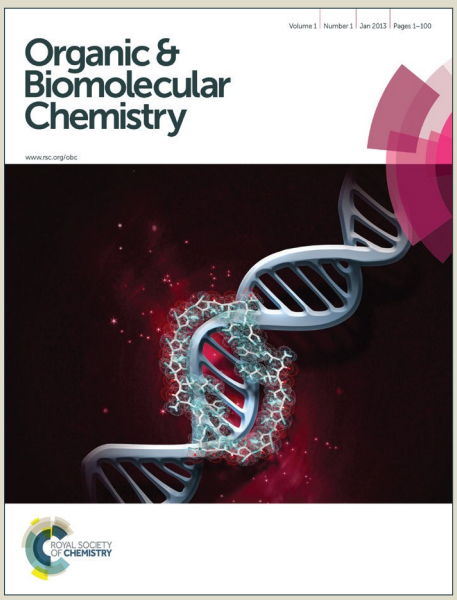
Organic & Biomolecular Chemistry

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/obc

An Investigation on the Interaction Modes of a Single-Strand DNA Aptamer and RBP4 Protein: A Molecular Dynamic Simulations Approach

Raheleh Torabi[a], Kowsar Bagherzadeh[b,c], Hedayatollah Ghourchian[a], Massoud Amanlou[c*]

[a]Laboratory of Microanalysis, Institute of Biochemistry & Biophysics, University of Tehran, P.O. Box 13145-1384, Tehran, Iran.

[b]Razi Drug Research Center, Iran University of Medical Sciences, Tehran, Iran.

[c]Department of Medicinal Chemistry, Faculty of Pharmacy and Drug Design and Development Research Center, Tehran University of Medical Sciences, Tehran, Iran.

* Address correspondence to this author Department of Medicinal Chemistry, Faculty of Pharmacy and Drug Design and Development Research Center, Tehran University of Medical Sciences, Tehran, Iran;

Tel: +98-21-66959067; Email: amanlou@tums.ac.ir

Abstract

Type two diabetes is one of the concerning health issues threatening public well-being worldwide. One of the pre-diagnosis biomarkers of this disease, Retinol binding protein 4 (RBP4) has been shown that can be detected with a 76-mer ssDNA aptamer instead of conventional antibodies. There is no structural information on RBP4 binding aptamer (RBA) and the mechanism of its binding to RBP4 still remains unexplored. The objective of the present study is to achieve a better understanding of specific binding interactions of the target protein (RBP4) and RBA, employing Molecular Dynamics simulations (MDs) to provide detailed information on fluctuations, conformational changes, critical bases and effective forces to develop regulated aptamers to be later employed in designing new aptamers for many useful recognition applications. RBA was designed according to its reported base pair sequence and secondary structure. HADDOCK on line docking program was used to predict a suitable RBP4-RBA mode of interaction to start MDs with. MDs methodology was used to analyze the final

Organic & Biomolecular Chemistry Accepted Manuscript

complex stability and detect interacting residues. Eventually, we conclude that single strand located bases are the key components that conduct the intercalation phenomenon with big targets rather than the ones involving loops and folded motifs, to encompass targets and probably inhibit their activity. Also, UV–visible, circular dichroism, and fluorescence spectroscopy measurements confirmed the interactions in between RBA and RBP4 and the RBP4-RBA complex formation.

Keywords: Diabetes, Insulin Resistance, Retinol binding protein 4, Aptamer, Haddock, Molecular dynamics simulations

Organic & Biomolecular Chemistry Accepted Manuscript

Introduction

Type two diabetes (T2D) has become a major challenge for public health worldwide. The increasing prevalence of this type of diabetes as a result of ageing, urbanization, increasing rate of obesity, physical inactivity and modern lifestyle lead to increased insulin resistance and abnormal glucose metabolism that eventually resulted in T2D[1]. These conditions change the levels of some biomarkers which can be prognosis for advent of T2D.

Retinol binding protein 4 (RBP4), adipocyte-secreted protein, directly depends on obesity and insulin resistance. Any increase in RBP4 level would alarm the advent T2D in a near future before observing any other significant changes in other diabetic biomarker levels [2 3]. Traditional clinical test to quantify concentrations of human serum RBP4 is enzyme-linked immunosorbent assay (ELISA) which apply mono and poly clonal antibodies as the detecting probes[4]. But antibodies are being replaced with novel molecules called aptamers that are recently introduced as diagnostic markers with similar affinity and specificity in different assays. Aptamers are single-stranded nucleic acids (ssDNA or RNA) with the ability of binding various types of molecular targets based on their three–dimensional structures. Aptamers have many advantages over antibodies such as being cheaper and less time-consuming to be synthetically produced and characterized. Further, aptamers are more stable in harsh conditions, have smaller molecule size, lack the large hydrophobic cores of proteins and as a matter do not aggregate, have lower immunogenicity and are easily modifiable chemically [5]. ssDNA aptamers are inherently more stable, cheaper, and easier to produce than RNA aptamers. RNA also requires reverse transcription[6], whereas ssDNA does not require this extra stage in SELEX process [7].

In 2008, Lee and collaborators reported a 76-mer ssDNA aptamer for detection, recognition and binding with high affinity (Kd: 0.2 ±0.03 μM) and specificity for RBP4 with an unknown

Organic & Biomolecular Chemistry Accepted Manuscript

mechanism of action[8]. Understanding of biophysical basis for the specificity of aptamer-target binding has been considered in some studies [9]. Strongly negatively charged nucleic acids within the folded macromolecule structure at physiological pH and in electronegative pockets are the ones responsible for electrostatic interactions with the targets [10]. But it should be considered that pure experimental studies are not capable of verifying all types of interactions that preset between aptamers and the target macromolecules precisely[11, 12].

Technical advances and advent of computational methodologies including classical molecular dynamic simulations (MDs) have eased the way for the scientists to have a better understanding of the phenomenon occurs during macromolecule interactions. Application of MDs technique to study nucleic acids behavior occurred a while after proteins, which is due to the charged and very flexible nature of nucleic acids in aqueous solutions. During the last decade, the technique have been widely used for monitoring systems with large number of nucleic acids [13-19] and predicting the dynamic nature of these macromolecules interactions based on initial biophysical data [20]. Accordingly, three different models of aptamer-target interactions have been proposed including; adaptive recognition (induced fit) [14, 21], preformed (lock and key) [22], and a combination of these two models.

The objective of the present study is to achieve a better understanding of specific binding interactions of the target protein (RBP4) and RBP4 binding aptamer (RBA), designed by Lee et.al.[8], employing MDs to provide detailed information on fluctuations, conformational changes, critical bases and effective forces to develop regulated aptamers to be later employed in designing new aptamers for many useful recognition applications. Further, UV–visible, circular dichroism, and fluorescence spectroscopy measurements were utilized to ensure the RBP4-RBA complex formation.

Organic & Biomolecular Chemistry Accepted Manuscript

Material and Methods

Computational Methodology

Three-dimensional structure preparation

The X-ray structure of RBP4 was taken from the structure of recombinant human serum RBP4 (PDB ID: 2wq9) [23]. The X-ray structure of RBP4 was analyzed with the aid of Swiss PDB viewer website [24] and the missed atoms of some residues (ARG2, ASN101, ARG153, GLU158, ILE 168) were fixed. Protonation state of various residues have been determined from the pKa calculation by PROPKA [25] and the residues neighborhood environment.

Due to the absence of an appropriate three-dimensional structure for 76-mer ssDNA, the three-dimensional structure of RBA was designed employing the software package Hyperchem 8.0 [26, 27] and Discovery Studio 4.1 [28]. First, a secondary structure was predicted according to RBA reported base pair sequence by means of the free-energy minimization algorithm according to Zuker with the application of the web server based mfold tool that showed a typical stem and loop motives [29]. Based on this predicted secondary structure, double stranded regions of RBA in loops were derived from other reported DNA structures (PDB IDs: 167D, 425D, and 119D). Then the rings were bonded to each loop followed by an energy minimization step in Hyperchem 8.0 with the employment of steepest descent algorithm. Then the loops were respectively added to linear stems of RBA and energy minimized to achieve a final structure. Since the conditions under which fold model of nucleic acids folding is not a reasonable physiologic condition, molecular dynamics simulation was run to ensure the modeled macromolecule.

Organic & Biomolecular Chemistry Accepted Manuscript

Docking studies

HADDOCK (High Ambiguity Driven protein-protein DOCKing) on line docking program was used to predict a suitable mode of RBP4 interacting residues and RBA bases in order to both calculate an approximate binding energy as well as obtaining an appropriate mode of interactions to start MDs with. The best complex was selected considering the type of interactions (mainly hydrogen bonds, hydrophobic and electrostatic interactions), the residues involved, stability scores, the binding energy and also HADDOCK scores[30, 31].

Molecular dynamics simulations studies

GROMACS 4.5.5 package[32] and Amber force field (amber99sb-ildn)[33] were used to analyze the final complex stability and detecting interacting residues. Amber force field have used in most of aptamer MDs studies, because of accurately reproducing the structural and dynamic properties of a large variety of NAs, describing complex conformational changes, performing well in simulations of DNAs in extreme environments and the excellent ability to reproduce very high-level QM data for hydrogen bond and stacking interactions [20].

The models were solvated with TIP-3 water model using dodecahedron box shape with the minimum distance of 10 Å between the protein surface and the box walls. The system was neutralized by replacing water molecules with appropriate number of counter ions, sodium and chloride. The particle mesh Ewald method was used for the electrostatic interaction Simulations. The van der Waals cutoff was 14 Å. Periodic boundary conditions were assigned in all directions. The whole system was then submitted to energy minimization employing steepest descent algorithm with tolerance of 1000 kJ/mol/nm. After convergence, the system went through NVT ensemble MD simulations for 20 ps. MD simulations were carried out employing

Organic & Biomolecular Chemistry Accepted Manuscript

NPT in a periodic boundary condition with constant number of particles in the systems, constant pressure, and constant temperature simulation criteria. The bond lengths that involved hydrogen atoms were restrained employing LINCS, allowing an integration step of 1 fs. The systems were coupled to external constant temperature (100, 200, 300 K with coupling time of τ_t = 0.1 ps and each for 1 ns) in three steps and external constant pressure (with coupling time of τp = 0.5). The final MD simulations were extended for extra 50 ns at constant pressure and temperature conditions. Berendsen barostat and thermostats were applied to keep the pressure and temperature constant at 1 bar and 300 K. The frequency of the snapshot was considered 10 to save every 10 frames of the trajectory file.

