.1%)	5.20%	53.7	62.4(12.4%)	16.90%	62.7	62.4(12.4%)	16.90%	71
16	(D)β8c_(A)Υ24a	4.00%	48.3(66.3%)	5.10%	46.5	54.6(11.9%)	7.70%	61.6	63.2(21.8%)	5.60%	58.5
17	(D)β14c_(A)Υ24a	3.50%	51.7(82.3%)	5.10%	47.3	44.9(6.8%)	5.80%	48.8	51.7(10.9%)	9.30%	54.2
18	(D)β27b_(A)Υ24a	3.50%	41.7(57%)	5.10%	42.2	48.7(22.1%)	6.70%	42.8	53.9(20.9%)	7.00%	54.7
19	(D)β5c_(A)Υ24a	3.50%	45.1(54.8%)	5.40%	44.4	41.4(20.7%)	6.30%	40.1	50.3(24.5%)	5.50%	47.3
20	(D)β5c_(A)Υ8b	4.80%	64.1(73.5%)	5.30%	65.6	64.1(12.2%)	11.40%	67	54.3(14.3%)	7.20%	58.4
21	(D)β8c_(A)Υ8b	4.80%	59.3(72.9%)	5.10%	60.2	59.3(12.7%)	13.30%	60.5	48.9(14.4%)	5.40%	51.9
22	(D)β11c_(A)Υ8b	4.00%	50.7(78.4%)	5.10%	50.7	50.7(10.4%)	14.40%	50.3	41.2(11.2%)	6.60%	40.1
23	(D)β14c_(A)Υ8b	4.70%	46.1(82.5%)	5.30%	43.5	53.1(6.7%)	7.20%	45	46.1(10.9%)	9.80%	40.6
24	(D)δ7a_(A)Υ8b	4.30%	41.2(67%)	5.00%	38.4	48.9(16.5%)	7.20%	47.5	48.9(16.5%)	7.20%	49.2
25	(D)β11c_(A)δ10a	4.50%	49.8(62.6%)	5.10%	52.5	56.4(21%)	6.50%	63.5	67.9(16.5%)	6.20%	67
26	(D)β8c_(A)δ10a	5.70%	51.6(61.8%)	6.00%	46.6	61.6(24%)	11.40%	63.7	51.6(14.2%)	15.90%	54.8
27	(D)β27b_(A)δ10a	3.50%	52.8(81.6%)	5.10%	48.3	46.5(8.4%)	8.30%	51.3	52.8(9.9%)	12.30%	51.8
28	(D)β14c_(A)δ10a	3.50%	41.9(58%)	5.10%	42.3	47.7(20.8%)	6.40%	43.9	53(21.2%)	6.00%	51.9
29	(D)β5c_(A)δ10a	3.50%	45.8(58%)	6.20%	46.4	42.2(16.2%)	7.30%	42.1	50.5(25.8%)	7.20%	48.1
30	(D)Υ8b_(A)δ10a	3.50%	46.9(77%)	5.10%	47.1	43.6(15.6%)	6.30%	40.2	52(7.3%)	8.00%	48.4

31	(D)β5c_(A)δ23d	4.40%	44.4(67.4%)	5.00%	41	53.7(10.7%)	8.60%	57.8	62.4(21.9%)	5.80%	63
32	(D)β8c_(A)δ23d	4.30%	45.5(68.5%)	5.10%	47	53(21.9%)	7.20%	56.1	58.1(9.6%)	9.60%	60.1
33	(D)β11c_(A)δ23d	3.80%	44.2(72.9%)	5.00%	44.8	51.2(7.9%)	5.60%	55.3	58.8(19.2%)	5.10%	56.9
34	(D)β14c_(A)δ23d	4.50%	39.2(62.5%)	5.20%	35.3	46.8(18.7%)	11.30%	44.8	46.8(18.7%)	11.30%	45.1
35	(D)β27b_(A)δ23d	4.60%	53.8(82.1%)	5.10%	55.3	46.2(7.4%)	8.50%	44.9	53.8(10.5%)	11.70%	58.5
36	(D)Υ8b_(A)δ23d	3.80%	51(79.3%)	5.10%	53.2	43.6(9%)	7.00%	46.9	57.2(11.8%)	7.80%	55.3
37	(D)Υ29a_(A)δ23d	4.60%	65.6(84.9%)	6.70%	68.1	65.6(10.2%)	9.70%	67.4	53.1(4.9%)	6.10%	57.4
38	(D)β5c_(A)δ26d	5.60%	56.6(78.2%)	5.10%	52.2	69.5(10.3%)	7.20%	69.9	85.3(11.4%)	10.10%	77
39	(D)β8c_(A)δ26d	5.50%	62.5(59.9%)	5.30%	60.1	71.2(20%)	10.30%	69.7	71.2(20%)	10.30%	71.9
40	(D)β11c_(A)δ26d	5.40%	58.8(79.4%)	5.10%	58.5	70.6(10.3%)	11.80%	67.1	70.6(10.3%)	11.80%	66.8
41	(D)β14c_(A)δ26d	5.50%	51.6(78%)	5.10%	47.1	62.7(11%)	13.40%	54.4	62.7(11%)	13.40%	55.6
42	(D)β27b_(A)δ26d	4.70%	63(74.4%)	5.40%	63.1	52.4(13.9%)	9.20%	48	63(11.7%)	13.10%	61.5
43	(D)Υ8b_(A)δ26d	6.00%	56.8(82.4%)	5.50%	61.1	44.4(4.3%)	18.20%	50.7	56.8(13.3%)	15.90%	58
44	(D)Υ29a_(A)δ26d	5.50%	78(69%)	5.90%	76.7	66(27.4%)	7.30%	73.8	52.6(3.6%)	10.80%	49.3
45	(D)β5c_(A)δ28d	5.60%	53.6(78.2%)	5.00%	58.3	67.6(9.9%)	8.10%	70.7	85(11.8%)	22.50%	80.9
46	(D)β8c_(A)δ28d	5.60%	58.1(76.8%)	5.10%	66	73.6(11.6%)	11.70%	72.4	73.6(11.6%)	11.70%	76.1
47	(D)β11c_(A)δ28d	5.80%	57.7(83.9%)	5.10%	62.5	73(8.1%)	20.60%	70.3	73(8.1%)	20.60%	69.3
48	(D)β14c_(A)δ28d	5.00%	52.1(69.4%)	5.20%	50.3	62.2(17.3%)	12.70%	56.8	71.2(13.4%)	13.00%	57.8
49	(D)β27b_(A)δ28d	5.60%	67.4(64.8%)	5.30%	62.2	56.8(23.1%)	5.40%	54.2	67.4(12%)	13.80%	60.6
50	(D)Υ8b_(A)δ28d	6.00%	61.4(81.6%)	5.30%	60.3	47.2(6.3%)	7.70%	56.7	61.4(12.1%)	21.10%	56.8
51	(D)Υ29a_(A)δ28d	6.00%	80.1(76.5%)	5.70%	79	64.9(11.7%)	15.30%	76.9	64.9(11.7%)	15.30%	50.7

5.5.4.3 Fluctuations of FRET state amplitudes

Table 7 Relative amplitudes *x* and errors due photon statistics Δx (see section 5.4.6) for those PDA fits with three FRET states (no overlaid states).

	$(ad)_a$		$(ad)_p$		$(ab)_a$	
DA-pair	x_1	Δx_1 (abs.)	x_2	Δx_2 (abs.)	x_3	Δx_3 (abs.)
(D)β11c_(A)α12d	85.6%	3.0%	9.7%	29.1%	4.7%	59.1%
(D)γ29a_(A)β14c	83.8%	13.3%	3.1%	49.5%	13.1%	112.2%
(D)δ7a_(A)β14c	88.0%	5.4%	6.2%	58.9%	5.8%	48.8%
(D)γ29a_(A)β27b	69.5%	7.6%	7.6%	44.5%	22.9%	29.3%
(D)δ7a_(A)β27b	66.4%	9.2%	13.5%	39.9%	20.1%	34.6%
(D)γ29a_(A)β33b	80.7%	9.0%	8.1%	40.5%	11.2%	52.3%
(D)δ7a_(A)β33b	67.9%	6.9%	12.6%	32.1%	19.4%	9.0%
(D)β8c_(A)γ24a	66.3%	11.5%	11.9%	48.0%	21.8%	26.8%
(D)β14c_(A)γ24a	57.0%	11.8%	22.1%	39.3%	20.9%	63.9%
(D)β27b_(A)γ24a	54.8%	21.8%	20.7%	68.3%	24.5%	43.4%
(D)β5c_(A)δ10a	62.6%	15.4%	21.0%	31.9%	16.5%	22.8%
(D)β14c_(A)δ10a	58.0%	11.7%	20.8%	28.0%	21.2%	58.1%

(D)β27b_(A)δ10a	58.0%	32.2%	16.2%	87.8%	25.8%	74.4%
(D)γ8b_(A)δ10a	77.0%	14.3%	15.6%	55.3%	7.3%	59.7%
(D)β5c_(A)δ23d	67.4%	10.6%	10.7%	42.3%	21.9%	21.8%
(D)β8c_(A)δ23d	68.5%	11.8%	21.9%	38.1%	9.6%	80.4%
(D)β11c_(A)δ23d	72.9%	4.1%	7.9%	29.4%	19.2%	12.7%
(D)γ8b_(A)δ23d	79.3%	10.0%	9.0%	52.7%	11.8%	63.5%
(D)β5c_(A)δ26d	78.2%	27.1%	10.3%	40.8%	11.4%	44.0%
(D)γ29a_(A)δ26d	69.0%	20.7%	27.4%	48.6%	3.6%	80.6%
(D)β5c_(A)δ28d	78.2%	11.7%	9.9%	37.5%	0.0%	55.9%
(D)β14c_(A)δ28d	69.4%	9.3%	17.3%	44.0%	13.4%	51.0%
	$\langle x_1 \rangle$	$\langle \Delta x_1 \rangle$	$\langle x_2 \rangle$	$\langle \Delta x_2 \rangle$	$\langle x_3 \rangle$	$\langle \Delta x_3 \rangle$
	70.8%	12.7%	13.8%	44.8%	14.8%	50.2%
	$\Delta\langle x_1 \rangle$ **(rel.)**		$\Delta\langle x_2 \rangle$ **(rel.)**		$\Delta\langle x_3 \rangle$ **(rel.)**	
	13.5%		46.0%		50.5%	

5.5.4.4 Fit parameters for all PDA fits

Table 8 PDA fit parameters for the datasets measured at 20 mM $MgCl_2$. See AppendixIV in [100] for all data and fit plots.

DA-pair	$\langle R_{DA}\rangle_{E,1}$, Å ($x_1$)	$\langle R_{DA}\rangle_{E,2}$, Å ($x_2$)	$\langle R_{DA}\rangle_{E,3}$, Å ($x_3$)	σ_{app}	D-only	impurities R_{DA}, Å (x)	$\langle B_G\rangle$, kHz	$\langle B_R\rangle$, kHz	α	g_G/g_R	χ_r^2
(D)β5c_(A)α12d	53.4(62.1%)	53.4(8%)	60.2(8.6%)	4.1%	20%	75.9(1.3%)	1.19	0.79	2.4%	0.4	1.48
(D)β8c_(A)α12d	58.7(59%)	48.3(11.6%)	58.7(9.4%)	3.9%	18.9%	74.6(1.1%)	1.13	0.75	2.5%	0.4	1.14
(D)β11c_(A)α12d	56.4(67.1%)	46.6(7.6%)	41(3.7%)	3.8%	20.8%	77.8(0.8%)	1.13	0.75	2.3%	0.4	0.97
(D)β14c_(A)α12d	44.2(62.1%)	44.2(8.6%)	44.2(8.6%)	5.9%	19%	80.5(1.8%)	1.13	0.75	2.9%	0.4	3.05
(D)Y8b_(A)α12d	46.8(56.2%)	36.9(6.1%)	36.9(6.1%)	4.9%	30.6%	57.9(1.1%)	0.99	0.67	2.8%	0.32	2.11
(D)Y29a_(A)α12d	55.1(66.5%)	55.1(8%)	47.7(3.1%)	3.6%	20%	74.8(2.5%)	1.44	0.99	2.3%	0.4	1.89
(D)δ7a_(A)α12d	39.8(51.1%)	49.1(10.7%)	49.1(10.7%)	4.6%	26.7%	69.9(8%)	1.12	0.73	3.2%	0.35	1.78
(D)Y8b_(A)β14c	44.3(60%)	50.5(5.6%)	44.3(8.7%)	4.3%	24.9%	67.3(0.7%)	1.13	0.75	2.7%	0.4	1.06
(D)Y29a_(A)β14c	52.2(59.5%)	42.2(2.2%)	56.1(9.3%)	4.5%	25.9%	72.3(3.1%)	1.27	0.53	1.5%	0.85	1.2
(D)δ7a_(A)β14c	43.9(57.9%)	38.9(4.1%)	51(3.8%)	4.3%	33.1%	62.4(1%)	1.42	0.61	1.6%	0.8	1.02
(D)Y29a_(A)β27b	42.1(44.6%)	50.8(4.9%)	60.5(14.7%)	3.7%	33.3%	68.9(2.5%)	1.43	0.62	2%	0.8	1.03
(D)δ7a_(A)β27b	39.3(38.4%)	46.3(7.8%)	51.8(11.6%)	3.7%	38.7%	71.1(3.5%)	1.43	0.62	1.9%	0.8	1.21
(D)Y29a_(A)β33b	49.5(55.5%)	59.8(5.6%)	75.1(7.7%)	4.3%	28.4%	90.7(2.8%)	1.43	0.62	1.9%	0.8	1.42
(D)δ7a_(A)β33b	52.7(57%)	59.7(10.6%)	73.4(16.3%)	4.4%	13.7%	95.1(2.3%)	1.50	0.56	1.5%	0.69	2.11
(D)β11c_(A)Y24a	48.7(51.4%)	62.4(8.5%)	62.4(8.5%)	5.5%	28.7%	78.9(2.9%)	1.54	0.59	1.5%	0.73	1.48
(D)β8c_(A)Y24a	48.3(42.5%)	54.6(7.6%)	63.2(14%)	4%	35.8%		1.81	0.75	1.4%	0.74	2.4
(D)β14c_(A)Y24a	51.7(59.4%)	44.9(4.9%)	51.7(7.9%)	3.5%	26.4%	65.6(1.3%)	1.74	0.64	1.4%	0.71	1.73
(D)β27b_(A)Y24a	41.7(32.5%)	48.7(12.6%)	53.9(11.9%)	3.5%	41.7%	64.8(1.3%)	1.73	0.72	1.4%	0.74	1.22
(D)β5c_(A)Y24a	45.3(38.6%)	41.6(17.1%)	50.4(16.7%)	3.5%	26.5%	64(1%)	1.54	0.59	1.3%	0.73	1.71
(D)β5c_(A)Y8b	64.1(39.6%)	64.1(6.6%)	54.3(7.7%)	4.8%	44%	77.4(2.1%)	1.13	0.75	2.2%	0.4	0.78
(D)β8c_(A)Y8b	59.3(47.1%)	59.3(8.2%)	48.9(9.3%)	4.8%	32.8%	75.9(2.6%)	1.29	0.55	1.3%	0.7	1.99
(D)β11c_(A)Y8b	50.7(46.9%)	50.7(6.2%)	41.2(6.7%)	4.8%	38.6%	63.8(1.6%)	1.57	0.69	1.3%	0.8	1.38
(D)β14c_(A)Y8b	46.1(47%)	53.1(3.8%)	46.1(6.2%)	4.7%	41.5%	66.5(1.5%)	1.57	0.69	1.3%	0.8	1.42
(D)δ7a_(A)Y8b	41.2(41.4%)	48.9(10.2%)	48.9(10.2%)	4.3%	37.1%	61.9(1%)	1.29	0.67	1.3%	0.75	2.74
(D)β11c_(A)δ10a	49.8(43.3%)	56.4(14.5%)	67.9(11.4%)	4.5%	28%	83.4(2.8%)	1.95	0.64	1.2%	0.71	2.32
(D)β8c_(A)δ10a	51.6(25.2%)	61.6(9.8%)	51.6(5.8%)	5.7%	57.6%	69.1(1.7%)	1.81	0.75	1.3%	0.74	0.78

(D)β27b_(A)δ10a	52.8(64.8%)	46.5(6.7%)	52.8(7.9%)	3.5%	19.7%	66.3(0.9%)	1.74	0.64	1.5%	0.71	1.32
(D)β14c_(A)δ10a	41.9(35.5%)	47.7(12.7%)	53(13%)	3.5%	37.7%	61.1(1%)	1.73	0.72	1.9%	0.74	1.69
(D)β5c_(A)δ10a	45.8(31.2%)	42.1(8.6%)	50.5(13.9%)	3.5%	43.1%	66(3.1%)	1.73	0.72	1.5%	0.74	1.74
(D)Y8b_(A)δ10a	46.9(56.6%)	52(5.4%)	43.6(11.5%)	3.5%	24%	70.9(2.5%)	1.73	0.72	1.5%	0.73	0.83
(D)β5c_(A)δ23d	44.4(32.7%)	53.7(5.2%)	62.4(10.6%)	4.4%	51.5%		1.19	0.79	2.4%	0.4	1.36
(D)β8c_(A)δ23d	45.5(38.5%)	53(12.3%)	58.1(5.4%)	4.3%	42%	73.2(1.8%)	1.30	0.54	1.4%	0.85	0.84
(D)β11c_(A)δ23d	44.2(45.1%)	51.2(4.9%)	58.8(11.9%)	3.8%	36.2%	71.5(1.9%)	1.66	0.65	1.3%	0.67	1.88
(D)β14c_(A)δ23d	39.2(31.7%)	46.8(9.5%)	46.8(9.5%)	4.5%	45.9%	72.6(3.5%)	1.29	0.54	1.5%	0.85	1.5
(D)β27b_(A)δ23d	53.8(50.9%)	46.2(4.6%)	53.8(6.5%)	4.6%	36.4%	66.5(1.5%)	1.27	0.53	1.2%	0.85	1.17
(D)Y8b_(A)δ23d	51(31%)	43.6(3.5%)	57.2(4.6%)	3.8%	60.1%	72(0.9%)	0.99	0.67	2.6%	0.32	0.68
(D)Y29a_(A)δ23d	65.6(65.7%)	65.6(7.9%)	53.1(3.8%)	4.6%	22.6%		1.57	0.69	1.3%	0.8	1.25
(D)β5c_(A)δ26d	56.6(56.7%)	69.5(7.5%)	85.3(8.3%)	5.6%	26.3%	90(1.3%)	0.99	0.67	3%	0.32	1.17
(D)β8c_(A)δ26d	62.5(46.1%)	71.2(15.4%)	71.2(15.4%)	5.5%	21%	101(2.2%)	0.99	0.67	2.6%	0.32	0.98
(D)β11c_(A)δ26d	58.8(64.7%)	70.6(8.4%)	70.6(8.4%)	5.4%	18.5%		0.99	0.67	2.6%	0.32	1.05
(D)β14c_(A)δ26d	51.6(39.1%)	62.7(5.5%)	62.7(5.5%)	5.5%	47.2%	73.5(2.8%)	1.19	0.79	2.7%	0.4	0.78
(D)β27b_(A)δ26d	63(53.6%)	52.4(10%)	63(8.4%)	4.7%	28%		2.19	1.15	0.7%	0.75	0.99
(D)Y8b_(A)δ26d	56.8(51.9%)	44.4(2.7%)	56.8(8.4%)	6%	33.6%	72.9(3.4%)	1.44	0.99	2.2%	0.4	1.29
(D)Y29a_(A)δ26d	78(51.4%)	66(20.4%)	52.6(2.7%)	5.5%	25.6%		1.44	0.99	2.3%	0.4	0.7
(D)β5c_(A)δ28d	53.6(58.2%)	67.6(7.4%)	85(8.8%)	5.6%	23.2%	103(2.3%)	1.44	0.91	2.4%	0.4	1.15
(D)β8c_(A)δ28d	58.1(58.2%)	73.6(8.8%)	73.6(8.8%)	5.6%	23%	82.3(1.2%)	1.30	0.54	1.4%	0.85	0.67
(D)β11c_(A)δ28d	57.7(62.5%)	73(6%)	73(6%)	5.8%	22.6%	84.2(3%)	1.29	0.55	1.2%	0.7	0.78
(D)β14c_(A)δ28d	52.1(49.8%)	62.2(12.4%)	71.2(9.6%)	5%	27%	86.3(1.3%)	1.30	0.54	1.3%	0.85	0.9
(D)β27b_(A)δ28d	67.4(46.3%)	56.8(16.5%)	67.4(8.6%)	5.6%	28.6%		1.29	0.67	1.2%	0.75	2.06
(D)Y8b_(A)δ28d	61.4(54%)	47.2(4.2%)	61.4(8%)	6%	32.1%	72.4(1.6%)	1.83	0.79	1.4%	0.67	1.51
(D)Y29a_(A)δ28d	80.1(58.1%)	64.9(8.9%)	64.9(8.9%)	5.5%	24.2%		1.57	0.69	1.3%	0.8	1.75

Table 9 PDA fit parameters for the datasets of the Mg^{2+}-titrations. See AppendixIV in [100] for all data and fit plots.

DA-pair	$[Mg^{2+}]$, mM	$\langle R_{DA}\rangle_{E,1}$, Å ($x_1$)	$\langle R_{DA}\rangle_{E,2}$, Å ($x_2$)	$\langle R_{DA}\rangle_{E,3}$, Å ($x_3$)	σ_{app}	D-only	impurities R_{DA}, Å (x)	$\langle B_G\rangle$, kHz	$\langle B_R\rangle$, kHz	α	g_G/g_R	χ_r^2
(D)b5c_(A)a12d	20	52.7(49.6%)	52.7(7.5%)	58.8(9%)	4%	32.6%	71.8(1.3 %)	2.66	1.56	0.7%	0.75	1.04
	1	55.8(56.8%)	55.8(2.7%)	58.8(10.4%)	3.6%	26.5%	73.6(3.6 %)	2.66	1.56	0.5%	0.75	2.02
	0.1	57.5(61.3%)	51(2.5%)	57.5(7.4%)	3.3%	27.5%	69.5(1.2 %)	2.66	1.46	0.5%	0.75	0.9
(D)b8c_(A)a12d	20	58.7(60%)	58.7(8.5%)	48.3(11.6%)	3.9%	18.9%	74.6(1.1 %)	1.13	0.75	2.5%	0.4	1.16
	1	61.6(49.4%)	61.6(2.5%)	51(20.1%)	4%	25.3%	77.8(2.7 %)	1.66	0.70	1.5%	0.73	1.71
	0.1	60.3(49.9%)	60.3(1%)	52.2(22.2%)	4%	24.1%	71.5(2.8 %)	1.44	0.60	1.5%	0.71	1.26
(D)b11c_(A)a12d	20	56.6(33.6%)	48.2(3.1%)	43.2(3.8%)	2.5%	41.9%	81.7(17.7 %)	1.91	0.63	1.5%	0.68	1.42
	1	56.5(42.8%)	49.9(9.2%)	44(9.7%)	3.72%	36.8%	68.9(1.4 %)	1.91	0.63	1.3%	0.68	1.7
	0.1	54.8(26.1%)	48(9.4%)	43.8(2.8%)	3.7%	61.6%		1.91	0.63	1.3%	0.68	1.17
(D)g29a_(A)a12d	20	54.1(53.3%)	49.6(9.7%)	54.1(7.4%)	2.7%	26.6%	66.5(3 %)	2.79	1.31	0.7%	0.76	1.24
	1	56(54.5%)	49.6(5.1%)	56(6.7%)	2.8%	29.2%	67(4.5 %)	2.79	1.31	0.7%	0.76	0.85
	0.1	56.6(34.3%)	49.6(2%)	56.6(4.5%)	3.2%	52.8%	71.2(6.3 %)	2.79	1.41	0.8%	0.76	1.37
(D)d7a_(A)a12d	20	39.8(51.1%)	49.1(10.7%)	49.1(10.7%)	4.6%	26.7%	69.9(0.8 %)	1.12	0.73	3.2%	0.35	1.78
	1	41.1(41.6%)	49.7(7.7%)	49.7(7.7%)	4.1%	42.3%	65.4(0.7 %)	1.44	0.60	1.6%	0.71	1.59
	0.1	42.5(36.9%)	49.6(8.6%)	49.6(8.6%)	4.5%	44.3%	66.3(1.6 %)	1.44	0.60	1.6%	0.71	1.25
(D)g29a_(A)b33b	20	49.8(49.8%)	55.8(7.6%)	72.1(6.6%)	3.2%	32.7%	80.4(3.3 %)	1.79	0.74	1.7%	0.71	1
	1	59.3(50.7%)	53.1(3.3%)	67.4(13.4%)	3.2%	29.6%	76.5(3 %)	1.79	0.74	1.7%	0.71	1.51
	0.1	62.7(45%)	55.8(2.2%)	68.6(22.5%)	3.2%	27.5%	74.7(2.8 %)	1.79	0.74	1.7%	0.71	1.43
(D)d7a_(A)b33b	20	51.9(48.8%)	59.8(8.6%)	73.1(14%)	4.2%	25.8%	86.6(2.7 %)	1.43	0.62	2%	0.8	1.43
	1	53.4(42.1%)	61.5(7%)	72(12.5%)	4%	38.5%		1.43	0.62	1.4%	0.8	1.12
	0.2	53.8(40.9%)	61.4(20.1%)	71.1(13%)	4.1%	26%		1.66	0.70	1.4%	0.73	1.04
(D)b5c_(A)g24a	20	48.7(51.4%)	62.4(8.5%)	62.4(8.5%)	5.5%	28.7%	78.9(2.9 %)	1.54	0.59	1.5%	0.73	1.66
	1	45.8(24.3%)	56.2(7.7%)	68.8(18.9%)	4.6%	44.6%	87.1(4.5 %)	1.64	0.62	1.6%	0.7	1.27
	0.1	46.9(23.4%)	55.2(11%)	66.8(14.1%)	5.1%	46.8%	77.8(4.7 %)	1.64	0.62	1.6%	0.7	0.94
(D)b8c_(A)g24a	20	48.3(42.5%)	54.6(7.6%)	63.2(14%)	4%	35.8%		1.81	0.75	1.4%	0.74	2.4
	1	47.6(23.7%)	56.4(10.6%)	65.7(12.9%)	4%	52.7%		1.66	0.70	1.2%	0.73	1.09
	0.2	48.3(18.3%)	56(18%)	64.3(15.3%)	4%	48.4%		1.66	0.70	1.5%	0.73	1.09
(D)b11c_(A)g24a	20	51.7(59.4%)	44.9(4.9%)	51.7(7.9%)	3.5%	26.4%	65.6(1.3 %)	1.74	0.64	1.4%	0.71	1.73
	1	51.6(63.4%)	46.1(1.7%)	51.6(7.6%)	2.9%	25.3%	61.5(2 %)	1.66	0.70	1.3%	0.73	1.11
	0.2	52.4(63.7%)	45.8(1.1%)	52.4(7.8%)	2.6%	26.1%	62.6(1.3 %)	1.66	0.70	1.4%	0.73	1.41

(D)b14c_(A)g24a	20	41.7(32.5%)	48.7(12.6%)	53.9(11.9%)	3.5%	41.7%	64.8(1.3 %)	1.73	0.72	1.4%	0.74	2.28
	1	42.4(11.4%)	50.3(10.8%)	57.2(6.7%)	3.5%	60.1%	74.7(11 %)	1.64	0.62	1.6%	0.7	1.49
	0.1	41.5(5.3%)	48.6(8.7%)	56.1(9.8%)	3.5%	63%	69.5(13.2 %)	1.64	0.62	1.8%	0.7	2.07
(D)b27b_(A)g24a	20	45.1(22.8%)	41.4(7.2%)	50.3(10.3%)	3.5%	58.8%	64.8(0.9 %)	2.79	1.41	0.6%	0.76	2.02
	1	45.1(27.9%)	41.4(6.6%)	50.3(9.5%)	3.5%	54.3%	66.7(1.7 %)	2.79	1.21	0.6%	0.76	1.47
	0.1	45.1(20.2%)	41.4(1%)	50.3(7.1%)	3.5%	68.7%	69.2(3 %)	2.79	1.51	0.8%	0.76	2.26
(D)b5c_(A)g8b	20	63.2(41.6%)	63.2(4.9%)	52.7(4.7%)	4.7%	48.9%		1.82	0.77	1%	0.73	0.81
	1	65.5(37.7%)	65.5(1.4%)	54.3(5%)	4.7%	55.9%		1.82	0.77	1%	0.73	1.05
	0.1	67.2(45.9%)	67.2(0.4%)	57.7(10.6%)	4.7%	43.1%		1.82	0.77	1%	0.73	1.7
DA-pair	**[Mg^{2+}], mM**	**$\langle R_{DA}\rangle_{E,1}$, Å ($x_1$)**	**$\langle R_{DA}\rangle_{E,2}$, Å ($x_2$)**	**$\langle R_{DA}\rangle_{E,3}$, Å ($x_3$)**	**σ_{app}**	**D-only**	**impurities R_{DA}, Å (x)**	**$\langle B_G\rangle$, kHz**	**$\langle B_R\rangle$, kHz**	**α**	**g_G/g_R**	**χ_r^2**
(D)b8c_(A)g8b	20	59.2(48.9%)	59.2(7.7%)	49(9.7%)	4.8%	30.8%	76.3(2.9 %)	1.29	0.55	1.3%	0.7	1.76
	0.1	59(42.2%)	59(0.7%)	51.2(19.3%)	4.8%	35%	72.2(2.7 %)	1.35	0.63	1.7%	0.75	2.06
(D)b11c_(A)g8b	20	52.2(49%)	52.2(6%)	42.8(5.5%)	4.4%	37.3%	68.5(2.3 %)	1.49	0.58	1.5%	0.69	1.3
	3	53.2(49.1%)	53.2(3.8%)	40.6(4.9%)	4.4%	37.7%	71(4.5 %)	1.49	0.58	1.5%	0.69	1.82
	1	53(49.1%)	53(2.7%)	41.2(4.8%)	4.4%	38.1%	73.4(5.3 %)	1.49	0.58	1.5%	0.69	1.37
	0.3	53.4(46.2%)	53.4(1.7%)	45(6.9%)	4.4%	42.2%	69(3 %)	1.69	0.62	1.6%	0.66	2.14
	0.1	52.6(40.8%)	52.6(0.5%)	44.7(7.3%)	5%	47.8%	71.7(3.6 %)	1.62	0.54	1.5%	0.71	2.01
	0.03	52.4(41.5%)	52.4(0.2%)	44.7(6.8%)	5%	46%	71(5.6 %)	1.62	0.54	1.5%	0.71	2.12
(D)b5c_(A)d10a	20	49.8(43.3%)	56.4(14.5%)	67.9(11.4%)	4.5%	28%	83.4(2.8 %)	1.95	0.64	1.2%	0.71	2.32
	1	46.6(8.9%)	55.7(10.6%)	68.9(3.7%)	4.5%	69.8%	74.8(7 %)	1.64	0.62	1.5%	0.7	1.26
	0.1	47.1(6.6%)	55.5(13.7%)	69.2(6.6%)	4.5%	69.6%	79.5(3.5 %)	1.64	0.62	1.7%	0.7	1.17
(D)b8c_(A)d10a	20	51.3(20.7%)	61.6(9.4%)	51.3(4.6%)	4.5%	65.3%		1.82	0.81	0.9%	0.75	1.19
	1	58(15.5%)	58(1.4%)	50(9.9%)	3.9%	69.7%	66.7(3.6 %)	1.82	0.81	1%	0.75	0.9
	0.1	58.7(19.5%)	58.7(0.6%)	51.1(5.7%)	4.7%	71.2%	66.8(3 %)	1.82	0.81	1.1%	0.75	1.51
(D)b11c_(A)d10a	20	52.8(64.8%)	46.5(6.7%)	52.8(7.9%)	3.5%	19.7%	66.3(0.9 %)	1.74	0.64	1.5%	0.71	1.32
	1	52.6(56.6%)	46.5(2%)	52.6(7%)	3.5%	30.3%	64.3(4.2 %)	1.44	0.60	1.6%	0.71	1.46
	0.1	53.7(43.7%)	46.5(1.5%)	53.7(5.9%)	3.5%	45.4%	65.3(3.6 %)	1.44	0.60	1.6%	0.71	0.92
(D)b14c_(A)d10a	20	41.9(35.5%)	47.7(12.7%)	53(13%)	3.5%	37.7%	61.1(1 %)	1.73	0.72	1.9%	0.74	1.69
	1	42(11.6%)	49.9(9%)	56.2(8.8%)	3.5%	56.4%	73.1(14.2 %)	1.64	0.62	2.1%	0.7	1.77
	0.1	43.2(6.5%)	50.9(10.9%)	59.6(7%)	3.5%	59.7%	74.5(15.8 %)	1.64	0.62	2.1%	0.7	3.57
(D)b27b_(A)d10a	20	46.7(29.8%)	42.8(14.4%)	51.4(9.3%)	3.5%	43.6%	64.3(2.9 %)	1.73	0.72	1.6%	0.74	1.51
	1	47.8(41.3%)	42.8(4.8%)	51.8(11%)	3.5%	41.2%	65.4(1.8 %)	1.73	0.72	1.4%	0.74	1.29
	0.1	47.9(34.4%)	42.8(2%)	52.4(7.5%)	3.5%	54.5%	65.7(1.6 %)	1.73	0.72	1.3%	0.74	1.99

(D)g8b_(A)d10a	20	46.4(59.3%)	50.8(9.1%)	42.4(4.9%)	3.5%	24.1%	69.7(2.6 %)	1.73	0.72	1.4%	0.74	0.86
	1	51.4(46.2%)	42.4(1.2%)	51.4(5.5%)	3.5%	46.2%	62.7(0.9 %)	1.73	0.72	1.6%	0.74	2.17
	0.1	52.8(65.4%)	42.4(0.9%)	52.8(7.5%)	7.5%	25%	64.8(1.1 %)	1.73	0.72	1.6%	0.74	1.15
(D)b5c_(A)d23d	20	44.1(28.3%)	51.1(5%)	62.5(15.7%)	4.4%	51%		1.56	0.65	1.6%	0.71	0.78
	1	50.5(36.1%)	50.5(1.5%)	62.6(13.6%)	4.4%	48.8%		1.56	0.65	1.6%	0.73	0.93
	0.1	53.5(30.5%)	53.5(0.6%)	61(13.5%)	4.4%	55.4%		1.56	0.65	1.5%	0.71	0.83
(D)b11c_(A)d23d	20	44.2(45.1%)	51.2(4.9%)	58.8(11.9%)	3.8%	36.2%	71.5(1.9 %)	1.66	0.65	1.3%	0.67	1.88
	1	43.1(40.5%)	50.9(9%)	58.7(12.1%)	3.7%	38.4%		1.82	0.81	0.9%	0.75	1.16
	0.1	44.3(25.4%)	51.2(15.7%)	58.5(12.4%)	4.1%	46.4%		1.82	0.81	1.1%	0.75	1.16
(D)b27b_(A)d23d	20	53.7(52.4%)	46.3(5.3%)	53.7(6.9%)	4.5%	34.3%	68.2(1.1 %)	1.27	0.53	1.4%	0.85	1.02
	0.1	53.4(60.2%)	46.3(1.8%)	53.4(7.1%)	4.5%	28.3%	75(2.7 %)	1.27	0.53	1.8%	0.85	0.83
(D)g8b_(A)d23d	20	49.9(34.5%)	42.9(3.8%)	57.5(2.8%)	3.1%	55.5%	69.9(3.4 %)	1.82	0.77	1.1%	0.73	1.19
	1	46.9(20.7%)	42.1(1.4%)	51.5(9.9%)	3%	65.6%	62(2.4 %)	1.82	0.69	1.1%	0.73	1.54
	0.1	47.5(20.1%)	43(1.8%)	51.4(14.5%)	3%	62.6%	62.1(1.1 %)	1.82	0.67	1.1%	0.73	0.88
DA-pair	**[Mg^{2+}], mM**	**$\langle R_{DA}\rangle_{E,1}$, Å ($x_1$)**	**$\langle R_{DA}\rangle_{E,2}$, Å ($x_2$)**	**$\langle R_{DA}\rangle_{E,3}$, Å ($x_3$)**	**σ_{app}**	**D-only**	**impurities R_{DA}, Å (x)**	**$\langle B_G\rangle$, kHz**	**$\langle B_R\rangle$, kHz**	**α**	**g_G/g_R**	**χ_r^2**
(D)b27b_(A)d26d	20	63.1(55.4%)	52.6(10%)	63.1(8%)	4.9%	26.7%		2.19	1.15	0.7%	0.75	1.03
	1	62.6(61.6%)	52.6(3%)	62.6(7.1%)	4%	28.3%		2.19	1.11	0.5%	0.75	0.69
	0.1	62.8(64.2%)	52.6(0.8%)	62.8(8.1%)	4.6%	26.9%		2.19	1.11	0.5%	0.75	1.8
(D)b14c_(A)d28d	20	52.2(52.4%)	63.1(12.8%)	71.1(8.2%)	5.1%	25.2%	89.6(1.4 %)	1.29	0.54	1.3%	0.85	1.26
	0.1	54.8(23.6%)	62(43.1%)	71.4(12.9%)	4.7%	20.4%		1.35	0.63	1.7%	0.75	0.73
(D)b27b_(A)d28d	20	68.4(57.7%)	56.8(14.1%)	68.4(8.1%)	5.5%	17.5%	90.7(2.7 %)	1.49	0.55	1.5%	0.68	1.27
	3	66.8(68.9%)	55.7(4.1%)	66.8(8.1%)	5.33%	16.2%	102(2.7 %)	1.69	0.62	1.6%	0.66	0.96
	1	67.2(62.4%)	57(2.8%)	67.2(7.4%)	5.4%	21%	86.5(6.3 %)	1.49	0.55	1.5%	0.68	1.23
	0.3	67.7(54.8%)	56.7(2.7%)	67.6(6.2%)	5.3%	15.5%	96.9(20.9 %)	1.49	0.55	1.5%	0.68	1.09
	0.1	66.9(68%)	56.9(3.8%)	66.9(7.9%)	4.4%	16.9%	96.6(3.4 %)	1.49	0.55	1.5%	0.68	0.93
	0.03	67.3(67.4%)	57.4(3.7%)	67.3(8%)	4.7%	19.1%	92.5(1.8 %)	1.49	0.55	1.5%	0.68	1.38
(D)g8b_(A)d28d	20	61.4(54.1%)	47.2(4.2%)	61.4(8%)	6%	32.1%	72.4(1.6 %)	1.83	0.79	1.4%	0.67	1.5
	1	59.5(54.7%)	51.5(1.8%)	59.5(8.7%)	5.4%	34.8%		2.00	0.90	0.9%	0.72	1.85
	0.1	59.7(52.7%)	48.9(0.7%)	59.7(7.8%)	4.5%	38.8%		2.00	0.90	0.9%	0.72	1.3
(D)g29a_(A)d28d	20	77.5(62.3%)	65.1(13.1%)	65.1(13.1%)	5.5%	11.4%		1.54	0.59	1.1%	0.73	1.15
	1	75.9(64.9%)	66.2(12.4%)	66.2(12.4%)	5.5%	10.3%		1.54	0.59	1.4%	0.73	1.49
	0.1	79.4(48.2%)	70(18.6%)	70(18.6%)	5.5%	14.7%		1.54	0.59	1.5%	0.73	0.9

5.5.4.5 **Assignment of the peaks**

Next, assignment of the FRET peaks to the three conformers should had been done. This was straightforward for the major FRET population. Using all 51 distances associated with the largest amplitude, structural modeling clearly yields the conformer $(ad)_a$ (Figure 5-1C and Figure 5-19A). However, the two minor populations have similar amplitudes, so that information was not very useful. Instead a chemical identification procedure was established by performing Mg^{2+} titration and assigning the conformers via their unique affinity to Mg^{2+}. It was observed that at low Mg^{2+} one minor peak significantly decreased in amplitude, whereas the other one remained stable (Figure 5-12A). This is due to fast species interconversion at low Mg^{2+} concentrations [98, 101, 102] that results in an apparent disappearance of short-lived FRET states within the time resolution of the experiment. A possible kinetic scheme is shown in Figure 5-12B. Referring to the geometric models of the RNA four-way junction in Figure 5-1C, a minor high FRET peak resembles $(ab)_a$ for labeling e.g. on junction arms *b* and *c* and $(ad)_p$ for labeling e.g. on arms *b* and *d*. Following this logic, it was decided that the minor FRET peak disappearing at low Mg^{2+} represents $(ad)_p$, whereas the stable one represents $(ab)_a$. This observation can be attributed to fast species interconversion at low Mg^{2+} concentrations [98, 101, 102]. Considering the time resolution of the experiment, FRET peaks cannot be resolved due to species with sub-ms lifetimes and apparently disappear. Thus, minor peaks can be unambiguously assigned according to their behavior at low Mg^{2+}.

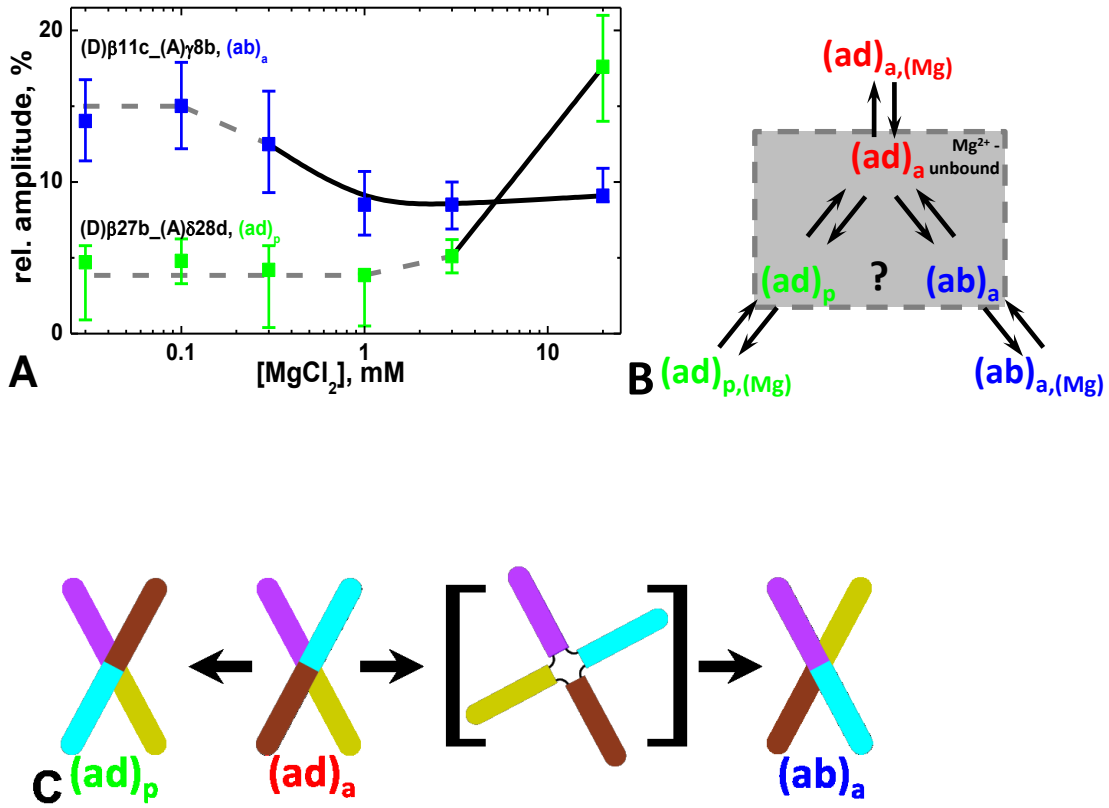


Figure 5-12 (A) Relative fractions of the state assigned to conformer *(ad)*$_p$ for *(D)β27b_(A)δ28d* and the state assigned to *(ab)*$_a$ for *(D)β11c_(A)γ8b* at various $MgCl_2$-concentrations (section 5.5.4.5). (B) Possible transitions between the three conformers and their respective Mg^{2+}-bound and unbound states. (C) The structural transitions from conformer *(ad)*$_a$ to *(ad)*$_p$ (left) or to *(ab)*$_a$ (right). The brackets indicate a transition state.

Two exemplary Mg^{2+} titrations are presented in Figure 5-13. At low Mg-concentrations, the minor high FRET peak slightly increases in amplitude for *(D)β11c_(A)γ8b* and significantly decreases for *(D)β27b_(A)δ28d*. Referring to the geometric model of J(abcd) (Figure 5-1C), the "stable" and "unstable" species resemble *(ab)*$_a$ and *(ad)*$_p$ conformers, respectively, which can be shown as follows. At low Mg-concentrations, the minor high-FRET peak slightly increases in amplitude for *(D)β11c_(A)γ8b* and significantly decreases for *(D)β27b_(A)δ28d* Figure 5-13. The geometric model suggests that, for labeling on helices *b* and *d* the minor state yielding the smaller distance can be assigned to *(ad)*$_p$ (Figure 5-1C). Thus, in the case of *(D)β27b_(A)δ28d* (Figure 5-13) the minor high FRET peak is due to the *(ad)*$_p$ state. Analogously, for *(D)β11c_(A)γ8b* it is due to *(ab)*$_a$. 14 titrations were performed which all confirmed the above assignment, that is, the minor FRET peak that disappears at low Mg^{2+} is consistent with the assumed geometry of *(ad)*$_p$.

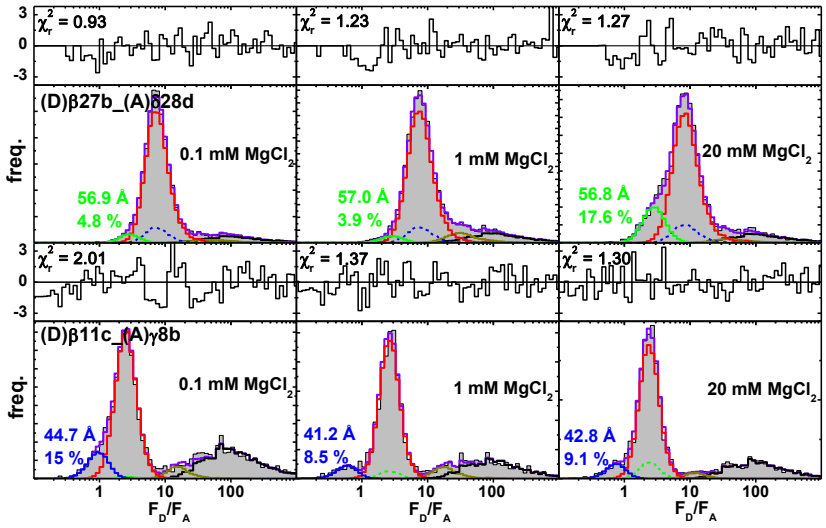


Figure 5-13 PDA (selected bursts) of Mg^{2+}- titrations (0.1 mM, left, 1 mM, middle and 20 mM, right) for samples *(D)β27b_(A)δ28d* (upper panel) and *(D)β11c_(A)γ8b* (lower panel). F_D/F_A histograms of experimental datasets (gray areas) are fitted (purple solid lines) with three FRET states ($\langle R_{DA}\rangle_{E(1)}$, green for $(ad)_a$, $\langle R_{DA}\rangle_{E(2)}$, red for $(ad)_p$ and $\langle R_{DA}\rangle_{E(3)}$, blue for $(ab)_a$), one D-only (black) and one impurity state (dark yellow) (0see section 5.5.5.3 for all PDA fit parameters). Weighted residuals are shown in the respective upper plots.

In addition, Mg^{2+} titrations were performed for all samples with *(A)α12d* which is close to the junction, and all samples with labeled Guanines (*(A)γ24a* and *(A)δ10a*) where the dye's mean position is far away from the RNA's helical axis (see section 5.4.2). In other 19 cases the minor peaks were assigned based on the geometric model (Figure 5-1C). Interestingly, the different stability of $(ab)_a$ and $(ad)_p$ at low Mg^{2+} concentrations can be rationalized considering the differences in transition pathways from $(ad)_a$ to $(ab)_a$ and $(ad)_p$, respectively. The $(ad)_a$ to $(ad)_p$ transition is accomplished by simple rotations of helix pairs that retain the ability to interact with each other whereas a $(ad)_a$ to $(ab)_a$ transition requires a destacking of helix pairs (Figure 5-12C), and thus can be expected to be associated with crossing higher energy barriers. At this stage of analysis the presence of conformer $(ab)_p$ cannot be completely excluded as its FRET populations overlap with those of other conformers (Figure 5-1C). All PDA results are summarized in0 (see AppendixIV in [100] for all data and fit plots).

5.5.4.6 **Modeling**

We obtained three unique sets of 51 distances $\langle R_{DA}\rangle_E$ and corresponding errors ΔR_{DA} (section 5.5.4.2) to generate accurate three-dimensional structural models of the three conformers by rigid body docking (RBD) and further refinement in the next step.

In RBD we, initially, treated the individual arms as rigid bodies having perfect A-RNA form as confirmed by hpFRET measurements of the ds helix arms in Figure 5-1D (for details see section 5.5.2). Significant deviations could be expected only in close proximity to the junction. In addition to the 51 FRET distances, the helices were kept together by four restraints representing the covalent bonds in the phosphate backbone of the four strands. In the RBD part of FPS structures that agree best with the measured distances are found by repeatedly modeling the explicit accessible volumes (AVs) of fluorophores during the RDB dynamics [9] (sections 5.4.1, 5.4.7 and 5.4.9 for details). In combination with the determined input uncertainties the weighted data-model deviation for a set of 51 + 4 distances is minimized by docking the four helices of dsRNA. This is repeated 1000 times with random start orientations to cover all local minima.

5.5.4.7 **Clustering of the RBD models**

With the help of χ_r^2 and RMSD vs previous structure plots one can identify the solutions clusters below and abowe the significance threshold (section 5.4.10).

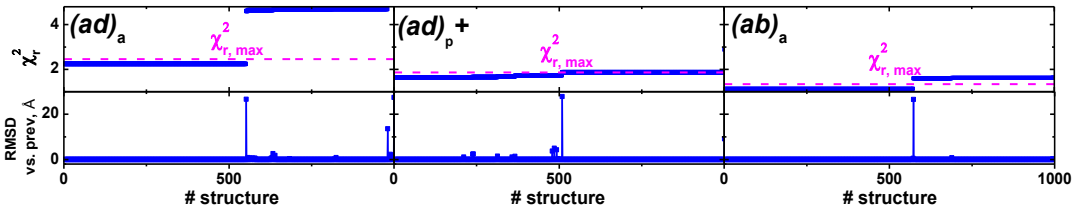


Figure 5-14 χ_r^2 (upper panels) and RMSD vs. the previous structure (lower panels) plotted against the structure ID found after docking in ascending order with respect to χ_r^2 for *(ad)*$_a$ (left), *(ad)*$_p$ (middle) and *(ab)*$_a$ (right). The dashed magenta lines represent 84 % confidence thresholds (see section 5.4.10):$\chi^2_{r,max}$ = 2.46, 1.86 and 1.34 for *(ad)*$_a$, *(ad)*$_p$+ and *(ab)*$_a$, respectively.

χ_r^2 and “RMSD vs. the previous structure” for all docked structures are plotted in Figure 5-14 for the three conformers. The values for χ_r^2 shown in Figure 5-14 were calculated considering deviations between input distances and model distances not only for DA-pairs ($\langle R_{DA} \rangle_E$ and R_{model}) but also for the four restraints representing the covalent bonds in the phosphate backbone of the four strands. Steps in the χ_r^2 plot and corresponding peaks in the RMSD plot (Figure 5-14) separate groups of very similar solutions (clusters). For *(ad)*$_a$ and *(ab)*$_a$ unique solutions are found below the 84 % confidence threshold ($\chi^2_{r,\max}$, see section 5.4.10 for details). Hence, they can be assumed to be unique. For *(ad)*$_p$ there are several solutions below $\chi^2_{r,\max}$. However, they deviate from the best one by only up to 2.9 Å RMSD which is comparable to the model’s precision of 2.3 Å (see Figure 5-19E). This would mean that the respective "best" structures can be considered unique with a confidence level of > 99.99 % for $(ad)_a$, 84 % for $(ad)_{p+}$ and 97 % for $(ab)_a$. It should be mentioned that FRET does not provide structural information in the junction regon, thus RBD is unable to distinguish between $(ad)_{p+}$ and $(ad)_{p-}$ conformers.

5.5.4.8 Uncertainties of the RBD models

To estimate the uncertainties of the RBD models a bootstrapping procedure [77] is applied according to [9] (section 5.4.11). In Figure 5-15 the respective best RBD models are overlaid with sets of structures representing regions of confidence. The average precision defined by root mean square deviations (r.m.s.d) for all P-atoms is 1.7 Å, 2.3 Å, and 2.6 Å, for conformers $(ad)_a$, $(ad)_{p+}$, and $(ab)_a$, respectively.

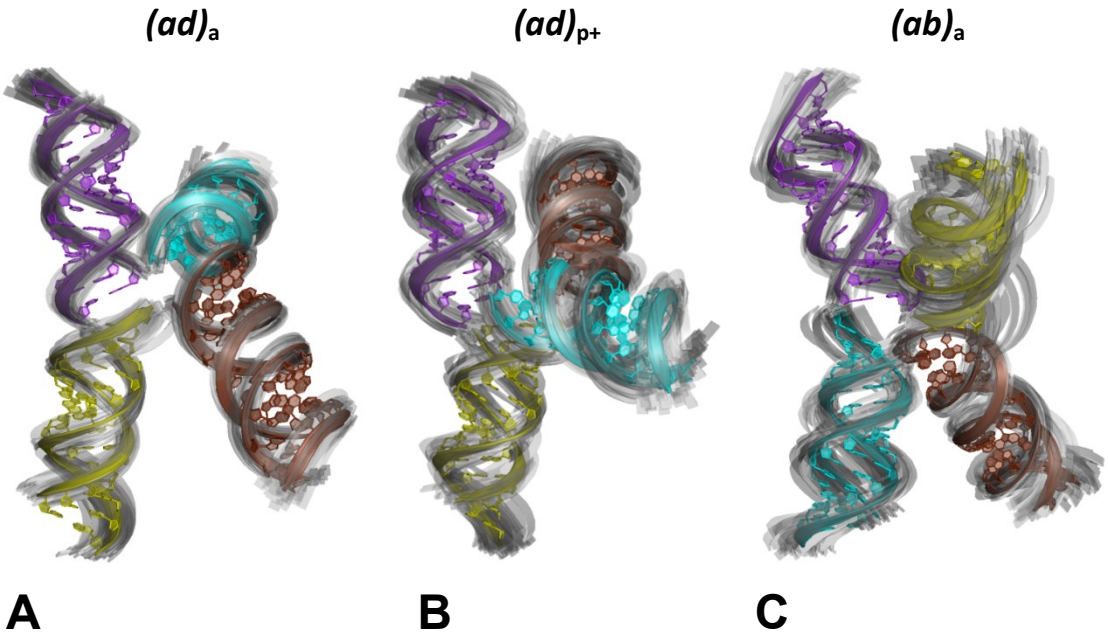


Figure 5-15 Structures (cartoon representation) of the rigid body models with the lowest χ_r^2 for conformer *(ad)*$_a$ (A), *(ad)*$_{p+}$ (B) and *(ab)*$_a$ (C) overlaid with 100 structures resulting from bootstrapping (gray transparent) that indicate the uncertainties of the helix positions and orientations (see section 5.4.11).

Table 10 Angles between helix pairs for rigid body models with lowest χ_r^2 of onformers *(ad)*$_a$ and (*ab*)$_a$.

	(ad)$_a$	(ab)$_a$
helix pair	angle [°]	
ab	70	144
ac	136	146
ad	159	62
bc	139	55
bd	101	123
cd	63	144

5.5.4.9 **“2nd best” docking results**

For all three conformers, approximately half of the solutions generated by the docking procedure deviate strongly from the respective best ones (RMSD = 26.6 Å for *(ad)*$_a$, 28.1 Å for *(ad)*$_{p+}$ and 26.6 Å for *(ab)*$_a$, see Figure 5-14) and can be excluded with a confidence of > 99.99 % (*(ad)*$_a$), 84 % (*(ad)*$_{p+}$) and 97 % (*(ab)*$_a$), respectively. Comparing these “second best” clusters of structures with the according best ones shows that they are quasi-mirrored images (see Figure 5-16). This is due to the fact that for most labeling positions (labeled uracils) the dye linker points into the major groove of dsRNA (see Figure 2-1). Thus, the mean positions of the dyes are close to the helix axis (see Figure 5-5) and are, therefore, less sensitive to mirror image transformation of the structure. For labeled guanines, dye linkers point into the minor groove (see Figure 2-1) resulting in AVs that are significantly displaced from the helical axis (seeFigure 5-5). Labeling positions *(A)γ24a* and *(A)δ10a* were specifically chosen to distinguish between the two quasi - mirrored solutions for conformer *(ad)*$_a$ resulting in a significantly increased χ_r^2 for the “second best” solution (Figure 5-14).

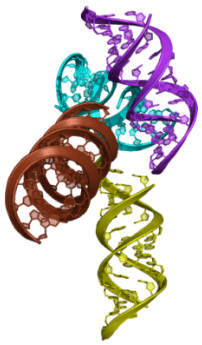
best solution	*2nd best solution*
***(ad)*$_a$**	
	
***(ad)*$_{p+}$**	
	
***(ab)*$_a$**	
	

Figure 5-16 Comparison of structures with lowest (left) and 2nd lowest (right) χ_r^2 resulting from rigid body docking for conformers (*ad*)$_a$ (top), (*ad*)$_{p+}$ (middle) and (*ab*)$_a$ (bottom).

5.5.4.10 **Conformer $(ad)_a$ at 0.1 mM $MgCl_2$**

For a few cases, the major FRET peak ($(ad)_a$) yielded a significant shift in distance at low Mg^{2+}-concentrations compared to the corresponding measurements at 20 mM (see Table 11). These changes were attributed to a slightly changed conformation of $(ad)_a$ in its Mg^{2+}-unbound state which has been predicted to resemble a 90° cross [103]. It exhibits a clear $(ad)_a$ conformation (see Figure 5-15) only with angles between the stacked helix pairs closer to 90° compared to the structure at 20 mM $MgCl_2$.

Table 11 $\langle R_{DA}\rangle_E$ measured at 0.1 mM $MgCl_2$ and corresponding ones at 20 mM.

DA pair	0.1 mM $MgCl_2$	20 mM $MgCl_2$	DA pair	0.1 mM $MgCl_2$	20 mM $MgCl_2$
	$<R_{DA}>_E$, Å	$<R_{DA}>_E$, Å		$<R_{DA}>_E$, Å	$<R_{DA}>_E$, Å
(D)β5c_(A)α12d	57.5	53.4	(D)β5c_(A)δ10a	47.1	49.8
(D)β8c_(A)α12d	60.3	58.7	(D)β8c_(A)δ10a	58.7	51.6
(D)β11c_(A)α12d	54.8	56.4	(D)β11c_(A)δ10a	53.7	52.8
(D)γ29a_(A)α12d	56.6	55.1	(D)β14c_(A)δ10a	43.2	41.9
(D)δ7a_(A)α12d	42.5	39.8	(D)β27b_(A)δ10a	47.9	45.8
(D)γ29a_(A)β33b	62.7	49.5	(D)γ8b_(A)δ10a	52.8	46.9
(D)δ7a_(A)β33b	53.8	52.7	(D)β5c_(A)δ23d	53.5	44.4
(D)β5c_(A)γ24a	46.9	48.7	(D)β11c_(A)δ23d	44.3	44.2
(D)β8c_(A)γ24a	48.3	48.3	(D)β27b_(A)δ23d	53.4	53.8
(D)β11c_(A)γ24a	52.4	51.7	(D)γ8b_(A)δ23d	47.5	51
(D)β14c_(A)γ24a	41.5	41.7	(D)β27b_(A)δ26d	62.8	63
(D)β27b_(A)γ24a	45.1	45.1	(D)β14c_(A)δ28d	54.8	52.1
(D)β5c_(A)γ8b	67.2	64.1	(D)β27b_(A)δ28d	66.9	67.4
(D)β8c_(A)γ8b	59	59.3	(D)γ8b_(A)δ28d	59.7	61.4
(D)β11c_(A)γ8b	52.6	50.7	(D)γ29a_(A)δ28d	79.4	80.1

Docking the J(abcd) with the 30 distances for conformer $(ad)_a$ measured at 0.1 mM $MgCl_2$ (see Table 11) yields two solutions (see Figure 5-17 and Figure 5-18) with very similar values for χ_r^2 (χ_r^2 = 1.9 and 2.0, respectively, $\chi_{r,\max}^2$ = 2.2). For the docking an error of 7 % was assumed for each distance.

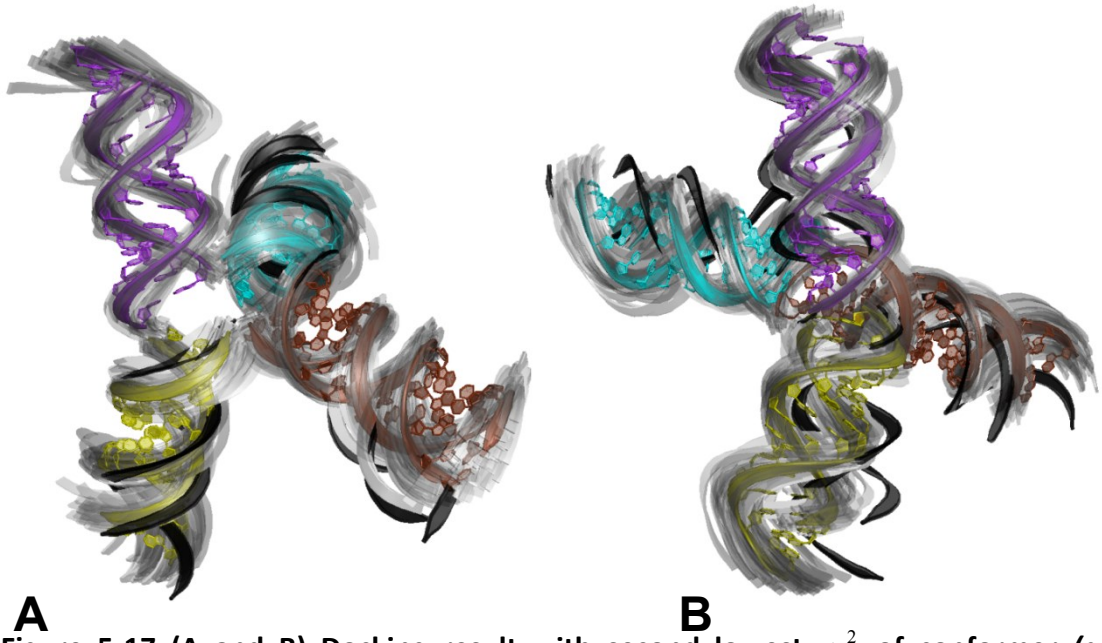


Figure 5-17 (A and B) Docking result with second lowest χ_r^2 of conformer *(ad)*$_a$ with distances measured at 0.1 mM $MgCl_2$ (colored) overlaid with 100 structures (grey transparent) indicating the uncertainties of the helix positions and orientations resulting from bootstrapping and with best solution for 20 mM $MgCl_2$ (black).

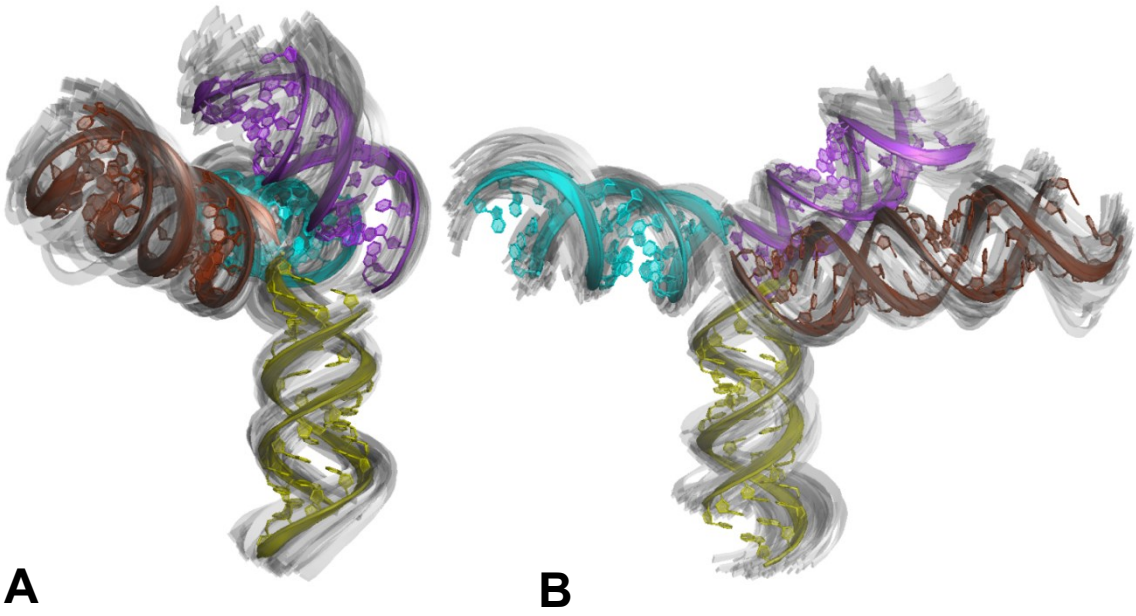


Figure 5-18 (A and B) Docking result with lowest χ_r^2 of conformer *(ad)*$_a$ with distances measured at 0.1 mM $MgCl_2$ (colored) overlaid with 100 structures (grey transparent) indicating the uncertainties of the helix positions and orientations resulting from bootstrapping.

The solution with the second lowest χ_r^2 exhibits a clear *(ad)*$_a$ conformation (RMSD over all P atoms is 7.8 Å, see Figure 5-17) while differing more than the confidence regions determined by bootstrapping (see Figure 5-18) and with angles between the helix pairs closer to 90° compared to the structure at 20 mM $MgCl_2$ (see Table 12).

Table 12 Angles between helix pairs for rigid body model with second lowest χ_r^2 of conformer *(ad)*$_a$ with distances measured at 0.1 mM $MgCl_2$ and best solution for 20 mM $MgCl_2$.

	20 mM $MgCl_2$	0.1 mM $MgCl_2$
helix pair	angle [°]	
ab	70	73
ac	136	119
ad	159	142
bc	139	129
bd	101	80
cd	63	98

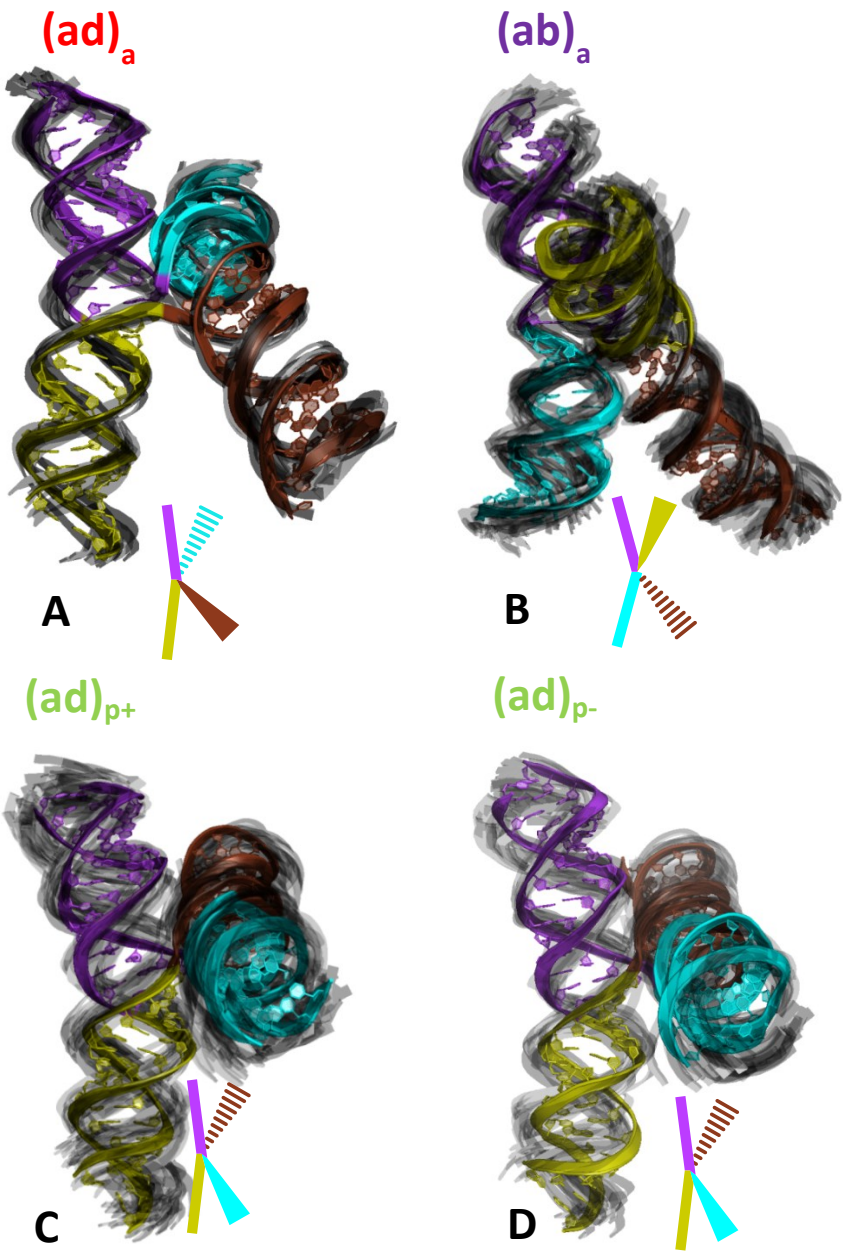
Noteworthy is that the solution with the lowest χ_r^2 (see Figure 5-18) resembles a “rough” quasi-mirrored image of the second “best” one with clearly distorted helix stacking. Whereas the second “best” solution looks very similar to the one, found with 20mM Mg. Possible reason for this can be the instability in the rigid body docking due to an insufficient amount of distance restraints.

