Indian Journal of Biochemistry & Biophysics
Vol.58, June 2021, pp. 208-218

Insights into the dynamics of cyclic diguanosine monophosphate I riboswitch using molecular dynamics simulation

Priyanka Kumari[1], Surya Pratap Singh[2]* & Anup Som[1]*

[1]Centre of Bioinformatics, Institute of Interdisciplinary Studies, University of Allahabad, Prayagraj-211 002, Uttar Pradesh, India

[2]Department of Bioscience and Biotechnology, Banasthali Vidyapith, Banasthali -304 022, Rajasthan, India

Received 03 July 2020; *revised 01 April 2021*

Riboswitches are key cis-regulatory elements, present at 5' UTRs of several prokaryotic mRNAs, involved in gene expression regulation by binding selectively to specific ligands followed by conformational changes. However, understanding of ligand-free riboswitch, conformational changes between the ligand-free and ligand-bound riboswitch and binding mechanism of ligand to the aptamer of riboswitch is limited. In the present paper we describe the structural and dynamical properties of ligand-free c-di-GMP I riboswitch aptamer and its possible binding mechanism with c-di-GMP by means of all-atom molecular dynamics (MD) simulations. Various analyses such as principal component analysis, cross correlation dynamics analysis, network analysis and trajectory analyses were carried out. The ligand-free structure shows stable conformation with folded P2, P3 and an unwind P1 helix with open binding pocket at helix-join-helix junction while the ligand-bound structure showed closed binding pocket structure compared to the ligand-free structure. The junction residues significantly showed anti-correlations with P1 helix and weak correlated motions with each other in the open conformation of the ligand-free aptamer of riboswitch. The networks analysis of binding pocket residues suggested interaction among the identified key residues of the binding pocket (G20, A47, C92) and P1 helix illustrating the role of P1 helix in binding of ligand. The structural insights, on c-di-GMP I riboswitch, presented in this paper can be useful for the development of therapeutics against *V. cholerae*. The understanding of the c-di-GMP I riboswitch at dynamic molecular level provides a potential solution for riboswitch drug design.

Keywords: Binding pocket, Dynamic cross correlation, Principal component analysis, Secondary and tertiary structures, Trajectory analysis

Riboswitches are cis-acting, gene regulatory structural domains usually present at the 5' UTR region of mRNA which selectively recognize specific small ligands/metabolites[1,2]. They are found in bacteria, archaea, fungi, and some plants. The structure of riboswitch consists of two regions; the aptamer and the expression platform. The aptamer binding to specific ligand results in the conformational changes which regulate expression of adjacent gene/genes by controlling transcription, translation and splicing mechanisms[3,4]. Thus riboswitches are critical to bacterial life processes and offers a new class of RNA antibiotic drug-target for infectious diseases especially bacterial diseases[5-9].

Based on the type of ligand bound to a riboswitch, more than 40 classes of riboswitches have been discovered[10]. One such class is Cyclic Diguanosine Monophosphate I (cyclic-di-GMP I) riboswitch which senses cyclic di-guanosine monophosphate (c-di-GMP/cyclic-di-GMP). It folds into secondary and tertiary configurations forming a stable three-dimensional structure. The cyclic-di-GMP I riboswitch regulates the expression of genes involved in the metabolism of cyclic-di-GMP which play crucial role in bacterial survival[11,12]. The cyclic-di-GMP is a secondary messenger involved in the mechanisms such as virulence, mobility, quorum sensing and biofilm formation[13]. It has been reported that the c-di-GMP regulates virulence factor genes via riboswitches in *C. difficile*[14]. In a recent study, Tamayo (2019) reported the role of c-di-GMP riboswitches in controlling bacterial pathogenesis mechanisms[15]. They are present in wide variety of bacteria specially the pathogenic strains of proteobacteria, and proposed to be potential drug targets[6,8].

*Correspondence:
E-mail: suryabioinfo@gmail.com (SPS); som.anup@gmail.com (AS)
Suppl. Data available on respective page of NOPR

The conformational transitions between ligand-free and ligand-bound riboswitch aptamer are responsible in regulating the adjacent gene expression. The role of ligand-bound aptamer structures can be studied by crystal structure data in RCSB-PDB, but ligand-free structure is lacking. Thus, it is necessary to get information about the initial ligand-free structure to study the mechanism of aptamer-ligand binding and gene regulation. Molecular Dynamics (MD) simulation studies have been proved valuable computational methods to gain insight into the structural and functional aspects of biomolecules including riboswitches[16,17]. The secondary structural orientations, motions for ligand binding and switching mechanisms, of adenosine deaminase A-riboswitch, have also been explored A similar work was also carried out by Sharma *et al*. (2009) and Zhou *et al*. (2020) for add A-riboswitch to show the role of secondary structural orientation and motions for binding and switching mehanism[18,19]. Further, in a study Li *et al*. (2018) showed the use of MD studies to analyse the binding of natural and inhibitor ligand to 3'-3'-cGAMP riboswitch[20].In this paper, we repost a MD simulation based study to predict the ligand-free cyclic-di-GMP I aptameric ribowsitch and the changes in the aptameric riboswitch upon ligand binding.

Structure of c-di-GMP I riboswitch aptamer

To study the conformational changes and mechanism, detailed knowledge of c-di-GMP I structure is important. The structure of cyclic-di-GMP I riboswitch aptamer shows three helices P1, P2 and P3 forming a three-way junction representing junctions J1/2, J2/3 and J1/3 respectively (Fig. 1). The helix P1 represents nucleotides from G9-C14 and G93-C97, the helix P2 from G21-C46 and the helix P3 from A48-A91. The c-di-GMP I riboswitch aptamer binding pocket is composed of residues from the junction of the three helices (P1, P2, P3) known as junction loops J1/2 (C15-G20) J1/3 (C92) and J2/3 (A47)[21].

Materials and Methods

All-atom MD simulations were carried out using GROMACS package (https://www.gromacs.org/). The details of the system and algorithms used are provided as below:

Initial structure and system preparation

The crystal structure of c-di-GMP I riboswitch aptamer, used in present study was retrieved from protein data bank (PDB) database [PDB code: 3IRW at 2.7Å]. Only the RNA chain of 90 nucleotides long is used as template to derive two systems for further studies[21].