Graphical representation softwares

The MD simulations results were analyzed with the aid of VMD software [34] and visualized using PyMol [35, 36]. The graphs were all represented with the application of Grace GUI toolkit 5.1.22 version [37], except those of RMSF which were plotted with the aid of Microsoft Office Excel 2010.

Experimental Studies

Materials and Reagents

Human RBP4 full length protein was provided by abcam (ab63267, abcam, USA). A 76-mer RBA were purchased from Bioneer (Germany). A 400 µM concentration of RBA was prepared in 0.2 M Saline phosphate buffer (pH 7.0). 100 µl of 0.4 µM RBA in 0.2 M Saline phosphate buffer (pH 7.0) was added into 100 µl of 400 µM RBP4 to achieve a 200 µM concentration of RBP4. 100 µl of 0.2 M Saline phosphate buffer (pH 7.0) was added into 100 µl of 400 µM RBP4

Organic & Biomolecular Chemistry Accepted Manuscript

to achieve a 200 μM concentration of RBP4 as the control sample and then both wells were incubated for 2 hrs at the room temperature (25±1 ºC) with mild shaking. The final solution was centrifuged at 9,000 rpm (7,000 g) for 10 minutes to collect at the bottom while the unconjugated RBA was left in the supernatant. Then the supernatant was discarded and the conjugates was collected from the bottom of the tubes. The whole washing step was repeated at least three times. 200 μl of 0.2 M Saline phosphate buffer (pH 7.0) was used as the blank sample in all experiments.

Devices

The UV–visible spectroscopic measurements were performed on Thermo Scientific NanoDrop 2000 spectrophotometer (USA). The spectral display shows data for the samples normalized to a 1 mm path. The absorption spectra for the free RBP4, free RBA and RBA–RBP4 complexes were recorded. Circular dichroism (CD) spectra of the RBP4 and RBA–RBP4 complexes were recorded with Aviv Model 215 Circular Dichroism Spectrometer (Lakewood, NJ) at 25 °C, using rectangular quartz cells with a path length of 1 mm for far-UV spectra. Three scans with a scan speed of 50 nm/min were performed. The spectrum of buffer solution was taken and subtracted from the spectra of RBP4 and RBA–RBP4 complexes. Intrinsic fluorescence spectra of the free RBP4, free RBA and RBA–RBP4 complexes were measured at 25 °C using a Varian Cary Eclipse fluorescence spectrophotometer with the excitation and emission slit widths of 5 nm. Fluorescence emission from Trp was measured using excitation at 295 nm to avoid the contribution of tyrosines. In intrinsic fluorescence studies, the concentration of protein was 200 μM in 0.2 M Saline phosphate buffer, pH 7.0.

Organic & Biomolecular Chemistry Accepted Manuscript

Results and Discussions

RBP4 is a useful biomarker to monitor insulin resistance diagnosis and also as a predictor of T2D risk. Lee et al. have reported a novel ssDNA aptamer (RBA) that has improved conventional methods for RBP4 detection and measurements. The reported aptamer specifically binds to RBP4 with a high specificity and affinity. They have analyzed RBA sequence and predicted the corresponding secondary structure with the application of web server based mfold tool that shows typical stems and loop motives, which are probably the binding regions for the target protein (Scheme1, *Supplementary data, Figure S1*). But they didn't study nor reported the modes through which RBA and RBP4 interact. Therefore no explanations have been suggested for RBP4-RBA specific binding. Three-dimensional structure of RBA was prepared based on the secondary by Lee. Et.al. that suggests two G-rich regions, ACGGT and AGGGG at bases 19-23 and 26-32 respectively within the aptamer sequence which are predicted to form G-quartet structures . As a matter of fact, RBA structure was designed through the method mentioned above (computational methodology). Then the obtained energy minimized structure went through classical molecular dynamic simulation(s) for 50 ns. The root mean square deviations (RMSD) plot of the aptamer backbone structure during the simulations time along with the changes in the aptamer conformation are graphed (Figure 1a) and visualized (*Supplementary data, Figure S2*). A considerable rearrangement of stem participant nucleic acids of loop2 was observed during the first 10 ns of MDs. The nucleic acid G42 that is pairing C36 is replaced by G43 and consequently G43 that is pairing C35 is replaced by G44 (Figure 2). Although the reorder does not make significant changes in the general structure of RBA (Table 1).

The aptamers need to be folded in order to bind to proteins, and interaction with proteins restrain the folded structures[38]. Also ligand/protein binding to an aptamers induces rearrangements in the

Organic & Biomolecular Chemistry Accepted Manuscript

aptamer folding towards more compact structures[39] .Therefore, a structure with the least radius of gyration and most compactness (*supplementary data, Figure S3* and named **RBA-a**) was chosen as one possible conformation of the target ssDNA to interact with RBP4. Additionally, a second conformation of the aptamer obtained after 50 ns of simulations (named **RBA-b**) was considered as a structure with higher radius of gyration in comparison with the former one to continue the Docking studies with, for the further evaluations of RBP4-RBA probable modes of binding. Actually the difference between the two chosen conformations (**RBA-a** and **RBA-b**) is mainly in the 5′ end which is very flexible. As a matter of fact, both conformations were Docked over RBP4 to see if there is any significant difference in the ways the two conformations interact with RBP4.

HADDOCK docking program web server was employed to predict an appropriate mode of RBA-RBP4 complex interaction to be later used as an input for the following molecular dynamic simulations to take the dynamic nature of MD simulations as well as solvent effects into account. Since no evidences were at hand on RBA modes of interactions with RBA4, BindN webserver [40] (http://bioinformatics.ksu.edu/bindn/) was used to get a better perception of residues (like surface ones, and those constructing the binding pocket) of RBP4 that are liable of binding with the aptamer. The web-based tool efficiently predict DNA/RNA binding sites in amino acid sequences based on three features, including the side chain pKa value, hydrophobicity index and molecular mass of an amino acid with 69.40% sensitivity and 70.47% specificity. Adaptive Poisson-Boltzmann Solver (APBS) software [41] has further confirmed BindN webserver results by evaluating the electrostatic properties of the nanoscale bio molecular system to predict the residues capable of binding DNA/RNA based on the protein side chain pKa values. According to the obtained results (*supplementary data, Figure S4*), the surface of the protein is arginine rich

Organic & Biomolecular Chemistry Accepted Manuscript

(12 out 30 surface residues) which definitely enhance the probability of hydrogen bond and salt-bridge formations as well as electrostatic interactions with phosphate deoxyribose backbone of the aptamer. The obtained residues as well as those reported as the active pocket residues in the literature [42] were defined as the residues capable of interacting with RBA to HADDOCK. In addition to the double strands bases, the single stranded bases of RBA which were not expected to participate in interactions with RBA4 were also defined as passive bases for HADDOCK web server.

HADDOCK calculations of **RBA-a** over RBP4 resulted in nine clusters, each containing four PDB files (*Supplementary data, Figure S5*). RBP4-(**RBA-a**) complexes modes of interactions were evaluated and the participant residues as well as nucleic acids were studied. Residues Arg2, Arg5, Arg60, Arg62, Leu64, Arg66, Trp67, Asp68, Glu81, Lys99, Asn101, Phe135, Arg163 and Arg166 are the key residues that conduct interactions with RBA through nucleic bases T11, A12, T13, T14, T18, T19, G30, G45, T46, T49, T50, T54, T57, and T63 (*Supplementary data, Figure S7-a*). The majority of residues RBA interacts with are those of surface residues that are mainly located in loops and turns structures with enough flexibility. The interacting residues and nucleic acids were observed to be very similar in the obtained clusters. The best mode of RBP4-RBA structure was then chosen among all the received ones based on its HADDOCK score which is a weighted sum of the four terms of Electrostatic energy (weight 0.2), Van der Waals energy (weight 1.0), Desolvation energy (weight 1.0), and Restraints violation energies, (distance, SANI, weight 0.1) that are presented in *supplementary data, Table S1*. The same procedures and analysis were repeated for those of RBA-b docked over RBP4 that resulted in seven clusters (*Supplementary data, Figure S6, Table S2*). Residues Glu1, Lys31, Arg60, Arg62, Leu64, Arg66, Trp67, Asp68, Asn124, Arg163 and Arg166 are the key residues that conduct

Organic & Biomolecular Chemistry Accepted Manuscript

interactions with RBA through nucleic bases T13, T14, C15, A16, A17, T18, T19, G30, G33, G52, and C55 (*Supplementary data, Figure S7-b*). Although the conformation of RBA is different in the two obtained structures (**RBA-a** and **RBA-b**), but their modes of interactions and the interacting residues and nucleic acids are identical. Further, as it was expected, the HADDOCK score of **RBA-a** docked over RBP4 is significantly better than that of **RBA-b** (with HADDOCK scores of -53.4 and 22.2 respectively) and more aptamer nucleic bases are participating the interactions with RBP4. Having in mind that if the observed interactions in the chosen RBP4-RBA complex structure to start molecular dynamic simulations are not consistent, they won't stay stable during the simulation time, RBP4-(**RBA-a**) complex was selected as the initial input for the subsequent MD simulations.

The backbone RMSD plots of both the protein and aptamer were plotted as a function of 50 ns of simulations time and with respect to their corresponding initial energy minimized structure (Figure 1). In order to have a clearer picture of RBA presence impact on the protein structure, 50 ns of MD simulations was performed for the individual protein as well. A comparison was done between the protein backbone root mean square deviations plots of lone RBP4 and the one in complex with RBA (Figure 1b) that shows the system containing lone RBP4 is equilibrated later with an increasing trend in RMSD value, while for that in complex with RBA, the system becomes stable after almost 10 ns of simulations and stays stable to the end of the simulation time. This observation indicates the production of promising interactions between the protein residues and nucleic bases of RBA that limits conformational changes in the protein structure. Additionally, because the backbone RMSD of RBP4 in complex with RBA is below 1.5 Å, it is concluded that no structural rearrangements involving refolding of a helix into a loop or a β-sheet to a turn occurs as a result of RBA binding. Considering RBA, noticeable ups and downs

Organic & Biomolecular Chemistry Accepted Manuscript

are observable in its RMSD plot that features the very flexible nature of single strand segments of the ss DNA, especially in the 5′ end. While in that of RBA in complex with RBP4, no considerable deviations are seen in the ssDNA backbone. Those Negligible deviations which are seen in the first nanoseconds of the simulations is due to the dynamic nature of the method which helps RBA to adapt a conformation in which it can best conduct favorable interactions with RBP4.

