STUDY OF WATER DYNAMICS IN THE DNA-DAUNOMYCIN INTERCALATION PATHWAY

A thesis submitted towards partial fulfillment of

5 year integrated BS-MS programme

by

Hutashan Vajpeyi

Under the guidance of

Dr. Arnab Mukherjee

Department of Chemistry

Indian Institute of Science Education and Research, Pune.

Certificate

This is to certify that this dissertation entitled "Study of water dynamics in the DNA-Daunomycin intercalation pathway " towards the partial fulfillment of the 5 year integrated BS-MS programme at the Indian Institute of Science Education and Research Pune, represents original research carried out by Hutashan Vajpeyi at IISER under the supervision of Dr. Arnab Mukherjee

Supervisor:	**Head Chemistry:**
Date:	**Date:**
Place:	**Place:**

Acknowledgements:

Firstly, I would like to acknowledge my guide Dr. Arnab Mukherjee for being very patient (patience being the greatest asset of a teacher) with a student like me, who had no prior experience of computational chemistry and for being one of the very few teachers I have seen, who love laughing. I owe whatever little knowledge I have of computational chemistry to him.

Secondly, I would like to acknowledge the Ph.D students (especially Wilbee DS and Anurag Sunda) working in my laboratory (if I can, after working for nearly an year, call this lab my own), who helped me in my work.

Thirdly, I would like to acknowledge IISER for providing me access to many scientific journals, which have proved invaluable in my work.

Fourthly, I would like to acknowledge my fellow students of HR-1, who more than anyone else, made me believe that I belonged in this college.

Introduction

Water is very important for the functioning of many biological macromolecules. It is extremely important for the structure of DNA, this is illustrated by the fact that the DNA duplex is surrounded by at least two hydration layers. Understanding water's role in the intercalation process involved in the complexation of daunomycin (daunorubicin)[1,2] can provide insight into the way the drug functions. Figure 1 shows the chemical structure of daunomycin.

A large amount of work has been done in trying to understand the intercalation[3,4,5]of daunomycin in DNA[6,7], owing to its ability to act as a chemotherapeutic drug. It is an antibiotic of the anthracycline family. It has a planar aromatic ring system and an amino sugar moiety. It is used in the treatment of cancer[8]. Its mode of action is intercalation into DNA by inserting its planar rings between two successive base pairs[9,10,11,12], thus causing the inhibition of DNA replication[13,14]. It is the planar, heteroaromatic moiety which slides between the DNA base pairs during intercalation, binding along the minor groove. In the present project WEhave studied the water dynamics involved in the intercalation process in an attempt to understand the role of water in the intercalation process, and the effect of intercalation on the dynamics of water molecules in proximity of the drug-DNA system.

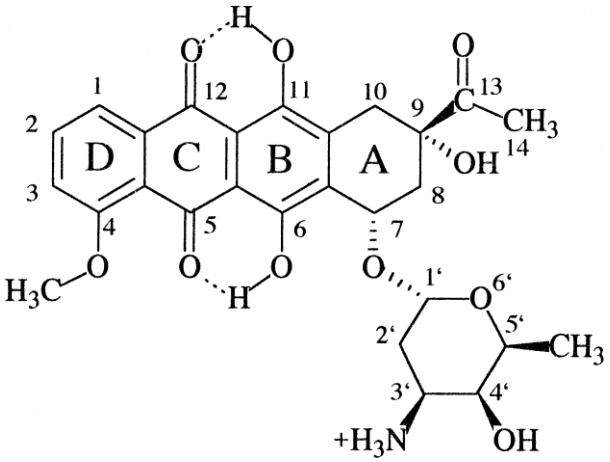

Figure 1: Daunomycin (formula $C_{27}H_{29}NO_{10}$)

Intercalation involves the bending of the DNA, opening of the rise and roll angle, widening of the minor groove etc[15], so it causes distortion in the shape of the strand. Water plays an

important role in the DNA molecule's complexation with daunomycin. To understand water's role in the intercalation process one must study water dynamics in the salvation layer around the drug-DNA system.

It has been found [16,17,18] that the preferred daunomycin triplet binding site contains adjacent GC base pairs of variable sequence, flanked by an AT base pair, hence the strand used in this project was the twelve base pair strand ***d(GCGCACGTGCG)$_2$*** with daunomycin intercalated between the base pairs C6-G18 and G7-C18.

The strand of DNA used was that of B-DNA. Deoxyribonucleic Acid (DNA) is a nucleic acid which consists of two long polymers of simple units called nucleotides, with backbones made of sugars and phosphate groups joined by ester bonds. These two strands run in opposite directions to each other and are therefore anti-parallel. Attached to each sugar is one of four types of molecules called bases (Adenine, Cytosine, Guanine, Thymine). DNA contains the genetic instructions used in the development and functioning of all known living organisms. The B-DNA molecule is a right-handed double helix, with anti-parallel backbone chains. It has about 10-10.5 nucleotides per turn. The base pairs are centered on the helix axis and nearly perpendicular to it. On an average the base pairs are rotated at 35.6˚ from the adjacent base pair and the average rise per base pair is 0.34nm [19]. Figure 2 shows the ribbon-slab representation of B-DNA structure where each residue has been colored using NDB format.

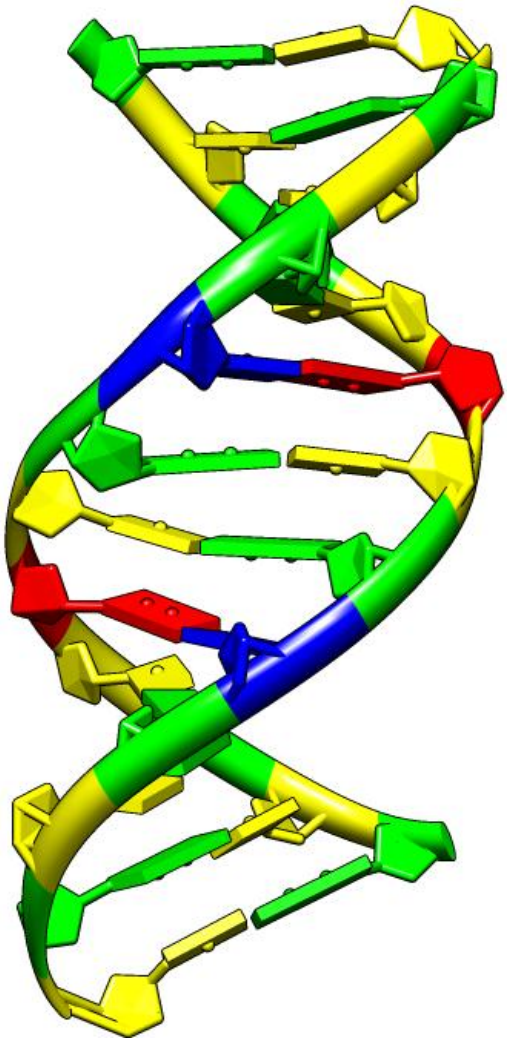

Figure 2: The above is a representation of BDNA generated from the Chimera software[20].Color of the residues follows Nucleic Acid Database (NDB) convention.

An important part of understanding water dynamics involved in the intercalation process is the hydrogen bond network which forms between the water molecules and the drug-DNA system. A hydrogen bond is the attractive interaction between a hydrogen atom and an electronegative atom (nitrogen, oxygen etc)[21]**,** when the hydrogen atom is covalently bonded to another electronegative atom. Hydrogen atom covalently bonded to a more electronegative atom carries a partial positive charge, which because of the hydrogen atom's small size translates into a large charge density, a hydrogen bond results when this strong positive charge density attracts a lone pair of electrons on another electronegative atom. The strength of a hydrogen bond is about one-tenth of a covalent bond. Every water molecule can make a maximum of four hydrogen bonds, the average number of hydrogen bonds between water molecules at room temperature is 3.5[(1)]. Many properties of water arise from the network of

hydrogen bonding, for example water has a high boiling point (100˚C) because of hydrogen bonding. The network of hydrogen bonds affects greatly the salvation dynamics when bio molecules like DNA are solvated in water. Hydrogen bonds play an important role in the structure of the DNA double helix[22-25]. Base pairing happens to a large part because of the specific requirements in the formation of hydrogen bonding between heterocyclic amines, so Adenine always pairs with Thymine and Cytosine always pairs with Guanine.

To get an accurate idea of the change which occurs in water dynamics, it is important to make calculations for as many structures as possible between the intercalated state of daunomycin and the "outside" binding of the drug.

In the current project we have used molecular dynamics (MD)[26,27] simulations for the study of the DNA-Daunomycin system. Over the years molecular dynamics simulations have developed as a strong tool for studying chemical systems. MD simulations have provided detailed information on the fluctuations and conformational changes of many biological molecules.

MD simulations generate information at the microscopic level, including atomic positions and velocities. MD simulations act as a bridge between theory and experiments, i.e we may test a theory by doing simulations using the same model and then test the model by comparing the results to theoretical results. These methods are regularly used, now, to study the structure, dynamics, and thermodynamics of molecules.

In this project we have used MD where the interactions between particles is described by a "force field", this is also called classical MD. Classical MD is deterministic, i.e once the positions and velocities of each atom are known, the state of the system can be predicted at any time in the future or the past. There also exist MD simulations where quantum chemical models or a mix of classical and quantum mechanical models are used.

There are many MD simulation program packages like AMBER[28], CHARMM[29], GROMACS[30,31] etc. In this project the molecular dynamics simulation package GROMACS (GROningen Machine for Chemical Simulations) was used for doing different simulations on the structures taken. The software PLUMED was used for umbrella sampling while constraining the structures.

In the given project we have taken the configurations generated by Mukherjee et al.[15], ranging from the intercalated state to the "outside" bound state. We have calculated different parameters (VACF, energy-autocorrelation etc) to get an idea of the water dynamics on three of the most important configurations from the above mentioned thirty, namely the intercalated state(IC), the first transition state(TS1) and the first minor groove bound state(MG1). Throughout the work, we have kept the initial conditions like the number of water molecules added to solvate he system, number and type of ions added to neutralize the system the same as that kept in the above mentioned work[15].