The ligand free riboswitch system was prepared by removing the protein part, c-di-GMP ligand, water molecules and unnecessary ions. The ligand bound Riboswitch system was prepared by deleting the protein chain, water molecules and iridium ions from the crystal structure[22]. Notable that protein chain does not participate in riboswitch function. Each system was placed at the centre of a cubic box, and then solvated using explicit water molecules (*i.e*.TIP3P). The charge of each system was neutralized by adding counter ions to them as detailed in (Table 1). The two systems were then passed to energy minimization,

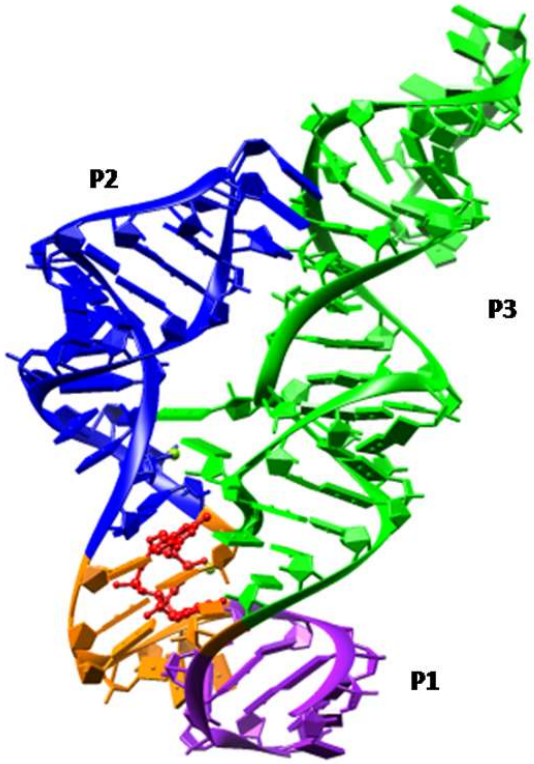

Fig. 1 — Structure of c-di-GMP I riboswitch. Helices P1, P2 and P3 are represented by blue, purple and green colour; the junction residues of J1/2, J2/3 and J1/3 interacting with ligand are represented in orange colour. The ligand is shown in red colour

Table 1 — Details of system preparation for MD simulations

System	Neutralizing counter ions	Total atoms (without water and sodium)	Length of simulation
Apo_RNA	89	2904	150 ns
RNA_Ligand_complex	87	2962	150 ns

using 5000 steps of the steepest descent algorithm, to avoid any steric clashes[23].

Molecular dynamics simulation

All MD simulations were carried out with GROMACS 5.1.2 package[24] using AMBER99SB-ildn force field[25].The ligand parameters were generated using Ante Chamber PYthon Parser interface (ACPYPE) server[26]. Before the production dynamics the minimized systems were passed for 100 ps of NVT equilibration to raise the systems temperature by 300K using the V-rescale thermostat[27]. This equilibration assigned initial velocities to the atoms of the system from Maxwell distribution. The heated systems were passed to 100 ps of NPT equilibration to maintain a constant pressure of 1 bar for the system, using Parinello-Rahman barostat[28]. It also maintains the homogeneous density across the systems. The Linear Constraint Solver (LINCS)[29] algorithm was used for constraining all the bonds and in order to compute the long-range electrostatics, Particle Mesh Ewald (PME)[30] method was incorporated with a cut off value of 1.0 nm. The MD simulations were run based on leapfrog algorithm[31]. A production run of 150 ns was carried out for each system under NPT ensemble with step size 2 fs.

Trajectory and structure analysis

VMD was used to visualise all the trajectory movements for the simulation time and extraction of co-ordinates for particular time frame and the average RMSDs.The structure comparison and visualization were done using PyMOL[32] and CHIMERA visualization tools[22]. The hydrogen bond interaction between the ligand and RNA was plotted using LigPlot[33]. The RMSD, RMSF, SASA etc. graphs of trajectories were obtained using terminal-shell commands and plotted using XMGRACE 2D plotting tool (https://plasma-gate.weizmann.ac.il/Grace/).

Principle component analysis and correlation dynamics analysis

Principle Component Analysis (PCA) was carried out using Bio3D library of R package[34]. The covariance matrices were constructed for the backbone atoms for each frame of the 150 ns long trajectory. The first three PCs (i.e. PC1, PC2 and PC3) were considered for the analysis, as they contributed most of the fluctuations. The dynamic cross correlation maps were created, using Bio3D package[35], to understand the correlated motions during the dynamics[36]. The extent of the correlated motion was calculated by correlation coefficient C(i,j) between the $i^{th}$and j^{th} atomsas follows:

$$C_{(i,j)} = \frac{(< \Delta r_i > \cdot < \Delta r_j >)}{\sqrt{(< \Delta r_i^2 > \cdot < \Delta r_j^2 >)}}$$

where Δr_i and Δr_j is the displacement of $i^{th}$and j^{th} atom from mean position and <> symbol represents the time average.

Correlation network analysis

The network analyses were carried out using the Bio3D and the igraph R packages. The networks were generated on the basis of dynamic cross-correlation matrices calculated for C1' atoms of nucleotide residues for the last 50ns of each simulation using as inputs the corresponding MD trajectory superimposed onto the initial structure.

We also calculated the relative importance of each node for communication using centrality measures and shortest path.

Results

We studied the RMSD, Rg and SASA evaluations of the ligand-free and ligand-bound c-di-GMP I riboswitch aptamer for general observation of properties from trajectory analyses. Li *et al.* (2019) and Gong *et al*. (2014) showed the importance of trajectory analyses to elucidate about the important dynamics of the structure[37, 38]. The RMSD calculations show small fluctuations after 75 ns indicating the systems have attained equilibrium. To estimate the conformational changes in the ligand-free and ligand-bound states of the riboswitch, the comparative plots of RMSDs for global (complete RNA) (Fig. 2A) and local (P1, P2, P3 helix and binding pocket region (Fig. 2B-E) RMSDs were obtained. The RMSD graph for the RNA system was generated using the starting minimised structure as reference structure. RMSDs of both the ligand-free and ligand-bound structures being flexible and dynamic show variations in RMSD values in an acceptable range, the average RMSD for ligand-free structure is 6.58 Å and for ligand-bound structure is 6.09 Å (as detailed in Table 2). As expected, the average RMSD value for ligand-bound structure is lower than ligand-free structure indicating the ligand-bound structure is more stable as compared to ligand-free structure. The helices, P2 and P3 are stabilized with lower fluctuations for both ligand-free and ligand-bound simulations (Fig. 2 and Table 2). However, the helix P1 contributed to the conformational change as the RMSDs of ligand-free structure (4.13 Å)