The root-mean-square fluctuations (RMSF) of the protein residues, especially those predicted to have an active role in RBA-RBP4 efficient binding, were plotted for the lone RBP4 and that in complex with RBA to better monitor their displacement and provide information on the interacting residues that affect the protein secondary structure changes. Considering Figure 3, large fluctuations at Arg2, Arg5, Asn66, and Arg163 that are located in loop and turn structures has decreased significantly which can be referred to their interplaying with RBA nucleotide bases that reduce pliability and therefore induce stability to RBP4 anatomy.

Also, the notable degree of fluctuations at Arg60, Arg62, and Arg166 residues that despite their presence in β-sheet formation are high (as a result of their being located in the protein surface area and water accessibility in the lone protein), have diminished as a result of their interactions with RBA bases. To sum up, majority of RBP4 residues show smaller degree of fluctuations in binding with RBA than those of the free RBP4 which is as a result of their interactions with the RBA bases as we can suggest that these interactions completely make RBP4 rigid. The RMSF plots of the lone RBA and that in complex with RBP4 were also plotted to investigate the nucleic bases fluctuations as well and as it was expected a considerable decrease is observed in the nucleic bases fluctuations in complex with retinol binding protein 4 (*supplementary data, Figure S8*). Radius of gyration (Rg) studies were performed as an indicator of structural changes occurs

Organic & Biomolecular Chemistry Accepted Manuscript

due to RBP4-RBA complex formation. As it is shown in s*upplementary data, Figure S3,* complex formation between RBP4 and RBA sustains stability and implies compactness into macromolecules structures in comparison to those of the lone aptamer and protein radius of gyrations.

The most efficient interactions observed during 50 ns molecular dynamic simulations of RBP4 in complex with RBA are displayed in Figure 4 (The evolution of RBP4-RBA interactions during the MD simulation time are shown every 10ns in Figures 9, 10, 11, 12, and 13 of *Supplementary data*, *section RBP4 and RBP interactions during 50 ns of MD simulations*). Also the hydrogen bond participant residues and bases and other types of nonbonding interactions during 50 ns of MD simulations are listed in Tables 2 and 3. The hydrogen binding percent of occurrence was also calculated as is presented in Table4. Accordingly, the interactive residues are arginine, asparagine and lysine which are polar and positively charged. Stacking of planar (sp2) resonance-stabilized chemical groups with nucleotide bases often involves non-canonical groups such as main chain peptide linkage and the side chains of asparagine and arginine residues.

Arginine residues are frequently numerous at ssDNA-protein interfaces, probably because these residues have several chemical characters suitable for interaction with ssDNA. The planar nature of this group makes it able to stack (arginine-base stacking as a type of cation–π interaction) with planar groups. The positively charged guanidinium group established favorable electrostatic field for attracting negatively charged poly anions and can form a hydrogen-bonded salt bridge when placed in close proximity with phosphate groups. The multivalent hydrogen-bonding capacity of arginine residues allows hydrogen bond formation either with base groups or with other protein moieties. Hydrogen bond formation strongly depends on geometry and requires a particular match of donor and acceptor groups. This type of electrostatic interaction establish more

Organic & Biomolecular Chemistry Accepted Manuscript

exacting recognition specifies than the other interaction types and ssDNA-binding proteins incorporate a rich and varied assortment of hydrogen bonds for recognition of phosphate, sugar and base groups. The high percent of hydrogen bond occurrence (Table 4) reveals firm interactions between the two macromolecules. The sequence-specific proteins use polar residues, especially asparagine, aspartate, glutamine, glutamate and arginine residues, to closely examine the Watson-Crick edges of bases in single-stranded nucleic acid ligands[43, 44]. Water molecules have a crucial role in DNA-protein binding interactions by providing an extension to side-chains to accomplish hydrogen bonding. These water molecules bridge hydrogen bond donor and acceptor atoms of interaction participants when they are not close enough and/or cannot make favorable hydrogen bond interactions due to packing/structural restrictions. Furthermore, water molecules serve to buffer electrostatic repulsions between electronegative atoms of the protein and the DNA[45, 46]. In certain setting, all of these chemical characters can be operational at the same time such that the arginine side chain simultaneously makes salt-bridging, base-stacking, and hydrogen bonding interactions.

As shown in Figure 4, thymidine and then guanine bases show critical roles in RBA-RBP4 binding interaction. Multiple contacts of the TT-loops, loops consist of a pair of thymidine bases (TT), with RBP4 anchors the structure of RBA including T10-T11, T13-T14, T18-T19 and T49-T50 (Tables 2 and 3). These TT-loops final rearrangement in the grooves exposed them to RBP4 surface. This evidence could be explained by a fascinating hypothesis supported by some studies in which the high affinity binding of TT loops to the target proteins which is derived from high conformational flexibility of the pyrimidine ring in thymidine have been claimed [47, 48].

Additionally, the loops (G-rich sequences) in RBA do not participate protein binding directly but the single strands intervals between them include bases that have considerable roles in RBA-

Organic & Biomolecular Chemistry Accepted Manuscript

RBP4 binding. While some studies report the significant function of G-quadruplex regions in single strand DNAs interplay[49-51], it seems that G-quadruplex motifs can discriminate between small molecules and bigger targets to a certain extent. While small molecules perch in these motifs, for a large target like RBP4, this regions do not establish face-to-face interactions like those observed in small molecule intercalations. Further, because of the interactions between base pairs in the loop regions, they probably show lower affinity for binding with the protein residues.

Also considerable direct hydrogen binding and water mediated ones are inevitable between bases especialy bases C15, A17, G30, G45, T46, T49, C55, T57, G58, G59, and and T77 and RBP4 residues, including Arg2, Arg5, Arg60, Asn66, Asp68, Lys99, Arg 163 and Arg166 (Table 4). The guanidinium groups of residues Arg60 and Arg62 form several direct hydrogen bonds as well as salt bridges with phosphate groups of T19 and T18. Also hydrophobic interactions between Leu64 aliphatic isobutyl side chain with single methyl group of T11 is considerable. Asn66 forms direct hydrogen bonds as well of water mediated ones with the carboxyl groups on pirimidine fragments of T14 and T18 and purine ring of A17. Additionaly, indole fragment of Trp67 interacts A12 purin ring through hydrogen bond formation and π-π electrostatic interactions. Also, an extra hydrogen bond is formed between the amide tail of Trp67 and pirimidine ring of T13. Finally, water mediated interactions between the carboxylate anion group of Asp68 and phosphate group of T18 in considerable.

Lys30 and Arg166 have hydrogen bond as well as salt bridge formations through their amino and guanidinium groups, respectively with the phosphate group of G30. Arg163 and Arg2 which were predicted as surface residues that would have interaction with the aptamer interact G45 and T46 by composing hydrogen bonds as well as salt bridges with the aid of guanidinium and

Organic & Biomolecular Chemistry Accepted Manuscript

phosphate groups, respectively. Residue Arg5 is stabilized in its position by forming hydrogen bonds and salt bridges with pirimidine fragments carboxyl groups of T49, T50, and G53 base phosphate group. Lys 99, Gly100 and Asn101 form hydrogen bonds as well as salt bridges like other mentioned residues with bases T63 and T77.

RBP4 normally binds to the transthyretin, forming a protein complex that reduces renal clearance of RBP4[52]. One point to be attentioned more to is the presence of residues Leu64, Trp67, Asn68 and Lys99 that construct the active pocket of RBP4 that ascertain the inhibition activity of the studied aptamer for the RBP4 interaction with transthyretin. In addition, the sphatial forbidance as a result of the RBP4-RBA adjency will further prevent the substrate attendance to the binding pocket and as a consequence result in a decreased affinity of RBP4 for transthyretin [42].

The protein secondary structure was investigated through DSSP analysis (database of secondary structure assignment) to attend conformational changes dictated as a result of RBA interactions and is presented in Figure 5. Since proteins unfold to some extent in water, a non-significant unfolding is observed in the lone protein secondary structure in water is acceptable (Figure 5a). According to Figure 5b, a notable decrease is observed in α-helix and 3-helix as well as β-sheet structures participant residues from 3%, 8% and 47 % to 1%, 7% and 44%, respectively. Reznik et al. also reported helical structure content of 11% in the secondary structure of RBP4 that is close to our observed results[53] (Table 5). Same time, the percent of residues in the bend and coil structure structures increases from 10% and 20% to 14% and 23%, respectively. These observations show that a re-folding is occurring in the protein over all secondary structure as a result of RBA binding. Considering the protein secondary structure in complex with RBA, it is perceived that the aptamer interactions induces significant conformational alternations in RBP4 that get along well with the radius of gyration analysis verdicts.

Organic & Biomolecular Chemistry Accepted Manuscript

Experimental observations further confirmed the interactions in between RBA and RBP4. The UV-visible spectra of the free RBA, free RBP4 and RBA- RBP4 complexes are presented in Figure 6 that ensures the complex formation. The effect of RBA binding on RBP4 secondary structure was investigated applying far-UV-CD. In the far-UV, the peptide bond is the principal absorbing group and studies in this region can give information on the secondary structure. Since the 3_10 and α-helical conformations cannot be distinguished by CD[54], we can judge the effect of binding of RBA to RBP4 on totality of helical structures. Figure 7 shows the far-UV-CD spectra of the free RBP4 and RBP4-RBA complex. For both, the spectrum was characterized by a negative band with double minima at 208 and 222 nm, indicative of proteins with high contents of helical structure[55]. Upon binding RBA, the overall shape of the spectrum did not change significantly, but the intensity of the negative band decreased. Such a qualitative description was confirmed by spectral deconvolution using CD spectrum deconvolution software CDNN version 2[56]. Having in mind that CD analysis often enhances alpha helix contributions, the helical structure content of RBP4 was obtained to be 15.9% and 12.55% in the free RBP4 and RBP4-RBA complex, respectively, with the helical content difference of 3.35% that get along well with the obtained computational data (Table 5). As it can see in Table 5, interestingly MD predicted secondary structures are much closed to the derived results from X-Ray Crystallography of RBP4 (PDB ID: 2wq9) [23] which can indicate the accuracy of this method. Our CD results are also confirmed by the Reznik *et al.* Study[53]. But there are some difference between percentage of each secondary structures in both CD results with the results of MDs and X-Ray Crystallography. As mentioned in literature [57] these differences can be explained by this fact that although CD is an excellent method for a rapid evaluation of proteins secondary structure, it does not give the residue-specific information like those obtained by X-ray crystallography or NMR studies. Furthermore, it should be noted that electronic transitions of the bases can alter the CD spectrum of the protein by decreasing the helix content of the protein which is quite evident from both CD and MDs results (Table 5).