Three configurations (MG1, TS1, IC) were chosen based on the free energy calculations done by Mukherjee et al [15]., i.e the intercalated state (IC)(first structure) was the most stable energetically, the first transition state (TS1)(second structure) was the least stable structure, the first minor groove bound state (MG1)(third state) was the most stable out-side bound state. The shape of the strand changes as we go from one structure to the next, so the study of water dynamics of the different structure can throw light on the effect water has on the stabilizing these structures.

Methods Used

X, θ and Y are the collective variables used in the present project are taken from Mukherjee et al [15]. These collective variables are the coordinates used in the present work. During simulations, when the configurations (the intercalated state, first transition state, first minor groove bound state) were constrained by an umbrella force these three (X, θ, Y) were kept constant at different specific values for every configuration. These three variables have been defined by using the vectors $\hat{b}, \hat{c}, \vec{d}, \hat{p}$. Here $\hat{b}$ is a body-fixed unit vector which points towards the minor groove**.** $\hat{\boldsymbol{c}}$ is the unit body-fixed vector which points towards the DNA axis, $\hat{c}$ is chosen such that it is almost orthogonal to $\hat{b}$. $\hat{p}$ is the unit vector on the planar part of the drug (three planar aromatic rings) pointing to the tip of the drug .$X = \hat{b}.\vec{d}$, so it determines the separation of the drug from DNA. $Y = \hat{c}.\vec{d}$ so it determines the displacement of the drug along the DNA axis. $\hat{a}$ is perpendicular to both $\hat{b}$ and $\hat{c}$. $\theta = \hat{b}.\hat{p}$, so it determines the orientation of the drug with respect to the intercalating set of base pairs. Figure 3 schematically shows the vectors mentioned above which constitute the collective variables

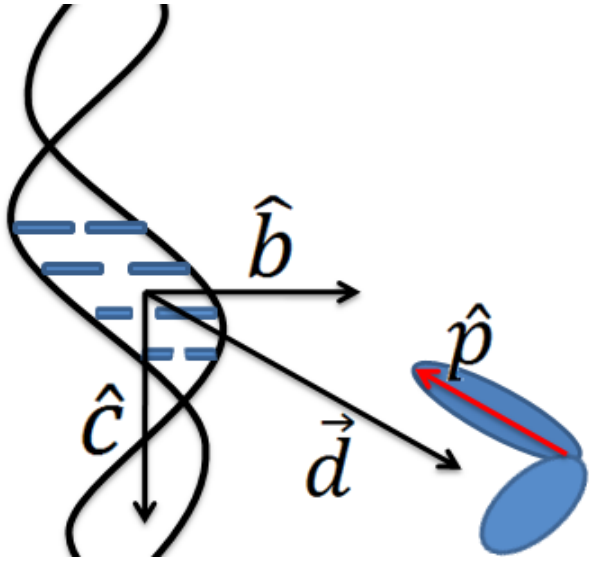

Figure 3: The collective variables used in the project, shown diagrammatically with respect to the DNA strand. The blue lines are the base pairs. $\hat{a}$ is perpendicular to both $\hat{b}$ and $\hat{c}$.

A: Initial Preparation:

Initially, configurations were taken with only the drug (daunomycin) and DNA. The drug-DNA system was placed in a cubic box. 8865 molecules of water were added as solvent. Twenty-two Na+ ions and one Cl- ion were added, to neutralize the system. These ions were

added as replacements for twenty three of the above added water molecules (one water molecule for every ion added). So, the total no. of water molecules left in the system was 8842.

The system was minimized, followed by a heating to 300K for 100ps with restraining of the positions of all heavy atoms (all atoms except the Hydrogen atoms were taken) in the drug-DNA complex with a force of *25kcal/mol/ Å* ^{2}. This was followed by five steps of energy minimization, each energy minimization followed by position restraints of progressively decreasing values of *5kcal/mol/ Å* ^{2}*, 2.5kcal/mol/ Å* ^{2}*, 1.75kcal/mol/ Å* ^{2}*, 1kcal/mol/ Å* ^{2}*, 0.5kcal/mol/ Å* ^{2}. The structure of the drug-DNA complex was checked at every step, to insure no distortions were creeping in.

All the above position restraints were done at a constant pressure of 1bar and a temperature of 300K with a coupling constant of 0.4ps. The thermostat used was v-rescale (velocity-rescaling)[32] and pressure coupling was done through the Parrinello-Rahman barostat[33,34]. The model used for water was TIP3P[35,36].

The Berendsen thermostat[37] was used earlier, but was creating distortions in the system as it suppresses the fluctuations in the kinetic energy of the system and therefore cannot produce trajectories consistent with the canonical ensemble[38], i.e it cannot be mapped onto a specific thermodynamic ensemble. So, the thermostat was changed from the Berendsen to the v-rescale

For the final equilibration, we constrained the value of X, θ and Y while applying an umbrella potential[39] and did a NVT simulation for 500ps, while storing the trajectories at every 1ps.

A harmonic umbrella potential $U = 1/2k(X - X_0)^2$ is used where k is spring constant and X_0 is the harmonic center[40], the value of k used was 5000 Nm^{-1}.

B: Calculation of Velocity Auto Correlation Function (VACF)

1) Production run

For calculating the velocity autocorrelation function of water molecules around the drug-DNA complex a NVT production run of 500ps was done, while constraining the values of X, θ and Y with an umbrella potential Trajectories were stored at every 4 femtoseconds (fs).

Velocity auto-correlation program of GROMACS (g_velacc) has been modified to calculate VACF for water molecules within any distance range from any group of the DNA and the drug.

2) Theory used

Velocity autocorrelation function is calculated using the following formula,

$$C(t) = \sum_{S} v(s)v(t+s) \qquad (1)$$

Velocity auto-correlation was calculated for water in all the three cases (intercalated state (IC), first transition state (TS1) and first minor groove bound state(MG1)). The VACF was calculated for water molecules at specified distance ranges from the system (DNA-drug complex). *This is done by considering velocity of all the water molecules in the VACF calculation if those water molecules reside for more than 10ps in the desired distance range from the selected parts of the DNA and the drug.*

The different groups taken were the minor groove of the three base pairs (A5-T20, C6-G19, G7-C18), which participate in the intercalation of daunomycin and the major groove of the above mentioned three base pairs. These groups were taken as daunomycin binds through the minor groove of the DNA strand, so the dynamics of water will be different around the major groove portion and the minor groove portion.

VACF was calculated for water molecules which were at distance ranges of 0-0.34nm and 1nm-20nm from the selected groups(DNA-drug complex, minor groove of the three base pairs, major groove of the three base pairs). These distance ranges were chosen in agreement with the widely accepted geometric definition of a hydrogen bond as having a length smaller than 3.5 Å .

By taking 0-0.34nm as a range of distance over which we calculate the VACF, we can calculate the VACF for only those water molecules which are part of the first hydration shell and the distance of 1.00nm-20.00nm can give water dynamics in bulk. Getting this bulk behavior is important as the VACF of bulk water molecules around the DNA have already been calculated [41] and by comparing our results we can make sure whether our results are correct or not, also the difference bulk behavior and the behavior of water molecules closer to the drug-DNA system can give us an idea of the effect the presence of the drug-DNA system has on the water dynamics.

While calculating the distances of the water molecules from the drug-DNA system, care had to be taken to choose a coordinate system which would eliminate the possibility of counting water molecules present in the system i.e. if we are calculating the VACF of water molecules at a distance range of 0-0.34nm from the major groove of the three base pairs (A5-T20, C6-G19, G7-C18), we should not take water molecules which, although in the above distance range, maybe present in the drug-DNA system itself, i.e. we want to take exclusively those water molecules which are outside the drug-DNA system. For this we would have to take a body-fixed coordinate system, i.e. a coordinate axis where the three orthogonal axis and the origin are fixed to the geometry of the body, in our case the DNA-drug complex.

3) The body-fixed coordinate system

The coordinate system has been shown schematically in figure 3. The first of the three orthogonal vectors defining our body-fixed coordinate system is $\vec{b}$, which is parallel to the axis of the DNA strand, the second is $\vec{c}$, which was parallel to the drug axis (Line passing through the three planar rings) and perpendicular to the DNA strand's axis, the third is $\vec{a}$, which is perpendicular to both the DNA's and drug's axis.

The VACF was decomposed to all the three coordinates for all three configurations in major and minor groove of the DNA.

C: Calculation of Energy auto-correlation.

1) Production run

For the calculation of energies, six bases (A5, C6, G7, C18, G19, T20) were taken as probes, for the calculation of the energy auto-correlations. Separate production runs of 5ns were done for each of the six probes for all three configurations taken (intercalated state, first transition state, first minor groove bound state). For every run, energies were stored at intervals of 100fs.

2) Theory used

The total energy was taken as the sum of the contributions of the total interaction energy of the probe with the rest of the DNA ($E_{p\text{-}DNA}$), with the ions ($E_{p\text{-}ion}$), with the water ($E_{p\text{-}w}$), with the drug ($E_{p\text{-}DM1}$)[42] as shown in Equation(2).

$$E_p(t) = E_{p-DNA}(t) + E_{p-ion}(t) + E_{p-w}(t) + E_{p-DM1}(t) \quad (2)$$

Energy auto-correlation is calculated using Equation(3)

$$C(t) = \frac{\langle \delta E(0) * \delta E(t) \rangle}{\langle \delta E(0) * \delta E(0) \rangle} \tag{3}$$

So, if for the auto-correlation of $E_{p\text{-}ion}$, the term would be:

$$C_{ion}(t) = \frac{\langle \delta E_{p-ion}(0) * \delta E_{p-ion}(t) \rangle}{\langle \delta E_{p-ion}(0) * \delta E_{p-ion}(0) \rangle} \tag{4}$$

Here, δE is the fluctuation from the equilibrium value.

D: Calculation of the number of H-Bonds as a function of distance

The number of hydrogen bonds was calculated as a function of distance from the selected group, i.e. from either the major groove or from the minor groove of the three base pairs (A5-T20, C6-G19, G7-C18). This can give us an idea about the existence of a possible second layer of hydration.