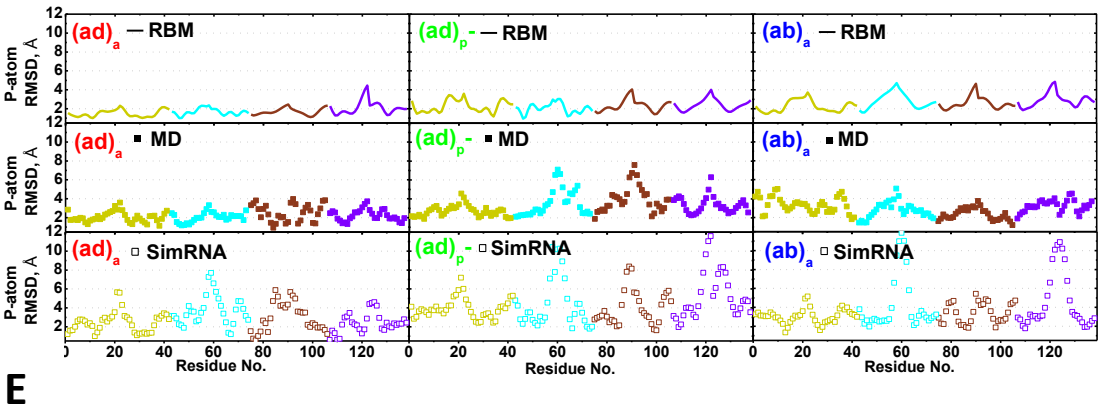
5.5.4.11 **Simulations**

We applied two alternative approaches to further refine the three RBD models by computer simulations using (i) all-atom molecular dynamics (MD) simulations in explicit solvent for a high-quality structural refinement in the vicinity of the RBD starting structure [78] (section 5.4.12) and as an alternative (ii) the coarse grained modeling program SimRNA [87] for fast *de novo* folding of the RNA four-way junction to cover a larger conformational space (section 5.4.13).

5.5.4.11.1 Trajectories of the constrained MD simulations

During MD simulations, positional restraints of the RBD model were applied to the outer helical regions for faster conformational sampling in the vicinity of the starting structures. Figure 5-19A-D show ensembles of structures with a good agreement with FRET data for conformers $(ad)_a$, $(ad)_{p+}$, $(ad)_{p-}$ and $(ab)_a$, . The resulting position uncertainties (r.m.s.d) for all P-atoms are plotted as squares in Figure 5-19E. The average precision of the phosphorus atoms are 2.3 Å, 2.9 Å, 3.3 Å and 2.4 Å, for conformers $(ad)_a$, $(ad)_{p+}$, $(ad)_{p-}$ and $(ab)_a$, respectively. The models of the J(abcd) from FRET-filtered molecular dynamics simulations have both the global geometry consistent with the FRET restraints and the local stereochemistry encoded in the MD force field. In particular, all bases at the junction are properly stacked after MD-refinement. The optimization of the local structure is achieved without significant violations of the global geometry as judged by comparison to respective rigid body models (will be discussed below).

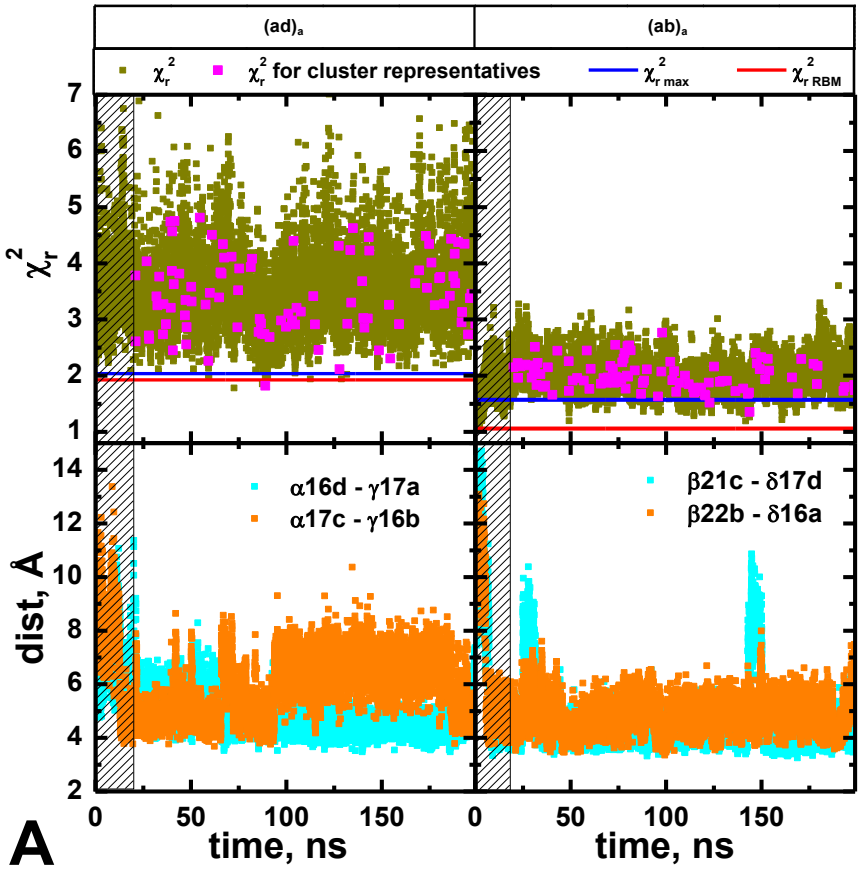

(ad)a
(ab)a
A
B
(ad)p+
(ad)p-
C
D


(ad)a — RBM
(ad)p- — RBM
(ab)a — RBM
(ad)a ▪ MD
(ad)p- ▪ MD
(ab)a ▪ MD
(ad)a □ SimRNA
(ad)p- □ SimRNA
(ab)a □ SimRNA
P-atom RMSD, Å
Residue No.

E

Figure 5-19 (A-D) Ensembles of the MD refined structures (cartoon representation) with good agreement with FRET data for conformers *(ad)*$_a$ (A), *(ab)*$_a$ (B) *(ad)*$_{p+}$ (C) and *(ad)*$_{p-}$ (D). For the respective structures with the best agreement (χ_r^2 = 1.78 for *(ad)*$_a$, 1.33 for *(ad)*$_{p+}$, 1.47 for *(ad)*$_{p-}$ and 1.19 for *(ab)*$_a$) helices *a*, *b*, *c* and *d* are colored purple, cyan, brown and yellow, respectively. Respective sketches depicting the mutual orientation of the helices are shown at the bottom of the structures. (E) Uncertainty of phosphorus atom positions for RBD (lines), MD refined models (filled squares) and SimRNA models (section 5.5.4.11.2).

To find the structures with best agreement to FRET data χ^2_r and the distances between the bases at the junction which belong to the same stacked helix pair but to different single strands are plotted over time (see Figure **5-20**). The small distance between them indicates proper base stacking. For each simulation the first 20 ns of each trajectory were disregarded in which the starting structure (rigid body model) equilibrate. For all three conformers the simulations yield structures with good agreement with FRET data and with similar χ_r^2 - values compared to the rigid body models (see Table 13). Structures were considered to have a good agreement with FRET if they have χ_r^2 - values below a $\chi^2_{r,\max}$ threshold (see Figure 5-20 and section 5.4.10). For calculation of $\chi^2_{r,\max}$ the respective cluster representatives with the best agreement with FRET are considered. Comparing the trajectories for $(ad)_{p+}$ and $(ad)_{p-}$, those for $(ad)_p^-$ yield a cluster representative with the better agreement with FRET than the trajectory for $(ad)_{p+}$. Therefore, $\chi^2_{r,\max}$ for $(ad)_p^-$ is used for $(ad)_{p+}$ and $(ad)_{p-}$. The trajectories of conformers $(ad)_a$, $(ad)_{p+}$, $(ad)_{p-}$ and $(ab)_a$ yield 11, 99, 356 and 392 structures which fulfill this criterion, respectively (see Figure 5-21,Figure 5-22).


(ad)a
(ab)a
χr2
χr2 for cluster representatives
χr2 max
χr2 RBM
χr2
dist, Å
α16d - γ17a
α17c - γ16b
β21c - δ17d
β22b - δ16a
time, ns
time, ns
A

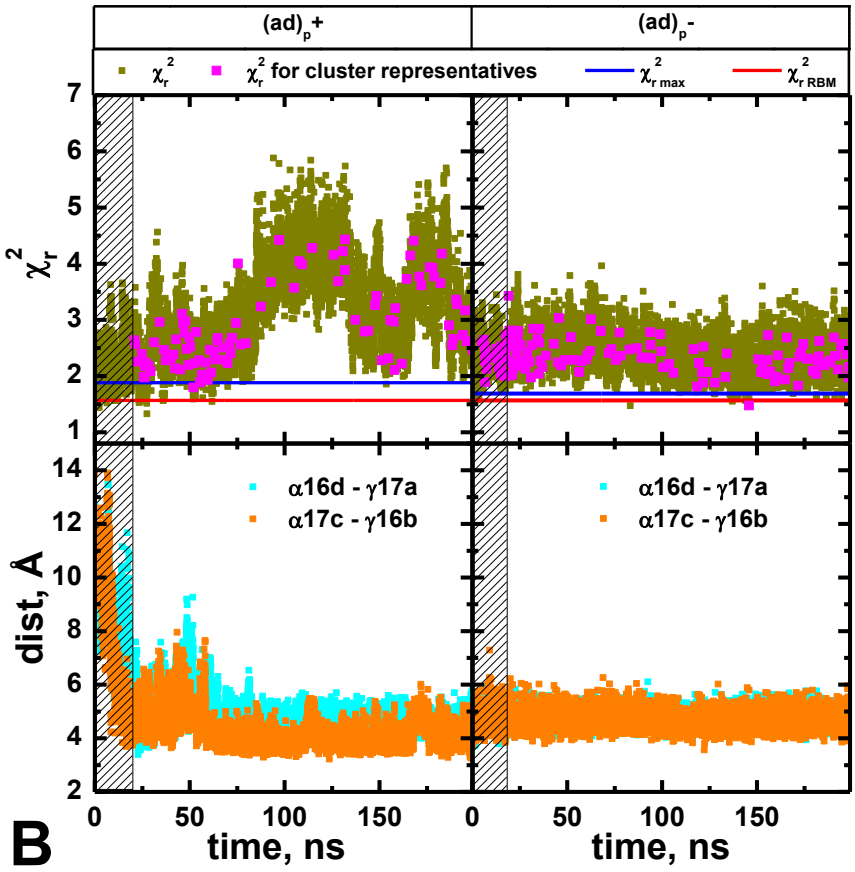

(ad)p+
(ad)p-
χr2
χr2 for cluster representatives
χr2 max
χr2 RBM
χr2
dist, Å
α16d - γ17a
α17c - γ16b
α16d - γ17a
α17c - γ16b
time, ns
time, ns
B

Figure 5-20 Trajectories of the MD simulations for (A) conformers $(ad)_a$ (left) and $(ab)_a$ (right) and (B) conformers $(ad)_{p+}$ (left) and $(ad)_{p-}$ (right). Lower panels: Distances between bases α16d and γ17a for conformers $(ad)_a$ and $(ad)_p$ and between β21c and δ17d for conformer $(ab)_a$ (cyan squares) and between bases α17c and γ16b for conformers $(ad)_a$ and $(ad)_p$ and between β22b and δ16a for conformer $(ab)_a$ (orange squares) plotted over time. Upper panels: χ^2_r values indicating the agreement with FRET data plotted over time for every member of the trajectory (dark yellow squares) and for the cluster representatives (magenta squares). The structures modeled during the first 20 ns of the trajectories (black meshes) are disregarded. The red horizontal lines indicate the χ^2_r values of the rigid body models. The blue horizontal lines indicate a $\chi^2_{r,max}$ value which is calculated via $\chi^2_{r,\max} = \chi^2_{r,\min} + [2/(n-p)]^{1/2}$ where *n* and *p* are number of distance constraints (51) the degrees of freedom (9), respectively (see section 5.4.10) and $\chi^2_{r,min}$.is the reduced chi-squared of the cluster representative with best agreement with FRET data.

To ensure that the MD trajectories explore a conformal space comparable to the one determined by the bootstrapping procedure for the rigid body model (see section 5.4.11) the mean of root mean square fluctuation (RMSF) over all constrained P atoms during the MD trajectory disregarding the first 10 ns is compared to the mean of root mean square deviation (RMSD) over the same P atoms resulting from the bootstrapping procedure of the rigid body model (see Table 13). Table 13 also shows the χ^2_r values of the MD and rigid body models with best agreement with FRET data and the deviations (RMSD) between them over all constrained P atoms.

Table 13 Comparison of rigid body models (RBM) and MD cluster representatives for the three RNA four-way junction conformers. Ratio between mean root mean square fluctuation (RMSF) over all restrained P atoms during the MD trajectories disregarding the first 20 ns and the mean root mean square deviation (RMSD) over the same P atoms resulting from the bootstrapping procedure of the rigid body model. χ^2_r values of the rigid body models and the MD cluster representatives with best agreement with FRET data ($\chi^2_{r,min}$), respectively. RMSDs over those P atoms which are restrained during the MD simulations between the rigid body models and the MD cluster representatives with best agreement with FRET data, respectively

	$(ad)_a$	$(ad)_{p+}$	$(ad)_{p-}$	$(ab)_a$
$\langle RMSF \rangle_{MD} / \langle RMSD \rangle_{RBD}$	1.33	1.49	1.44	0.98
$\chi^2_{r,min}$ RBD	1.93	1.57	-	1.06
$\chi^2_{r,min}$ MD	1.82	1.80	1.48	1.35
RMSD rigid body model vs MD model	3.8 Å	6.1 Å	6.02Å	3.5 Å

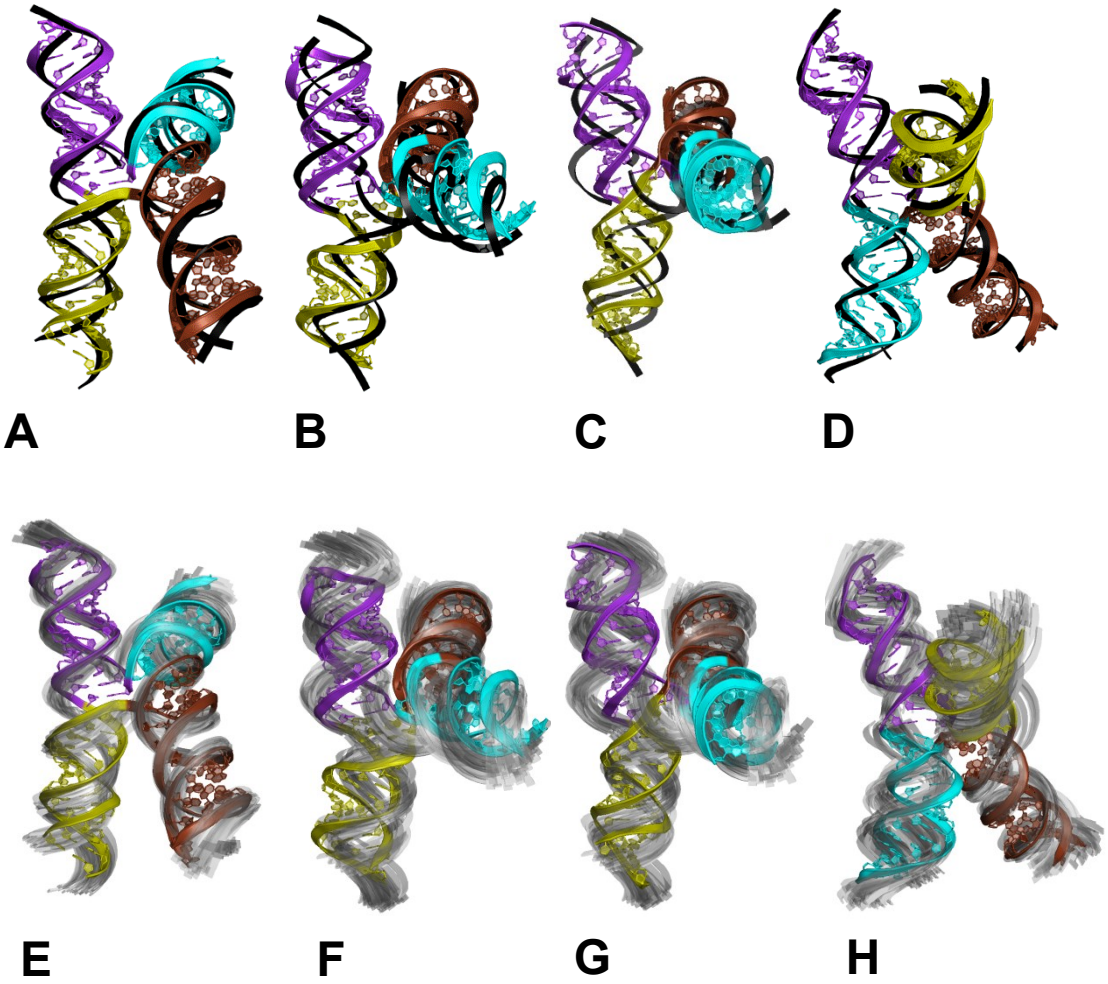


Figure 5-21 Structures (cartoon representation) of the MD refined models (cluster representatives) with the lowest χ_r^2 for conformer *(ad)*$_a$ (A and E), *(ad)*$_{p+}$ (B and F), *(ad)*$_{p-}$ (C and G) and *(ab)*$_a$ (D and H) overlaid with the respective rigid body models with lowest χ_r^2 (black, A - D) and with 100 structures (E – H, grey transparent) indicating the uncertainties of the helix positions and orientations of the rigid body models resulting from bootstrapping (see section 5.4.11).

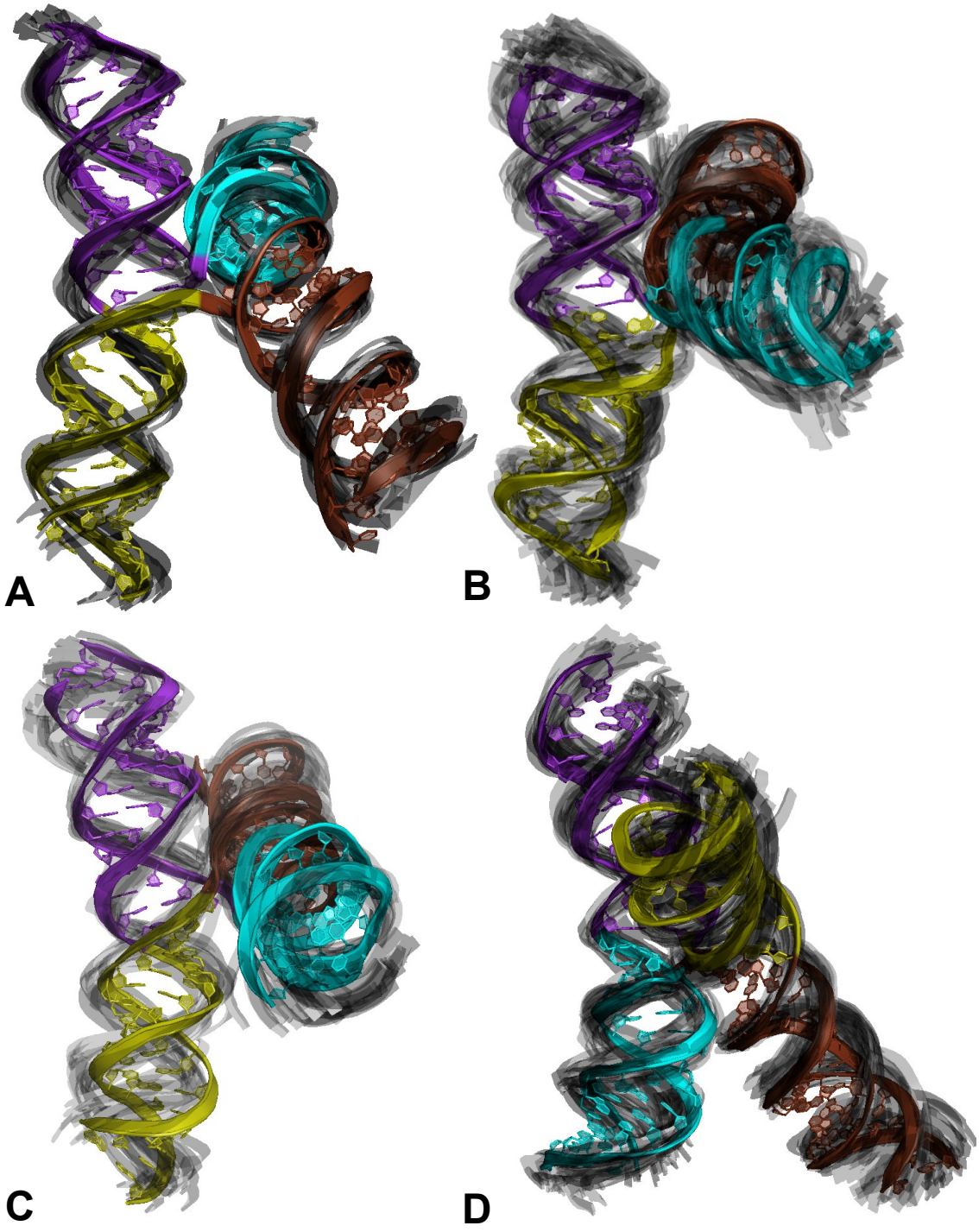


Figure 5-22 Structures from the constrained MD trajectories with acceptable agreement with FRET data (see Figure 5-20) for conformers $(ad)_a$ (A, 11 structures), $(ad)_{p+}$ (B, 99 structures), $(ad)_{p-}$ (C, 34 structures) and $(ab)_a$ (D, 392 structures). The mean RMSDs (over all P atoms) versus respective cluster representative with the best agreement with FRET (colored representation) are 2.2 Å for $(ad)_a$, 3.4 Å for $(ad)_{p+}$, 3.0 Å for $(ad)_{p-}$ and 2.9 Å for $(ab)_a$.

5.5.4.11.2 SimRNA

To investigate the formation and the equilibrium geometry of the RNA junction, two independent computational strategies were used, all utilizing the SimRNA program (software and details of simulations are described in section 5.4.13). The difference between these strategies lies in the selection of the restraints imposed on the simulation.

In the first strategy, no FRET data were used. The simulation started from isolated RNA chains and used secondary structure restraints only. 140 runs of Replica Exchange Monte Carlo method were performed (nine replicas, 16 000 000 iterations in each simulations). Models were selected from these simulations by combination of energy criteria and conformational clustering (see section 5.4.13 for details).

Using solely secondary structure restraints as initial knowledge, prior to any experimental data (strategy 1, see section 5.4.13), independent SimRNA simulations generated lowest-energy structures with topologies corresponding to all four possible conformers. 41 simulations out of 140 independent simulations converged into models with correctly folded secondary structure and were analyzed further.

Out of the six theoretically possible conformers of the J(abcd), four emerged in the unrestrained simulations: $(ad)_{a,}$ $(ad)_{p-}$, $(ab)_{a}$, and $(ab)_{p-}$. Two remaining conformers $(ad)_{p+}$ and $(ab)_{p+}$ that are theoretically possible, did not emerge at the final stage of any unrestrained simulations. The major conformer $(ad)_{a}$ did appear as the lowest energy structures in 20 out of 41 above-mentioned simulations. The two minor conformers $(ab)_{a}$ and $(ad)_{p-}$ did appear in 9 and 8 simulation runs. Theoretically possible, but experimentally unobserved $(ab)_{p-}$ conformer is also the least abundant one in the SimRNA simulations and appeared in 4 simulations only. For the three conformers observed in FRET measurements [$(ad)_{a}$, $(ab)_{a}$ and $(ad)_{p-}$], RMSD values were calculated between the structures from the unconstrained simulations and the structures refined using MD with FRET restraints. The median RMSD values obtained with that procedure are: 15.68 Å, 20.04 Å, and 13.97 Å for conformers $(ad)_{a}$, $(ab)_{a}$, and $(ad)_{p-}$, respectively.

The second strategy was very similar to the first one, and additionally included restraints from FRET experiments. In the preliminary simulations employing FRET restraints, the deviation of the helical arms of the J(abcd) from the ideal RNA A-helix due to distortions caused by “pulling” forces of the restraints was observed. In order to reduce these artificial

effects without influencing the dynamics of the junction, an additional set of distance restraints have been imposed on the backbone in regions predicted to form helices, outside of the immediate vicinity of the junction region. Hence, simulations with FRET data included three types of restraints: on base pairing, those from FRET, and restraints rigidifying the RNA helices. Here, nine independent simulations were performed for each set of FRET restraints corresponding to one of the three conformers $(ad)_{p+-}$, $(ab)_a$ and $(ad)_a$. These simulations have led to conformations characteristic for each of these conformers; the resulting models were very similar to those obtained by rigid body modeling followed by MD optimization (Table 14).

Table 14 Parameters of SimRNA models of the 3 conformers of the J(abcd). The values of the SimRNA energy are given both with and without the penalty for violation of the restraints (which include all three types of restraints used). RMSDs have been calculated for all atoms present in the SimRNA representation (see text).

Conformer	Weight of the FRET restraints	SimRNA energy including restraints (internal units)	SimRNA energy without restraints (internal units)	RMSD vs rigid body models [Å]	RMSD vs MD models [Å]
$(ad)_a$	0.1	-1551.617	-1569.517	4.58	3.28
	0.2	-1546.654	-1566.621	4.41	3.86
	0.5	-1542.210	-1556.418	3.01	3.65
$(ad)_{p-}$	0.1	-1539.210	-1552.001	5.07	5.54
	0.2	-1541.547	-1554.714	3.49	4.12
	0.5	-1540.894	-1556.402	3.81	4.59
$(ab)_a$	0.1	-1546.829	-1567.670	5.33	4.50
	0.2	-1544.277	-1565.470	4.71	5.80
	0.5	-1541.672	-1566.479	3.81	4.44

The modeling of J(abcd) by Monte Carlo dynamics, with and without FRET restraints allowed not only the determination of the structure of the three coexisting conformers of the RNA

four-way junction molecule, but also numerical observation of the transitions between conformers. Simulations of transitions between the folded forms were performed by starting from the structure of each conformer and imposing the set of FRET restraints from each of the other conformers. Nine independent isothermal simulations of 16 000 000 iterations each were performed for 6 possible transitions, at each of the following temperatures: 0.55, 0.60, 0.65, 0.70, 0.75 and 0.80. The weight of the FRET restraints used was set to 0.2 (increased from 0.1 used for folding, and equal to the one used for refinement). As a result, 5 out of 6 possible transitions between conformers were reproduced. Analysis of the pathways of all transitions revealed intermediate structures in which destacking of helices occurred.

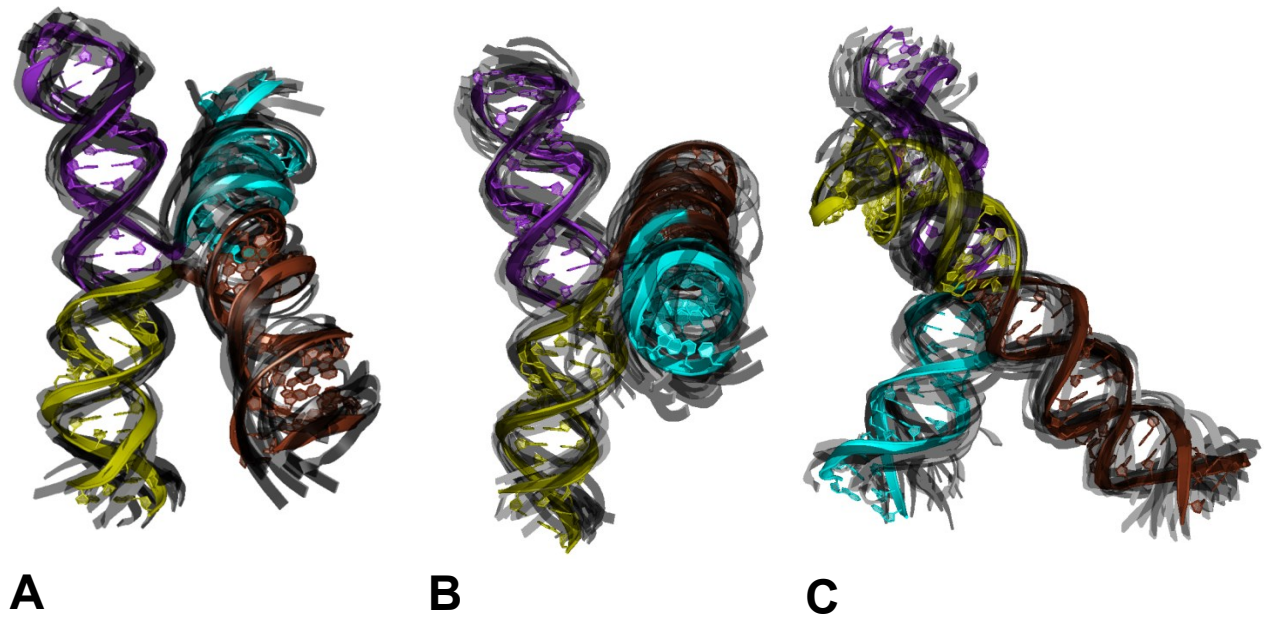


Figure 5-23 Structures obtained by coarse-grained modeling with SimRNA (in color, cartoon representation) superimposed onto conformations obtained by rigid-body docking (in gray); for *(ad)*$_a$ (A, 11 structures from rigid-body docking), *(ad)*$_p$ (B, 4 structures from rigid-body docking with *(ad)*$_{p+}$ junction topology, 12 structures from rigid-body docking with *(ad)*$_{p-}$ junction topology) and *(ab)*$_a$ (C, 17 structures from rigid-body docking). The mean RMSDs (over all P atoms) vs. the respective structure with the best agreement with FRET (colored representation, χ_r^2 = 1.68, 1.54 and 1.04 for *(ad)*$_a$, *(ad)*$_p$ and *(ab)*$_a$, respectively) are 3.9 Å for (ad)$_a$, 3.5 Å for (ad)$_p$ and 4.0 Å for (ab)$_a$.

Structural comparison (RMSD overall phosphorus atoms) of the models found with RBD (section 5.5.4.8), SimRNA (section 5.5.4.11.2) and SAXS (section 5.5.4.11.5) techniques versus MD cluster centroid with best agreement with the FRET data (section 5.5.4.11.1) is presented in Table 15.

Table 15 Summary of RMSD values.

	RMSD over all P atoms vs best MD cluster centroid, [Å]			
	$(ad)_a$	$(ad)_{p+}$	$(ad)_{p-}$	$(ab)_a$
Model found by Rigid Body Docking	3.8	6.3	7.6	3.6
MD model within error limit with largest deviation from the centroid	3.8	6.8	5.0	4.7
Best according to SAXS out of MD ensembles $(ad)_{p+}$	2.5	4.6	-	3.0
Best according to SAXS out of MD ensembles $(ad)_{p-}$	2.6	-	4.6	4.4
simRNA (cluster centroids)	4.6	6.9	6.2	5.1

5.5.4.11.3 Comparison of MD and crystallographic models of RNA four-way junctions

Models of all three conformers of the J(abcd) obtained using FRET-driven molecular modeling can be compared with known crystallographic and NMR structures of other four-way junctions. As of September 2013, according to the RNA Bricks database [104] there are 4370 four-way junctions, which can be identified in the RNA structures deposited in the PDB database. Among them 150 (142 according to RNA Frabase [105]) do not contain internal loops in the junction region, so that they have secondary structure that is equivalent to the junction studied in this work. This smaller subset can be further clustered into 9 non-redundant classes based on the sequence and geometry of the junction (see Table 16). Single representatives of each class were selected and their structures were compared with each other and with the structural models for the J(abcd) studied in this work (see Table 16, the selected structures are highlighted in color in

Table 17). For the structural comparison of the structures RMSD between P, C4' and N1/N9 (for pyrimidines/purines respectively) from all residues in the junction region up to 3 residues away from the junction along each of the chains (24 residues in total) are used. This allows one to calculate the RMSD between structures having non-identical sequences or secondary structures.

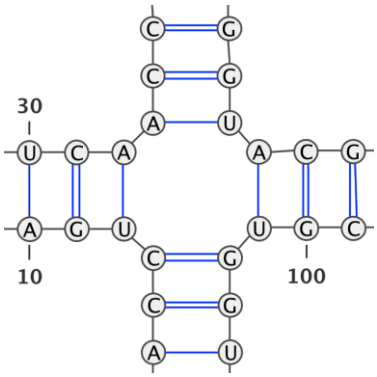
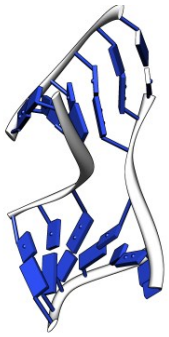
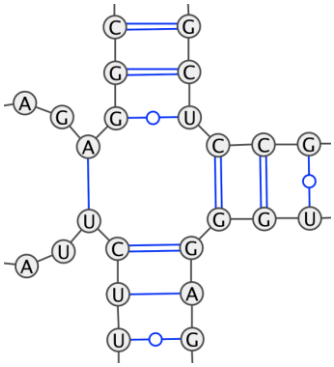
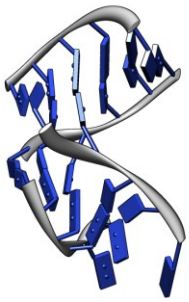
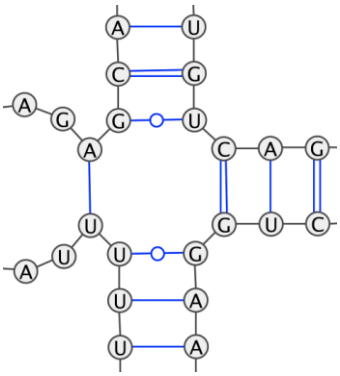
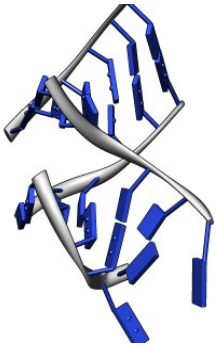
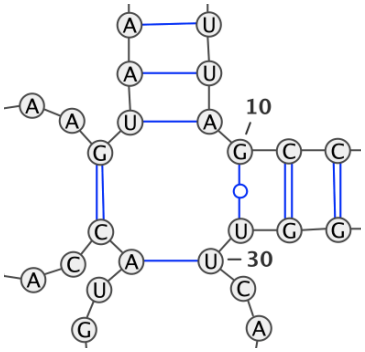
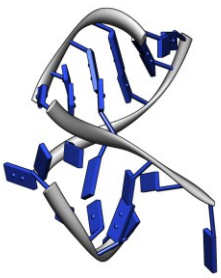
Structural clusters in

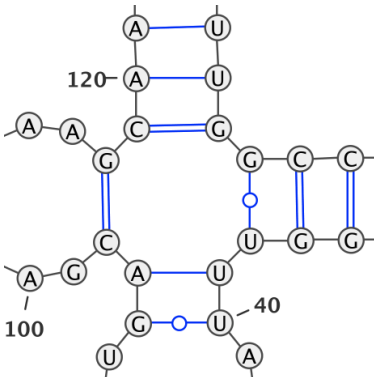
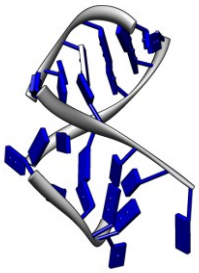
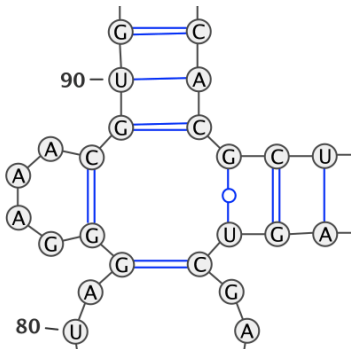
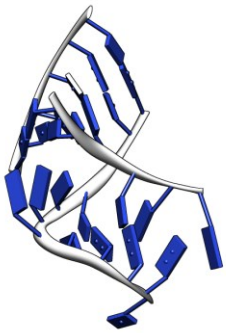
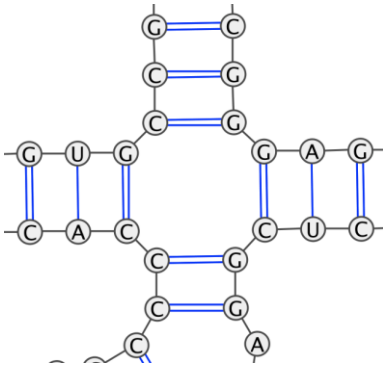
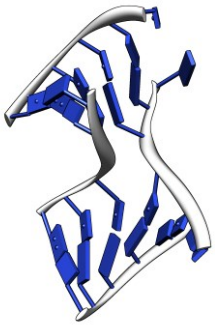
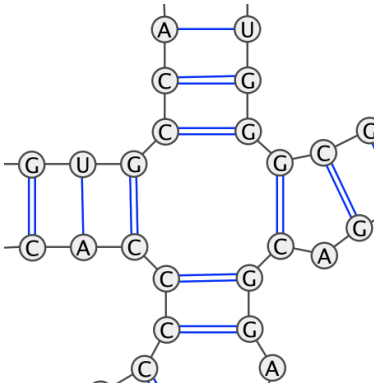
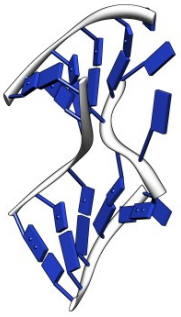
Table 17 are highlighted by red and green shaded boxes, grouping classes of structures in parallel and anti-parallel conformations respectively. Five of the classes contain highly conserved four-way junctions from large ribosomal RNA (rRNA) from various organisms. What makes the rRNA junctions different from the one studied here, is that they do not have four fully developed helices and one of these helices is reduced to a single Watson–Crick base-pair followed by non-canonical base-pairs. These junctions exhibit the parallel topology (cH family in the nomenclature of [106], H corresponds to antiparallel family), and their geometry is similar to the $(ad)_{p-}$ and not to that of any other conformers. It must be emphasized that the $(ad)_{p+}$ conformer [which would be very close geometrically, but not topologically, to $(ad)_{p-}$] was not observed in any of the experimentally determined structures.

The non-rRNA junction structures available in the PDB form four non-redundant classes, originating from: the hairpin ribozyme, Ribonuclease P (*B. subtilis*, *T. maritima*) and the pseudo–knot domain of the hepatitis C virus, respectively. The conformations of all the identified four-way junctions are highly influenced by tertiary contacts formed by the junction arms. In the hairpin ribozyme and Ribonuclease P interactions between junction arms lead to asymmetry in the junction geometry, while in the junction from the hepatitis C virus IRES domain the arms are connected by a single RNA strand, forming a pseudo–knotted structure. For all four non–rRNA four–way junctions, conformations are close to the anti–parallel conformers ($(ad)_a$ or $(ab)_a$) observed in current study.

Summarizing, $(ad)_a$, $(ab)_a$ and $(ad)_{p-}$ RNA four-way junction conformers obtained by modeling with SimRNA using FRET restraints have counterparts in RNA structures determined by X-ray crystallography, while the theoretically possible conformers, for which structural models could not be obtained using that methodology, have no counterpart in RNA X-ray structures determined so far.

Table 16 PDB IDs, secondary structure, and 3D geometry of the central region of all four-way junctions present in the PDB database (deposited before September 2013).

	PDB ID and resolution	Secondary structure	Junction geometry
Hairpin ribozyme	**1M5O** (2.2Å), 1M5K (2.4Å), 1M5V (2.4Å), 1M5P (2.6Å)		
Saccharomyces cerevisiae, large ribosomal subunit	**3U5D**, 3U5H (3Å), 3O58, 3O5H (4Å)		
Tetrahymena thermophila, large ribosomal subunit	**4A18**, 4A19, 4A1B, 4A1D (3.52Å)		
Spinacia oleracea, large ribosomal subunit	**3BBO** (9.4Å)		

Escherichia coli, large ribosomal subunit	3IZT, 3IZU, **3R8S**, 3R8T (3Å), 3OFQ, 3OFR (3.1Å), 3I1N, 3I1P, 3OFC, 3OFD (3.19Å), 3SGF (3.2Å), 2QAM, 2QAO (3.21Å), 2I2T, 2I2V (3.22Å), 3OAS, 3OAT (3.25Å), 3OFZ (3.29Å), 3OG0 (3.29Å), 1VT2, 2QBE, 2QBG, 3ORB, 4GAR, 4GAU (3.3Å), 1VS6, 1VS8 (3.46Å), 2AW4, 2AWB (3.46Å), 2QOZ, 2QP1 (3.5Å), 3DF2, 3DF4 (3.5Å), 2QBA, 2QBC (3.54Å), 3UOS (3.7Å), 3I20, 3I22 (3.71Å), 3I1R, 3I1T (3.81Å), 2QOV, 2QOX (3.93Å), 2QBI, 2QBK (4Å), 2Z4L, 2Z4N (4.45Å), 2WWQ (5.8Å), 3FIK (6.7Å), 3J01 (7.1Å), 2J28 (8Å), 3KCR (9.6Å), 3BBX (10Å)		
Haloarcula marismortui, large ribosomal subunit	**1VQ8**, 1VQO (2.2Å), 1VQK (2.3Å), 1JJ2, 1S72, 1VQ9, 1VQN, 1YHQ, 3CC2 (2.4Å), 3CCM (2.55Å), 1YIJ (2.6Å), 1YI2 (2.65Å), 1VQ6, 2OTL, 3CC4, 3CC7, 3CCJ, 3CPW, 3G6E, 3OW2 (2.7Å), 3CCE, 3CD6 (2.75Å), 1M90, 1YIT, 3CCU, 3CMA (2.8Å), 3G71 (2.85Å), 1QVG, 1YJ9, 1YJW, 2OTJ, 2QEX, 3CCL, 3CCQ, 3CCV, 3I56 (2.9Å), 1Q81, 3CCS, 3CME (2.95Å), 1Q82 (2.98Å), 1K8A, 1K9M, 1KD1, 1N8R, 1NJI, 1Q86, 1YJN, 2QA4, 3CCR, 3CXC, 1K73 (3Å), 1KQS, 1QVF (3.1Å), 1M1K, 1Q7Y (3.2Å), 1W2B (3.5Å)		
Thermotoga maritima, Ribonuclease P	**3OK7**, 3Q1Q (3.8Å), 3OKB (4.21Å), 3Q1R (4.21Å)		
Bacillus subtilis, Ribonuclease P	**1U9S** (2.9Å), 1NBS (3.15Å)		

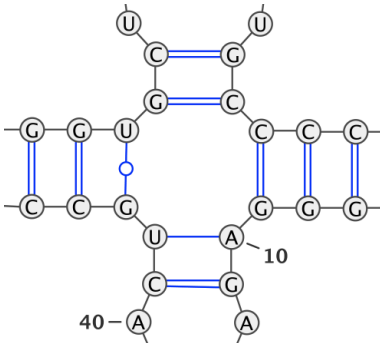
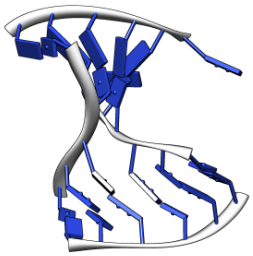
Hepatitis C virus IRES Pseudo–knot domain	**3T4B** (3.55Å)		

Table 17 RMSD in Å between RNA four-way junctions from the PDB (representatives of clusters from Table 4-11) and the obtained models of four-way junctions. Colored regions group known junctions with mutual RMSD (P, C4' and N1/N9 atoms from 24 residues in the junction center, see text) smaller than 5 Å. The red shaded cluster groups junctions correspond to parallel (cH) conformation and the green shaded one groups those in antiparallel (H) one. Lowest RMSD values for comparison between FRET-based models of the RNA four-way junction studied in this work and crystallographic models of other RNA four-way junctions are shown in bold.

		RNA four-way junctions from PDB									This work – SimRNA				This work – MD			
		1M5O	3OK7	1U9S	3T4B	3U5D	4A18	3BBO	3R8S	1VQ8	$(ad)_a$	$(ab)_a$	$(ad)_{p+}$	$(ad)_{p-}$	$(ad)_a$	$(ab)_a$	$(ad)_{p+}$	$(ad)_{p-}$
RNA four-way junctions fr omPDB	1M5O		2.23	3.39	5.35	12.41	12.45	12.40	12.27	12.04	**3.96**	4.93	8.66	9.88	**3.12**	7.07	8.53	10.33
	3OK7	2.23		2.33	3.95	11.49	11.50	11.32	11.26	11.45	5.21	**3.60**	7.59	8.94	**4.24**	5.68	7.44	9.29
	1U9S	3.39	2.33		4.56	11.41	11.45	10.98	10.83	11.10	5.23	**3.13**	7.63	8.86	**4.30**	5.09	7.63	9.19
	3T4B	5.35	3.95	4.56		10.35	10.38	9.42	9.36	9.62	8.10	**3.23**	6.16	7.13	6.96	**4.81**	5.11	7.42
	3U5D	12.41	11.49	11.41	10.35		0.97	2.32	2.39	3.11	10.32	11.11	9.15	**5.63**	10.90	10.67	7.96	**4.68**
	4A18	12.45	11.50	11.45	10.38	0.97		2.51	2.45	3.12	10.38	11.13	9.12	**5.65**	10.92	10.75	7.95	**4.64**
	3BBO	12.40	11.32	10.98	9.42	2.32	2.51		0.84	3.04	11.09	10.34	8.87	**5.27**	11.55	9.75	7.41	**4.28**
	3R8S	12.27	11.26	10.83	9.36	2.39	2.45	0.84		2.90	10.94	10.23	8.78	**5.15**	11.39	9.71	7.28	**4.14**
	1VQ8	12.04	11.45	11.10	9.62	3.11	3.12	3.04	2.90		10.30	10.58	8.61	**4.76**	10.57	10.52	7.00	**4.27**
This work	$(ad)_a$	**3.96**	5.21	5.23	8.10	10.32	10.38	11.09	10.94	10.30		7.11	10.50	12.09	3.35	9.01	10.89	12.31
	$(ab)_a$	4.93	**3.60**	**3.13**	**3.23**	11.11	11.13	10.34	10.23	10.58	7.11		7.06	8.41	5.63	3.21	6.82	8.56
	$(ad)_p^+$	8.66	7.59	7.63	6.16	9.15	9.12	8.87	8.78	8.61	10.50	7.06		6.61	10.18	8.07	4.30	7.64
	$(ad)_p^-$	9.88	8.94	8.86	7.13	**5.63**	**5.65**	**5.27**	**5.15**	**4.76**	12.09	8.41	6.61		11.73	9.35	3.49	2.28
	$(ad)_a$	**3.12**	**4.24**	**4.30**	6.96	10.90	10.92	11.55	11.39	10.57	3.35	5.63	10.18	11.73		7.28	10.42	12.11
	$(ab)_a$	7.07	5.68	5.09	**4.81**	10.67	10.75	9.75	9.71	10.52	9.01	3.21	8.07	9.35	7.28		8.04	9.03
	$(ad)_p^+$	8.53	7.44	7.63	5.11	7.96	7.95	7.41	7.28	7.00	10.89	6.82	4.30	3.49	10.42	8.04		4.67
	$(ad)_p^-$	10.33	9.29	9.19	7.42	**4.68**	**4.64**	**4.28**	**4.14**	**4.27**	12.31	8.56	7.64	2.28	12.11	9.03	4.67	

5.5.4.11.4 Comparison of the $(ad)_a$ conformer with the crystal structure of the hairpin ribozyme

The conformation of the junction region of the $(ad)_a$ conformer is similar to the one in the crystal structure of the hairpin ribozyme (PDB-ID 1M5K, see Figure 5-1A). Figure 5-24 presents the superposition of the structure of the $(ad)_a$ conformer from the constrained MD-trajectory with the best agreement to FRET data and analogous structure generated based on the junction region of the hairpin ribozyme (PDB: 1M5K, chains AB). In order to enable the comparison, the central region of 1M5K was extracted, and its helices were extended in ideal A-helix geometry with sequence identical to that of the RNA four-way junction studied in this work; modeling was performed with the ModeRNA modeling method [107]. The presented superposition of the $(ad)_a$ conformer from FRET-guided modeling and an idealized version of the hairpin ribozyme four-way junction minimizes RMSD between all phosphorus atoms. The similarity of the coaxial stacking between both structures is clearly visible. However, the global geometries of both junctions, including the angles between helical arms are different. The structure of the $(ad)_a$ conformer reconstructed based on FRET data has an approximate D_2 symmetry. The minimal RMSD computed on the backbone atoms between the obtained models and models rotated around each of the three approximate C_2 axes are: 11.9 Å, 12.0 Å, 4.3 Å. In contrast, the structure reconstructed from the four-way junction in the hairpin ribozyme has an approximate C_2 symmetry (the minimal RMSD values between models and rotated ones around the three axes are: 16.4 Å, 13.5 Å, 9.4 Å). This difference can be explained by additional tertiary contacts formed in the active center of the ribozyme, which are absent in the molecule studied in this work (see Figure 5-1A).

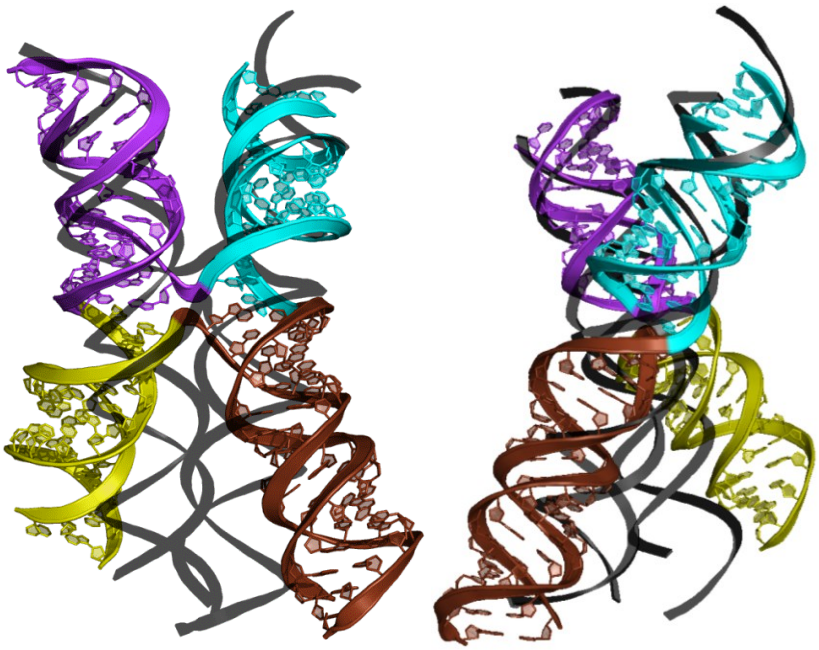


Figure 5-24 Comparison of structures of the major conformers $(ad)_a$ of J(abcd) (colored) with one reconstructed from the crystal structure (PDB-ID: 1M5K) of the hairpin ribozyme.

In conclusion, RNA folding simulations with SimRNA were capable of identifying the FRET-supported topologies with restraints on secondary structure, without other experimental data; however the inclusion of FRET-derived restraints drives the RNA modeling toward unequivocal determination of conformations that agree with structural models determined by a combination of rigid-body modeling with FRET data, followed by MD refinement.

5.5.4.11.5 SAXS

As an independent test of the FRET restrained modeling, small-angle X-ray scattering (SAXS) data of the RNA four-way junction construct was obtained (see Figure 5-9, section 5.4.16). SAXS can probe the low and intermediate resolution structure of macromolecules in solution and has been applied previously to determine the conformations of functional RNAs and nucleic acid junctions [92, 96, 108]. Comparing the experimental scattering data to the structure with the helix orientation from the crystal structure of the hairpin ribozyme (see Figure 5-24), poor agreement was found, suggesting that this structure is not representative of the solution conformation of the J(abcd) (Figure 5-25). Next, the experimental SAXS data

was compared to FRET-restrained structures. For conditions used in the performed experiments, the scattering intensity of a mixture of different species is simply the sum of the individual scattering contributions weighted by their relative populations.

However, with current data it is not possible to select structures from an ensemble of structures for each conformation and at the same time determine the relative population of the conformations. Therefore, all trajectories have been clustered to get 100 clusters. Out of each cluster one structure was selected which best represented a centroid of its respective cluster. SAXS data was screened according to FRET restraints, versus all possible combinations (with correct weighting 71 %, 15 % and 14 % for $(ad)_a$, $(ab)_a$, and $(ad)_p$-, respectively) of the 3 x 100 centroids.

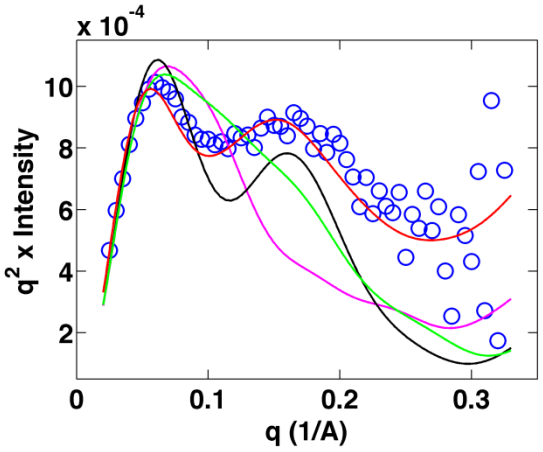


Figure 5-25 Results of cluster screening by SAXS. Blue, green, red, magenta and black curves represent experimental data, chrystal structure derived model, based combination of profiles from the "1 sigma" ensembles, using 70-20-10 weights, H-shape and X-shape, respectively.

Many combinations of structures from these FRET-derived ensembles provide good fits to the experimental data, indicating that the FRET restraints bias the simulations towards conformations that are populated in solution. It was observed that the fit improves if a higher weight is assigned to the minor component $(ab)_a$ than what was determined by FRET. These small differences might be caused by imperfections in the calculations of the scattering profiles, e.g. due to the neglect of ions scattering, or by differences in the experimental conditions between FRET and SAXS measurements or other experimental errors. Overall, the SAXS data are in good agreement with the FRET predictions and confirm

the power of the FRET-derived restraints to bias the ensembles toward the structures that are populated in solution.

5.5.4.12 **Conclusions and outlook**

FRET restrained high-precision structural modeling allowed establishing highly accurate structural models of three coexisting conformers of the RNA four-way junction. Even though distance uncertainties for the minor populations were often significantly larger compared to the major conformer, the precisions of the minor models are only slightly worse. In comparison to NMR spectroscopy larger RNAs can be studied.

Due to dynamic behavior and the resulting structural heterogeneity of the J(abcd), its structure determination by traditional structural biology approaches is very difficult.

Making use of the advantage, three conformers were structurally resolves in parallel by FRET including the transient minor conformers. Additionally, both FRET-restrained molecular dynamics simulations and coarse-grained structure predictions provided meaningful and consistent all-atom models, which additionally were proven by SAXS. For the first time, the structure of an J(abcd) was determined in a state uninfluenced by, e.g., interactions between the arms or with other macromolecules.

As demonstrated in this work, experimental restraints from FRET can be used in computational modeling of RNA structures in post-filtering of, e.g., MD-derived ensembles, or in restrained folding simulations e.g. with SimRNA. As a next step, AV modeling could be implemented into a computational method. This work demonstrates that FRET restraints can guide the macromolecular folding simulations to identify unambiguous solutions that are essentially identical to those obtained with the rigid body assumptions. Moreover, FRET-restrained high-precision structural modeling is also applicable to structurally heterogeneous and flexible proteins, whose overall structures are notoriously difficult to determine. Together with filtered fluorescence correlation spectroscopy [109], FRET harbors the potential to study conformational control of biomolecular function in complex systems within a nanosecond to minute time range and associate it to detailed dynamic structures without spatial averaging.

The procedure described in this work can be used to determine structures of large structured RNAs that exhibit a dynamic behavior and cannot be determined with X-ray crystallography or NMR. It is compatible with different methods for modeling, including rigid-body assembly, molecular dynamics refinement, and folding of complex molecules starting from extended sequence.

5.5.5 RNA three-way junction

5.5.5.1 Fitting the data

In most of the cases one big and an additional small (typically ~ 3 – 10 %, and only in some cases > 10 %) population was required to reach a satisfactory fit quality in PDA analysis (Figure 5-10C). However, it was shown that those minor populations could be explained with the presence of not fully hybridized (incomplete, section 5.3.1) molecules in the measurement solution (section 5.5.5.4) or with acceptor photophysics (section 5.5.5.5). Noteworthy is that with or without this minor state, the distance for major state may change for not more than 0.2 Å. Hence, throughout this work only the major population will be taken into account for structural modeling. A typical example of PDA is shown in Figure 5-10C.

5.5.5.2 Distances and errors

Table 18 Values for measured distances for J(abc) molecules $\langle R_{DA}\rangle_E$, their relative amplitudes A and the values of apparent σ, measurement errors ΔR_{DA} (resulting from $\Delta R_{DA}(\kappa^2)$ and $\Delta R_{DA}(E)$, see section 5.4.6) and model distances R_{model} resulting from rigid body docking. See 7.5.1 for all data and fit plots.

#	DA-pair	σ_{app},%	$\langle R_{DA}\rangle_{E1}$, Å	A_1, %	ΔR_{DA}, %	R_{model}, Å	$\langle R_{DA}\rangle_{E2}$, Å	A_2, %
1	**(D)γ29a/(A)β14c/δα–Δd**	5.0	57.3	83.9	5.0	61.0	48.9	16.1
2	**(D)γ7b/(A)β14c/δα–Δd**	5.2	47.6	91.2	5.1	46.4	38	8.8

3	**(D)γ8b/(A)β14c/δα–Δd**	6.0	43	89.5	5.1	38.7	34.2	10.5
4	**(D)γ29a/(A)β27b/δα–Δd**	4.6	52.3	92.8	5.1	51.0	60	7.2
5	**(D)γ29a/(A)β33b/δα–Δd**	5.1	63.5	96.3	5.1	61.3	53.2	3.7
6	**(A)γ12b/(D)β11c/δα–Δd**	5.3	46.1	94.8	5.1	49.1	52.4	5.2
7	**(A)γ12b/(D)β14c/δα–Δd**	5.5	45.3	92.5	5.1	44.1	52	7.5
8	**(A)γ12b/(D)β5c/δα–Δd**	5.0	58.4	100.0	5.0	61.3		
9	**(A)γ12b/(D)β8c/δα–Δd**	5.3	53	94.1	5.1	54.9	46.5	5.9
10	**(A)γ24a/(D)β11c/δα–Δd**	4.7	54.1	100.0	5.0	55.0		
11	**(A)γ24a/(D)β14c/δα–Δd**	4.9	49	100.0	5.0	49.8		
12	**(A)γ24a/(D)β27b/δα–Δd**	4.5	44.5	84.0	5.2	44.0	51.6	16.0
13	**(A)γ24a/(D)β29b/δα–Δd**	4.7	47.2	92.2	5.0	44.1	57.6	7.8
14	**(A)γ24a/(D)β5c/δα–Δd**	4.2	60.9	93.4	5.0	61.9	52.6	6.6
15	**(A)γ24a/(D)β8c/δα_Δd**	4.2	55.3	100.0	5.0	55.8		
16	**(A)γ8b/(D)β11c/δα–Δd**	5.0	42.8	80.9	5.4	43.0	49	19.1
17	**(A)γ8b/(D)β5c/δα–Δd**	5.1	58.3	95.6	5.0	58.6	69	4.4
18	**(A)γ8b/(D)β8c/δα–Δd**	5.0	52.1	89.5	5.1	53.2	61.5	10.5
19	**(A)δα_Δd10a/(D)β11c/γ**	4.2	53.6	95.2	5.0	53.0	60.3	4.8
20	**(A)δα_Δd10a/(D)β14c/γ**	4.7	48.9	85.2	5.1	48.4	57.9	14.8
21	**(A)δα_Δd10a/(D)β27b/γ**	5.0	46.4	100.0	5.0	45.8		
22	**(A)δα_Δd10a/(D)β29b/γ**	4.6	46.7	95.5	5.0	45.4	55.2	4.5
23	**(A)δα_Δd10a/(D)β5c/γ**	5.4	59.7	90.2	5.1	58.4	51.3	9.8
24	**(A)δα_Δd10a/(D)β8c/γ**	4.1	54.2	89.6	5.0	52.5	63	10.4
25	**(A)δα_Δd10a/(D)γ7b/β**	5.3	49.2	88.6	5.0	55.2	57.4	11.4
26	**(A)δα_Δd10a/(D)γ8b/β**	5.4	45.5	98.0	5.0	47.6	55.5	2.0
27	**(D)β27b/(A)δα–Δd26c/γ**	5.0	55.4	86.2	5.4	54.6	49	13.8
28	**(D)β29b/(A)δα–Δd26c/γ**	4.4	58.7	100.0	5.1	54.2		
29	**(D)γ29a/(A)δα–Δd26c/β**	5.0	67.3	89.8	5.1	71.0	57.9	10.2
30	**(D)γ7b/(A)δα–Δd26c/β**	4.6	55.5	88.9	5.2	56.4	48.9	11.1
31	**(D)γ8b/(A)δα–Δd26c/β**	4.6	50.6	95.0	5.1	52.0	44.9	5.0
32	**(D)β27b/(A)δα–Δd28c/γ**	4.0	58.7	86.5	5.0	59.0	49.8	13.5
33	**(D)β29b/(A)δα–Δd28c/γ**	4.6	63.7	93.7	5.1	61.7	53.8	6.3
34	**(D)γ29a/(A)δα–Δd28c/β**	5.2	68.7	100.0	5.2	73.1		
35	**(D)γ7b/(A)δα–Δd28c/β**	4.0	60.9	81.7	5.3	63.2	54.9	18.3
36	**(D)γ8b/(A)δα–Δd28c/β**	3.3	54.6	88.6	5.0	56.5	49.2	11.4
37	**(A)δα_Δd7a/(D)β11c/γ**	4.3	55.7	100.0	5.0	53.6		
38	**(A)δα_Δd7a/(D)β14c/γ**	4.6	43.5	84.5	5.1	44.4	49.8	15.5
39	**(A)δα_Δd7a/(D)β27b/γ**	4.7	48.6	95.5	5.0	45.4	40.7	4.5
40	**(A)δα_Δd7a/(D)β29b/γ**	4.4	53.9	100.0	5.0	54.7		
41	**(A)δα_Δd7a/(D)β5c/γ**	4.8	57.6	100.0	5.0	55.1		
42	**(A)δα_Δd7a/(D)β8c/γ**	4.3	57.9	100.0	5.0	54.4		
43	**(A)δα_Δd7a/(D)γ7b/β**	5.2	52.6	90.6	5.2	60.1	58.7	9.4
44	**(A)δα_Δd7a/(D)γ8b/β**	4.6	46.5	92.1	5.0	47.4	51.4	7.9

Table 19 Values for J(abc(C2)).

#	DA-pair	σ_{app} ,%	$\langle R_{DA}\rangle_{E1}$, Å	A_1, %	ΔR_{DA}, %	R_{model} , Å	$\langle R_{DA}\rangle_{E2}$, Å	A_2, %
1	(D)γ29a/(A)β14c/δα(C2)–Δd	3.5	56.1	97.0	5.1	58.1	47.2	3.0
2	(D)γ7b/(A)β14c/δα(C2)–Δd	4.1	47.9	93.4	5.0	47.2	41.6	6.6
3	(D)γ8b/(A)β14c/δα(C2)–Δd	3.6	44	94.4	5.0	46.0	50.3	5.6
4	(D)γ29a/(A)β27b/δα(C2)–Δd	5.4	47.8	75.8	5.2	44.0	57.1	24.2
5	(D)γ29a/(A)β33b/δα(C2)–Δd	4.5	58	78.5	5.6	62.5	66.8	21.5
6	(A)γ12b/(D)β11c/δα(C2)–Δd	5.3	48.5	96.4	5.0	48.2	55.4	3.6
7	(A)γ12b/(D)β14c/δα(C2)–Δd	3.6	45.6	95.6	5.0	46.7	52.4	4.4
8	(A)γ12b/(D)β5c/δα(C2)–Δd	3.2	57.5	100.0	5.0	58.6		
9	(A)γ12b/(D)β8c/δα(C2)–Δd	4.6	49.7	92.0	5.0	49.4	58.9	8.0
10	(A)γ24a/(D)β11c/δα(C2)–Δd	3.1	56.5	97.5	5.0	57.1	49.1	2.5
11	(A)γ24a/(D)β14c/δα(C2)–Δd	3.2	48.1	91.3	5.0	49.2	54.2	8.7
12	(A)γ24a/(D)β27b/δα(C2)–Δd	6.4	43.2	92.7	5.1	39.6	52.0	7.3
13	(A)γ24a/(D)β29b/δα(C2)–Δd	4.1	51.2	96.8	5.0	53.7	61.2	3.2
14	(A)γ24a/(D)β5c/δα(C2)–Δd	3.5	58.2	95.8	5.0	58.6	68.6	4.2
15	(A)γ24a/(D)β8c/δα(C2)_Δd	4.2	55.2	88.4	5.1	56.0	62.8	11.6
16	(A)γ8b/(D)β11c/δα(C2)–Δd	4.3	49.8	94.2	5.0	51.3	41.9	5.8
17	(A)γ8b/(D)β5c/δα(C2)–Δd	4.0	60.7	95.0	5.1	64.3	53.9	5.0
18	(A)γ8b/(D)β8c/δα(C2)–Δd	3.6	57.3	95.9	5.0	57.6	49.5	4.1
19	(A)δα(C2)_Δd10a/(D)β11c/γ	3.6	56.5	97.2	5.0	55.2	65.0	2.8
20	(A)δα(C2)_Δd10a/(D)β14c/γ	3.8	48.4	90.6	5.0	47.5	54.5	9.4
21	(A)δα(C2)_Δd10a/(D)β27b/γ	6.3	44.4	96.1	5.0	43.7	51.5	3.9
22	(A)δα(C2)_Δd10a/(D)β29b/γ	3.4	51.4	100.0	5.1	56.2		
23	(A)δα(C2)_Δd10a/(D)β5c/γ	3.5	57.2	90.8	5.0	54.2	66.5	9.2
24	(A)δα(C2)_Δd10a/(D)β8c/γ	4.7	55.4	89.6	5.1	52.8	63.9	10.4
25	(A)δα(C2)_Δd10a/(D)γ7b/β	4.1	52.9	100.0	5.0	57.0		
26	(A)δα(C2)_Δd10a/(D)γ8b/β	4.6	46.5	97.6	5.0	44.2	53.3	2.4
27	(D)β27b/(A)δα(C2)–Δd30c/γ	2.9	62.6	96.0	5.0	62.9	54.4	4.0
28	(D)β29b/(A)δα(C2)-Δd30c/γ	4.0	66.2	88.1	5.0	57.4	59.1	11.9
29	(D)γ29a/(A)δα(C2)–Δd30c/β	3.9	64	96.5	5.1	69.9	55.5	3.5
30	(D)γ7b/(A)δα(C2)–Δd30c/β	4.2	62.4	84.0	5.0	65.0	54.3	16.0
31	(D)γ8b/(A)δα(C2)–Δd30c/β	2.5	58.8	93.0	5.1	61.0	53.5	7.0
32	(A)δα(C2)_Δd7a/(D)β11c/γ	4.4	53	96.1	5.0	52.1	62.5	3.9
33	(A)δα(C2)_Δd7a/(D)β14c/γ	5.3	41.7	90.5	5.1	41.3	47.8	9.5
34	(A)δα(C2)_Δd7a/(D)β27b/γ	4.5	44.8	91.9	5.0	45.7	51.8	8.1
35	(A)δα(C2)_Δd7a/(D)β29b/γ	3.3	54.9	95.3	5.0	56.1	49.1	4.7
36	(A)δα(C2)_Δd7a/(D)β5c/γ	3.3	52	94.0	5.0	52.8	60.0	6.0
37	(A)δα(C2)_Δd7a/(D)β8c/γ	3.2	54.3	97.9	5.0	54.7	62.7	2.1
38	(A)δα(C2)_Δd7a/(D)γ7b/β	3.3	51.1	100.0	5.0	52.4		
39	(A)δα(C2)_Δd7a/(D)γ8b/β	3.7	43.3	86.7	5.0	43.3	50.0	13.3

Table 20 Values for J(abc(C5)).