Organic & Biomolecular Chemistry Accepted Manuscript

RBP4 has some Trp residues among which Trp67 interacts A12 purine ring through hydrogen bond formation and π-π electrostatic interactions and also, an extra hydrogen bond between the amide tail of Trp67 and pyrimidine ring of T13. When excited at 295 nm, free RBP4 shows a typical Trp fluorescence emission spectrum, with a peak at 335 nm (Figure 8). After binding RBA, the emission spectrum undergoes a drastic decrease in intensity and a blue shift to 330 nm. The observed blue shift (5 nm) in Trp fluorescence reflects reduced accessibility of Trp to the bulk solvent. As seen in Figure 8, RBP4 exhibits a significant emission band cantered at 330 nm, which closely overlaps the emission band of Trp67. Therefore, in RBP4-RBA complex, Trp emission is strongly quenched through the energy transfer to the attached RBA molecules.

Organic & Biomolecular Chemistry Accepted Manuscript

Conclusions

Structural information of aptamers and elucidation of detailed interactions between them and their targets will enhance our understanding of molecular and structural basis of protein–DNA interaction at the atomic level in general that facilitates effective design and improvement of DNA aptamers for therapeutic applications.

Also, the advent of computational studies has arisen the opportunity of predicting/monitoring mechanisms of aptamer-target recognition/interaction in a way that cannot be evident from static 3D-structures and or experimental investigations. Aptamers have a limited number of interactions to make with a protein target and therefore aptamers that 'fit' into a crevice on a protein, such as an active site, are more likely to be selected [57]. Modification of RBP4 activity (the ability of RBP4 to bind transthyretin, stability (e.g. structural or half-life) of RBP4 in tissues or in circulation) is a target for methods of treatment for insulin resistance and for conditions related to insulin resistance[52].

In the present study, a three dimensional model of RBA was constructed and docked over RBP4. MDs were carried out on RBP4-RBA complex in order to explore the interactions involved their binding and the participant residues of the protein and the aptamer nucleotide bases. The MDs results show that polar and positively charged residues including arginine, asparagine and lysine from RBP4 as well as thymine nucleic acid from RBA are the key residues leading RBP4-RBA binding interactions. According to the obtained results, the RBP4-RBA interaction model is considered as a combination of recognition (induced fit) and preformed (lock and key) models. This study has also showed that RBA binding to RBP4 probably inhibits RBP4 interactions with transthyretin by detention of RBP4 active pocket residues. The RBP4-RBA complex formation was ensured experimentally and with the application of UV–visible, circular dichroism, and fluorescence spectroscopy measurements. Since modification of RBP4 activity (the ability of

Organic & Biomolecular Chemistry Accepted Manuscript

RBP4 to bind transthyretin, stability (e.g. structural or half-life) of RBP4 in tissues or in circulation) is a target for methods to treat insulin resistance, this ability of the studied aptamer can be considered as a competence. Eventually, we can conclude that in the interactions in between long ssDNA oligoes and big targets, single strand located bases are those conducting the intercalation phenomenon rather than the ones involving loops and folded motifs to encompass targets and probably inhibit their catalytic reaction (if possible). As a future perspective, the inhibition potency of the studied aptamer in forfending transthyretin -RBP4 interaction can be used as a potential diagnostic and therapeutic agent to reduce insulin resistance effects of RBP4 in high risk individuals for type two diabetes and also other associated diseases and syndromes such as metabolic syndromes and cardiovascular diseases.

Acknowledgment

The financial support of the Research Council of the Tehran University of Medical Sciences is gratefully acknowledged.

References

1. F. B. Hu, *Diabetes Care*, 2011, **34**, 1249-1257.
2. M. Rocha, C. Banuls, L. Bellod, S. Rovira-Llopis, C. Morillas, E. Sola, V. M. Victor and A. Hernandez-Mijares, *PLoS One*, 2013, **8**, e78670.
3. D. M. Muoio and C. B. Newgard, *Nature*, 2005, **436**, 337-338.
4. T. E. Graham, C. J. Wason, M. Bluher and B. B. Kahn, *Diabetologia*, 2007, **50**, 814-823.
5. S. D. Jayasena, *Clin Chem*, 1999, **45**, 1628-1650.
6. K. A. Marshall and A. D. Ellington, *Methods Enzymol*, 2000, **318**, 193-214.
7. C. H. Lin and D. J. Patel, *Chem Biol*, 1997, **4**, 817-832.
8. S. J. Lee, B. S. Youn, J. W. Park, J. H. Niazi, Y. S. Kim and M. B. Gu, *Anal Chem*, 2008, **80**, 2867-2873.
9. W. James, *Curr Opin Pharmacol*, 2001, **1**, 540-546.
10. T. Hermann and E. Westhof, *J Med Chem*, 1999, **42**, 1250-1261.
11. T. Hermann and D. J. Patel, *Science*, 2000, **287**, 820-825.
12. J. A. Cowan, T. Ohyama, D. Wang and K. Natarajan, *Nucleic Acids Res*, 2000, **28**, 2935-2942.

Organic & Biomolecular Chemistry Accepted Manuscript

13. D. S. Shcherbinin and A. V. Veselovslii, *Biofizika*, 2013, **58**, 415-424.
14. P. H. Lin, C. W. Tsai, J. W. Wu, R. C. Ruaan and W. Y. Chen, *Biotechnol J*, 2012, **7**, 1367-1375.
15. Z. Gong, Y. Zhao, C. Chen and Y. Xiao, *J Biomol Struct Dyn*, 2011, **29**, 403-416.
16. R. V. Reshetnikov, A. V. Golovin and A. M. Kopylov, *Biochemistry (Mosc)*, 2010, **75**, 1017-1024.
17. A. Villa, J. Wohnert and G. Stock, *Nucleic Acids Res*, 2009, **37**, 4774-4786.
18. P. Jayapal, G. Mayer, A. Heckel and F. Wennmohs, *J Struct Biol*, 2009, **166**, 241-250.
19. C. Schneider and J. Suhnel, *Biopolymers*, 1999, **50**, 287-302.
20. A. Perez, I. Marchan, D. Svozil, J. Sponer, T. E. Cheatham, 3rd, C. A. Laughton and M. Orozco, *Biophys J*, 2007, **92**, 3817-3829.
21. F. Pitici, D. L. Beveridge and A. M. Baranger, *Biopolymers*, 2002, **65**, 424-435.
22. B. E. Eaton, *Current opinion in chemical biology*, 1997, **1**, 10-16.
23. M. Nanao and T. J. Stout, Crystal structure of rbp4 bound to oleic acid : pdb2wq9, http://www.rcsb.org/pdb/explore.do?structureId=2wq9, DOI: 10.2210/pdb2wq9/pdb).
24. N. Guex and M. C. Peitsch, *Electrophoresis*, 1997, **18**, 2714-2723.
25. M. Rostkowski, M. H. Olsson, C. R. Sondergaard and J. H. Jensen, *BMC Struct Biol*, 2011, **11**, 6.
26. T. Fujita, *Hepatology*, 2008, **48**, 1725; author reply 1725-1726.
27. M. Froimowitz, *Biotechniques*, 1993, **14**, 1010-1013.
28. S. Souza, *Advances in Pharmaceutics*, 2014, **2014**, 12.
29. M. Zuker, *Nucleic acids research*, 2003, **31**, 3406-3415.
30. C. Dominguez, R. Boelens and A. M. Bonvin, *J Am Chem Soc*, 2003, **125**, 1731-1737.
31. S. J. de Vries, A. D. van Dijk, M. Krzeminski, M. van Dijk, A. Thureau, V. Hsu, T. Wassenaar and A. M. Bonvin, *Proteins*, 2007, **69**, 726-733.
32. S. Pronk, S. Pall, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess and E. Lindahl, *Bioinformatics*, 2013, **29**, 845-854.
33. A. E. Aliev, M. Kulke, H. S. Khaneja, V. Chudasama, T. D. Sheppard and R. M. Lanigan, *Proteins*, 2014, **82**, 195-215.
34. W. Humphrey, A. Dalke and K. Schulten, *J Mol Graph*, 1996, **14**, 33-38, 27-38.
35. T. Makarewicz and R. Kazmierkiewicz, *J Chem Inf Model*, 2013, **53**, 1229-1234.
36. L. Schrödinger, *(Schrödinger,. LLC, New York)*, 2010.
37. E. Vigmond, http://plasma-gate.weizmann.ac.il/Grace/).
38. J. Ashby, S. Schachermeyer, Y. Duan, L. A. Jimenez and W. Zhong, *Journal of Chromatography A*, 2014, **1358**, 217-224.
39. C. D. Stoddard, R. K. Montange, S. P. Hennelly, R. P. Rambo, K. Y. Sanbonmatsu and R. T. Batey, *Structure*, 2010, **18**, 787-797.
40. L. Wang and S. J. Brown, *Nucleic Acids Res*, 2006, **34**, W243-248.
41. S. Unni, Y. Huang, R. M. Hanson, M. Tobias, S. Krishnan, W. W. Li, J. E. Nielsen and N. A. Baker, *J Comput Chem*, 2011, **32**, 1488-1491.
42. H. M. Naylor and M. E. Newcomer, *Biochemistry*, 1999, **38**, 2647-2653.
43. S. Jones, D. T. Daley, N. M. Luscombe, H. M. Berman and J. M. Thornton, *Nucleic acids research*, 2001, **29**, 943-954.
44. C. C. C. Phoebe A. Rice, *Protein-nucleic Acid Interactions: Structural Biology*, RSC Publishing, 2008.