Results and Discussions:

A. Velocity Auto-Correlation Function (VACF): In this section we describe the VACF decomposition of water molecules in three orthogonal body-fixed axes attached to DNA within a distance of 0.34nm from the intercalating base pair set (A5-G7). This is performed for three structures IC, TS1, and MG1 described above.

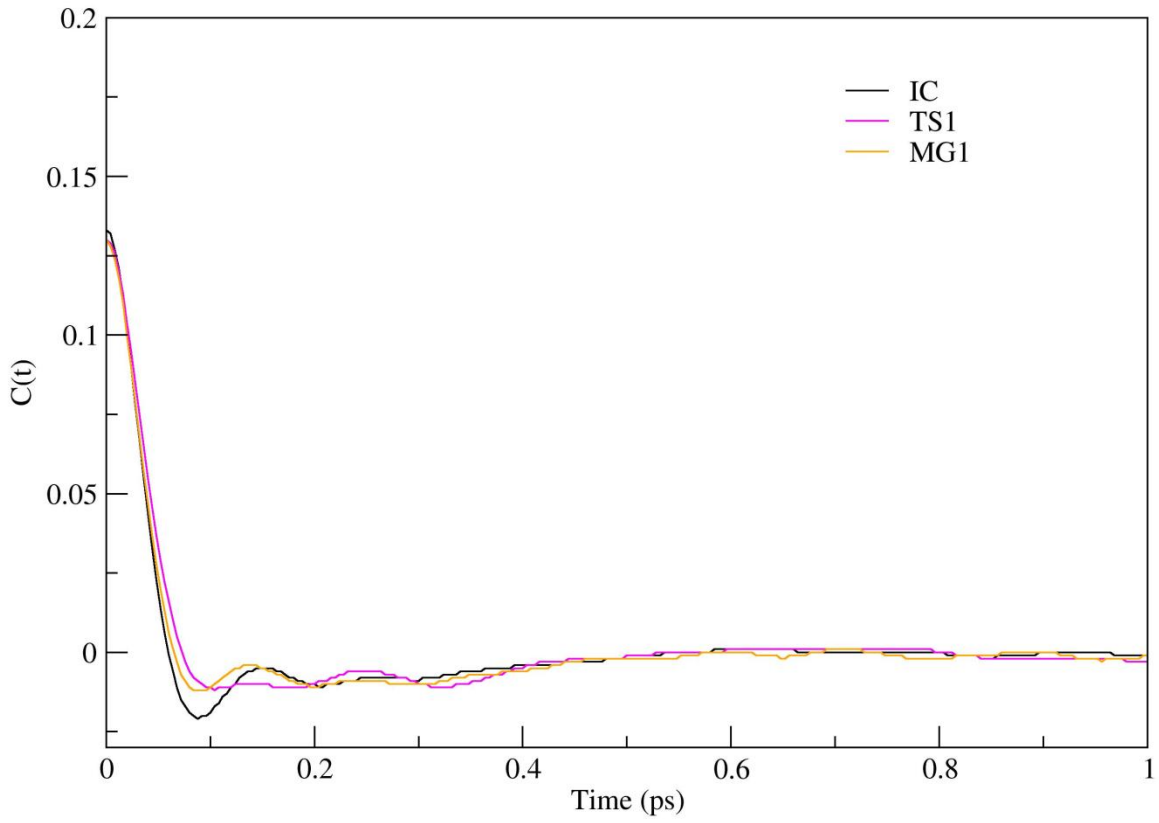

Figure 4: VACF for the major groove along $\vec{a}$.

Figure 4 shows the VACF of the three states mentioned above (IC, MG1, TS1) along the major groove for $\vec{a}$. IC is the only structure in which the drug protrudes to the major groove. Therefore, the slowing of VACF for IC could indicate the hindrance of water movement perpendicular to the DNA and Drug axis by the drug.

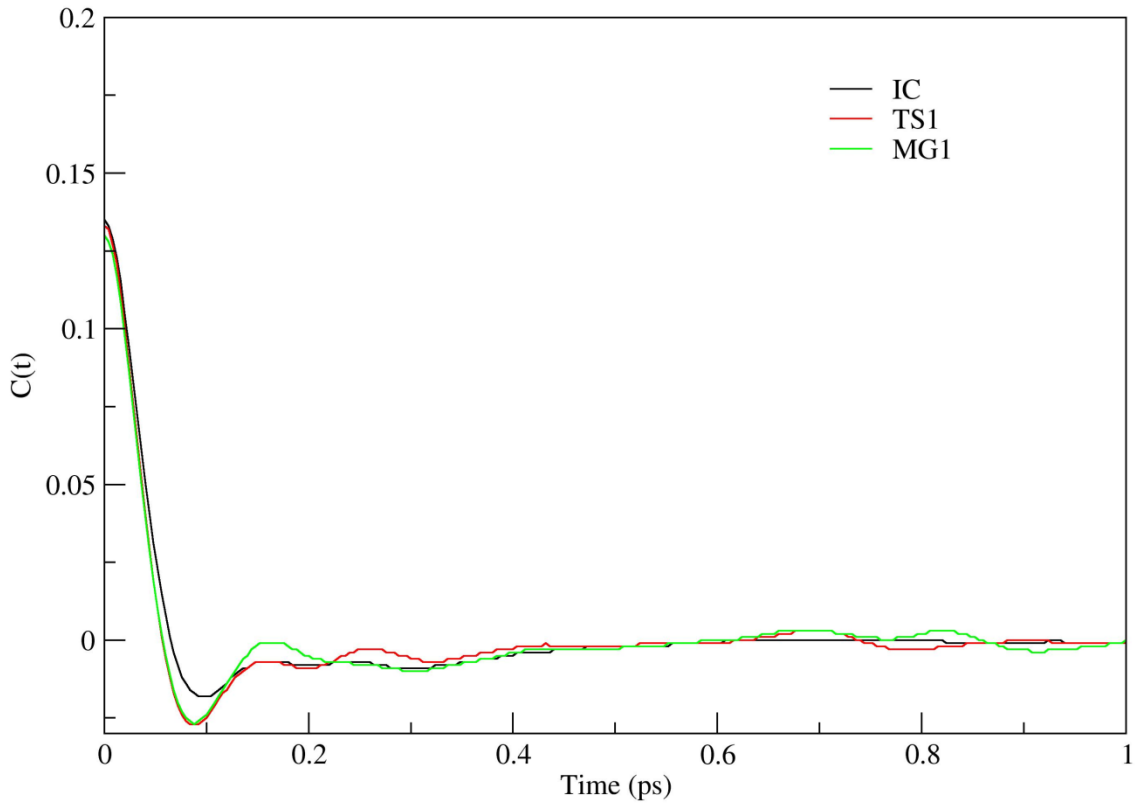

Figure 5: VACF for the major groove along $\vec{b}$.

Figure 5 shows the VACF of the three states mentioned above (IC, MG1, TS1) along the major groove for $\vec{b}$. Again IC is different from MG1 and TS1 since both MG1 and TS1 the drug does not appear in the major groove side. However, the dynamics is faster in the case of IC along the body fixed vector. This reflects that the water molecules chosen for the VACF calculation for IC lies farther from the DNA due to excluded volume of the drug appearing in the major groove side.

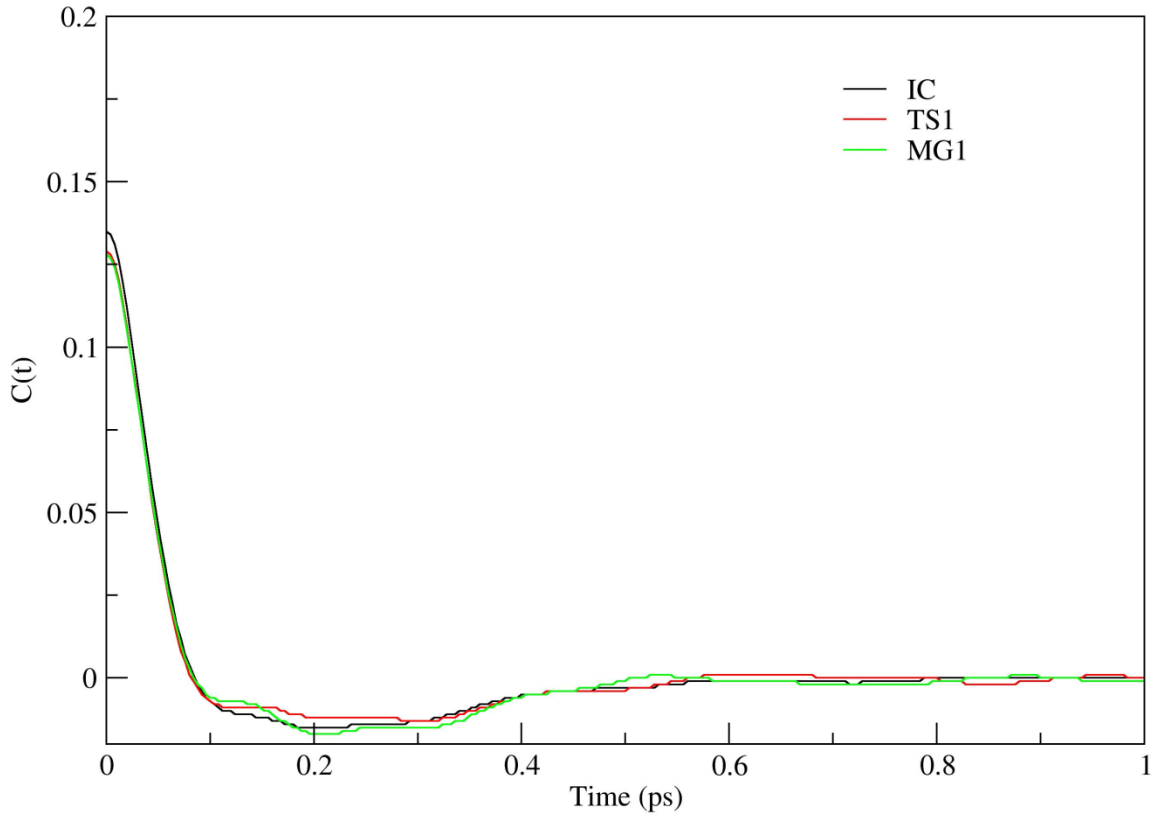

Figure 6: VACF for the major groove along $\vec{c}$

Figure 6 shows the VACF of water along the major groove for the DNA axis. This figure shows the interesting result that all the three structures show shallow minima of similar nature in the VACF in the major groove. This reflects the shallow nature of the major groove allowing random exchange of velocities resulting in similar profile.