#	DA-pair	σ_{app}, %	$\langle R_{DA}\rangle_{E1}$, Å	A_1, %	ΔR_{DA}, %	R_{model}, Å	$\langle R_{DA}\rangle_{E2}$, Å	A_2, %
1	**(D)γ29a/(A)β14c/δα(C5)–Δd**	4.6	54.2	91.7	5.2	59.3	66.1	8.3
2	**(D)γ7b/(A)β14c/δα(C5)–Δd**	3.8	46.5	94.6	5.0	41.7	56.5	5.4
3	**(D)γ8b/(A)β14c/δα(C5)–Δd**	4.0	46.4	96.3	5.0	45.4	55.3	3.7
4	**(D)γ29a/(A)β27b/δα(C5)–Δd**	5.5	57.4	84.6	5.4	55.3	51.2	15.4
5	**(D)γ29a/(A)β33b/δα(C5)–Δd**	6.0	68.1	77.7	5.9	75.7	60.7	22.3
6	**(A)γ12b/(D)β11c/δα(C5)–Δd**	5.6	51.1	91.7	5.0	50.3	42.2	8.3
7	**(A)γ12b/(D)β14c/δα(C5)–Δd**	5.7	45.7	95.3	5.1	48.9	52.9	4.7
8	**(A)γ12b/(D)β5c/δα(C5)–Δd**	4.7	60.5	92.7	5.0	64.6	47.2	7.3
9	**(A)γ12b/(D)β8c/δα(C5)–Δd**	5.4	54.9	81.2	5.0	54.6	46.9	18.8
10	**(A)γ24a/(D)β11c/δα(C5)–Δd**	2.9	55.2	90.8	5.1	60.2	63.7	9.2
11	**(A)γ24a/(D)β14c/δα(C5)–Δd**	4.2	52.1	94.8	4.9	53.1	59.8	5.2
12	**(A)γ24a/(D)β27b/δα(C5)–Δd**	4.9	47.6	100.0	5.0	44.8		
13	**(A)γ24a/(D)β29b/δα(C5)–Δd**	3.5	53.1	94.6	5.0	55.2	46.7	5.4
14	**(A)γ24a/(D)β5c/δα(C5)–Δd**	3.5	63.6	97.4	5.0	65.5	54.6	2.6
15	**(A)γ24a/(D)β8c/δα(C5)_Δd**	4.1	63	99.6	5.1	60.9	55.8	0.4
16	**(A)γ8b/(D)β11c/δα(C5)–Δd**	5.0	48.3	97.2	5.2	49.6	56.7	2.8
17	**(A)γ8b/(D)β5c/δα(C5)–Δd**	3.1	60.7	93.9	4.9	65.5	68	6.1
18	**(A)γ8b/(D)β8c/δα(C5)–Δd**	3.5	54	95.3	5.0	58.3	60.4	4.7
19	**(A)δα(C5)_Δd10a/(D)β11c/γ**	4.1	55.8	88.1	4.8	57.5	63	11.9
20	**(A)δα(C5)_Δd10a/(D)β14c/γ**	5.1	52	89.8	5.1	51.1	58.2	10.2
21	**(A)δα(C5)_Δd10a/(D)β27b/γ**	5.7	46.7	100.0	5.0	45.4		
22	**(A)δα(C5)_Δd10a/(D)β29b/γ**	3.7	51.4	93.6	5.0	53.7	58.9	6.4
23	**(A)δα(C5)_Δd10a/(D)β5c/γ**	4.5	63	89.0	5.4	62.1	70.9	11.0
24	**(A)δα(C5)_Δd10a/(D)β8c/γ**	4.9	63	94.4	5.1	57.3	55.6	5.6
25	**(A)δα(C5)_Δd10a/(D)γ7b/β**	3.8	54.1	94.7	5.0	57.1	46.5	5.3
26	**(A)δα(C5)_Δd10a/(D)γ8b/β**	3.6	48.5	98.1	4.9	45.9	56.7	1.9
27	**(D)β27b/(A)δα(C5)–Δd33c/γ**	4.4	65.6	85.5	5.3	64.9	58.6	14.5
28	**(D)β29b/(A)δα(C5)-Δd33c/γ**	5.0	58.7	100.0	5.0	52.7		
29	**(D)γ29a/(A)δα(C5)–Δd33c/β**	6.0	62.5	100.0	5.1	69.5		
30	**(D)γ7b/(A)δα(C5)–Δd33c/β**	5.0	57.6	95.6	5.0	58.3	47.3	4.4
31	**(D)γ8b/(A)δα(C5)–Δd33c/β**	3.8	59.8	88.4	5.2	62.2	53.2	11.6
32	**(A)δα(C5)_Δd7a/(D)β11c/γ**	4.8	48.6	87.2	5.1	49.5	58.3	12.8
33	**(A)δα(C5)_Δd7a/(D)β14c/γ**	5.4	44	82.6	5.0	40.4	52	17.4
34	**(A)δα(C5)_Δd7a/(D)β27b/γ**	3.6	48.8	96.0	4.9	48.1	43.2	4.0
35	**(A)δα(C5)_Δd7a/(D)β29b/γ**	3.3	52.6	95.3	5.0	53.6	46.9	4.7
36	**(A)δα(C5)_Δd7a/(D)β5c/γ**	3.2	53.8	95.4	5.0	50.9	63.5	4.6
37	**(A)δα(C5)_Δd7a/(D)β8c/γ**	3.7	52.9	90.9	5.1	50.9	62.1	9.1
38	**(A)δα(C5)_Δd7a/(D)γ7b/β**	3.8	49.4	100.0	5.1	53.0		
39	**(A)δα(C5)_Δd7a/(D)γ8b/β**	3.4	44.6	91.3	4.9	46.1	50.7	8.7

5.5.5.3 Fit parameters for all PDA analysis

Table 21 Full PDA fit parameters for the J(abc) molecule measured at 20 mM $MgCl_2$. See section 7.5.1 for all data and fit plots.

DA-pair	σ_{app},%	$\langle R_{DA} \rangle_{E1}$, Å	A_1, %	$\langle R_{DA} \rangle_{E2}$, Å	A_2, %	impurities R_{DA}, Å	impurities A, %	A_{Donly}, %	χ^2_r
(D)γ29a/(A)β14c/δα–Δd	5.0	57.3	68.6	48.9	13.2	77.2	1.8	16.4	1.11
(D)γ7b/(A)β14c/δα–Δd	5.2	47.6	32.1	38	3.1	75.5	14.5	50	1.14
(D)γ8b/(A)β14c/δα–Δd	6.0	43	50.4	34.2	5.9	85.6	1.2	42.6	1.62
(D)γ29a/(A)β27b/δα–Δd	4.6	52.3	60.3	60	4.7	80	5.4	29.5	1.49
(D)γ29a/(A)β33b/δα–Δd	5.1	63.5	73.7	53.2	2.8	78.2	2.7	20.7	0.58
(A)γ12b/(D)β11c/δα–Δd	5.3	46.1	66.8	52.4	3.7	70.6	2.6	26.8	1
(A)γ12b/(D)β14c/δα–Δd	5.5	45.3	60.2	52	4.9	74.8	5.1	29.9	1.12
(A)γ12b/(D)β5c/δα–Δd	5.0	58.4	74.8			71.6	2.8	22.4	1.42
(A)γ12b/(D)β8c/δα–Δd	5.3	53	67.1	46.5	4.2	67	3.7	25	1.33
(A)γ24a/(D)β11c/δα–Δd	4.7	54.1	69.8					30	1.04
(A)γ24a/(D)β14c/δα–Δd	4.9	49	53.4			77.6	5.6	41	1.47
(A)γ24a/(D)β27b/δα–Δd	4.5	44.5	35.1	51.6	6.7	67	2.5	55.7	1.37
(A)γ24a/(D)β29b/δα–Δd	4.7	47.2	61.2	57.6	5.2	71.4	2.3	31.3	1
(A)γ24a/(D)β5c/δα–Δd	4.2	60.9	60.9	52.6	4.3			36.9	1.21
(A)γ24a/(D)β8c/δα_Δd	4.2	55.3	61.1					38.9	1.52
(A)γ8b/(D)β11c/δα–Δd	5.0	42.8	26.7	49	6.3			66.9	1.5
(A)γ8b/(D)β5c/δα–Δd	5.1	58.3	48	69	2.2			49.8	0.86
(A)γ8b/(D)β8c/δα–Δd	5.0	52.1	34.1	61.5	4	78.3	2.3	59.6	1.26
(A)δα_Δd10a/(D)β11c/γ	4.2	53.6	80.9	60.3	4.1			14.1	1.02
(A)δα_Δd10a/(D)β14c/γ	4.7	48.9	64.5	57.9	11.2	79.6	5.1	19.2	0.8
(A)δα_Δd10a/(D)β27b/γ	5.0	46.4	72.3					27.7	1.71
(A)δα_Δd10a/(D)β29b/γ	4.6	46.7	65.9	55.2	3.1	77.2		30.1	0.97
(A)δα_Δd10a/(D)β5c/γ	5.4	59.7	69.7	51.3	7.6	71.9	3.8	18.9	1.28
(A)δα_Δd10a/(D)β8c/γ	4.1	54.2	73	63	8.5			18.5	1.24
(A)δα_Δd10a/(D)γ7b/β	5.3	49.2	47.6	57.4	6.1	76.6	10.1	36.2	1.53
(A)δα_Δd10a/(D)γ8b/β	5.4	45.5	74.2	55.5	1.5	77.5	2.3	22.6	1.17

(D)β27b/(A)δα–Δd26c/γ	5.0	55.4	26.8	49	4.3	76.6	3.7	65.2	1.26
(D)β29b/(A)δα–Δd26c/γ	4.4	58.7	47.7			76.2	6.3	46.1	1.37
(D)γ29a/(A)δα–Δd26c/β	5.0	67.3	48.4	57.9	5.5			46	0.75
(D)γ7b/(A)δα–Δd26c/β	4.6	55.5	46.4	48.9	5.8	67.7	2.1	48.7	1.16
(D)γ8b/(A)δα–Δd26c/β	4.6	50.6	58.6	44.9	3.1	75.7	2.3	35.9	1.43
(D)β27b/(A)δα–Δd28c/γ	4.0	58.7	54.3	49.8	8.5			37.2	1.15
(D)β29b/(A)δα–Δd28c/γ	4.6	63.7	62.5	53.8	4.2			33.3	1.04
(D)γ29a/(A)δα–Δd28c/β	5.2	68.7	53.5					46.5	1.29
(D)γ7b/(A)δα–Δd28c/β	4.0	60.9	17.4	54.9	3.9	76	9.3	69.4	0.74
(D)γ8b/(A)δα–Δd28c/β	3.3	54.6	60.8	49.2	7.8			31.5	1.08
(A)δα_Δd7a/(D)β11c/γ	4.3	55.7	84.2			73.9	1.4	14.4	1.21
(A)δα_Δd7a/(D)β14c/γ	4.6	43.5	59.4	49.8	10.9	76.7	1.4	28.3	1.34
(A)δα_Δd7a/(D)β27b/γ	4.7	48.6	59.2	40.7	2.8	69.7	2.1	35.9	1.05
(A)δα_Δd7a/(D)β29b/γ	4.4	53.9	33.8			74.6	16.5	49.7	1.5
(A)δα_Δd7a/(D)β5c/γ	4.8	57.6	78.4			87.9	3.2	18.4	1.23
(A)δα_Δd7a/(D)β8c/γ	4.3	57.9	82.4			80.3	2.4	15.1	1.6
(A)δα_Δd7a/(D)γ7b/β	5.2	52.6	34.5	58.7	3.6	71.3	1.8	60.2	1.28
(A)δα_Δd7a/(D)γ8b/β	4.6	46.5	65.5	51.4	5.6	83	1.9	27	1.36

Table 22 Full PDA fit parameters for the J(abc(C2)) molecule measured at 20 mM $MgCl_2$. See section 7.5.2 for all data and fit plots.

DA-pair	σ_{app} ,%	$\langle R_{DA}\rangle_{E1}$, Å	A_1, %	$\langle R_{DA}\rangle_{E2}$, Å	A_2, %	impurities R_{DA}, Å	impurities A, %	A_{Donly}, %	χ^2_r
(D)γ29a/(A)β14c/δα(C2)–Δd	3.5	56.1	48.2	47.2	1.5			50.3	1.17
(D)γ7b/(A)β14c/δα(C2)–Δd	4.1	47.9	56.7	41.6	4	73.2	1.8	37.5	1.23
(D)γ8b/(A)β14c/δα(C2)–Δd	3.6	44	49.3	50.3	2.9	68	0.9	46.9	1.13
(D)γ29a/(A)β27b/δα(C2)–Δd	5.4	47.8	21.6	57.1	6.9			71.5	0.78
(D)γ29a/(A)β33b/δα(C2)–Δd	4.5	58	28.9	66.8	7.9			63.2	1.05
(A)γ12b/(D)β11c/δα(C2)–Δd	5.3	48.5	72.3	55.4	2.7	72.4	2.2	22.8	1.05
(A)γ12b/(D)β14c/δα(C2)–Δd	3.6	45.6	67.4	52.4	3.1	70.7	4.2	25.2	1.49
(A)γ12b/(D)β5c/δα(C2)–Δd	3.2	57.5	77.6			80.2	7.8	14.6	1.72
(A)γ12b/(D)β8c/δα(C2)–Δd	4.6	49.7	72	58.9	6.3			21.7	1.2
(A)γ24a/(D)β11c/δα(C2)–Δd	3.1	56.5	66.9	49.1	1.7	75.6	1.6	29.7	0.93
(A)γ24a/(D)β14c/δα(C2)–Δd	3.2	48.1	64	54.2	6.1	68.9	1	29	1.27
(A)γ24a/(D)β27b/δα(C2)–Δd	6.4	43.2	55.7	52	4.4	76.5	1.8	38.1	0.92
(A)γ24a/(D)β29b/δα(C2)–Δd	4.1	51.2	72.1	61.2	2.4	75.8	4.4	21	1.06
(A)γ24a/(D)β5c/δα(C2)–Δd	3.5	58.2	64	68.6	2.8			33.2	1.35
(A)γ24a/(D)β8c/δα(C2)_Δd	4.2	55.2	58.6	62.8	7.7			33.7	1.15
(A)γ8b/(D)β11c/δα(C2)–Δd	4.3	49.8	45.3	41.9	2.8	60.8	1	50.9	1.47
(A)γ8b/(D)β5c/δα(C2)–Δd	4.0	60.7	48	53.9	2.5			49.5	1.07
(A)γ8b/(D)β8c/δα(C2)–Δd	3.6	57.3	53.5	49.5	2.3			44.2	1.12
(A)δα(C2)_Δd10a/(D)β11c/γ	3.6	56.5	82.3	65	2.4			15.3	1.32
(A)δα(C2)_Δd10a/(D)β14c/γ	3.8	48.4	67.8	54.5	7	73	1	24.2	1.15
(A)δα(C2)_Δd10a/(D)β27b/γ	6.3	44.4	76.3	51.5	3.1	69.2	0.9	19.7	1.21
(A)δα(C2)_Δd10a/(D)β29b/γ	3.4	51.4	77.2			72.9	2.5	20.3	1.33
(A)δα(C2)_Δd10a/(D)β5c/γ	3.5	57.2	72.1	66.5	7.3			20.6	1.18
(A)δα(C2)_Δd10a/(D)β8c/γ	4.7	55.4	71.8	63.9	8.3			19.9	0.93
(A)δα(C2)_Δd10a/(D)γ7b/β	4.1	52.9	78.5			85.1	3.7	17.8	1.08
(A)δα(C2)_Δd10a/(D)γ8b/β	4.6	46.5	56.3	53.3	1.4	70.9	1.3	40.9	1.14
(D)β27b/(A)δα(C2)–Δd30c/γ	2.9	62.6	60.4	54.4	2.5			37.1	1.1
(D)β29b/(A)δα(C2)-Δd30c/γ	4.0	66.2	69	59.1	9.3	82.6	3.7	18	1.42
(D)γ29a/(A)δα(C2)–Δd30c/β	3.9	64	60.3	55.5	2.2			37.6	0.99
(D)γ7b/(A)δα(C2)–Δd30c/β	4.2	62.4	61.9	54.3	12			26.3	1.15
(D)γ8b/(A)δα(C2)–Δd30c/β	2.5	58.8	67.5	53.5	5.1			27.4	1.13
(A)δα(C2)_Δd7a/(D)β11c/γ	4.4	53	63.5	62.5	2.6	86.3	4.9	29	1.28
(A)δα(C2)_Δd7a/(D)β14c/γ	5.3	41.7	54.6	47.8	5.7	63.3	0.7	39	1.26
(A)δα(C2)_Δd7a/(D)β27b/γ	4.5	44.8	59	51.8	5.2	73.3	1.8	33.9	1.78
(A)δα(C2)_Δd7a/(D)β29b/γ	3.3	54.9	75.3	49.1	3.7	73.8	2.6	18.5	0.68
(A)δα(C2)_Δd7a/(D)β5c/γ	3.3	52	55.1	60	3.5			41.3	1.25
(A)δα(C2)_Δd7a/(D)β8c/γ	3.2	54.3	70.4	62.7	1.5	74.8	4.1	24	1
(A)δα(C2)_Δd7a/(D)γ7b/β	3.3	51.1	67.9			72.5	1.9	28.4	1.26
(A)δα(C2)_Δd7a/(D)γ8b/β	3.7	43.3	57.9	50	8.9	64.7	0.7	32.4	1.38

Table 23 Full PDA fit parameters for the J(abc(C5)) molecule measured at 20 mM $MgCl_2$. See section 7.5.3 for all data and fit plots.

DA-pair	σ_{app}, %	$\langle R_{DA}\rangle_{E_1}$, Å	A_1, %	$\langle R_{DA}\rangle_{E_2}$, Å	A_2, %	impurities R_{DA}, Å	impurities A, %	A_{Donly}, %	χ^2_r
(D)γ29a/(A)β14c/δα(C5)–Δd	4.6	54.2	52.7	66.1	4.8			42.5	1.19
(D)γ7b/(A)β14c/δα(C5)–Δd	3.8	46.5	49.4	56.5	2.8	72.7	3.4	44.4	1.1
(D)γ8b/(A)β14c/δα(C5)–Δd	4.0	46.4	28.6	55.3	1.1			70.4	1.54
(D)γ29a/(A)β27b/δα(C5)–Δd	5.5	57.4	21.4	51.2	3.9			72.9	1.89
(D)γ29a/(A)β33b/δα(C5)–Δd	6.0	68.1	26.5	60.7	7.6			66	1.41
(A)γ12b/(D)β11c/δα(C5)–Δd	5.6	51.1	65.3	42.2	5.9	69.3	2.5	26.3	1.09
(A)γ12b/(D)β14c/δα(C5)–Δd	5.7	45.7	55	52.9	2.7	69.7	1.8	40.6	1.51
(A)γ12b/(D)β5c/δα(C5)–Δd	4.7	60.5	70.1	47.2	5.5	85.9	8.3	16.1	0.84
(A)γ12b/(D)β8c/δα(C5)–Δd	5.4	54.9	60.8	46.9	14.1	72.7	2.4	22.7	1.24
(A)γ24a/(D)β11c/δα(C5)–Δd	2.9	55.2	65.4	63.7	6.6			28.1	1.04
(A)γ24a/(D)β14c/δα(C5)–Δd	4.2	52.1	58.8	59.8	3.2			38	1.12
(A)γ24a/(D)β27b/δα(C5)–Δd	4.9	47.6	75.6					24.4	1.47
(A)γ24a/(D)β29b/δα(C5)–Δd	3.5	53.1	69.6	46.7	4	68.9	1.1	25.3	1.14
(A)γ24a/(D)β5c/δα(C5)–Δd	3.5	63.6	59.5	54.6	1.6			38.8	1.32
(A)γ24a/(D)β8c/δα(C5)_Δd	4.1	63	628	55.8	2.4			34.9	0.95
(A)γ8b/(D)β11c/δα(C5)–Δd	5.0	48.3	45.3	56.7	1.3			53.4	1.05
(A)γ8b/(D)β5c/δα(C5)–Δd	3.1	60.7	58	68	3.8			38.2	1.6
(A)γ8b/(D)β8c/δα(C5)–Δd	3.5	54	53.1	60.4	2.6			44.3	1.66
(A)δα(C5)_Δd10a/(D)β11c/γ	4.1	55.8	64.7	63	8.7	80.3	3.7	22.9	1.22
(A)δα(C5)_Δd10a/(D)β14c/γ	5.1	52	67.1	58.2	7.6	77.7	1.7	23.7	1.41
(A)δα(C5)_Δd10a/(D)β27b/γ	5.7	46.7	70.8			79.3	4	25.3	1.23
(A)δα(C5)_Δd10a/(D)β29b/γ	3.7	51.4	61.7	58.9	4.2	74.3	7.7	26.4	1.11
(A)δα(C5)_Δd10a/(D)β5c/γ	4.5	63	68.3	70.9	8.4			23.2	1.19
(A)δα(C5)_Δd10a/(D)β8c/γ	4.9	63	72.8	55.6	4.3			22.9	1.08
(A)δα(C5)_Δd10a/(D)γ7b/β	3.8	54.1	72	46.5	4			24	1.07
(A)δα(C5)_Δd10a/(D)γ8b/β	3.6	48.5	67.4	56.7	1.3	73.9	0.6	30.6	1.39
(D)β27b/(A)δα(C5)–Δd33c/γ	4.4	65.6	44.9	58.6	7.6			47.5	1.07
(D)β29b/(A)δα(C5)-Δd33c/γ	5.0	58.7	63.8			76.2	4.5	31.7	1.1
(D)γ29a/(A)δα(C5)–Δd33c/β	6.0	62.5	34.7			73.9	2.6	62.7	1.1
(D)γ7b/(A)δα(C5)–Δd33c/β	5.0	57.6	67.8	47.3	3.1	75.2	4.4	24.6	1.08
(D)γ8b/(A)δα(C5)–Δd33c/β	3.8	59.8	50.2	53.2	6.6			43.2	1
(A)δα(C5)_Δd7a/(D)β11c/γ	4.8	48.6	63.5	58.3	9.3			27.2	1.56
(A)δα(C5)_Δd7a/(D)β14c/γ	5.4	44	55	52	11.6			33.4	1.39
(A)δα(C5)_Δd7a/(D)β27b/γ	3.6	48.8	65.6	43.2	2.7			31.7	1.72
(A)δα(C5)_Δd7a/(D)β29b/γ	3.3	52.6	72.7	46.9	3.6	69.8	3.8	19.9	1.46
(A)δα(C5)_Δd7a/(D)β5c/γ	3.2	53.8	74	63.5	3.6	74.5	2.9	19.5	0.92
(A)δα(C5)_Δd7a/(D)β8c/γ	3.7	52.9	71.5	62.1	7.2			21.3	1.72
(A)δα(C5)_Δd7a/(D)γ7b/β	3.8	49.4	78.1					21.9	1.41
(A)δα(C5)_Δd7a/(D)γ8b/β	3.4	44.6	60.6	50.7	5.8	72.4	1.5	32.1	1.24

5.5.5.4 Incomplete molecules

To figure out the origin of minor populations mentioned in section 5.5.5.1 incomplete (consisting of only two strands) (section 5.3.1) variants of three-way junction molecules were investigated. This should help to understand how much incomplete molecules, possibly appearing in the measurement solution, can influence the outcome of analysis for J(abc).

For each three-way junction molecule an incomplete variant was hybridized by leaving the unlabeled strand out resulting in a two strand molecule (see Figure 5-26). Hybridization of incomplete molecules is described in section 5.3.1. Investigations of incomplete variants were performed only for the J(abc) molecule.

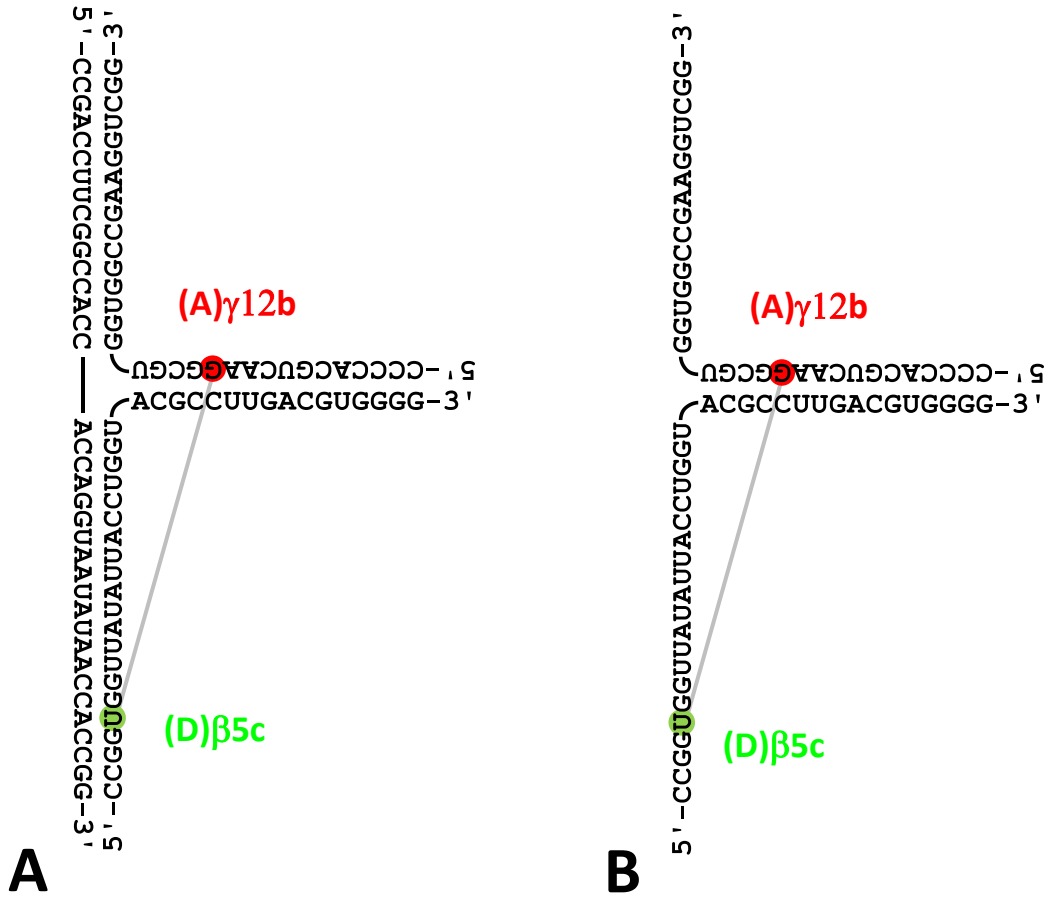


Figure 5-26 (A) Example of fully complementary three-way junction molecule ((A)γ12b/(D)β5c/δα–Δd) and (B) its correspoding incomplete molecule ((A)γ12b/(D)β5c).

Incomplete molecules were measured and further analyzed under the same conditions as normal three-way junctions which made the comparison of both molecules possible. As an

example the comparison of δα–Δd7a / (D)γ8b / β sample with its incomplete molecule is presented. Figure 5-27 clearly demonstrates that the minor peak on the upper plot (red line) is located at the same area on S_G/S_R axes as the major peak on the lower plot (blue line). This shows that the minor population in the complete molecule (upper plot) is not a state of J(abc) molecule but rather an incomplete molecule in the measurement solution.

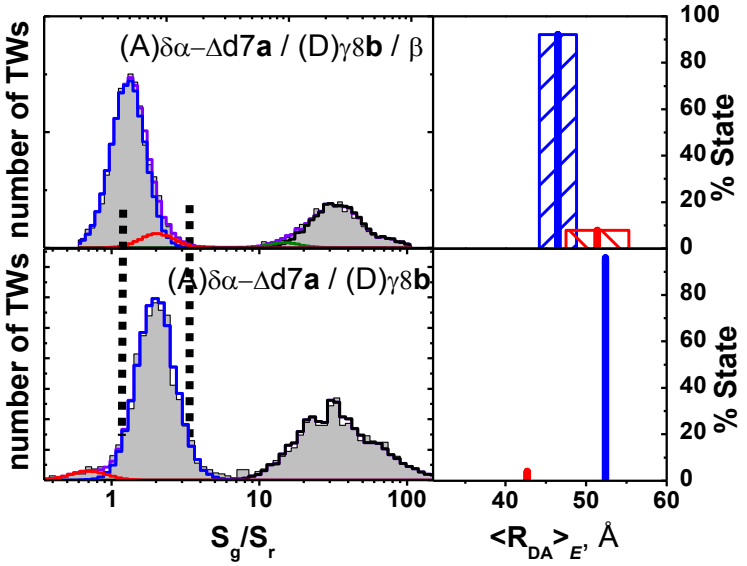


Figure 5-27 PDA analysis of the three-way junction δα–Δd7a / (D)γ8b / β sample (upper) with its corresponding incomplete δα–Δd7a / (D)γ8b (lower) sample. Dashed lines are used to emphasize the similar regions on the S_g/S_r axes of upper and lower plots.

Summarizing the results one can see that only 33 out of 44 three-way junction molecules had minor peak in PDA analysis. 16 out of 33 are likely due to incomplete molecules in the measurement solution.

In other cases incomplete molecules were either yielding the same DA distance as the corresponding fully complementary molecule, or they were not present in the measurement solution. In the first case the major peak of the incomplete molecule in the PDA analysis overlapped with the major peak of the corresponding peak for the J(abc). In the second case distances yielding from incomplete molecules were strongly shifted to the shorter or longer distance range and were not observed in three way junction measurements.

See section 5.5.5.2 for the obtained distances and section 7.5.4 in appendix for all data and fit plots.

5.5.5.5 Acceptor photophysics

To study effects of acceptor photophysics on my measurements I decided to exclude the bright bursts from the analysis and compare results of cut and full data sets (Figure 5-28 A,B). By examining samples which had only one conformation it was found that they all have τ_{red} (time obtained from the red decay tail fitting) typically more than 1.5ns and less than 3.5ns (Figure 5-28 C) which justifies the cutting criteria (blue dashed lines A and B).

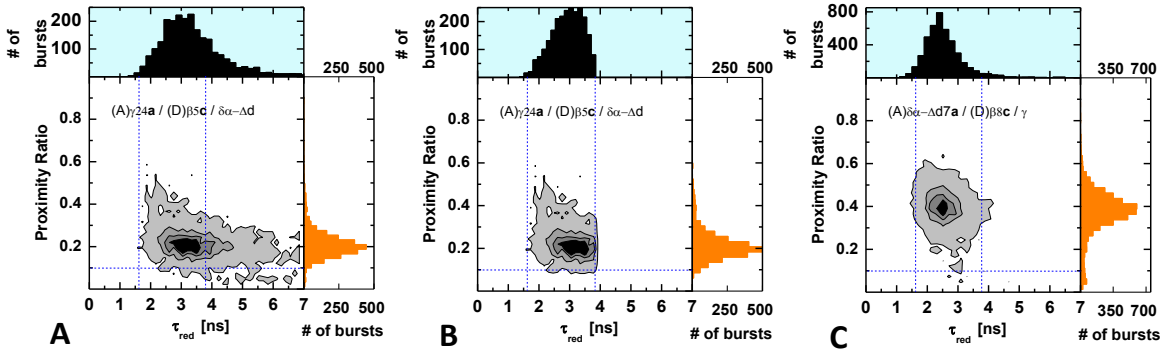


Figure 5-28 (A-C) The proximity ratio (ratio of the acceptor and the sum of donor and acceptor fluorescence ($S_A/(S_D + S_A)$)) versus acceptor lifetime τ_{red} is shown. Four dashed lines represent the cutting thresholds. The area outside the dashed box was excluded from further analysis. (A) Frequency histogram of the (A)δα–Δd7a / (D)β8c / γ sample before applying a "red cut". (B) Frequency histogram of the (A)δα–Δd7a / (D)β8c / γ sample after applying a "red cut". (C) An example of a sample with only one major population and with the acceptor lifetime in the range of 1.6-3.7 ns.

PDA analysis shows clearly that after performing a "red cut" (τ_{red} is used to set up a threshold to exclude bursts) the minor peak has disappeared. Additionally, value of χ_r^2 = 0.53 suggests that no minor peak is necessary for obtaining a good fit (Figure 5-29 A,B).

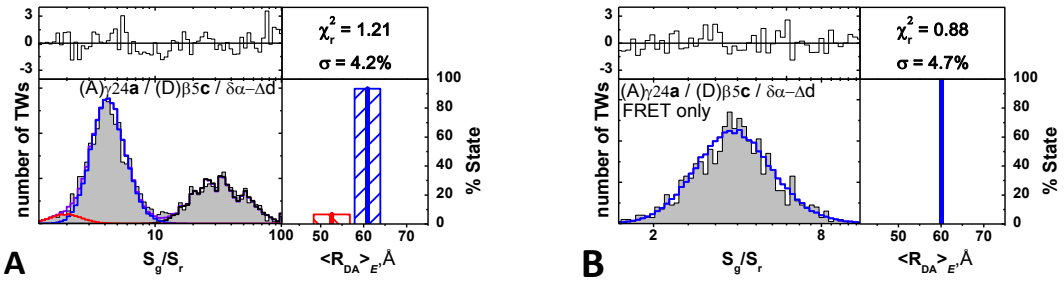


Figure 5-29 (A,B) PDA analysis for sample (A)δα–Δd7a / (D)β8c / γ. (A) One major and one minor FRET states as well as one D-only population was needed to fit the experimental data. (B) The same dataset after performing a "red cut". Minor as well as D-only populations disappear.

From sections 5.5.5.4 and 5.5.5.5 one can make a conclusion that the minor peak in the PDA analysis can be explained either with presence of the incomplete molecules in the measurement solutions or with the acceptor photophysics. This allows me to concentrate my structural studies solely on the major population of the studied three-way junctions.

5.5.5.6 **Modeling**

Rigid body docking (RBD) (section 5.4.7), which is a part of the FPS approach, was performed to generate and refine three-dimensional structural models. As in case of J(abcd), RNA helices were assumed to be rigid bodies with perfect A-RNA form (section 5.5.2). However, in the junction region some base pairs may be split due to steric tension similar to three-way DNA junction [110].

To encounter this effect in modeling, the two base pairs of each helix closest to the junction region were cut into four separate nucleotide fragments and were held together by elastic hydrogen and standard covalent bond restraints (Figure 5-30). For justification of this approach see section 5.5.5.6.1.3.

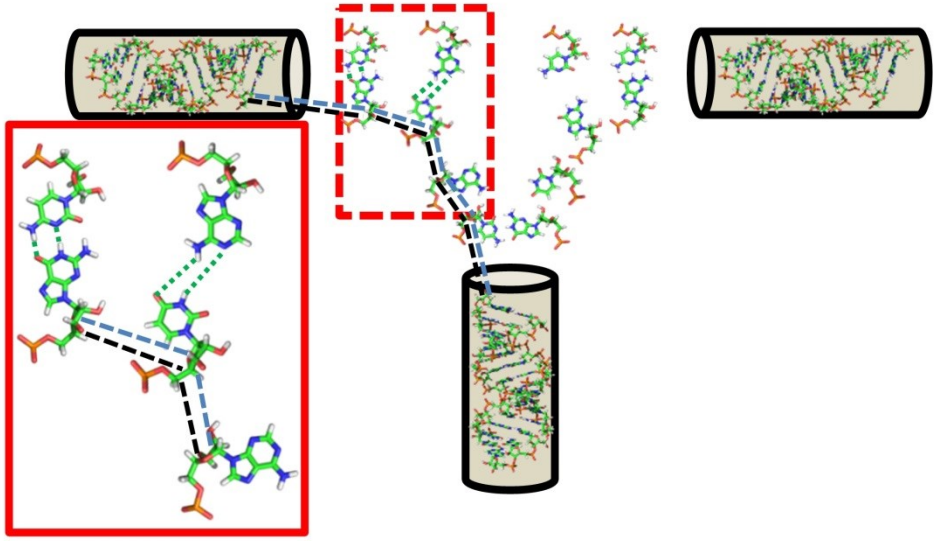


Figure 5-30 RNA fragments used in the modeling procedure. Rigid fully paired RNA helices in cylinders and single nucleotide fragments are shown. Connecting hydrogen, covalent and artificial bonds are shown as green, black and blue dashed lines, respectively (see section 5.4.7 for more information)

To ensure a proper stacking of the cut nucleotides belonging to the same strand, artificial bonds between two sugar oxygen atoms of neighboring nucleotides were introduced. Bulges were modeled by addition of single nucleotides as separate fragments connected with covalent and artificial bond restraints with each other (Figure 5-8). FRET restraints between fully paired parts of helices, and bonds between fragments were used to keep all parts together. The influence of the covalent, hydrogen and especially artificial "bonds", introduced for the simulation procedure, on the 3D structures of the obtained models will be discussed in the section 5.5.5.12.

Several flatness parameters were introduced to describe the planarity of the obtained structures which will be discussed in details in 5.5.5.8.1.

Two independent strategies described below were used to obtain 3D structures with the help of FPS.

5.5.5.6.1 First strategy – Rigid Body Docking

FRET and bond restraints were applied on the helices in the docking and refinement procedures (section 5.4.7). Obtained RBD structures with the best agreement with FRET data (the structure with the lowest χ^2 value) are the desirable solutions ("best" family of structures).

"Best" structure for J(abc) is flat (small deviation from complete planarity, Table 28) and has a Y-shape (Figure 5-31A). Cluster analysis was performed to estimate the uniqueness of this structure. It showed that the "best" structure is unique with > 99.9 % of confidence (see section 5.5.5.6.1.2 for cluster analysis results). The precisions of the rigid body models were estimated via bootstrapping procedure as described in [111] and in section 5.4.11. The average precision of the overall helix geometry of the fully paired RNA molecule is 2.1 Å defined by RMSD for all phosphorus atoms of the helices (two nucleotides of each helix closest to the junction were not taken into account) (Figure 5-31E).

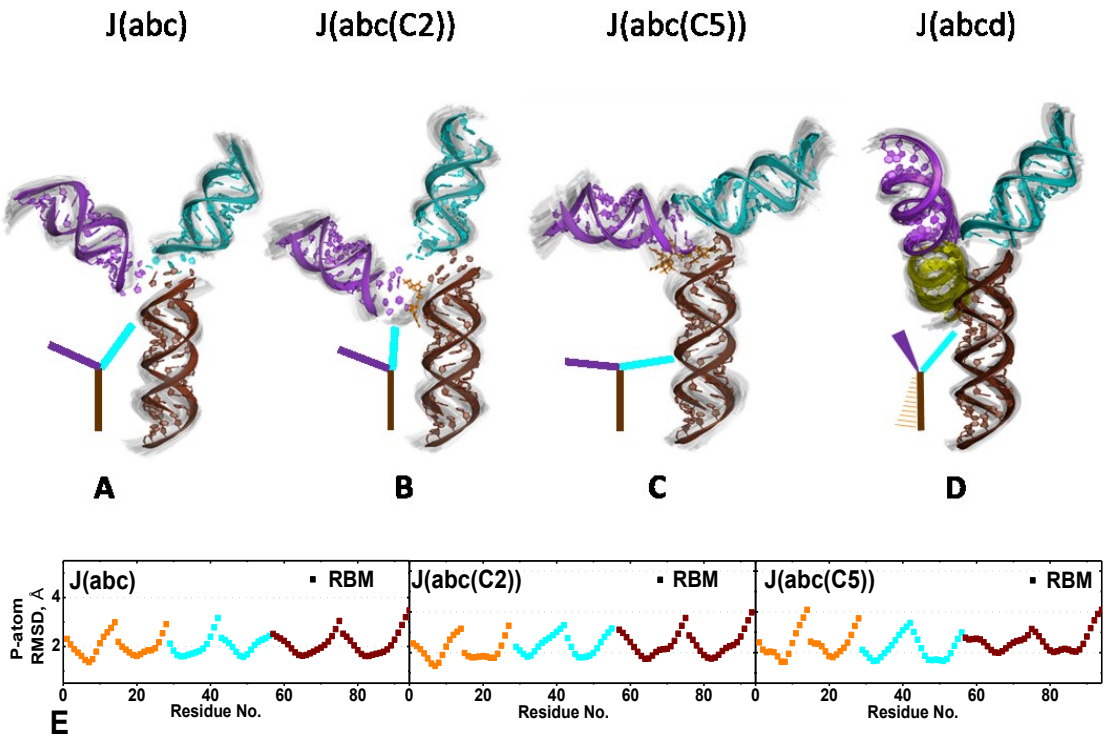


Figure 5-31 3D RBD "best" models (cartoon representation) with their schematic representations of J(abc) (A), J(abc(C2)) (B), J(abc(C5)) (C) and J(abcd) (D) overlaid with 100 structures resulting from bootstrapping (gray transparent) that indicate the uncertainties of the helix positions and orientations (see section 5.4.11). For consistency with Figure 5-3, helices *a, b, c* and *d* are depicted in purple, cyan, brown and yellow, respectively. The strong structural impact of the bulge is observed. (E) Uncertainty of phosphorus atom positions of the helical region (no junction included) for RBD models.

To get the 3D structure of 2C bulged molecule J(abc(C2)), 39 FRET distances of major populations were used. The overall shape is flat (small deviation from complete planarity, Table 28), consists of two coaxially stacked helices (helix *c* and *b*) and the third one (helix *a*) with an acute angle to the helix *b* (Figure 5-31B). The "best" structure obtained for this conformer is considered unique with > 99.9 % confidence (see section 5.5.5.6.1.2 for cluster analysis results) and has a precision of 2.3 Å of the overall helix geometry for phosphorus atoms of the helical part (Figure 5-31E). However, the cluster of "best" solutions contains structures with different conformations in the junction region. The bulge appears on different sides of the junction, "left" and "right", looking from helix *b* towards helix *a* (Figure 5-32A). In Figure 5-32C,D two "competing" structures are shown where helices *a* and *c* are kept in the same orientation. Helix *b* and a few nucleotides around the junction were omitted for clarity. Additionally, 2C bulge (shown in cyan) and 8 nucleotides (shown in red, yellow, blue, purple) belonging to the same strand $\delta\alpha(\mathrm{C2})-\Delta\delta$, from helices *a* and *c,* are

shown as “sticks”. Comparing the junction regions of two structures, it is apparent that in the “right” case (Figure 5-32D) the bulged strand continues from helix *a* to helix *c* performing a natural right turn, whereas in the “left” case (Figure 5-32C) an additional unnatural turn occurs. If I had to disregard one of the two structures based on this, my favorite structure would be the right one (Figure 5-32D). Furthermore, MD simulation studies were performed in order to resolve the “bulge side” problem. Results will be presented in section 5.5.5.7.2.

In case of J(abc(C5)), as in the previous molecule, 39 distances of major populations were used in 3D structural determination. Here, again, the structure is flat (small deviation from complete planarity, Table 28) and two of helices are stacked coaxially, but this time helix *a* and *b* are stacked and helix *c* is perpendicular to the other two (Figure 5-31C). The uniqueness of the “best” structure, provided by FPS, is 84 % confident (see section 5.5.5.6.1.2 for cluster analysis results). Resulting precision of the helical part of the RBD model is 2.4 Å, obtained from the bootstrapping (Figure 5-31E). Due to the lack of structural information in the bulge region, the whole bulge family is shown in Figure 5-32B.

To compare the major conformation of J(abcd) (section 5.5.4) with the three-way junctions, in Figure 5-31 helix *c* was kept in the same position in all cases. Additionally helix *b* and helix *c* of each molecule were kept always in the same plane.

In the section 5.5.5.7 results of RBD will be presented in an alternative way, defining the conformational space occupied by the obtained models.

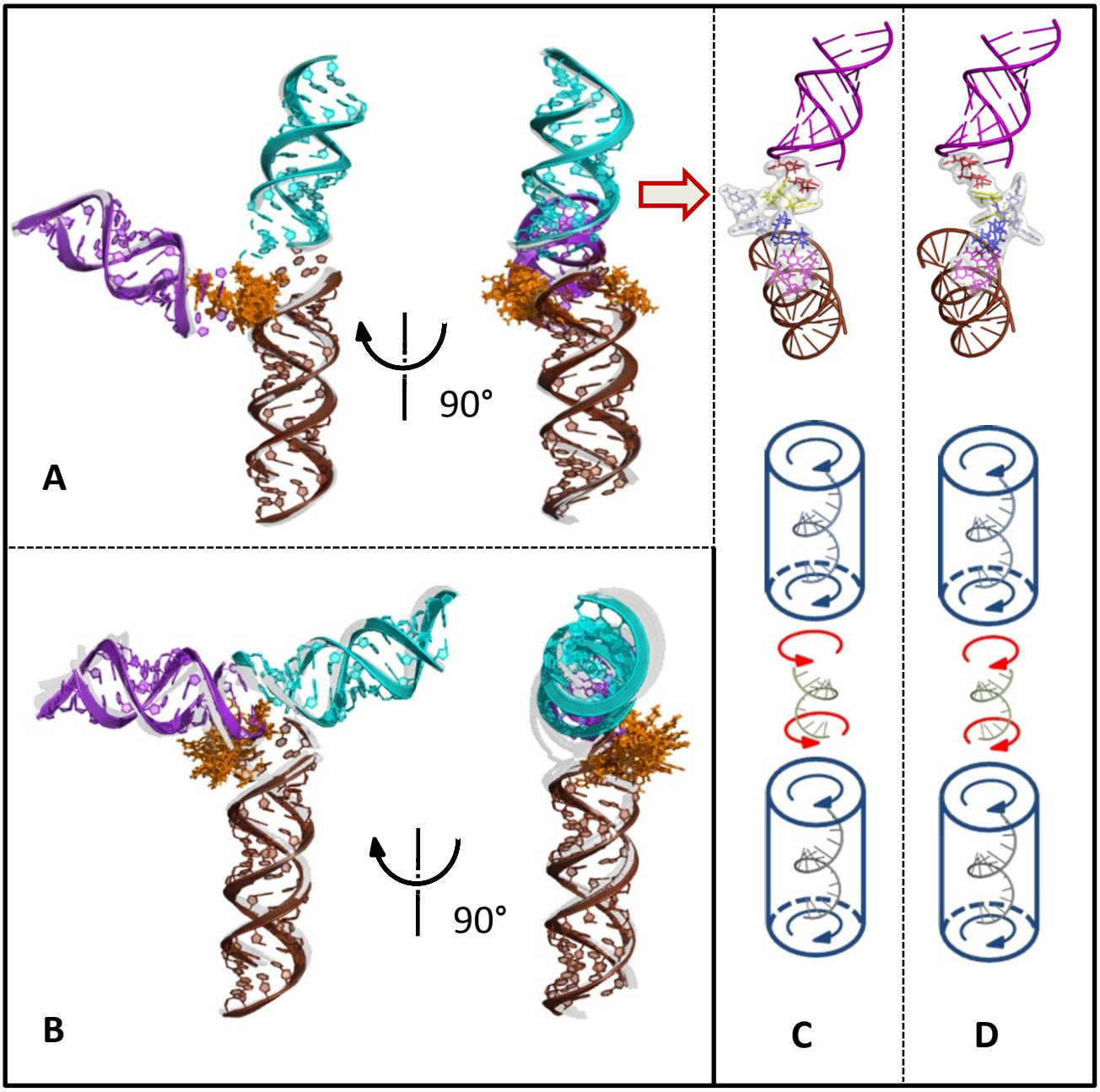


Figure 5-32 (A,B) Overlay of 20 structures of J(abc(C2)) (A) and J(abc(C5)) (B) compatible with experimental FRET data. Helices of "best" structures are colored, whereas other overlaid structures are shown in gray. Bulge nucleotides are shown in orange. (C,D) In the upper part, two possible turns for the strands containing bulge are shown for the J(abc(C2)). A schematic representation of the left turned and right turned helices is shown in the lower part. Comparing (C) and (D) versions one can notice an unnatural left turn of the strand containing the bulge in (C) and continuous natural right turn on (D).

5.5.5.6.1.1 Deviations of model distances from measured distances

To get an overview over the agreement of the rigid body models with experimental FRET data, deviations of model distances from the corresponding measured distances weighted by error from photon statistics, $\Delta R_{DA}(E)$ (Figure 5-33 A) and overall error (error from photon statistics + κ^2 error), ΔR_{DA} (Figure 5-33 B) (see section 5.4.6 for description of errors) for all studied three-way junctions together were calculated. In both cases it was possible to fit the data relatively fine with Gauss fit, which demonstrates that no explicit outliers were present.

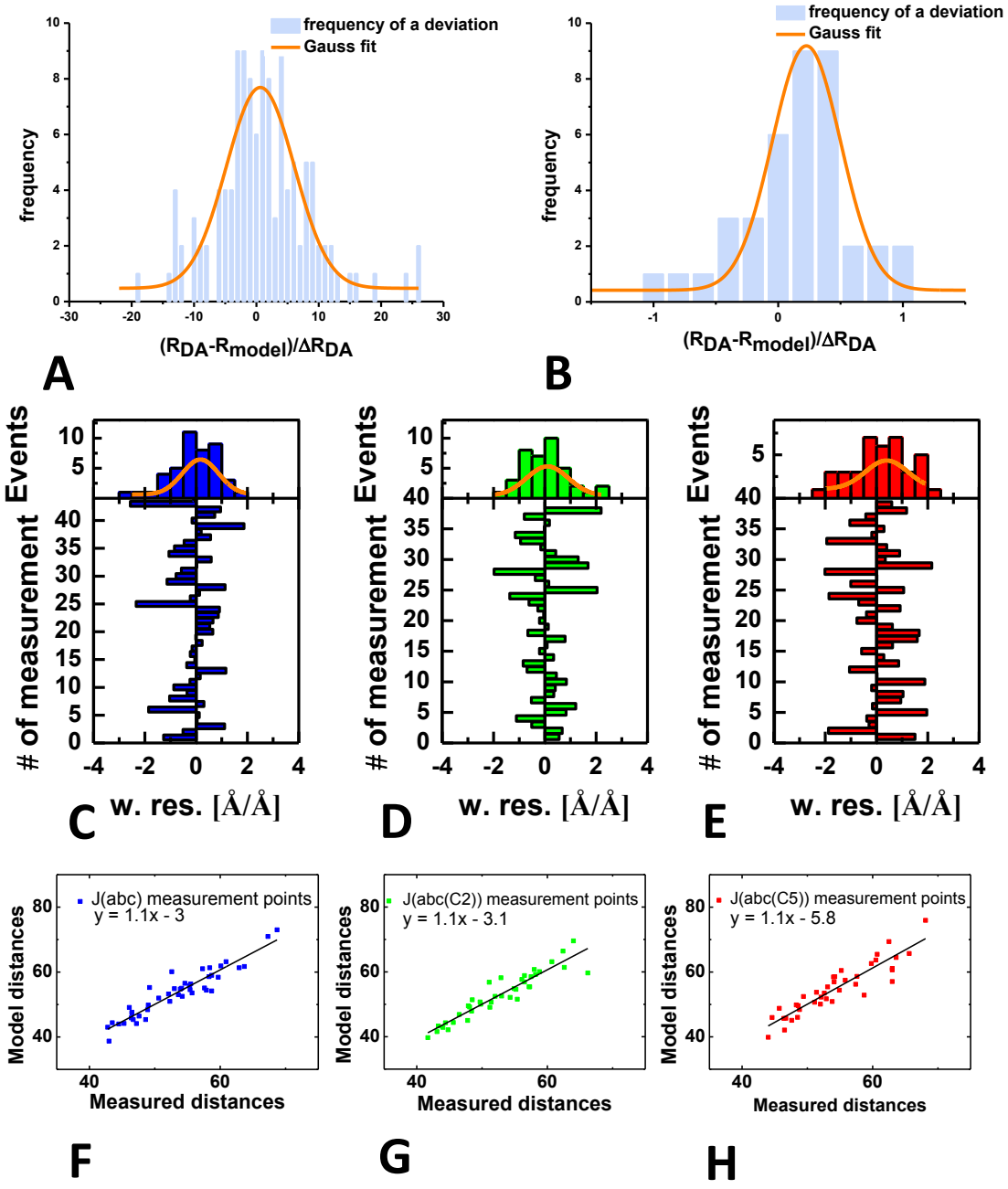


Figure 5-33 (A,B) Frequency histograms of summed up deviations between measured and model distances for all studied three-way junctions, weighted by the error (section 5.4.6) resulting from photon statistics ($\Delta R_{DA}(E)$) (A) and the overall error resulting from photon statistics and κ^2 errors (ΔR_{DA}) (B). Orange lines represent Gauss fits with x_{center} = 0.63, sigma = 5.45 for (A) and x_{center} = 0.3, sigma = 0.91 for (B). In (C-E) distance deviations are weighted by the overall error for J(abc)(C), J(abc(C2))(D) and J(abc(C5))(E). On the upper part of plots (C-E) resulting frequency histograms are shown, fitted with Gauss fits (x_{center} = 0.16, sigma = 0.75 for (C), x_{center} = 0.16, sigma = 0.85 for (D), x_{center} = 0.52, sigma = 1.27 for (E). (C-E) Measurement numbers correspond to the numbers mentioned in Table 18 for J(abc)(C), Table 19 for J(abc(C2))(D) and Table 20 for J(abc(C5))(E). Measured distances for J(abc), J(abc(C2)) and J(abc(C5)) are plotted against their corresponding model distances on (F), (G) and (H) respectively.

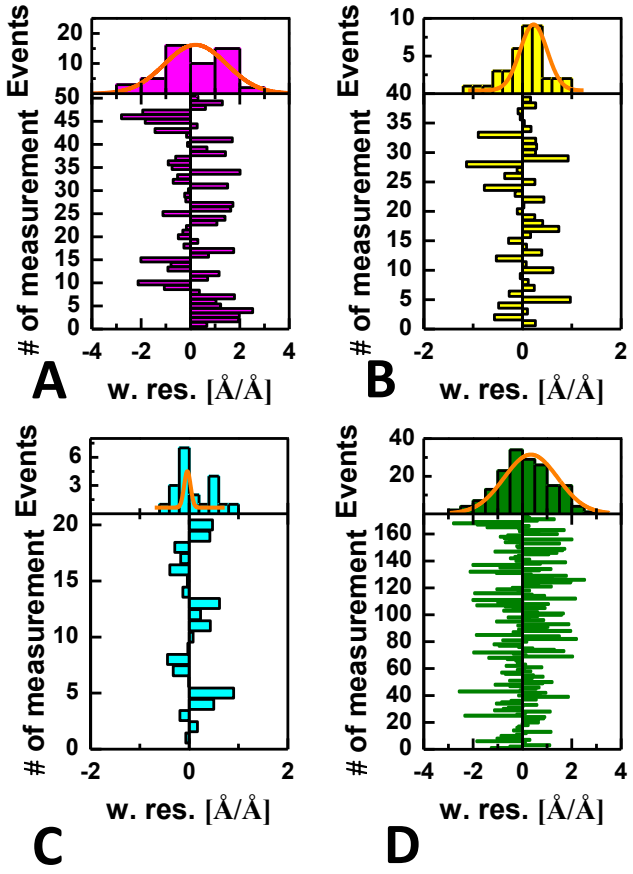


Figure 5-34 Frequency histograms of deviations between measured and model distances of J(abcd) major conformation (A), average deviations of common DA pairs for all studied three-way junctions only (B) and together with major conformation of four-way junction (C), weighted by the overall error (section 5.4.6) resulting from photon statistics and κ^2 errors (ΔR_{DA}). (D) Distance deviations of all studied molecules (major conformation for J(abcd)) piled up (173 measurements together), weighted by the overall error are presented. Orange lines represent Gauss fits with x_{center} = 0.53, sigma = 1.77 for (A), x_{center} = 0.22, sigma = 0.27 for (B), x_{center} = -0.03, sigma = 0.06 for (C) and x_{center} = 0.34, sigma = 1.08 for (D). Measurement numbers in (A-C) correspond to the numbers mentioned in Table 6, Table 19 and Table 41 in appendix, respectively. In (D) the measurement numbers from 1 – 44, 45 – 83, 84 – 121 and 122 – 173 represent the measurements for J(abc), J(abc(C2)), J(abc(C5)) and J(abcd) molecules,respectively.

Measured distance deviation from the model distances for each DA pair is shown separately for each of the three-way junction molecules (Figure 5-33C-E, lower parts). From the resulting histograms it is possible to see that the fluctuations of the distance deviations are around 0 (x_{center} is close to 0, Figure 5-33C-E, upper parts). Another representation of the distance deviations are shown on the Figure 5-33F-H, where measured distances are plotted against their corresponding model distances. In an ideal case one would expect the linear regression fit of the data points to go through 0. Fits for our data have only small deviations from the ideal case fit. Additionally, measured and model distance deviations for J(abcd) major conformation (Figure 5-34A), all studied three-way junctions (Figure 5-34B), all common (Figure 5-34C) and piled up (Figure 5-34D) DA distances for studied three- and four-way junctions (major conformation for J (abcd)) were calculated. Here, as in case of three-way junction molecules, fluctuations of distance deviations are around 0.

Summarzing the results, I conclude that in my studies were no systematic errors and the obtained models are correct.

5.5.5.6.1.2 Cluster analysis of the RBD results

With the help of χ^2_r (reduced energy of bonds, clashes and FRET restraints) and $\chi^2_{r\_FRET}$ plots (reduced energy of FRET restraints only) (Figure 5-35A,B, Figure 5-36A,B) one can clearly identify two clusters of solutions below and above the significance threshold (section 5.4.10) for J(abc) and J(abc(C2)) molecules, respectively. In case of J(abc(C5)) molecule, however, χ^2_r and $\chi^2_{r\_FRET}$ values increase gradually (Figure 5-37A,B) and in order to demonstrate the difference between the structures below and above the threshold, RMSD of helices versus the "best" structure should be additionally presented (Figure 5-37C). Only with the help of the latter one can identify the "best" cluster of solutions with structures from 1-375.

A big number of separate bodies in the junction region used for the docking procedure for all three molecules not only provides better flexibility in the junction but also introduces larger uncertainty in the overall structure. That is the reason why the structures belonging to the same cluster still have slightly different values of χ^2_r.

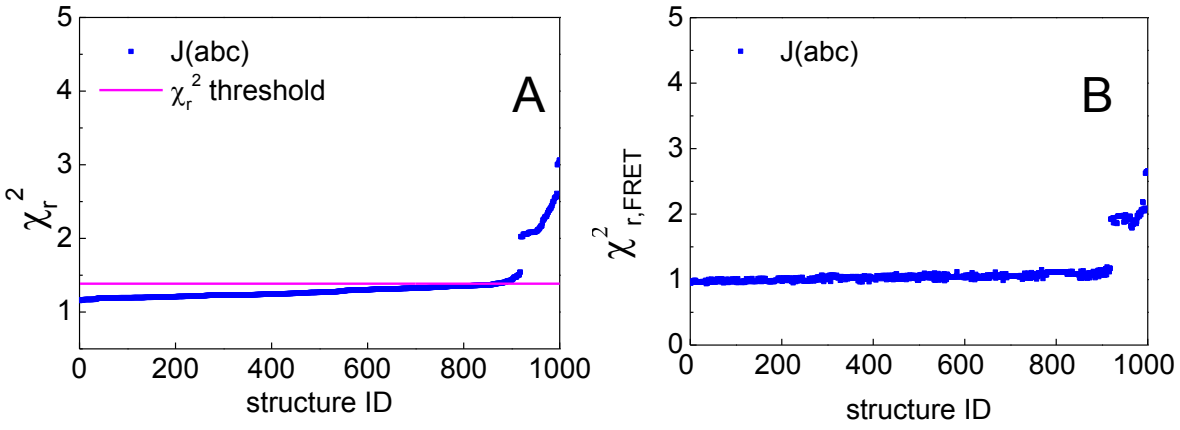


Figure 5-35 (A)χ^2_r and (B)$\chi^2_{r,FRET}$ are plotted against the structure ID after docking for J(abc) molecule. The magenta line represents 84 % confidence threshold (see section 5.4.10): $\chi^2_{r,max}$ = 1.38.

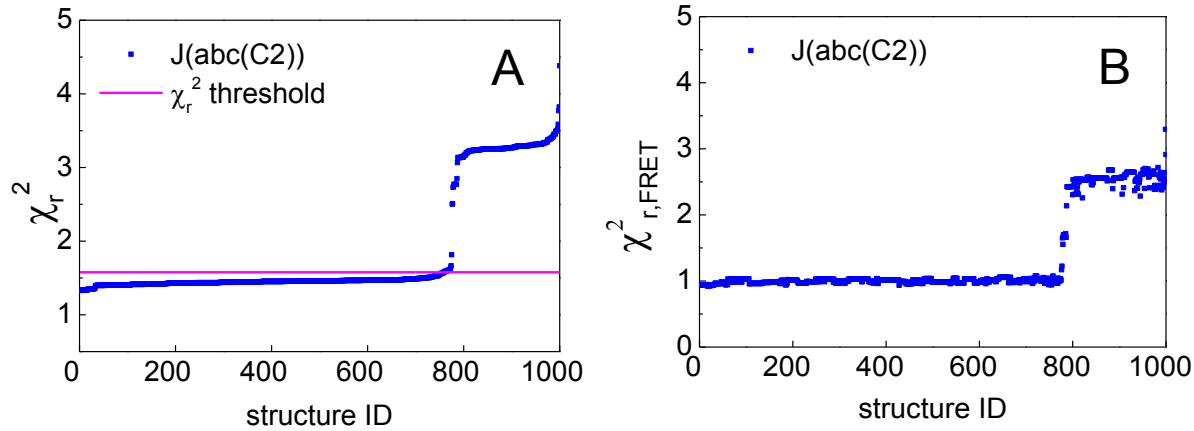
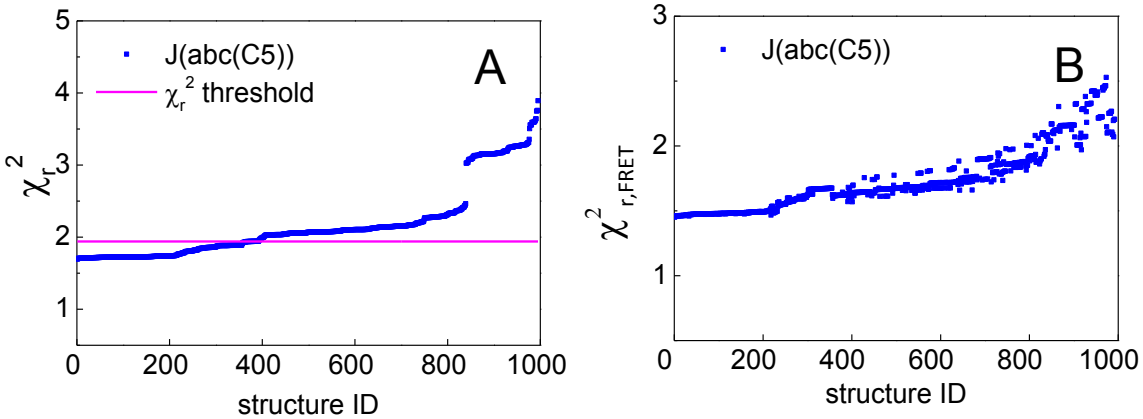


Figure 5-36 (A)χ^2_r and (B)$\chi^2_{r,FRET}$ are plotted against the structure ID after docking for J(acb(C2)) molecule: $\chi^2_{r,max}$ = 1.69.



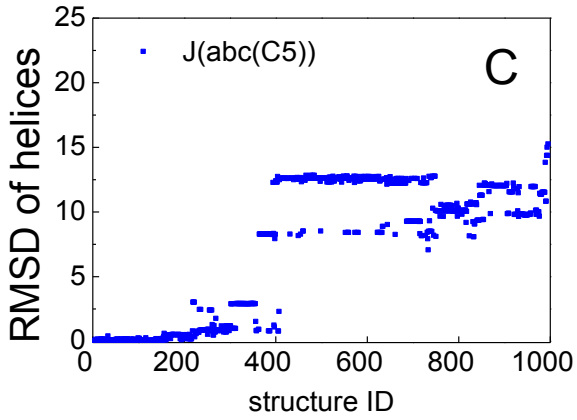


Figure 5-37 (A)χ^2_r, (B)$\chi^2_{r,\mathrm{FRET}}$ and (C) RMSD of helices versus the "best" structure (junction region is not involved) are plotted against the structure ID found after docking for J(abc(C5)) molecule: $\chi^2_{r,\max}$ = 1.95.

5.5.5.6.1.3 Stiff vs flexible helices

In order to justify the cutting of the helices into several fragments (section 5.5.5.6) the RBD procedure with the helices before (stiff helices) and after (flexible helices) the cut was performed. Cluster analysis of the docking results for stiff helices was done and compared with the results for flexible helices. From Figure 5-38 one can see that the $\chi^2_{r,\mathrm{FRET}}$ values for J(abc) and J(abc(C5)) are considerably higher in case of docking with stiff helices. Whereas for J(abc(C2)) molecule there is almost no difference. For better understanding of these results the "best" structures from both docking procedures were overlayed in Figure 5-39A-D. For the complete comparison of the structures $\chi^2_{r,\mathrm{FRET}}$ and RMSD values were calculated and presented in Table 24.

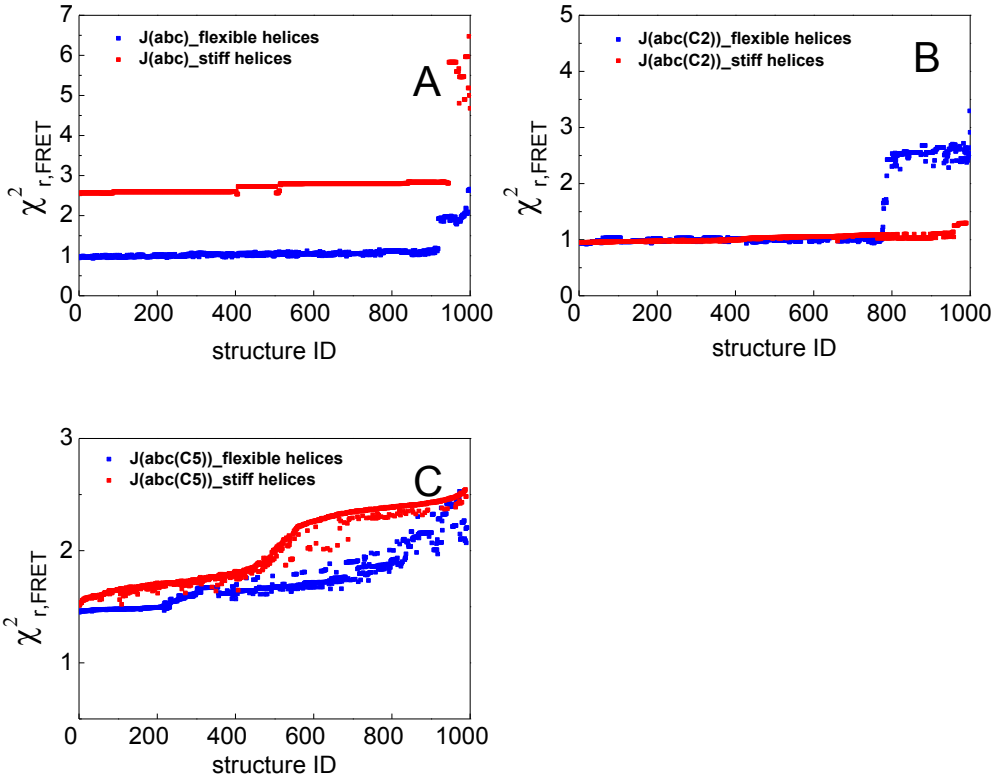


Figure 5-38 $\chi^2_{r,FRET}$ against the structure ID after docking with flexible (blue) and stiff (red) helices is plotted for J(abc) (A), J(abc(C2)) (B) and J(abc(C5)) (C) molecules.

From visual inspection of the generated structures one can see that docking with stiff helices result in perfect Y shape structure for the J(abc) molecule (Figure 5-39A) which, however, has much worse agreement with FRET then the structure generated with flexible helices (Table 24).

For molecules with bulges $\chi^2_{r,FRET}$ values of the structures with stiff helcies are low and comparable with flexible models' $\chi^2_{r,FRET}$ (Table 24), yet the junction regions of these models do not look natural and no proper stacking of the nucleotides is observed (Figure 5-39E-G). Moreover, the structure for J(abc(C5)) is not flat (Table 28) anymore (Figure 5-39D).