Organic & Biomolecular Chemistry Accepted Manuscript

45. C. K. Reddy, A. Das and B. Jayaram, *J Mol Biol*, 2001, **314**, 619-632.
46. S. Roy and A. R. Thakur, *Journal of biomolecular structure & dynamics*, 2010, **27**, 443-456.
47. A. M. Varizhuk, V. B. Tsvetkov, O. N. Tatarinova, D. N. Kaluzhny, V. L. Florentiev, E. N. Timofeev, A. K. Shchyolkina, O. F. Borisova, I. P. Smirnov, S. L. Grokhovsky, A. V. Aseychev and G. E. Pozmogova, *Eur J Med Chem*, 2013, **67**, 90-97.
48. M. Scuotto, M. Persico, M. Bucci, V. Vellecco, N. Borbone, E. Morelli, G. Oliviero, E. Novellino, G. Piccialli, G. Cirino, M. Varra, C. Fattorusso and L. Mayol, *Org Biomol Chem*, 2014, **12**, 5235-5242.
49. A. Verdian-Doghaei, M. R. Housaindokht, M. R. Bozorgmehr and K. Abnous, *Journal of biomolecular structure & dynamics*, 2014, DOI: 10.1080/07391102.2014.935482, 1-11.
50. M. Marusic, R. N. Veedu, J. Wengel and J. Plavec, *Nucleic acids research*, 2013, **41**, 9524-9536.
51. A. De Rache, T. Doneux, I. Kejnovska and C. Buess-Herman, *J Inorg Biochem*, 2013, **126**, 84-90.
52. N. Mody, T. E. Graham, Y. Tsuji, Q. Yang and B. B. Kahn, *Am J Physiol Endocrinol Metab*, 2008, **294**, E785-793.
53. G. O. Reznik, Y. Yu, G. E. Tarr and C. R. Cantor, *J Proteome Res*, 2003, **2**, 243-248.
54. T. S. Sudha, E. K. Vijayakumar and P. Balaram, *Int J Pept Protein Res*, 1983, **22**, 464-468.
55. N. Mogharrab, H. Ghourchian and M. Amininasab, *Biophys J*, 2007, **92**, 1192-1203.
56. G. Bohm, R. Muhr and R. Jaenicke, *Protein Eng*, 1992, **5**, 191-195.
57. A. D. Keefe, S. Pai and A. Ellington, *Nat Rev Drug Discov*, 2010, **9**, 537-550.

Organic & Biomolecular Chemistry Accepted Manuscript

Legend of the Tables

Table 1. H-bonds monitored in loop regions of the free RBA during 50ns and during its interactions with RBP4 (for 50 ns).

Table 2. Evolution of the hydrogen bond interactions in between RBP4 and RBA during 50ns of MD simulations**.**

Table 3.Evolution of the non-bonded interactions (except hydrogen bonds) in between RBP4 and RBA during 50ns of MD simulations.

Table 4. Hydrogen bond occupancy during 50ns of RBP4-RBA complex MDs. The criterion for hydrogen bonding was set at ≤ 3.5 Å distance between electron donor atom and hydrogen of electron acceptor atom with 120-degree angle cut-off.

Table 5. Comparison between the percent of residues present in different secondary structures of the free RBP4 and RBP4 in complex with RBA.

Organic & Biomolecular Chemistry Accepted Manuscript

Table 1.

Nucleic acid HB* Donor	Nucleic acid HB Acceptor	Free RBA (0-10 ns) A°	Free RBA (10-50 ns) A°	During 50 ns interacting with RBP4 A°
loop1				
A20:H62	**T27**: O4	2.13 ± (0.37)	2.01± (0.23)	1.97 ± (0.15)
T27:H3	**A20**: N1	1.99 ± (0.15)	1.97± (0.13)	1.94 ± (0.10)
G26:H1	**C21**:N3	2.03 ± (0.19)	2.08± (0.43)	1.96 ± (0.09)
G26:H21	**C21**:O2	1.98 ± (0.26)	2.15± (0.83)	1.88 ± (0.12)
C21:H42	**G26**: O6	2.13 ± (0.54)	1.99± (0.20)	1.96 ± (0.16)
loop2				
G43:H1	**C35**:N3	2.20 ± (0.71)	4.54± (0.77)	4.93 ± (0.59)
G43:H21	**C35**:O2	2.19 ± (0.84)	5.13± (0.55)	5.62 ± (0.54)
C35:H42	**G43**: O6	2.16 ± (0.63)	3.71± (1.15)	4.11 ± (0.64)
G44:H1	**C35**:N3	1.36 ± (0.79)	2.00± (0.13)	1.99 ± (0.10)
G44:H21	**C35**:O2	1.32 ± (0.789)	1.93 ± (0.15)	1.87 ± (0.13)
C35:H42	**G44**:O6	1.36 ± (0.74)	2.28 ± (0.62)	1.99 ± (0.18)
G43:H1	**C36**:N3	3.44 ± (0.58)	2.30 ± (0.56)	2.00 ± (0.12)
G43:H21	**C36**:O2	3.44 ± (0.66)	1.89 ± (0.21)	1.88 ± (0.12)
C36:H42	**G43**:O6	3.79 ± (0.88)	2.76 ± (1.19)	2.03 ± (0.24)
G42:H1	**C36**:N3	2.04 ± (0.31)	3.08 ± (0.87)	2.84 ± (0.34)
G42:H21	**C36**:O2	1.96 ± (0.35)	3.24 ± (0.94)	2.74 ± (0.50)
C36:H42	**G42**:O6	2.09 ± (0.43)	3.26 ± (0.87)	3.04 ± (0.38)
loop3				
A60:H62	**T77**:O4	2.10 ± (0.39)	2.06 ± (0.24)	1.95 ± (0.15)
T77:H3	**A60**:N1	2.05 ± (0.37)	1.97 ± (0.12)	1.98 ± (0.11)
G61:H1	**C76**:N3	1.97 ± (0.10)	1.97 ± (0.10)	1.96 ± (0.09)
G61:H21	**C76**:O2	1.90 ± (0.13)	1.91 ± (0.14)	2.85 ± (0.23)
C76:H42	**G61**:O6	1.97 ± (0.17)	1.96 ± (0.17)	1.92 ± (0.14)
A62:H62	**T75**:O4	2.04 ± (0.25)	2.06 ± (0.23)	2.00 ± (0.18)
T75:H3	**A62**:N1	2.01 ± (0.23)	1.96 ± (0.12)	1.97 ± (0.11)
A74:H62	**T63**:O4	2.11 ± (0.49)	2.11 ± (0.42)	2.09 ± (0.20)
T63:H3	**A74**:N1	2.04 ± (0.46)	2.01 ± (0.23)	1.93 ± (0.10)

*HB: Hydrogen bond

Organic & Biomolecular Chemistry Accepted Manuscript

Hydrogen Bond Interactions After 10ns of MDs	Hydrogen Bond Interactions After 20ns of MDs	Hydrogen Bond Interactions After 30ns of MDs	Hydrogen Bond Interactions After 40ns of MDs	Hydrogen Bond Interactions After 50ns of MDs

Table 2.

Organic & Biomolecular Chemistry Accepted Manuscript

Don.	Acc.	Dist.	Don.	Acc.	Dist.	Don.	Acc.	Dist.	Don.	Acc.	Dist.	Don.	Acc.	Dist.
ASN66:ND2	DC15:O2	2.6	ASN65:ND2	DT14:O_4	2.8	TRP67:NE1	DT11:O4	2.8	DA2:N6	TYR173: O	2.9	GLY172:N	DA2:O3'	2.8
ASN66:ND2	DA16:O4'	2.8	ASN66:ND2	DC15:O_2	2.8	ASN66:ND2	DC15:O2	2.8	DA12:N6	ASP68: O	2.7	GLY172:N	DT3:O2P	2.8
ARG163:NH2	DG28:O1P	2.8	ASN171:ND2	DA17:O_1P	2.9	ARG60:NH1	DA17:O2P	2.6	ARG166:NH2	DT19:O1P	2.7	DT11:N3	TYR173: O	2.8
ARG153:NH1	DG28:O2P	2.9	ARG60:NH1	DA17:O_1P	2.7	ARG62:NH1	DT18:O2P	2.9	ARG153:NH2	DT27:O1P	2.9	DA12:N6	ASP68: O	2.8
ARG163:NE	DG28:O5'	2.7	ARG166:NH2	DT19:O_1P	2.9	ARG166:NH2	DT19:O1P	2.7	ARG163:NH2	DG37:O2P	2.7	DA12:N1	TYR173:OH	2.7
ARG166:NH2	DG30:O2P	2.7	ARG166:NE	DA29:O_3'	2.9	ARG166:NH1	DG30:O2P	2.9	ARG163:NH1	DG37:O5'	2.9	DA16:N6	ASN65:OD1	2.8
GLU1:N	DT46:O1P	2.8	ARG166:NH2	DG30:O_2P	3.0	ARG163:NH1	DG30:O6	2.6	DA47:N6	GLU1:OE2	2.9	ARG60:NH1	DA17:O2P	2.8
ARG2:NH1	DA47:O1P	2. 9	ARG163:NH2	DG37:O_2P	2.9	ARG163:NH1	DG37:O2P	2.7	DG48:N1	GLU1:OE2	2.8	ARG60:NH2	DA17:O2P	2.7
ARG5:NH1	DT49:O4	2. 8	GLU1:N	DT46:O_1P	2.7	ARG163:NH2	DG37:O2P	2.7	DG48:N2	GLU1:OE1	2.7	ASN171:ND2	DA17:O2P	2.7
ARG5:NH2	DT50:O4	2.8	ARG2:N	DT46:O1P	2.9	ARG2:NE	DT46:O2P	2.7	ARG5:NH1	DT49:O4	2.8	ARG166:NH1	DA29:O3'	2.8
DG56:N1	ASP03:OD1	2.8	ARG2:NE	DT46:O2P	2.7	ARG2:NH2	DT46:O2P	2.9	SER7:OG	DG52:O1P	2.6	ARG163:NH1	DG30:O6	2.9
LYS87:NZ	DT57:O2	2.7	ARG2:NH2	DT46:O2P	2.9	DA47:N6	GLU1:OE1	2.9	ARG5:NE	DG53:O1P	2.7	ARG163:NH2	DG37:O2P	2.8
THR78:OG1	DG58:O2P	2.6	DA47:N6	GLU1:OE1	2.8	DG48:N1	GLU1:OE2	2.8	GLU81:N	DT57:O2P	2.8	ARG2:N	DT46:O_1P	2.9
-	-	-	DG48:N1	GLU1:OE2	2.7	DG48:N2	GLU1:OE2	2.9	LYS87:NZ	DT57:O2	2.8	ARG2:NH1	DT46:O2P	2.8
-	-	-	ARG5:NH1	DT49:O4	2.7	ARG5:NH1	DT49:O4	3.0	LYS99:NZ	DA62:O2P	2.8	DA47:N6	GLU1:OE1	2.8
-	-	-	ARG5:NH2	DT49:O4	2.7	ARG5:NH2	DG53:O1P	2.8	-	-	-	ARG5:NH2	DT49:O4	2.9
-	-	-	ARG5:NH2	DG53:O1P	2.9	GLU81:N	DT57:O2P	2.8	-	-	-	ARG5:NH2	DG53:O1P	2.7
-	-	-	DC55:N4	ASP103:OD1	2.8	ASP82:N	DT57:O2P	2.9	-	-	-	DG53:N7	SER7:OG	2.7
-	-	-	LYS87:NZ	DT57:O2	2.8	LYS87:NZ	DT57:O2	2.8	-	-	-	LYS85:NZ	DC55:O2	2.8
-	-	-	THR78:OG1	DG58:O2P	2.8	ASN101:ND2	DG58:O3'	2.9	-	-	-	GLU81:N	DT57:O2P	2.8
-	-	-	THR76:OG1	DG59:O2P	2.8	LYS99:NZ	DG61:O3'	3.0	-	-	-	LYS87:NZ	DT57:O2	2.8
-	-	-	SER95:OG	DA64:O1P	2.6	LYS99:NZ	DA62:O1P	2.8	-	-	-	THR78:OG1	DG58:O2P	2.5
-	-	-	LYS89:NZ	DT77:O2	2.89	ARG121:NH2	DT77:O1P	2.9	-	-	-	ASN101:ND2	DG58:O3'	2.8
-	-	-	-	-	-	ARG:121:NE	DT77:O2P	2.8	-	-	-	LYS89:NZ	DG59:O2P	2.7
-	-	-	-	-	-	ARG121:NE	DT77:O5'	2.9	-	-	-	LYS99:NZ	DA62:O1P	2.8
-	-	-	-	-	-	ASN101:N	DT77:O3'	2.9	-	-	-	ARG121:NH2	DT77:O1P	2.8
-	-	-	-	-	-	-	-	-	-	-	-	ASN101:N	DT77:O3'	2.8