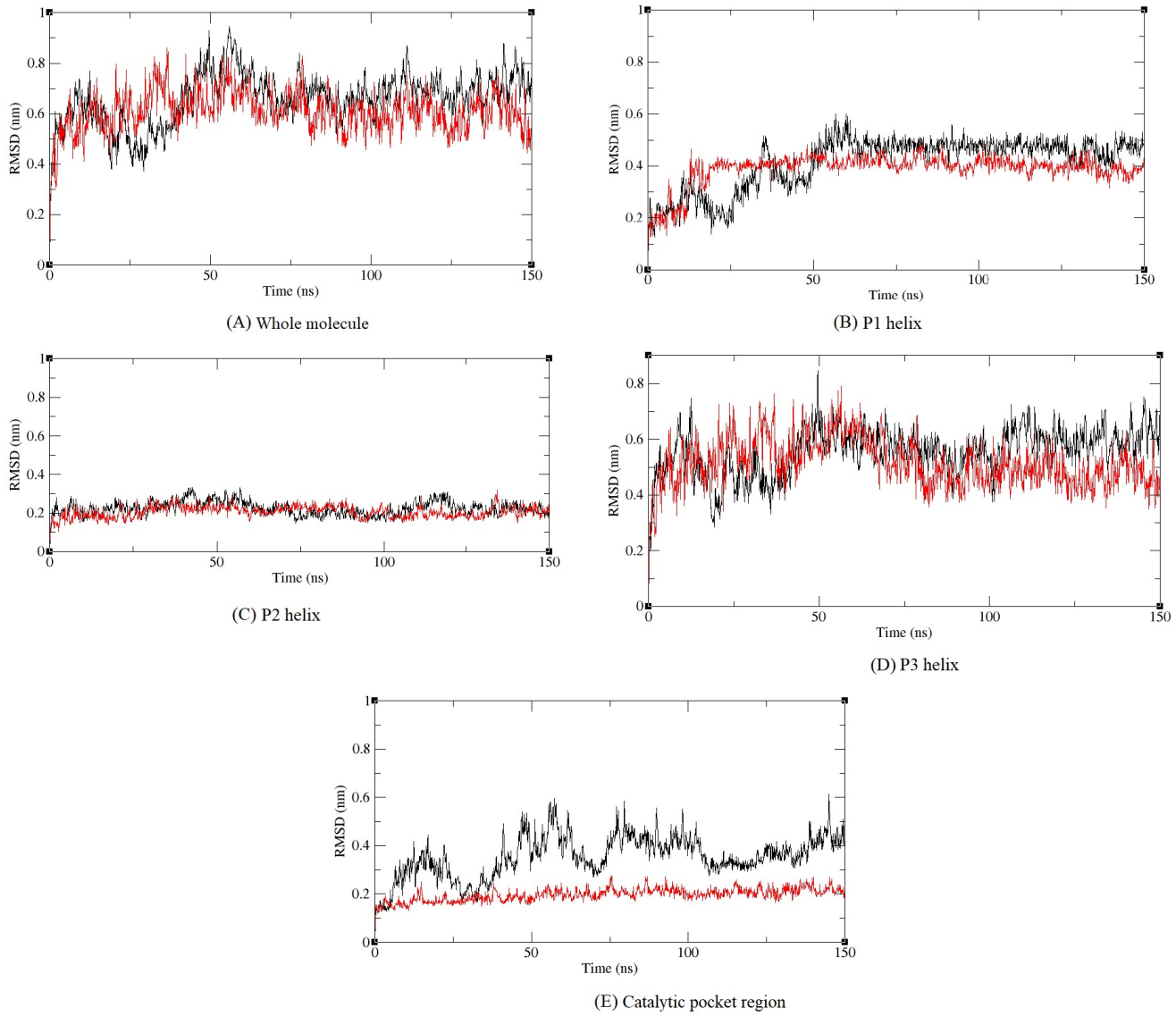

Fig. 2 — Backbone atoms RMSD plots of ligand-bound (red) and ligand-free (black) systems during 150 ns simulation. (A) The graph shows stability in structures after 75 ns and represents whole system RMSD comparison while; (B) to (E) represent RMSD values of the different regions of the c-di-GMP I riboswitch aptamer with respect to the starting minimised structure at the production run

Table 2 — Average RMSD values of the various structures (complete systems and their different regions)

Systems	Average RMSD for ligand-free (nm)	Average RMSD for ligand-bound (nm)
Whole	0.658	0.608
P1 helix	0.413	0.388
P2 helix	0.223	0.205
P3 helix	0.554	0.513
Catalytic pocket	0.354	0.196

were higher and stable as compared to ligand-bound structure (3.88 Å). The increase in the RMSD values of binding pocket of ligand-free aptamer (3.54 Å) with respect to the ligand-bound aptamer (1.96 Å) which is relatively stable predict conformational changes occurring in binding pocket. The Rg and SASA values also verify conformational changes in the binding pocket and indicate its opening.

The radius of gyration (Rg) plot for the catalytic pocket clearly demonstrated that the binding pocket is not compact rather expanded or open in the ligand-free aptamer, but the binding pocket remained closed for ligand-bound structure (Fig. 3).

SASA provides insights into the surface area of the biomolecules (protein or RNA) interacting with water or the area exposed to water. The graph shows a difference (6.09 nm^2) in the average surface area as a function of time for both ligand-free and ligand-bound

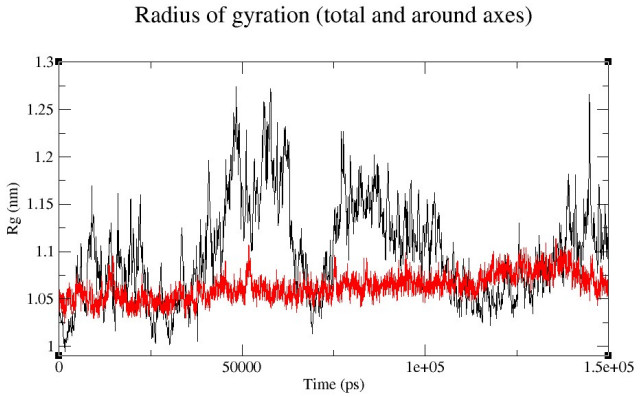

Fig. 3 — The radius of gyration graph for the binding pocket of ligand-free (black) and ligand-bound (red) aptamer structure

structures (Fig. 4A). This difference might be due to the expansion of ligand-free structure supporting the results of the radius of gyration, shown above (Fig. 3). Moreover, the binding pocket SASA graph revealed that the ligand-free structure has relatively more solvent accessible surface area till 110 ns however, post 110 ns SASA of ligand-free and ligand-bound structure becomes almost identical that indicates the MD simulation of ligand-free structure is able to sample the closed conformation of the binding pocket. (Fig. 4B). The average SASA difference between ligand-free and ligand-bound structures for binding pocket is 0.27 nm^2, this minor difference may be due to exposure of part of A47 residue which gets capped, in ligand bound structure, due to the U shape structure of ligand molecule.