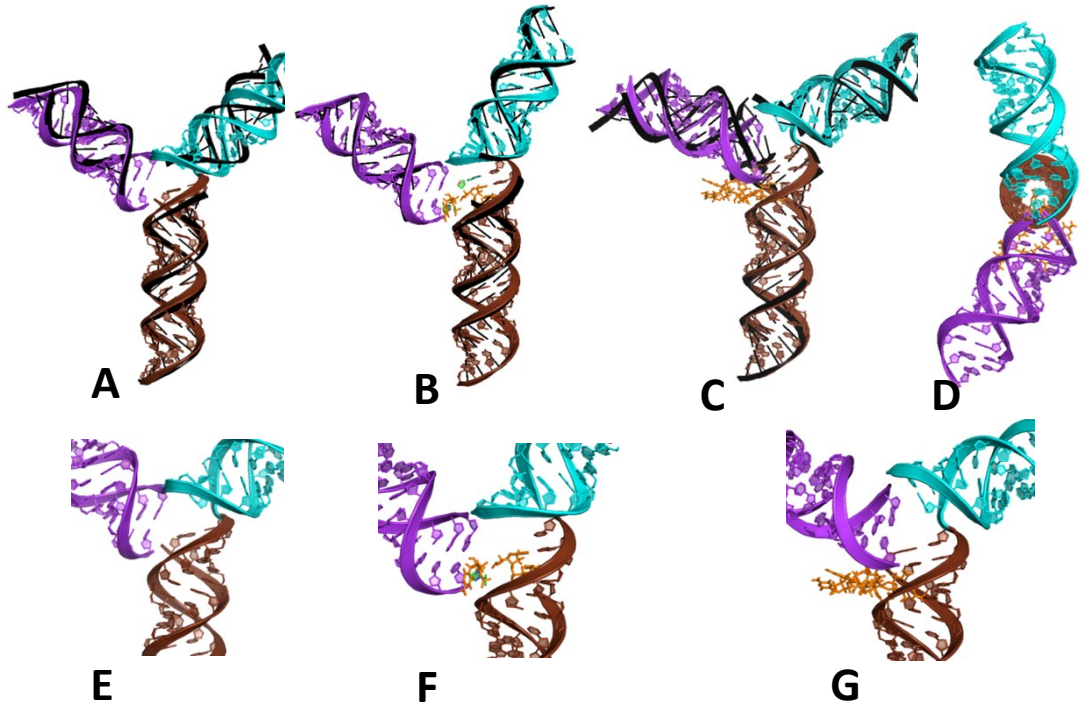


Figure 5-39 (A-D) Comparison of the "best" structures obtained by the RBD with flexible (black) and stiff (colored) helices for J(abc)(A), J(abc(C2))(B) and J(abc(C5))(C). Junction regions for the structures generated with RBD with stiff helices for respective molecules are shown on E-G.

Table 24 Comparison of the best rigid body models (RBM) generated with flexible and stiff helices for the three-way junction molecules. $\chi^2_{r,FRET}$ (no bond and clash energy is included) values of the rigid body models with best agreement with FRET data ($\chi^2_{r,min,FRET}$), as well as RMSDs over P atoms on the helical region between the respective rigid body models are calculated.

	J(abc)	J(abc(C2))	J(abc(C5))
$\chi^2_{r,min}$ RBD	0.93	0.94	1.45
$\chi^2_{r,min}$ RBD_stiff	2.56	0.95	1.55
RMSD of RBD model vs RBD_stiff over P atoms of the helical region	6.2 Å	1.7 Å	3.15 Å

Summarizing the results one can see that the usage of the flexible helices was particularly helpful in case of fully-paired three-way junction molecule. For J(abc(C2)) and J(abc(C5)) the biggest difference appears in the junction region and for the latter one the overall geometry

changes, which contradicts to the literature [5]. Hence, cutting the helices into fragments was a clever idea to provide structures with good agreement with FRET and possibly best geometry of the junction region.

5.5.5.6.2 Second strategy – Metropolis Monte Carlo Sampling

Rigid body docking was performed with bonds between helices and without FRET restraints (unrestrained docking). Afterwards, using the obtained structures as starting points, Metropolis Monte Carlo (MMC) (section 5.4.8) sampling was performed; exploring the allowed conformational space by generating possible meaningful structures of studied molecules (no clashes and overlays of helices with each other). Sampling procedure was followed by the screening of the structures according to FRET distances, herewith acquiring the desirable structures with $\chi^2_{r,\text{min\_FRET}}$ = 1.05, 1.84 and 1.41 for J(abc), J(abc(C2)), J(abc(C5)) molecules, respectively (Figure 5-40). Conformational space occupied by the obtained structures will be discussed in section 5.5.5.9.

5.5.5.6.3 Comparison of the "best" structures generated with two strategies described above

The comparison of the respective "best" structures (Figure 5-40) obtained by two strategies (sections 5.5.5.6.1 and 5.5.5.6.2) reveals RMSD 2.7 Å, 2.5 Å, 0.3 Å (Table 25) over phosphorus atoms on the helical region of J(abc), J(abc(C2)), J(abc(C5)) molecules, respectively, demonstrating the similarity of the obtained structures.

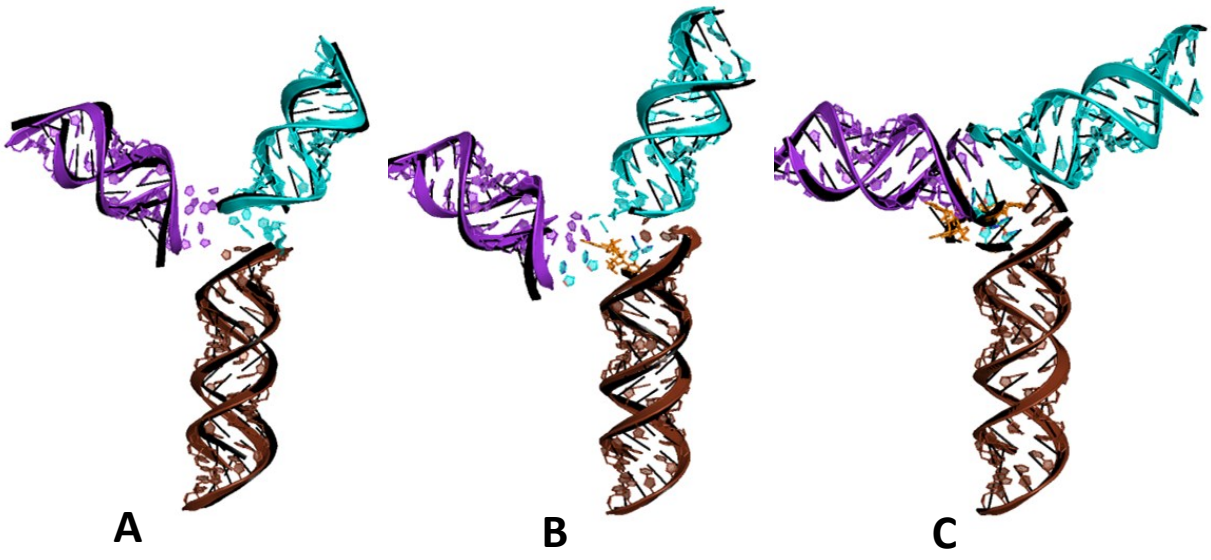


Figure 5-40 Comparison of the "best" structures obtained by the RBD (black) and MMC sampling (colored) for J(abc)(A), J(abc(C2))(B) and J(abc(C5))(C).

Table 25 Comparison of the RBM and MMC best structure for the three-way junction molecules. $\chi^2_{r,\mathrm{FRET}}$ (no bond and clash energy is included) values of the rigid body model and the MMC structure with best agreement with FRET data ($\chi^2_{r,\ \mathrm{min,FRET}}$), respectively. RMSDs over P atoms on the helical region between the rigid body models and the MMC structure with best agreement with FRET data, respectively.

	J(abc)	J(abc(C2))	J(abc(C5))
$\chi^2_{r,\mathrm{min\_FRET}}$ RBD	0.93	0.94	1.45
$\chi^2_{r,\mathrm{min\_FRET}}$ MMC	1.05	1.84	1.41
RMSD of MMC model vs rigid body model over P atoms of the helical region	2.71 Å	2.53 Å	0.29 Å

Additionally, the structures with χ^2_r below the significance threshold, $\chi^2_r < \chi^2_{r\ \mathrm{max}}$ ("best" family , section 5.4.10) from docking and sampling procedures, were compared with each other, demonstrating identical structural behavior in the potentially achievable conformational space of the models (see section 5.5.5.9.). Here it is important to mention that $\chi^2_{r\ \mathrm{max}}$ in both cases was calculated separately from respective $\chi^2_{r\ \mathrm{min}}$ values.

5.5.5.7 Molecular Dynamics Simulations

5.5.5.7.1 Three way junctions

All-atom molecular dynamics (MD) simulations were used to refine the RBD models. Simulations were performed in explicit solvent in order to obtain a high-quality structural refinement of the RBD starting structures [78] (section 5.4.12).

In contrary to MD simulations for J(abcd) (section 5.5.4.11.1), no positional restraints of RBD model were applied on the outer helical regions for more flexibility of the whole system (section 5.4.12). In total four, four and six trajectories were simulated in parallel for J(abc), J(abc(C2)) and J(abc(C5)) molecules, respectively. Simulation results for each molecule were added up, becoming 650 ns, 400 ns and 640 ns long free simulations for J(abc), J(abc(C2)) and J(abc(C5)), respectively. In Figure 5-41 cluster analysis of the MD data is shown where structures are sorted according to their χ_r^2 values. It is noteworthy that only few structures for each molecule were below the $\chi_{r,\max}^2$ threshold (section 5.4.10), calculated from the lowest χ_r^2 value for the MD structures and there were no structures below the threshold, calculated from the lowest χ_r^2 of the RBM (see 5.5.5.11 for more details).

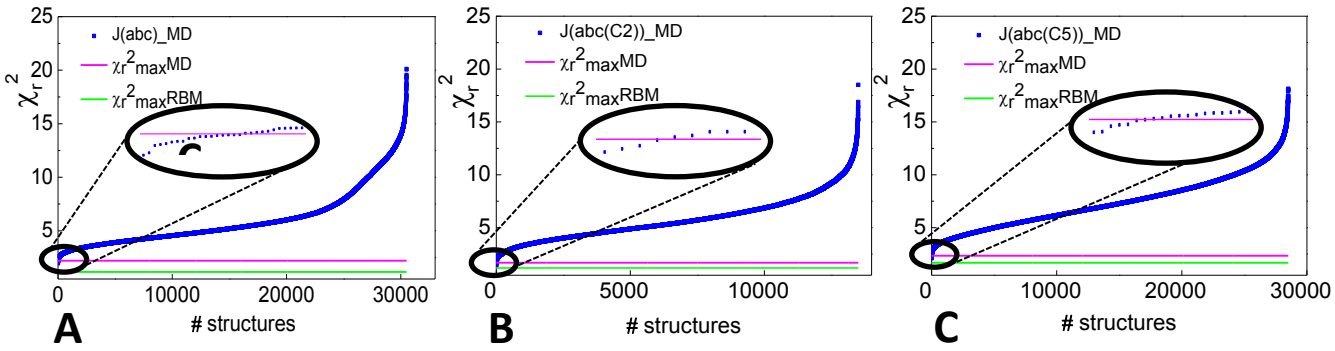


Figure 5-41 χ_r^2 is plotted against the structure ID after MD simulation for J(abc) (A) J(abc(C2)) (B) and J(abc(C5)) (C) molecules. The magenta line represents 84 % confidence threshold (see section 5.4.10): $\chi_{r,\max}^2$ MD = 2.2, 1.68, 2.34 for J(abc), J(abc(C2)) and J(abc(C5)) molecules, respectively.

The overlay of the respective "best" structures from MD and RBD is shown in Figure 5-42. Yielding RMSD as well as χ_r^2 values of the compared structures are presented in Table 26.

Figure 5-42 MD refined models with the best agreement with FRET (colored) overlayed with respective best rigid body models (black) for (A) J(abc), (B) Jabc(C2)) and (C) J(abc(C5)) molecules.

Table 26 Comparison of rigid body models (RBM) and MD best structure for the three-way junction molecules. χ^2_r values of the rigid body models and the MD structures with best agreement with FRET data (χ^2_{r,min_FRET}), respectively. RMSDs overall P atoms of the helical region between the rigid body models and the MD structures with best agreement with FRET data, respectively. Here, helical region is considered part of helices without first two and the last two nucleotides on each strand.

	3WJ	2C	5C
$\chi^2_{r,\ min\_FRET}$ RBD	0.93	0.94	1.45
$\chi^2_{r,\ min\_FRET}$ MD	2	1.45	2.12
RMSD of MD model vs rigid body model overall P atoms of the helical region	5.9Å	5.4Å	10.11Å

Representation of the MD simulation results in Euler angles will be presented in section 5.5.5.10.

5.5.5.7.2 Bulge side determination

As mentioned before, FRET provides no structural information in the junction region. Hence, the orientation of the bulge with respect to the helices in the rigid body models is arbitrary and is not supported by the experiment. In order to gain more information about the preferred orientation of the bulge it was decided to run several unrestrained simulations in parallel (section 5.4.12). First, for the calculation of the angle describing the bulge (orange) orientation shown in Figure 5-43, the center of mass of each of the three helices, helix *a* (red ribbon), helix *b* (green ribbon) and helix *c* (yellow ribbon) (COM_a, COM_b, COM_c, respectively) was calculated.

Then the center of helix B was chosen to be the origin. From this, the two vectors from COM_b to COM_a and from COM_b to COM_c (shown in green) were calculated. These two vectors span a plane (shown in blue). From this, the normal vector of the plane can be calculated (red line) using the cross product. The orientation of the normal vector with respect to the plane can be checked using the triple product.

Then, the vector from COM_b to the center of mass of the bulge (cyan line) was calculated. The angle can then be calculated between this vector and the normal of the plane. If it is 90°, then the center of mass of the bulge is in the plane of the three helices. If it is smaller than 90°, it is one the same side of the plane as the normal vector. If it is larger than 90°, it is on the other side of the plane.

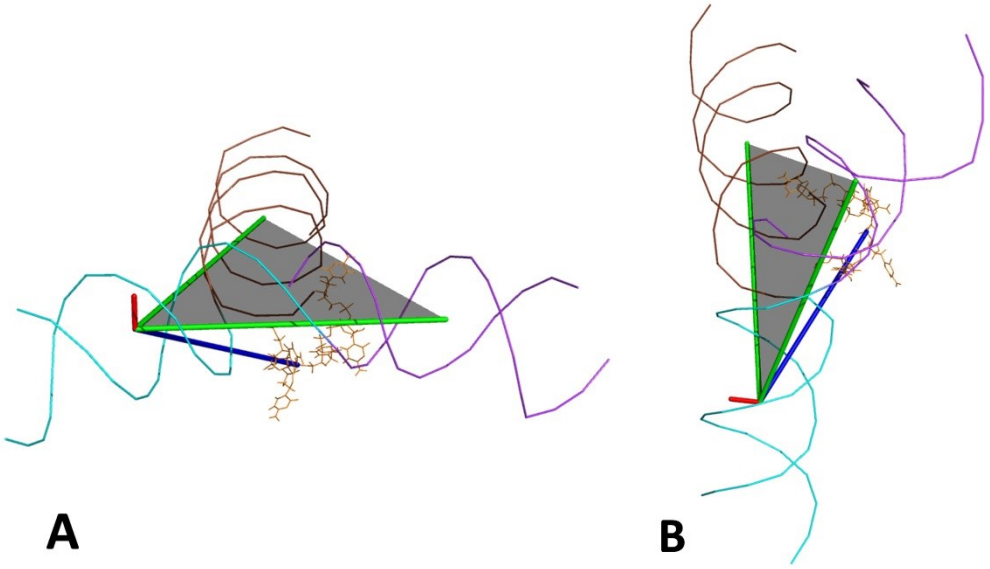


Figure 5-43 (A,B) Determination of the bulge orientation with respect to the helices. a,b,c helices are colored in bluepurple, cyan and chocolate, respectively. Red, green and blue vectors represent the normal to the plane, connection between COMs of the helices and the connection of COM_b to the COM of the bulge. Black transparent plane appears by connection between COMs of the helices.

For J(abc(C2)) two different initial structures (variant 1 and 2) with the bulge on different sides were used (Figure 5-32A). Simulations 40 ns long for each structure were performed. In case of variant 1 in 4 out of 5 simulations no bulge transition was observed. In the 5^{th} simulation the bulge performed a "round trip", moving from one side to the other and returning back. For the variant2, however, in 2 cases the bulge performed a full transition, in 2 cases there was no transition at all, and in the last simulation a "round trip" was observed (Table 27).

In RBD procedure for the J(abc(C5)) molecule all generated structures had their bulge on the same side (Figure 5-32B), hence, two initial structures with the same bulge side but with slighly different orientation were used to investigate bulge motion. 25 and 18 simulations with the length of 40 ns and 20 ns each were performed for variant1 and variant2, respectively (Table 27).

For variant1 15 out of 25 simulations showed no transition of the bulge to the other side, 3 out of 25 showed transition to the other side, and in 7 out of 25 simulations the bulge either stuck in the middle or went to the other side and got back. In case of variant2 in the majority of simulations a full transition was observed (11 out of 18). In 7 cases no transition was observed and only in one case the bulge performed a "round trip" (Table 27). Examples of the observed motions of the 5C bulge are demonstated in Figure 5-44.

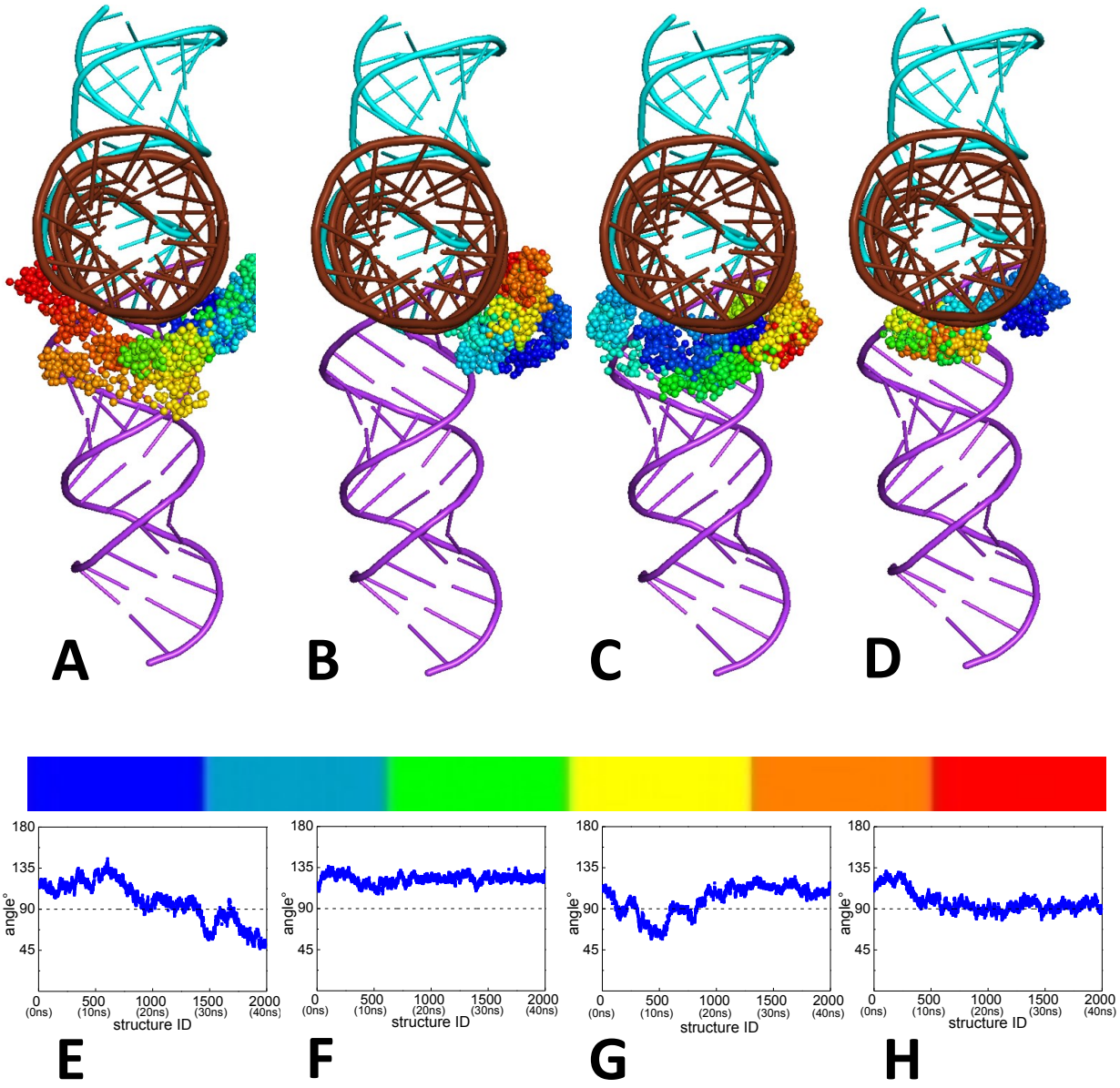


Figure 5-44 Motion of the 5C bulge within 40 ns simulation time for variant1. (A,B,C,D) Center of the mass of each bulge is shown as a colored sphere. From the beginning to the end of the simulations color of the spheres is changing from blue to red following the rainbow colors order. (E,F,G,H) Demonstrate the change in the angle between the red and blue vectors (see Figure 5-43), characterizing bulge side. Dashed line represents the conventional "border" between two sides of the molecule. (A,E) show full and (B,F) no transition of the bulge, on (C,F) "return trip" occurs. On (D,H) bulge moves to the middle point between two sides and sticks there.

Table 27 Summary of simultaneous simulations performed to study the orientation of the bulges. 10 (5 for variant1 and 5 for variant2) for J(abc(C2)) and 44 (25 for variant1 and 19 for variant2) for J(abc(C5)) simulations were performed. Transition of the bulge to the other side (Full transition), no transition (No transition), transition of the bulge to the other side and back ("return trip") and sticking of the bulges in the middle of both sides were observed.

		2C		
Name	**Simulation length**	**Full transition**	**No transition**	**"round trip"**
variant 1	**40ns**	**0**	**4**	**1**
variant 2	**40ns**	**2**	**2**	**1**
		5C		
		Full transition	**No transition**	**"round trip" or sticking in the middle**
variant 1	**40ns**	**3**	**15**	**7**
variant 2	**20ns**	**7**	**11**	**1**

Summary of the results show that in majority of the simulations for J(abc(C5)) molecule the 5C bulge did not change its side during the observation time. This could mean that the RBD model is correct in terms of bulge side. However, the side of 2C bulge is still unclear as both variants seem to be meaningful for MD. Noteworthy, the variant1 of J(abc(C2)), which seems to be preferred by MD, is actually my less favorable structure due to less natural left turn of the strand carrying the bulge (discussed in section 5.5.5.6.1).

5.5.5.8 Overview of the structure generating methods described above

In this work four strategies were used for FRET based structure generating. Each strategy has its advantages and disadvantages which are summarized below:

1. Rigid body docking with flexible helices (section 5.5.5.6.1)
   - Positive
     a. Generated structures are in a very good agreement with FRET data.
     b. Proper stacking of nucleotides in the junction region.
     c. Generated structures can be used as starting structures for MD.

   - Negative
     a. Generated structures are a simplified models and do not take into account the stereochemistry of the helices.

b. Large number of fragments, which makes the parameterization of the input files in FPS (section 5.4.7) for docking procedure time concuming.

2. Rigid body docking with stiff helices (section 5.5.5.6.1.3)
   - Positive
     a. Due to a small number of fragments, parameterization of the input files takes less time than for the case with flexible helices.
     b. "Best" structure generated for J(abc(C2)) is comparable with the "best" structure generated with flexible helices.

   - Negative
     a. Worse agreement with FRET than in case of flexible helices for J(abc) and J(abc(C5)).
     b. Absence of basepairing and stacking of the nucleotides in the junction region.
     c. Inapropriate as starting structures for MD.

3. Metropolis Monte Carlo sampling (MMC) (section 5.5.5.6.2)
   - Positive
     a. Obtained structures are in a good agreement with FRET, comparable with RBD.
     b. Possibility to explore a large variety of possible solutions (see conformational space studies in section 5.5.5.9).
     c. Proves that structures obtained with RBD are unique and there are no other possibilities to have different geometry of the helices with good agreement with FRET (see also section 5.5.5.9).
     d. Demonstrates the independence (at least not fully dependence) of the obtained structures from the bonds used in the junction region (see also section 5.5.5.9).

   - Negative
     a. Long sampling time

4. MD simulation (section 5.5.5.7)
   - Positive
     a. Provides accurate structures where all types of chemical interactions including long-range interactions, interections with ions in the solution and interaction with the solvent are taken into account.

   - Negative

a. Time consuming simulations.
b. It is hard to assess the completeness of sampling.

5.5.5.8.1 Flatness description of our molecules

For the 3D structure comparison of the "best" structures generated employing RBD with stiff and flexible helices, MMC and MD, flatness parameter η was calculated. From Table 28 one can see that all except MD models for J(abc) are practically perfectly flat. For the J(abc(C2)) all strategies prove that the molecule is less flat than J(abc), however the deviations from the $\eta = 360°$ are only 3° - 9°**.** RBD with flexible helices and MMC "best" models are nearly flat, whereas RBD with stiff helices and MD "best" structures are far less flat.

Taking into account that the three-way junctions with bulges in the junction region are always flat [5] and comparing the flatness of structures, one can conclude that RBD with flexible helices and MMC strategies suit the most to the structure determination studies.

Table 28 Comparison of the flatness parameter calculated for all "best" structures generated with the four strategies discussed in section 5.5.5.8.

	η			
	RBD_flexible helices	**RBD_stiff helices**	**MMC**	**MD**
J(abc)	**359.321°**	**359.999°**	**359.374°**	**348.269°**
J(abc(C2))	**350.747°**	**355.874°**	**356.763°**	**354.53°**
J(abc(C5))	**355.047°**	**339.75°**	**354.982°**	**344.764°**

Additionally, two alternative descriptions of the flatness were introduced: "helix to the plane" and "normalized volume", which were calculated for the models in the "best" family of solutions. "Helix to the plane" corresponds to the angle between the helix and the plane of other two helices. This angle was calculated for each of three helices. "Normalized volume" corresponds to the volume of parallelepiped spanned on the unit vectors along axes of helices of the structure. Zero value corresponds to complete flatness of the structure and 1 – when three helices are orthogonal to each other. Flatness parameters for the "best" models as well as their distributions within the "best" family are presented in (Table 29).

Table 29 Values of "helix to the plane" and "normalized volume" are presented for all three-way junction molecules.

Flatness measures						
	J(abc)		J(abc(2C))		J(abc(5C))	
Euler angles	For the best structure	Mean ± error	For the best structure	Mean ± error	For the best structure	Mean ± error
helix *a*_plane (*bc*)	9.9°	7° ± 15°	39.1°	28° ± 12°	19.4°	28° ± 13°
helix *b*_plane (*ac*)	7.8°	5° ± 11°	24.4°	22° ± 9°	17.6°	27° ± 13°
helix *c*_plane (*ab*)	5.8°	4° ± 10°	20.1°	17° ± 8°	30.2°	33.9° ± 10°
norm. volume	0.10	0.05° ± 0.2°	0.34	0.25° ± 0.2°	0.30	0.42° ± 0.2°

From calculated parameters of "normalized volume" one can see that "best" structure of J(abc) is almost completely flat, whereas "best" structures of bulged molecules have slight deviations from complete flatness.

5.5.5.9 Conformational Space-RBD

For complete description of the obtained 3D structures, mutual and Euler angles for helices were calculated. Mutual angles between helix pairs allow to make a comparison of the studied molecules according to their shape, whilst Euler angles give a comprehencive structural description and provide exact orientation of helices in a structure fixed coordinate system Figure 5-46A,B. The system of coordinates was chosen in such way that the rotation matrix of helix *a* is an identity matrix. In other words in case of Euler representation of rotation, helix *a* lies along z axis ($\varphi_a, \theta_a, \Psi_a = 0°$), and the rotation angle of the P atom of the first nucleotide equals Ψ_{p_atom} = -50.86° (Figure 5-45).

The angles for each structure in the "best" family of solutions generated with docking procedure were calculated. Calculated angles were used for demonstration of the conformational space, occupied by the structures within the "best" family, which in its turn

corresponds to the precision of the model (colored areas outlined with the black lines) (Figure 5-46C,D).

As mentioned in the section 5.5.5.6.2, describing the second strategy of generating the structures, unrestrained MMC simulations at different "temperatures" were performed. Logically, by increasing the "temperature", the ability of the system to jump out of the local energy minima increases, allowing the system to explore a larger conformational space.

For visualizing the results of unrestrained simulations the mutual and Euler angles for all structures generated at different "temperatures" were calculated and plotted together (Figure 5-46E,F,G) (see for Euler representation). The landscape of the enthalpy (bond + clash energy) changing from the light gray (high values of enthalpy) to black (low values of enthalpy) is demonstrating the confinement of the conformational space depending on the threshold applied to the enthalpy (see 5.5.5.12 for more details). The biggest values of the enthalpy obtained for each simulation (performed at different "temperatures") are used as thresholds in the Figure 5-46E,F,G.

Structures from the "best" families generated by RBD (shown in magenta) and MMC sampling (shown in green) were displayed in the same Figure 5-46E,F,G, making the comparison of generated data easier.

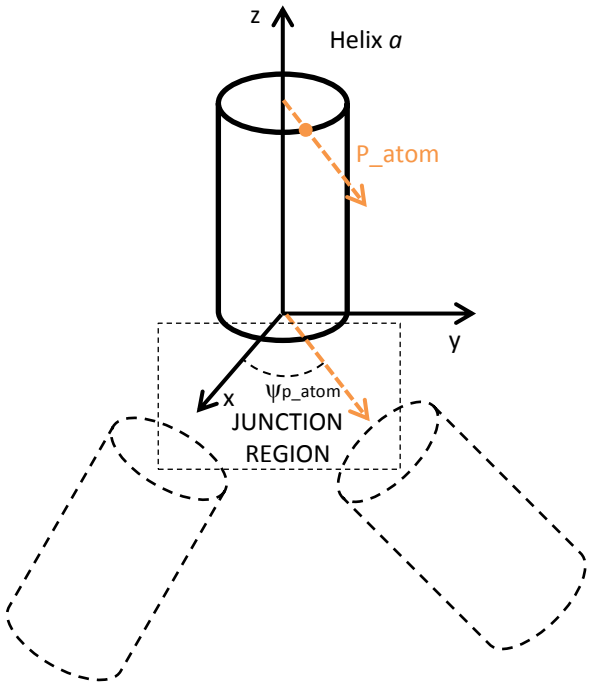
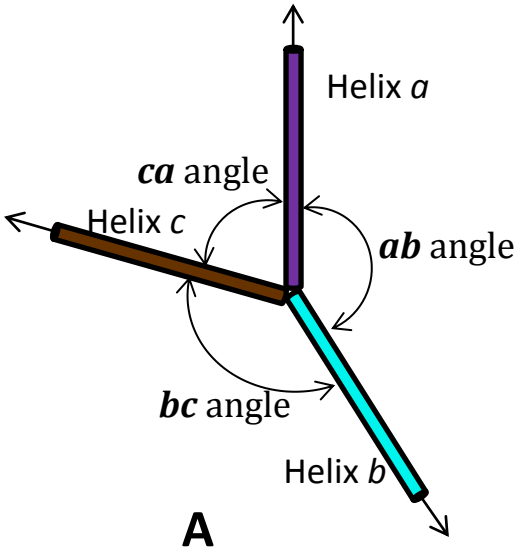
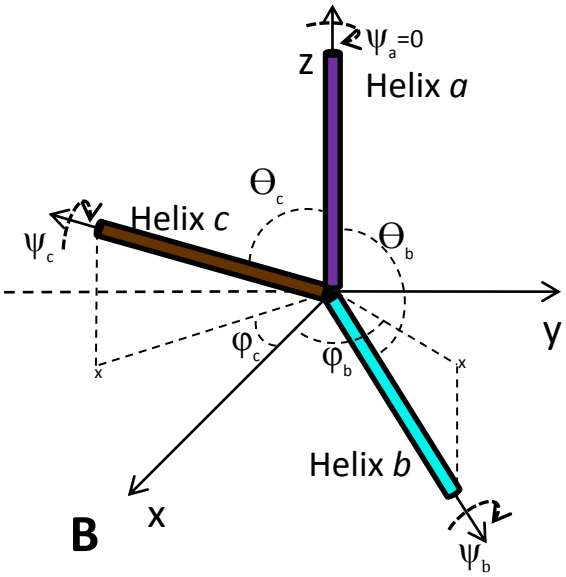


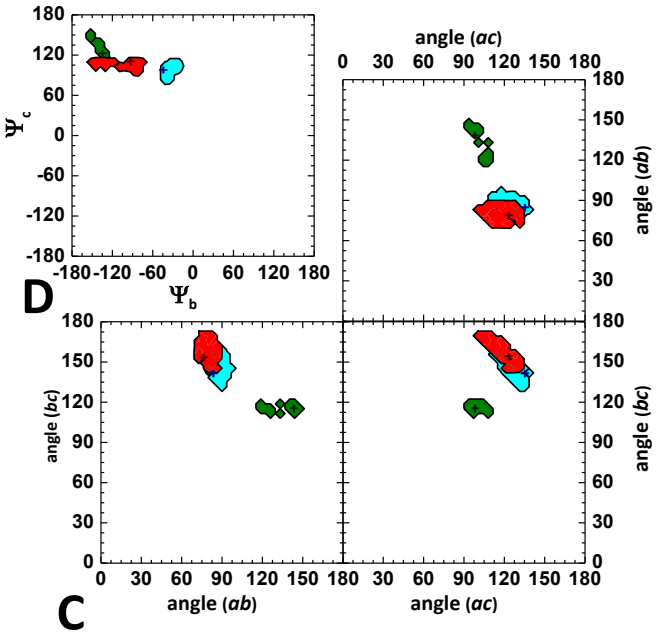
Figure 5-45 Exact position of helix *a* used in the coordinate system fixed to the structure. Position of P atom of first nucleotide from the free end of the helix *a*, as well as its rotation angle Ψ_{p_atom} = -50.86° are demonstrated. Positions of two other helices is determined with respect to the helix *a*.



A

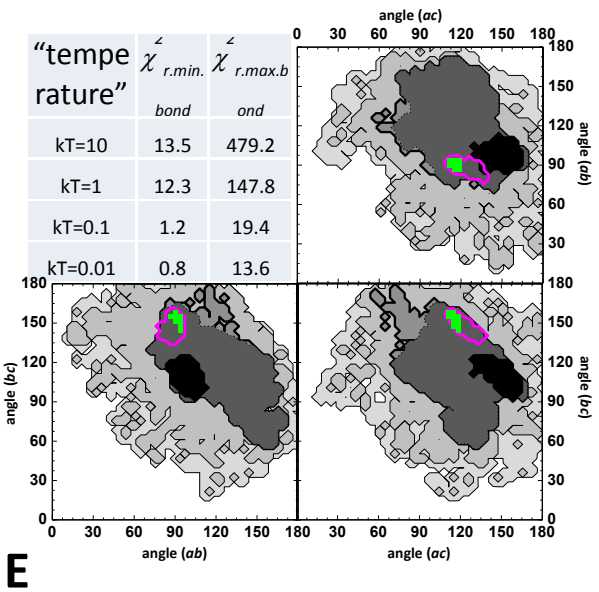


B



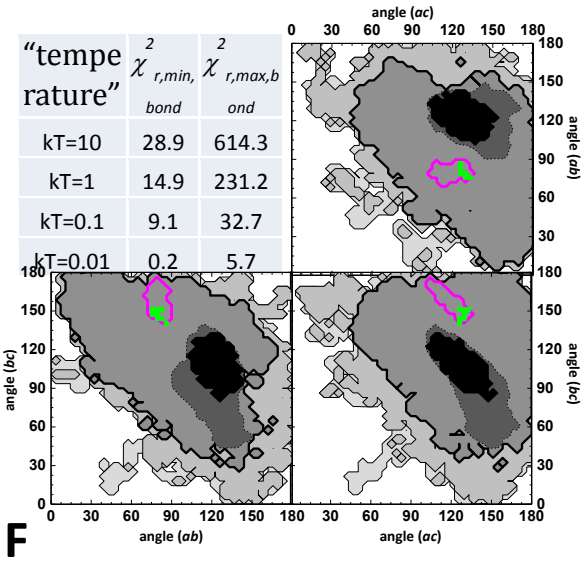
D

C



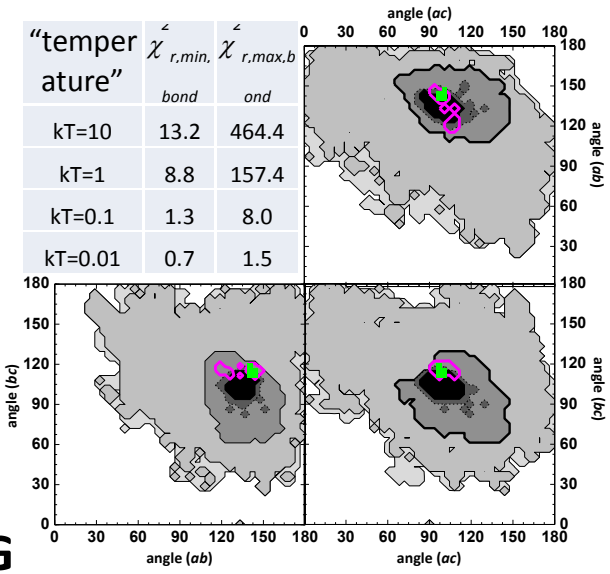
"temperature"	$\chi^2_{r,min,bond}$	$\chi^2_{r,max,bond}$
kT=10	13.5	479.2
kT=1	12.3	147.8
kT=0.1	1.2	19.4
kT=0.01	0.8	13.6

E



"temperature"	$\chi^2_{r,min,bond}$	$\chi^2_{r,max,bond}$
kT=10	28.9	614.3
kT=1	14.9	231.2
kT=0.1	9.1	32.7
kT=0.01	0.2	5.7

F



"temperature"	$\chi^2_{r,min,bond}$	$\chi^2_{r,max,bond}$
kT=10	13.2	464.4
kT=1	8.8	157.4
kT=0.1	1.3	8.0
kT=0.01	0.7	1.5

G

Figure 5-46 Visualization of mutual (A) and Euler (B) angles describing steric orientation of the helices. (B) The helix *a* (purple) of each structure was oriented along the z axis with rotation angle $\Psi_a = 0$ (for more details see text above). Helix *b* and helix *c* are depicted in cyan and brown, respectively. (C,D) Confined conformational space allowed for the helices of structures in the "best" family of solutions after docking procedure of each molecule expressed in mutual (C) and Euler (D) angles. Crosses correspond to the position of "best" structures of the respective molecule. Unrestrained simulations (gray scale) together with "best" families after docking (magenta) and sampling (green) are shown on (E), (F), (G) for J(abc), J(abc(C2)) and J(abc(C5)), respectively. Change of the gray tones from light to dark (black) corresponds to the change in "bonds" enthalpy from bigger to smaller values, respectively.

With the help of an *in house* developed script mutual and Euler angles were recovered from the structures within the best clusters. Angles of the "best" structures as well as the distribution of the angular values within the cluster are presented in Table 30.

Table 30 Mutual and Euler angles (Figure 5-46A,B) of the "best" rigid body structures as well as mean values and their variation ranges within the "best" cluster of solutions for three three-way junction molecules. Values for Ψ_b are marked with red to demonstrate the biggest structural difference between J(abc) and J(abc(C2)).

Shape description						
	J(abc)		J(abc(2C))		J(abc(5C))	
Mutual angles	For the best structure	Mean ± error	For the best structure	Mean ± error	For the best structure	Mean ± error
angle (*ab*)	82.2°	87.9° ± 7°	77°	80° ± 6°	143.4°	133° ± 12°
angle (*ac*)	132.5°	124.7° ± 12°	123.2°	114° ± 6°; 129° ± 6°	96.1°	100.8° ± 8°
angle (*bc*)	144.6°	147.2° ± 9°	151°	150° ± 6°; 164° ± 5°	115.2°	114.5° ± 3°
Full geometric description						
	J(abc)		J(abc(2C))		J(abc(5C))	
Euler angles	For the best structure	Mean ± error	For the best structure	Mean ± error	For the best structure	Mean ± error
φ_b	53.9°	44.8° ± 12°	61.8°	56° ± 10°	40.8°	44.2° ± 4°
θ_b	82.2°	87.9° ± 7°	77°	80° ± 6°	143.4°	133° ± 12°
Ψ_b	-39.8°	-32.3° ± 8°	-85.6°	-135° ± 7°;-86° ± 8°	-133.5	-138.8° ± 13°
φ_c	-118.2°	-125.8° ± 13°	-141.1°	-135° ± 10°	-108.7°	-100.5° ± 9°
θ_c	132.5°	124.7° ± 12°	123.2°	114° ± 6°; 129° ± 6°	96.1°	100.8° ± 8°
Ψ_c	96°	95.8° ± 12°	110.7°	105.4° ± 7°	124.2°	131° ± 20°

5.5.5.10 **FRET and bond driven structures**

MMC unrestrained sampling was performed at the kT = 0.1 “temperature” as described in section 5.4.8. Here, however, initial structures were chosen from the “best” family models generated with RBD. Mutual angles of the generated structures were calculated and plotted in Figure 5-47.

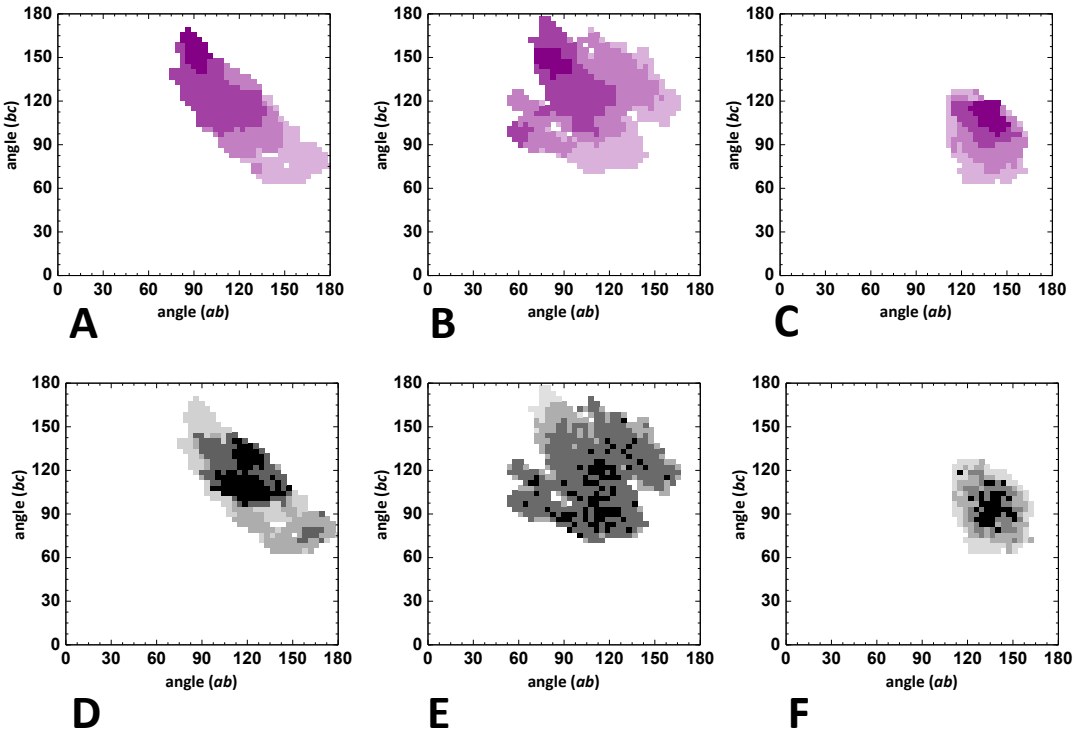


Figure 5-47 Unrestrained sampling with the “best” structures after RBD as an initial structures are shown on (A,D), (B,E) and (C,F) for J(abc), J(abc(C2)) and J(abc(C5)), respectively. Dark magenta (A-C) and dark gray (D-F) areas correspond to the lower $\chi^2_{r\_FRET}$ and enthalpy values, respectively.

Two different representation of the same data are shown on the upper (A-C) and lower (D-F) plots. In the first case the landscape of $\chi^2_{r,FRET}$ values are plotted with magenta tones, where the change of the color from dark to light corresponds to the change in the $\chi^2_{r,FRET}$ from low to high values. On the lower three plots gray colorcode represents the enthalpy of the structures, where the darker tones correspond to the lower values of enthalpy with the black being the areas with the lowest enthalpy values. Taking into account that the starting

structures for each molecule were located in the dark magenta areas (upper plots), one can see that the unrestrained sampling explores the area around the initial structures for a while and then moves away from that space, exploring areas with significantly larger $\chi^2_{r,\mathrm{FRET}}$ values (light magenta). Moreover, from the pairwise comparison of the upper and lower plots Figure 5-47 (A,D), (B,E) and (C,F) for J(abc), J(abc(C2)) and J(abc(C5)), respectively, one can clearly see the different locations of the dark magenta and black areas in the conformational space. This is another demonstration of the fact that FRET and bond driven structures occupy different conformational space.

5.5.5.11 **Conformational Space – MD**

For demonstration of the conformational space sampled by the MD, "best" MD structures according to FRET, as well as for comparison with RBD results, Euler and mutual angles for the simulated structures were calculated, as described in 5.5.5.7.2.

In Figure 5-48 the histogram of the structure appearences is ploted as gray area. Light gray was used to demonstrate least populated areas. Darker tones of gray were used to demonstrate an increased number of structures in a certain area. MD structures with the best agreement with FRET data (structures below the threshold from Figure 5-41, "best" MD family) and the contour of the area with the "best" family structures after docking are shown in green and magenta, respectively (see Figure 7-24 for conformational space represented in Euler angles).

Evaluating the results for J(abc) (Figure 5-48A) one can see that the MD favours the region (black area) far from the FRET preferred region, and only few structures from the "best" MD family explore the conformational space inside the magenta contour. For J(abc(C2)) (Figure 5-48B) this looks slightly better as MD and FRET prefer quite similar areas in conformational space and the "best" MD stuctures also lie in the same area. In case of J(abc(C5)) (Figure 5-48C) neither MD preferred structures nor "best" MD family structures overlap with the conformational space of the "best" family structures after docking.

It is important to mention that $\chi^2_{r,\mathrm{min}}$ of the best MD structure is already significantly bigger than for the RBD structure (see Table 26). Hence, the "best" family structures of MD are already quite different from the RBD "best" structure and are different between each other. This is the reason why there is such a big scattering of green dots in the

Figure **5-48**. Another point is that from visual inspection of the "best" MD and RBD structures, particularly for J(abc(C5)) molecule (Figure 5-42,C), one could expect that mutual angles should be similar for the respective structures. However, from the large differences in RMSD values (see Table 26) and flatness parameters (Table 28) one can conclude that the respective structures are quite different from each other and only visual view of the 3D structures might be deceptive.

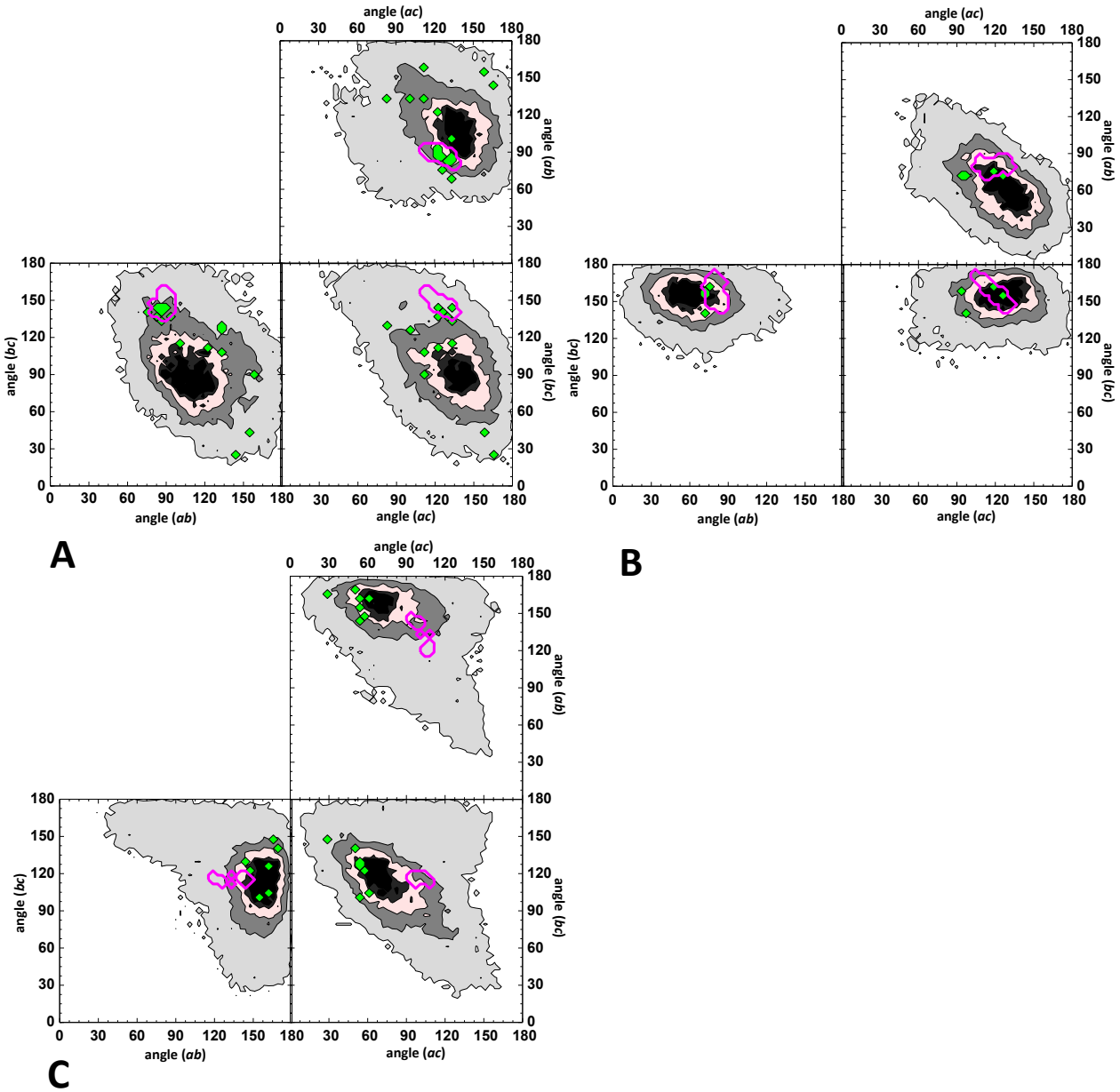


Figure 5-48 Mutual angles calculated for the structures of free MD simulations (gray scale) together with MD "best" family structures according to FRET (green) and the contour of the area with the "best" family structures after docking are shown on (A), (B), (C) for J(abc), J(abc(C2)) and J(abc(C5)), respectively. It should be noted that as the presented data consist of several trajectories with slightly different initial structures, reweighting of the data is necessary. Hence, the appearences histogram might be slightly changed.

From RMSD values and from the visual inspection of the structures (sections 5.5.5.7 and 5.5.5.10) one can conclude that the "best" structures from MD and RBD are not identical although in some cases they explore the same conformational space. If for cases of J(abc) and J(abc(C5)) the rotation of some helices and consequent displacement of the other helices seem to be the main difference between the structures, for J(abc(C2)) the main difference is at the edges of the helices, which is explained with the flexibility of the helices during MD simulations.

For the "best" MD and RBD structures, their χ^2_r values, as well as RMSD values between them see Figure 5-42A,B,C and Table 26.

5.5.5.12 **Conclusion**

Comparison of RBD and MMC strategies described in sections 5.5.5.6.1 and 5.5.5.6.2.

RMSD values of the "best" structures generated with two strategies demonstrate the similarities of the obtained models (Table 25). Particularly for structures obtained for J(abc(C5)) molecule RMSD difference of 0.3 Å makes them almost identical.

Comparison of the obtained 3D structures.

3D structure comparison of three- and four-way junctions (section 5.5.4) reveals that after omitting helix *d*, the structure becomes nearly planar (flat) (Figure 5-31A-D). Noteworthy is that the resulting fully paired three-way junction has a Y shape, and only after addition of the bulges coaxial stacking occurs. This can probably be explained with the lower tension in the junction region for J(abc(C2)), where stacking of helices *a* and *c* occurs (Figure 5-31A,B). Similar results were obtained for DNA three-way junction molecules after addition of bulges in the junction region [112]. Moreover, Stühmeier *et al*. found that the acute angle between two helices becomes even smaller by increasing the number of unpaired nucleotides in the bulge. In contrary to them, current results show that addition of a bigger 5C bulge in the same position causes a much more complex change in the structure. Now, coaxially stacked

helix pair is changed (helices *a* and *b*) and the third one (helix *c*) is perpendicular to the other two (Figure 5-31C). Tertiary contacts between the relatively large bulge and one of the helices could be a possible explanation. The reasons for this rearrangement may be rather complex and cannot be elucidated by smFRET experiments.

Noteworthy, the deviation from the planarity of the studied three-way junction molecules is very small irrespective whether there is bulge or not in the junction. Another notable fact is that two helices are coaxially stacked for bulged molecules, which is in a good agreement with solid phase data of natural RNA three-way junction systems from literature [5, 10]. For fully paired RNAs no studies were performed so far to be compared with.

Conformational space studies.

By drawing a conclusion only from Figure 5-46C, one could say that models obtained with the first strategy for J(abc) and J(abc(C2)) occupy similar conformational space and therefore have similar 3D structures. However, the largest difference is in the rotation of the helices along their own axis, which is particularly big for helix *b*. From Figure 5-46D one can clearly see that after addition of the 2C bulge in the junction region helix *b* rotates during the modeling for about 50° - 100° and for the "best" structures Ψ_b (shown as crosses in Figure 5-46C,D) changes from -39.8° to -85.6° (see Table **30** for the values of all mutual and Euler angles and mean values of angular distributions).

Another notable fact is that the "best" structures of each molecule are far away from each other and do not "belong" to the each other's conformational space, despite the fact that later can be overlapping.

Additionally, one can notice that Ψ_b for the model obtained with the RBD for J(abc(C2)) varies in quite a big range of values which can be explained with the fact that the FRET measurements are in general not sensitive enough to the rotation of the helices around their axes, which in its turn leads to uncertainty in respect to the bulge side discussed in 5.5.5.6.1 and 5.5.5.7.2.

Figure 5-46E,F,G shows the landscape of potential "bond" energy (enthalpy), which consists of the "bond" and clash energies (between two atoms, connected with the bonds introduced

for simulation procedure) of the structures, generated by unrestrained simulation (no FRET restraints).

Applying different thresholds on the enthalpy values gives possibility to differentiate less clashing structures with small violations of the bonds (dark gray) from more clashing structures with larger violation of the bonds (light gray) shown in Figure 5-46E,F,G. Black area, corresponding to the structures with the strictest threshold on the enthalpy (lowest "bond" + clash energy), is not overlaying with the area preferred by the FRET driven structures (magenta and green areas).

This demonstrates that even when the violations of the "bonds" are low and there are minimal clashes between atoms in the junction region, the obtained structures are still far from being the "best" structures. This, in its turn, indicates that FRET information is mandatory for obtaining the "best" structures.

It is particularly important that the green area ("best" family after MMC) completely lies in the magenta ("best" family after RBD) contour for each of the three-way junctions (Figure 5-46E,F,G). The significance of this fact is that after FRET screening of the MMC data, which was performed over a wide "temperature" range, allowing big violations of the covalent, hydrogen and artificial "bonds" at higher "temperatures", obtained structure are still similar to the RBD structures.

This demonstrates one more time the unambiguity of the obtained models. Another conclusion is that the "bonds" used in the simulations are not too strong and do not play a decisive role in achievement of the "best" structures.

Conclusion MD

Summarizing the results (section 5.5.5.7) one can clearly note that the "best" RBD structures are in better agreement with the FRET data then the "best" MD structures. However, in contrast to MD, in the docking procedure a simplified stereochemistry of the helices is used and processes like stacking interactions, long range interactions, interactions with solvent are not taken into account. In this respect MD structures are more sophisticated and presumably more similar to the natural ones.

5.5.6 Incomplete molecules

As mentioned in section 5.5.5.4 incomplete (consisted of two strands) variants of the fully paired thre-way junction molecules were investigated Figure 5-26 (see Table 33 and Table 34 for fitting results). First reason for these investigations was to understand whether the minor population in analysis of fully paired molecules can be explained with the presence of incomplete molecules in the measurement solutions. Second reason was to study the behavior of two strands, which are only partially complementary to each other, whether proper basepairing occurs inbetween of the strands and whether the overall system is dynamic or static.

As already mentioned in section 5.5.5.4, in the analysis of the 33 out of 44 three way junctions a minor peak was observed. 16 out of 33 cases could be explained by the presence of incomplete molecules during the measurements. In other 16 cases origin is still unknown, however in 6 out of these 16 molecules, minor population was too small (< 5 %) to be considered as a minor state. Additionally in 1 case minor peak was shown to appear due to the acceptor photophysical properties, discussed in section 5.5.5.5. In the rest of 11 cases no minor peak was observed, and the distances obtained from the PDA analysis of the respective incomplete molecules were shifted from the fully paired cases, hence, it is very unlikely that minor peaks in three-way junction molecules are overlapped by the major peaks (see Table 31 for the possible origin of the minor peaks for each J(abc) molecule).

Table 31 Full list of J(abc) molecules and the probable origin of the minor peak observed in PDA analysis (see section 7.5.4 for all data and fit plots)

#	DA-pair	origin of the minor peak
1	**(D)γ29a/(A)β14c/δα–Δd**	unknown
2	**(D)γ7b/(A)β14c/δα–Δd**	unknown
3	**(D)γ8b/(A)β14c/δα–Δd**	unknown
4	**(D)γ29a/(A)β27b/δα–Δd**	minor is incomplete
5	**(D)γ29a/(A)β33b/δα–Δd**	unknown, too small(< 5 %)
6	**(A)γ12b/(D)β11c/δα–Δd**	minor is incomplete
7	**(A)γ12b/(D)β14c/δα–Δd**	minor is incomplete
8	**(A)γ12b/(D)β5c/δα–Δd**	no minor peak (and incomplete is far away, hence, there is probably no minor-major overlapping occurs

9	**(A)γ12b/(D)β8c/δα–Δd**	minor is incomplete
10	**(A)γ24a/(D)β11c/δα–Δd**	no minor peak (and incomplete is far away, hence, there is probably no minor major overlapping occurs
11	**(A)γ24a/(D)β14c/δα–Δd**	no minor peak (and incomplete is far away, hence, there is probably no minor major overlapping occurs
12	**(A)γ24a/(D)β27b/δα–Δd**	minor is incomplete
13	**(A)γ24a/(D)β29b/δα–Δd**	minor is incomplete
14	**(A)γ24a/(D)β5c/δα–Δd**	unknown
15	**(A)γ24a/(D)β8c/δα_Δd**	no minor peak (and incomplete is far away, hence, there is probably no minor major overlapping occurs
16	**(A)γ8b/(D)β11c/δα–Δd**	unknown
17	**(A)γ8b/(D)β5c/δα–Δd**	unknown, too small (< 5 %)
18	**(A)γ8b/(D)β8c/δα–Δd**	minor is incomplete
19	**(A)δα_Δd10a/(D)β11c/γ**	unknown, too small (< 5 %)
20	**(A)δα_Δd10a/(D)β14c/γ**	unknown
21	**(A)δα_Δd10a/(D)β27b/γ**	no minor peak (and incomplete is far away, hence, there is probably no minor major overlapping occurs
22	**(A)δα_Δd10a/(D)β29b/γ**	minor is incomplete (< 5 %)
23	**(A)δα_Δd10a/(D)β5c/γ**	unknown
24	**(A)δα_Δd10a/(D)β8c/γ**	minor is incomplete
25	**(A)δα_Δd10a/(D)γ7b/β**	minor is incomplete
26	**(A)δα_Δd10a/(D)γ8b/β**	unknown, too small(< 5 %)
27	**(D)β27b/(A)δα–Δd26c/γ**	minor is incomplete
28	**(D)β29b/(A)δα–Δd26c/γ**	no minor peak (if there were a peak, it would be due to incomplete)
29	**(D)γ29a/(A)δα–Δd26c/β**	minor is incomplete
30	**(D)γ7b/(A)δα–Δd26c/β**	unknown
31	**(D)γ8b/(A)δα–Δd26c/β**	unknown, too small (< 5 %)
32	**(D)β27b/(A)δα–Δd28c/γ**	minor is incomplete
33	**(D)β29b/(A)δα–Δd28c/γ**	minor is incomplete
34	**(D)γ29a/(A)δα–Δd28c/β**	no minor peak (if there were a peak, it would be due to incomplete)
35	**(D)γ7b/(A)δα–Δd28c/β**	unknown
36	**(D)γ8b/(A)δα–Δd28c/β**	unknown
37	**(A)δα_Δd7a/(D)β11c/γ**	no minor peak (no incomplete molecule was present in the measurement)
38	**(A)δα_Δd7a/(D)β14c/γ**	minor is incomplete
39	**(A)δα_Δd7a/(D)β27b/γ**	unknown, too small (< 5 %)
40	**(A)δα_Δd7a/(D)β29b/γ**	no minor peak (and incomplete is far away, hence, there is probably no minor major overlapping occurs
41	**(A)δα_Δd7a/(D)β5c/γ**	no minor peak (no incomplete molecule was present in the measurement)

42	**(A)δα_Δd7a/(D)β8c/γ**	no minor peak (no incomplete molecule was present in the measurement)
43	**(A)δα_Δd7a/(D)γ7b/β**	unknown
44	**(A)δα_Δd7a/(D)γ8b/β**	minor is incomplete

For the structural invesigations of the incomplete molecules prediction of the possible conformations was performed with the mfold web server [113]. Predictions were performed for each pair of the strands, separately. Mfold computes the free energies of the predictied structures which indicate more and less favourable structures in the nature (Figure 5-49A-N). The only thing is clear that in the predicted structures complementary parts of the strands form fully paired helices. No other conlculsion is possible, since dynamic or static behavior of the molecule would depend on which part of the molecule was labelled.

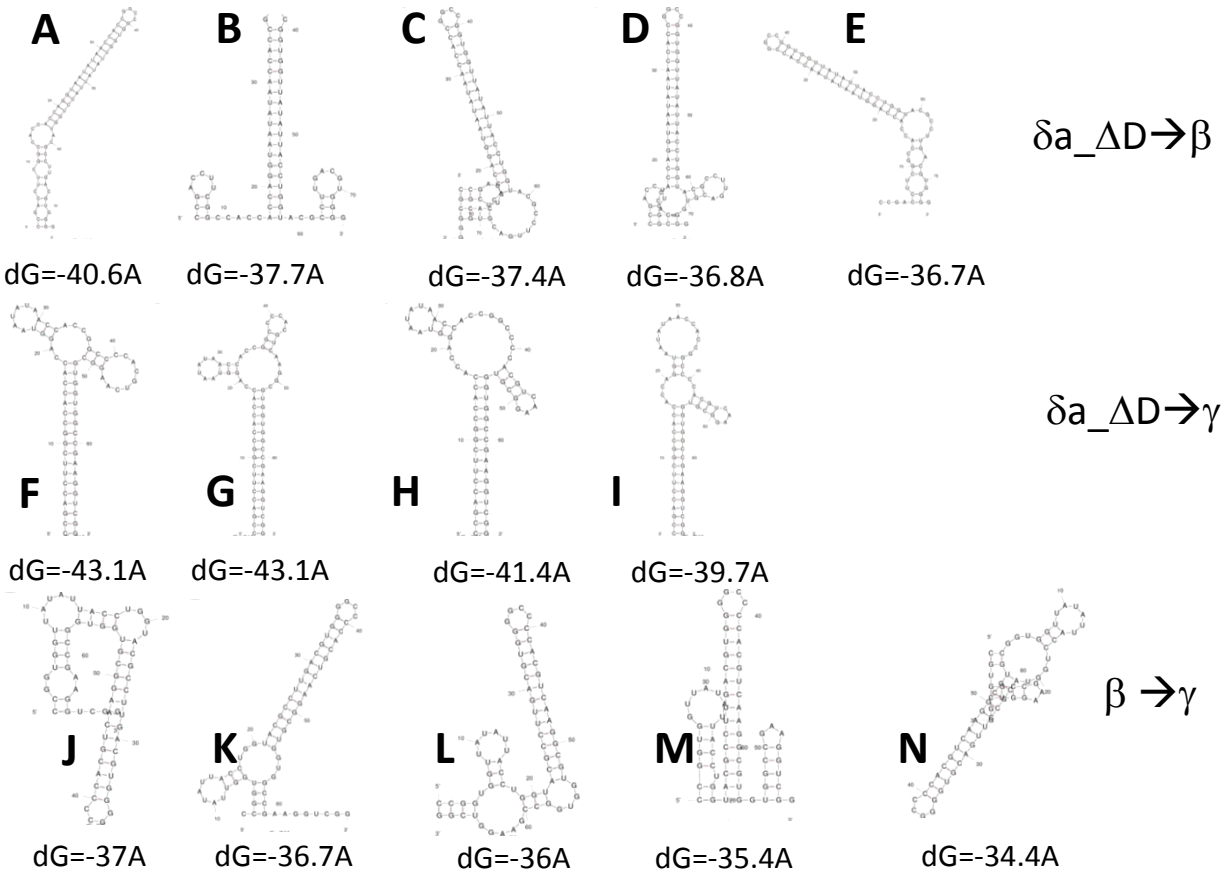


Figure 5-49 Prediction results fot J(abc) molecule performed with mfold. dG is the free energy of the predictied structure. The lower dG the more probable is the structure. First, second and third rows represent the structures obtained from δa_ΔD and β stands, δa_ΔD and γ strands and β and γ strands, respectively.

From the overview of the results for incomplete molecules (Table 32) one can see that only in 5 cases out of 44 static behaviour of the molecule was observed, in all other cases the deviation from the static FRET line (section 5.4.4) on the F_D/F_A vs τ_D plots was observed (Figure 5-50 A,B). This allows one to conclude that the incomplete molecules demonstrate dynamical behavior within the observation time.

Table 32 List of molecules, labelling scheme (labelled helices and strands) and the position with respect to static FRET line.

DA-pair	labelled helices	labelled strands	on/off the static FRET line
(D)γ29a/(A)β14c	A->C	γ–>β	off
(D)γ7b/(A)β14c	B->C	γ–>β	off
(D)γ8b/(A)β14c	B->C	γ–>β	off
(D)γ29a/(A)β27b	A->B	γ–>β	off
(D)γ29a/(A)β33b	A->B	γ–>β	yes
(A)γ12b/(D)β11c	B->C	β–>γ	off
(A)γ12b/(D)β14c	B->C	β–>γ	yes
(A)γ12b/(D)β5c	B->C	β–>γ	off
(A)γ12b/(D)β8c	B->C	β–>γ	on
(A)γ24a/(D)β11c	A->C	β–>γ	off
(A)γ24a/(D)β14c	A->C	β–>γ	off
(A)γ24a/(D)β27b	A->B	β–>γ	off
(A)γ24a/(D)β29b	A->B	β–>γ	off
(A)γ24a/(D)β5c	A->C	β–>γ	off
(A)γ24a/(D)β8c	A->C	β–>γ	off
(A)γ8b/(D)β11c	B->C	β–>γ	off
(A)γ8b/(D)β5c	B->C	β–>γ	yes
(A)γ8b/(D)β8c	B->C	β–>γ	yes
(A)δα_Δd10a/(D)β11c	A->C	β–> δα_ΔD	off
(A)δα_Δd10a/(D)β14c	A->C	β–> δα_ΔD	off
(A)δα_Δd10a/(D)β27b	A->B	β–> δα_ΔD	off
(A)δα_Δd10a/(D)β29b	A->B	β–> δα_ΔD	off
(A)δα_Δd10a/(D)β5c	A->C	β–> δα_ΔD	off
(A)δα_Δd10a/(D)β8c	A->C	β–> δα_ΔD	off
(A)δα_Δd10a/(D)γ7b	A->B	γ–> δα_ΔD	on
(A)δα_Δd10a/(D)γ8b	A->B	γ–> δα_ΔD	off
(D)β27b/(A)δα–Δd26c	B->C	β–> δα_ΔD	off
(D)β29b/(A)δα–Δd26c	B->C	β–> δα_ΔD	off
(D)γ29a/(A)δα–Δd26c	A->C	γ–> δα_ΔD	off
(D)γ7b/(A)δα–Δd26c	B->C	γ–> δα_ΔD	off
(D)γ8b/(A)δα–Δd26c	B->C	γ–> δα_ΔD	off
(D)β27b/(A)δα–Δd28c	B->C	β–> δα_ΔD	off
(D)β29b/(A)δα–Δd28c	B->C	β–> δα_ΔD	off
(D)γ29a/(A)δα–Δd28c	A->C	γ–> δα_ΔD	off
(D)γ7b/(A)δα–Δd28c	B->C	γ–> δα_ΔD	off

(D)γ8b/(A)δα–Δd28c	B->C	γ–> δα_ΔD	Off
(A)δα_Δd7a/(D)β11c	A->C	β–> δα_ΔD	off
(A)δα_Δd7a/(D)β14c	A->C	β–> δα_ΔD	off
(A)δα_Δd7a/(D)β27b	A->B	β–> δα_ΔD	off
(A)δα_Δd7a/(D)β29b	B->A	β–> δα_ΔD	yes
(A)δα_Δd7a/(D)β5c	A->C	β–> δα_ΔD	off
(A)δα_Δd7a/(D)β8c	A->C	β–> δα_ΔD	off
(A)δα_Δd7a/(D)γ7b	B->A	γ–> δα_ΔD	yes
(A)δα_Δd7a/(D)γ8b	A->B	γ–> δα_ΔD	off

More detailed investigation is necessery to undestand the structural behaviour of incomplete molecules; however this is out of the scope of current work.

