



OPEN

# Probing the Structural Dynamics of the Catalytic Domain of Human Soluble Guanylate Cyclase

Rana Rehan Khalid<sup>1,2,3</sup>, Arooma Maryam<sup>1,4</sup>, Osman Ugur Sezerman<sup>3</sup>, Efstratios Mylonas<sup>5</sup>, Abdul Rauf Siddiqi<sup>1</sup>✉ & Michael Kokkinidis<sup>2,5</sup>✉

In the nitric oxide (NO) signaling pathway, human soluble guanylate cyclase (*hsGC*) synthesizes cyclic guanosine monophosphate (cGMP); responsible for the regulation of cGMP-specific protein kinases (PKGs) and phosphodiesterases (PDEs). The crystal structure of the inactive *hsGC* cyclase dimer is known, but there is still a lack of information regarding the substrate-specific internal motions that are essential for the catalytic mechanism of the *hsGC*. In the current study, the *hsGC* cyclase heterodimer complexed with guanosine triphosphate (GTP) and cGMP was subjected to molecular dynamics simulations, to investigate the conformational dynamics that have functional implications on the catalytic activity of *hsGC*. Results revealed that in the GTP-bound complex of the *hsGC* heterodimer, helix 1 of subunit  $\alpha$  ( $\alpha$ :h1) moves slightly inwards and comes close to helix 4 of subunit  $\beta$  ( $\beta$ :h4). This conformational change brings loop 2 of subunit  $\beta$  ( $\beta$ :L2) closer to helix 2 of subunit  $\alpha$  ( $\alpha$ :h2). Likewise, loop 2 of subunit  $\alpha$  ( $\alpha$ :L2) comes closer to helix 2 of subunit  $\beta$  ( $\beta$ :h2). These structural events stabilize and lock GTP within the closed pocket for cyclization. In the cGMP-bound complex,  $\alpha$ :L2 detaches from  $\beta$ :h2 and establishes interactions with  $\beta$ :L2, which results in the loss of global structure compactness. Furthermore, with the release of pyrophosphate, the interaction between  $\alpha$ :h1 and  $\beta$ :L2 weakens, abolishing the tight packing of the binding pocket. This study discusses the conformational changes induced by the binding of GTP and cGMP to the *hsGC* catalytic domain, valuable in designing new therapeutic strategies for the treatment of cardiovascular diseases.

Guanylate cyclase (GC) exists as membrane-bound particulate guanylate cyclase (pGC) and cytoplasmic soluble guanylate cyclase (sGC) in mammalian tissues. Both these forms of the sGC utilize guanosine triphosphate (GTP) as a substrate and metal ions ( $Mg^{2+}$ ,  $Mn^{2+}$ ) as cofactors to stimulate the release of inorganic pyrophosphate and promote cyclization at the  $\alpha$ -phosphorus atom to form 3',5'-cyclic guanosine monophosphate (cGMP)<sup>1</sup>. Nitric oxide (NO) and hormones modulate the activity of sGC and pGC respectively. It is now established that cGMP binds to various downstream signal transduction effector proteins and ligand-gated ion channels to regulate a multitude of physiological processes and pathological conditions such as cerebellar motor control, cardiac failure, pulmonary hypertension, erectile dysfunction, gut peristalsis and neurodegeneration<sup>2–5</sup>.

Soluble guanylate cyclase (sGC) is a heterodimeric protein consisting of  $\alpha$  and  $\beta$  subunits. In human, four different subunits ( $\alpha_1$ ,  $\alpha_2$ ,  $\beta_1$  and  $\beta_2$ ) exist which assemble to make  $\alpha_1\beta_1$ ,  $\alpha_2\beta_1$  and  $\alpha_1\beta_2$  heterodimer isoforms while  $\alpha_1\alpha_1$  &  $\beta_1\beta_1$  homodimers also exist. Out of all these isoforms, only two sGC isozyme,  $\alpha_1\beta_1$  and  $\alpha_2\beta_1$ , are well characterized in human<sup>1,6</sup>.  $\alpha_1\beta_1$  is an abundantly expressed cytosolic protein and a well-established member in the NO signaling cascade of mammals and insects. In contrast, the  $\alpha_2\beta_1$  isoform is membrane-associated and despite common ligand-binding features they respond differently and play distinct roles in the body. Although sGC is most ubiquitously expressed in lung, brain, muscle, spleen, brain and heart its distribution across tissues is isoform-specific. Levels of cGMP at a particular site in the human body can be regulated by modulating the expression of specific guanylyl cyclase isoforms<sup>7</sup>.