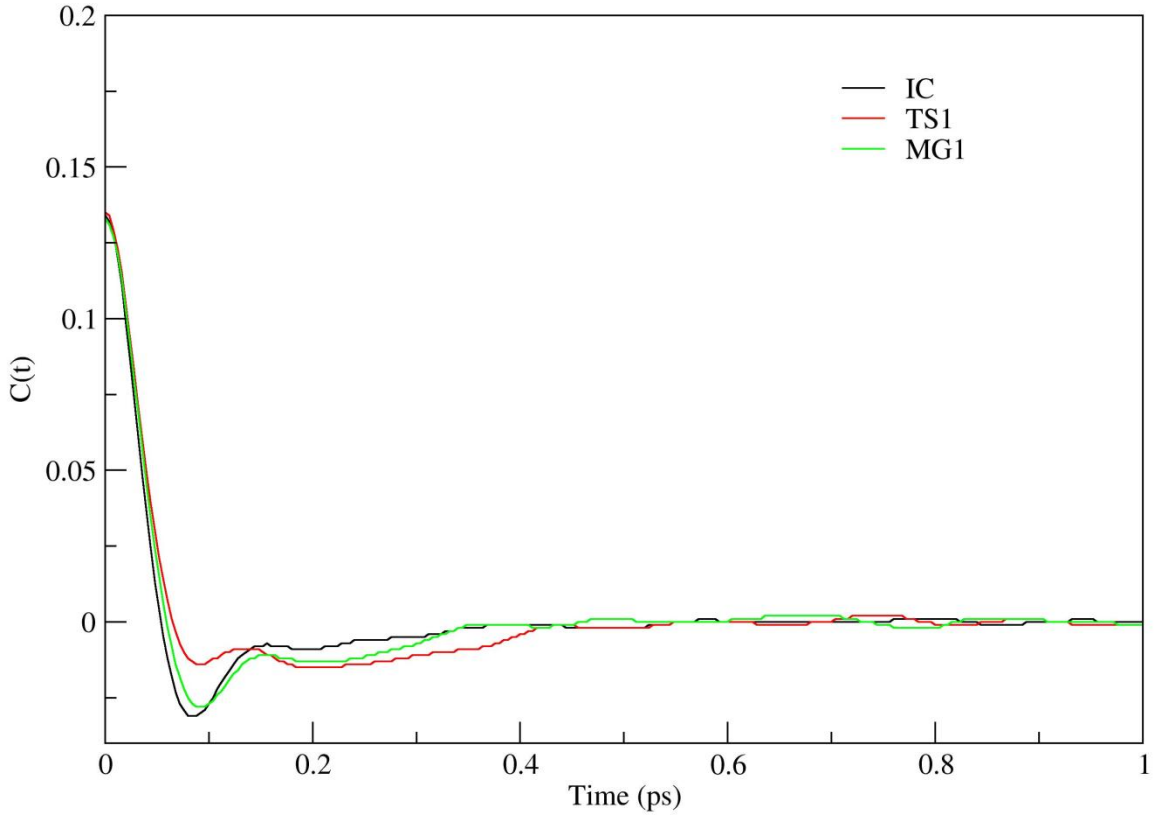

Figure 7: VACF for the minor groove along $\hat{a}$. Notice the fast water dynamics around TS1 compared to the stable state.

Figure 7 shows the VACF of water along the minor groove for $\vec{a}$. VACF shows slow down for IC and MG1. The drug is about to be inserted through the minor groove side of the DNA in the TS1 structure. In this configuration water shows faster dynamics probably due to coupled motion of the fast dynamics of the drug.

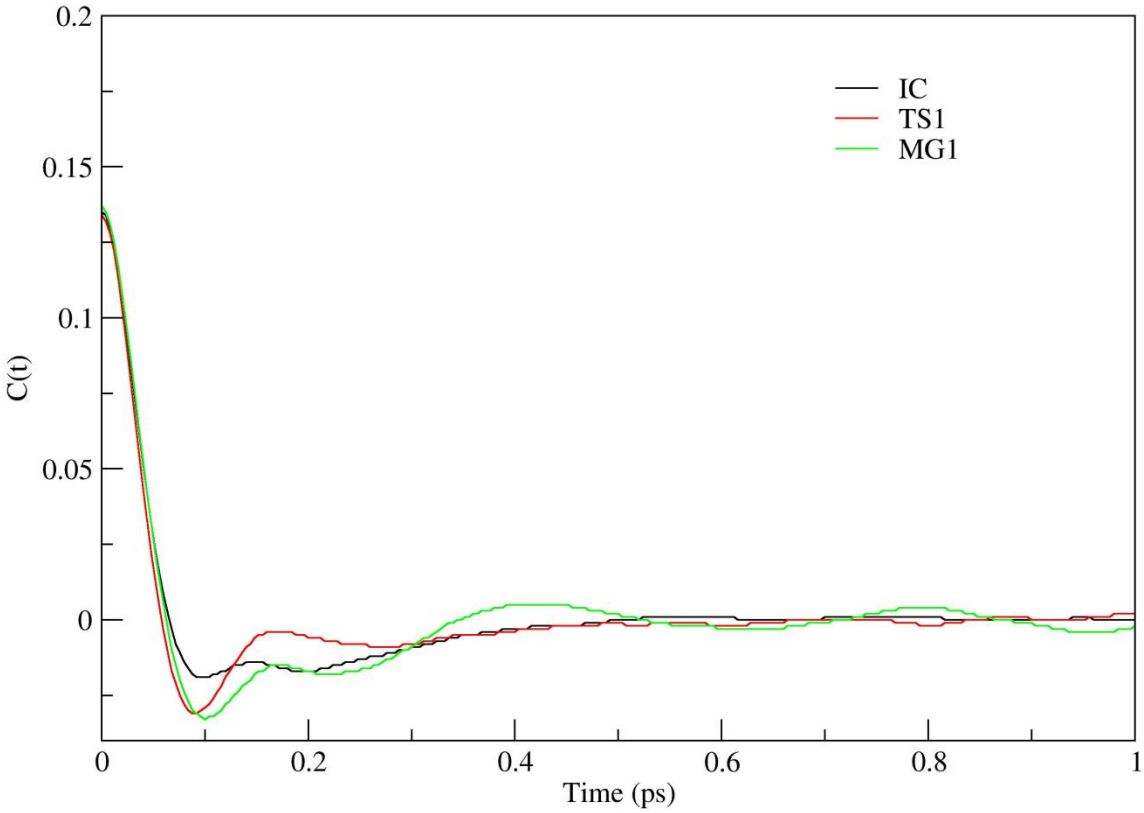

Figure 8: VACF for the minor groove along $\vec{b}$.

Figure 8 shows the VACF along the minor groove for the drug axis. VACF shows slow down for TS1 and MG1, which is very interesting as it implies faster water dynamics along the drug's axis when it is intercalated. In TS1 and MG1 as the drug is comparatively 'outside' as compared to IC, water molecules probed by the VACF calculation are lying within the DNA and the drug for TS1 and MG1 whereas they are outside the drug surface in case of IC.

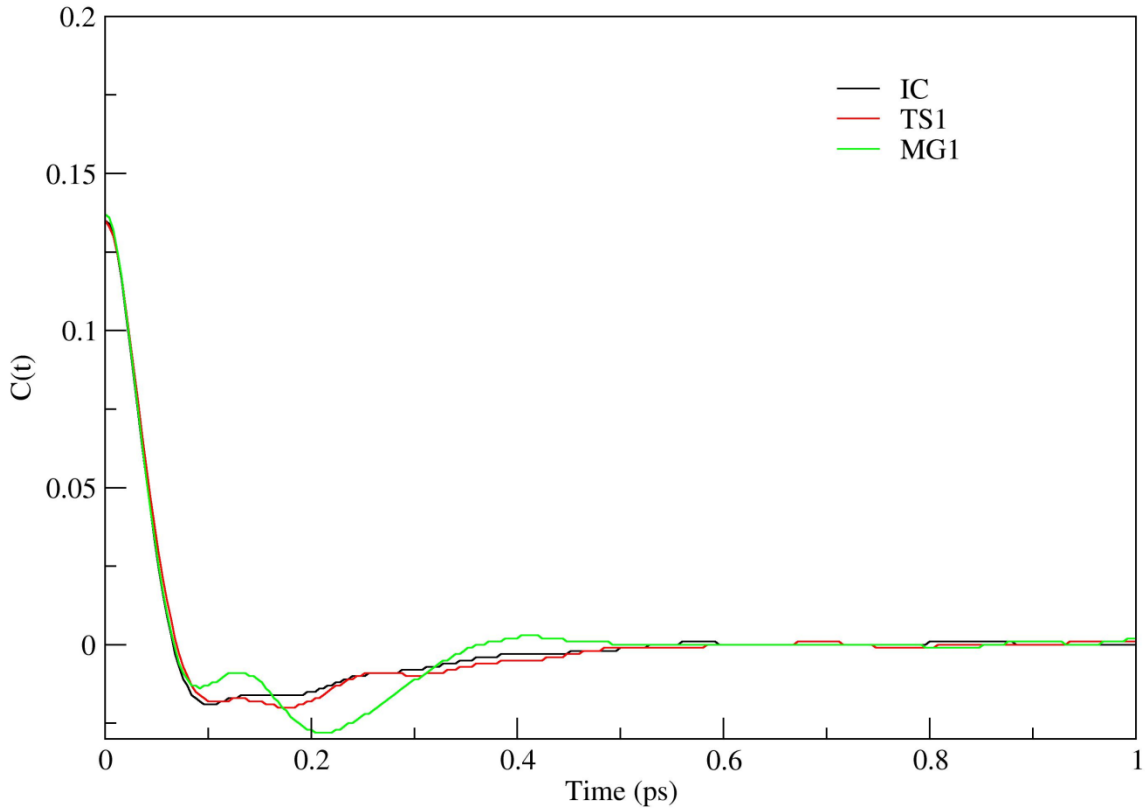

Figure 9: VACF for the minor groove along $\hat{c}$. Notice the appearance of the minima for the MG1 structure.