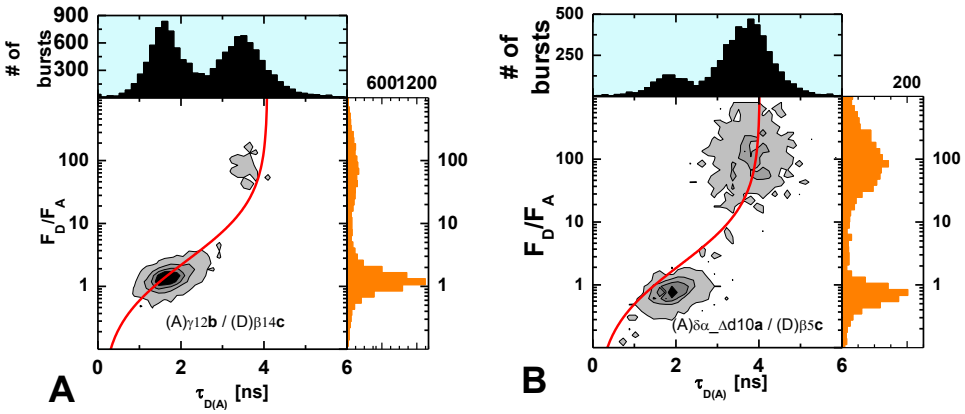


Figure 5-50 A typical examples of static (A) and dynamic (B) behaviour of the incomplete (A)γ12b/(D)β14c (A) and (A)δα_Δd10a/(D)β5c (B) molecules.

DA-pair	Equation type	Color	Equation number	Equation
(A)γ12b/(D)β14c	Static FRET line corrected for linker movement	red	Eq. 5.4-5	F_D/F_A = 0.7723/0.2860)/((4.0791/((0.0372*x^3)+(0.2617*x^2)+0.5629*x-0.0518))-1)
(A)δα_Δd10a/(D)β5c	Static FRET line corrected for linker movement	red	Eq. 5.4-5	F_D/F_A =(0.7807/0.3450)/((4.0159/((-0.0388*x^3)+(0.2754*x^2)+0.5300*x-0.0482))-1)

5.5.6.1 Distances and errors for measurements of incomplete molecules at 20mM $MgCl_2$

Table 33 Values for measured distances for incomplete molecules $\langle R_{DA}\rangle_E$, their relative amplitudes A and the values of apparent σ, measurement errors ΔR_{DA} (resulting from $\Delta R_{DA}(\kappa^2)$ and $\Delta R_{DA}(E)$, see section 5.4.6) and model distances R_{model} resulting from rigid body docking. See section 7.5.4 for all data and fit plots.

DA-pair	σ_{app},%	$\langle R_{DA}\rangle_{E1}$, Å	A_1, %	$\langle R_{DA}\rangle_{E2}$, Å	A_2, %	$\langle R_{DA}\rangle_{E3}$, Å	A_3, %
(D)γ29a/(A)β14c	5.0	43.7	62.1	35.4	37.9		
(D)γ7b/(A)β14c	4.5	51	100.0				
(D)γ8b/(A)β14c	5.0	48.3	65.1	43.4	34.9		
(D)γ29a/(A)β27b	4.4	49.6	54.8	43.9	24.0	58.3	21.2
(D)γ29a/(A)β33b	4.5	62.2	61.2	74.4	36.6	52.5	2.2
(A)γ12b/(D)β11c	4.5	50.4	55.6	74.9	44.4		
(A)γ12b/(D)β14c	5.2	47.5	96.3	60	3.7		
(A)γ12b/(D)β5c	4.9	44.8	100.0				
(A)γ12b/(D)β8c	5.1	47.2	93.2	57.2	6.8		
(A)γ24a/(D)β11c	4.0	44	65.9	37.3	34.1		
(A)γ24a/(D)β14c	5.0	40.9	66.0	33.7	19.1	47.2	14.9
(A)γ24a/(D)β27b	4.0	45.4	56.8	51.8	43.2		
(A)γ24a/(D)β29b	4.6	51.9	88.7	38.5	11.3		
(A)γ24a/(D)β5c	4.0	43.5	80.2	54.1	10.1	33.4	9.7
(A)γ24a/(D)β8c	4.0	47.5	77.4	55.1	12.2	38.3	10.4
(A)γ8b/(D)β11c	5.0	54.4	87.1	46.9	12.9		
(A)γ8b/(D)β5c	4.5	47	76.3	56.1	12.1	40.7	11.6
(A)γ8b/(D)β8c	4.5	54.2	94.6	44.5	5.4		
(A)δα_Δd10a/(D)β11c	4.5	48.5	77.4	53.7	22.6		
(A)δα_Δd10a/(D)β14c	4.5	46.6	71.5	41.7	25.0	56.8	3.5
(A)δα_Δd10a/(D)β27b	4.0	28	76.5	40.2	23.5		
(A)δα_Δd10a/(D)β29b	4.3	53.4	68.0	43.2	32.0		
(A)δα_Δd10a/(D)β5c	4.0	43	90.8	48.3	9.2		
(A)δα_Δd10a/(D)β8c	4.2	52.7	95.3	64.5	4.7		
(A)δα_Δd10a/(D)γ7b	4.8	52.2	92.5	45	7.5		
(A)δα_Δd10a/(D)γ8b	4.5	48.9	88.5	40.9	11.5		
(D)β27b/(A)δα–Δd26c	5.0	55.7	57.5	49.8	42.5		

(D)β29b/(A)δα–Δd26c	5.0	59.4	74.0	52.3	26.0		
(D)γ29a/(A)δα–Δd26c	5.0	58.6	61.5	68.2	38.5		
(D)γ7b/(A)δα–Δd26c	4.4	40.9	82.6	47	17.4		
(D)γ8b/(A)δα–Δd26c	4.5	46.3	53.1	39.8	46.9		
(D)β27b/(A)δα–Δd28c	4.5	60.9	63.6	53.7	36.4		
(D)β29b/(A)δα–Δd28c	5.0	64	81.4	53.5	18.6		
(D)γ29a/(A)δα–Δd28c	4.0	68.4	56.3	60.6	35.2	56.2	8.5
(D)γ7b/(A)δα–Δd28c	5.0	39.5	62.7	47.3	37.3		
(D)γ8b/(A)δα–Δd28c	4.0	46.7	64.9	38.4	35.1		
(A)δα_Δd7a/(D)β11c	4.0	56	52.4	48.9	47.6		
(A)δα_Δd7a/(D)β14c	4.5	49.3	64.5	43.7	35.5		
(A)δα_Δd7a/(D)β27b	4.0	45.8	47.8	31.5	36.6	38	15.6
(A)δα_Δd7a/(D)β29b	5.5	40.9	59.2	27.7	40.8		
(A)δα_Δd7a/(D)β5c	4.0	63.4	58.1	56.5	39.2	48.4	2.7
(A)δα_Δd7a/(D)β8c	4.0	56.7	84.4	64.5	10.7	49.7	4.9
(A)δα_Δd7a/(D)γ7b	3.9	50.8	100.0	0			
(A)δα_Δd7a/(D)γ8b	4.5	52.4	96.0	42.7	4.0		

5.5.6.2 Fit parameters for all PDA analysis for incomplete molecules

Table 34 Full PDA fit parameters for the incomplete molecules measured at 20 mM $MgCl_2$. See section 7.5.4 for all data and fit plots.

DA-pair	σ_{app},%	$\langle R_{DA}\rangle_{E1}$, Å	A_1, %	$\langle R_{DA}\rangle_{E2}$, Å	A_2, %	$\langle R_{DA}\rangle_{E3}$, Å	A_3, %	impurities R_{DA}, Å	impurities A, %	A_{Donly}, %	χ^2_r
(D)γ29a/(A)β14c	5.0	43.7	17.7	35.4	10.8			66.1	3.8	67.7	1.74
(D)γ7b/(A)β14c	4.5	51	12							88	1.06
(D)γ8b/(A)β14c	5.0	48.3	14	43.4	7.5			66.1	4.2	74.3	1.53
(D)γ29a/(A)β27b	4.4	49.6	19.4	43.9	8.5	58.3	7.5	75	2.7	61.9	1.03
(D)γ29a/(A)β33b	4.5	62.2	18.9	74.4	11.3	52.5	0.7			69.1	0.69
(A)γ12b/(D)β11c	4.5	50.4	47.4	74.9	37.8					14.8	2.55
(A)γ12b/(D)β14c	5.2	47.5	65.8	60	2.5			82.7	7.5	24.2	1.47
(A)γ12b/(D)β5c	4.9	44.8	11.9					66	4.1	84.1	1.14
(A)γ12b/(D)β8c	5.1	47.2	64.3	57.2	4.7			73.7	4.1	27.1	1.97
(A)γ24a/(D)β11c	4.0	44	8.9	37.3	4.6			69.9	1.1	85.4	1.72
(A)γ24a/(D)β14c	5.0	40.9	16.9	33.7	4.9	47.2	3.8	74.1	3.3	71.2	1.12

(A)γ24a/(D)β27b	4.0	45.4	21.3	51.8	16.2			66.5	0.8	61.7	1.22
(A)γ24a/(D)β29b	4.6	51.9	1.26	38.5	0.161			71-2	2-Feb	40.4	1.25
(A)γ24a/(D)β5c	4.0	43.5	18.2	54.1	2.3	33.4	2.2	67.7	1.9	75.4	1.64
(A)γ24a/(D)β8c	4.0	47.5	26.1	55.1	4.1	38.3	3.5	69.6	2.3	64	1.27
(A)γ8b/(D)β11c	5.0	54.4	33.9	46.9	5					61.2	0.96
(A)γ8b/(D)β5c	4.5	47	30.2	56.1	4.8	40.7	4.6	69.8	3.2	57.3	1.16
(A)γ8b/(D)β8c	4.5	54.2	37	44.5	2.1			69.1	3.6	57.3	1.35
(A)δα_Δd10a/(D)β11c	4.5	48.5	46.9	53.7	13.7			75.7	7.7	31.6	1.26
(A)δα_Δd10a/(D)β14c	4.5	46.6	48.7	41.7	17	56.8	2.4	74.3	6.6	25.3	1.61
(A)δα_Δd10a/(D)β27b	4.0	28	15.3	40.2	4.7			57.4	0.6	79.4	1.81
(A)δα_Δd10a/(D)β29b	4.3	53.4	5.1	43.2	2.4			70.8	2.7	89.7	2.11
(A)δα_Δd10a/(D)β5c	4.0	43	26.6	48.3	2.7			62.6	1.8	68.9	1.59
(A)δα_Δd10a/(D)β8c	4.2	52.7	69.2	64.5	3.4					27.4	1.46
(A)δα_Δd10a/(D)γ7b	4.8	52.2	64.5	45	5.2			77.4	4	26.3	1.74
(A)δα_Δd10a/(D)γ8b	4.5	48.9	46.8	40.9	6.1			74.3	2.2	44.9	1.32
(D)β27b/(A)δα–Δd26c	5.0	55.7	16.8	49.8	12.4			61.8	1.5	66.6	1.8
(D)β29b/(A)δα–Δd26c	5.0	59.4	21.6	52.3	7.6			78.7	11.9	59.4	1.71
(D)γ29a/(A)δα–Δd26c	5.0	58.6	11.5	68.2	7.2			90.8	13.6	67.8	1.31
(D)γ7b/(A)δα–Δd26c	4.4	40.9	21.4	47	4.5			75.8	9.1	65	1.77
(D)γ8b/(A)δα–Δd26c	4.5	46.3	6.9	39.8	6.1			78.5	9.5	77.5	1.67
(D)β27b/(A)δα–Δd28c	4.5	60.9	43.1	53.7	24.7					32.2	1.68
(D)β29b/(A)δα–Δd28c	5.0	64	25.8	53.5	5.9			81.7	11.4	56.9	1.26
(D)γ29a/(A)δα–Δd28c	4.0	68.4	26.4	60.6	16.5	56.2	4			52.2	1.17
(D)γ7b/(A)δα–Δd28c	5.0	39.5	17.8	47.3	10.6			75.4	10.7	68	0.64
(D)γ8b/(A)δα–Δd28c	4.0	46.7	10	38.4	5.4					84.6	1.77
(A)δα_Δd7a/(D)β11c	4.0	56	35.1	48.9	31.9			68.6	3.6	29.4	1.94
(A)δα_Δd7a/(D)β14c	4.5	49.3	9.8	43.7	5.4			65.6	1	83.8	1.04
(A)δα_Δd7a/(D)β27b	4.0	45.8	15.3	31.5	11.7	38	5	61.8	1.5	66.6	1.8
(A)δα_Δd7a/(D)β29b	5.5	40.9	28.4	27.7	19.6					52	2.19
(A)δα_Δd7a/(D)β5c	4.0	63.4	32.2	56.5	21.7	48.4	1.5	71.9	2.3	42.4	0.54
(A)δα_Δd7a/(D)β8c	4.0	56.7	58.3	64.5	7.4	49.7	3.4	73.9	1.4	29.5	1.15
(A)δα_Δd7a/(D)γ7b	3.9	50.8	62.5	0	0					37.5	1.33
(A)δα_Δd7a/(D)γ8b	4.5	52.4	55.6	42.7	2.3					42.1	1.47

6 Outlook

For studies of biomolecular dynamics by FCS on nanosecond to second timescale, a low cost hardware correlator, as an alternative to the current commercial devices, was developed. In the second part the structures of RNA branched molecules were determined by state of the art measurement and analysis techniques. Here, single-molecule techniques were applied as a quantitative tool providing detailed structural insights along with associated precision of the structural models.

Hardware correlator

Here a fast hardware photon correlator implemented in a field-programmable gate array (FPGA) combined with a compact confocal fluorescence setup is presented. The results of test measurements demonstrate that the correlator is comparable with the current generation of commercial devices. It has two independent units with a time resolution of 4 ns while utilizing less than 15 % of a low-end FPGA. The number of applications of FCS for life science [1, 114, 115] and material sciences [116] has increased dramatically in the last years. FCS plays a particularly important role in mobility studies of proteins and DNA or RNA fragments within the cytosol or other cell organelles [117]. This technique is also applied in studies of molecular (e.g. ligand receptor) interactions or translocation processes [2]. Therefore it is obvious that flexible and low cost instrumentation with high performance will facilitate the use of FCS for even more applications in education, applied sciences and basic research.

RNA studies

Branched RNA junctions are often dynamic and heterogeneous. Thus, the application of traditional methods (NMR, X-ray crystallography etc.) for structural determination is extremely difficult. Compared to traditional methods single-molecule fluorescence provides the advantage of avoiding averaging of distinct states. In my studies spectroscopic results were combined with advanced computer simulations in order to obtain a detailed molecular description of the structures based on FRET information. The single-molecule FRET (smFRET) experiments performed in this work were free of problems due to molecule immobilization

and surface artifacts as the measurements were performed on freely diffusing molecules. Along with the single-molecule advantage this makes smFRET particularly suitable for investigation of heterogeneous and dynamic systems. In this work a high Mg^{2+}-concentration (20 mM $MgCl_2$) was chosen for the measurement buffer, to avoid structural transitions within the observation time (~ 1 ms). FRET-averaged mean DA distances $\langle R_{DA} \rangle_E$ were obtained from the FRET measurement. Even though no dynamics was expected for our molecules, no TCSPC or FCS measurements were performed to study possible dynamics.

We demonstrated that with the FRET restrained high-precision structural modeling, structures of three coexisting conformers of RNA four-way junction and three three-way junction molecules can be resolved with precision of ~ 2 Å. Together with MD and coarse-grained simulations meaningful full atom structural models were obtained. Furthermore, this is the first time that coexisting transient minor conformers of an RNA four-way junction (J(abcd)) molecule were structurally solved. The sequence of the J(abcd) molecule was designed based on the sequence of the hairpin ribozyme. Therefore the presented results are structurally highly relevant.

Moreover, structural changes of the RNA three-way junction molecules, induced by the insertion of the bulges with different lengths in the junction region, were studied. Finally, conformational space, occupied by the RNA three-way junction models, was investigated, demonstrating the sterically accessible and FRET preferred conformational space.

Noteworthy, structural determination of fully paired RNA four- and three-way junction molecules, without tertiary contacts between helices or with other molecules, was never done before. Also, the influence of the bulge, as a single structural element, on the overall 3D structure of the RNA three-way junction molecules, was successfully investigated.

It is important to note that in order to achieve an atomistic picture of structural models, FRET has to be combined with computational methods [9, 118]. This combination of FRET with other techniques, e.g. SAXS or MD, makes each technique stronger in a hybrid approach [119, 120]. So it is obvious that the future of the structural biology is hybrid.

7 APPENDIX

7.1 Hardware correlator manuscript

Note: A 4 ns hardware photon correlator based on a general-purpose FPGA development board implemented in a compact setup for fluorescence correlation spectroscopy

Stanislav Kalinin[a)], Ralf Kühnemuth, Hayk Vardanyan, and Claus A.M. Seidel[a)]

Institut für Physikalische Chemie, Lehrstuhl für Molekulare Physikalische Chemie, Heinrich-Heine-Universität, Universitätsstraße 1, Geb. 26.32.02, 40225 Düsseldorf, Germany

(Received 2 July 2012; accepted 5 September 2012; published online 21 September 2012)

We present a fast hardware photon correlator implemented in a field-programmable gate array (FPGA) combined with a compact confocal fluorescence setup. The correlator has two independent units with a time resolution of 4 ns while utilizing less than 15 % of a low-end FPGA. The device directly accepts TTL signals from two photon counting detectors and calculates two auto- or cross-correlation curves in real time. Test measurements demonstrate that the performance of our correlator is comparable with the current generation of commercial devices. The sensitivity of the optical setup is identical or even superior to current commercial devices. The FPGA design and the optical setup both allow for a straightforward extension to multi-color applications. This inexpensive and compact solution with a very good performance can serve as versatile platform for uses in education, applied sciences and basic research.

Fluorescence correlation spectroscopy (FCS)[34, 121] is a powerful method to study diffusion and dynamics of biomolecules[122-125]. Fluctuations of the fluorescence signal $F(t)$ are characterized by auto- or cross-correlation functions $G_{ij}(t_c)$:

$$G_{ij}(t_c) = \left\langle F_i(t+t_c)F_j(t)\right\rangle / \left(\left\langle F_i(t)\right\rangle\left\langle F_j(t)\right\rangle\right) \quad (1),$$

where t_c is the correlation (or lag) time, and subscripts i and j refer to distinct detection channels. For calculation of correlation functions, hardware photon correlators are often employed[126]. Nowadays, real-time software correlation is also possible[57-60], provided that data acquisition hardware with sufficient time resolution is available. However, in any case, a dedicated data acquisition or processing board is typically required to build even a simple FCS setup.

In this work we propose an FPGA-based implementation of a hardware photon correlator. Specifically, it is based on a general-purpose Xilinx SP605 evaluation board (Xilinx, USA) equipped with a value-line Spartan 6 FPGA chip (XC6SLX45T). Photon detectors can be

directly connected to the SP605 board requiring no additional custom-built hardware. Test measurements of fluorescence fluctuations of Rhodamine 110 show that the time resolution of 4 ns is easily achieved in practice. In this respect our design is comparable with "fast" versions of commercial devices, such as the ALV 6010/200. Two correlator units utilize less than 15 % of the FPGA's resources, which suggests that the design should fit even into low-end FPGAs. Alternatively, more correlators can be implemented in parallel without sacrificing time resolution, which is for instance useful for FRET-FCS experiments. Additional features of our correlator include a real-time display of photon count rates, and a display of the intensity trace with millisecond time resolution. FPGA firmware and operating software are available from the authors (http://www.mpc.uni-duesseldorf.de/seidel/software.htm).

While this manuscript was in preparation, Mocsár et al.[127] published an independent implementation of an FPGA correlator based on a National Instruments 7833R board. The design of Mocsár et al. is optimized for computing very large number of correlation functions; however, this feature is implemented at the cost of time resolution (100 ns for single curve, lower for multiple curves). Thus, our approach is rather complementary to that of Mocsár et al, being optimized for time resolution, low FPGA utilization (Supplementary Material, Table II), and applications such as confocal detection of freely diffusing molecules.

Performance of the correlator is demonstrated using a compact home-built optical setup employing only the minimum of required optical components, ensuring maximized optical throughput and stability of alignment.

Our photon correlator has a classical multi-tau architecture, as described in detail previously[57, 59, 128] (see also Supplementary Material, section 4). Specifically, it has 28 internal time cascades, each consisting of 8 equally spaced time bins. The time bin width for the first two cascades is Δt_c = 4 ns whereas it doubles for every next cascade starting from the third one. The initial part of the correlation function (~4.3 s) is thereby calculated internally using 224 time bins. Every 268 ms (synchronously with updating the last cascade), current states of all photon pair counters, total number of registered photons, and state of the last time bin are transferred to a PC using the provided USB (virtual serial port) interface. Higher cascades (here up to 128.8 s correlation time) are processed on the PC. These data are updated at most once per received frame, thus generating only a minimal load on the PC. To compute FCS curves, normalization is performed according to Eq. (2),

$$G_{ij}(t_c) = N_{\text{pairs}}(t_c) / \left[S_i S_j (t_{\max} - t_c) \Delta t_c \right] \qquad (2)$$

In Eq. (2) $N_{\text{pairs}}(t_c)$ is the number of photon pairs accumulated for the time bin t_c, S_i and S_j are the average signals in channels *i* and *j*, respectively, $t_{\max}$ is the total experiment time, and Δt_c is the bin width for t_c.

The detectors must be connected to the "USER_SMA_GPIO_N" and "USER_SMA_GPIO_P" inputs of the SP605, which by design form a differential input. This can potentially result in distortion of the signal and crosstalk between the channels. We found that for 3.3 V input signals (e.g. from modern single photon avalanche diodes) 6 dB attenuators are sufficient to reduce crosstalk and other artifacts below detection limits (see test measurements below). For other signal levels and/or pulse shapes optimal conditions may vary.

The inputs are sampled with an effective frequency of 500 MHz. The dead time is artificially set to 16 ns for each channel to avoid false detections due to e.g. pulse overshoot. There is no dead time between the channels and $G(t_c = 0)$ is also calculated in the case of cross-correlation. The maximum count rate is mainly limited by the size of photon pair counters,

which should not overflow more than once per transmitted frame (268 ms). We estimate that this could happen at >16 MHz per channel (average over 268 ms), which is well above optimal count rates of modern single photon detectors.

At first we correlated a pseudorandom test signal generated within the same FPGA chip using the Linear Feedback Shift Register algorithm[129]. We simulated interconversion of two states “1” and “2”, for which we expect an exponential term in the correlation function. The fit of the computed correlation curve recovers expected[130] parameters with an accuracy of <0.3 % (see Supplemental Material section 1 for details).

Next, FCS measurements of Rhodamine 110 (Rh110) diffusion and photophysics were performed using a home-built confocal setup (Fig. 1). The excitation source is a tunable Ar-ion laser (35-LAP-431-220, Melles-Griot, Bensheim, Germany) set to 496 nm and coupled via single mode fiber (Schäfter&Kirchhoff, Hamburg, Germany) to a modular system consisting of galvanized aluminum cubes connected via dove-tail adapters. The main dichroic (BS 500, AHF Tübingen, Germany) is mounted on a micrometer driven rotation and tilt manipulator inside the first cube. The objective (UPlanSApo 60x/1.2w, Olympus Hamburg, Germany) is attached to a z-micrometer (SM1Z, Thorlabs Dachau, Germany) at the exit port. A tube lens (f = 160 mm achromatic lens, Linos, Göttingen, Germany) focuses the fluorescence light leaving the second exit port onto a pinhole (Plano Wetzlar, Germany), which is mounted on an 8-position wheel. Pinhole sizes can be varied between 25 µm and 5 mm (70 µm was used in the presented experiments). The spatially filtered light is then collimated by a second lens (f = 100 mm achromatic lens, Linos) before being divided by a polarizing beam splitter (TSWP 633, Linos). After passing bandpass filters (HC525/39, AHF) to remove scattered laser light and limit the detected fluorescence range the two beams are finally focused by two plano-convex lenses (f = 10 mm, Linos) onto two single photon avalanche diodes (PD200A, MPD Bolzano, Italy or PerkinElmer SPCM AQR-14/AQRH-14). Detectors are attached via xy-micrometer manipulators. All optical components are broadband AR coated and can be exchanged quickly. The modular design allows for easy extension, i.e. to set up four, six or eight detection channels to record multiparameter data (polarization and various spectral ranges). The sensitivity of the whole setup was found to be at least as sensitive as commercial microscopes like the Olympus IX71 equipped with equivalent detectors.

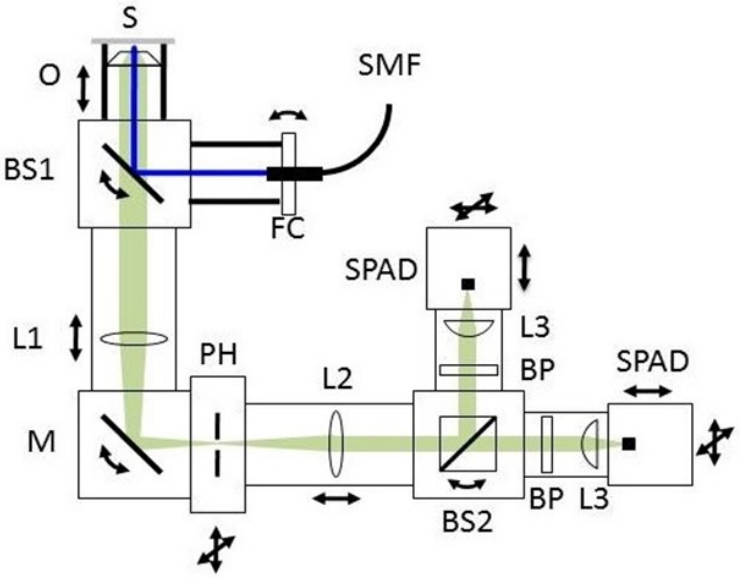


FIG. 1. Schematic drawing of the FCS setup used to perform test measurements. BS1 dichroic beamsplitter, BS2 polarizing beamsplitter, L1 tube lens f=160 mm, L2 collimating lens f= 100 mm, L3 lens f=10 mm, SMF single

mode fiber, FC fiber coupler with 2-axis tilt, O objective, S sample plate, M broadband mirror, PH pinhole on 8 position wheel, BP band pass filter, SPAD single photon avalanche diode.

In Fig. 2, the FCS curve generated using our SP605-based design is compared with that obtained using a commercial ALV 6010/200 correlator with a specified time resolution of 5 ns. In the ~100ns – 1s range, the curves are hardly distinguishable. At shorter correlation times the photon antibunching term[16] is slightly more pronounced in the curve computed using our correlator. We then fitted the correlation curves using Eq. (3), which accounts for translational diffusion, triplet and reaction kinetics, and photon antibunching:

$$G(t_c) = 1 + \frac{1}{N}\left(\frac{1}{1+t_c/t_d}\right)\left(\frac{1}{1+t_c/\left[(z_0/\omega_0)^2 t_d\right]}\right)^{1/2}\left(1-T+T\exp(-t_c/t_T)\right)\times\left(1-K+K\exp(-t_c/t_K)\right) \quad (3)$$
$$\times\left(1-\exp(-t_c/t_A)\right)$$

In Eq. (3), N is an average number of bright molecules in the detection volume, t_d is the diffusion time, z_0/ω_0 is aspect ratio of the confocal volume, T is the equilibrium fraction of the triplet state, t_T is the triplet relaxation time, and t_A is the antibunching time. The second kinetic term with an amplitude K and a characteristic time t_K is caused by saturation and complex triplet and/or radical kinetics at high irradiance[62, 63]. Detailed investigation of this effect and its origin is beyond the scope of this work.

The fitted parameters for the FCS curves generated using our design and the ALV 6010/200 agree typically within a few percent (see figure text for details). In particular, N and t_D show almost perfect agreement (N: 0.4%; t_D: 0.2 %). The deviations can be attributed to limited statistics of the measurements and to instability of the fit with respect to K and t_K because this process is slower than diffusion. Weighted residuals (Fig. 2A, upper panel) show no features specific for either curve.

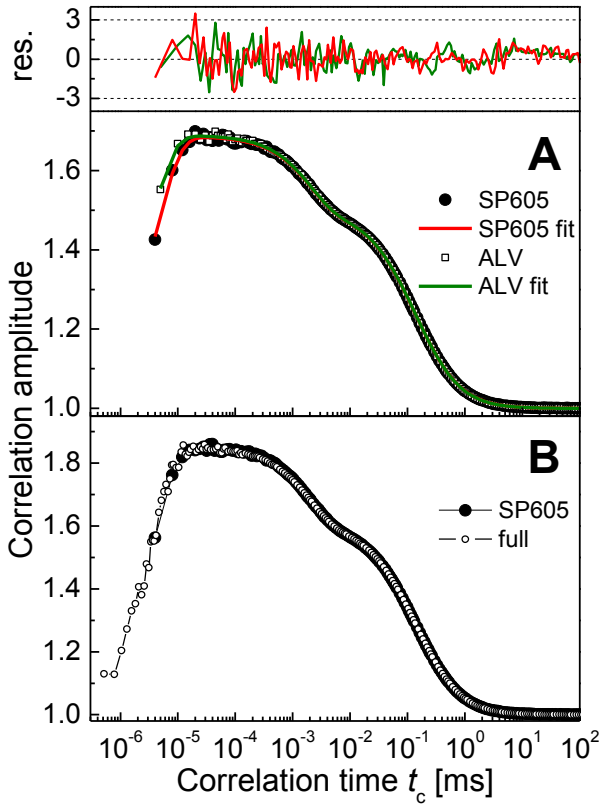


FIG. 2. A: Comparison of the FCS curves of Rhodamine 110 generated by ALV 6010/200 (□) and the SP605-based correlator presented in this work (●). The curves are fitted by Eq. (3) with the following parameters: ALV: N = 1.450; t_D = 0.144 ms; z_0/ω_0 = 4.1; T = 0.274; t_T = 2.03 µs; K = 0.217; t_K = 0.591 ms; t_A = 3.1 ns; SP605: N = 1.455; t_D = 0.144 ms; z_0/ω_0 = 4.0; T = 0.272; t_T = 2.03 µs; K = 0.186; t_K = 0.52 µs; t_A = 4.0 ns. Weighted residuals are shown above. **B**: Comparison with the FCS curve calculated by software correlation of photon traces recorded with TCSPC cards (○). In all cases two cross-correlation functions (detector1→detector2 and detector2→detector1) were calculated and the average was taken. In plot (B) the curves were measured with slightly different concentrations of Rh110 and were scaled to the same amplitude. The mean irradiance was ca. 50 kW/cm^2 (450 µW at the objective).

To ensure that the observed anticorrelation in the ns range is not due to an artifact, we recorded photon traces with picosecond time resolution using two TCSPC cards (SPC832, Becker & Hickl, Germany) on the same experimental setup. Then we performed full software correlation of the recorded signals[59]. Figure 2B shows that our FCS curve agrees very well with the computed full correlation curve over the whole time range. This comparison convincingly demonstrates that the electronic time resolution of 4 ns is readily achieved in practice.

We presented a free implementation of a fast hardware correlator with features and performance similar to current commercial devices. To our knowledge, significantly better (sub-ns) time resolution can be achieved only by using high-end electronics with TCSPC capabilities (for instance, Becker&Hickl SPC series or DPC-230; PicoQuant PicoHarp or HydraHarp modules). Considering the low cost of the SP605 board and of the presented compact FCS setup, and also the fact that it can be easily extended to more than two detection channels and several parallel correlator units, our design should find more uses in research, general analytic applications, and education. An increasing number of applications of FCS has been reported for life[1, 114, 115] and material sciences[2, 116]. The mobility of proteins and DNA- or RNA-fragments within the cytosol or other cell organelles belong to the most prominent measurement parameters of intracellular measurements[117]. Numerous studies are also devoted to the analysis of molecular (e.g. ligand receptor) interactions or translocation processes[2]. Therefore flexible and low cost instrumentation[25] with high performance will facilitate the use of FCS for even more applications.

We are grateful to Dr. Suren Felekyan and to Denis Dörr for valuable discussions and for help with measurements. C.S. thanks the German Science Foundation (SE 1195/13-1) and EU (FP7-Health-2007-A-201837) for funding during this work.

1D. Magde, E. L. Elson, and W. W. Webb, Phys. Rev. Lett. **29**, 705 (1972).
2R. Rigler, Ü. Mets, J. Widengren, and P. Kask, Eur. Biophys. J. **22**, 169 (1993).
3P. Schwille, F. J. Meyer-Almes, and R. Rigler, Biophys. J. **72**, 1878 (1997).
4M. Branden, T. Sanden, P. Brzezinski, and J. Widengren, Proc. Natl. Acad. Sci. U.S.A. **103**, 19766 (2006).
5D. Nettels, I. V. Gopich, A. Hoffmann, and B. Schuler, Proc. Natl. Acad. Sci. U.S.A. **104**, 2655 (2007).
6C. Eggeling, C. Ringemann, R. Medda, G. Schwarzmann, K. Sandhoff, S. Polyakova, V. N. Belov, B. Hein, C. von Middendorff, A. Schonle, and S. W. Hell, Nature (London) **457**, 1159 (2009).
7K. Bacia, S. A. Kim, and P. Schwille, Nat. Methods **3**, 83 (2006).
8D. Magatti and F. Ferri, Rev. Sci. Instrum. **74**, 1135 (2003).
9M. Wahl, I. Gregor, M. Patting, and J. Enderlein, Opt. Express **11**, 3583

(2003).
10S. Felekyan, R. Kühnemuth, V. Kudryavtsev, C. Sandhagen, W. Becker, and C. A. M. Seidel, Rev. Sci. Instrum. **76**, 083104 (2005).
11T. A. Laurence, S. Fore, and T. Huser, Opt. Lett. **31**, 829 (2006).
12G. Mocsár, B. Kreith, J. Buchholz, J. W. Krieger, J. Langowski, and G. Vamosi, Rev. Sci. Instrum. **83**, 046101 (2012).
13K. Schätzel, M. Drewel, and S. Stimac, J. Mod. Opt. **35**, 711 (1988).
14P. Alfke, "Efficient shift registers, LFSR counters, and long pseudorandom sequence generators," Xilinx Application Note No. 052 (1996); see http://www.xilinx.com/support/documentation/application_notes/xapp052.pdf.
15W. Al-Soufi, B. Reija, M. Novo, S. Felekyan, R. Kühnemuth, and C. A. M. Seidel, J. Am. Chem. Soc. **127**, 8775 (2005).
16Ü. Mets, *Fluorescence Correlation Spectroscopy*, Springer Series in Chemical Physics Vol. 65, edited by R. Rigler and E. L. Elson (Springer-Verlag, 2001),Chap. 16, p. 346.
17J. Widengren, Ü. Mets, and R. Rigler, J. Phys. Chem. **99**, 13368 (1995).
18J. Widengren, A. Chmyrov, C. Eggeling, P. A. Lofdahl, and C. A. M. Seidel, J. Phys. Chem. A **111**, 429 (2007).
19E. Haustein and P. Schwille, Annu. Rev. Biophys. Biomol. Struct. **36**, 151 (2007).
20Z. Petrasek and P. Schwille, J. R. Soc., Interface **6**, S15 (2009).
21A. Shahzad and G. Kohler, Appl. Spectrosc. Rev. **46**, 166 (2011).
22S. Salzinger, S. Huber, S. Jaksch, P. Busch, R. Jordan, and C.M. Papadakis, Colloid Polym. Sci. **290**, 385 (2012).
23W. Al-Soufi, B. Reija, S. Felekyan, C. A. M. Seidel, and M. Novo, Chem. Phys. Chem. **9**, 1819 (2008).
24S. R. Yu, M. Burkhardt, M. Nowak, J. Ries, Z. Petrasek, S. Scholpp, P. Schwille, and M. Brand, Nature (London) **461**, 533 (2009).
25See supplementary material at http://dx.doi.org/10.1063/1.4753994 for further implementation details and test measurements.

A 4 ns hardware photon correlator based on a general-purpose FPGA development board implemented in a compact setup for fluorescence correlation spectroscopy

Stanislav Kalinin[a)], Ralf Kühnemuth, Hayk Vardanyan, and Claus A.M. Seidel[a)]

Supplemental Material

Correlation of a pseudo-random test signal

At first we correlated a pseudorandom test signal generated within the same FPGA chip using the Linear Feedback Shift Register algorithm (ref. 14 of the main text). We simulated interconversion of two states "1" and "2", with characteristic signals S_1 = 488.3 kHz and S_2 = 976.6 kHz, respectively. The times spent in each state were exponentially distributed with t_1 = t_2 = 1.0486 ms and thus an exponential term with a relaxation time of $t_R = 1/(t_1^{-1} + t_2^{-1})$ = 524.3 µs is expected in the correlation function (Eq. (S1)).

$$G(t_c) = 1 + A\exp(-t_c / t_R) \qquad \text{(S1)}$$

Figure S1 shows the computed correlation curve and the best fit to it (Eq. (S1)). The fit yields t_R = 523.1 µs with only 0.23% deviation from the expected value. The correlation amplitude A is 0.1085, which is slightly below A = 0.1111 expected for $\langle S_2\rangle/\langle S_1\rangle$ = 2. This can be rationalized by considering that two or more consecutive (merged) pulses may occur in the simulated signal. These pulses are not resolved by the correlator, effectively decreasing the contrast to $\langle S_2\rangle/\langle S_1\rangle \cong 1.984$ (instead of the assumed $\langle S_2\rangle/\langle S_1\rangle$ = 2). For this signal ratio we expect $A \cong 0.1087$, which is in perfect agreement with the fit results.

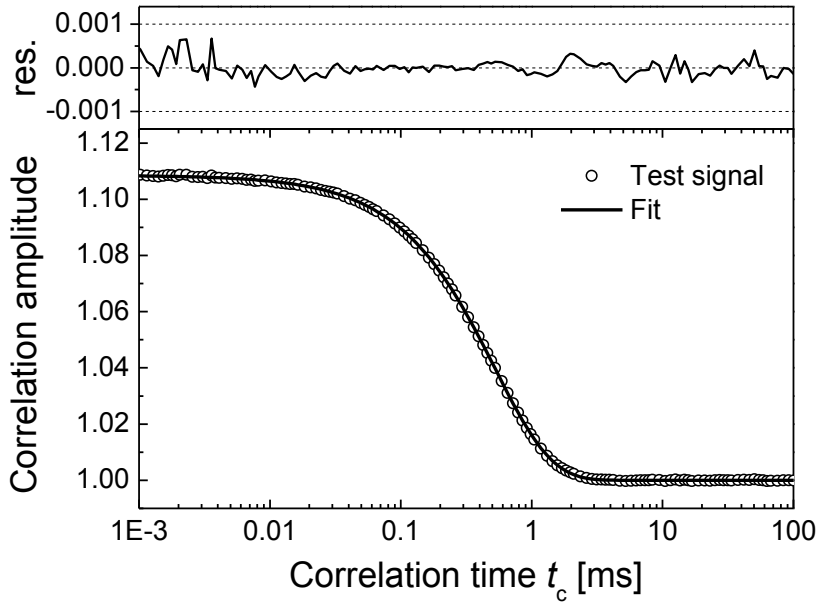


FIG. S1. Correlation of pseudorandom test signal with exponential distribution of switching times with $t_1 = t_2 = 1.0486$ ms^{-1}. The black solid line represents the least-squares fit to the data (Eq. (2)), resulting in $t_R = 523.1$ µs and $A = 0.1085$. Residuals are shown in the upper panel.

TABLE I. Part number list for the FCS setup

type	specification	part#	supplier
lens (1x)	f=160 mm	32 2270 322	Linos
lens (1x)	f=100 mm	32 2288 322	Linos
lens (2x)	f=10 mm	31 2011 322	Linos
pinholes	Pt/Ir, d=25µm - 200µm	A0404P - A0411P	Plano
micrometer screws (12x)	Dz = 6.5 mm	148-201	Mitutoyo
dichroic beamsplitter	500 nm longpass	BS 500	AHF
z-translator	Dz = 1.5 mm	SM1Z	Thorlabs
objective	1.2 NA water immersion	UPlanSApo 60x/1.2w	Olympus
50/50 beamsplitter	500-750 nm	TSWP 633	AHF

bandpass (2x)	HC525/39		AHF
detectors (2x)	SPAD, d=20µm	PD200A	MPD
laser	Multiline Ar+	35-LAP-431-220	MellesGriot
fiber&coupler	single mode fiber		SuK
fiber mount	Kinematic Cage Mount	KC1/M	Thorlabs
iris	d=18mm	G061009000	Linos
rods (8x)	d=6mm, l=3"(min)	ER3	Thorlabs
Home-built components:			
Beamsplitter cubes (3x)			
Tube lens mount (2x)			
Detector mount (2x)			
Pinhole module (1x)			
Sample plate (1x)			
Base plate (1x)			

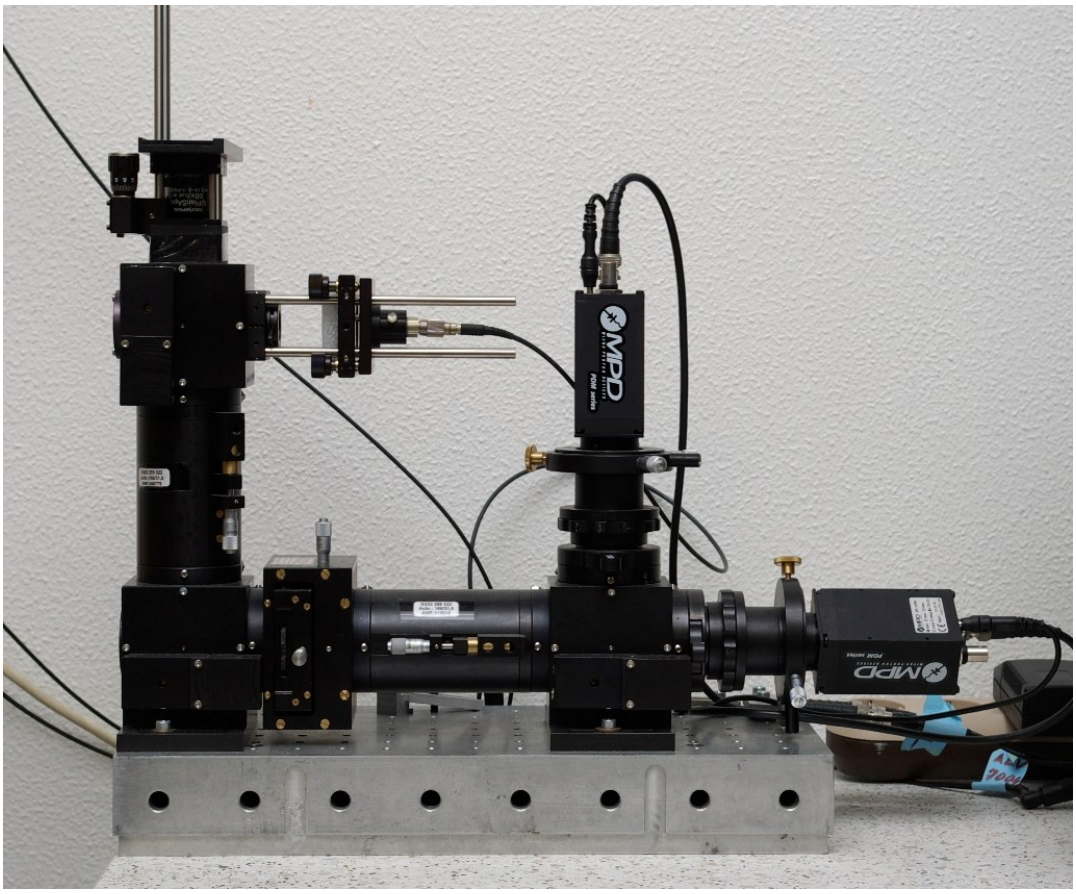


FIG. S2. Photo of the FCS setup here equipped with MPD detectors (PD200A, MPD Bolzano, Italy).

TABLE II. FPGA utilization summary (whole design, two independent correlator blocks, mapped).

FPGA resource	Units used (% of available)
Logic slices	854 (12%)
Memory (BRAM)	19 (8%)
DSP blocks	6 (10%)

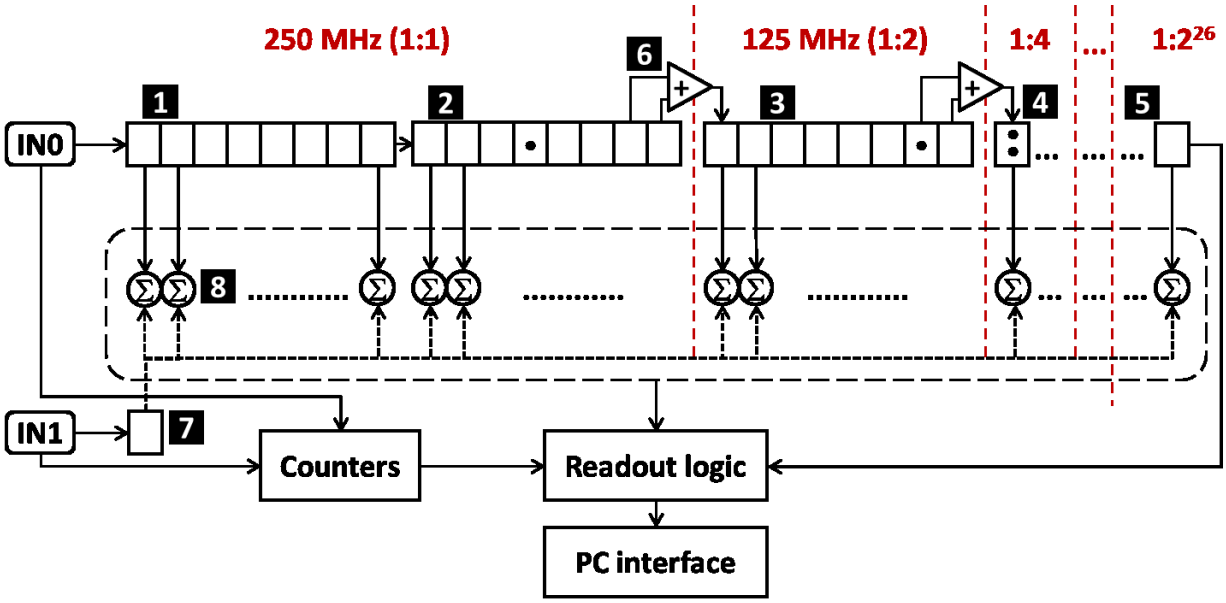


FIG. S3. Multi-tau scheme. Cascades are implemented as 8-element shift registers, where each element can store 1 bit for cascades 1, 2 and 3, 2 bits for cascade 4, and then at least *n*-2 bits for cascade *n*. Cascades 1-4 (**1**-**4**) and 28 (**5**) are shown explicitly. Every clock cycle, conditioned input signal of detection channel 0 is loaded into the first bin of cascade 1. Starting from cascade 3, the time bin length is doubled for every next cascade, and the respective initial bins are loaded with the sum of bins 7 and 8 of the previous cascade (e.g. adder **6** provides the data for bin 1 of cascade 3). Input data of detection channel 1 is first registered (**7**). Each cascade bin has an assigned accumulator operating essentially as a photon pair counter (circles with Σ, **8**). Accumulators are triggered by data in register **7**, in which case the cascade data are added to the accumulator current state and the result is stored. Clock domains separated by red dashed lines with respective operating frequencies shown in red. Readout logic takes care of transferring data to PC, which include current states of all photon pair counters (N_{pairs} in Eq. (2) of the main text), signals of both input channels (S_i and S_j), and the state of the last bin. The normalized correlation function is calculated by PC software (Eq. (2) in the main text).

Cascades 1-4 are implemented in hardware according to the presented scheme with data stored in registers and fully parallel accumulators, which allows for high speed operation. For higher cascades, a logically equivalent scheme is implemented using built-in RAM and DSP blocks (Fig. S4).

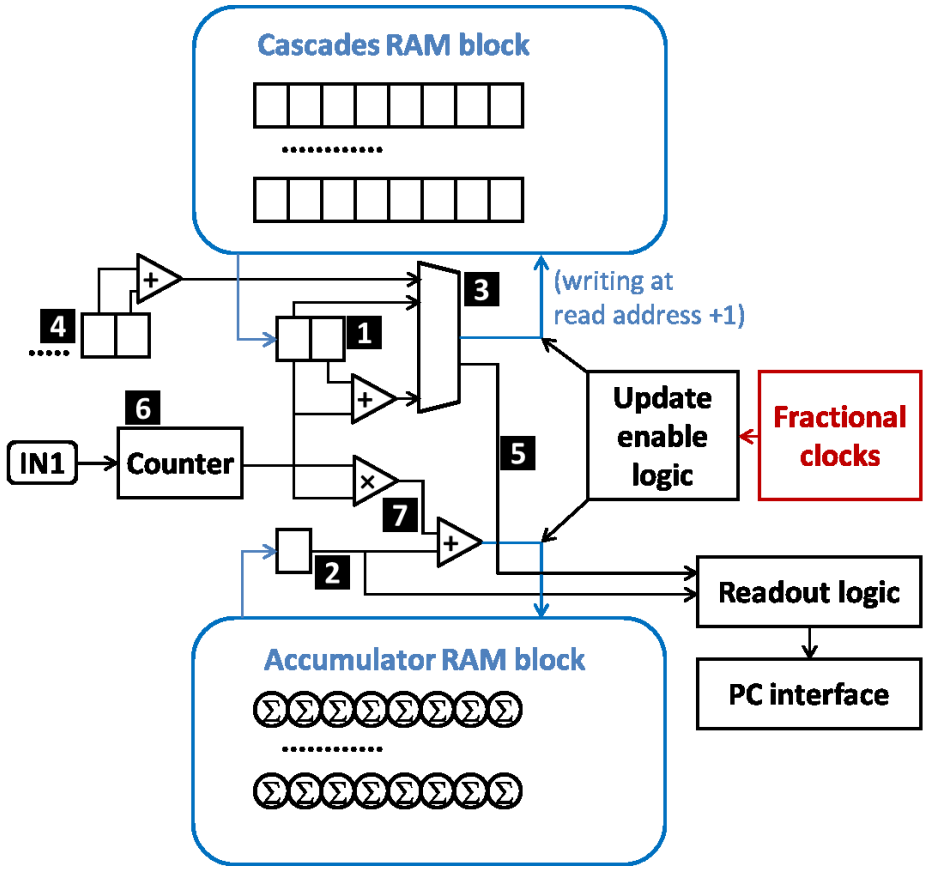


FIG. S4. Implementation details for higher cascades. In this case the data are updated relatively rarely and sequential execution of arithmetic operations is more rational than the fully parallel scheme shown in Fig. S3. Cascade and accumulator data are now stored in two separate memory blocks (RAMB8BWER) shown in blue. The data are read into a two-element shift register **1** and a register **2**, respectively, constantly iterating over all available memory addresses. Multiplexer **3** is responsible for providing data to the next updated cascade bin. The data are selected from three options: data of previous cascades **4**; previously read element; or the sum of two previously read elements. The last element of the last cascade is also sent (**5**) to the readout block to be transferred to the PC. A product of the first element of the shift register **1** and a value of the photon counter of input 1 (**6**) is calculated, representing newly accumulated photon pairs. This value is then added to the old state of the corresponding photon counter data **2**. These operations are conveniently performed with DSP block(s) **7**. Writing to both memory blocks is enabled only if one or several cascades must be updated at a given time, which is dictated by fractional clocks (i.e. cascade n is updated once per $4\times2^{n-2}$ ns, $n>3$). Memory address increment logic, special cases for read and write addresses, and the reset logic of counter **6** are not explicitly shown.

7.2 Input files for FPS

Table 35 Input "labeling positions" file for FPS for J(abc) molecule.

lablling position	helix *a*/*b*/*c*	Acceptor/Donor	Determination method of AV	L_{linker},Å	w_{linker},Å	$R_{dye(1)}$,Å	$R_{dye(2)}$,Å	$R_{dye(3)}$,Å	Attachment atom number
a7AA	rna4wjA_HV01	A	AV3	22	4.5	11	3	1.5	210
c29DA	rna4wjA_HV01	D	AV3	20	4.5	5	4.5	1.5	918
c8DB	rna4wjB_HV01	D	AV3	20	4.5	5	4.5	1.5	241
c8AB	rna4wjB_HV01	A	AV3	22	4.5	11	3	1.5	241
b27DB	rna4wjB_HV01	D	AV3	20	4.5	5	4.5	1.5	692
b27AB	rna4wjB_HV01	A	AV3	22	4.5	11	3	1.5	692
b33AB	rna4wjB_HV01	A	XYZ	22	4.5	11	3	1.5	884
b5DC	rna4wjC_HV01	D	AV3	20	4.5	5	4.5	1.5	149
b8DC	rna4wjC_HV01	D	AV3	20	4.5	5	4.5	1.5	247
b11DC	rna4wjC_HV01	D	AV3	20	4.5	5	4.5	1.5	340
b14DC	rna4wjC_HV01	D	AV3	20	4.5	5	4.5	1.5	433
b14AC	rna4wjC_HV01	A	AV3	22	4.5	11	3	1.5	433
a28AC	rna4wjC_HV01	A	AV3	22	4.5	11	3	1.5	1042
c24AA	rna4wjA_HV01	A	AV3	22	4.5	11	3	1.5	750
a10AA	rna4wjA_HV01	A	AV3	22	4.5	11	3	1.5	296
c7DB	rna4wjB_HV01	D	AV3	20	4.5	5	4.5	1.5	203
b29DB	rna4wjB_HV01	D	AV3	20	4.5	5	4.5	1.5	748
c12AB	rna4wjB_HV01	A	AV3	22	4.5	11	3	1.5	364
a26AC	rna4wjC_HV01	A	AV3	22	4.5	11	3	1.5	979
labelling position	separate fragments	Attachment atom number							
hAnCC1P	hAnCC1	446							
hAnCC1O3	hAnCC1	454							
hAnCC1N461	hAnCC1	461							
hAnCC1H467	hAnCC1	467							
hAnCC1O4	hAnCC1	452							
hAnCC2P	hAnCC2	477							
hAnCC2O3	hAnCC2	485							

hAnCC2N492	hAnCC2	492							
hAnCC2H498	hAnCC2	498							
hAnCC2O4	hAnCC2	483							
hAnGG1O3	hAnGG1	551							
hAnGG1P	hAnGG1	543							
hAnGG1H567	hAnGG1	567							
hAnGG1O562	hAnGG1	562							
hAnGG1O4	hAnGG1	549							
hAnGG2P	hAnGG2	509							
hAnGG2O3	hAnGG2	517							
hAnGG2H533	hAnGG2	533							
hAnGG2O528	hAnGG2	528							
hAnGG2O4	hAnGG2	515							
hBnUG1P	hBnUG1	449							
hBnUG1O3	hBnUG1	457							
hBnUG1H473	hBnUG1	473							
hBnUG1O468	hBnUG1	468							
hBnUG1O4	hBnUG1	455							
hBnUG2O3	hBnUG2	490							
hBnUG2P	hBnUG2	483							
hBnUG2H504	hBnUG2	504							
hBnUG2O500	hBnUG2	500							
hBnUG2O4	hBnUG2	489							
hBnAC1O3	hBnAC1	555							
hBnAC1P	hBnAC1	547							
hBnAC1N562	hBnAC1	562							
hBnAC1H568	hBnAC1	568							
hBnAC1O4	hBnAC1	553							
hBnAC2P	hBnAC2	514							
hBnAC2O3	hBnAC2	522							
hBnAC2N526	hBnAC2	526							
hBnAC2H538	hBnAC2	538							
hBnAC2O4	hBnAC2	520							

hCnAC1O3	hCnAC1	710							
hCnAC1P	hCnAC1	702							
hCnAC1N717	hCnAC1	717							
hCnAC1H723	hCnAC1	723							
hCnAC1O4	hCnAC1	708							
hCnAC2P	hCnAC2	669							
hCnAC2O3	hCnAC2	677							
hCnAC2N681	hCnAC2	681							
hCnAC2H693	hCnAC2	693							
hCnAC2O4	hCnAC2	675							
hCnUG1P	hCnUG1	604							
hCnUG1O3	hCnUG1	612							
hCnUG1H628	hCnUG1	628							
hCnUG1O623	hCnUG1	623							
hCnUG1O4	hCnUG1	610							
hCnUG2O3	hCnUG2	646							
hCnUG2P	hCnUG2	638							
hCnUG2H659	hCnUG2	659							
hCnUG2O655	hCnUG2	655							
hCnUG2O4	hCnUG2	644							
BO3	rna4wjB_HV01	426							
BP	rna4wjB_HV01	578							
AO3	rna4wjA_HV01	421							
AP	rna4wjA_HV01	577							
CO3	rna4wjC_HV01	578							
CP	rna4wjC_HV01	733							
AO3O4	rna4wjA_HV01	419							
APO4	rna4wjA_HV01	583							
BO3O4	rna4wjB_HV01	424							
BPO4	rna4wjB_HV01	584							
CO3O4	rna4wjC_HV01	576							
CPO4	rna4wjC_HV01	739							

Table 36 Input "distances" file for FPS for J(abc) molecule.

l.p.1	l.p.2	$\langle R_{DA} \rangle_E$,Å	error(+), Å	error(-), Å
FRET distances				
b11DC	a7AA	55.7	2.8	2.8
b11DC	a10AA	53.6	2.7	2.7
b11DC	c8AB	42.8	2.3	2.3
b11DC	c24AA	54.1	2.7	2.7
b11DC	c12AB	46.1	2.3	2.3
b14DC	a7AA	43.5	2.2	2.2
b14DC	a10AA	48.9	2.4	2.4
b14DC	c24AA	49	2.5	2.5
b14DC	c12AB	45.3	2.3	2.3
b27DB	a7AA	48.6	2.4	2.4
b27DB	a28AC	58.7	2.9	2.9
b27DB	a26AC	55.4	3	3
b27DB	a10AA	46.4	2.3	2.3
b27DB	c24AA	44.5	2.3	2.3
b29DB	a7AA	53.9	2.7	2.7
b29DB	a28AC	63.7	3.2	3.2
b29DB	a26AC	58.7	2.9	2.9
b29DB	a10AA	46.7	2.3	2.3
b29DB	c24AA	47.2	2.4	2.4
b5DC	a7AA	57.6	2.9	2.9
b5DC	a10AA	59.7	3	3
b5DC	c8AB	58.3	2.9	2.9
b5DC	c24AA	60.1	3	3
b5DC	c12AB	58.4	2.9	2.9
b8DC	a7AA	57.9	2.9	2.9

b8DC	a10AA	54.2	2.7	2.7
b8DC	c8AB	52.1	2.7	2.7
b8DC	c24AA	55.3	2.8	2.8
b8DC	c12AB	53	2.7	2.7
c29DA	a28AC	68.7	3.6	3.6
c29DA	a26AC	67.3	3.4	3.4
c29DA	b33AB	62.9	3.2	3.2
c29DA	b27AB	52.3	2.7	2.7
c29DA	b14AC	57.3	2.9	2.9
c7DB	a7AA	52.6	2.7	2.7
c7DB	a28AC	60.9	3.2	3.2
c7DB	a26AC	55.5	2.9	2.9
c7DB	a10AA	49.2	2.5	2.5
c7DB	b14AC	47.6	2.4	2.4
c8DB	a7AA	46.5	2.3	2.3
c8DB	a28AC	54.6	2.7	2.7
c8DB	a26AC	50.6	2.5	2.5
c8DB	a10AA	46.5	2.7	2.7
c8DB	b14AC	43	3.1	3.1
Covalent bonds				
AP	hAnGG1O3	2	1	20
hAnGG1P	hAnGG2O3	2	1	20
hAnGG2P	hBnUG2O3	2	1	20
hBnUG2P	hBnUG1O3	2	1	20
hBnUG1P	BO3	2	1	20
BP	hBnAC1O3	2	1	20
hBnAC1P	hBnAC2O3	2	1	20

hBnAC2P	hCnUG2O3	2	1	20
hCnUG2P	hCnUG1O3	2	1	20
hCnUG1P	CO3	2	1	20
CP	hCnAC1O3	2	1	20
hCnAC1P	hCnAC2O3	2	1	20
hCnAC2P	hAnCC2O3	2	1	20
hAnCC2P	hAnCC1O3	2	1	20
AO3	hAnCC1P	2	1	20
Hydrogen bonds				
hAnGG2H533	hAnCC2N492	1.8	1	20
hAnCC2H498	hAnGG2O528	1.8	1	20
hAnGG1H567	hAnCC1N461	1.8	1	20
hAnCC1H467	hAnGG1O562	1.8	1	20
hBnUG1H473	hBnAC1N562	1.8	1	20
hBnUG1O468	hBnAC1H568	1.8	1	20
hBnUG2O500	hBnAC2H538	1.8	1	20
hBnUG2H504	hBnAC2N526	1.8	1	20
hCnAC2N681	hCnUG2H659	1.8	1	20
hCnAC2H693	hCnUG2O655	1.8	1	20
hCnAC1N717	hCnUG1H628	1.8	1	20
hCnAC1H723	hCnUG1O623	1.8	1	20
Artificial bonds				
AO3O4	hAnCC1O4	5.9	1	20
hAnCC1O4	hAnCC2O4	5.9	1	20
hAnCC2O4	hCnAC2O4	5.9	1	20
hCnAC2O4	hCnAC1O4	5.9	1	20

hCnAC1O4	CPO4	5.9	1	20
CO3O4	hCnUG1O4	5.9	1	20
hCnUG1O4	hCnUG2O4	5.9	1	20
hCnUG2O4	hBnAC2O4	5.9	1	20
hBnAC2O4	hBnAC1O4	5.9	1	20
hBnAC1O4	BPO4	5.9	1	20
BO3O4	hBnUG1O4	5.9	1	20
hBnUG1O4	hBnUG2O4	5.9	1	20
hBnUG2O4	hAnGG2O4	5.9	1	20
hAnGG2O4	hAnGG1O4	5.9	1	20
hAnGG1O4	APO4	5.9	1	20

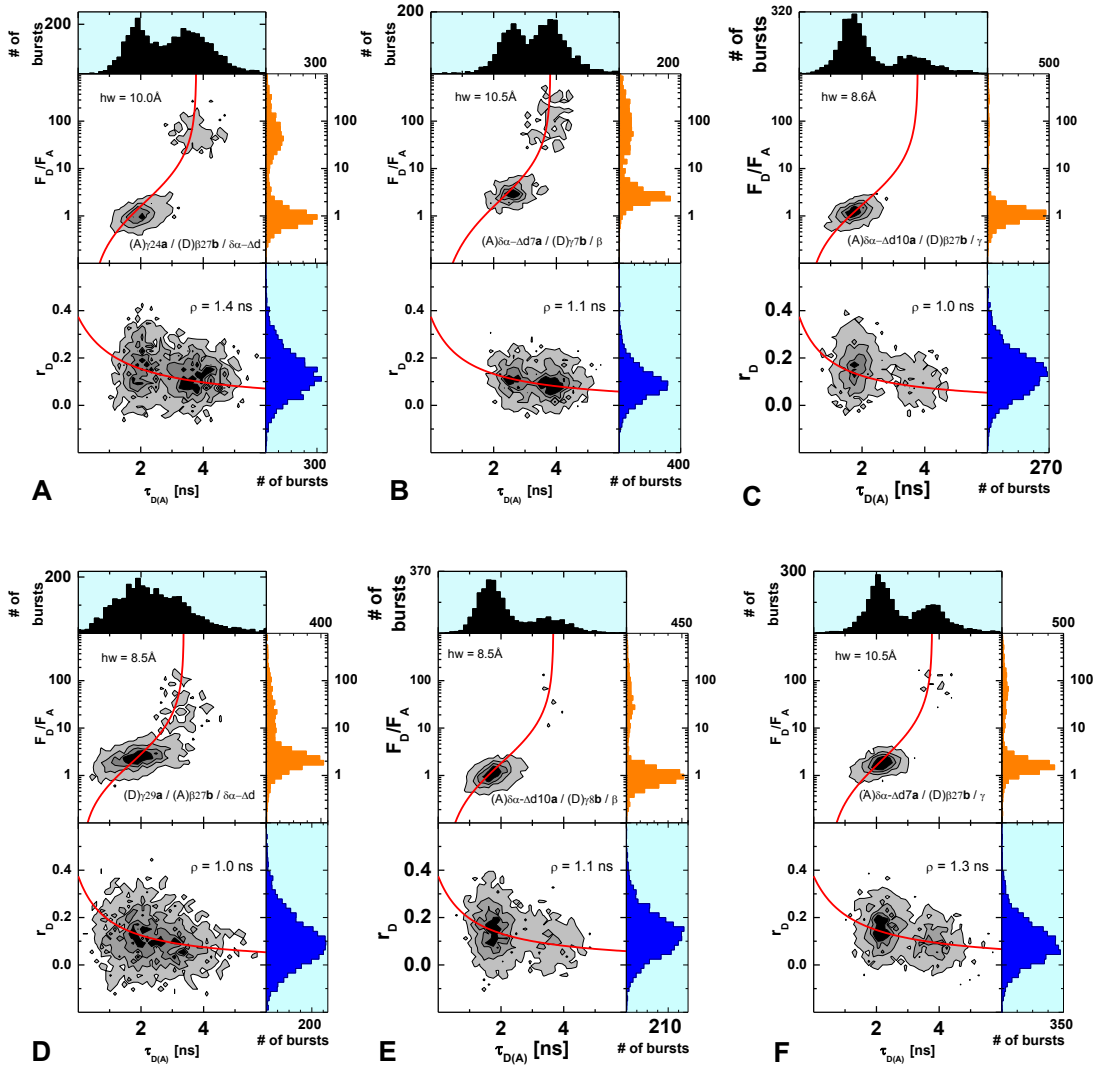
Table 37 Parameters used in FPS for docking and refinement procedurs.

	docking	**refinement**
viscosity	**1**	**0.7**
dt adjustement	**1**	**0.5**
max iterations(k)	**200**	**500**
max force	**400**	**10000**
clash tolerance(Å)	**1**	**0.5**

7.3 Complementary measurement data and PDA fits for thesis

7.3.1 2D histograms

7.3.1.1 Measurements at 20 mM $MgCl_2$ for J(abc)



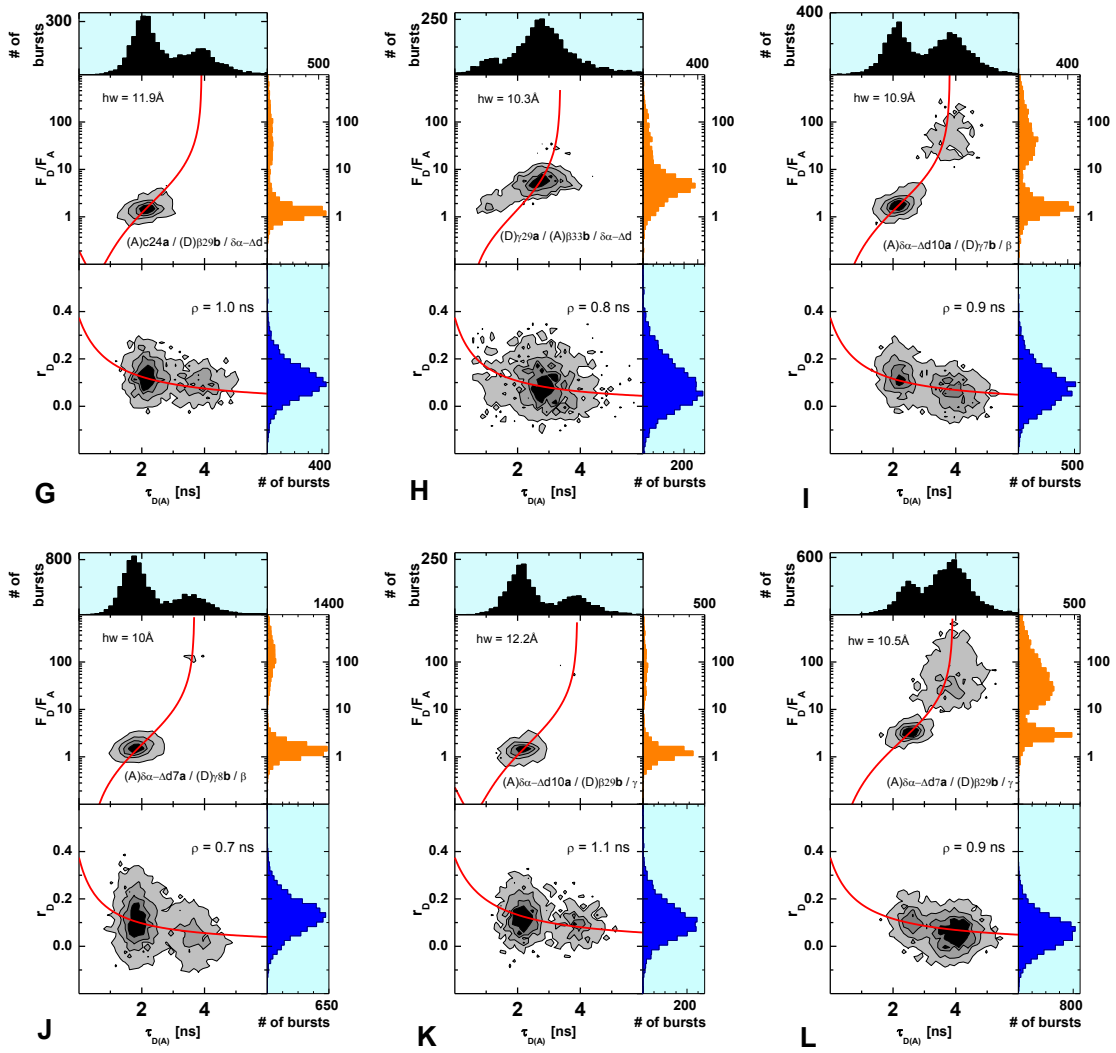
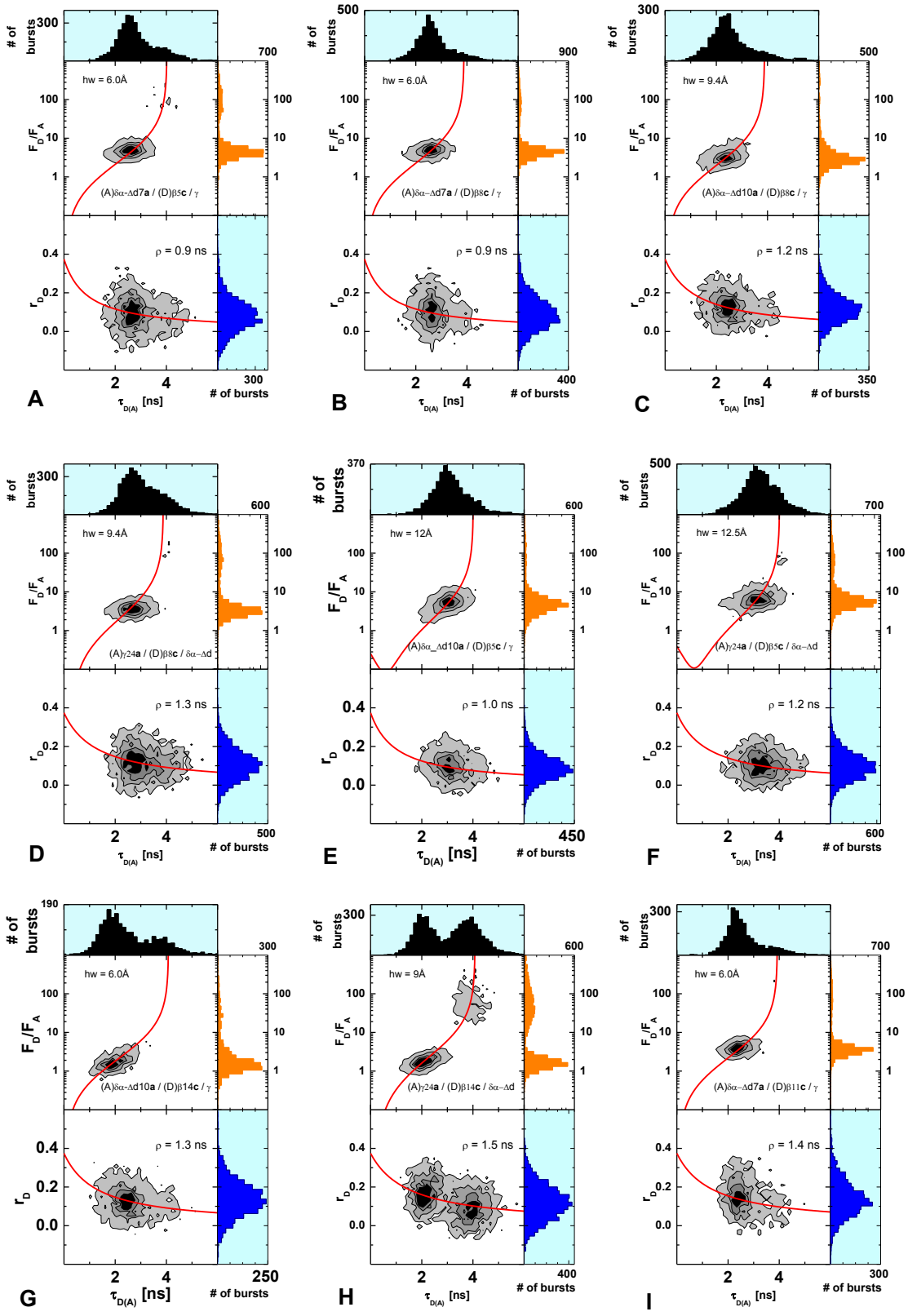


Figure 7-1 2D burst frequency histograms of F_D/F_A versus the donor fluorescence lifetime $\tau_{D(A)}$ (upper panel) and the donor fluorescence anisotropy r_D versus $\tau_{D(A)}$ (lower panel) for samples with fluorescence labels on the helices *a* and *b*. The number of molecules (fluorescence bursts) in each bin is gray scale, shaded from white (lowest) to black (highest). 1D histograms are shown as projections. In the F_D/F_A vs $\tau_{D(A)}$ plot, the theoretical relationship between F_D/F_A and $\tau_{D(A)}$ (static FRET line; red) is overlaid. The solid red line in the r_D-$\tau_{D(A)}$ diagram is given by the Perrin equation. Values of half with (hw) and rotation correlation time (ρ) are shown in the plots.