*Don.; stands for donor groups, Acc.; stands for acceptor groups and Dist.; stands for distance which are in Angstroms.

Organic & Biomolecular Chemistry Accepted Manuscript

Table 3.

Non-bonded interactions After 10ns of MDs			Non-bonded interactions After 20ns of MDs			Non-bonded interactions After 30ns of MDs			Non-bonded interactions After 40ns of MDs			Non-bonded interactions After 50ns of MDs		
Prot.	**DNA**	**Dist.**	**Prot.**	**DNA**	**Dist.**	**Prot.**	**DNA**	**Dist.**	**protein**	**DNA**	**Dist.**	**protein**	**DNA**	**Dist.**
ARG62:NE	DA17:O2P	3.3	ARG62:NH1	DA17:O3'	3.0	LEU167:CD1	DG23:N7	3.3	ASN66: O	DA12:C2	3.1	TYR25:CE2	DA2:N1	3.3
GLN164:NE2	DG31:N7	3.0	GLN38:NE2	DT18:O2P	3.2	ARG153:NH2	DG23:O6	3.0	ARG60:NH1	DA17:O2P	3.0	TYR25:CD2	DA2:C2	3.4
ASP126:CB	DT54:O5'	3.2	ARG62:NH1	DT18:O2P	3.2	TYR165: O	DA29:C5'	3.3	LYS85:NZ	DT54:O4	3.2	CYS70:SG	DT11:O4	2.9
ASP126: O	DT54:N1	3.1	ARG153:NH2	DT27:O1P	3.3	ASP126: O	DG53:C2'	3.1	LEU125: O	DC55:N4	3.0	TRP67:CH2	DA12:N7	3.2
ASP126: O	DT54:C6	3.2	GLN164:OE1	DG30:N7	3.3	LYS85:NZ	DT54:O4	3.3	LYS85:NZ	DC55:N3	2.9	TRP67:CZ3	DA12:C5	3.2
ASP126: O	DT54:C5	3.2	ASP126:C	DT54:O4'	3.1	LYS85:NZ	DC55:N3	2.9	LYS85:NZ	DC55:O2	3.1	TRP67:CE2	DA12:N6	3.3
ASP126: O	DT54:C4	3.4	ASP126: O	DT54:O4'	3.2	LYS85:NZ	DC55:O2	3.4	LEU125:CD1	DG56:O6	3.3	ASN66:ND2	DT14:C6	3.3
ASP126: O	DT54:C2	3.3	GLY127:CA	DT54:C2	3.4	THR78:CG2	DT57:O3'	3.2	THR78:CG2	DT57:O3'	3.1	LYS29:NZ	DT18:O2P	3.1
ASN101:ND2	DG58:O4'	3.0	LEU123:CD1	DC55:N4	3.3	THR78:OG1	DG58:C5'	3.0	THR78:OG1	DG58:C5'	2.9	ASP126: O	DG53:C2'	3.1
ASN101:OD1	DG58:N3	3.1	LYS85:NZ	DC55:O2	3.3	ASP102:OD1	DT77:C3'	3.1	THR78:OG1	DG58:C4'	3.2	THR80:CB	DT57:O2P	3.3
LYS89:NZ	DG59:O2P	3.1	GLU81:N	DT57:O2P	3.2	ASP102:CG	DT77:O3'	3.2	LYS89:CE	DG59:O2P	3.3	TRP91:CH2	DA62:O1P	3.3
LYS99:NZ	DT63:O1P	3.0	ASN101:OD1	DG58:C1'	3.3	ASP102:OD1	DT77:O3'	2.6	LYS89:NZ	DG59:O2P	3.1	GLN98:NE2	DT63:O3'	3.1
-	-	-	ASN101:ND2	DG58:N3	3.1	-	-	-	GLN98:CB	DT63:O1P	3.3	GLN98:NE2	DA64:O2P	3.3
-	-	-	GLY100: O	DT77:O3'	2.8	-	-	-	ARG121:NH2	DT77:O1P	3.0	GLY34:CA	DT77:O1P	3.2
-	-	-	-	-	-	-	-	-	ASP102:OD1	DT77:C3'	3.2	ASP102:OD1	DT77:C3'	3.3
-	-	-	-	-	-	-	-	-	ASN101:N	DT77:O3'	3.0	GLY100:CA	DT77:O3'	3.2
-	-	-	-	-	-	-	-	-	ASP102:CG	DT77:O3'	3.1	ASP102:OD1	DT77:O3'	2.7
-	-	-	-	-	-	-	-	-	ASP102:OD1	DT77:O3'	2.5	-	-	-

*Prot.; stands for protein, Dist.; stands for distance which are in Angstroms.

Organic & Biomolecular Chemistry Accepted Manuscript

Table 4.

Donor	Acceptor	Occupancy* (%)	Donor	Acceptor	Occupancy (%)
ARG2: HE	DT46:O_2P	67.8	DG56: H21	ASP103: OD2	27.6
ARG60: H_{11}	DA17: O_2P	66.8	ARG153: H21	DG28: O2P	25.2
LYS87: HZ1	DT57: O2	62.5	DG48: H21	GLU1: OE2	24.7
ARG163: H11	DG30: O6	60.2	ARG153: H11	DG28: O2P	24.5
LYS85: HZ1	DC55: O2	59.6	LYS99: HZ1	DT63: O1P	24.4
THR78: HG	DG58: O_2P	57.9	ARG62: H11	DT18: O2P	24.1
ASN66: D21	DC15: O2	53.9	ARG60: H21	DA17: O2P	21.3
GLU81: H	DT57: O_2P	53.2	DT11: H3	TYR173: O	21.3
GLU1: H1	DT46: O1P	52.8	ARG2: H	DT46: O1P	21.0
ARG163: H21	DG37: O2P	51.5	ASN101: D21	DG58: O4'	20.8
LYS89: HZ1	DG59: O2P	48.9	DC55: H41	LEU125: O	19.6
DG48: H1	GLU1: OE2	47.7	ARG163: H11	DG37: O1P	19.5
ARG5: H21	DG53: O1P	47.2	DT50: H3	ARG5: NH2	19.4
ARG2: H21	DT46: O2P	47.2	ARG121: HE	DT77: O2P	17.9
ARG166: H11	DT19: O1P	47.1	ASN101: D21	DG58: N3	17.8
ARG166: H21	DT19: O1P	43.7	LYS99: H	DT63: O1P	17.8
DA47: H61	GLU1: OE1	43.1	ASP68: H	DA12: N1	16.4
ARG166: H11	DG30:O2P	43.0	GLN164: E21	DG30: N7	16.2
ARG5: H11	DT49:O4	41.6	ARG121: HE	DT77: O1P	15.7
ARG166: H21	DG30: O2P	40.0	LYS85: HZ1	DC55: N3	15.6
DT77: H3T	ASP102: OD1	39.8	ASN65: D21	DT14: O4	14.4
ARG5: H21	DT49: O4	38.9	ARG153: H21	DG28: O1P	13.6
DT77: H3T	ASN101: N	38.6	DA16: H61	ASN65: OD1	13.2
ARG5: HE	DG53: O1P	37.8	ARG163: H11	DG37: O5'	13.1
ARG121: H21	DT77: O1P	35.3	ARG153: HE	DG28: O1P	12.6
DG48: H1	GLU1: OE1	34.5	ARG153: H11	DG28: O1P	12.5
ASN66: H	DA16: N1	34.1	DG48: H21	GLU1: OE1	12.4
ARG166: HE	DG30: O1P	31.9	GLN98: E21	DA64: O2P	11.8
ARG163: H11	DG37: O2P	31.5	ARG5: H21	DT50: O4	11.6
DA47: H61	GLU1: OE2	31.1	ASN171: D21	DA17: O1P	11.0
ASN101: D21	DG58: O3'	30.6	SER7: HG	DG53: N7	10.7
DA12: H61	ASP68: O	29.1	LYS29: HZ1	DT18: O2P	10.6
LYS99: HZ1	DA62: O1P	28.7	ASP82: H	DT57: O2P	10.2
LYS85: HZ1	DT54: O4	27.7			

*Occupancy (%): hydrogen bond occupancy

Organic & Biomolecular Chemistry Accepted Manuscript

Table 5.