Figure 9 shows the VACF along the minor groove for the DNA axis.VACF of MG1 shows two minima. First minima shows faster water motions with normal correlation time whereas the second minimum shows slower water molecules with long correlation time. This needs to be further investigated to identify which water molecules showing faster and which are showing slower water dynamics. Remember that the VACF is calculated along the axis of the DNA and the alignment of the drug in the minor groove is along the axis.

Energy Autocorrelation:

Energy autocorrelation was fitted to triexponential curve. The values of the time constants and amplitudes were calculated for the entire run and the first 50ps.

The equation used was:

$$Y = a_0 * e^{(-x/a_1)} + a_2 * e^{(-x/a_3)} + a_4 * e^{(-x/a_5)}$$

Time constants and amplitudes:

1). Fitting coefficients of the triexponential function for the intercalated state (IC) within first 50ps.

	a_0	$a_{2(}$ps)	a_3	a_4(ps)	a_5	a_6(ps)
A5	0.62	0.07	0.26	1.10	0.12	8.67
C6	0.53	0.09	0.29	3.42	0.16	68.63
G7	0.65	0.07	0.23	1.10	0.11	10.78
C18	0.65	0.07	0.29	1.38	0.06	36.88
G19	0.49	0.08	0.26	1.88	0.25	53.14
T20	0.64	0.07	0.24	2.03	0.12	51.12

2) Fitting coefficients of the triexponential function for the intercalated state (IC) for the entire run of 5ns.

	a_0	a_1(ps)	a_2	a_3(ps)	a_4	a_5(ps)
A5	0.61	0.066	0.26	1.1	0.12	8.88
C6	0.52	0.084	0.3	3.05	0.17	68.60
G7	0.66	0.07	0.23	1.09	0.11	10.69

C18	0.71	0.09	0.23	2.4	0.04	252.2
G19	0.55	0.38	0.34	21.5	0.1	161.37
T20	0.7	0.1	0.22	4.66	0.07	356.83

3). Fitting coefficients of the triexponential function for the transition state (TS1) within first 50ps.

	a_0	a_1(ps)	a_2	a_3(ps)	a_4	a_5(ps)
A5	0.60	0.088	0.27	2.4	0.12	148.8
C6	0.65	0.067	0.18	2.18	0.14	50.64
G7	0.65	0.1	0.25	2.6	0.09	50.20
C18	0.59	0.067	0.28	1.2	0.13	20.70
G19	0.58	0.08	0.24	1.45	0.17	45.07
T20	0.62	0.08	0.22	2.14	0.16	87.50

3). Fitting coefficients of the triexponential function for the transition state (TS1) for the entire run.

	a_0	a_1(ps)	a_2	a_3(ps)	a_4	a_5(ps)
A5	0.58	0.077	0.28	1.84	0.12	139.83
C6	0.68	0.08	0.17	3.73	0.17	75.9
G7	0.64	0.09	0.25	2.35	0.10	43.81
C18	0.58	0.07	0.28	1.2	0.13	20.7
G19	0.61	0.10	0.22	2.00	0.16	55.06

T20	0.62	0.08	0.22	2.3	0.15	171.5

3). For the first minor groove bound state (MG1):for first 50ps.

	a_0	a_1(ps)	a_2	a_3(ps)	a_4	a_5(ps)
A5	0.66	0.06	0.26	1.14	0.08	14.01
C6	0.76	0.07	0.18	2.32	0.05	92.82
G8	0.69	0.07	0.23	1.9	0.07	44.57
C18	0.68	0.08	0.22	2.91	0.09	53.56
G19	0.73	0.07	0.19	2.38	0.07	46.57
T20	0.62	0.09	0.26	3.21	0.1	34.21

4). For the first minor groove bound state (MG1) for the entire run.

	a_0	a_1(ps)	a_2	a_3(ps)	a_4	a_5(ps)
A5	0.67	0.06	0.26	1.26	0.07	16.14
C6	0.78	0.07	0.18	2.87	0.04	102.96
G7	0.77	0.1	0.18	5.32	0.03	296.94
C18	0.71	0.09	0.23	4.74	0.05	473.56
G19	0.75	0.08	0.19	3.28	0.05	80.32
T20	0.61	0.08	0.25	2.51	0.13	25.52

4). For the normal strand of BDNA (Only A5, C6 and G7 were used as probes in this case) for first 50ps

	a_0	a_1(ps)	a_2	a_3(ps)	a_4	a_5(ps)
A5	0.66	0.083	0.26	1.71	0.08	37.15
C6	0.57	0.08	0.23	2.22	0.2	84.22
G7	0.62	0.071	0.29	2.05	0.09	37.00

5). For the normal strand of BDNA (Only A5, C6 and G7 were used as probes in this case) for the entire run

	a_0	a_1(ps)	a_2	a_3(ps)	a_4	a_5(ps)
A5	0.71	0.11	0.23	3.05	0.05	117.95
C6	0.58	0.085	0.22	2.7	0.19	102.25
G7	0.69	0.12	0.25	4.87	0.04	123.00

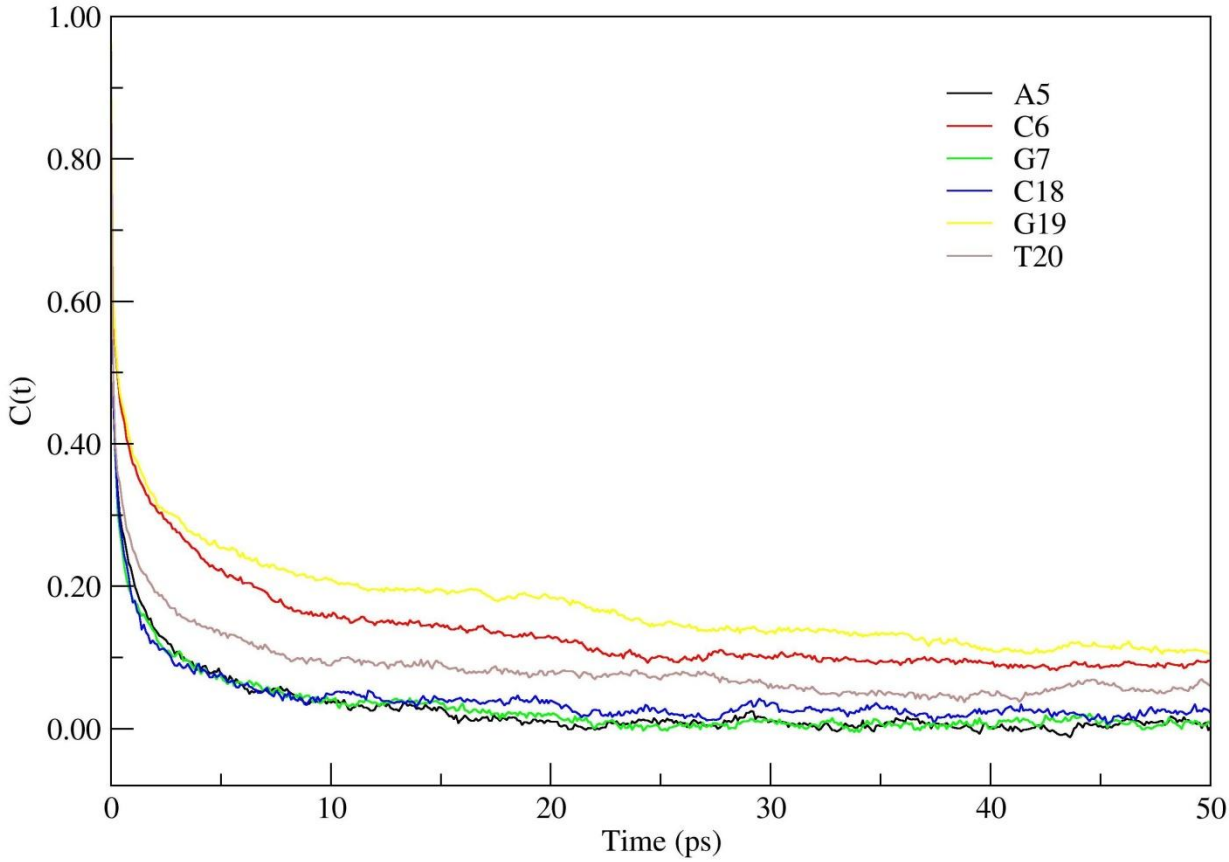

Figure 100: Energy autocorrelation over first 50ps for IC.

Figure 10 shows the values for the energy autocorrelation of the six different probes for IC for the first 50ps of the run. The fastest components were in the range of 1-10fs for each of the probes as they originated mainly from the solvent's contribution. The fast components are all in the range of 1-5ps, but there was significant difference between the slowest components of all other probes and A5, G7, thus implying relatively fast dynamics for these two probes for IC. It is interesting to note that although the drug is intercalated between C6 and G7, but we get a significant difference between the long time scales of both probes.