A
hw = 6.0Å
(A)δα–Δd7a / (D)β5c / γ
ρ = 0.9 ns
B
hw = 6.0Å
(A)δα–Δd7a / (D)β8c / γ
ρ = 0.9 ns
C
hw = 9.4Å
(A)δα–Δd10a / (D)β8c / γ
ρ = 1.2 ns
D
hw = 9.4Å
(A)γ24a / (D)β8c / δα–Δd
ρ = 1.3 ns
E
hw = 12Å
(A)δα_Δd10a / (D)β5c / γ
ρ = 1.0 ns
F
hw = 12.5Å
(A)γ24a / (D)β5c / δα–Δd
ρ = 1.2 ns
G
hw = 6.0Å
(A)δα–Δd10a / (D)β14c / γ
ρ = 1.3 ns
H
hw = 9Å
(A)γ24a / (D)β14c / δα–Δd
ρ = 1.5 ns
I
hw = 6.0Å
(A)δα–Δd7a / (D)β11c / γ
ρ = 1.4 ns
# of bursts
F_D/F_A
r_D
τ_D(A) [ns]

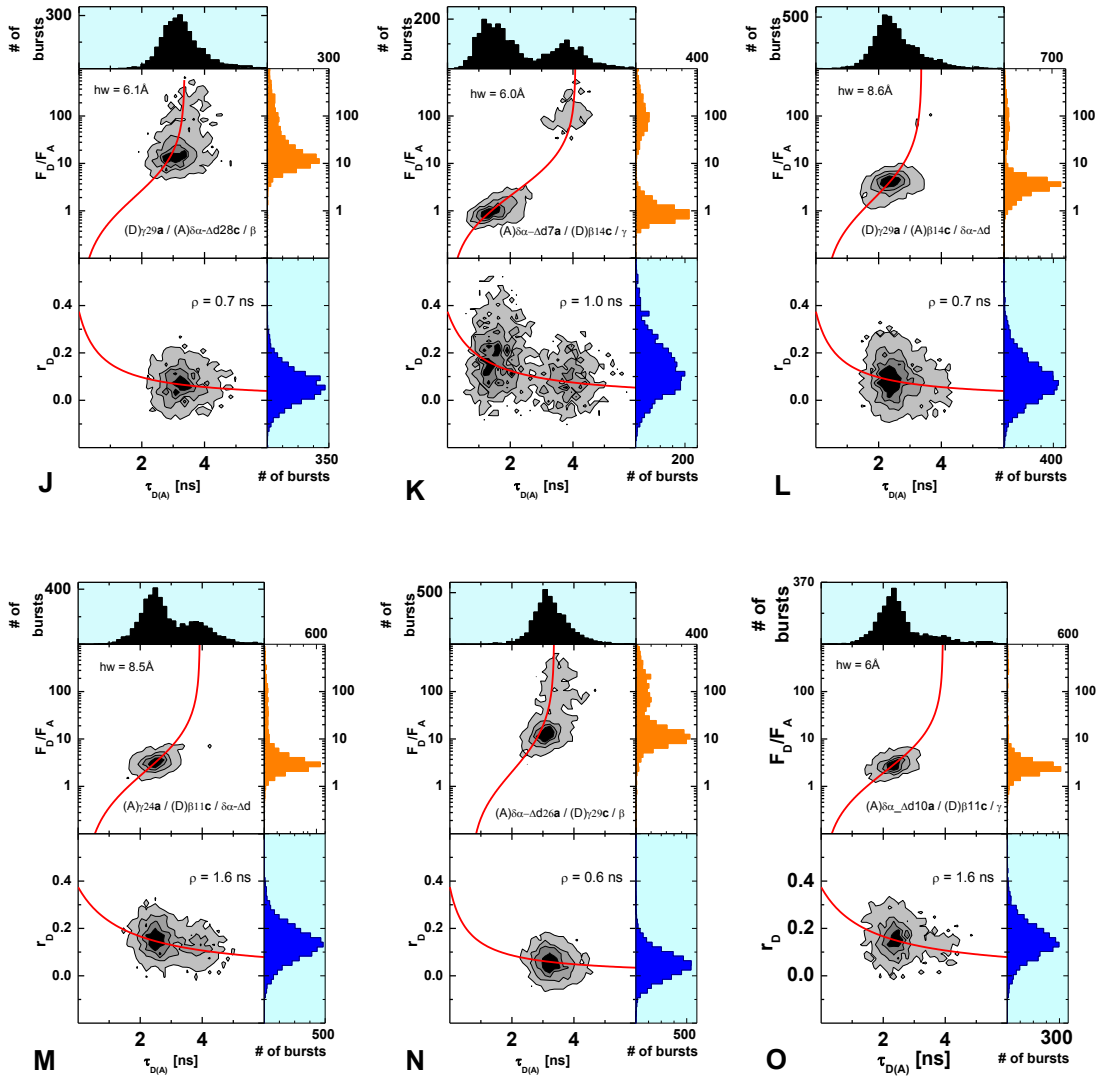
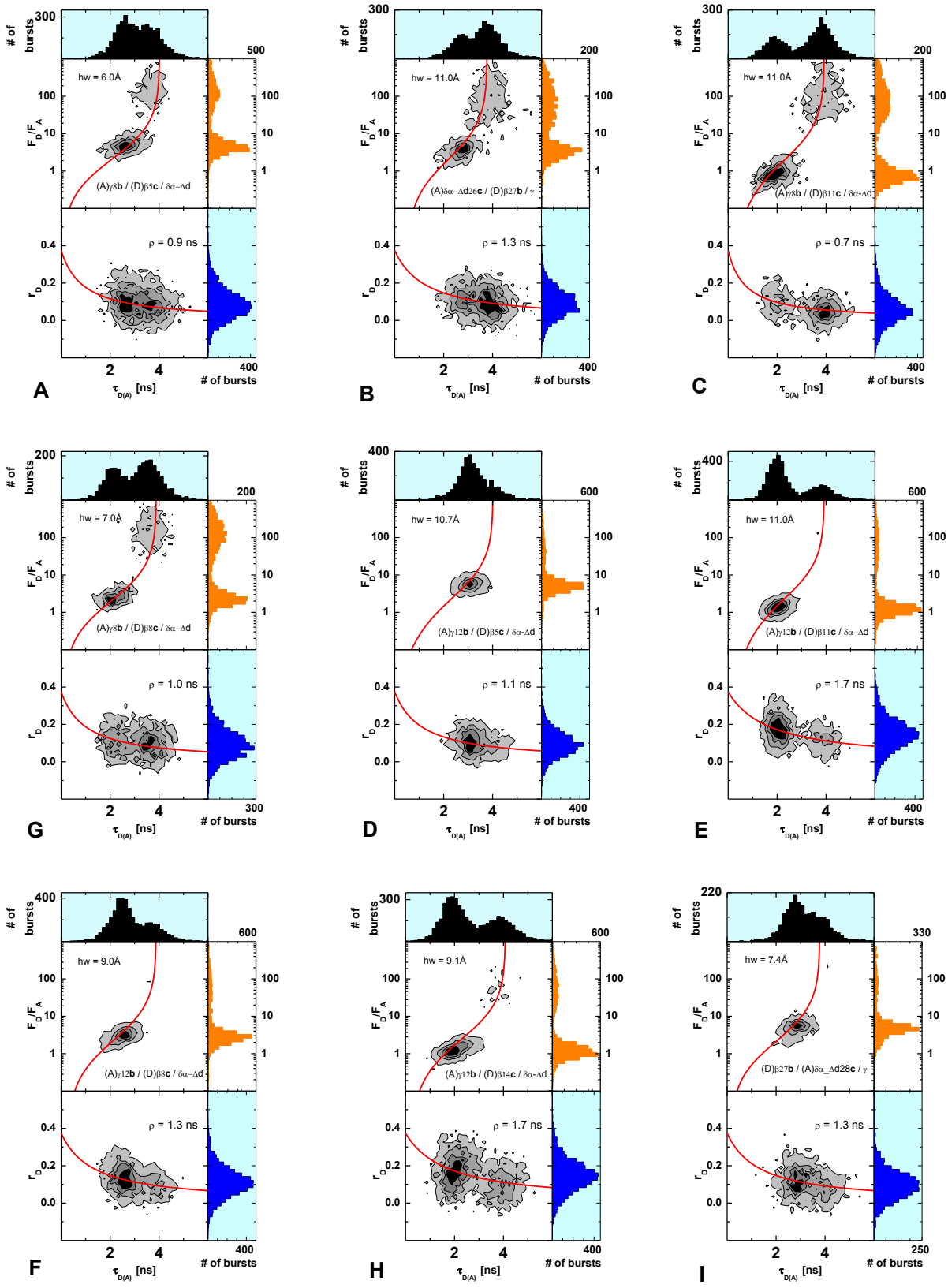


Figure 7-2 2D burst frequency histograms of F_D/F_A versus the donor fluorescence lifetime $\tau_{D(A)}$ (upper panel) and the donor fluorescence anisotropy r_D versus $\tau_{D(A)}$ (lower panel) for samples with fluorescence labels on the helices *a* and *c*. The number of molecules (fluorescence bursts) in each bin is gray scale, shaded from white (lowest) to black (highest). 1D histograms are shown as projections. In the F_D/F_A vs $\tau_{D(A)}$ plot, the theoretical relationship between F_D/F_A and $\tau_{D(A)}$ (static FRET line; red) is overlaid. The solid red line in the r_D-$\tau_{D(A)}$ diagram is given by the Perrin equation. Values of half with (hw) and rotation correlation time (ρ) are shown in the plots.


A
hw = 6.0Å
(A)γ8b / (D)β5c / δα–Δd
ρ = 0.9 ns
B
hw = 11.0Å
(A)δα–Δd26c / (D)β27b / γ
ρ = 1.3 ns
C
hw = 11.0Å
(A)γ8b / (D)β11c / δα–Δd
ρ = 0.7 ns
G
hw = 7.0Å
(A)γ8b / (D)β8c / δα–Δd
ρ = 1.0 ns
D
hw = 10.7Å
(A)γ12b / (D)β5c / δα–Δd
ρ = 1.1 ns
E
hw = 11.0Å
(A)γ12b / (D)β11c / δα–Δd
ρ = 1.7 ns
F
hw = 9.0Å
(A)γ12b / (D)β8c / δα–Δd
ρ = 1.3 ns
H
hw = 9.1Å
(A)γ12b / (D)β14c / δα–Δd
ρ = 1.7 ns
I
hw = 7.4Å
(D)β27b / (A)δα_Δd28c / γ
ρ = 1.3 ns
# of bursts
# of bursts
τD(A) [ns]
rD
FD/FA

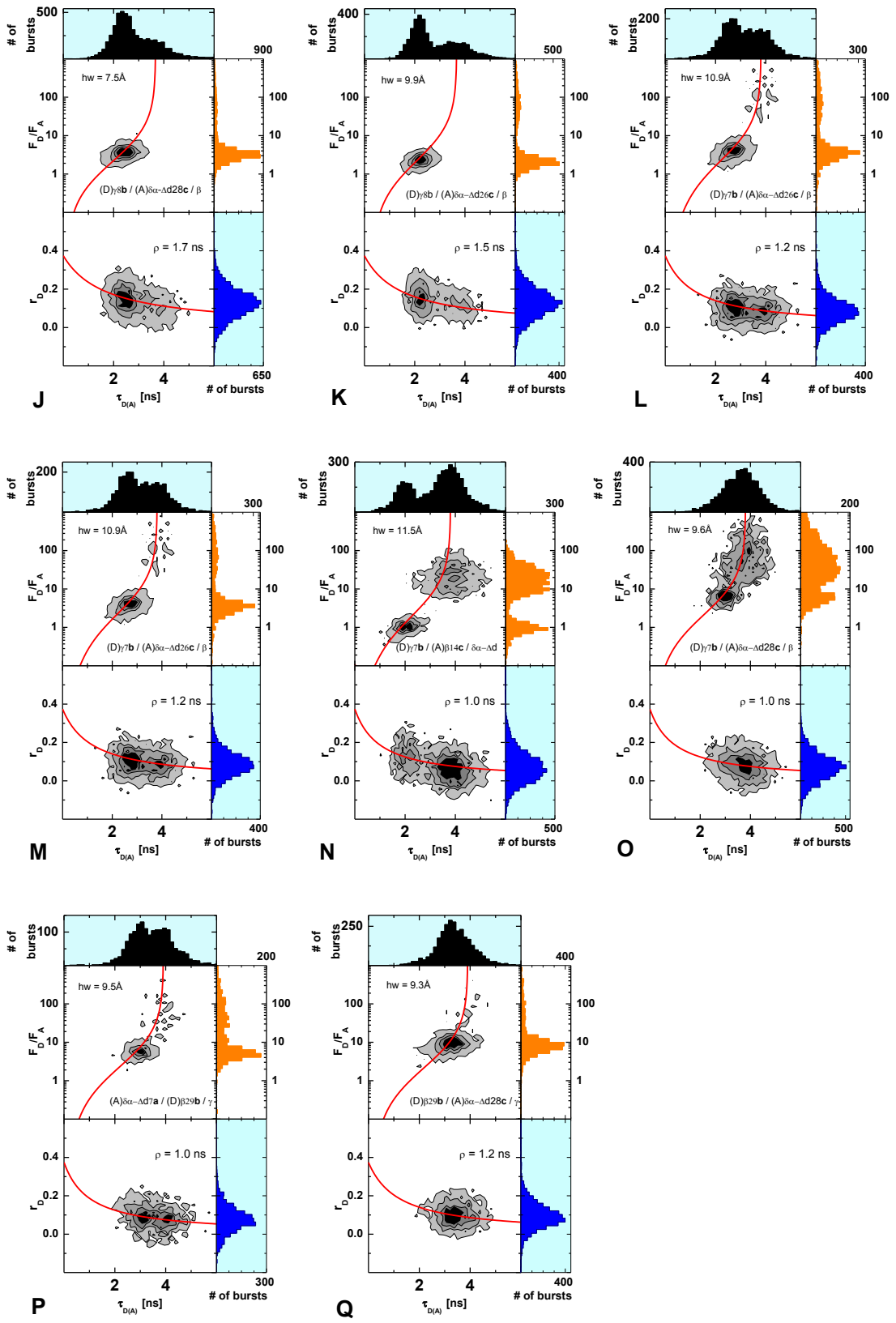
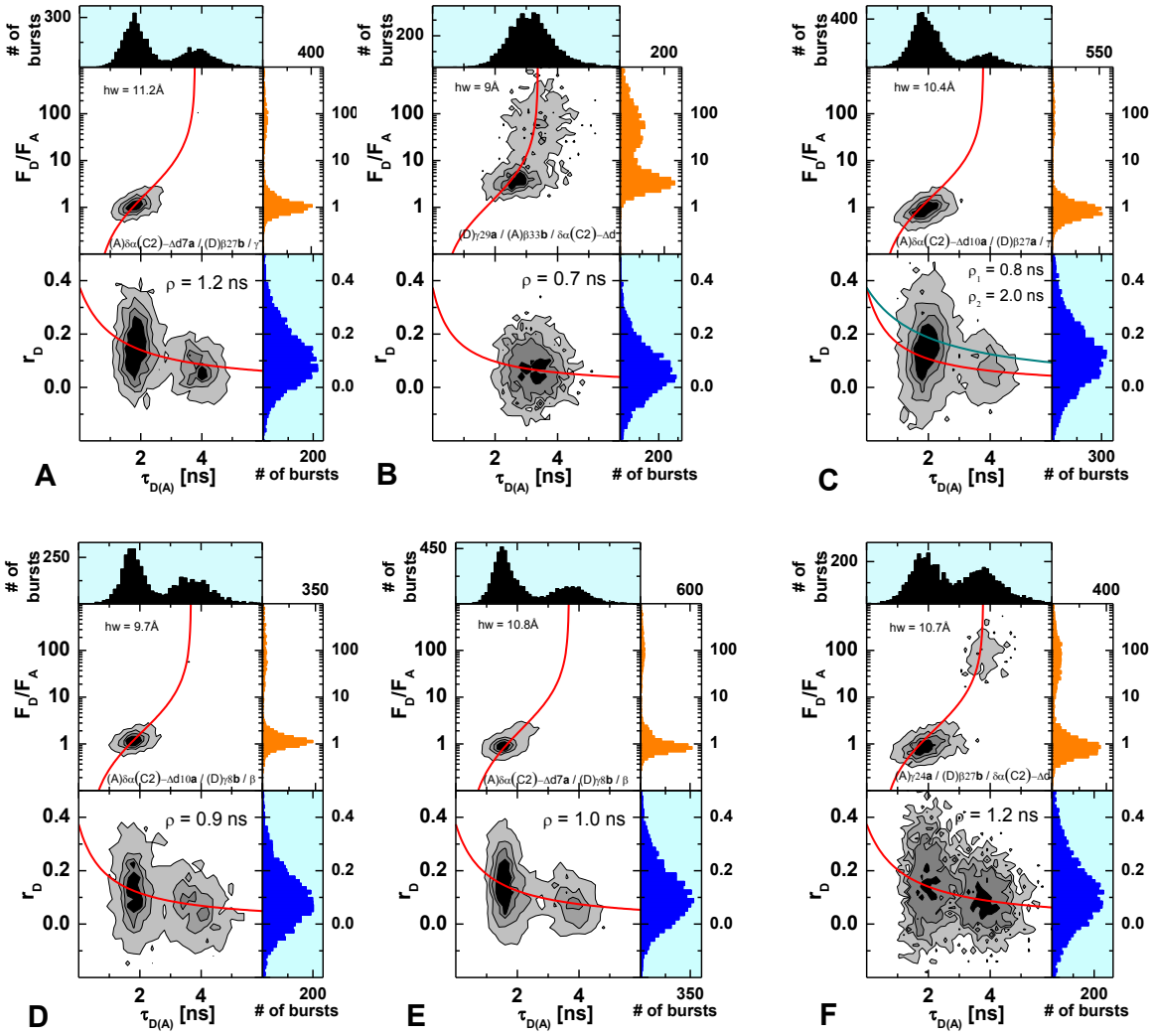


Figure 7-3 2D burst frequency histograms of F_D/F_A versus the donor fluorescence lifetime $\tau_{D(A)}$ (upper panel) and the donor fluorescence anisotropy r_D versus $\tau_{D(A)}$ (lower panel) for samples with fluorescence labels on the helices *b* and *c*. The number of molecules (fluorescence bursts) in each bin is gray scale, shaded from white (lowest) to black

(highest). 1D histograms are shown as projections. In the F_D/F_A vs $\tau_{D(A)}$ plot, the theoretical relationship between F_D/F_A and $\tau_{D(A)}$ (static FRET line; red) is overlaid. The solid red line in the r_D-$\tau_{D(A)}$ diagram is given by the Perrin equation. Values of half with (hw) and rotation correlation time (ρ) are shown in the plots.

7.3.1.2 Measurements at 20 mM $MgCl_2$ for J(abc(C2))



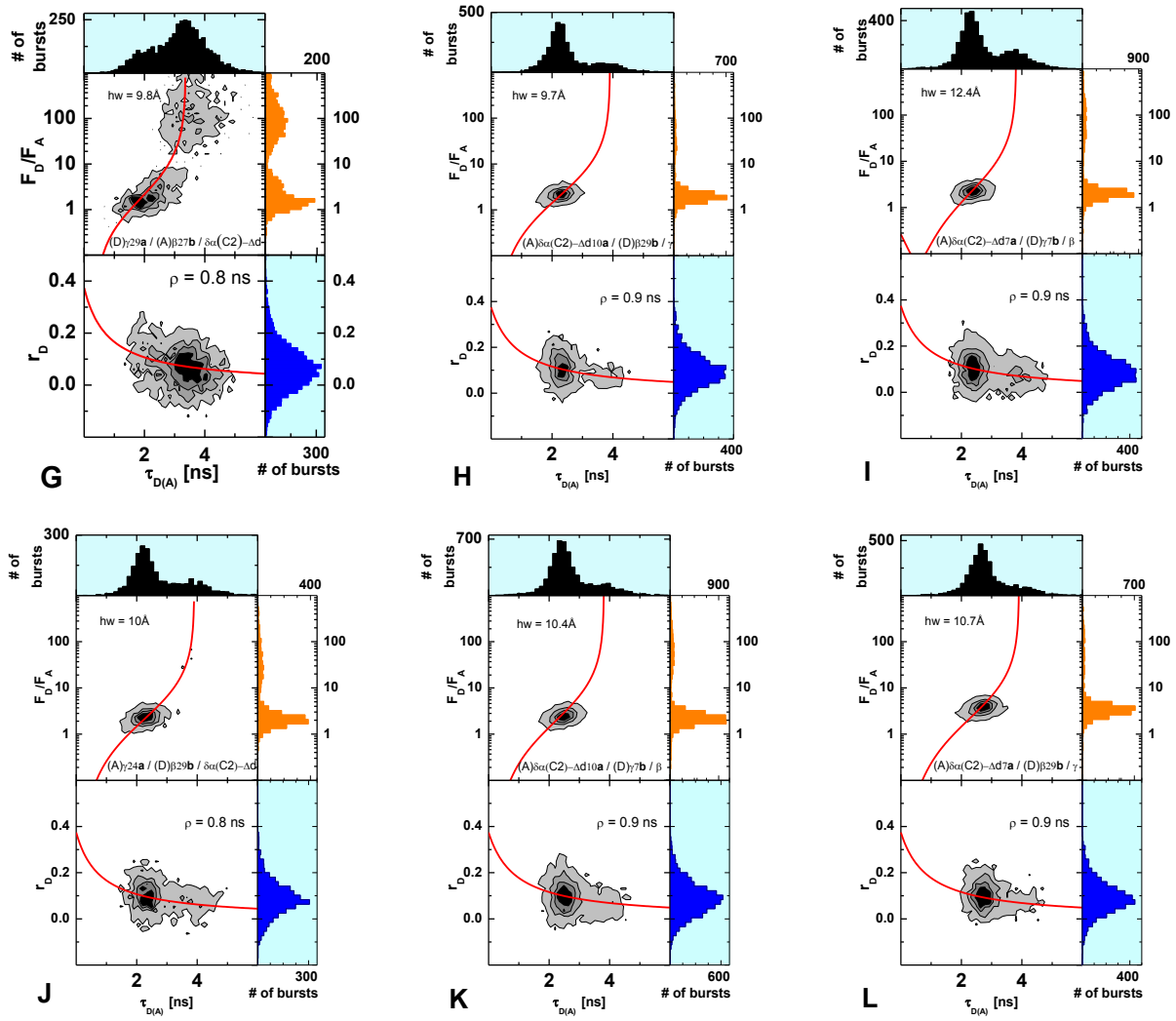
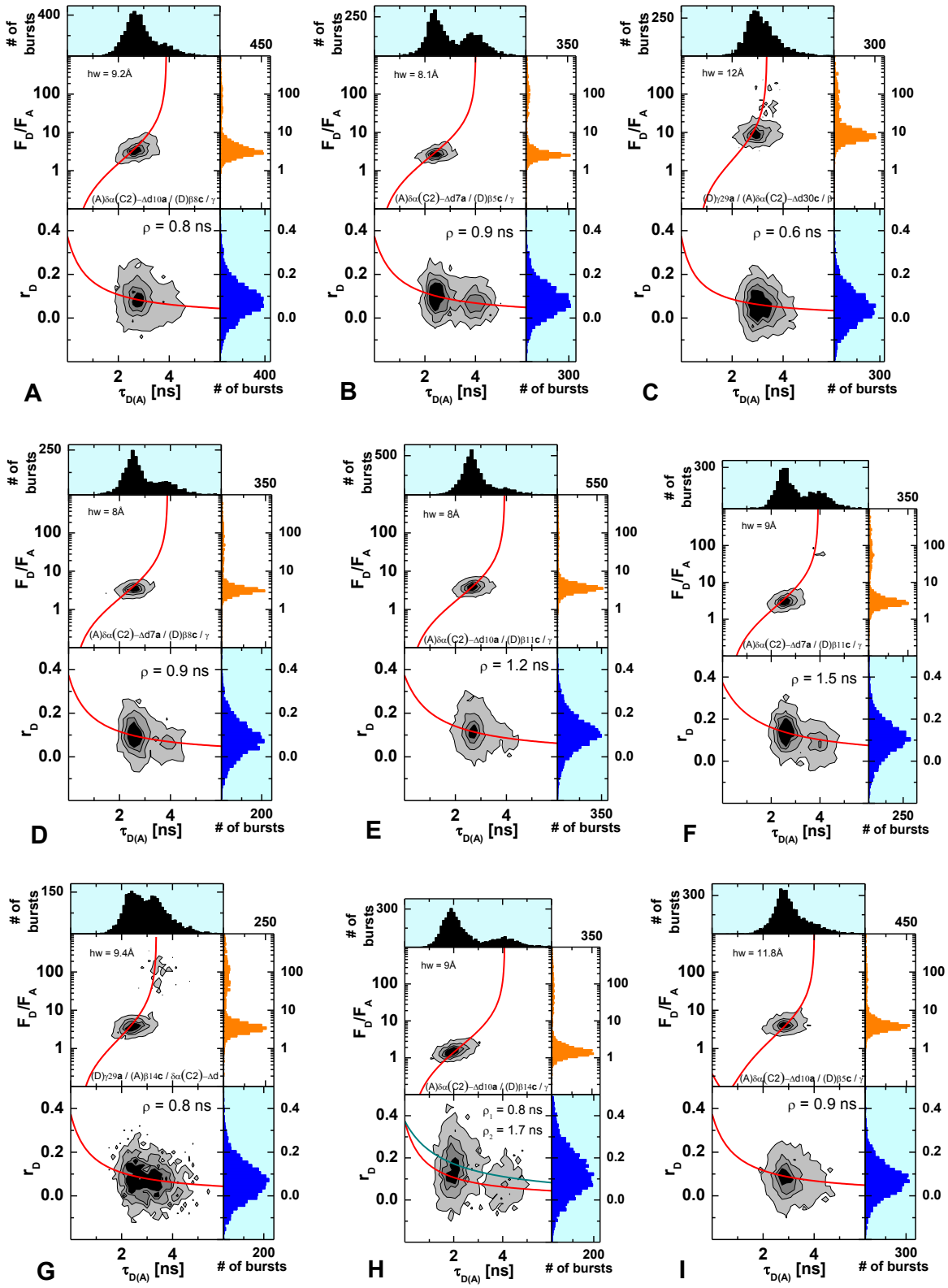


Figure 7-4 2D burst frequency histograms of F_D/F_A versus the donor fluorescence lifetime $\tau_{D(A)}$ (upper panel) and the donor fluorescence anisotropy r_D versus $\tau_{D(A)}$ (lower panel) for samples with fluorescence labels on the helices *a* and *b*. The number of molecules (fluorescence bursts) in each bin is gray scale, shaded from white (lowest) to black (highest). 1D histograms are shown as projections. In the F_D/F_A vs $\tau_{D(A)}$ plot, the theoretical relationship between F_D/F_A and $\tau_{D(A)}$ (static FRET line; red) is overlaid. The solid red line in the r_D-$\tau_{D(A)}$ diagram is given by the Perrin equation. Values of half with (hw) and rotation correlation time (ρ) are shown in the plots.


A
hw = 9.2Å
ρ = 0.8 ns
B
hw = 8.1Å
ρ = 0.9 ns
C
hw = 12Å
ρ = 0.6 ns
D
hw = 8Å
ρ = 0.9 ns
E
hw = 8Å
ρ = 1.2 ns
F
hw = 9Å
ρ = 1.5 ns
G
hw = 9.4Å
ρ = 0.8 ns
H
hw = 9Å
ρ1 = 0.8 ns
ρ2 = 1.7 ns
I
hw = 11.8Å
ρ = 0.9 ns
# of bursts
F_D/F_A
r_D
τ_D(A) [ns]

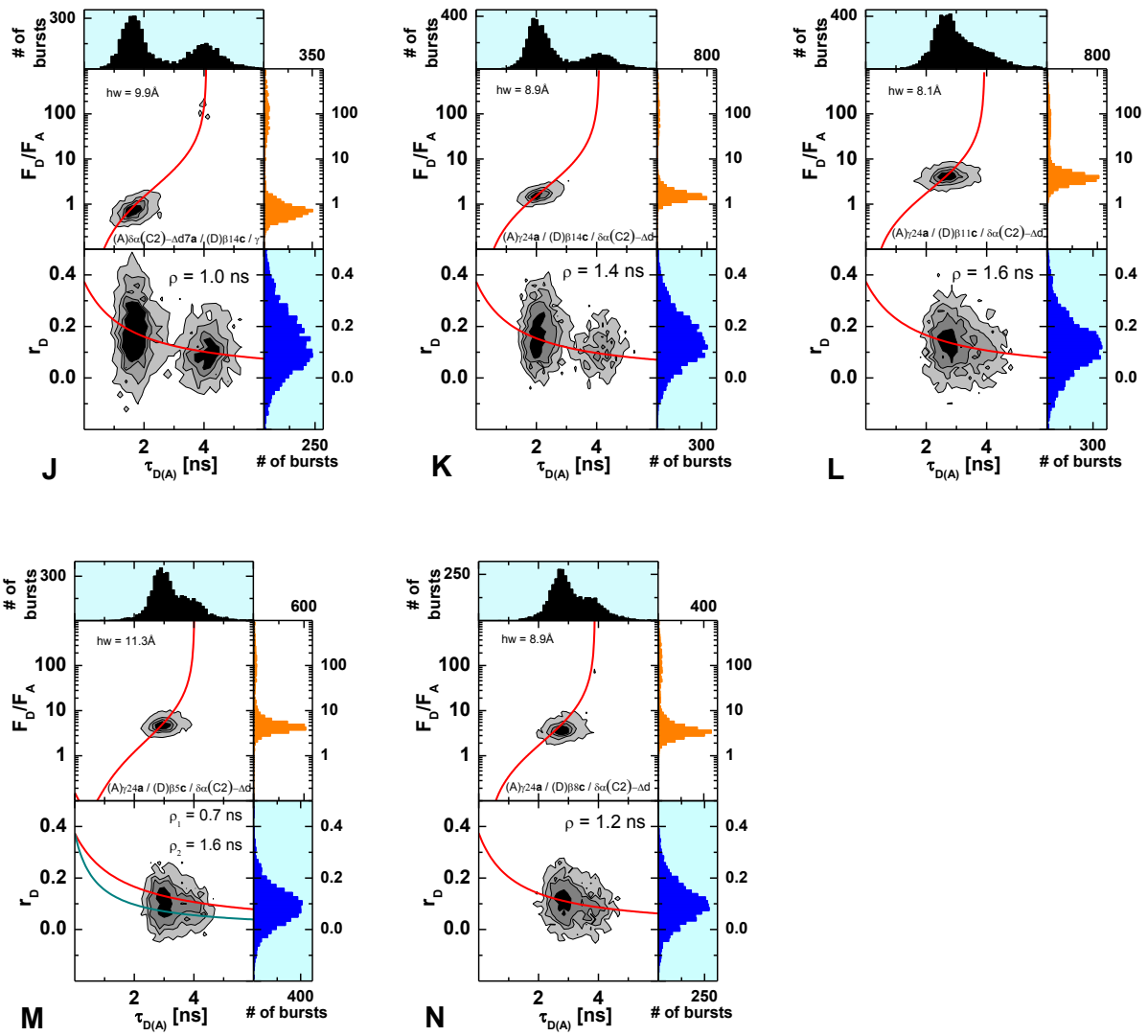
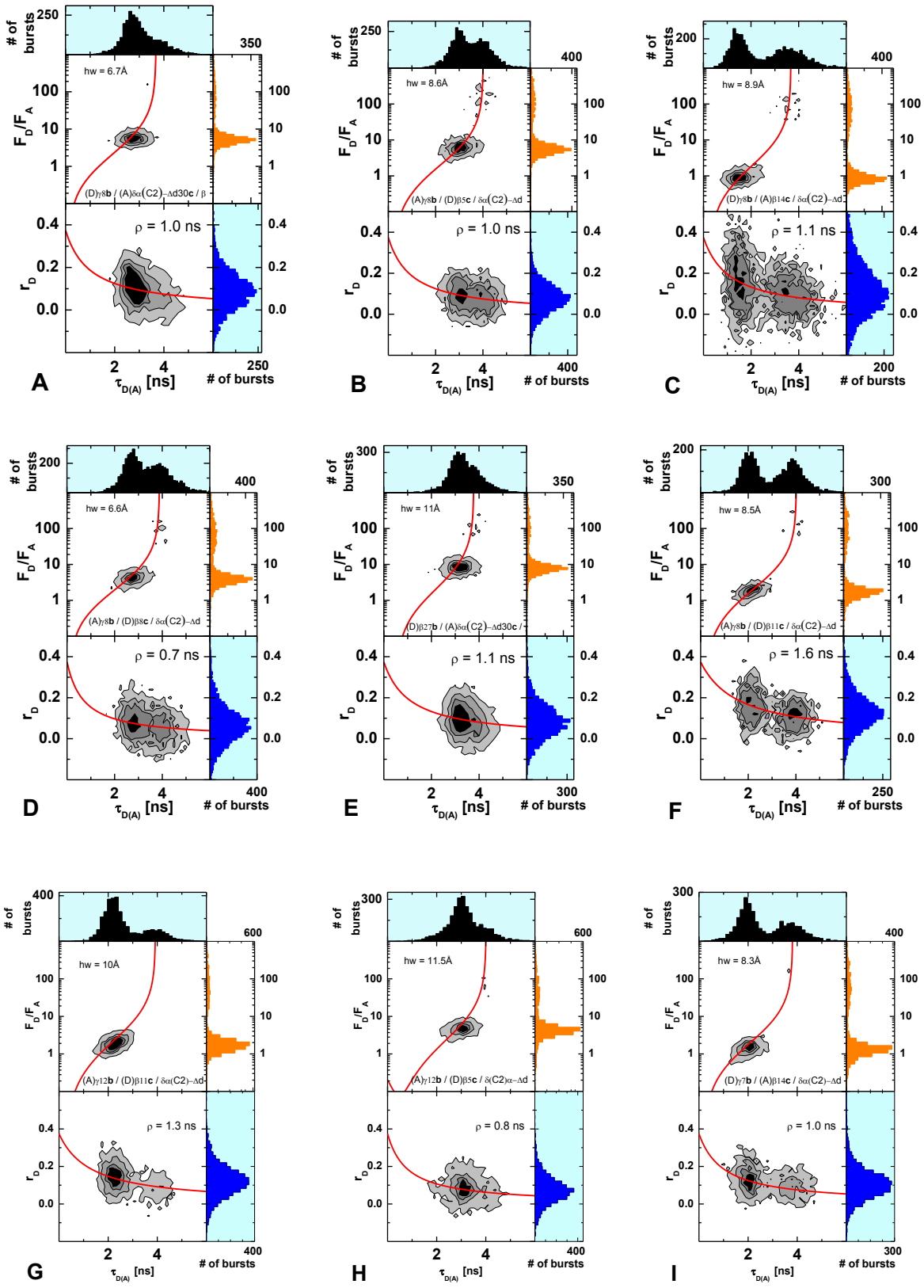


Figure 7-5 2D burst frequency histograms of F_D/F_A versus the donor fluorescence lifetime $\tau_{D(A)}$ (upper panel) and the donor fluorescence anisotropy r_D versus $\tau_{D(A)}$ (lower panel) for samples with fluorescence labels on the helices *a* and *c*. The number of molecules (fluorescence bursts) in each bin is gray scale, shaded from white (lowest) to black (highest). 1D histograms are shown as projections. In the F_D/F_A vs $\tau_{D(A)}$ plot, the theoretical relationship between F_D/F_A and $\tau_{D(A)}$ (static FRET line; red) is overlaid. The solid red line in the r_D-$\tau_{D(A)}$ diagram is given by the Perrin equation. Values of half with (hw) and rotation correlation time (ρ) are shown in the plots.


A
hw = 6.7Å
ρ = 1.0 ns
B
hw = 8.6Å
ρ = 1.0 ns
C
hw = 8.9Å
ρ = 1.1 ns
D
hw = 6.6Å
ρ = 0.7 ns
E
hw = 11Å
ρ = 1.1 ns
F
hw = 8.5Å
ρ = 1.6 ns
G
hw = 10Å
ρ = 1.3 ns
H
hw = 11.5Å
ρ = 0.8 ns
I
hw = 8.3Å
ρ = 1.0 ns
# of bursts
F_D/F_A
r_D
τ_D(A) [ns]

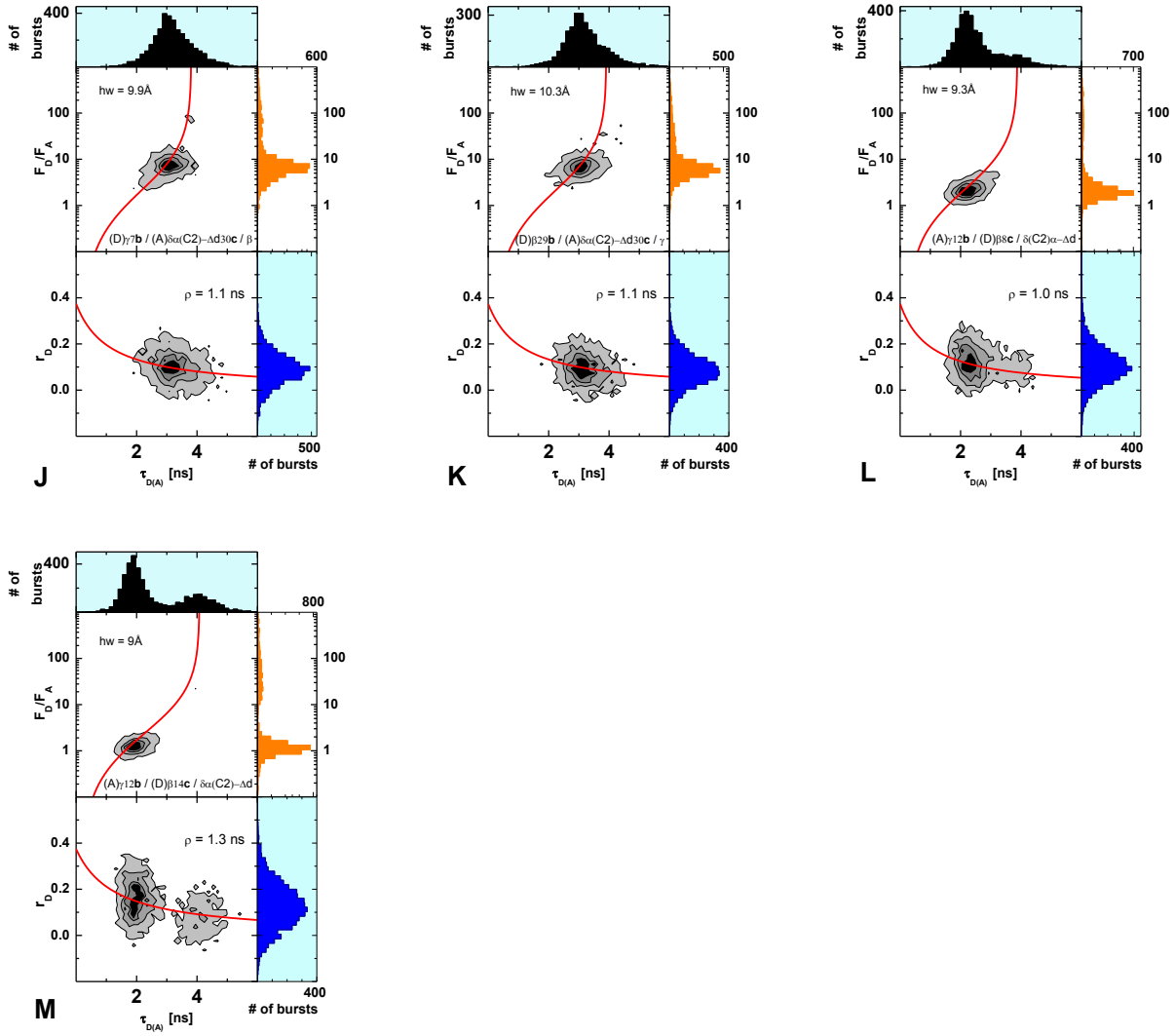
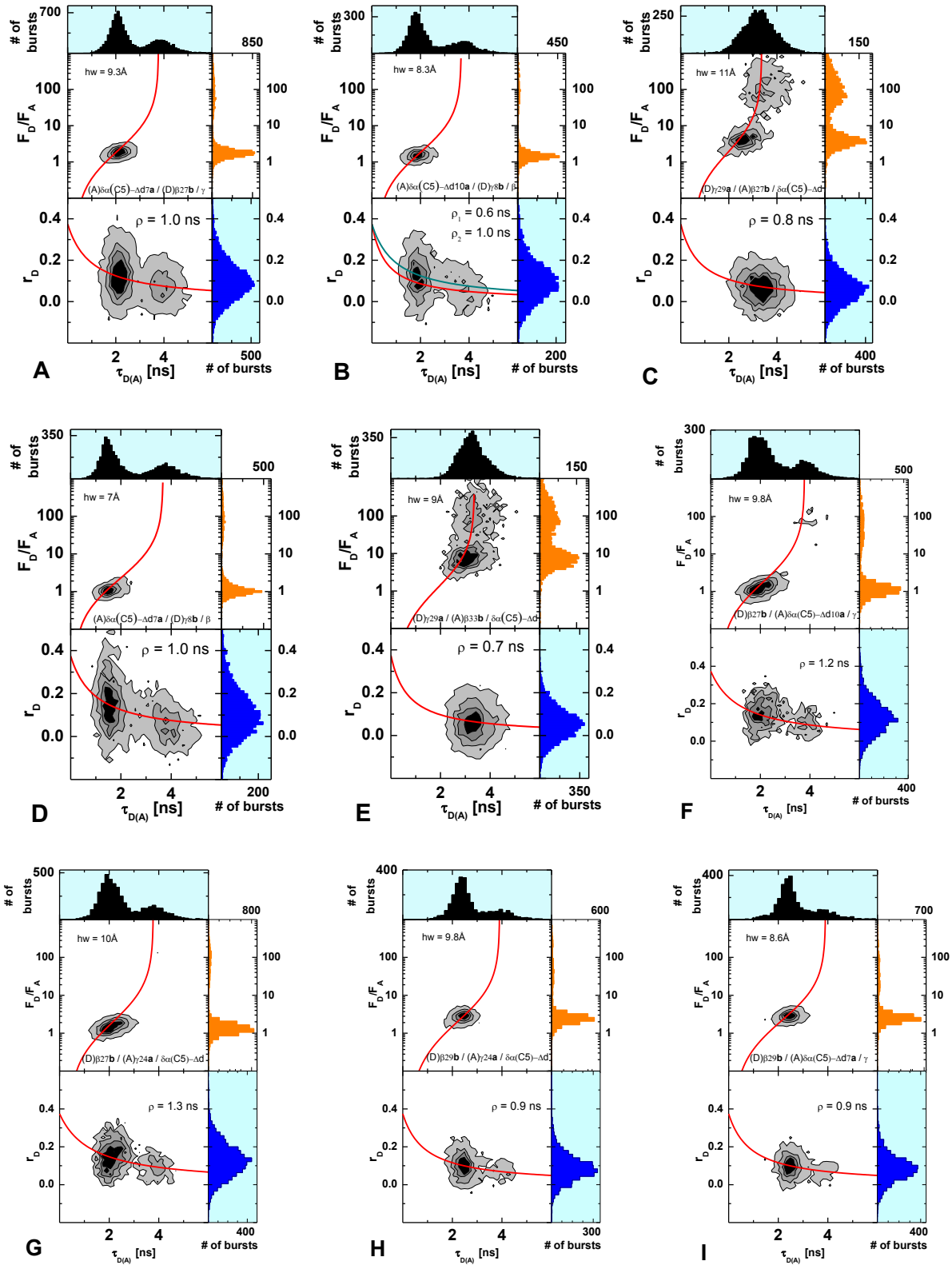


Figure 7-6 2D burst frequency histograms of F_D/F_A versus the donor fluorescence lifetime $\tau_{D(A)}$ (upper panel) and the donor fluorescence anisotropy r_D versus $\tau_{D(A)}$ (lower panel) for samples with fluorescence labels on the helices *b* and *c*. The number of molecules (fluorescence bursts) in each bin is gray scale, shaded from white (lowest) to black (highest). 1D histograms are shown as projections. In the F_D/F_A vs $\tau_{D(A)}$ plot, the theoretical relationship between F_D/F_A and $\tau_{D(A)}$ (static FRET line; red) is overlaid. The solid red line in the r_D-$\tau_{D(A)}$ diagram is given by the Perrin equation. Values of half with (hw) and rotation correlation time (ρ) are shown in the plots.

7.3.1.3 Measurements at 20 mM $MgCl_2$ for J(abc(C5))



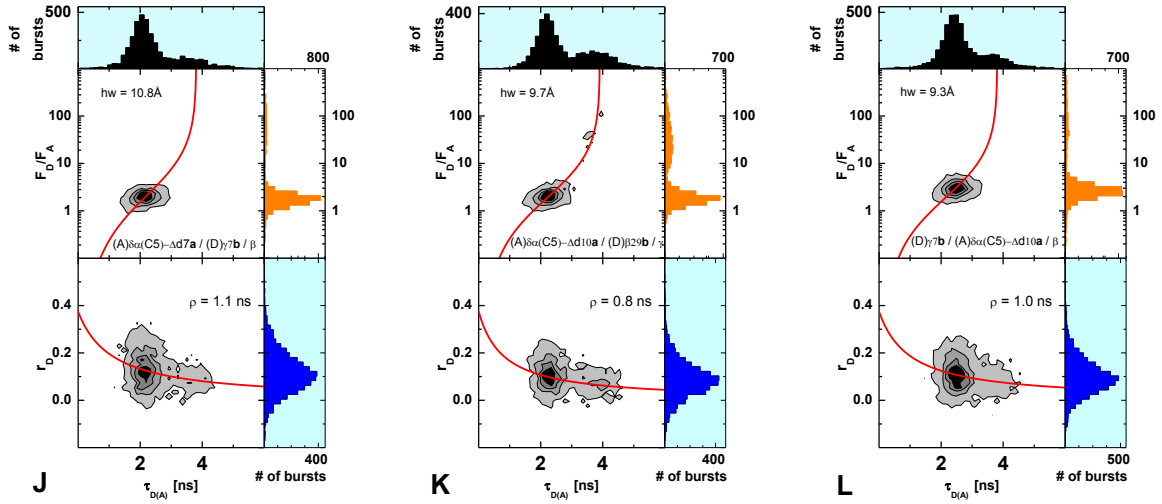
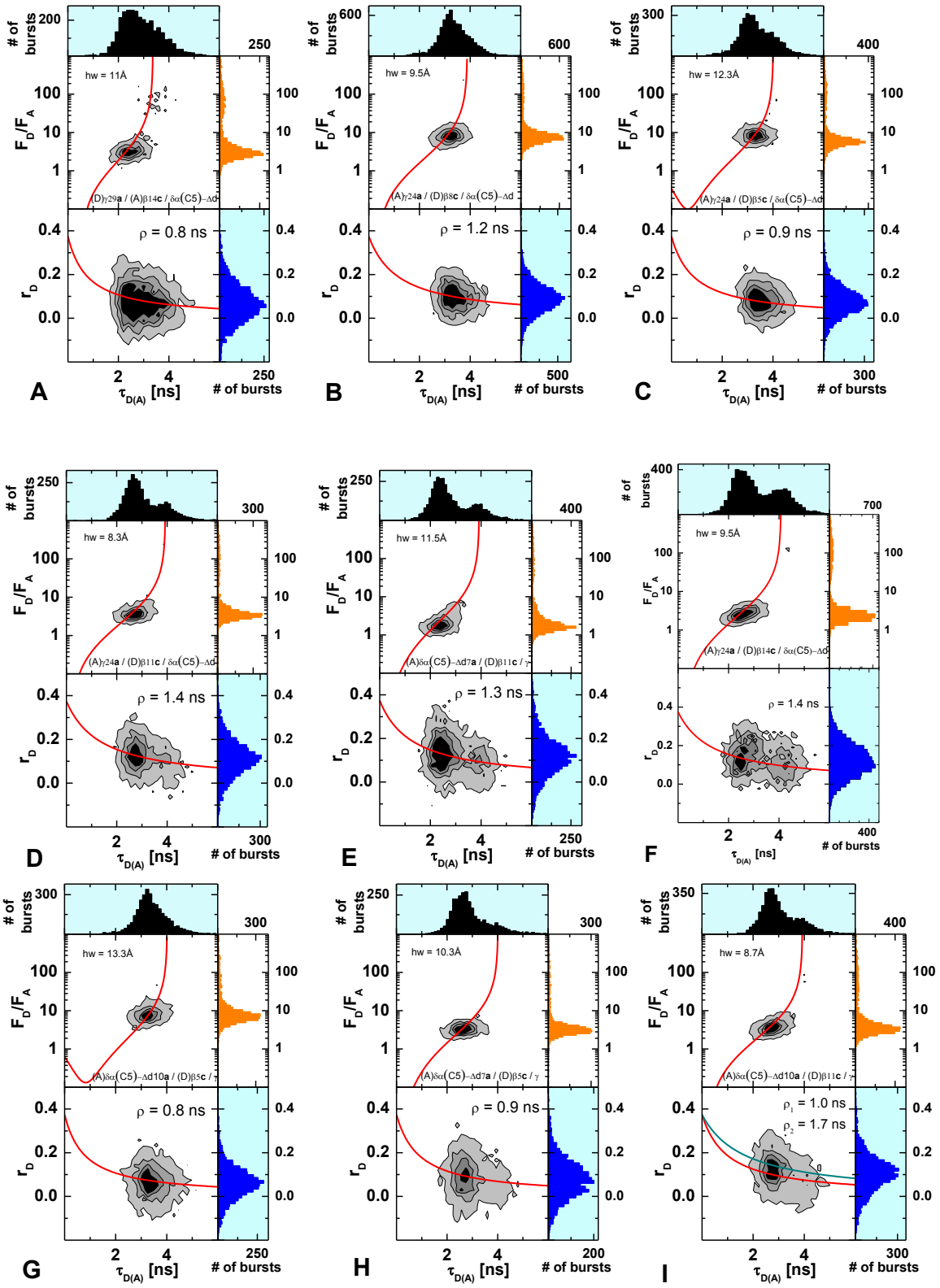


Figure 7-7 2D burst frequency histograms of F_D/F_A versus the donor fluorescence lifetime $\tau_{D(A)}$ (upper panel) and the donor fluorescence anisotropy r_D versus $\tau_{D(A)}$ (lower panel) for samples with fluorescence labels on the helices *a* and *b.* The number of molecules (fluorescence bursts) in each bin is gray scale, shaded from white (lowest) to black (highest). 1D histograms are shown as projections. In the F_D/F_A vs $\tau_{D(A)}$ plot, the theoretical relationship between F_D/F_A and $\tau_{D(A)}$ (static FRET line; red) is overlaid. The solid red line in the r_D-$\tau_{D(A)}$ diagram is given by the Perrin equation. Values of half with (hw) and rotation correlation time (ρ) are shown in the plots.


A
hw = 11Å
ρ = 0.8 ns
B
hw = 9.5Å
ρ = 1.2 ns
C
hw = 12.3Å
ρ = 0.9 ns
D
hw = 8.3Å
ρ = 1.4 ns
E
hw = 11.5Å
ρ = 1.3 ns
F
hw = 9.5Å
ρ = 1.4 ns
G
hw = 13.3Å
ρ = 0.8 ns
H
hw = 10.3Å
ρ = 0.9 ns
I
hw = 8.7Å
ρ1 = 1.0 ns
ρ2 = 1.7 ns
# of bursts
F_D/F_A
r_D
τ_D(A) [ns]

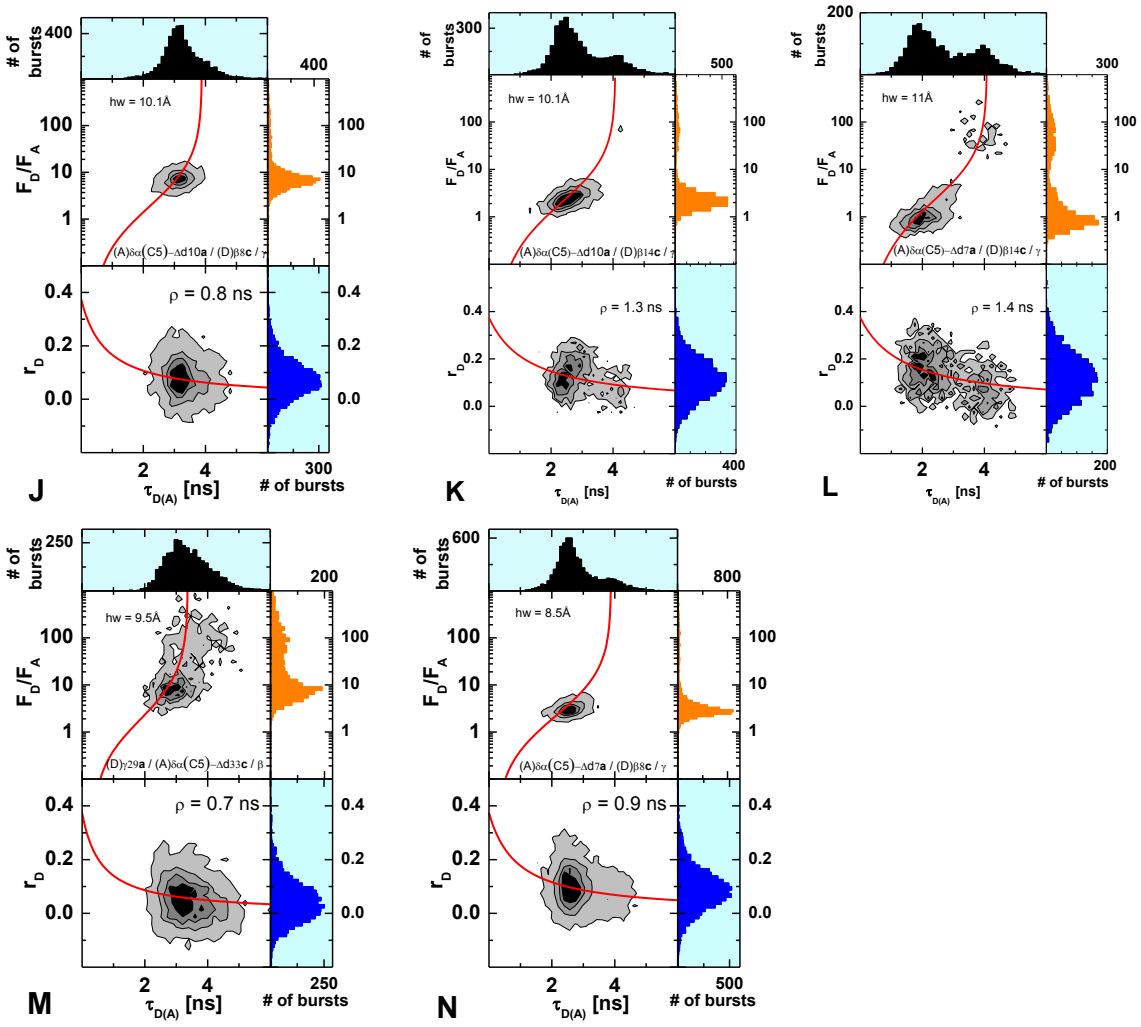
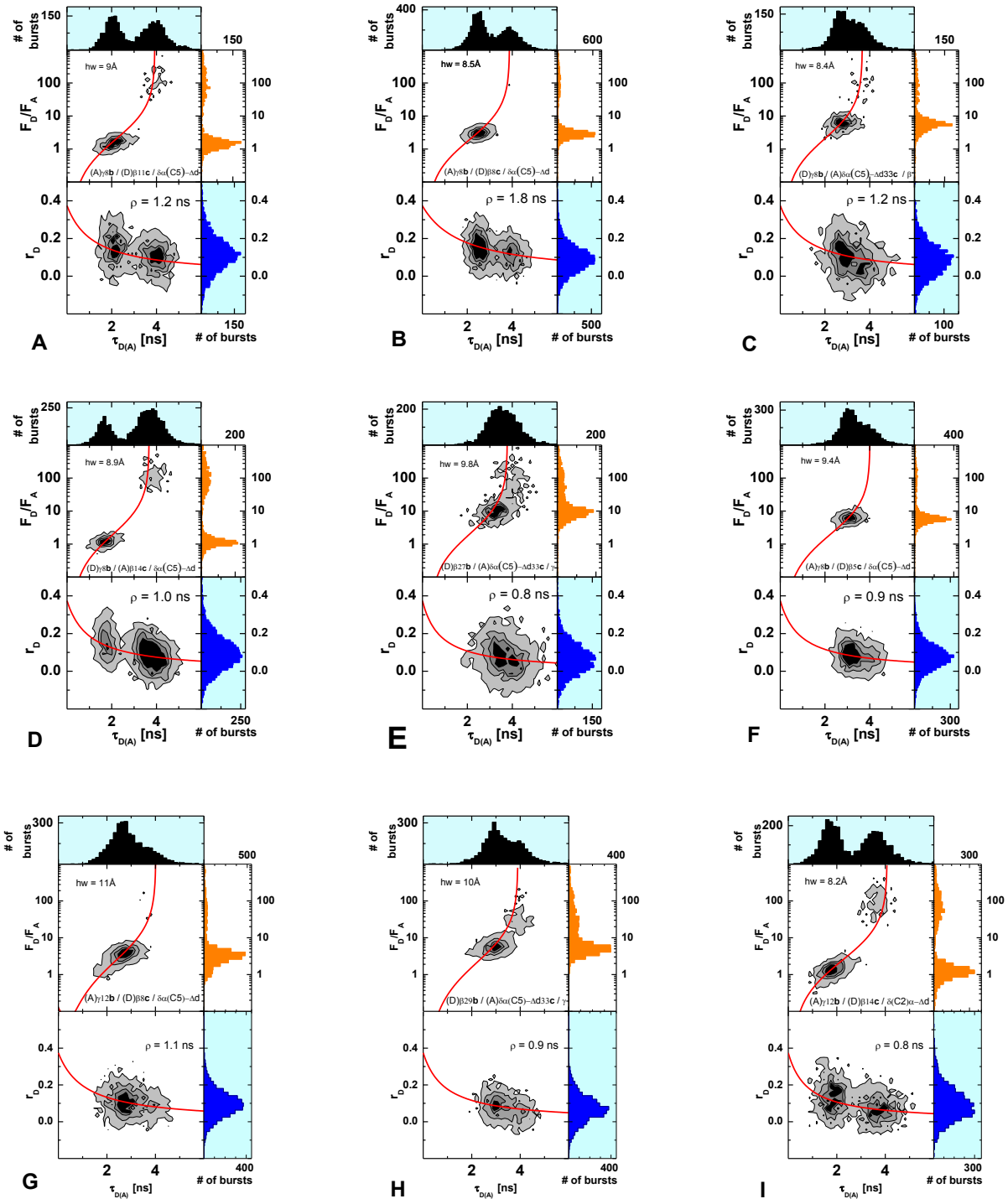


Figure 7-8 2D burst frequency histograms of F_D/F_A versus the donor fluorescence lifetime $\tau_{D(A)}$ (upper panel) and the donor fluorescence anisotropy r_D versus $\tau_{D(A)}$ (lower panel) for samples with fluorescence labels on the helices *a* and *c*. The number of molecules (fluorescence bursts) in each bin is gray scale, shaded from white (lowest) to black (highest). 1D histograms are shown as projections. In the F_D/F_A vs $\tau_{D(A)}$ plot, the theoretical relationship between F_D/F_A and $\tau_{D(A)}$ (static FRET line; red) is overlaid. The solid red line in the r_D-$\tau_{D(A)}$ diagram is given by the Perrin equation. Values of half with (hw) and rotation correlation time (ρ) are shown in the plots.


A
hw = 9Å
ρ = 1.2 ns
B
hw = 8.5Å
ρ = 1.8 ns
C
hw = 8.4Å
ρ = 1.2 ns
D
hw = 8.9Å
ρ = 1.0 ns
E
hw = 9.8Å
ρ = 0.8 ns
F
hw = 9.4Å
ρ = 0.9 ns
G
hw = 11Å
ρ = 1.1 ns
H
hw = 10Å
ρ = 0.9 ns
I
hw = 8.2Å
ρ = 0.8 ns
# of bursts
F_D/F_A
r_D
τ_D(A) [ns]

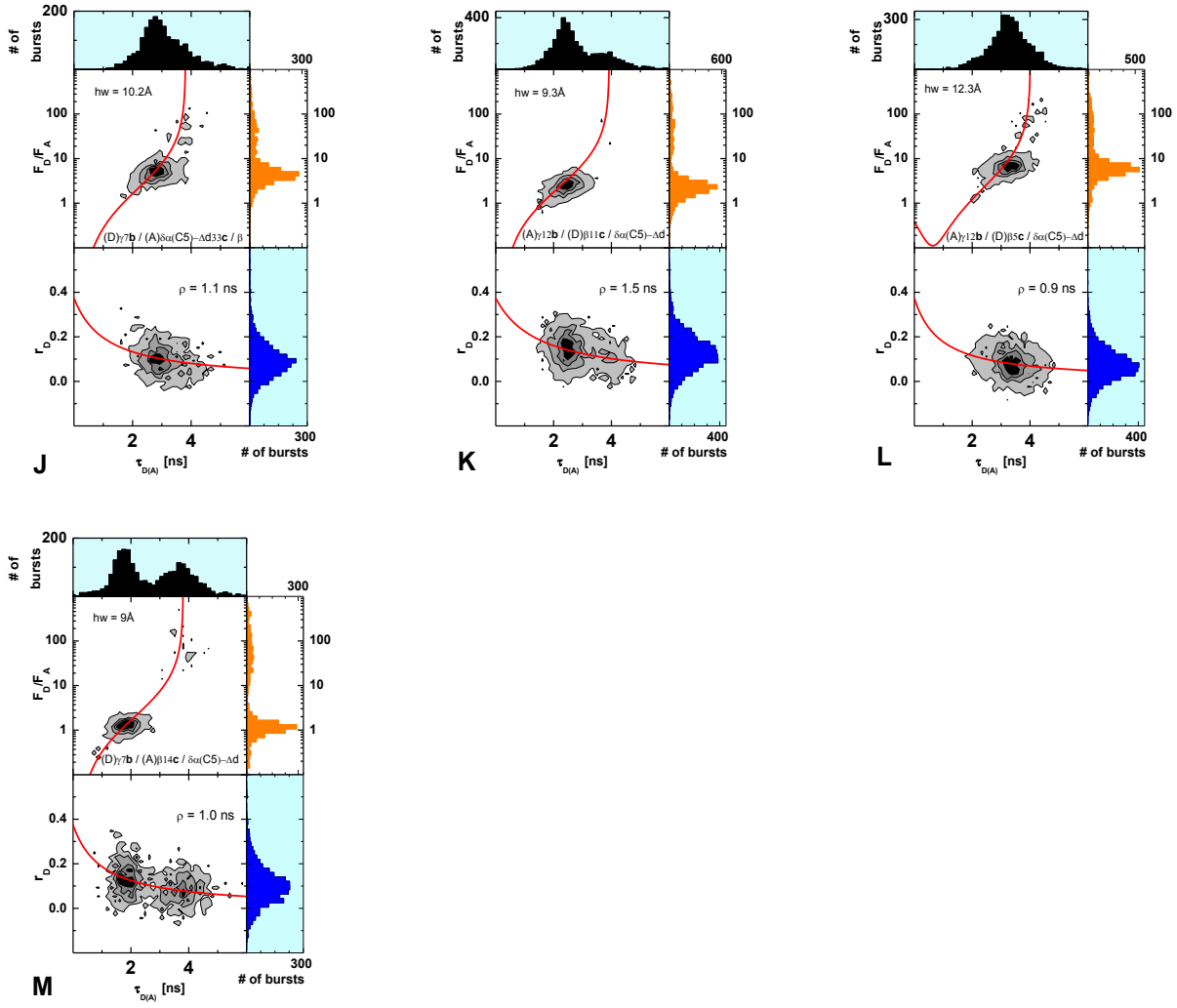


Figure 7-9 2D burst frequency histograms of F_D/F_A versus the donor fluorescence lifetime $\tau_{D(A)}$ (upper panel) and the donor fluorescence anisotropy r_D versus $\tau_{D(A)}$ (lower panel) for samples with fluorescence labels on the helices *b* and *c*. The number of molecules (fluorescence bursts) in each bin is gray scale, shaded from white (lowest) to black (highest). 1D histograms are shown as projections. In the F_D/F_A vs $\tau_{D(A)}$ plot, the theoretical relationship between F_D/F_A and $\tau_{D(A)}$ (static FRET line; red) is overlaid. The solid red line in the r_D-$\tau_{D(A)}$ diagram is given by the Perrin equation. Values of half with (hw) and rotation correlation time (ρ) are shown in the plots.