	MDs[#]						CD_1[$]			CD_2[&]			X-Ray Crys[§]		
Protein	β-Sheet	Bend	Coil	Turn	α -Helix	3-Helix	Helix	β-Sheet	Other	Helix	β-Sheet	Other	Helix	β-Sheet	Other
RBP4	47*	10	20	12	8	3	15.65	33.22	51.13	11	31	59	11	47	42
RBP4 in complex with RBA	44	13	23	11	7	1	12.55	31.15	56.45	-	-	-	-	-	-

*Numbers are percent (%) of all residues, [#] The results obtained from molecular dynamic simulation results, [$] Circular dichroism results calculated in this article by CD spectrum deconvolution software, [&] Circular Dichroism results from Reznik et al. study[53], [§]X-Ray Crystallographic structure of RBP4 (PDB ID: 2wq9) [23].

Organic & Biomolecular Chemistry Accepted Manuscript

Legends of the schemes

Scheme1. The two dimensional representative of RBA predicted via the web server based mfold tool, based on the reported sequence by Lee et al. [8]

Legend of the figures

Figure1. RMSD plots of, a. the lone RBA (Orange) and RBA in complex with RBP4 (Green), b. the lone RBP4 backbone (Magenta) and RBP4 backbone in complex with RBA (Blue), as a function of simulation time.

Figure2. Rearrangement of nucleic bases in loop 2 after 10 ns of MDs, a. After energy minimization, and b. After 10 ns of simulations.

Figure3. RMSF plot of lone RBP4 (violet) and that in complex with RBA (blue), **a**. residues 1–15, **b**. residues 55–69, and **c**. residues 160-172.

Figure4. An average structure of RBP4-RBA complex during 50ns of MD simulations. The most prominent interactions between RBP4 in complex with RBA are presented. In order to avoide complication, only direct hydrogen bonds and water mediated ones are highlighted employing yellow dashes.

Figure5. Secondary structure analysis of RBA4 in, **a**. absence and, **b**. presence of RBA. Structure is presented in black, β-sheet in green, Coil in orange, Turn in magenta, α-helics in cyan and 3-helics in violet colours.

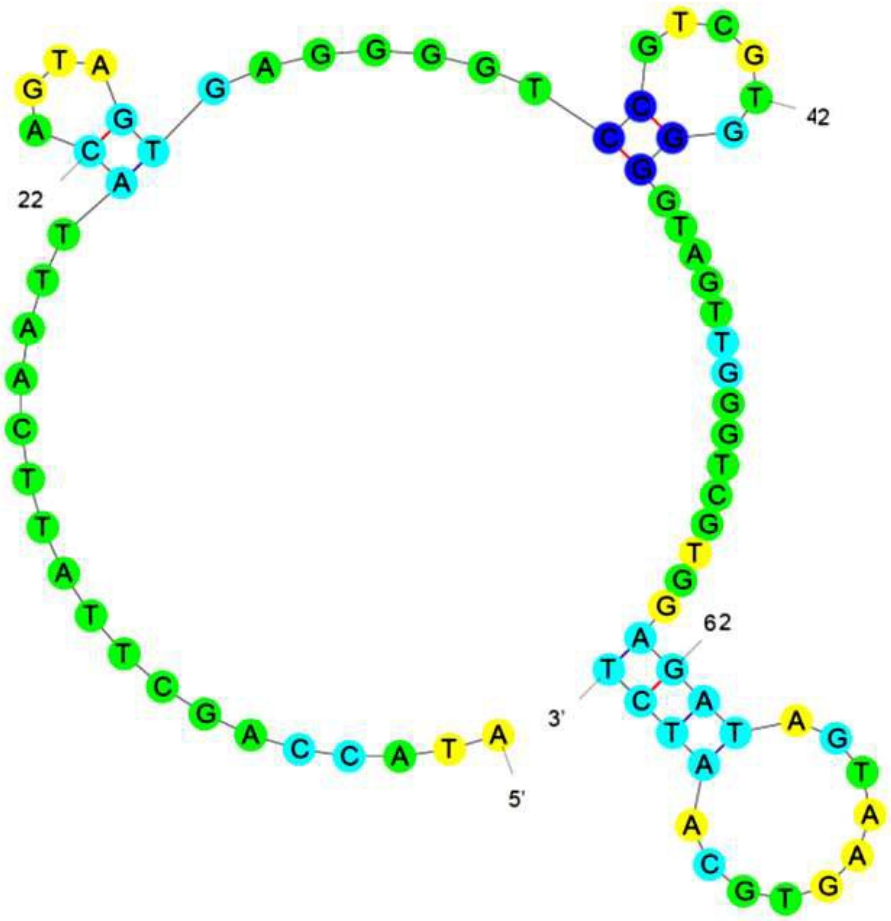
Figure 6. Ultraviolet-visible spectra of the free RBA (grey line), free RBP4 (red dashes) and RBA-RBP4 complex (blue line).

Figure 7. Far-UV CD spectra of 20 μM free RBP4 (red dashes) and RBP4-RBA complex (blue line). In these measurements, RBP4 was dissolved in a 0.2 M saline phosphate buffer (pH 7.0) and the temperature was set to 25°C.

Organic & Biomolecular Chemistry Accepted Manuscript

Figure 8. Trp emission fluorescence spectra of free RBP4 (red dashes) and RBA-RBP4 complex (blue line) in a 0.2 M saline phosphate buffer (black dots) with pH 7.0, at 25°C. In both cases RBP4 concentration was 0.2 μM and the excitation wavelength was assigned to 295 nm.