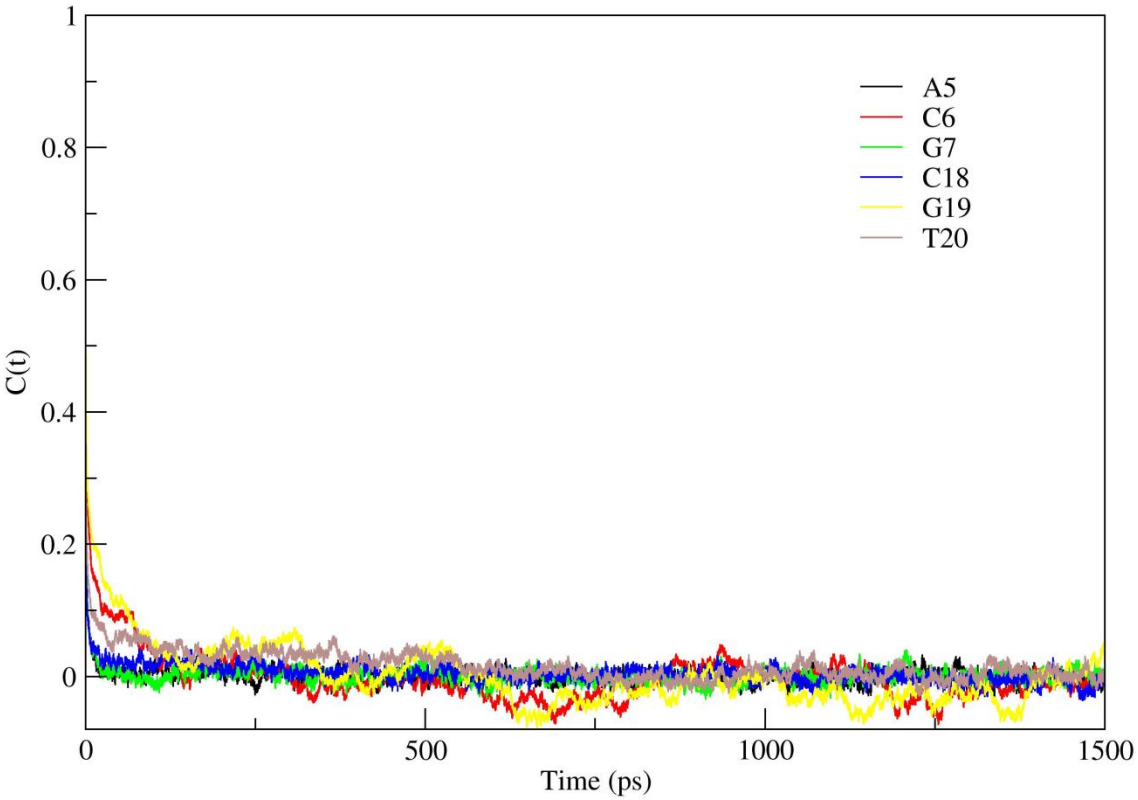

Figure 111: Energy autocorrelation over entire run for IC.

Figure 11 shows the values for the energy autocorrelation of the six different probes for IC for the entire run of the simulation. The fastest components are again of the same range for all the six probes. But now in the fast components the values for G19 are a lot higher than the other bases, this value was significantly higher for the entire run than for the first 50ps. The slow components were again very different for A5, G7 and the other four probes, but now the difference is very large.

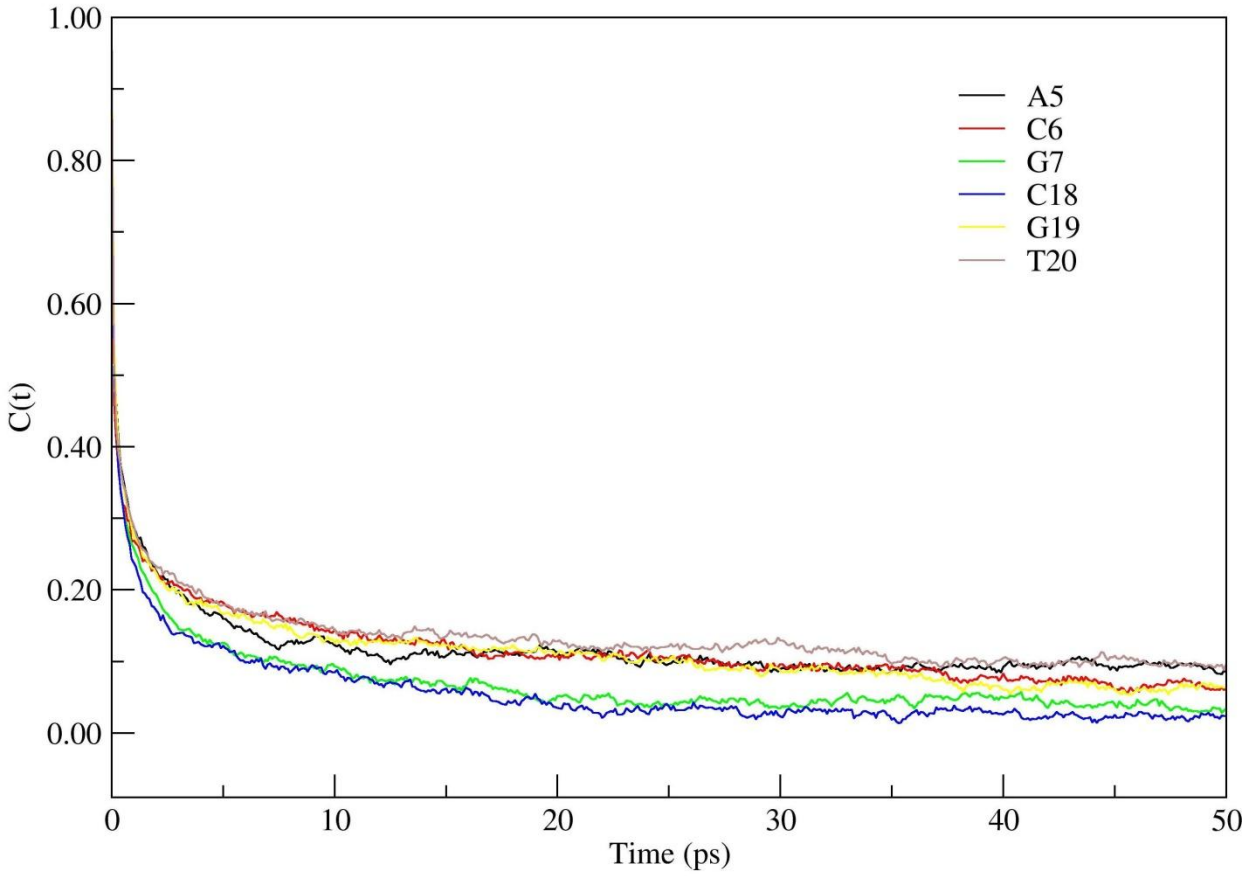

Figure 122: Energy autocorrelation over first 50ps for TS1.

Figure 12 shows the values for the energy autocorrelation of the six different probes for TS1 for the first 50ps of the run. Again the fastest components were all of the same range. The fast components were all of the same range, but in the slowest components the values were much higher for A5 and T20 as compared to others. The value was nearly similar for C6 and G7, implying their being in a similar environment for the first 50ps.

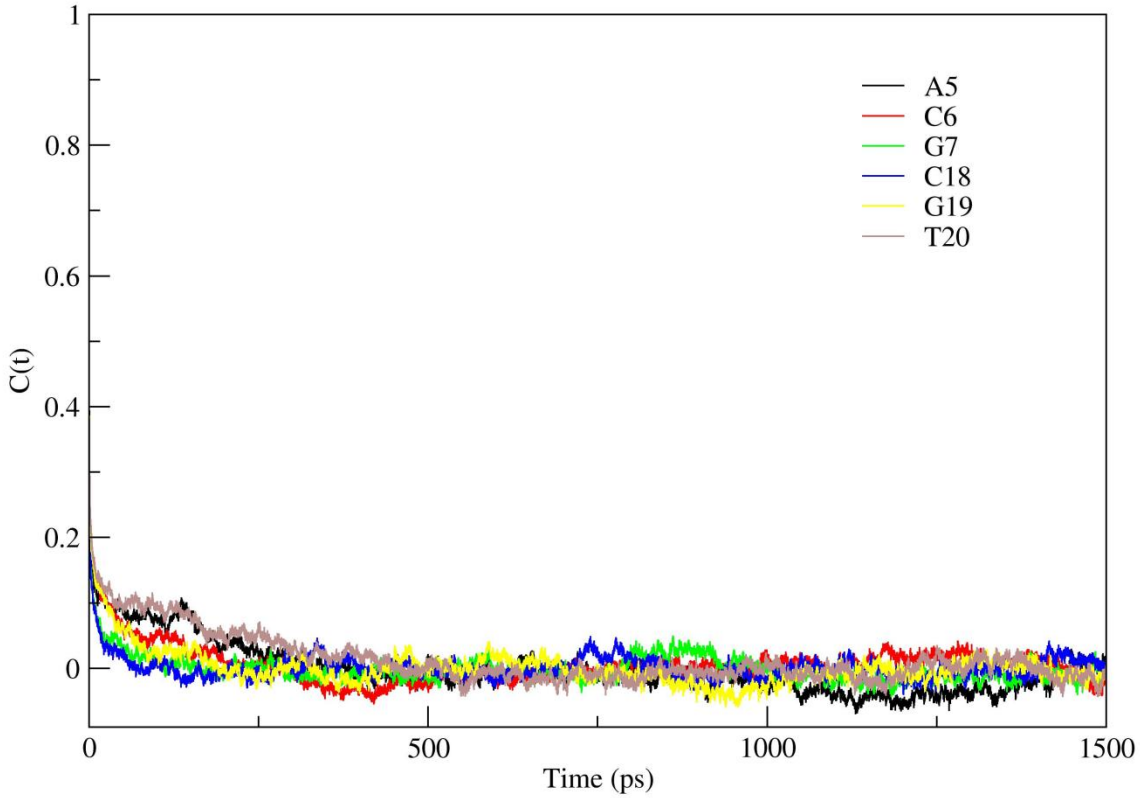

Figure 133: Energy autocorrelation over entire run for TS1.

Figure 13 shows the values for the energy autocorrelation of the six different probes for IC for the entire run. The fastest components were in the same range for all the six probes. The middle components were also in the same range, but in the slowest components the values for A5 and T20 were much larger than for the other probes, the value for T20 being much larger than for the first 50ps, implying that as the simulation proceeds it's interactions with the ions, drug change significantly.