7.4 Static FRET lines

Table 38 Static FRET lines used in fits for J(abc) molecule (section 7.3.1.1)

DA-pair	formula	half with
(D)β27b/(A)δα–Δd28c/γ	(0.6294/0.3010)/((3.7781/((-0.0582*x^3)+(0.3479*x^2)+0.5423*x+-0.1042))-1)	hw=7.4
(A)δα_Δd10a/(D)β5c/γ	(0.7807/0.3450)/((4.0159/((-0.1144*x^3)+(0.9200*x^2)+-0.9495*x+0.4021))-1)	hw=12
(A)δα_Δd10a/(D)β8c/γ	(0.6936/0.3450)/((3.8831/((-0.0424*x^3)+(0.2606*x^2)+0.6418*x+-0.0605))-1)	hw=9.4
(D)γ8b/(A)δα–Δd28c/β	(0.6731/0.3010)/((3.6737/((-0.0646*x^3)+(0.4043*x^2)+0.4091*x+-0.0836))-1)	hw=7.5
(A)δα_Δd7a/(D)γ8b/β	(0.6731/0.3020)/((3.6737/((-0.0965*x^3)+(0.6462*x^2)+-0.0589*x+-0.0507))-1)	hw =10
(D)γ29a/(A)δα–Δd28c/β	(0.6730/0.3010)/((3.3649/((-0.0563*x^3)+(0.3536*x^2)+0.4583*x+-0.0393))-1)	hw=6.1
(A)δα_Δd10a/(D)β27b/γ	(0.6294/0.3450)/((3.7781/((-0.0695*x^3)+(0.4337*x^2)+0.3873*x+-0.1309))-1)	hw=8.6
(A)δα_Δd7a/(D)β8c/γ	(0.6936/0.3020)/((3.8831/((-0.0424*x^3)+(0.2606*x^2)+0.6418*x+-0.0605))-1)	hw=6
(A)δα_Δd7a/(D)β27b/γ	(0.6294/0.3020)/((3.7781/((-0.0932*x^3)+(0.6162*x^2)+0.0415*x+-0.1510))-1)	hw=10.5
(A)δα_Δd7a/(D)β11c/γ	(0.7276/0.3450)/((3.9274/((-0.0407*x^3)+(0.2686*x^2)+0.5860*x+-0.0536))-1)	hw=6
(A)δα_Δd10a/(D)β14c/γ	(0.7723/0.3450)/((4.0791/((-0.0372*x^3)+(0.2617*x^2)+0.5629*x+-0.0518))-1)	hw=6
(A)δα_Δd7a/(D)β14c/γ	(0.7723/0.3020)/((4.0791/((-0.0372*x^3)+(0.2617*x^2)+0.5629*x+-0.0518))-1)	hw=6
(A)δα_Δd7a/(D)β5c/γ	(0.7807/0.3020)/((4.0159/((-0.0388*x^3)+(0.2754*x^2)+0.5300*x+-0.0482))-1)	hw=6
(D)γ8b/(A)β14c/δα–Δd	(0.6731/0.3310)/((3.6737/((-0.0809*x^3)+(0.5270*x^2)+0.1786*x+-0.0852))-1)	hw=8.9
(D)γ29a/(A)β14c/δα–Δd	(0.5328/0.3310)/((3.3649/((-0.0875*x^3)+(0.4539*x^2)+0.5141*x+-0.1739))-1)	hw=8.6
(D)γ29a/(A)β27b/δα–Δd	(0.5328/0.3180)/((3.3649/((-0.0559*x^3)+(0.2470*x^2)+0.8234*x+-0.0763))-1)	hw=8.5
(D)γ29a/(A)β33b/δα–Δd	(0.5328/0.4600)/((3.3649/((-0.1094*x^3)+(0.6008*x^2)+0.2906*x+-0.2492))-1)	hw=10.3
(A)γ24a/(D)β11c/δα–Δd	(0.7276/0.3240)/((3.9274/((-0.0644*x^3)+(0.4564*x^2)+0.2200*x+-0.0748))-1)	hw=8.5
(A)γ24a/(D)β8c/δα_Δd	(0.6936/0.3210)/((3.8831/((-0.0763*x^3)+(0.5267*x^2)+0.1294*x+-0.0939))-1)	hw=9.4
(A)γ24a/(D)β27b/δα–Δd	(0.6294/0.3210)/((3.7781/((-0.0868*x^3)+(0.5667*x^2)+0.1371*x+-0.1505))-1)	hw=10
(A)γ24a/(D)β5c/δα–Δd	(0.7807/0.3210)/((4.0159/((-0.1251*x^3)+(1.0159*x^2)+-1.2001*x+0.5546))-1)	hw=12.5
(A)γ24a/(D)β14c/δα–Δd	(0.7723/0.3020)/((4.0791/((-0.0644*x^3)+(0.4873*x^2)+0.0952*x+-0.0513))-1)	hw=9
(A)γ8b/(D)β5c/δα–Δd	(0.7807/0.3350)/((4.0159/((-0.0388*x^3)+(0.2754*x^2)+0.5300*x+-0.0482))-1)	hw=6
(A)γ8b/(D)β8c/δα–Δd	(0.6936/0.3350)/((3.8831/((-0.0515*x^3)+(0.3299*x^2)+0.5154*x+-0.0825))-1)	hw=7
(A)γ8b/(D)β11c/δα–Δd	(0.7276/0.3350)/((3.9274/((-0.0984*x^3)+(0.7378*x^2)+-0.4011*x+0.0820))-1)	hw=11
(A)γ12b/(D)β14c/δα–Δd	(0.7723/0.2860)/((4.0791/((-0.0655*x^3)+(0.4969*x^2)+0.0738*x+-0.0473))-1)	hw=9.1
(A)γ12b/(D)β11c/δα–Δd	(0.7276/0.2860)/((3.9274/((-0.0984*x^3)+(0.7378*x^2)+-0.4011*x+0.0820))-1	hw=11
(A)γ12b/(D)β8c/δα–Δd	(0.6936/0.2860)/((3.8831/((-0.0716*x^3)+(0.4889*x^2)+0.2066*x+-0.0999))-1)	hw=9
(A)γ12b/(D)β5c/δα–Δd	(0.7807/0.2860)/((4.0159/((-0.0907*x^3)+(0.7104*x^2)+-0.4231*x+0.1288))-1)	hw=10.7
(A)δα_Δd7a/(D)γ7b/β	(0.7202/0.3020)/((3.8118/((-0.0976*x^3)+(0.7062*x^2)+-0.2860*x+0.0447))-1)	hw=10.5
(A)δα_Δd10a/(D)γ7b/β	(0.7202/0.3450)/((3.8118/((-0.1045*x^3)+(0.7627*x^2)+-0.4135*x+0.0921))-1)	hw=10.9
(A)δα_Δd10a/(D)β11c/γ	(0.7276/0.3450)/((3.9274/((-0.0407*x^3)+(0.2686*x^2)+0.5860*x+-0.0536))-1)	hw=6
(D)γ7b/(A)δα–Δd28c/β	(0.7202/0.3010)/((3.8118/((-0.0837*x^3)+(0.5939*x^2)+-0.0400*x+-0.0284))-1)	hw=9.6
(D)γ7b/(A)β14c/δα–Δd	(0.7202/0.3310)/((3.8118/((-0.1157*x^3)+(0.8556*x^2)+-0.6286*x+0.1839))-1)	hw=11.5
(A)δα_Δd7a/(D)β29b/γ	(0.7411/0.3020)/((3.9091/((-0.0921*x^3)+(0.6902*x^2)+-0.3066*x+0.0633))-1)	hw=10.5
(A)δα_Δd10a/(D)β29b/γ	(0.7411/0.3450)/((3.9091/((-0.1238*x^3)+(0.9607*x^2)+-0.9585*x+0.3713))-1)	hw=12.2
(D)β29b/(A)δα–Δd28c/γ	(0.7411/0.3010)/((3.9091/((-0.0747*x^3)+(0.5455*x^2)+0.0178*x+-0.0350))-1)	hw=9.3
(A)γ24a/(D)β29b/δα–Δd	(0.7411/0.3210)/((3.9091/((-0.1176*x^3)+(0.9065*x^2)+-0.8239*x+0.2988))-1)	hw=11.9
(D)β29b/(A)δα–Δd26c/γ	(0.7411/0.2810)/((3.9091/((-0.0774*x^3)+(0.5673*x^2)+-0.0298*x+-0.0238))-1)	hw=9.5
(D)β27b/(A)δα–Δd26c/γ	(0.6294/0.3020)/((3.7781/((-0.0999*x^3)+(0.6685*x^2)+-0.0612*x+-0.1479))-1)	hw =11
(D)γ7b/(A)δα–Δd26c/β	(0.7202/0.2810)/((3.8118/((-0.1045*x^3)+(0.7627*x^2)+-0.4135*x+0.0921))-1)	hw=10.9
(D)γ8b/(A)δα–Δd26c/β	(0.6731/0.2810)/((3.6737/((-0.0949*x^3)+(0.6344*x^2)+-0.0348*x+-0.0555))-1)	hw=9.9
(D)γ29a/(A)δα–Δd26c/β	(0.5328/0.2810)/((3.3649/((-0.1345*x^3)+(0.7690*x^2)+0.0339*x+-0.3313))-1)	hw=12
(A)δα_Δd10a/(D)γ8b/β	(0.6731/0.3450)/((3.6737/((-0.0759*x^3)+(0.4891*x^2)+0.2515*x+-0.0888))-1)	hw=8.5

Table 39 Static FRET lines used in fits for J(abc(C2)) molecule (section 7.3.1.2).

DA-pair	formula	half with
(A)γ12b/(D)β14c/δα(C2)–Δd	(0.7723/0.2860)/((4.0791/((-0.0644*x^3)+(0.4873*x^2)+0.0952*x+-0.0513))-1)	hw=9
(A)γ12b/(D)β11c/δα(C2)–Δd	(0.7276/0.2860)/((3.9274/((-0.0831*x^3)+(0.6097*x^2)+-0.1091*x+-0.0155))-1)	hw=10
(A)γ12b/(D)β8c/δα(C2)–Δd	(0.6936/0.2860)/((3.8831/((-0.0751*x^3)+(0.5170*x^2)+0.1493*x+-0.0958))-1)	hw=9.3
(A)γ12b/(D)β5c/δα(C2)–Δd	(0.7807/0.2860)/((4.0159/((-0.1046*x^3)+(0.8331*x^2)+-0.7271*x+0.2778))-1)	hw=11.5
(A)δα(C2)_Δd7a/(D)γ7b/β	(0.7202/0.3020)/((3.8118/((-0.1348*x^3)+(1.0159*x^2)+-1.0130*x+0.3773))-1)	hw=12.4
(A)δα(C2)_Δd10a/(D)γ7b/β	(0.7202/0.3450)/((3.8118/((-0.0960*x^3)+(0.6928*x^2)+-0.2559*x+0.0343))-1)	hw=10.4
(D)γ7b/(A)δα(C2)–Δd30c/β	(0.7202/0.3010)/((3.8118/((-0.0881*x^3)+(0.6292*x^2)+-0.1161*x+-0.0088))-1)	hw=9.9
(D)γ7b/(A)β14c/δα(C2)–Δd	(0.7202/0.3310)/((3.8118/((-0.0672*x^3)+(0.4623*x^2)+0.2322*x+-0.0709))-1)	hw=8.3
(A)δα(C2)_Δd7a/(D)β29b/γ	(0.7411/0.3020)/((3.9091/((-0.0954*x^3)+(0.7177*x^2)+-0.3705*x+0.0879))-1)	hw=10.7
(A)δα(C2)_Δd10a/(D)β29b/γ	(0.7411/0.3450)/((3.9091/((-0.0801*x^3)+(0.5900*x^2)+-0.0799*x+-0.0108))-1)	hw=9.7
(D)β29b/(A)δα(C2)-Δd30c/γ	(0.7411/0.3010)/((3.9091/((-0.0890*x^3)+(0.6637*x^2)+-0.2457*x+0.0412))-1)	hw=10.3
(A)γ24a/(D)β29b/δα(C2)–Δd	(0.7411/0.3210)/((3.9091/((-0.0844*x^3)+(0.6258*x^2)+-0.1597*x+0.0127))-1)	hw=10
(D)γ29a/(A)β14c/δα(C2)–Δd	(0.5328/0.3310)/((3.3649/((-0.0975*x^3)+(0.5210*x^2)+0.4122*x+-0.2083))-1)	hw=9.4
(D)γ29a/(A)β33b/δα(C2)–Δd	(0.5328/0.4600)/((3.3649/((-0.0925*x^3)+(0.4870*x^2)+0.4638*x+-0.1909))-1)	hw=9
(A)γ8b/(D)β8c/δα(C2)–Δd	(0.6936/0.3350)/((3.8831/((-0.0478*x^3)+(0.3019*x^2)+0.5671*x+-0.0739))-1)	hw=6.6
(A)γ24a/(D)β11c/δα(C2)–Δd	(0.7276/0.3210)/((3.9274/((-0.0602*x^3)+(0.4220*x^2)+0.2900*x+-0.0781))-1)	hw=8.1
(A)γ24a/(D)β27b/δα(C2)–Δd	(0.6294/0.3210)/((3.7781/((-0.0958*x^3)+(0.6367*x^2)+0.0013*x+-0.1502))-1)	hw=10.7
(D)γ8b/(A)β14c/δα(C2)–Δd	(0.6731/0.3310)/((3.6737/((-0.0809*x^3)+(0.5270*x^2)+0.1786*x+-0.0852))-1)	hw=8.9
(A)γ24a/(D)β14c/δα(C2)–Δd	(0.7723/0.3210)/((4.0791/((-0.0633*x^3)+(0.4779*x^2)+0.1161*x+-0.0549))-1)	hw=8.9
(A)γ8b/(D)β5c/δα(C2)–Δd	(0.7807/0.3350)/((4.0159/((-0.0623*x^3)+(0.4686*x^2)+0.1353*x+-0.0490))-1)	hw=8.6
(A)γ24a/(D)β5c/δα(C2)–Δd	(0.7807/0.3210)/((4.0159/((-0.1010*x^3)+(0.8006*x^2)+-0.6455*x+0.2352))-1)	hw=11.3
(A)γ8b/(D)β11c/δα(C2)–Δd	(0.7807/0.3350)/((4.0159/((-0.0613*x^3)+(0.4595*x^2)+0.1550*x+-0.0519))-1)	hw=8.5
(D)γ29a/(A)β27b/δα(C2)–Δd	(0.5328/0.3180)/((3.3649/((-0.1027*x^3)+(0.5558*x^2)+0.3591*x+-0.2261))-1)	hw=9.8
(A)δα(C2)_Δd7a/(D)β5c/γ	(0.7807/0.3020)/((4.0159/((-0.0571*x^3)+(0.4249*x^2)+0.2289*x+-0.0600))-1)	hw=8.1
(D)β27b/(A)δα(C2)–Δd30c/γ	(0.6294/0.3010)/((3.7781/((-0.0999*x^3)+(0.6685*x^2)+-0.0612*x+-0.1479))-1)	hw=11
(A)δα(C2)_Δd7a/(D)β8c/γ	(0.6936/0.3020)/((3.8831/((-0.0610*x^3)+(0.4042*x^2)+0.3745*x+-0.0987))-1)	hw=8
(A)δα(C2)_Δd7a/(D)β11c/γ	(0.7276/0.3020)/((3.9274/((-0.0702*x^3)+(0.5029*x^2)+0.1229*x+-0.0643))-1)	hw=9
(A)γ24a/(D)β8c/δα(C2)_Δd	(0.6936/0.3210)/((3.8831/((-0.0705*x^3)+(0.4799*x^2)+0.2250*x+-0.1008))-1)	hw=8.9
(A)δα(C2)_Δd7a/(D)γ8b/β	(0.6731/0.3020)/((3.6737/((-0.1096*x^3)+(0.7484*x^2)+-0.2720*x+0.0036))-1)	hw=10.8
(D)γ8b/(A)δα(C2)–Δd30c/β	(0.6731/0.3010)/((3.6737/((-0.0555*x^3)+(0.3365*x^2)+0.5298*x+-0.0684))-1)	hw=6.7
(D)γ29a/(A)δα(C2)–Δd30c/β	(0.5328/0.3010)/((3.3649/((-0.1345*x^3)+(0.7690*x^2)+0.0339*x+-0.3313))-1)	hw=12
(A)δα(C2)_Δd7a/(D)β27b/γ	(0.6294/0.3020)/((3.7781/((-0.1027*x^3)+(0.6903*x^2)+-0.1043*x+-0.1455))-1)	hw=11.2
(A)δα(C2)_Δd10a/(D)β27b/γ	(0.6294/0.3450)/((3.7781/((-0.0919*x^3)+(0.6060*x^2)+0.0611*x+-0.1511))-1)	hw=10.4
(A)δα(C2)_Δd7a/(D)β14c/γ	(0.7723/0.3020)/((4.0791/((-0.0753*x^3)+(0.5810*x^2)+-0.1179*x+0.0011))-1)	hw=9.9
(A)δα(C2)_Δd10a/(D)γ8b/β	(0.6731/0.3450)/((3.6737/((-0.0919*x^3)+(0.6114*x^2)+0.0118*x+-0.0642))-1)	hw=9.7
(A)δα(C2)_Δd10a/(D)β5c/γ	(0.7807/0.3450)/((4.0159/((-0.1104*x^3)+(0.8842*x^2)+-0.8573*x+0.3492))-1)	hw=11.8
(A)δα(C2)_Δd10a/(D)β8c/γ	(0.6936/0.3450)/((3.8831/((-0.0740*x^3)+(0.5075*x^2)+0.1688*x+-0.0975))-1)	hw=9.2
(A)δα(C2)_Δd10a/(D)β11c/γ	(0.7276/0.3450)/((3.9274/((-0.0592*x^3)+(0.4138*x^2)+0.3066*x+-0.0783))-1)	hw=8
(A)δα(C2)_Δd10a/(D)β14c/γ	(0.7723/0.3450)/((4.0791/((-0.0644*x^3)+(0.4873*x^2)+0.0952*x+-0.0513))-1)	hw=9

Table 40 Static FRET lines used in fits for J(abc(C5)) molecule (section7.3.1.3).

DA-pair	formula	half with
(A)γ12b/(D)β11c/δα(C5)–Δd	(0.7276/0.2860)/((3.9274/((-0.0738*x^3)+(0.5329*x^2)+0.0589*x+-0.0540))-1)	hw=9.3
(A)γ12b/(D)β14c/δα(C5)–Δd	(0.7723/0.2860)/((4.0791/((-0.0560*x^3)+(0.4168*x^2)+0.2488*x+-0.0700))-1)	hw=8.2
(A)γ12b/(D)β8c/δα(C5)–Δd	(0.7807/0.2860)/((4.0159/((-0.0957*x^3)+(0.7541*x^2)+-0.5302*x+0.1781))-1)	hw=11
(A)γ12b/(D)β5c/δα(C5)–Δd	(0.7807/0.2860)/((4.0159/((-0.1207*x^3)+(0.9764*x^2)+-1.0962*x+0.4900))-1)	hw=12.3
(A)δα(C5)_Δd10a/(D)γ7b/β	(0.7202/0.3450)/((3.8118/((-0.0796*x^3)+(0.5605*x^2)+0.0308*x+-0.0440))-1)	hw=9.3
(A)δα(C5)_Δd7a/(D)γ7b/β	(0.7202/0.3020)/((3.8118/((-0.1027*x^3)+(0.7482*x^2)+-0.3805*x+0.0793))-1)	hw=10.8
(D)γ7b/(A)δα(C5)–Δd33c/β	(0.7202/0.3010)/((3.8118/((-0.0927*x^3)+(0.6666*x^2)+-0.1980*x+0.0155))-1)	hw=10.2
(D)γ7b/(A)β14c/δα(C5)–Δd	(0.7202/0.3310)/((3.8118/((-0.0756*x^3)+(0.5291*x^2)+0.0964*x+-0.0557))-1)	hw=9
(A)δα(C5)_Δd7a/(D)β29b/γ	(0.7411/0.3020)/((3.9091/((-0.0661*x^3)+(0.4755*x^2)+0.1667*x+-0.0607))-1)	hw=8.6
(A)δα(C5)_Δd10a/(D)β29b/γ	(0.7411/0.3450)/((3.9091/((-0.0801*x^3)+(0.5900*x^2)+-0.0799*x+-0.0108))-1)	hw=9.7
(D)β29b/(A)δα(C5)-Δd33c/γ	(0.7411/0.3010)/((3.9091/((-0.0844*x^3)+(0.6258*x^2)+-0.1597*x+0.0127))-1)	hw=10
(A)γ24a/(D)β29b/δα(C5)–Δd	(0.7411/0.3210)/((3.9091/((-0.0815*x^3)+(0.6017*x^2)+-0.1058*x+-0.0035))-1)	hw=9.8
(A)γ24a/(D)β14c/δα(C5)–Δd	(0.7723/0.3210)/((4.0791/((-0.0702*x^3)+(0.5373*x^2)+-0.0173*x+-0.0270))-1)	hw=9.5
(A)δα(C5)_Δd10a/(D)β14c/γ	(0.7723/0.3450)/((4.0791/((-0.0779*x^3)+(0.6041*x^2)+-0.1721*x+0.0183))-1)	hw=10.1
(A)γ24a/(D)β27b/δα(C5)–Δd	(0.6294/0.3210)/((3.7781/((-0.0868*x^3)+(0.5667*x^2)+0.1371*x+-0.1505))-1)	hw=10
(A)δα(C5)_Δd10a/(D)β27b/γ	(0.6294/0.3450)/((3.7781/((-0.0844*x^3)+(0.5477*x^2)+0.1736*x+-0.1493))-1)	hw=9.8
(A)γ8b/(D)β8c/δα(C5)–Δd	(0.6936/0.3350)/((3.8831/((-0.0661*x^3)+(0.4450*x^2)+0.2947*x+-0.1018))-1)	hw=8.5
(A)γ24a/(D)β8c/δα(C5)_Δd	(0.6936/0.3210)/((3.8831/((-0.0776*x^3)+(0.5365*x^2)+0.1091*x+-0.0918))-1)	hw=9.5
(D)γ8b/(A)β14c/δα(C5)–Δd	(0.6731/0.3310)/((3.6737/((-0.0809*x^3)+(0.5270*x^2)+0.1786*x+-0.0852))-1)	hw=8.9
(D)γ29a/(A)β14c/δα(C5)–Δd	(0.5328/0.3310)/((3.3649/((-0.1193*x^3)+(0.6668*x^2)+0.1902*x+-0.2827))-1)	hw=11
(D)γ29a/(A)β27b/δα(C5)–Δd	(0.5328/0.3180)/((3.3649/((-0.1193*x^3)+(0.6668*x^2)+0.1902*x+-0.2827))-1)	hw=11
(A)γ8b/(D)β5c/δα(C5)–Δd	(0.7807/0.3350)/((4.0159/((-0.0718*x^3)+(0.5485*x^2)+-0.0423*x+-0.0104))-1)	hw=9.4
(A)γ24a/(D)β5c/δα(C5)–Δd	(0.7807/0.3210)/((4.0159/((-0.1207*x^3)+(0.9764*x^2)+-1.0962*x+0.4900))-1)	hw=12.3
(A)γ8b/(D)β11c/δα(C5)–Δd	(0.7276/0.3350)/((3.9274/((-0.0702*x^3)+(0.5029*x^2)+0.1229*x+-0.0643))-1)	hw=9
(A)γ24a/(D)β11c/δα(C5)–Δd	(0.7276/0.3210)/((3.9274/((-0.0623*x^3)+(0.4389*x^2)+0.2558*x+-0.0769))-1)	hw=8.3
(D)γ29a/(A)δα(C5)–Δd33c/β	(0.6730/0.4600)/((3.3649/((-0.0989*x^3)+(0.6503*x^2)+-0.0725*x+0.0107))-1)	hw=9
(A)δα(C5)_Δd7a/(D)γ8b/β	(0.6731/0.3020)/((3.6737/((-0.0595*x^3)+(0.3659*x^2)+0.4780*x+-0.0759))-1)	hw=7
(A)δα(C5)_Δd7a/(D)β11c/γ	(0.7276/0.3020)/((3.9274/((-0.1070*x^3)+(0.8109*x^2)+-0.5736*x+0.1537))-1)	hw=11.5
(A)δα(C5)_Δd7a/(D)β27b/γ	(0.6294/0.3020)/((3.7781/((-0.0784*x^3)+(0.5019*x^2)+0.2604*x+-0.1444))-1)	hw=9.3
(A)δα(C5)_Δd7a/(D)β5c/γ	(0.7807/0.3020)/((4.0159/((-0.0844*x^3)+(0.6560*x^2)+-0.2922*x+0.0739))-1	hw=10.3
(A)δα(C5)_Δd7a/(D)β8c/γ	(0.6936/0.3020)/((3.8831/((-0.0661*x^3)+(0.4450*x^2)+0.2947*x+-0.1018))-1)	hw=8.5
(D)γ29a/(A)β33b/δα(C5)–Δd	(0.5328/0.3010)/((3.3649/((-0.0988*x^3)+(0.5296*x^2)+0.3991*x+-0.2127))-1)	hw=9.5
(D)β27b/(A)δα(C5)–Δd33c/γ	(0.6294/0.3010)/((3.7781/((-0.0844*x^3)+(0.5477*x^2)+0.1736*x+-0.1493))-1)	hw=9.8
(D)γ8b/(A)δα(C5)–Δd33c/β	(0.6731/0.3010)/((3.6737/((-0.0747*x^3)+(0.4800*x^2)+0.2687*x+-0.0892))-1)	hw=8.4
(A)δα(C5)_Δd10a/(D)β5c/γ	(0.7807/0.3450)/((4.0159/((-0.1443*x^3)+(1.1901*x^2)+-1.6694*x+0.8689))-1)	hw=13.3
(A)δα(C5)_Δd10a/(D)β8c/γ	(0.6936/0.3450)/((3.8831/((-0.0853*x^3)+(0.5990*x^2)+-0.0222*x+-0.0720))-1)	hw=10.1
(A)δα(C5)_Δd10a/(D)β11c/γ	(0.7276/0.3450)/((3.9274/((-0.0667*x^3)+(0.4745*x^2)+0.1825*x+-0.0716))-1)	hw=8.7
(A)δα(C5)_Δd10a/(D)γ8b/β	(0.6731/0.3450)/((3.6737/((-0.0736*x^3)+(0.4711*x^2)+0.2856*x+-0.0893))-1)	hw=8.3
(A)δα(C5)_Δd7a/(D)β14c/γ	(0.7723/0.3020)/((4.0791/((-0.0912*x^3)+(0.7198*x^2)+-0.4506*x+0.1269))-1)	hw=11

7.5 PDA analysis

7.5.1 Measurements at 20 mM $MgCl_2$ for J(abc)

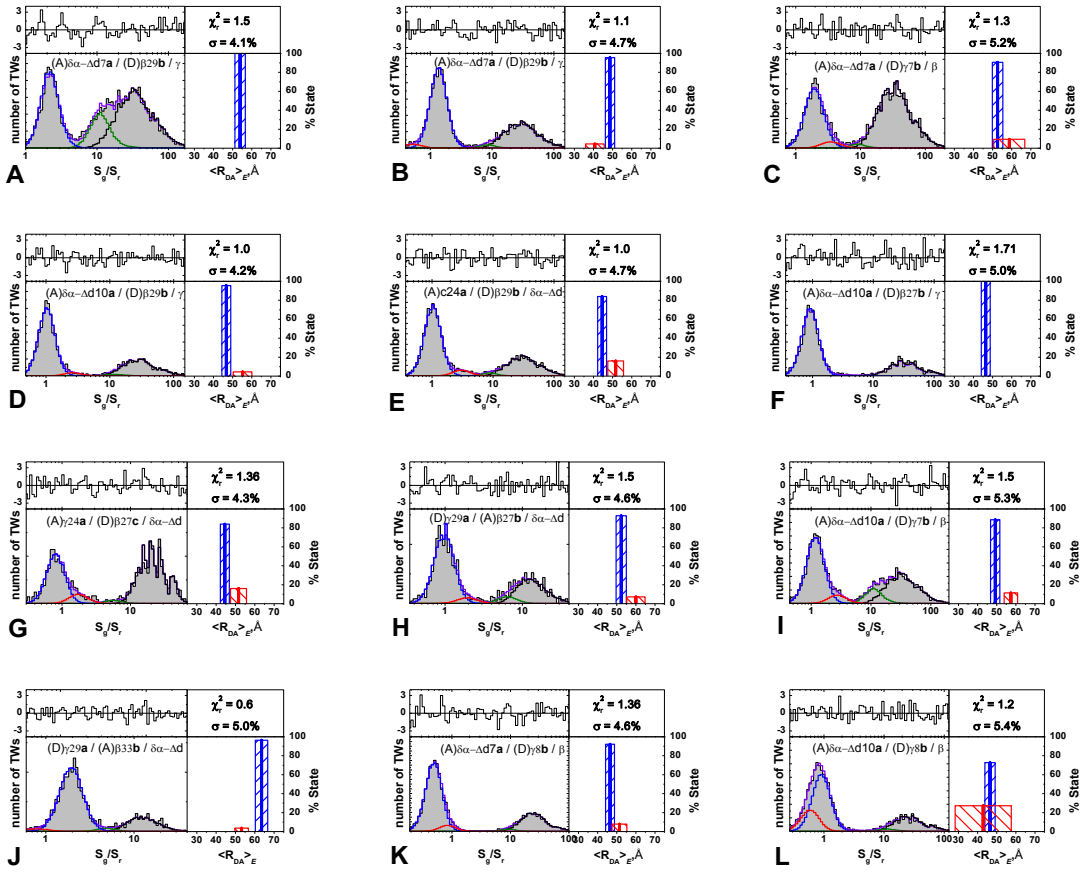


Figure 7-10 PDA analysis for the samples with fluorescence labels on the helices *a* and *b*. S_g/S_r histograms of experimental data (gray area) are fitted (violet solid line) with two FRET states ($\langle R_{DA}\rangle_{E(1)}$, blue solid line for major population and $\langle R_{DA}\rangle_{E(2)}$, red solid line for minor population, a D-only state (black solid line) and, if necessary, a state accounting for impurities (dark yellow solid line). Weighted residuals are shown in the upper plot. The right panel shows the distances, relative amplitudes (solid lines) and the confidence intervals (striped boxes) of the three FRET states (blue for major population, red for minor population). See Table 21 for all parameters used.

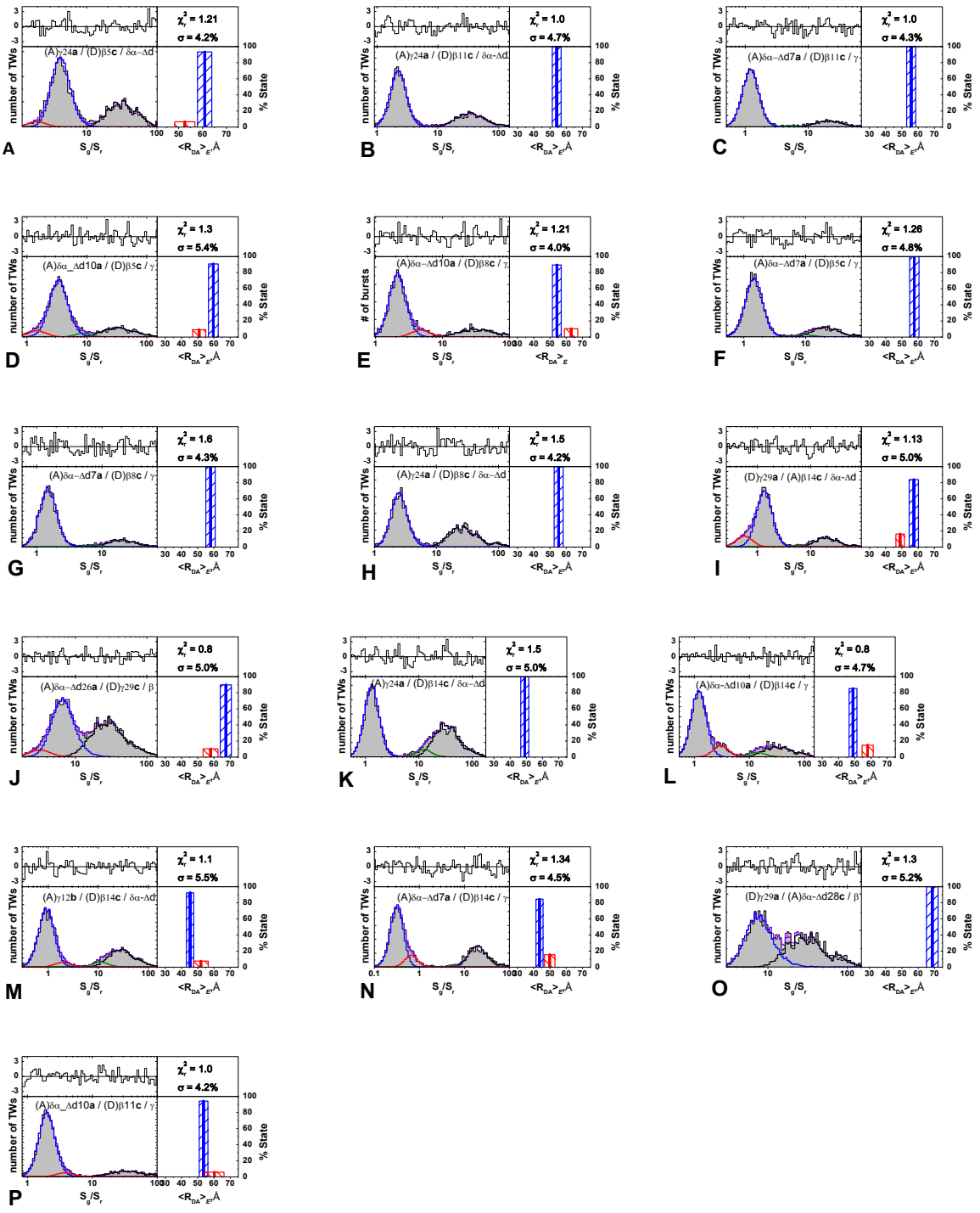


Figure 7-11 PDA analysis for the samples with fluorescence labels on the helices *a* and *c*. S_g/S_r histograms of experimental data (gray area) are fitted (violet solid line) with two FRET states ($\langle R_{DA}\rangle_{E(1)}$, blue solid line for major population and $\langle R_{DA}\rangle_{E(2)}$, red solid line for minor population, a D-only state (black solid line) and, if necessary, a state accounting for impurities (dark yellow solid line). Weighted residuals are shown in the upper plot. The right panel shows the distances, relative amplitudes (solid lines) and the confidence intervals (striped boxes) of the three FRET states (blue for major population, red for minor population). See Table 21 for all parameters used.

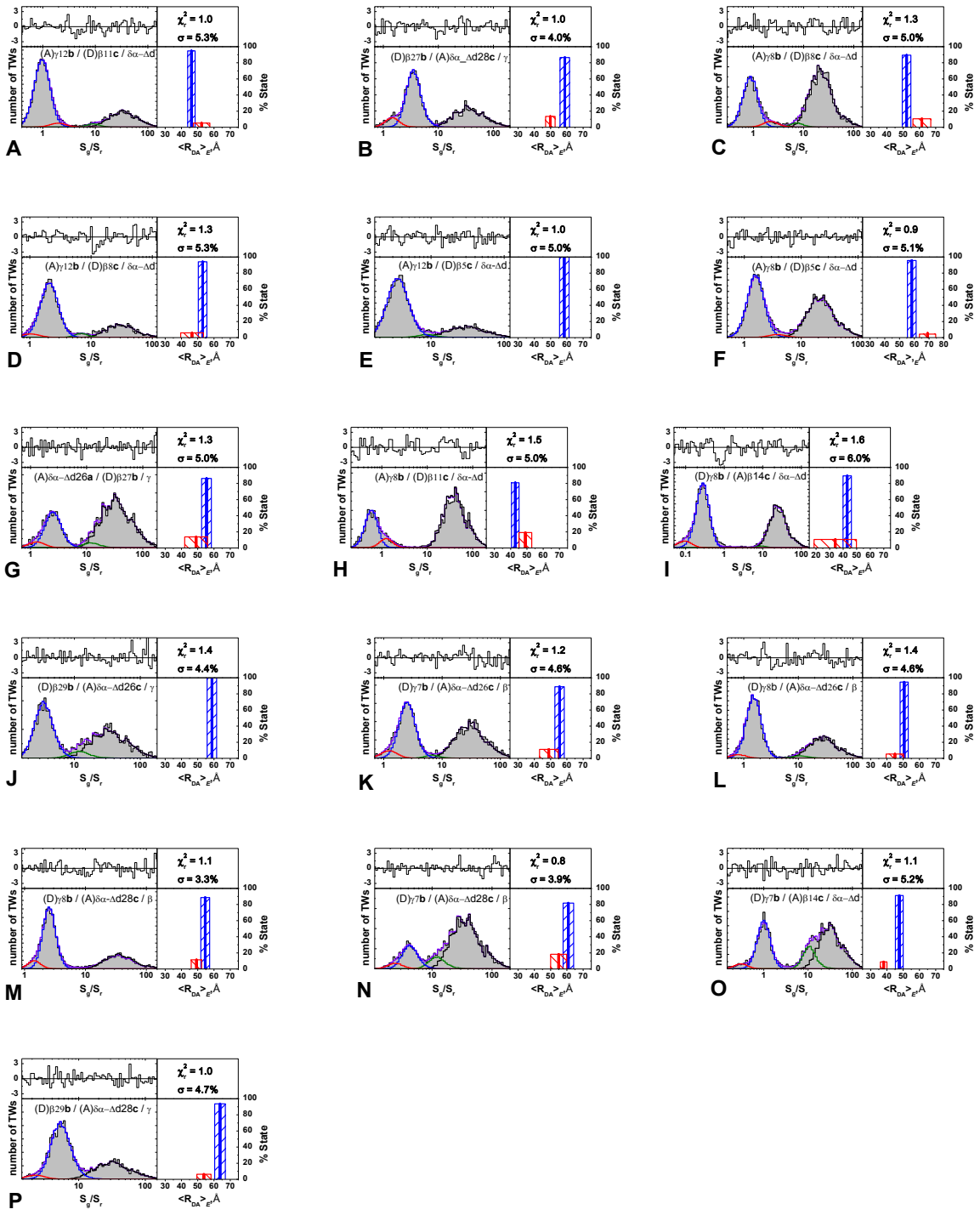


Figure 7-12 PDA analysis for the samples with fluorescence labels on the helices *b* and *c*. S_g/S_r histograms of experimental data (gray area) are fitted (violet solid line) with two FRET states ($\langle R_{DA}\rangle_{E(1)}$, blue solid line for major population and $\langle R_{DA}\rangle_{E(2)}$, red solid line for minor population, a D-only state (black solid line) and, if necessary, a state accounting for impurities (dark yellow solid line). Weighted residuals are shown in the upper plot. The right panel shows the distances, relative amplitudes (solid lines) and the confidence intervals (striped boxes) of the three FRET states (blue for major population, red for minor population). See Table 21 for all parameters used.

7.5.2 Measurements at 20 mM $MgCl_2$ for J(abc(C2))

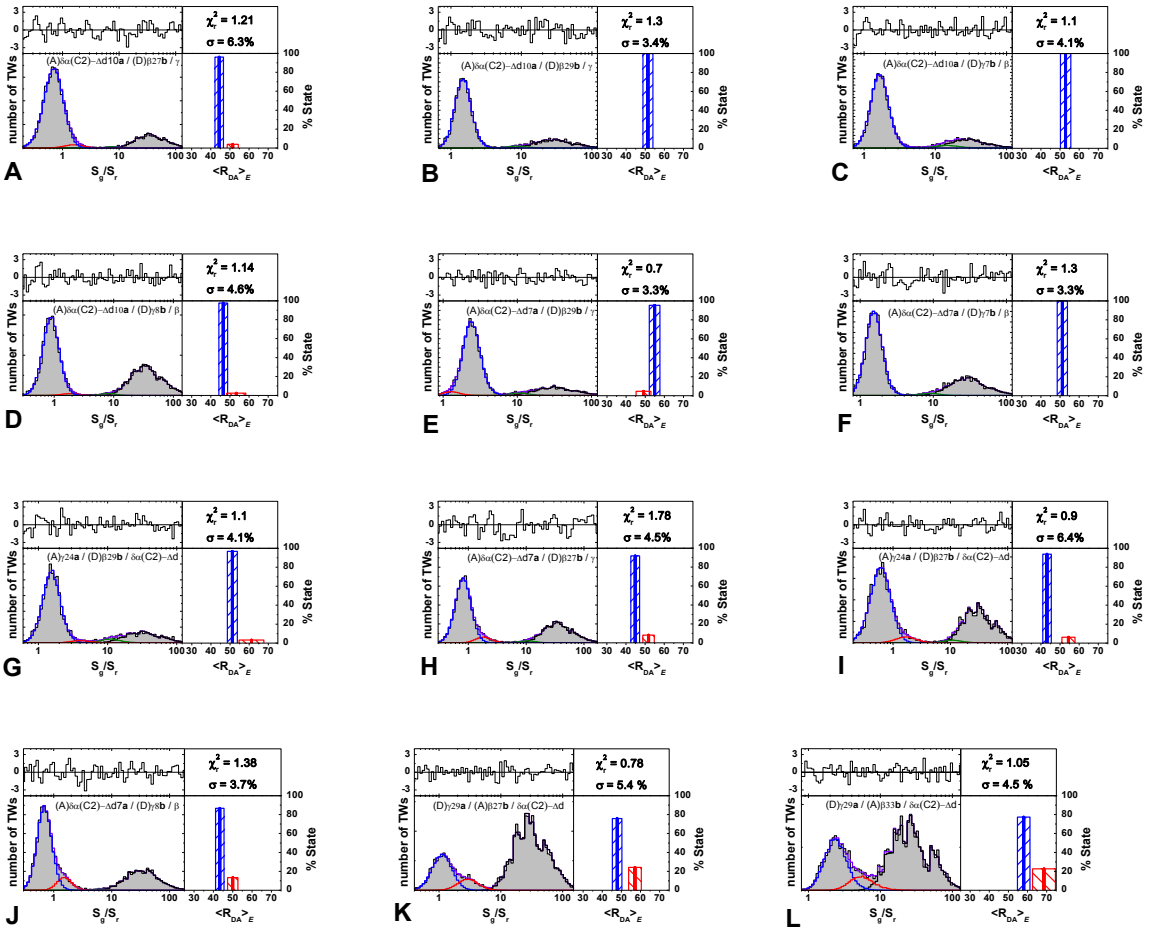


Figure 7-13 PDA analysis for the samples with fluorescence labels on the helices *a* and *b*. S_g/S_r histograms of experimental data (gray area) are fitted (violet solid line) with two FRET states ($\langle R_{DA} \rangle_{E(1)}$, blue solid line for major population and $\langle R_{DA} \rangle_{E(2)}$, red solid line for minor population, a D-only state (black solid line) and, if necessary, a state accounting for impurities (dark yellow solid line). Weighted residuals are shown in the upper plot. The right panel shows the distances, relative amplitudes (solid lines) and the confidence intervals (striped boxes) of the three FRET states (blue for major population, red for minor population). See Table 22 for all parameters used.

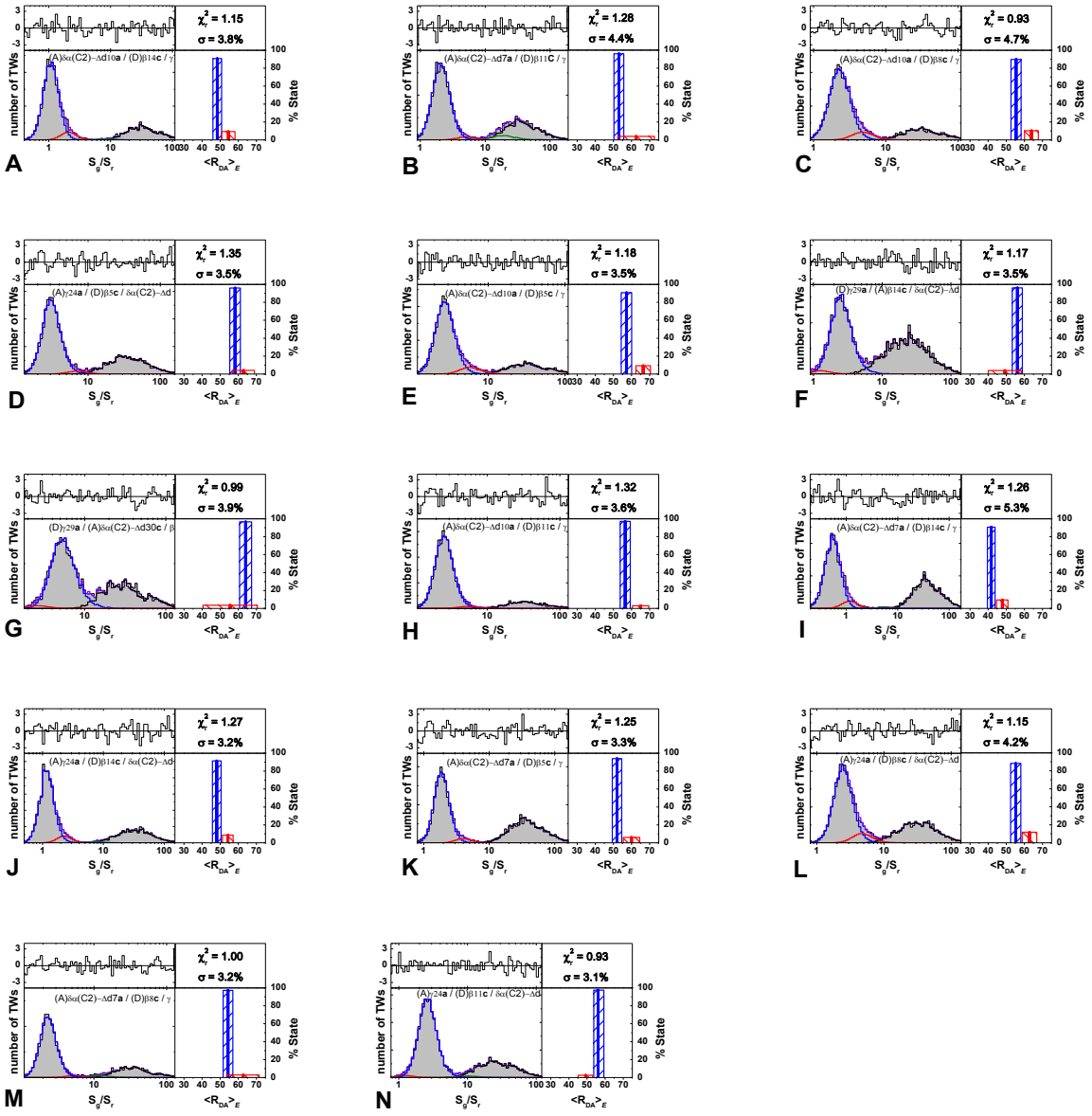


Figure 7-14 PDA analysis for the samples with fluorescence labels on the helices *a* and *c*. S_g/S_r histograms of experimental data (gray area) are fitted (violet solid line) with two FRET states ($\langle R_{DA}\rangle_{E(1)}$, blue solid line for major population and $\langle R_{DA}\rangle_{E(2)}$, red solid line for minor population, a D-only state (black solid line) and, if necessary, a state accounting for impurities (dark yellow solid line). Weighted residuals are shown in the upper plot. The right panel shows the distances, relative amplitudes (solid lines) and the confidence intervals (striped boxes) of the three FRET states (blue for major population, red for minor population). See Table 22 for all parameters used.

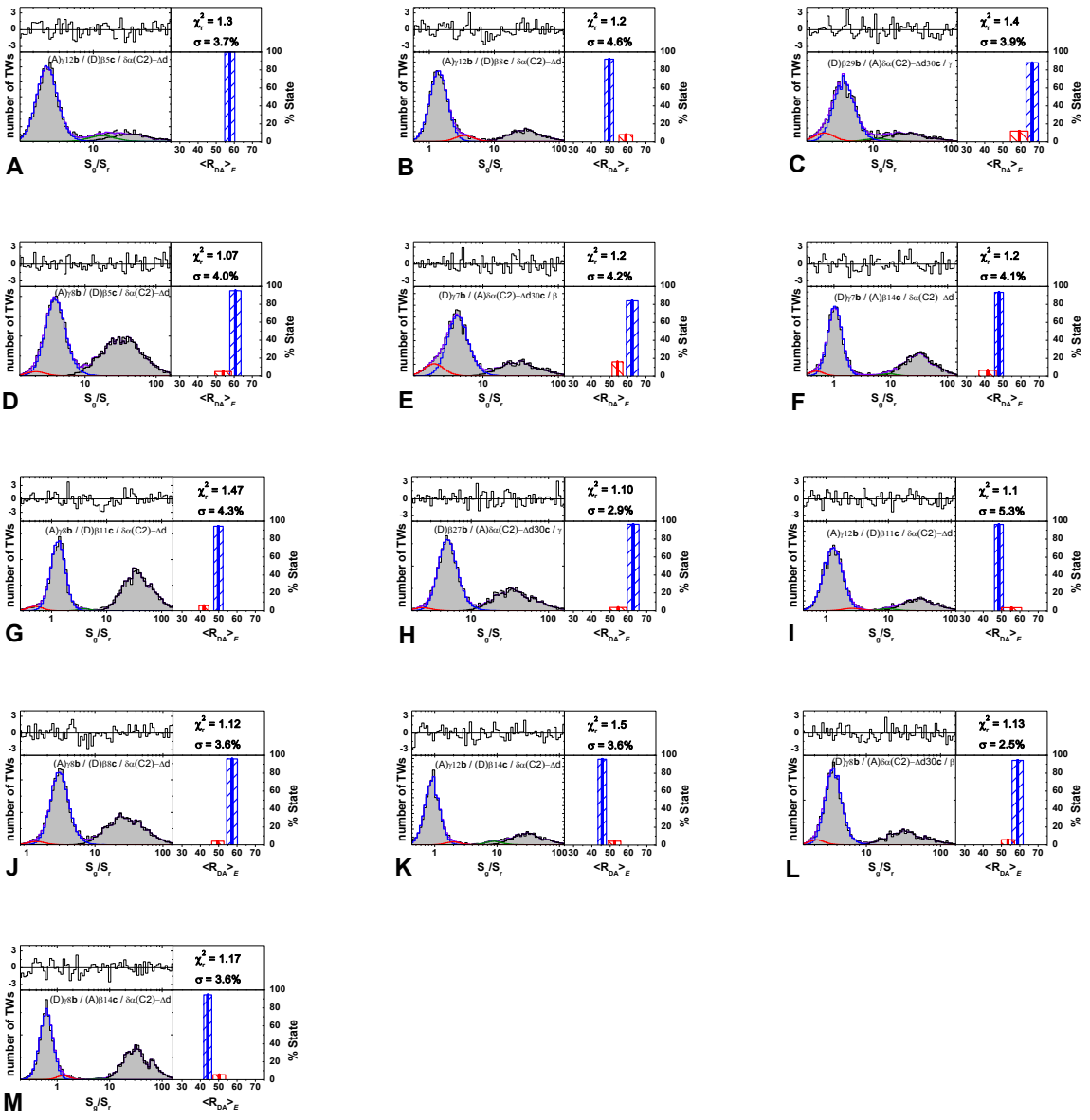


Figure 7-15 PDA analysis for the samples with fluorescence labels on the helices *b* and *c*. S_g/S_r histograms of experimental data (gray area) are fitted (violet solid line) with two FRET states ($\langle R_{DA}\rangle_{E(1)}$, blue solid line for major population and $\langle R_{DA}\rangle_{E(2)}$, red solid line for minor population, a D-only state (black solid line) and, if necessary, a state accounting for impurities (dark yellow solid line). Weighted residuals are shown in the upper plot. The right panel shows the distances, relative amplitudes (solid lines) and the confidence intervals (striped boxes) of the three FRET states (blue for major population, red for minor population). See Table 22 for all parameters used.

7.5.3 Measurements at 20 mM $MgCl_2$ for J(abc(C5))

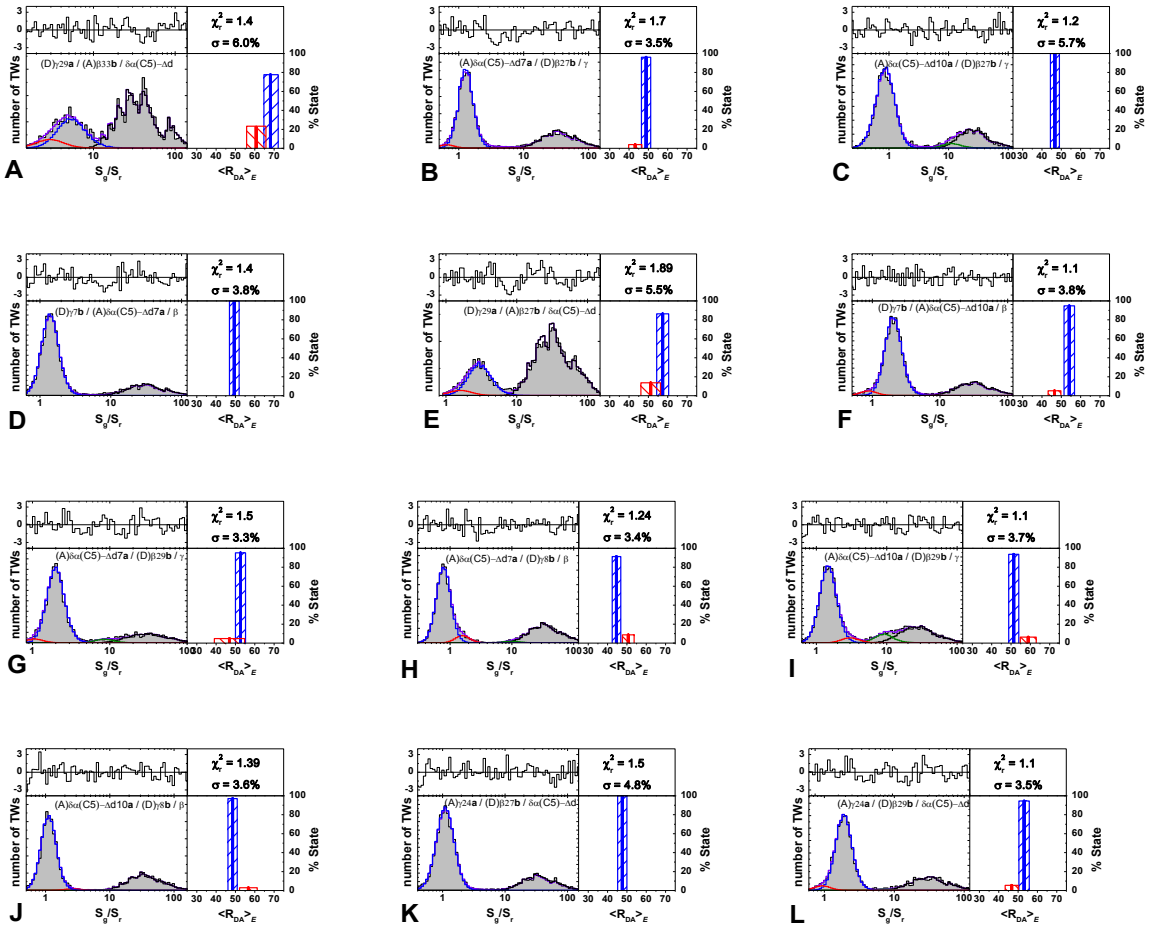


Figure 7-16 PDA analysis for the samples with fluorescence labels on the helices *a* and *b*. S_g/S_r histograms of experimental data (gray area) are fitted (violet solid line) with two FRET states ($\langle R_{DA}\rangle_{E(1)}$, blue solid line for major population and $\langle R_{DA}\rangle_{E(2)}$, red solid line for minor population, a D-only state (black solid line) and, if necessary, a state accounting for impurities (dark yellow solid line). Weighted residuals are shown in the upper plot. The right panel shows the distances, relative amplitudes (solid lines) and the confidence intervals (striped boxes) of the three FRET states (blue for major population, red for minor population). See Table 23 for all parameters used.

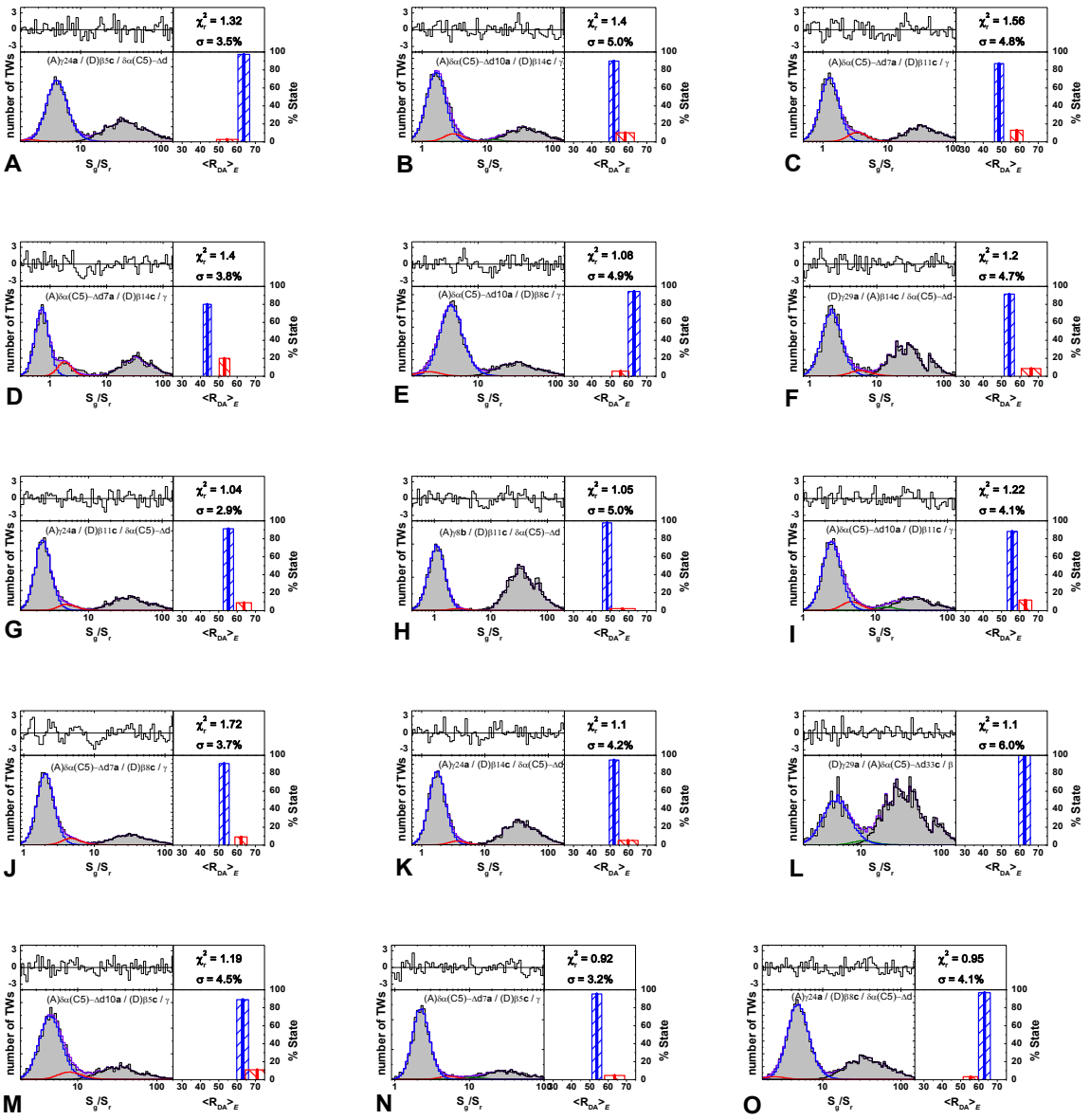


Figure 7-17 PDA analysis for the samples with fluorescence labels on the helices *a* and *c*. S_g/S_r histograms of experimental data (gray area) are fitted (violet solid line) with two FRET states ($\langle R_{DA}\rangle_{E(1)}$, blue solid line for major population and $\langle R_{DA}\rangle_{E(2)}$, red solid line for minor population, a D-only state (black solid line) and, if necessary, a state accounting for impurities (dark yellow solid line). Weighted residuals are shown in the upper plot. The right panel shows the distances, relative amplitudes (solid lines) and the confidence intervals (striped boxes) of the three FRET states (blue for major population, red for minor population). See Table 23 for all parameters used.

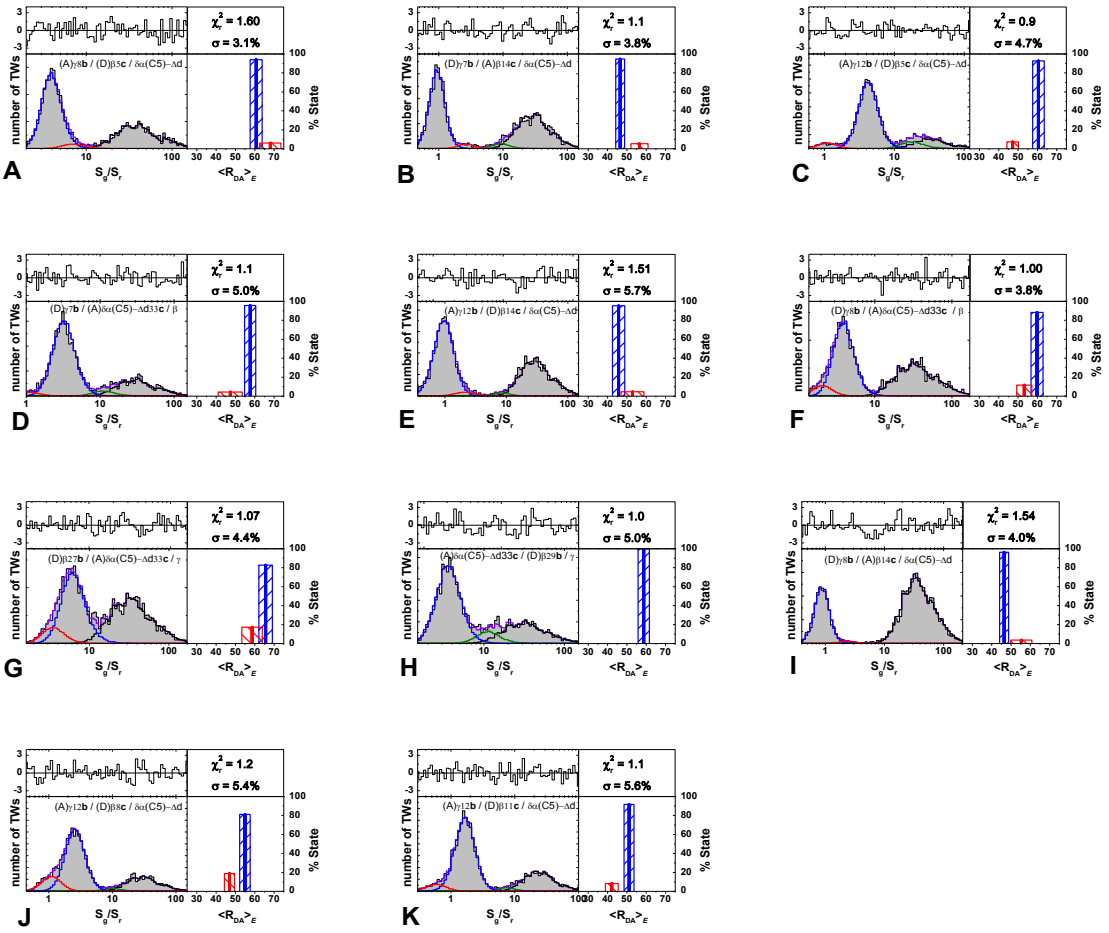


Figure 7-18 PDA analysis for the samples with fluorescence labels on the helices *b* and *c*. S_g/S_r histograms of experimental data (gray area) are fitted (violet solid line) with two FRET states ($\langle R_{DA} \rangle_{E(1)}$, blue solid line for major population and $\langle R_{DA} \rangle_{E(2)}$, red solid line for minor population, a D-only state (black solid line) and, if necessary, a state accounting for impurities (dark yellow solid line). Weighted residuals are shown in the upper plot. The right panel shows the distances, relative amplitudes (solid lines) and the confidence intervals (striped boxes) of the three FRET states (blue for major population, red for minor population). See Table 23 for all parameters used.

7.5.4 Measurements at 20 mM $MgCl_2$ for incomplete molecules

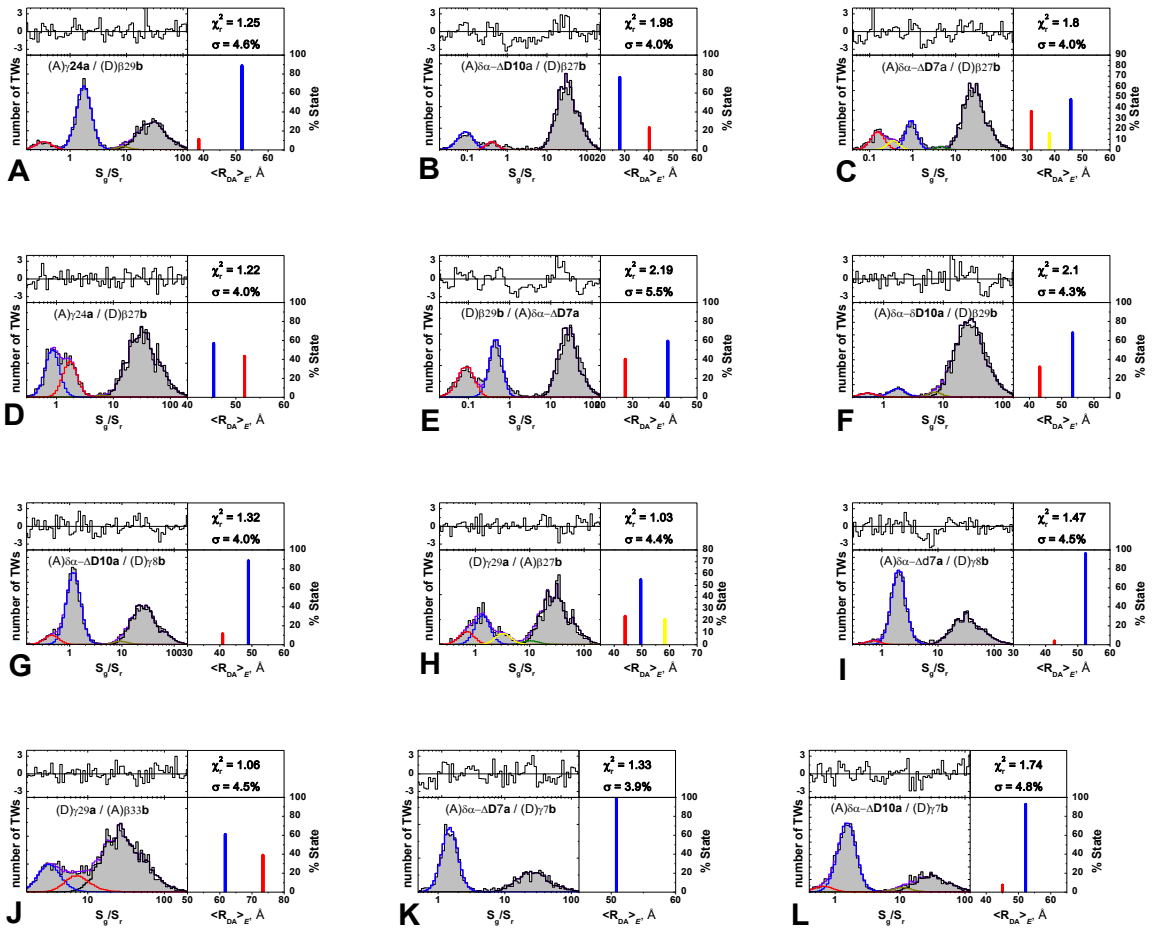


Figure 7-19 PDA analysis for the samples with fluorescence labels on the helices *a* and *b*. S_g/S_r histograms of experimental data (gray area) are fitted (violet solid line) with two or three FRET states ($\langle R_{DA}\rangle_{E(1)}$, blue solid line for major population, $\langle R_{DA}\rangle_{E(2)}$, red solid line for the first minor population and $\langle R_{DA}\rangle_{E(3)}$, yellow solid line for the second minor population a D-only state (black solid line) and, if necessary, a state accounting for impurities (dark yellow solid line). Weighted residuals are shown in the upper plot. The right panel shows the distances, relative amplitudes (solid lines) and the confidence intervals (striped boxes) of the three FRET states (blue for major population, red for minor population). See Table 34 for all parameters used.