Binding pocket Analysis

The residues within 5 Å of ligand, were selected as binding pocket residues using chimera tool (Table 3 and Fig. 5A). The specific binding of ligand to the riboswitch is due to intricate three-dimensional structure of riboswitch with pockets and cavities. This structure is formed by the polymers of nucleotides forming base pairs with hydrogen bond and base stacking. The binding pocket forms a cup shaped structure where the distance between G14 and C46 residues determine the opening and closing of binding pocket (Fig. 5A). Therefore, the average distances between the two residues' centre of masses were plotted over the trajectory for both apo and complex simulation (Fig. 5B). The plot clearly shows the opening of binding pocket for ligand-free structure, the highest distance is illustrated at 65 ns. The average distance between the stacked A47 and C92 is also higher for apo configuration (1.45 nm) as

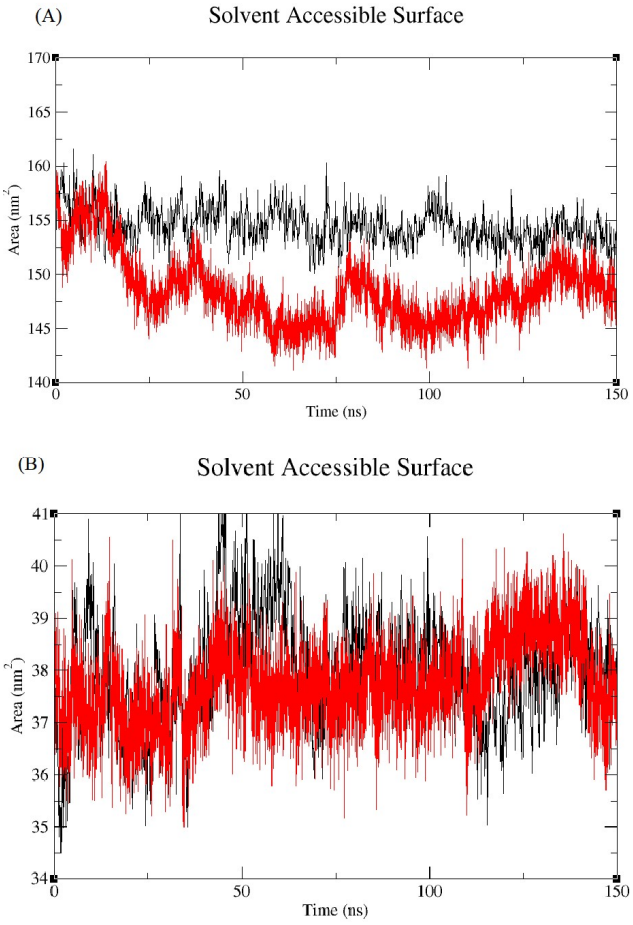

Fig. 4 — SASA plots for ligand-free (black) and ligand-bound (red) systems during 150 ns simulation (A) shows SASA for complete RNA system; and (B) represents SASA calculations for the binding pocket nucleotides

Table 3 — The catalytic pocket residues generated using Chimera tool

S. No	Nucleotide residue	Nucleotide residue number
1	G	14
2	C	15
3	A	16
4	C	17
5	A	**18**
6	G	19
7	G	**20**
8	G	21
9	C	**46**
10	A	**47**
11	A	**48**
12	A	49
13	G	50
14	C	**92**
15	C	93

Bold faced residues represent key residues participating in hydrogen bond formation with ligand

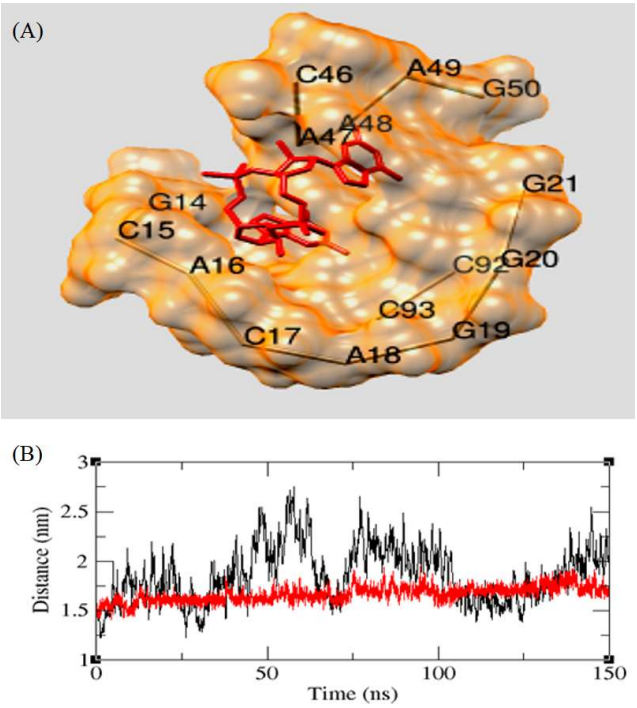

Fig. 5 — (A) Image representing the catalytic cavity formed by the surrounding nucleotides and the ligand; and (B) Average distance between G14 and C46 residues over the trajectory for apo (black) and complex (red) systems

compared to the bound configurations (1.07 nm) and the average distance between the G20 and A47 slightly lower for apo configuration (1.19 nm) than that of bound configuration (1.24 nm).

In the crystal structure the ligand c-di-GMP (C2E) forms 11 bonds with the surrounding nucleic bases (Fig. 6A shows hydrogen bond with surrounding nucleotide residues). The two guanosines of the c-di-GMP form hydrogen bond with G20 and C92 nucleotide residues which are highly conserved and play an important role in ligand binding mechanism. The O6 and N7 of Gα hydrogen bond with the exocyclic amine of G20. Other hydrogen bonds are made between the N2 of Gα and the 3′-O of A48 and the 2′-OH of C46, and between the Gα2′-OH and nonbridging phosphate oxygen of A47. The second guanosine of c-di-GMP, Gβ, forms a standard Watson Crick base pair with C92, a highly conserved nucleotide 3′ of P3 (Fig. 6B). The interaction is further supported by hydrogen bonds between the 2′-OH of A47 and the N7 of Gβ, between the N1 of A18 and the N2 of Gβ, and between the Gβ 2′-OH and the exocyclic amine of A18[21]. These interactions are in good agreement with the hydrogen bond interactions map of key binding pocket residues obtained from

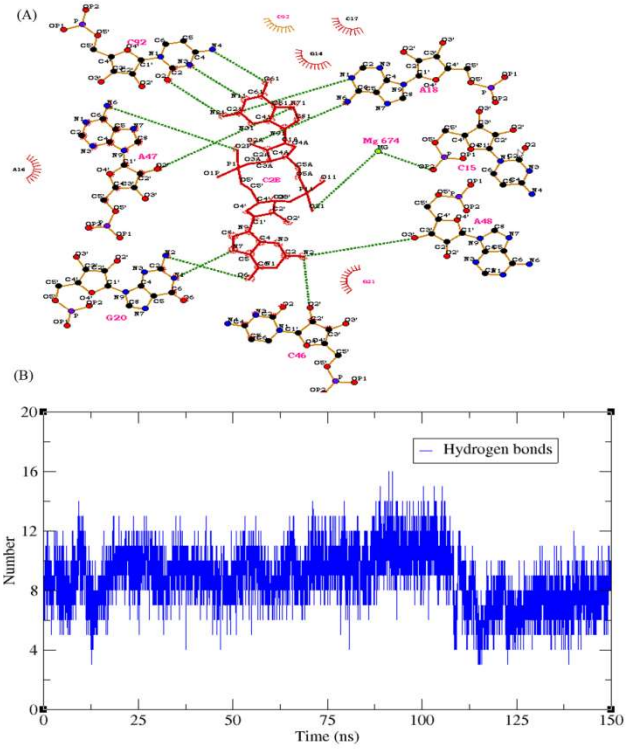

Fig. 6 — (A) Hydrogen bonding and nucleotide details of the binding pocket of c-di-GMP I riboswitch aptamer. Hydrogen bonds are represented by green dotted and red colour shows the C2E ligand; and (B) Temporal change in number of stable hydrogen bond (blue) between the c-di-GMP I riboswitch aptamer and C2E ligand for 150 ns MD simulation