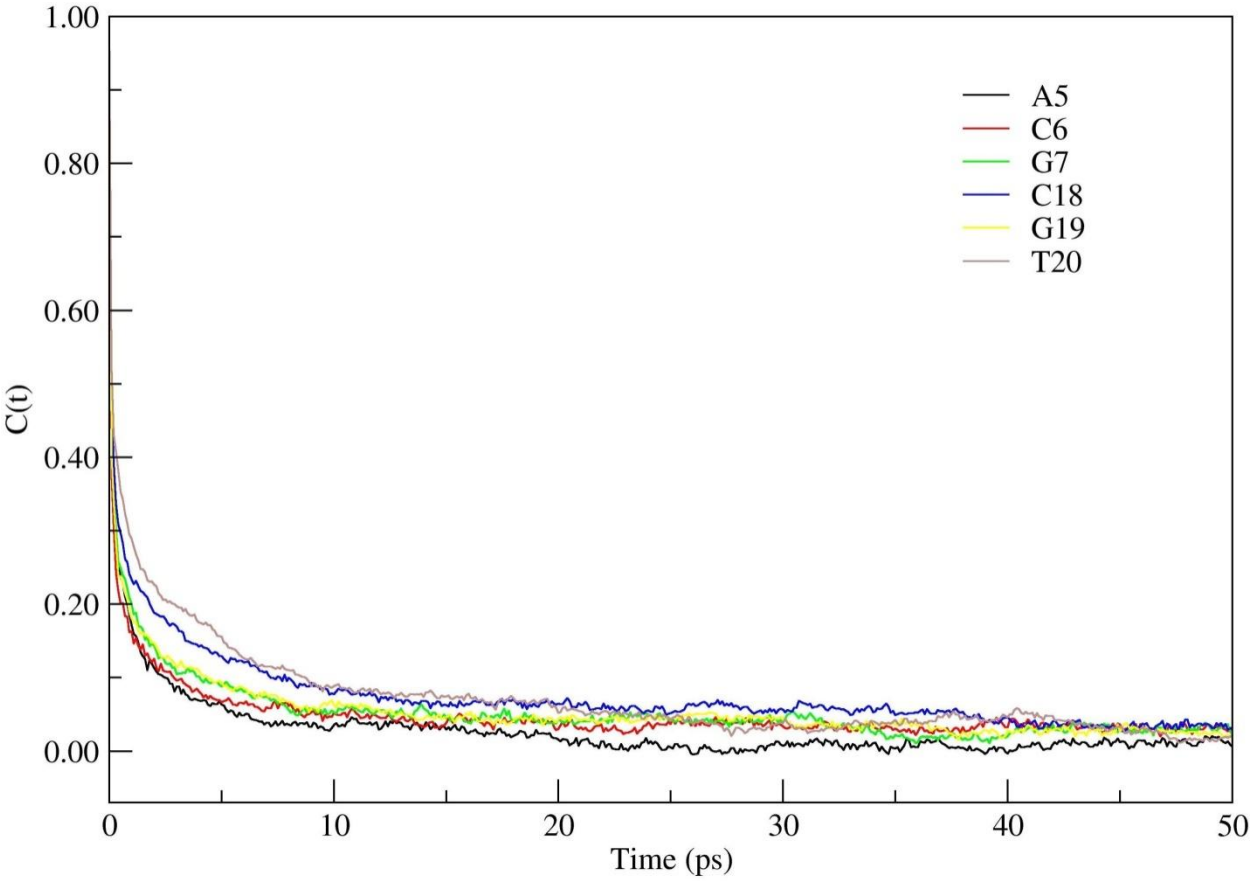

Figure 14: Energy autocorrelation over first 50ps for MG1.

Figure 14 shows the values for the energy autocorrelation of the six different probes for MG1 for the first 50ps of the run. The fastest components are of the same range, the fast components are also of the same range. In the slow components the value for C6 was larger and was much slower for A5. The distance between the drug and A5 is the greatest in this configuration and this relatively 'fast' slow component is arising from the interaction between the A5 probe and the drug.

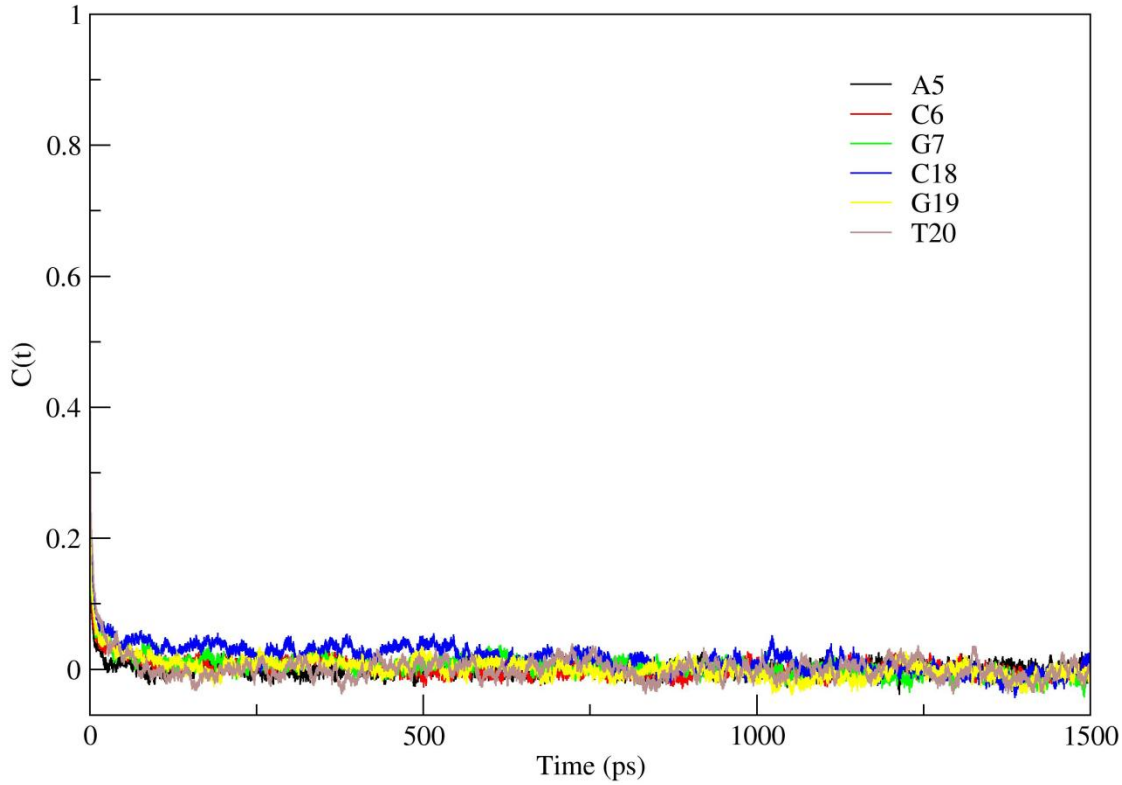

Figure 15: Energy autocorrelation over entire run for MG1.

Figure 15 shows the values for the energy autocorrelation of the six different probes for MG1 for the entire run. The fastest components wee gain of the same range for all six probes. The fast components were also of the same range, but the slowest components were very different with the 'fastest' being A5 (16.14ps) and the 'slowest' being C18 (473.56ps). It can be concluded from the above six graphs that the long time-scales are dominated by the drug's interaction with the probes. It can also be concluded that the fastest components are mainly dominated by the probe-water interaction, as they are of the same range for all cases for the different probes.

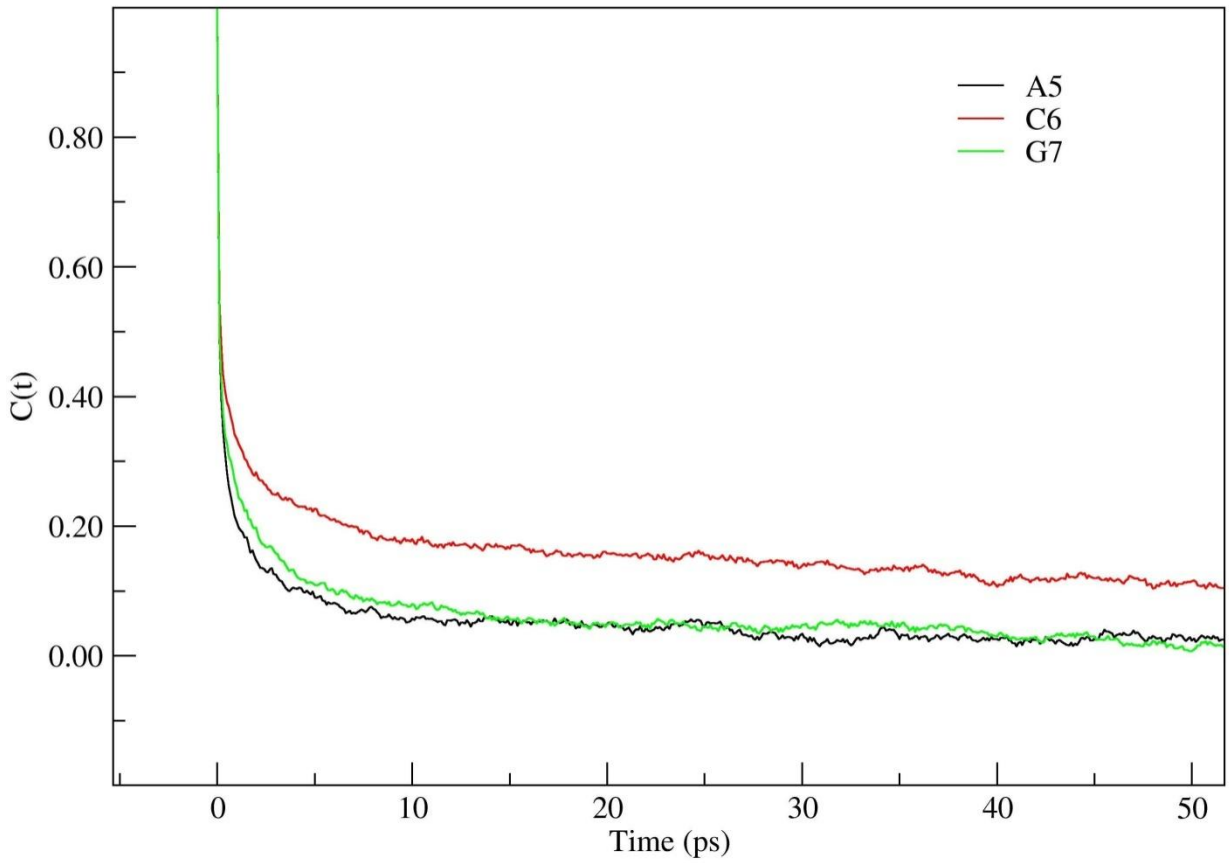

Figure 16: Energy autocorrelation over first 50ps for normal BDNA strand.