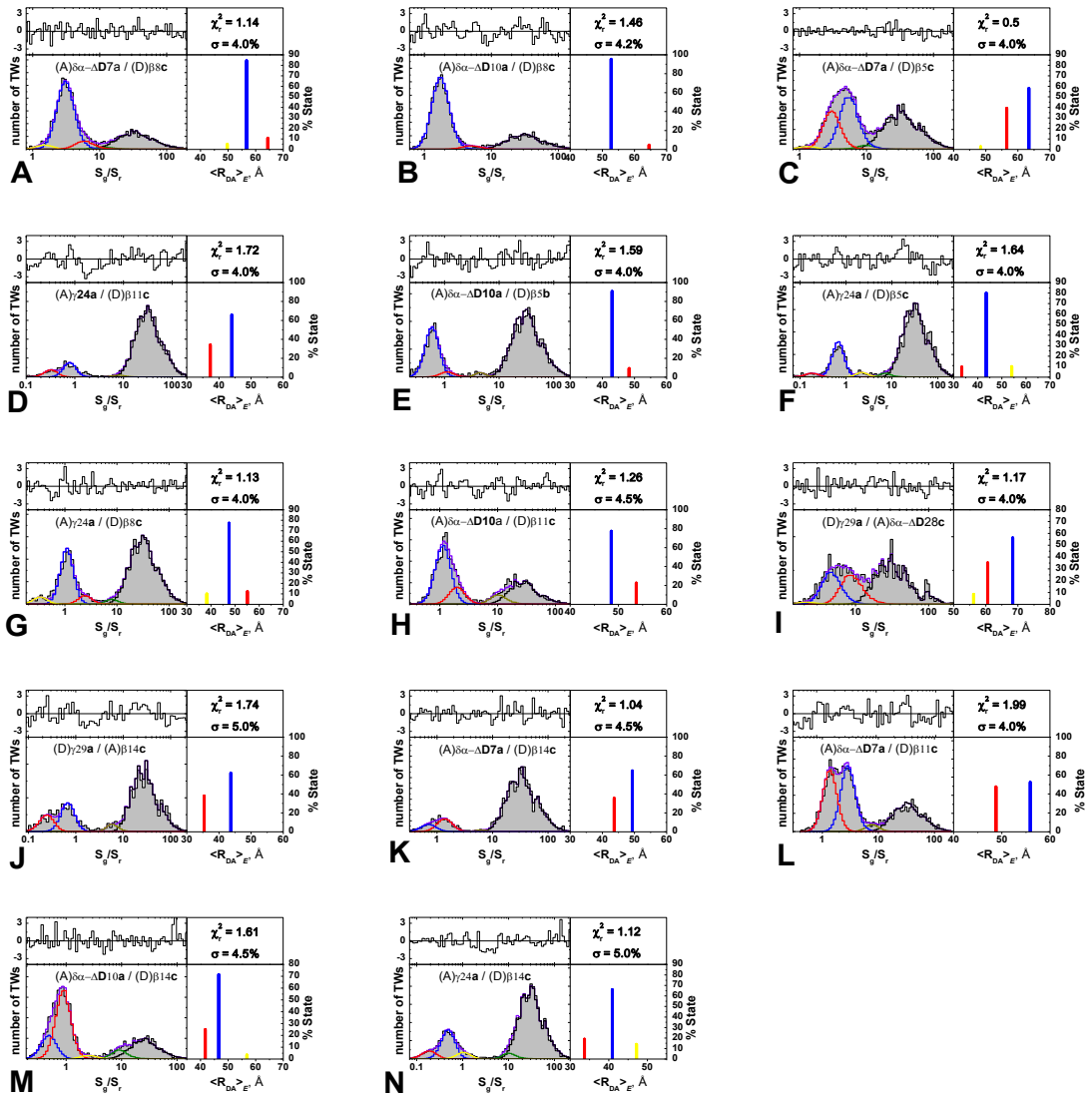


Figure 7-20 PDA analysis for the samples with fluorescence labels on the helices *a* and *c*. S_g/S_r histograms of experimental data (gray area) are fitted (violet solid line) with two or three FRET states ($\langle R_{DA}\rangle_{E(1)}$, blue solid line for major population, $\langle R_{DA}\rangle_{E(2)}$, red solid line for the first minor population and $\langle R_{DA}\rangle_{E(3)}$, yellow solid line for the second minor population a D-only state (black solid line) and, if necessary, a state accounting for impurities (dark yellow solid line). Weighted residuals are shown in the upper plot. The right panel shows the distances, relative amplitudes (solid lines) and the confidence intervals (striped boxes) of the three FRET states (blue for major population, red for minor population). See Table 34 for all parameters used.

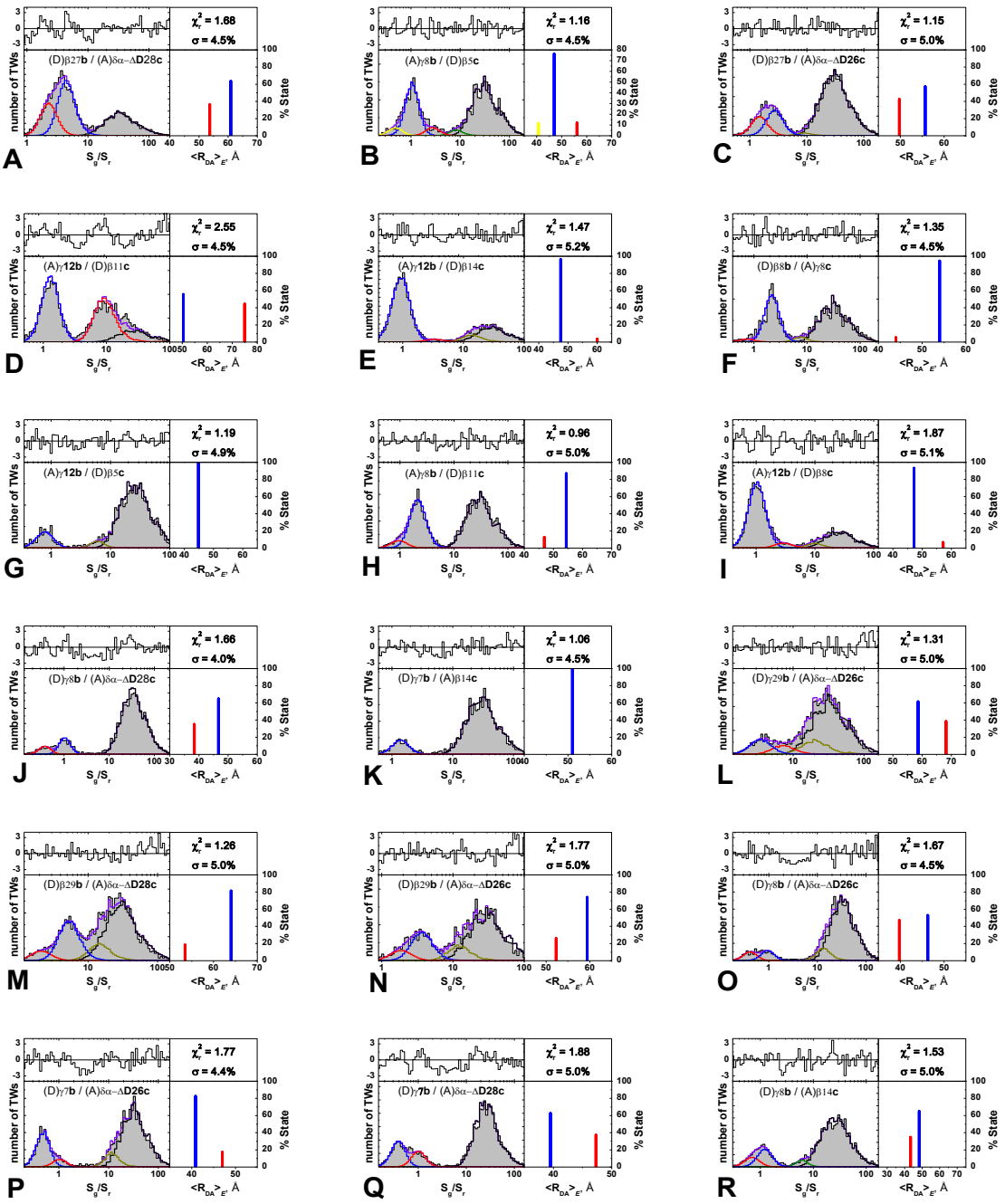


Figure 7-21 PDA analysis for the samples with fluorescence labels on the helices *b* and *c*. S_g/S_r histograms of experimental data (gray area) are fitted (violet solid line) with two or three FRET states ($\langle R_{DA}\rangle_{E(1)}$, blue solid line for major population, $\langle R_{DA}\rangle_{E(2)}$, red solid line for the first minor population and $\langle R_{DA}\rangle_{E(3)}$, yellow solid line for the second minor population a D-only state (black solid line) and, if necessary, a state accounting for impurities (dark yellow solid line). Weighted residuals are shown in the upper plot. The right panel shows the distances, relative amplitudes (solid lines) and the confidence intervals (striped boxes) of the three FRET states (blue for major population, red for minor population). See Table 34 for all parameters used.

7.6 eTCSPC fluorescence decay measurements of single labeled RNA three-way junctions

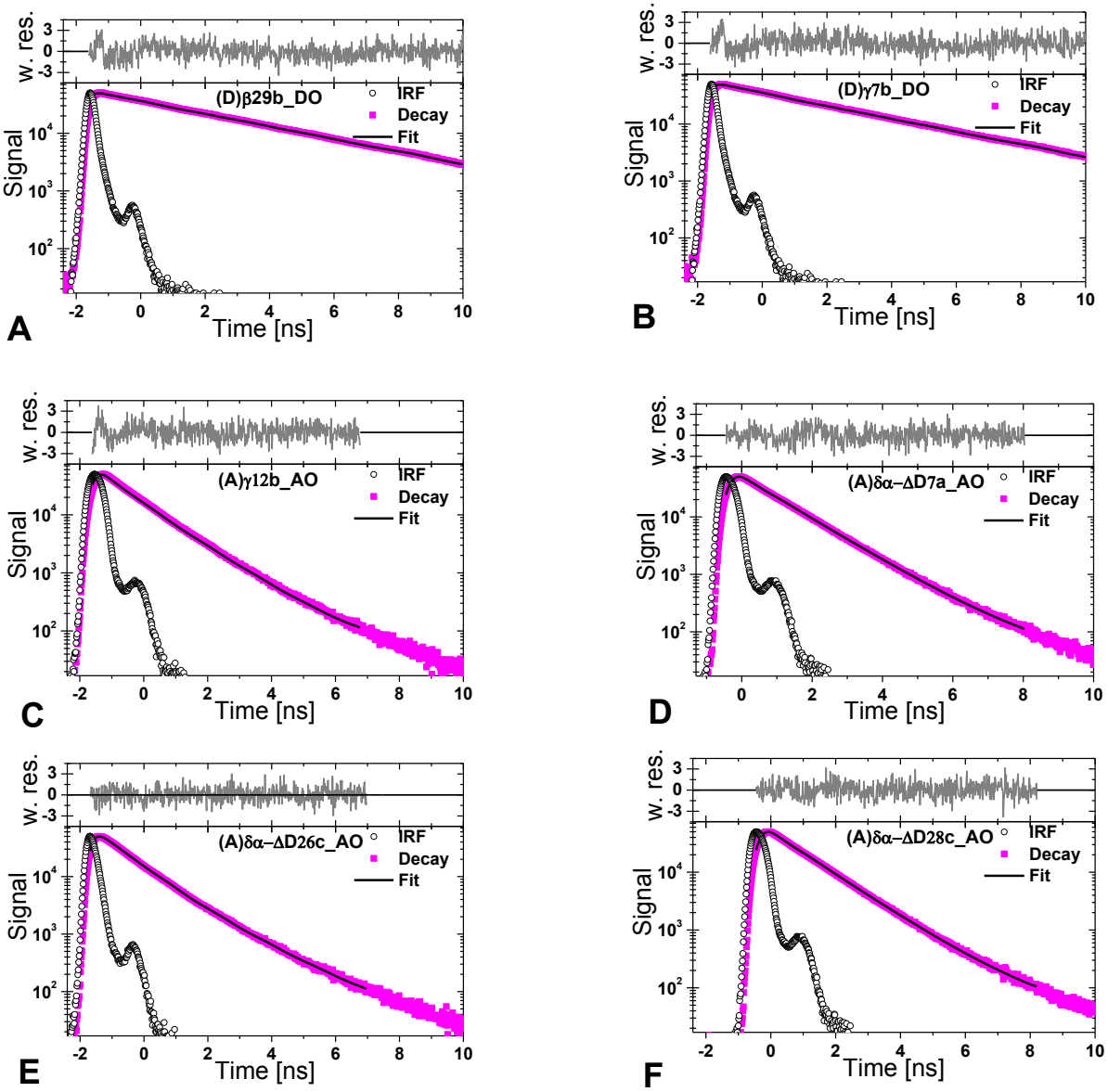


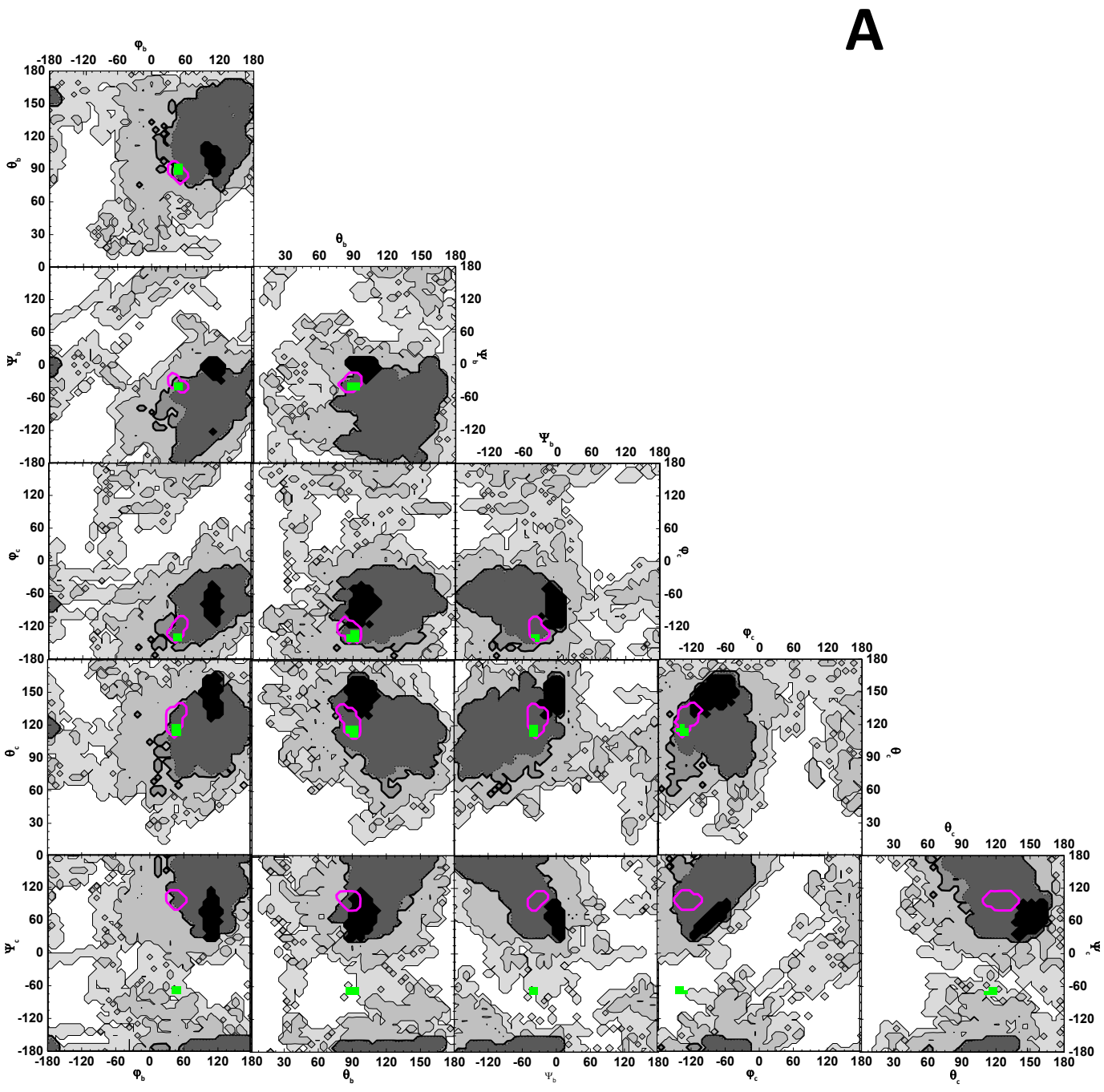
Figure 7-22 eTCSPC measurements of fluorescence decay for RNA three-way junctions single-labeled with Alexa488 (A and B) and Cy5 (C - F). Experimental data (purple filled squares), instrument response function (IRF, black open circles) and fits to the data (black solid lines) are shown. Weighted residuals are presented above each plot (gray solid lines). See Table 3 for all fit parameters.

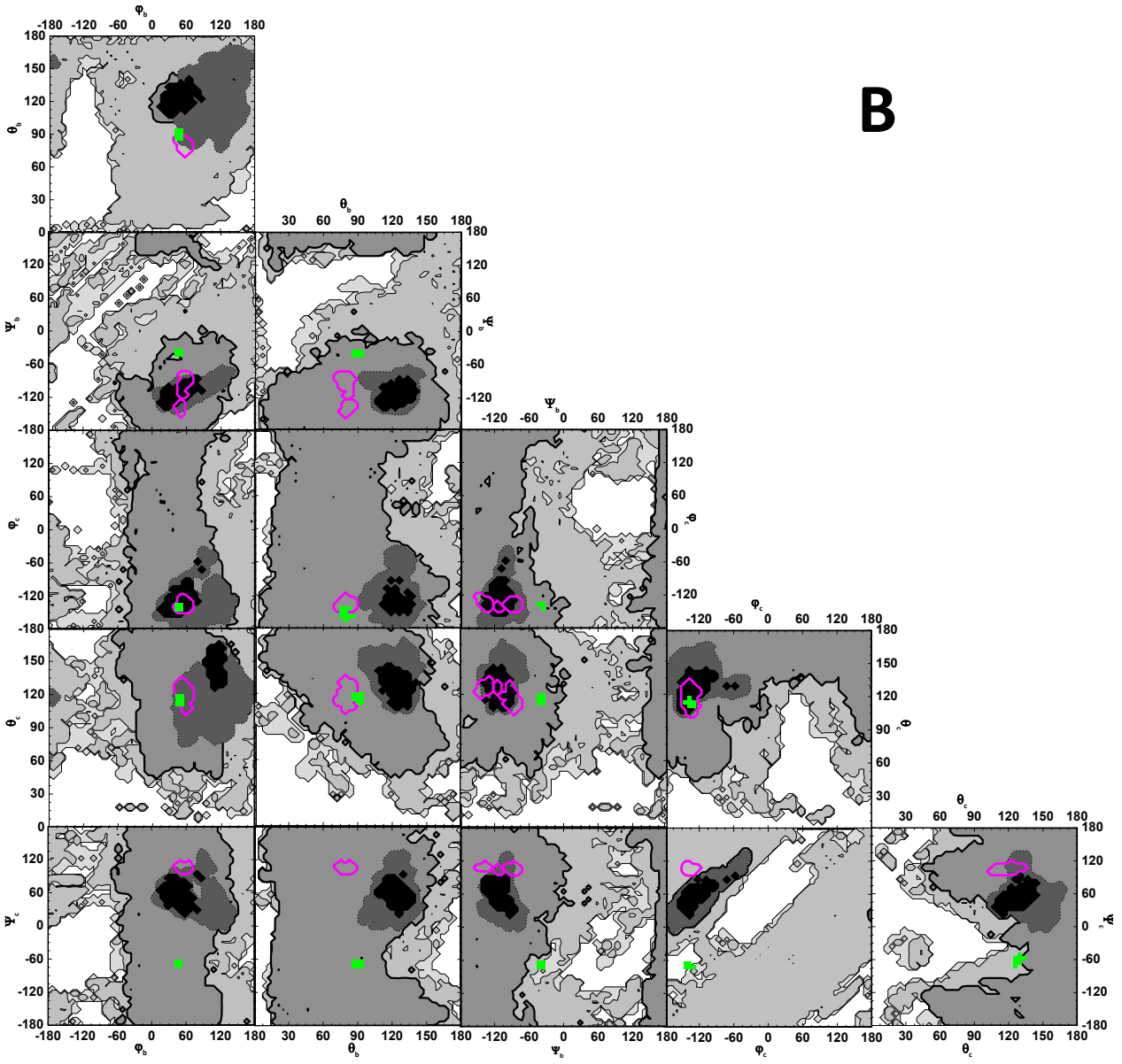
7.7 **DA-pairs**

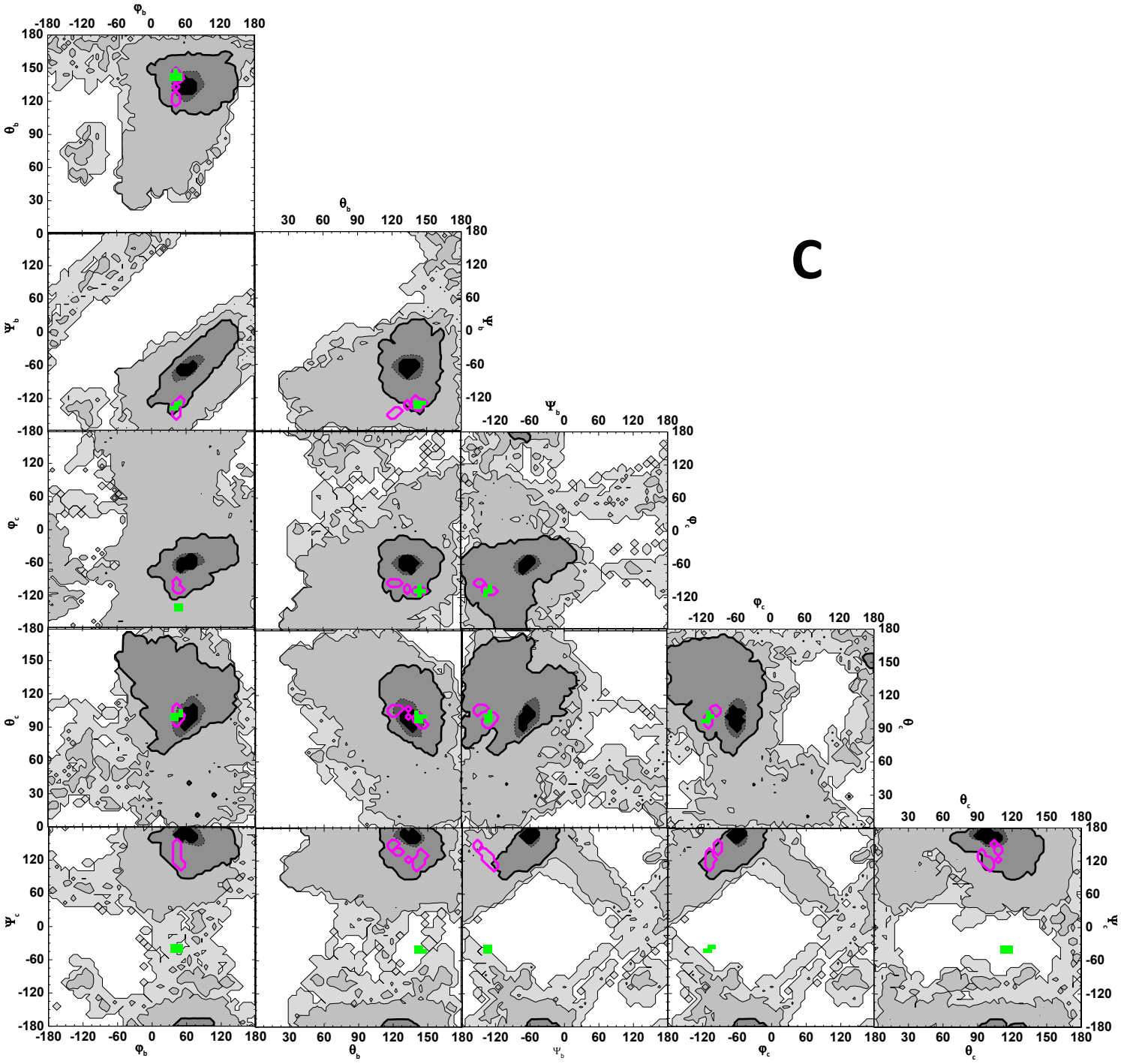
Table 41 List of common samples for three-and four way junction molecules.

#	DA-pair
1	(D)ϒ29a_(A)β14c
2	(D)ϒ8b_(A)β14c
3	(D)ϒ29a_(A)β27b
4	(D)ϒ29a_(A)β33b
5	(D)β11c_(A)ϒ24a
6	(D)β14c_(A)ϒ24a
7	(D)β27b_(A)ϒ24a
8	(D)β5c_(A)ϒ24a
9	(D)β8c_(A)ϒ24a
10	(D)β11c_(A)ϒ8b
11	(D)β5c_(A)ϒ8b
12	(D)β8c_(A)ϒ8b
13	(D)β11c_(A)δ10a
14	(D)β14c_(A)δ10a
15	(D)β27b_(A)δ10a
16	(D)β5c_(A)δ10a
17	(D)β8c_(A)δ10a
18	(D)ϒ8b_(A)δ10a
19	(D)δ7a_(A)β27b
20	(D)δ7a_(A)ϒ8b

7.8 Simulation data represented in Euler angles




B


C
φb
θb
Ψb
φc
θc
Ψc

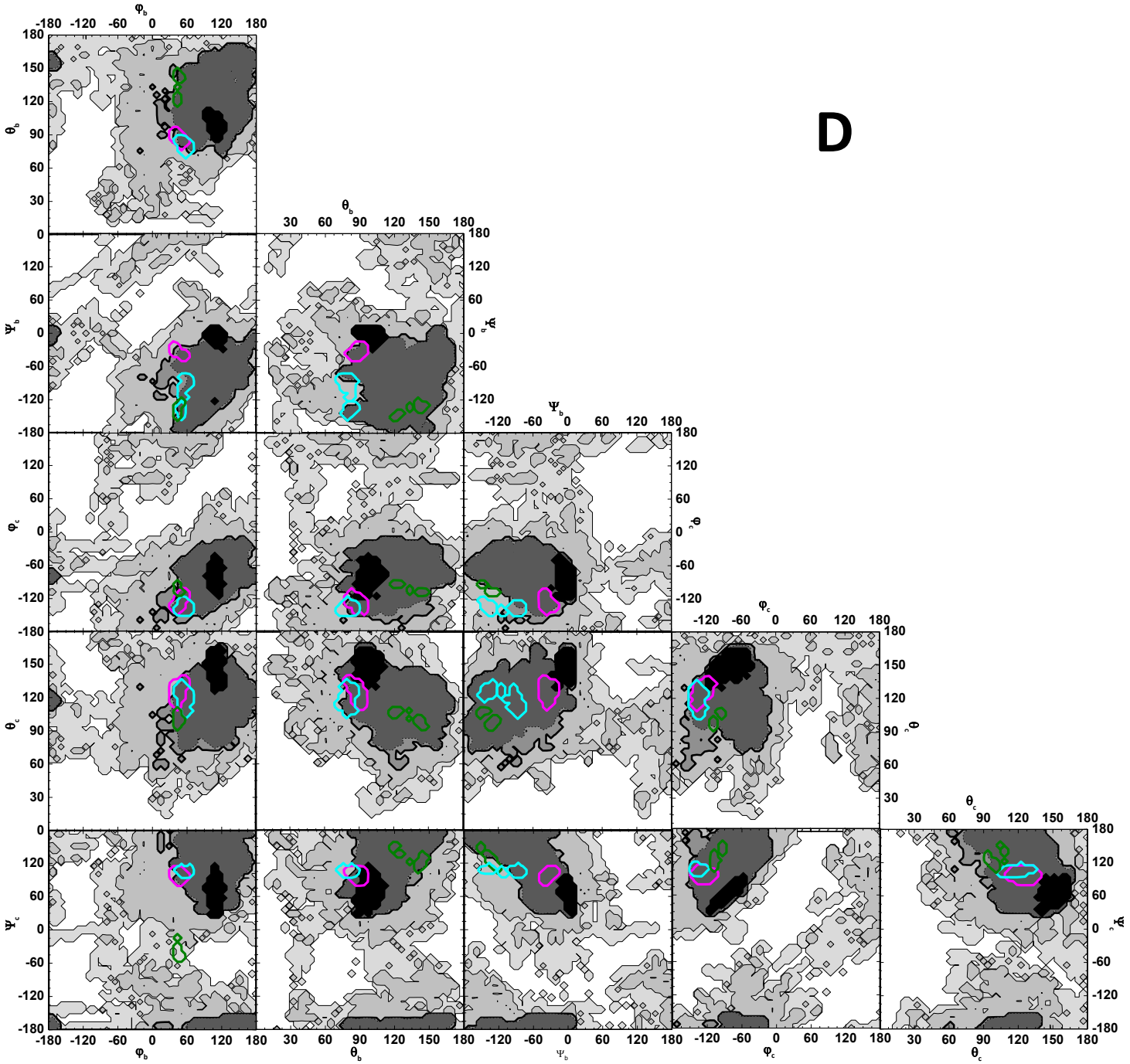
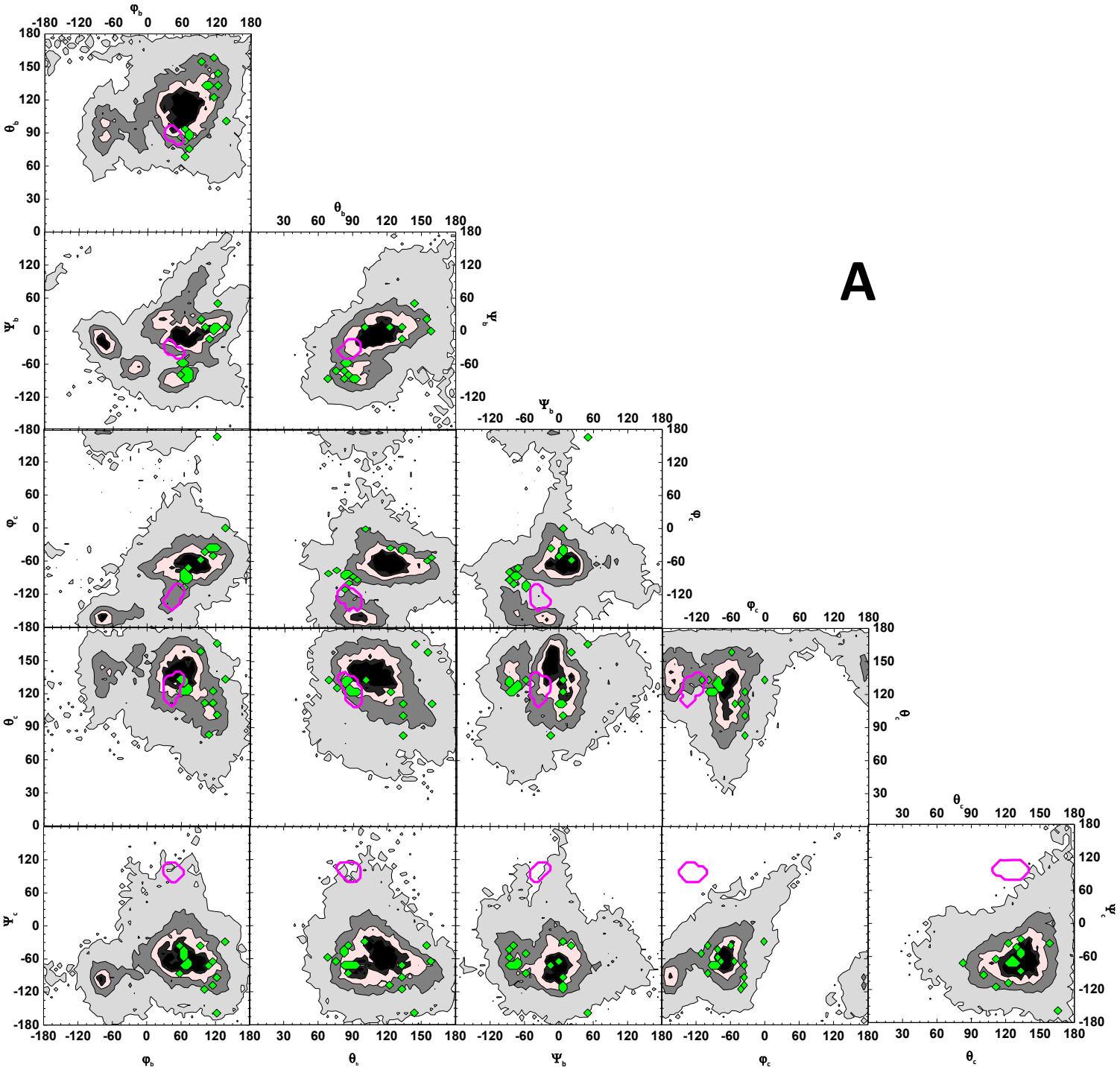
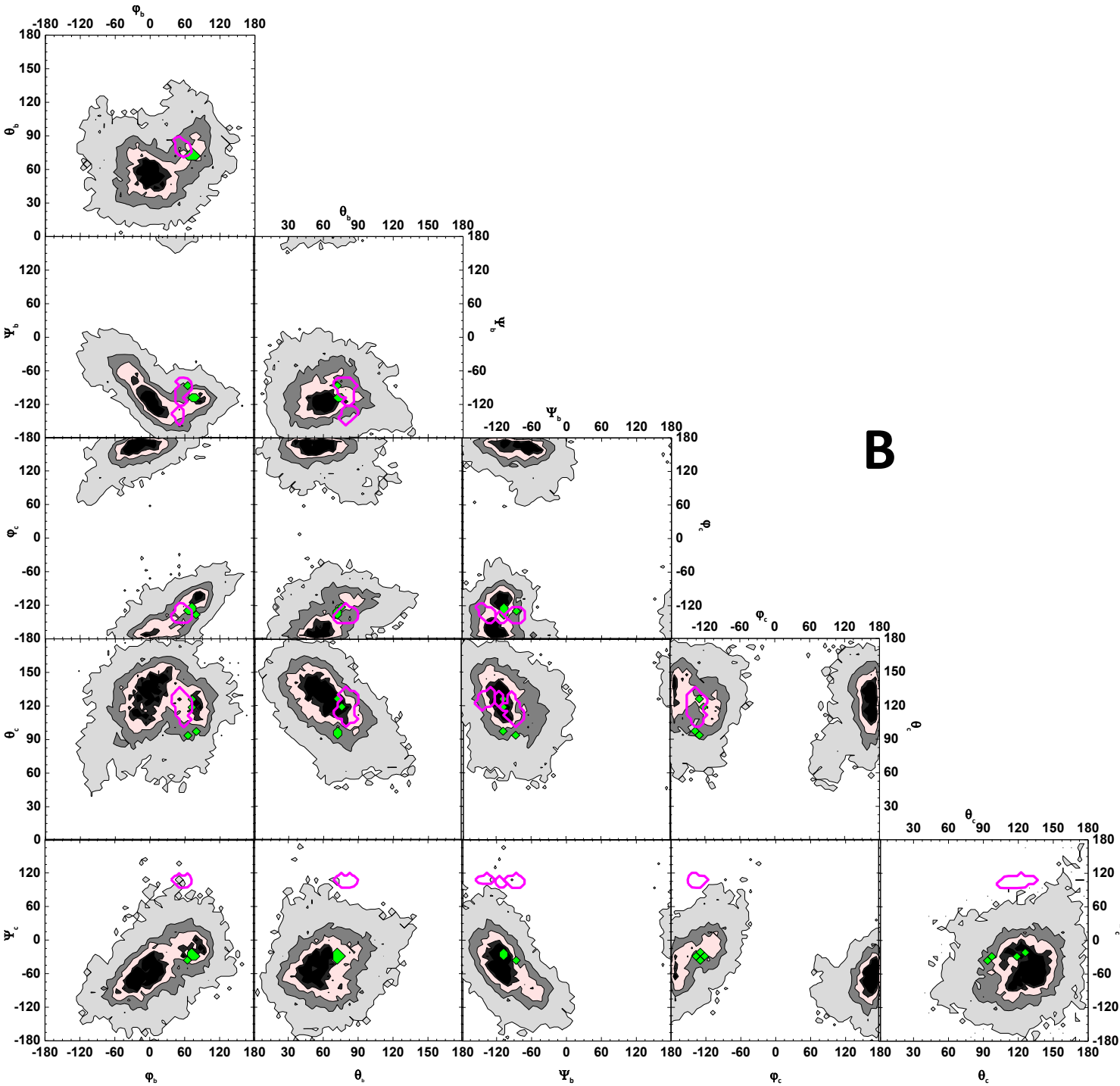

D
φb
θb
Ψb
φc
θc
Ψc

Figure 7-23 Euler angles calculated for the structures of free MMC simulations (gray scale) together with MMC "best" family structures according to FRET (green) and the contour of the area with the "best" family structures after docking (magenta) are shown on (A), (B), (C) for J(abc), J(abc(C2)) and J(abc(C5)), respectively. Overlay of the Euler angles calculated for all three molecules is shown on (D). Magenta, cyan, green and grey correspond to RBD results for J(abc), J(abc(C2)), J(abc(C5)) and free MMC simulations for J(abc), respectively.




B
φb
θb
Ψb
φc
θc
Ψc

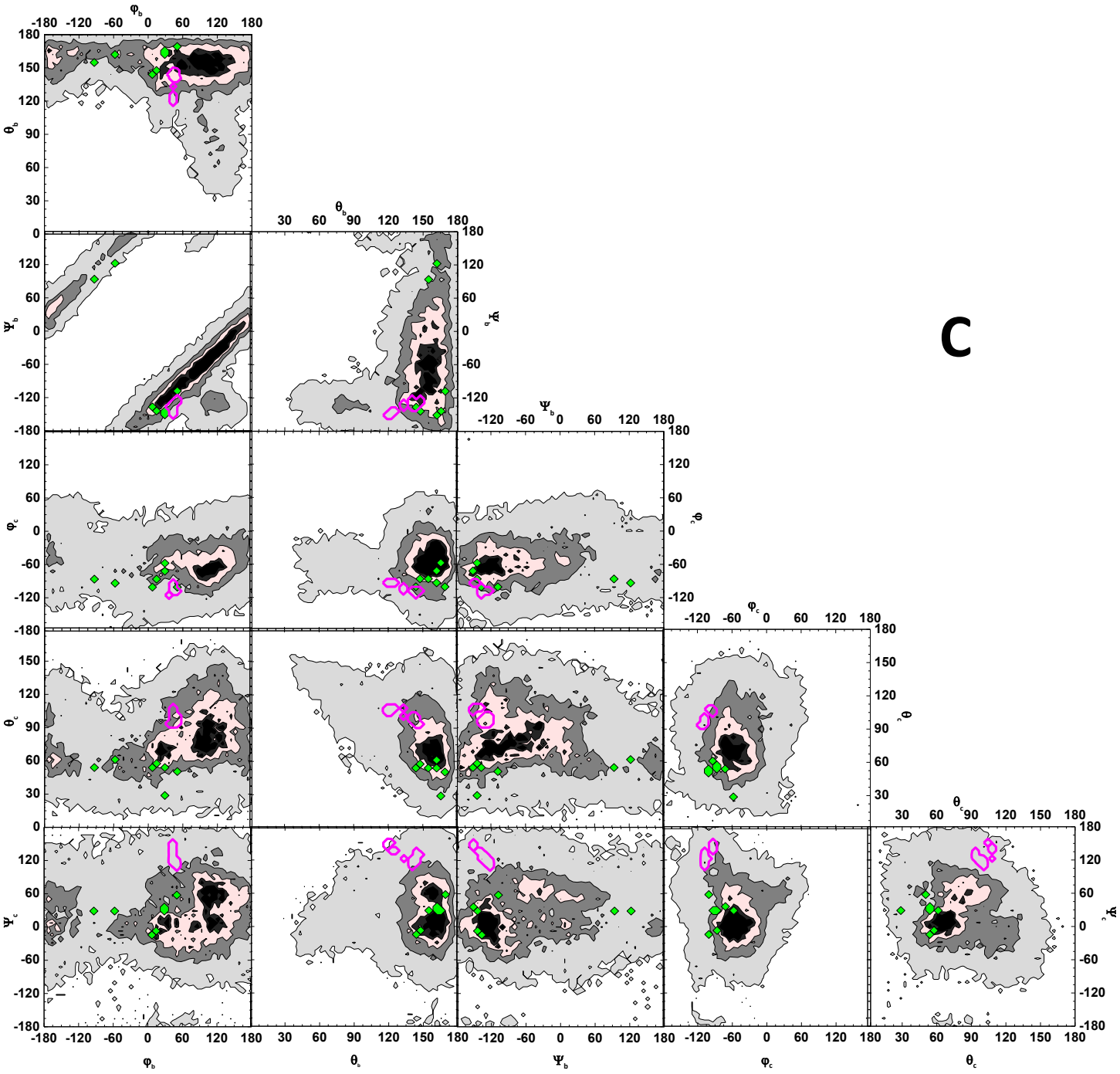


Figure 7-24 Euler angles calculated for the structures of free MD simulations (gray scale) together with MD "best" family structures according to FRET (green) and the contour of the area with the "best" family structures after docking are shown on A, B, C for J(abc), J(abc(C2)) and J(abc(C5)), respectively.

8 References

1. Petrasek, Z. and P. Schwille, *Fluctuations as a source of information in fluorescence microscopy.* Journal of the Royal Society Interface, 2009. **6**: p. S15-S25.
2. Al-Soufi, W., et al., *Dynamics of supramolecular association monitored by fluorescence correlation spectroscopy.* Chemphyschem, 2008. **9**(13): p. 1819-1827.
3. Amaral, P.P., et al., *The eukaryotic genome as an RNA machine.* Science, 2008. **319**(5871): p. 1787-1789.
4. Walter, N.G., S.A. Woodson, and R.T. Batey, *Non-protein coding RNAs*2009, Heidelberg: Springer.
5. Lescoute, A. and E. Westhof, *Topology of three-way junctions in folded RNAs.* RNA, 2006. **12**(1): p. 83-93.
6. Lilley, D.M.J., *Structures of helical junctions in nucleic acids.* Quarterly Reviews of Biophysics, 2000. **33**(2): p. 109-159.
7. Förster, T., *Zwischenmolekulare Energiewanderung und Fluoreszenz.* Annalen der Physik, 1948. **437**(2): p. 55-75.
8. Stryer, L., *Fluorescence energy transfer as a spectroscopic ruler.* Annual Review of Biochemistry, 1978. **47**: p. 819-846.
9. Kalinin, S.P., T.; Sindbert, S.; Rothwell, P. J.; Berger, S.; Restle, T.; Goody, R. S.; Gohlke, H.; Seidel, C. A. M., *A toolkit and benchmark study for FRET-restrained high-precision structural modeling.* Nature Methods, 2012. **9**: p. 1218-1225.
10. Lamiable, A., et al., *Automated prediction of three-way junction topological families in RNA secondary structures.* Computational Biology and Chemistry, 2012. **37**: p. 1-5.
11. Moazed, D., *Small RNAs in transcriptional gene silencing and genome defence.* Nature, 2009. **457**(7228): p. 413-420.
12. Mercer, T.R., M.E. Dinger, and J.S. Mattick, *Long non-coding RNAs: insights into functions.* Nature Reviews Genetics, 2009. **10**(3): p. 155-159.
13. Kortmann, J. and F. Narberhaus, *Bacterial RNA thermometers: molecular zippers and switches.* Nature Reviews Microbiology, 2012. **10**(4): p. 255-265.
14. Dethoff, E.A., et al., *Visualizing transient low-populated structures of RNA.* Nature, 2012. **491**(7426): p. 724-28.
15. Reining, A., et al., *Three-state mechanism couples ligand and temperature sensing in riboswitches.* Nature, 2013. **499**(7458): p. 355-9.
16. Laing, C. and T. Schlick, *Analysis of four-way junctions in RNA structures.* Journal of Molecular Biology, 2009. **390**(3): p. 547-59.
17. Al-Hashimi, H.M. and N.G. Walter, *RNA dynamics: it is about time.* Current Opinion in Structural Biology, 2008. **18**(3): p. 321-329.
18. Lilley, D.M.J., *Folding of branched RNA species.* Biopolymers, 1998. **48**(2-3): p. 101-112.
19. Hirschfeld, T., *Quantum efficiency independence of the time integrated emission from a fluorescent molecule.* Applied Optics, 1976. **15**(12): p. 3135-3139.
20. Moerner, W.E. and L. Kador, *Optical Detection and Spectroscopy of Single Molecules in a Solid.* Physical Review Letters, 1989. **62**(21): p. 2535-2538.
21. Moerner, W.E. and D.P. Fromm, *Methods of single-molecule fluorescence spectroscopy and microscopy.* Review of Scientific Instruments, 2003. **74**(8): p. 3597-3619.
22. Kühnemuth, R. and C.A.M. Seidel, *Principles of Single Molecule Multiparameter Fluorescence Spectroscopy.* Single Molecules, 2001. **2**(4): p. 251-254.

23. Gonçalves, M.S.T., *Fluorescent Labeling of Biomolecules with Organic Probes.* Chemical Reviews, 2009. **109**(1): p. 190-212.
24. Orrit, M. and J. Bernard, *Single Pentacene Molecules Detected by Fluorescence Excitation in a p-Terphenyl Crystal.* Physical Review Letters, 1990. **65**(21): p. 2716-2719.
25. Shera, E.B., et al., *Detection of single fluorescent molecules.* Chemical Physics Letters, 1990. **174**(6): p. 553-557.
26. Ha, T., et al., *Probing the interaction between two single molecules: Fluorescence resonance energy transfer between a single donor and a single acceptor.* Proc.Natl.Acad.Sci.USA., 1996. **93**: p. 6264-6268.
27. Deniz, A.A., et al., *Single-pair fluorescence resonance energy transfer on freely diffusing molecules: Observation of Förster distance dependence and subpopulations.* Proc.Natl.Acad.Sci.USA., 1999. **96**(7): p. 3670-3675.
28. Weiss, S., *Fluorescence spectroscopy of single biomolecules.* Science, 1999. **283**(5408): p. 1676-1683.
29. Margittai, M., et al., *Single-molecule fluorescence resonance energy transfer reveals a dynamic equilibrium between closed and open conformations of syntaxin 1.* Proc.Natl.Acad.Sci.USA., 2003. **100**: p. 15516-15521.
30. Borgia, M.B., et al., *Single-molecule fluorescence reveals sequence-specific misfolding in multidomain proteins.* Nature, 2011. **474**(7353): p. 662-U142.
31. Sakon, J.J. and K.R. Weninger, *Detecting the conformation of individual proteins in live cells.* Nature Methods, 2010. **7**(3): p. 203-U56.
32. Henzler-Wildman, K. and D. Kern, *Dynamic personalities of proteins.* Nature, 2007. **450**(7172): p. 964-972.
33. Tokuriki, N. and D.S. Tawfik, *Protein Dynamism and Evolvability.* Science, 2009. **324**(5924): p. 203-207.
34. Magde, D., E.L. Elson, and W.W. Webb, *Thermodynamic fluctuations in a reacting system - measurement by fluorescence correlation spectroscopy.* Physical Review Letters, 1972. **29**: p. 705-708.
35. Magde, D., E.L. Elson, and W.W. Webb, *Fluorescence Correlation Spectroscopy. II. An Experimental Realization.* Biopolymers, 1974. **13**: p. 29-61.
36. Mets, Ü. and R. Rigler, *Submillisecond Detection of Single Rhodamine Molecules in Water.* Journal of Fluorescence, 1994. **4**(3): p. 259-264.
37. Elson, E.L. and D. Magde, *Fluorescence Correlation Spectroscopy. I. Conceptual Basis and Theory.* Biopolymers, 1974. **13**: p. 1-27.
38. Palmer III, A.G. and N.L. Thompson, *Theory of sample translation in fluorescence correlation spectroscopy.* Biophysical Journal, 1987. **51**: p. 339-343.
39. Widengren, J. and Ü. Mets, *Conceptual Basis of Fluorescence Correlation Spectroscopy and Related Techniques as Tools in Bioscience* in *Single Molecule Detection in Solution: Methods and Applications*, C. Zander, J. Enderlein, and R.A. Keller, Editors. 2002. p. 69-120.
40. Lakowicz, J.R., *Principles of Fluorescence Spectroscopy*1999, New York: Kluwer Academic/ Plenum Publishers.
41. Kasha, M., *Characterization of electronic transitions in complex molecules.* Discussion of the Faraday Society, 1950. **9**: p. 14.
42. Andrews, D.L., *A Unified Theory of Radiative and Radiationless Molecular-Energy Transfer.* Chemical Physics, 1989. **135**(2): p. 195-201.
43. Andrews, D.L. and D.S. Bradshaw, *Virtual photons, dipole fields and energy transfer: a quantum electrodynamical approach.* European Journal of Physics, 2004. **25**(6): p. 845-858.

44. Braslavsky, S.E., et al., *Pitfalls and limitations in the practical use of Förster's theory of resonance energy transfer.* Photochemical & Photobiological Sciences, 2008. **7**(12): p. 1444-1448.
45. van der Meer, B.W., G. Cooker, and S.Y. Chen, *Resonance Energy Transfer: Theory and Data*1994, New York: VCH Publishers.
46. Kalinin, S., et al., *On the origin of broadening of single-molecule FRET efficiency distributions beyond shot noise limits.* Journal of Physical Chemistry B, 2010. **114**(18): p. 6197-206.
47. Schwille, P., F.J. Meyer-Almes, and R. Rigler, *Dual-color fluorescence cross-correlation spectroscopy for multicomponent diffusional analysis in solution.* Biophysical Journal, 1997. **72**(4): p. 1878-86.
48. Branden, M., et al., *Localized proton microcircuits at the biological membrane-water interface.* Proc Natl Acad Sci U S A, 2006. **103**(52): p. 19766-70.
49. Nettels, D., et al., *Ultrafast dynamics of protein collapse from single-molecule photon statistics.* Proc Natl Acad Sci U S A, 2007. **104**(8): p. 2655-60.
50. Eggeling, C., et al., *Direct observation of the nanoscale dynamics of membrane lipids in a living cell.* Nature, 2009. **457**(7233): p. 1159-62.
51. Bacia, K., S.A. Kim, and P. Schwille, *Fluorescence cross-correlation spectroscopy in living cells.* Nat Methods, 2006. **3**(2): p. 83-9.
52. O'Connor, D.V. and D. Phillips, *Time-correlated Single Photon Counting*1984, New York: Academic Press.
53. Eggeling, C., et al., *Data registration and selective single-molecule analysis using multi-parameter fluorescence detection.* Journal of Biotechnology, 2001. **86**(3): p. 163-80.
54. Eggeling, C., et al., *Analysis of Photobleaching in Single-molecule Multicolor Excitation and Förster Resonance Energy Transfer Measurements.* Journal of Physical Chemistry A, 2006. **110**(9): p. 2979-2995.
55. Rothwell, P.J., et al., *Multi-parameter single-molecule fluorescence spectroscopy reveals heterogeneity of HIV-1 Reverse Transcriptase:primer/template complexes.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**: p. 1655-1660.
56. Sisamakis, E., et al., *Accurate single-molecule FRET studies using multiparameter fluorescence detection.* Methods in Enzymology, 2010. **475**: p. 456-514.
57. Magatti, D. and F. Ferri, *25 ns software correlator for photon and fluorescence correlation spectroscopy.* Review of Scientific Instruments, 2003. **74**(2): p. 1135-1144.
58. Wahl, M., et al., *Fast calculation of fluorescence correlation data with asynchronous time-correlated single-photon counting.* Optics Express, 2003. **11**(26): p. 3583-3591.
59. Felekyan, S., et al., *Full correlation from picoseconds to seconds by time-resolved and time-correlated single photon detection.* Review of Scientific Instruments, 2005. **76**(8): p. 083104.
60. Laurence, T.A., S. Fore, and T. Huser, *Fast, flexible algorithm for calculating photon correlations.* Optics Letters, 2006. **31**(6): p. 829-831.
61. Mets, Ü., *Antibunching and rotational diffusion in FCS.*, in *Fluorescence Correlation Spectroscopy*, R. Rigler and E.L. Elson, Editors. 2001, Springer-Verlag: Berlin Heidelberg New York. p. 346-359.
62. Widengren, J., Ü. Mets, and R. Rigler, *Fluorescence Correlation Spectroscopy of Triplet States in Solution: A Theoretical and Experimental Study.* Journal of Physical Chemistry, 1995. **99**: p. 13368-13379.

63. Widengren, J., et al., *Strategies to improve photostabilities in ultrasensitive fluorescence spectroscopy.* Journal of Physical Chemistry A, 2007. **111**(3): p. 429-440.
64. Rupert, P.B., et al., *Transition state stabilization by a catalytic RNA.* Science, 2002. **298**(5597): p. 1421-1424.
65. Rasnik, I., S.A. McKinney, and T. Ha, *Nonblinking and longlasting single-molecule fluorescence imaging.* Nature Methods, 2006. **3**(11): p. 891-893.
66. Lipfert, J., D. Herschlag, and S. Doniach, *Riboswitch Conformations Revealed by Small-Angle X-Ray Scattering*, in *Riboswitches: Methods and Protocols*, A. Serganov, Editor 2009. p. 141-159.
67. Cai, Q., et al., *Nanometer distance measurements in RNA using site-directed spin Labeling.* Biophysical Journal, 2007. **93**(6): p. 2110-2117.
68. Muschielok, A., et al., *A nano-positioning system for macromolecular structural analysis.* Nature Methods, 2008. **5**(11): p. 965-971.
69. Sindbert, S., et al., *Accurate Distance Determination of Nucleic Acids via Forster Resonance Energy Transfer: Implications of Dye Linker Length and Rigidity.* Journal of the American Chemical Society, 2011. **133**(8): p. 2463-2480.
70. Rothwell, J.P., *Structural Investigation of the HIV-1 RT using single pair Fluorescence Energy Transfer*, in *Max-Planck-Institut für Molekulare Physiologie*2002, University of Dortmund: Dortmund.
71. Kalinin, S., et al., *Characterizing multiple molecular states in single-molecule multiparameter fluorescence detection by probability distribution analysis.* Journal of Physical Chemistry B, 2008. **112**(28): p. 8361-8374.
72. Kalinin, S., et al., *Probability distribution analysis of single-molecule fluorescence anisotropy and resonance energy transfer.* Journal of Physical Chemistry B, 2007. **111**(34): p. 10253-10262.
73. Antonik, M., et al., *Separating structural heterogeneities from stochastic variations in fluorescence resonance energy transfer distributions via photon distribution analysis.* Journal of Physical Chemistry B, 2006. **110**(13): p. 6970-6978.
74. Kalinin, S., et al., *Detection of structural dynamics by FRET: A photon distribution and fluorescence lifetime analysis of systems with multiple states.* J. Phys. Chem. B, 2010. **114**: p. 7983-7995.
75. Soong, T.T., *Fundamentals of probability and statistics for engineers*2004, West Sussex, England: John Wiley & Sons.
76. Nicholas Metropolis, A.W.R., Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller, *Equation of state calculations by fast computing machines.* The Journal of Chemical Physics, 1953: p. 21(6):1087-1092.
77. Efron, B., *Jackknife, Bootstrap and Other Resampling Methods in Regression-Analysis - Discussion.* Annals of Statistics, 1986. **14**(4): p. 1301-1304.
78. Case, D.A., et al., *The Amber biomolecular simulation programs.* Journal of Computational Chemistry, 2005. **26**(16): p. 1668-1688.
79. Hornak, V., et al., *Comparison of multiple amber force fields and development of improved protein backbone parameters.* Proteins, 2006. **65**(3): p. 712-725.
80. Perez, A., et al., *Refinenement of the AMBER force field for nucleic acids: Improving the description of alpha/gamma conformers.* Biophysical Journal, 2007. **92**(11): p. 3817-3829.
81. Banas, P., et al., *Performance of Molecular Mechanics Force Fields for RNA Simulations: Stability of UUCG and GNRA Hairpins.* Journal of Chemical Theory and Computation, 2010. **6**(12): p. 3836-3849.

82. Joung, I.S. and T.E. Cheatham, *Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations.* Journal of Physical Chemistry B, 2008. **112**(30): p. 9020-9041.
83. Jorgensen, W.L., et al., *Comparison of Simple Potential Functions for Simulating Liquid Water.* Journal of Chemical Physics, 1983. **79**(2): p. 926-935.
84. Darden, T., D. York, and L. Pedersen, *Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems.* Journal of Chemical Physics, 1993. **98**(12): p. 10089-10092.
85. Ryckaert, J.P., G. Ciccotti, and H.J.C. Berendsen, *Numerical-Integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of N-Alkanes.* Journal of Computational Physics, 1977. **23**(3): p. 327-341.
86. D.A. Case, V.B., J.T. Berryman, R.M. Betz, Q. Cai, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke, H. Gohlke, A.W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T.S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K.M. Merz, F. Paesani, D.R. Roe, A. Roitberg, C. Sagui, R. Salomon-Ferrer, G. Seabra, C.L. Simmerling, W. Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu and P.A. Kollman. *AMBER 14*. 2014.
87. Rother, K., et al., *Template-based and template-free modeling of RNA 3D structure: Inspirations from protein structure modeling*, in *In RNA 3D structure analysis and prediction*, N.B. Leontis and E. Westhof, Editors. 2012, Springer-Verlag: Berlin.
88. Lu, X.J. and W.K. Olson, *3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures.* Nucleic Acids Res, 2003. **31**(17): p. 5108-21.
89. Aduri, R., et al., *AMBER Force Field Parameters for the Naturally Occurring Modified Nucleosides in RNA.* Journal of Chemical Theory and Computation, 2007. **3**(4): p. 1464-1475.
90. Smolsky, I.L., et al., *Biological small-angle x-ray scattering facility at the Stanford synchrotron radiation laboratory.* Journal of Applied Crystallography, 2007. **40**: p. S453-S458.
91. Guinier, A., *La diffraction des rayons X aux très petits angles: Application à l'étude de phénomènes ultramicroscopiques.* Ann. Phys. (Paris), 1939. **12**: p. 161-237.
92. Lipfert, J. and S. Doniach, *Small-angle X-ray scattering from RNA, proteins, and protein complexes*, in *Annual Review of Biophysics and Biomolecular Structure*2007, Annual Reviews: Palo Alto. p. 307-327.
93. Koch, M.H.J., P. Vachette, and D.I. Svergun, *Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution.* Quarterly Reviews of Biophysics, 2003. **36**(2): p. 147-227.
94. Svergun, D., C. Barberato, and M.H.J. Koch, *CRYSOL - A program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates.* Journal of Applied Crystallography, 1995. **28**: p. 768-773.
95. Schneidman-Duhovny, D., M. Hammel, and A. Sali, *FoXS: a web server for rapid computation and fitting of SAXS profiles.* Nucleic Acids Research, 2010. **38**: p. W540-W544.
96. Rambo, R.P. and J.A. Tainer, *Improving small-angle X-ray scattering data for structural analyses of the RNA world.* Rna-a Publication of the Rna Society, 2010. **16**(3): p. 638-646.
97. Ali, M., et al., *The Ligand-Free State of the TPP Riboswitch: A Partially Folded RNA Structure.* Journal of Molecular Biology, 2010. **396**(1): p. 153-165.
98. Hohng, S., et al., *Conformational flexibility of four-way junctions in RNA.* Journal of Molecular Biology, 2004. **336**(1): p. 69-79.

99. Woźniak, A.K., et al., *Single molecule FRET measures bends and kinks in DNA.* Proc.Natl.Acad.Sci.USA., 2008. **105**: p. 18337-18342.
100. Sindbert, S., *FRET restrained high-precision structural modeling of biomolecules*, in *Institut für Physikalische Chemie II*2012, Heinrich-Heine-Universität Düsseldorf.
101. Buck, J., et al., *Influence of ground-state structure and Mg(2+) binding on folding kinetics of the guanine-sensing riboswitch aptamer domain.* Nucleic Acids Research, 2011. **39**(22): p. 9768-9778.
102. Buck, J., et al., *Dissecting the influence of Mg(2+) on 3D architecture and ligand-binding of the guanine-sensing riboswitch aptamer domain.* Nucleic Acids Research, 2010. **38**(12): p. 4143-4153.
103. Walter, F., A.I.H. Murchie, and D.M.J. Lilley, *Folding of the four-way RNA junction of the hairpin ribozyme.* Biochemistry, 1998. **37**(50): p. 17629-17636.
104. Chojnowski, G., T. Walen, and J.M. Bujnicki, *RNA Bricks--a database of RNA 3D motifs and their interactions.* Nucleic Acids Res, 2014. **42**(Database issue): p. D123-31.
105. Popenda, M., et al., *RNA FRABASE 2.0: an advanced web-accessible database with the capacity to search the three-dimensional fragments within RNA structures.* BMC Bioinformatics, 2010. **11**: p. 231.
106. Laing, C. and T. Schlick, *Analysis of Four-way Junctions in RNA Structures.* Journal of Molecular Biology, 2009. **390**(3): p. 547-559.
107. Rother, M., et al., *ModeRNA: a tool for comparative modeling of RNA 3D structure.* Nucleic Acids Research, 2011. **39**(10): p. 4007-4022.
108. Bai, Y., et al., *Critical assessment of nucleic acid electrostatics via experimental and computational investigation of an unfolded state ensemble.* Journal of the American Chemical Society, 2008. **130**(37): p. 12334-12341.
109. Felekyan, S., et al., *Filtered FCS: species auto- and cross-correlation functions highlight binding and dynamics in biomolecules.* ChemPhysChem, 2012. **13**: p. 1036-1053.
110. Sabir, T., et al., *Branchpoint expansion in a fully complementary three-way DNA junction.* Journal of the American Chemical Society, 2012. **134**(14): p. 6280-5.
111. Kalinin, S., et al., *A toolkit and benchmark study for FRET-restrained high-precision structural modeling.* Nat Methods, 2012. **9**(12): p. 1218-25.
112. Stühmeier, F., et al., *Global structure of three-way DNA junctions with and without additional unpaired bases: a fluorescence resonance energy transfer analysis.* Biochemistry, 1997. **36**: p. 13530-13538.
113. Zuker, M., *Mfold web server for nucleic acid folding and hybridization prediction.* Nucleic Acids Research, 2003. **31**(13): p. 3406-15.
114. Haustein, E. and P. Schwille, *Fluorescence correlation spectroscopy: Novel variations of an established technique.* Annual Review of Biophysics and Biomolecular Structure, 2007. **36**: p. 151-169.
115. Shahzad, A. and G. Kohler, *Fluorescence Correlation Spectroscopy (FCS): A Promising Tool for Biological Research.* Applied Spectroscopy Reviews, 2011. **46**(2): p. 166-173.
116. Salzinger, S., et al., *Aggregation behavior of thermo-responsive poly(2-oxazoline)s at the cloud point investigated by FCS and SANS.* Colloid and Polymer Science, 2012. **290**(5): p. 385-400.
117. Yu, S.R., et al., *Fgf8 morphogen gradient forms by a source-sink mechanism with freely diffusing molecules.* Nature, 2009. **461**(7263): p. 533-U100.
118. Brunger, A.T., et al., *Three-dimensional molecular modeling with single molecule FRET.* Journal of Structural Biology, 2011. **173**(3): p. 497-505.

119. Cowieson, N.P., B. Kobe, and J.L. Martin, *United we stand: combining structural methods.* Current Opinion in Structural Biology, 2008. **18**(5): p. 617-22.
120. Boura, E., et al., *Solution structure of the ESCRT-I complex by small-angle X-ray scattering, EPR, and FRET spectroscopy.* Proceedings of the National Academy of Sciences of the United States of America, 2011. **108**(23): p. 9437-9442.
121. Rigler, R., et al., *Fluorescence correlation spectroscopy with high count rate and low background: analysis of translational diffusion.* European Biophysics Journal, 1993. **22**: p. 169-175.
122. Schwille, P., F.J. Meyer-Almes, and R. Rigler, *Dual-color fluorescence cross-correlation spectroscopy for multicomponent diffusional analysis in solution.* Biophysical Journal, 1997. **72**(4): p. 1878-1886.
123. Branden, M., et al., *Localized proton microcircuits at the biological membrane-water interface.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(52): p. 19766-19770.
124. Nettels, D., et al., *Ultrafast dynamics of protein collapse from single-molecule photon statistics.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(8): p. 2655-2660.
125. Eggeling, C., et al., *Direct observation of the nanoscale dynamics of membrane lipids in a living cell.* Nature, 2009. **457**(7233): p. 1159-1162.
126. Bacia, K., S.A. Kim, and P. Schwille, *Fluorescence cross-correlation spectroscopy in living cells.* Nature Methods, 2006. **3**(2): p. 83-89.
127. Mocsar, G., et al., *Note: Multiplexed multiple-tau auto- and cross-correlators on a single field programmable gate array.* Review of Scientific Instruments, 2012. **83**(4).
128. Schätzel, K., M. Drewel, and S. Stimac, *Photon Correlation Measurements at Large Lag Times: Improving Statistical Accuracy.* Journal of Modern Optics, 1988. **35**(4): p. 711-718.
129. Alfke, P. *Xilinx Application Note 052: Efficient Shift Registers, LFSR Counters, and Long Pseudo-Random Sequence Generators*. 1996; Available from: http://www.xilinx.com/support/documentation/application_notes/xapp052.pdf.
130. Al-Soufi, W., et al., *Fluorescence correlation spectroscopy, a tool to investigate supramolecular dynamics: Inclusion complexes of pyronines with cyclodextrin.* Journal of the American Chemical Society, 2005. **127**(24): p. 8775-8784.

9 Acknowledgements

I would like to thank all the people who supported and helped me during studies.

Particularly:

I thank **Prof. Dr. Claus Seidel** for the opportunity to work on this exciting project, to use the state of the art equipment in his lab and for teaching me to be more enthusiastic and creative in everything what I do. Also, I am thankful to him for high stress-tolerance, gained during the last few years.

I am especially greatful to **Dr. Stanislav Kalinin** for being a great advisor, teacher and colleague; for always having a correct and shortest solution to my scientific questions and for extensive help with my research and thesis even after leaving the group.

I am very thankful to **Dr. Suren Felekyan** for introducing me to the experimental part of my studies. I especially appreciated his ability to find useful information in all kind of measurement results, even when they did not make any sence to me. His constant support from the very first day of my arrival in Germany was invaluable for me.

Dr. Simon Sindbert for introducing me to the RNA community in our group. For the help in the lab, and for being a great colleague.

Sascha Fröbel for his significant contribution in this project; for being my best student and a very cool guy.

Dr. Oleg Opanasyuk for writing matlab scripts for me, for helpful discussions.

Mykola Dimura for computational part of my project, for correcting my thesis and for being a great guy to share the office with.

Olga Doroshenko for performing some measurements for me and for her positive aura.

Peter Zentis for reading and correcting my thesis, for being the best “Steuer berater” and a very nice person.

Dr. Ralf Kühnemuth and Thomas Peulen for fruitful discussions and always a valuable advices.

My second supervisor **Prof. Dr. Gerhard Steger.**

Prof. Dr. Holger Gohlke and **Christian Hanke** for MD simulations.

All my **colleagues**, for great working atmosphere.

A friend of mine, **Dr. Ara Ghazaryan** from TUM, for correcting the English grammar in my thesis.

I especially thankful to the best family in the world: **my family**. My parents and my sister, they have managed to be so close to me being over 4000 km away from Düsseldorf.

Finally I want to thank my girlfriend, **Isabel Dörnberger**, for always believing in me, for her patience, and for being my best friend.

Hiermit erkläre ich, dass ich die Promotion mit dem Thema

“FRET restrained high-precision structural modeling of RNA junctions”

am Institut für Physikalische Chemie II der Heinrich-Heine- Universität Düsseldorf unter der Leitung von Prof. Dr. Claus A.M. Seidel eigenständig und ohne unerlaubte Hilfe angefertigt und in der vorgelegten oder in ähnlicher Form noch bei keiner anderen Institution eingereicht habe.

Es existieren keine vorherigen Promotionsversuche.

Düsseldorf, den 03.07.2014

Hayk Vardanyan