Table 4 — The hydrogen bond occupancies (%) between the ligand and binding pocket residues for ligand-bound c-di-GMP I aptamer

Donor	Acceptor	Hbond occupancy
LIG675-Side	C92-Side	17.71%
LIG675-Side	A47-Side	69.04%
G20-Side	LIG675-Side	9.19%
A18-Side	LIG675-Side	2.13%
LIG675-Side	A18-Side	48.14%
LIG675-Side	C46-Side	28.10%
LIG675-Side	A49-Side	31.03%
A47-Side	LIG675-Side	18.31%
LIG675-Side	C17-Side	0.20%
G19-Side	LIG675-Side	0.07%
LIG675-Side	A16-Side	0.93%
LIG675-Side	G50-Side	0.13%
G14-Side	LIG675-Side	0.13%
C92-Side	LIG675-Side	0.33%
LIG675-Side	A91-Side	3.20%
A91-Side	LIG675-Side	1.60%
C17-Side	LIG675-Side	0.07%
LIG675-Side	C93-Side	0.13%

ligplot (Fig. 6A), chimera and hydrogen bond occupancy data (Table 4).

The temporal changes in the number of hydrogen-bonds formed between the binding pocket and the

C2E ligand is shown in the (Fig. 6B). The average hydrogen bonds are 9 per frames during the entire simulation time. However, the hydrogen bond number after 110ns shows relatively less hydrogen bonds. The hydrogen bond obtained from the PDB structure using chimera tool and ligplot showed 12 hydrogen bonds between binding pocket nucleotide residues and the c-di-GMP I ligand. The hydrogen bonds show the strength of binding of ligand to the riboswitch. The residues forming the hydrogen bonds are A18, G20, C46, A47, A48 and C92 (bold faced in Table 3).

Correlation network analysis was performed by constructing graphs in which each residue was represented by a single node, and the weight of the connection between pairs of nodes was proportional to their respective previously calculated correlation coefficients. We computed the topological parameters such as centrality measures (degree, betweenness and closeness) for apo and bound simulated system. The closeness and betweenness centrality used to identify critical communication nodes over the network. Residues presenting high betweenness values are considered important as they show high number of unique-shortest crossing them. The betweenness and closeness centrality for ligand-bound structure clearly shows high values for P1 helix correlating its role in ligand binding. To get a deeper understanding about the critical residues governing P2-P1 interactions, we computed the 50 shortest paths between G20 (located in P2) and G97 (located in P1) (Supplementary Fig. S1). The shortest path was found to be from G20-A47-C92 and then to P1 helix.

P1 helix dynamics

To support the conformational changes in P1 helix as evidenced by the trajectory analysis we also carried out hydrogen bond formation and torsion angle analyses. The hydrogen bond analysis shows that the ligand-free aptamer completely lacks the hydrogen bond formation between Gβ of ligand and C92 (which is considered as first hydrogen bonded pair) and the hydrogen bond between G14 and C93 is unstable and indicates the helix opening (Fig. 7).

The study of χ torsion angles, are important determinants of structural changes, illustrates the local conformational space over the MD trajectories. The nucleobases are linked with glycosidic bond to sugar of phosphate-sugar backbone. We analysed χ torsion angles, rotation of glycosidic bond, for binding pocket conserved key residues and residues from P1 helix using barnaba software[39]. Different rotamer states were obtained for P1 helix unpaired residues where the syn state dominates for G9 and U10 in the ligand-free structure (Fig. 8).

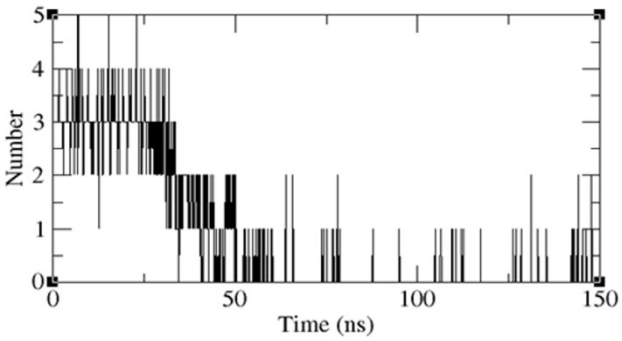

Fig. 7 — Hydrogen bond numbers between G14 and C93 residue over ligand-free aptamer MD production run

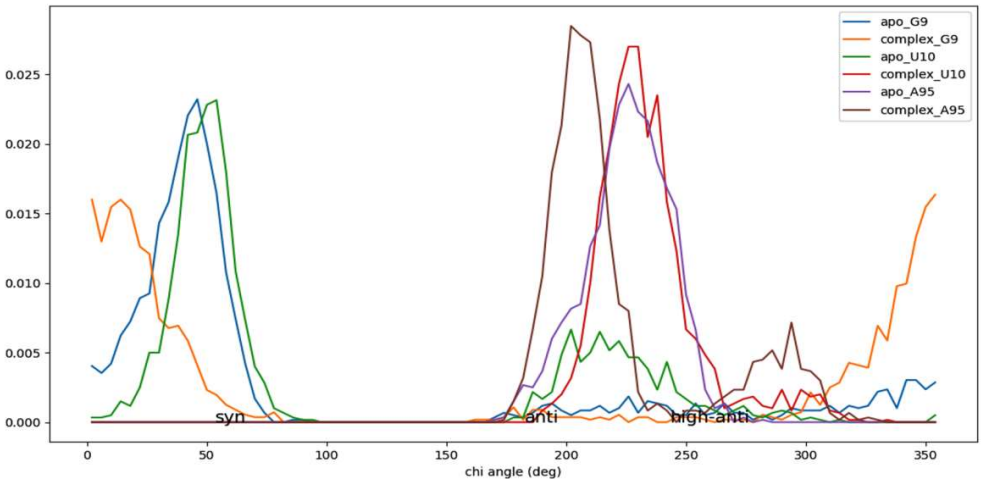

Fig. 8 — Graph showing chi torsion angle distribution of unpaired residues of P1 helix for MD production run for both ligand-free and ligand-bound systems