Figure 16 shows the values for the energy autocorrelation of the three different probes for B-DNA for the first 50ps of the run. The fastest components were of the same range, the fast components were also of the same range, but the slow component of C6 was larger than that of A5 and G7.

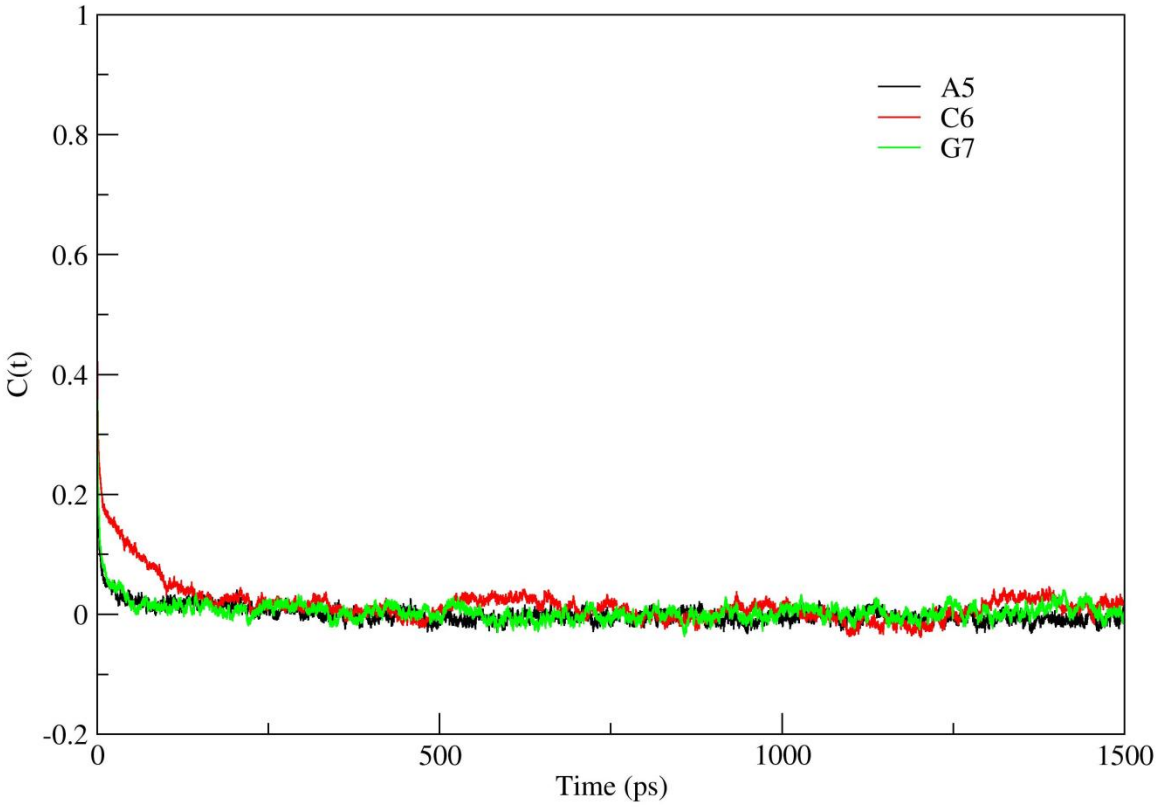

Figure17: Energy autocorrelation over entire run for normal BDNA strand.

Figure 17 shows the values for the energy autocorrelation of three different probes for the entire run. The fastest components were of the same range, the fast components were of the same range. It can be concluded that all the three probes would, in the case of BDNA exhibit similar time scales over al long time run, again validating the assumption that the drug (Daunomycin) present in the above three structures was caused by the interaction of the drug and the probe.

Conclusion:

We have performed several MD simulations, calculating water dynamics around DNA for different structures along the intercalation pathway in order to understand the dynamical role of water in the intercalation process. We have mainly performed velocity autocorrelation function along different body fixed axes of the DNA to see the anisotropic effect of water dynamics due to DNA and also due to the different positions of the drug. In order to achieve this, we took three representative configurations: intercalated state, minor groove-bound state and the transition state joining the two above. Previous studies focused on water dynamics around DNA, and the anisotropic effect using body-fixed axes and effect of water dynamics around states along intercalation pathway were not considered.

A GROMACS analysis code was modified to perform the calculations presented here. The results show interesting behavior for the different systems in major and minor groove. Some of the results are anomalous and require through study.

We found in our calculations of the energy autocorrelation for the three structures (IC, TS1, MG1) that the difference for the probes over the short runs was smaller than that for the long runs, interestingly this was not the case for BDNA. It was concluded that the slow time constants were arising from the drug's interaction with the probes.

Further study on the water dynamics along a more detailed intercalation pathway (involving more structures) would lead to an understanding about the exact role of water on this type of recognition processes.

References:

1) DiMarco, A.; Gaetani, M.; Orezzi, P.; Scarpinato, B.; Silvestrini, R.;Soldati, M.; Dasdia, T.; Valentini, L. *Nature* **1964**, *201*, 706.

2) Myers, C. E.; Chabner, B. A. Anthracyclines. In *Cancer Chemotherapy: Principles and Practice*; Chabner, B. A., Collins, J. M., Eds.; Lippincott: Philadelphia, 1990; p 356.

3) Lerman, L. S. J . Mol . Eiol. 1961, 3, 18.

4) Baginski, M.; Fogolari, F.; Briggs, J. M. *J. Mol. Biol.* **1997**, *274*, 253.

5)Trieb, M.; Rauch, C.; Wibowo, F. R.; Wellenzohn, B.; Liedl, K. R. *Nucleic Acids Res.* **2004**, *32*, 4696.

6) DiMarco, A.; Gaetani, M.; Orezzi, P.; Scarpinato, B.; Silvestrini, R.;Soldati, M.; Dasdia, T.; Valentini, L. *Nature* **1964**, *201*, 706.

7) Myers, C. E.; Chabner, B. A. Anthracyclines. In *Cancer Chemotherapy:Principles and Practice*; Chabner, B. A., Collins, J. M., Eds.; Lippincott: Philadelphia, 1990; p 356

8) Weiss, R. B. *Semin. Oncol.* **1992**, *19*, 670.

9) Lerman, L. S. *J. Mol. Biol.* **1961**, *3*, 18.

10) Chaires, J. B.; Satyanarayana, S.; Suh, D.; Fokt, I.; Przewloka, T.;Priebe, W. *Biochemistry* **1996**, *35*, 2047.

11) Chaires, J. B.; Dattagupta, N.; Crothers, D. M. *Biochemistry* **1985**,*24*, 260.

12) Wang, A. H.-J.; Ughetto, G.; Quigley, G. J.; Rich, A. *Biochemistry***1987**, *26*, 1152.

13) Denny, W. A.; Baguley, B. C. *Curr. Top. Med. Chem.* **2003**, *3*, 339.

14) Denny, W. A. *Expert Opin. In*V*estig. Drugs* **1997**, *6*, 1845.

15) Mukherjee A and Hynes J T. (*unpublished work*).

16) Chaires, J. B.; et al. Site and sequence specificity of the daunomycin-DNA interaction. Biochemistry 1987, 26 (25), 8227.

17) Chaires, J. B.; Herrera, J. E.; Waring, M. J. *Biochemistry* **1990**, *29*, 6145.

18) Cieplak, P.; Rao, S. N.; Grootenhuis, P. D. J.; Kollman, P. A. *Biopolymers* **1990**, *29*, 717.

19) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984.

20) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J. Comput. Chem.* **2004**, *25*, 1605.

21) Jeffrey, G. A. An Introduction to Hydrogen bonding; Oxford University Press: New York, Oxford, 1997, Chapter 10.

22) Guerra, C. F.; Bickelhaupt, F. M.; Snijders J. G.; Berends. J. Am. Chem. Soc. **2000**, *122*, 4117.

23) Jeffrey, G. A.; Saenger, W. Hydrogen Bonding in Biological Structures; Springer-Verlag: Berlin, New York, Heidelberg, 1991.

24) Saenger, W. Principlesof Nucleic Acid Structure; Springer-Verlag: New York, Berlin, Heidelberg, Tokyo, 1984.

25) Watson, J. D.; Crick, F. H. C. Nature 1953, 171, 737.

26) M.P. Allen and D.J. Tildesley, Computer Simulations of Liquids. , Clarendon Press, Oxford (1987).

27) D. Frenkel and B. Smit, Understanding Molecular Simulation, 2nd ed. Academic, New York, 2002.

28) Case, D. A. *AMBER7*; University of California: San Francisco, 2002.

29) Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983). "CHARMM: A program for macromolecular energy, minimization, and dynamics calculations". *J Comp Chem* **1983,** *4*, 187.

30) Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ (2005). "GROMACS: fast, flexible, and free". *J Comput Chem* **2005,** *26*, 1701.

31) Berendsen, H. J. C.; Spoel, D. v. d.; Vandrunen, R. *Comput. Phys. Commun.* **1995**, *91*, 43.

32) G. Bussi, D. Donadio, and M. Parrinello, J. Chem. Phys. **2007,** *126*, 014101.

33) Andersen, H. C. *Journal of Chemical Physics* **1980**, *72*, 2384.

34) Parrinello, M.; Rahman, A. *Journal of Applied Physics* **1981**, *52*, 7182.

35) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.;Klein, M. L. J. Chem. Phys. **1983**, *79*, 926.

36) Neria, E.; Fischer, S.; Karplus, M. J. Chem. Phys. **1996**, *105*, 1902.

37) H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, andJ. R. Haak, J. Chem. Phys. **1984**, *81*, 3684.

38) L. D. Landau and E. M. Lifshitz, "Statistical Physics, 3rd Edition Part 1", Butterworth-Heinemann, Oxford, 1996.

39) Torrie, G. M.; Valleau, J. P. J. Comp. Phys. **1977**, *23*, 187.

40) Mukherjee, A.; Lavery, R.; Bagchi, B.; Hynes, J. T. J. Am. Chem. Soc. **2008**, *130*, 9747

41) Pal, S.; Maiti, P. K.; Bagchi, B.; Hynes, J. T. J. Phys. Chem. B **2006**, *110*, 26396.