Organic & Biomolecular Chemistry Accepted Manuscript



Scheme1

Organic & Biomolecular Chemistry Accepted Manuscript

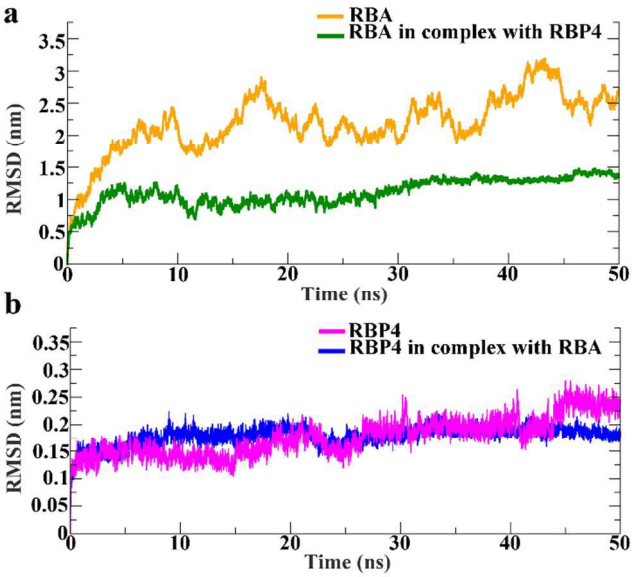


Figure1

Organic & Biomolecular Chemistry Accepted Manuscript

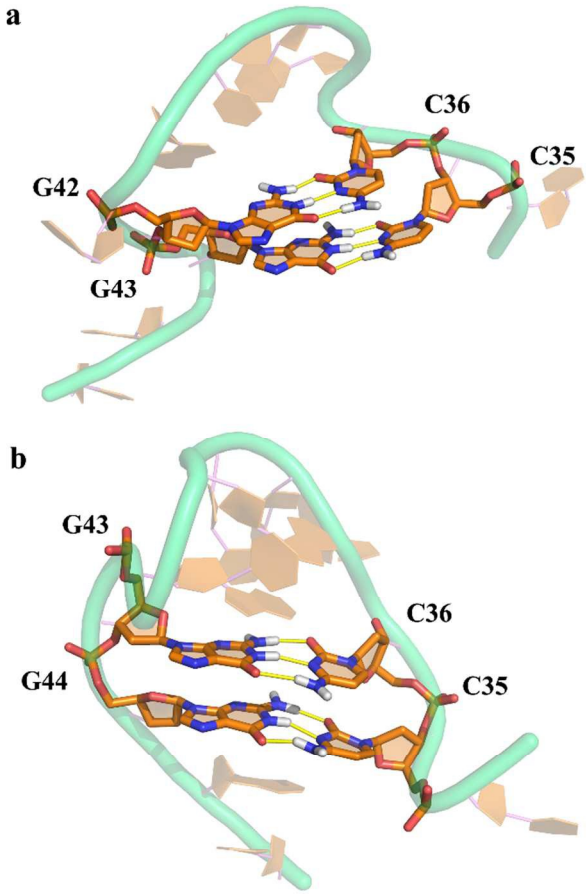


Figure2

Organic & Biomolecular Chemistry Accepted Manuscript

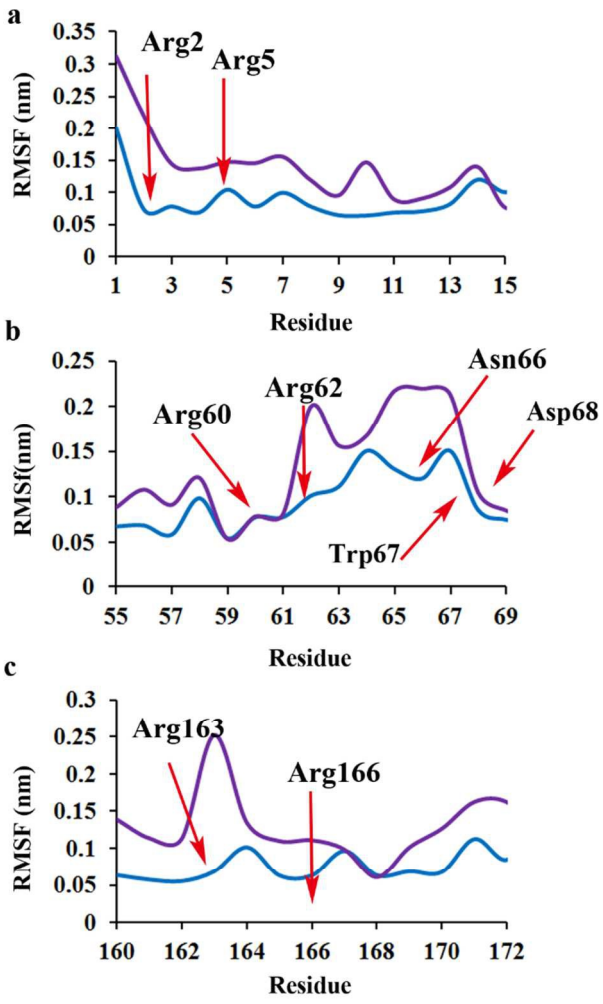


Figure3

Organic & Biomolecular Chemistry Accepted Manuscript

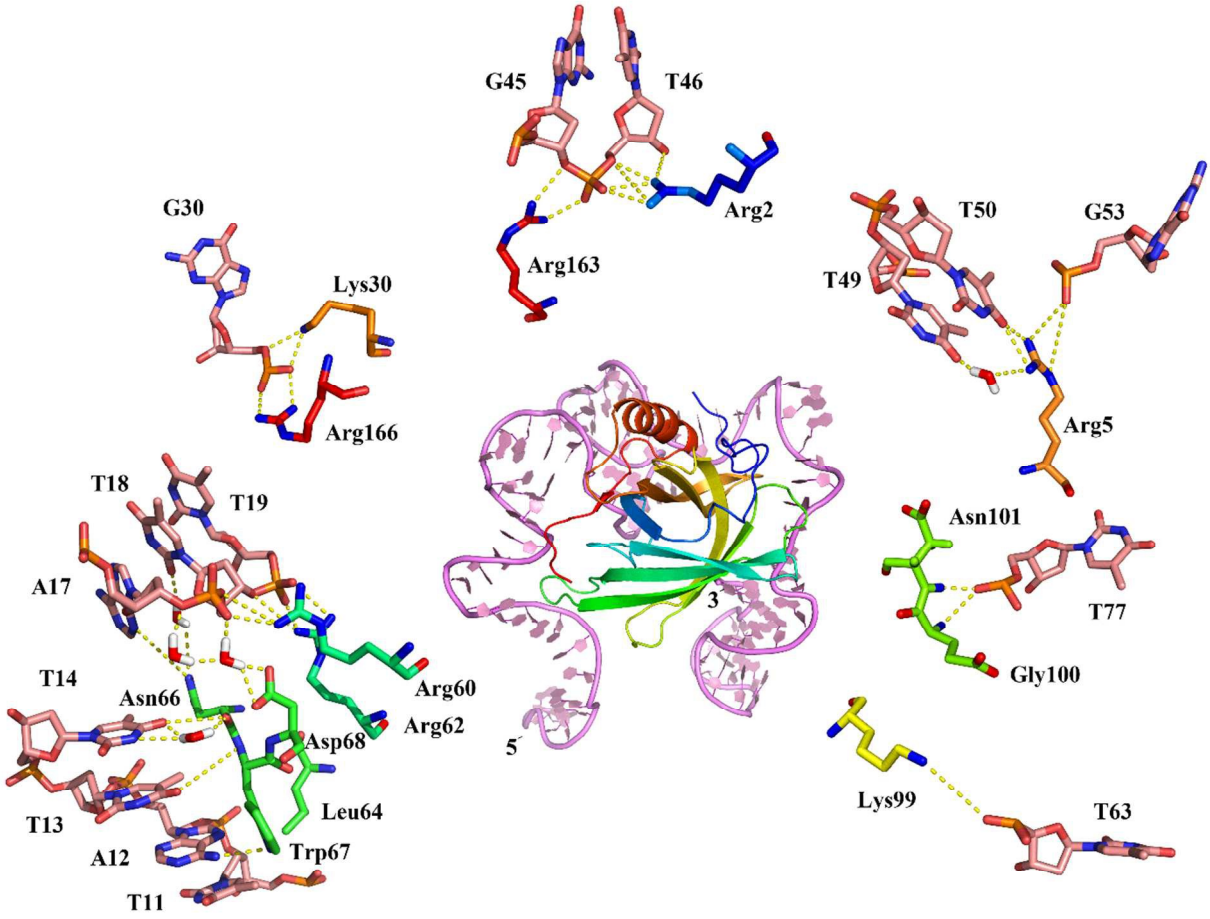


Figure4

Organic & Biomolecular Chemistry Accepted Manuscript

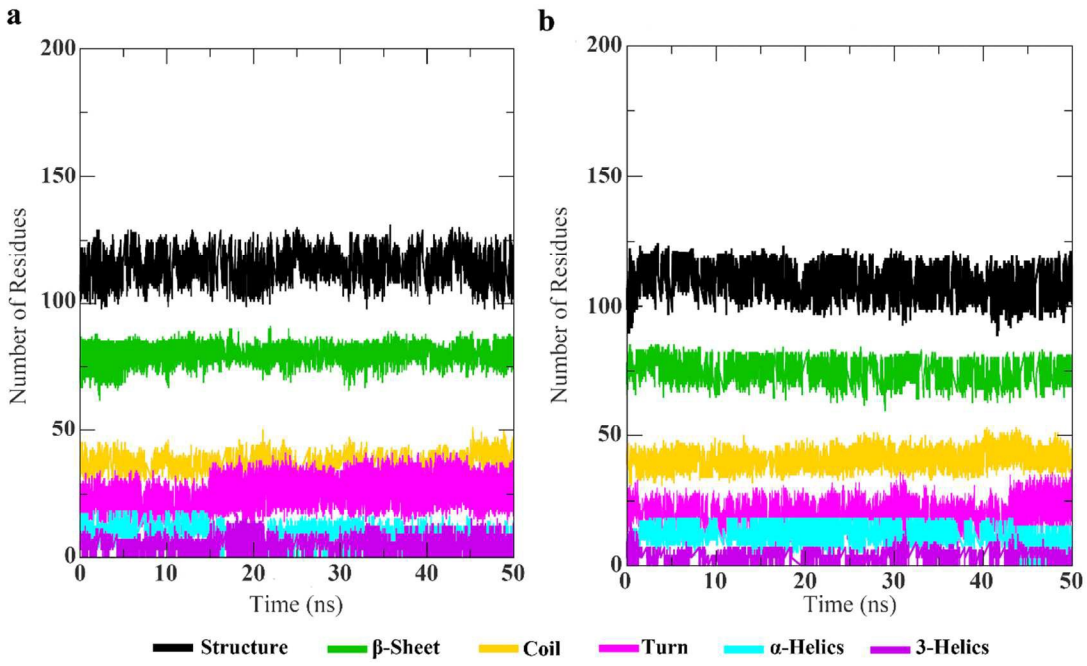


Figure5

Organic & Biomolecular Chemistry Accepted Manuscript

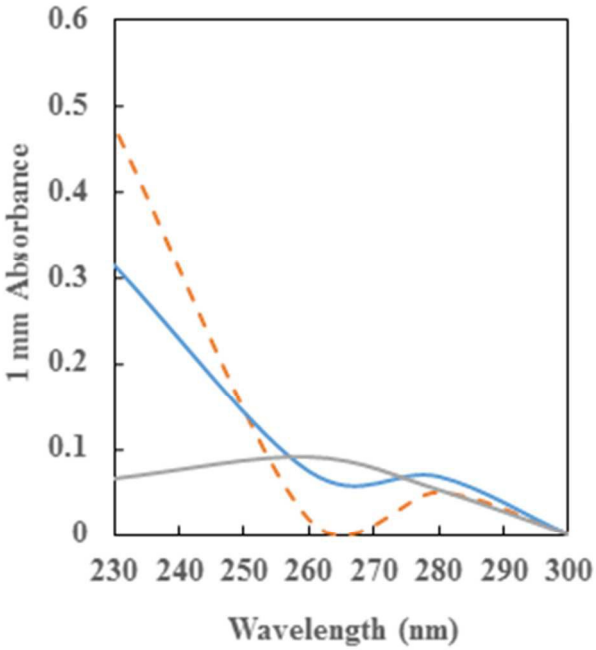


Figure 6.

Organic & Biomolecular Chemistry Accepted Manuscript

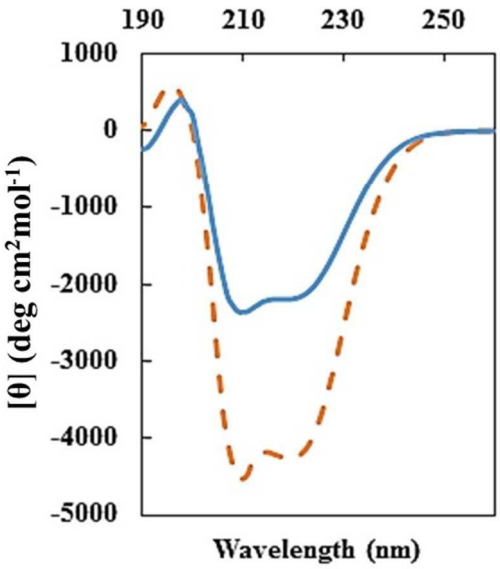


Figure 7

Organic & Biomolecular Chemistry Accepted Manuscript

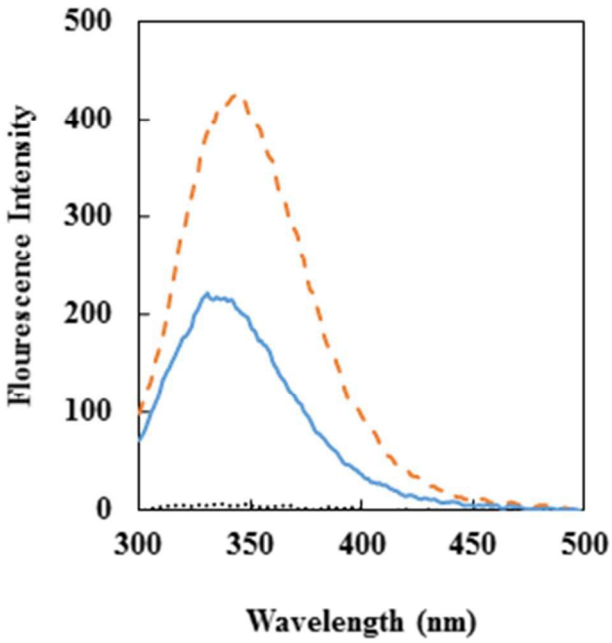


Figure 8.

Organic & Biomolecular Chemistry Accepted Manuscript