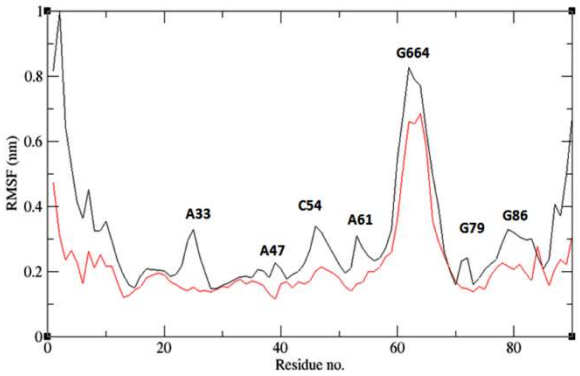

Fig. 9 — RMSF plot of ligand-free (black) and ligand-bound (red) systems during 150 ns simulation for their nucleotide residue

Characteristic features of ligand-free and ligand-bound c-di-GMP I riboswitch aptamer

To study the structural and dynamical properties of ligand-free and ligand-bound states of the c-di-GMP I riboswitch, the (root mean square fluctuations) RMSF fluctuations and superimposed structures were analysed. The RMSF were calculated over C1' atoms and plotted against the simulation time (Fig. 9). The fluctuations of nucleotides for ligand-free structure are more than ligand-bound for each nucleotide due to the less stability of ligand-free structure. The free motion of the terminal regions of the aptamer results in relatively higher fluctuations, further, a sharp peak around the 60-70 region is observed which constitutes the upper portion of P3 helix in the crystal structure. The lower half of the P3 helix is hydrogen bonded to the P2 helix while the upper half of the P3 helix is free, therefore, highly dynamic, and flexible and resulting in the high fluctuation.The binding site region shows higher fluctuations for ligand-free structure as compared to the ligand-bound structure indicates that the nucleotide atoms of this region of the ligand-free structure are more dynamic to move. The smaller peaks at 25, 46, 53 and 62 for ligand-bound structure show slightly higher fluctuations than their corresponding nucleotides or atoms in ligand-bound structure. The residue at 25 correspond to apex of P2 helix which is free to move as it do not participate in forming pseudo knot with P3 helix, similarly residues 46, 53 and 62 belong to P3 helix which have free movement due to no hydrogen bond formation with other nucleotide residue. An interesting observation is captured at 39th position which corresponds to the stacking residue A47 which shows higher fluctuations in ligand free system and its movement gets restricted in ligand bound state where the presence of two guanines stack with A47 into it.

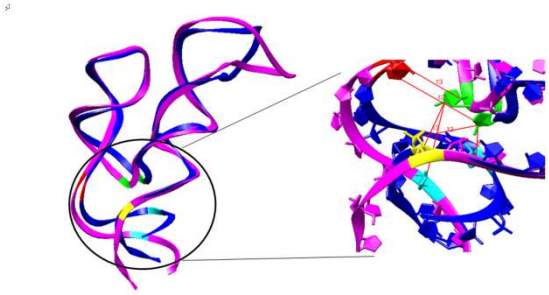

Fig. 10 — Superimposed image of the average ligand-free (magenta) and crystal ligand-bound (Blue) c-di-GMP I riboswitch aptamer structures. It shows changes at P1, P2 and P3 helices. Blue colour represents ligand-bound structure and magenta colour represents ligand-free structure. The zoomed region shows distance differences between residues G20 [red], A47 [green], C92 [yellow] and G14 [cyan] where the distance (Å) is shown in red

Table 5 — Distance data for junction residues of superimposed average ligand-free structure on the stable crystal structure

Residues	Distance (Å)
G14_free and G14_bound	10Å
G20_free and G20_bound	6Å
A47_free and A47_bound	6Å
C92_free and C92_bound	2Å
G20_free and A47_free	13Å
G14_bound and A47_bound	12Å
G14_free and A47_free	13Å
G20_bound and A47_bound	8Å
C92_free and A47_free	12Å
C92_bound and A47_bound	11Å

Further, to investigate the structural changes at tertiary level the average ligand-free structure was superimposed on the stable crystal structure (Fig. 10). The binding pocket, P1 helix and upper portion of P3 helix show displacement in their positions. The distances where calculated and analyzed for the key residues (G20[red], A47[green] and C92[yellow]) and pocket structure related residue (G14[cyan]). These residues show major displacement between each other and with each other for both the structures as tabulated below (Table 5). The calculations confirm the results of average distance calculated over simulation data in binding pocket analysis section. Moreover, this analysis revealed that in ligand-bound aptamer, the base pair G14 and C93 is aligned to form hydrogen bonds but in ligand-free aptamer C93 moves upwards.

Essential dynamics

To identify the dominant ligand-free motions of the apo_RNA, PCA was applied to all backbone atoms of

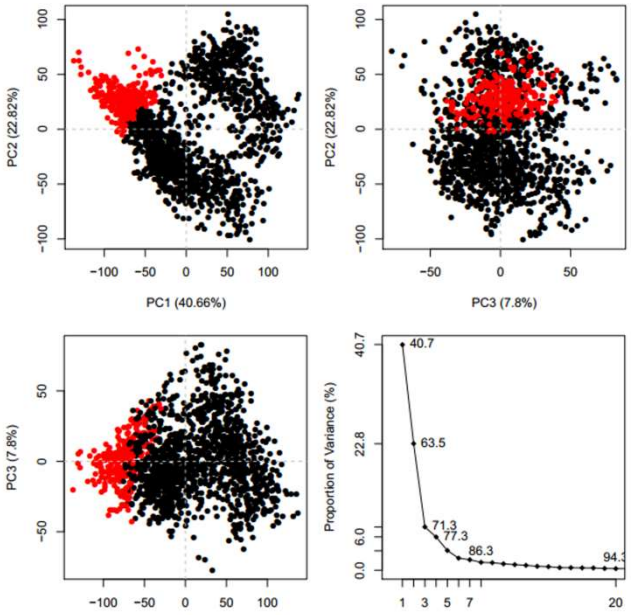

Fig. 11 — Principle component analysis of apo_RNA MD trajectories. 2D plots for the first three principal components PC1 *vs* PC2, PC2 Vs PC3 and PC1 *vs* PC3 respectively are represented by first three plots. Here the red and black colour shows two conformational subspaces. The last plot shows the proportion of variance contributed by first 20 PCs which covers 94% internal motions

the apo_RNA riboswitch aptamer. The percentage contribution of the total variance of top twenty eigenvectors is shown in (Fig. 11) down right. It is evident that first 20 principal components (PCs) capture over 94% of internal motions. Further, only first three PCs are describing about 71% of all the motions. The PC1 has the highest value (40.70%) and strongly dominates the overall variance showing the global dynamics. An assessment of the internal motions along first three PCs revealed that PC1and PC3 is prominently correlated.

Communication between binding pocket and P1 helix using correlated motions:

Dynamic cross correlation (DCC) represents the dynamical changes of the system over the time, it computes the correlated atomistic movements within the molecule (*i.e*. the degree to which they move together)[40]. The DCC map for ligand-free aptamer representing correlation coefficients calculated as time average over the simulation time is shown in (Fig. 12). The range of correlation is between +1 to −1 representing both positive and negative correlation and their variations are shown by the colour gradations. The blue colour represents positive correlation, pink represent negative correlation and

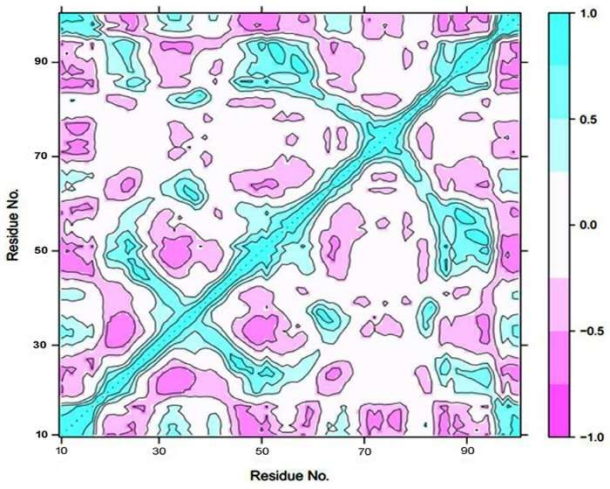

Fig. 12 — DCC map of last 50 ns simulation for ligand-free aptamer. Blue and pink colour corresponds to high correlations and high anti-correlations respectively, while white regions represent no correlation zone

white represent uncorrelated fluctuations. The map revealed that the junction residues are negatively correlated with P1 helix, however the junctions J1/2, J2/3 and J1/3 show weak or non-correlated movements with each other. The two arms of the P1 helix shows positive correlated fluctuations illustrating the nucleotides are hydrogen bonded and tend to move together. The blue islands which indicate strong positively correlated fluctuations of nucleotides show the formation of the aptamer fold. This supports the existence of a preformed aptamer fold in ligand-free structure.

Discussion

In this paper we have explored the possible stable structure of ligand-free aptamer of c-di-GMP I riboswitch that can sense the ligand c-di-GMP and inspected the riboswitch-ligand binding mechanism.

Various trajectory analyses indicated that the ligand-free structure of the c-di-GMP I riboswitch aptamer have a stable conformation with open binding pocket. This is in agreement with the existing literature where it had been shown that the ligand-free riboswitch existed in wide open form and has its P2 and P3 stems far apart[41]. However, in case of ligand-bound structure both the stems are close to each other and form pseudoknot like structure, i.e., the junction J1/2, J2/3 and J1/3 are close enough to make bonds with ligand and enclose it in the pocket. Thus, it is being anticipated that these conformational changes should regulate the mRNA translations the bacteria. A chemical probing-based study by Inuzuka *et al*.

(2018)[42], on full length c-di-GMP I riboswitch structure showed translational repression due to masking of ribosome binding site because of conformation change in the expression platform on binding of ligand to the aptamer of the riboswitch.

The Essential dynamics-based study of correlated motions of a biomolecule to understand the most fundamental fluctuations related to the biomolecular functional activity partitions conformational spaces into an essential subspace with a few degrees of freedom. The first few PCs were indicated showed two major distinct conformational sub-states of ligand-free c-di-GMP I riboswitch.

The ligand binding to the c-di-GMP I riboswitch results in the closure of the binding pocket by the movement of the J1/2, J2/3 and J1/3 junction residues. The DCC analyses of the ligand-free aptamer too suggested anti-correlated movements between the junction and P1 helix which agrees to the studies reported by Li *et al.* (2018 and 2019)[20,37]. The correlation network results also suggested that the bound riboswitch interactions are distinct from that of the apo riboswitch.

Conclusion

MD simulations were performed for both the ligand-bound and ligand-free aptamer of the c-di-GMP I riboswitch to identify ligand-free stable structure and the role of aptamer in recognition of specific ligand for binding[43,44]. The simulations resulted in identification of ligand-free aptamer structure with stable conformation which has folded P2, P3, open catalytic pocket and an unwind P1 helix. The results showed that the binding of ligand to the aptamer plays an important role in stabilizing the overall structure. In the absence of c-di-GMP I, the aptamer showed changes in the binding pocket region showing a shift in the residues A18, G20, C46, A47 and C92. These represent the key residues which play critical role in ligand binding and are found to be strongly conserved in the bacterial system. High and stable fluctuations were seen at the P1 helix region compared to the P2 and P3 helices for the trajectories dynamics of the ligand-free structure. Thus, the trajectory and correlation analyses demonstrated the interactions in the secondary structures to illustrate the binding mechanism. The data presented in this study could be further useful in finding high affinity inhibitors for c-di-GMP I riboswitch. These results might also be useful in understanding the conformational interactions between the aptamer and expression platform of c-di-GMP I riboswitch. Further, this study can be expanded to study mutations in the key residues and their effect on the binding affinity as well as in designing potential drug target[45].

Acknowledgement

The authors thank the anonymous reviewers for their valuable comments. We thank Dr. Dhananjay Bhattacharya, Dr. Jens Carlsson and Dr. RamaSubbu Sankararamakrishnan for various useful suggestions. PK thanks Dr. Mark E Tuckerman and Dr. Nishant Nair for useful discussions. PK thanks the UGC, India for providing financial assistant to carry out research work. Funding by the Department of Biotechnology (India) to AS is gratefully acknowledge.

Conflicts of interest

The authors declare no conflict of interest.

References

1 Barrick JE & Breaker RR, The distributions, mechanisms, and structures of metabolite-binding riboswitches. *Genome Biol*, 8 (2007) 11.

2 Guil S & Esteller M, RNA-RNA interactions in gene regulation: The coding and noncoding players. *Trends Biochem Sci*, 40 (2015) 248.

3 Mandal M & Breaker RR, Gene regulation by riboswitches. *Nat Rev Mol Cell Biol*, 5 (2004) 451.

4 Winkler WC & Breaker RR, Genetic control by metabolite-binding riboswitches. *Chembiochem*, 4 (2003) 1024.

5 Blount KF & Breaker RR, Riboswitches as antibacterial drug targets. *Nat Biotechnol*, 24 (2006) 1558.

6 Kumari P & Som A, *In silico* Identification of riboswitches in the human gut microbiome for therapeutic applications. *J RNA Genom*, 15 (2019) 610.

7 Mehdizadeh E, Saeid M & Barzegar A, Riboswitches : From living biosensors to novel targets of antibiotics. *Gene*, 529 (2016) 244.

8 Rekand IH, Ligand design for riboswitches, an emerging target class for novel antibiotics. *Future Med Chem*, 9 (2017) 1649.

9 Chan H, Ho J, Liu X, Zhang L, Wong SH, Chan MT & Wu WK, Potential and use of bacterial small RNAs to combat drug resistance: a systematic review. *Infect Drug Resist*, 10 (2017) 521.

10 McCown PJ, Corbino KA, Stav S, Sherlock ME & Breaker RR, Riboswitch diversity and distribution. *RNA*, 7 (2017) 995.

11 Sudarsan N, Lee ER, Weinberg Z, Moy RH, Kim JN, Link KH & Breaker RR, Riboswitches in eubacteria sense the second messenger cyclic di-GMP. *Science*, 321 (2008) 411.

12 Fernandez NL, Srivastava D, Ngouajio AL & Waters CM, Cyclic di-GMP positively regulates DNA repair in *Vibrio cholerae*. *J Bacteriol*, 200 (2018) e00005.

13 Li X, Chen F, Xiao J & He J, Structure and function of c-di-GMP riboswitches. *Sheng Wu Gong Cheng Xue Bao= Chinese J Biotechnol*, 33 (2017) 1357.

14 McKee RW, Harvest CK & Tamayo R, Cyclic Diguanylate Regulates Virulence Factor Genes via Multiple Riboswitches in Clostridium difficile. *mSphere*, 3 (2018) e00423.

15 Tamayo R, Cyclic diguanylate riboswitches control bacterial pathogenesis mechanisms. *PLoS Pathog*, 15 (2019) e1007529.

16 Stoddard CD, Montange RK, Hennelly SP, Rambo RP, Sanbonmatsu KY & Batey RT, Free state conformational sampling of the SAM-I riboswitch aptamer domain. *Structure*, 18 (2010) 787.

17 Kumar GA & CS S, *In silico* and *in vitro* validation of some benzimidazole derivatives as adenosine deaminase inhibitors. *Indian J Chem-Sec B*, 59 (2020) 1175.

18 Sharma M, Bulusu G & Mitra A, MD simulations of ligand-bound and ligand-free aptamer: molecular level insights into the binding and switching mechanism of the add A-riboswitch. *RNA*, 15 (2009) 1673.

19 Zhou T, Wang H, Song L, Zhao Y. Computational study of switching mechanism in add A-riboswitch. *J Theor Compt Chem*, 19 (2020) 2040001.

20 Li C, Zhao X, Zhu X, Xie P & Chen G, Structural studies of the 3′, 3′-cGAMP riboswitch induced by cognate and noncognate ligands using molecular dynamics simulation. *Int J Mol Sci*, 19 (2018) 3527.

21 Smith KD, Lipchock SV, Ames TD, Wang J & Breaker RR, Strobel SA,Structural basis of ligand binding by a c-di-GMP riboswitch. *Nat Struct Mol Biol*, 16 (2009) 1218.

22 Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC & Ferrin TE, UCSF Chimera-A visualization system for exploratory research and analysis. *J Comput Chem*, 25 (2004) 1605.

23 Freund RM, The steepest descent algorithm for unconstrained optimization and a bisection line-search method. Journal of Massachusetts Institute of Technology, *United States of America*, (2004).

24 Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B & Lindahl E, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *Software X*, 1 (2015) 19.

25 Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO & Shaw DE, Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins*, 78 (2010) 1950.

26 da Silva AWS & Vranken WF, ACPYPE-Antechamber python parser interface. *BMC Res Notes*. 5 (2012) 367.

27 Bussi G, Donadio D & Parrinello M, Canonical sampling through velocity rescaling. *J Chem Phys*, 126 (2007) 014101.

28 Parrinello M & Rahman A, Polymorphic transitions in single crystals: A new molecular dynamics method. *J Appl Phys*, 52 (1981) 7182.

29 Hess B, Bekker H, Berendsen HJ & Fraaije JG, LINCS: A linear constraint solver for molecular simulations. *J Comput Chem*, 12 (1997) 1463.

30 Darden T, Perera L, Li L & Pedersen L, New tricks for modelers from the crystallography toolkit: the particle mesh Ewald algorithm and its use in nucleic acid simulations. *Structure*, 7 (1999) R55.

31 Van Gunsteren WF & Berendsen HJ, A leap-frog algorithm for stochastic dynamics. *Mol Simul*, 1 (1988) 173.

32 Schrodinger LLC, The PyMOL molecular graphics system. 2010. Version, 1, r1.

33 Wallace AC, Laskowski RA & Thornton JM, LIGPLOT: A program to generate schematic diagrams of protein-ligand interactions. *Protein Eng Des Sel*, 8 (1995) 127.

34 David CC & Jacobs DJ, Principal component analysis: A method for determining the essential dynamics of proteins. *Methods Mol Biol*, 1084 (2014) 1084193.

35 Grant BJ, Rodrigues APDC, Elsawy KM, Mccammon AJ & Caves LSD, Bio3d: An R Package for the Comparative Analysis of Protein Structures. *Bioinformatics*, 22 (2006) 2695.

36 Antunes D, Jorge NAN, de Souza Costa MG, Passetti F & Caffarena ER, Unraveling RNA dynamical behavior of TPP riboswitches: a comparison between *Escherichia coli* and *Arabidopsis thaliana* L. *Sci Rep*, 9 (2019) 1.

37 Li C, Zhao X, Xie P, Hu J & Bi H, Molecular dynamics simulation on the allosteric analysis of the c-di-GMP class I riboswitch induced by ligand binding. *J Mol Recog*, 32 (2019) e2756.

38 Gong Z, Zhao Y, Chen C, Duan Y & Xiao Y, Insights into ligand binding to PreQ1 Riboswitch Aptamer from molecular dynamics simulations. *PloS One*, 9 (2014) e92247.

39 Bottaro S, Bussi G, Pinamonti G, Reißer S, Boomsma W & Lindorff-Larsen K, Barnaba: software for analysis of nucleic acid structures and trajectories. *RNA*, 25 (2019) 219.

40 Kasahara K, Fukuda I & Nakamura H, A novel approach of dynamic cross correlation analysis on molecular dynamics simulations and its application to Ets1 dimer–DNA complex. *PLoS One*, 9 (2014) e112419.

41 Wood S, Ferré-D'Amaré AR & Rueda D, Allosteric tertiary interactions preorganize the c-di-GMP riboswitch and accelerate ligand binding. *ACS Chem Biol*, 7 (2012) 920.

42 Inuzuka S, Kakizawa H, Nishimura KI, Naito T, Miyazaki K, Furuta H, Matsumura S & Ikawa Y, Recognition of cyclic-di-GMP by a riboswitch conducts translational repression through masking the ribosome binding site distant from the aptamer domain. *Genes Cells*, 6 (2018) 435.

43 Hu G, Li H, Xu S & Wang J, Ligand binding mechanism and its relationship with conformational changes in Adenine Riboswitch. *Int J Mol Sci*, 21 (2020) 1926.

44 Sun A, Gasser C, Li F, Chen H, Mair S, Krasheninina O, Micura R & Ren A, SAM-VI riboswitch structure and signature for ligand discrimination. *Nat Commun*, 16 (2019) 1.

45 Srivastava S & Pandey A, Computational screening of anticancer drugs targeting miRNA155 synthesis in breast cancer. *Indian J Biochem Biophys,* 57 (2020) 389.