Each subunit forms a multi-domain protein consisting of four globular domains. The N-terminus of the  $\beta_1$  subunit harbors a heme-NO binding (H-NOX) domain; a conserved structural region critical for NO and O<sub>2</sub>

<sup>1</sup>Department of Biosciences, COMSATS University, Islamabad, 45550, Pakistan. <sup>2</sup>Department of Biology, University of Crete, 70013, Heraklion, Greece. <sup>3</sup>Department of Biostatistics and Medical Informatics, Acibadem M. A. A. University, Istanbul, 34752, Turkey. <sup>4</sup>Department of Pharmaceutical Chemistry, Biruni Universitesi, Istanbul, 34010, Turkey. <sup>5</sup>Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology-Hellas (IMBB-FORTH), 70013, Heraklion, Greece. ✉e-mail: [abdulraufsher@gmail.com](mailto:abdulraufsher@gmail.com); [kokkinid@imbb.forth.gr](mailto:kokkinid@imbb.forth.gr)

binding. The corresponding region on the  $\alpha_1$  subunit is known as pseudo-HNOX domain because it lacks the heme-binding motif<sup>8–10</sup>. Binding of NO to the heme group of the HNOX domain induces a conformational switch in sGC. The NO-binding signal is disseminated to the catalytic region through the Per/Arnt/Sim (PAS) and long helical coiled coil domains that form the dimeric backbone of the sGC quaternary structure. Both domains are known to be involved in NO-induced signal transduction and heterodimerization<sup>8,11,12</sup>. The C-terminal catalytic domains of the  $\alpha_1$  and  $\beta_1$  subunits are responsible for the cyclization and synthesis of cGMP from GTP<sup>13,14</sup>.

The cyclase domain of sGC is a closely related homolog of the catalytic domains of soluble adenylate cyclase (sAC) and particulate guanylate cyclase (pGC). Soluble guanylate cyclase (sGC) is categorized into the class III purine nucleotidyl cyclase family like all aforementioned closely related homologs<sup>15,16</sup>. All these proteins convert purine-based ribonucleotide triphosphates (GTP, ATP) to their respective cyclic ribonucleotide monophosphates (cGMP, cAMP) in a similar manner<sup>17</sup>.

The three dimensional (3D) structures of the catalytic domains of sGC in apo-form (Protein Data Bank (PDB) ID: 3UVJ)<sup>14</sup> and the ATP-bound complex of sAC in *Rattus norvegicus* (PDB ID: 1CJK)<sup>18</sup> reveal a highly conserved secondary structure organization of the binding pockets. The catalytic domain of both proteins adopts a wreath-like head to tail orientation. It lies at the dimeric interface of the  $\alpha_1$  and  $\beta_1$  subunits and it is composed of residues from both catalytic subdomains<sup>19</sup>. The substrate binding site of the cyclase domain consists of the conserved D486 and D530 residues from the  $\alpha_1$  subunit and 548 N, 551 S, 552 R residues from the  $\beta_1$  subunit that establish electrostatic interactions with the phosphate moiety of GTP and metals<sup>14</sup>.

In human, both homodimeric and heterodimeric isoforms of the sGC cyclase domain in the apo-form have been resolved crystallographically but not the GTP-bound heterodimeric form<sup>1,14</sup>. The catalytic mechanism of the sGC cyclase domain is inferred from the crystal structures of the active, ATP-bound sAC complexes from human and *Rattus norvegicus* but no direct information of the active hsGC cyclase domain in complex with GTP is available.

To investigate the comparative dynamics behavior of the apo-, GTP- and cGMP-bound hsGC cyclase- $\alpha_1\beta_1$  heterodimers, we carried out *in silico* structural analysis and molecular dynamics (MD) simulations. Extensive structural information concerning conserved active site motifs and ribonucleotide triphosphate substrate binding residues was gleaned from the sAC catalytic domain by superposing it with the heterodimeric hsGC catalytic domain complexes. The binding conformation of ATP from sAC was used to model the GTP- and cGMP-bound hsGC catalytic domain complexes. The complexes were then subjected to a comprehensive MD simulation analysis in order to monitor the stability and binding modes pattern of catalytic interface residues. Furthermore, principal component analysis was performed to identify structural transitions in the presence of GTP or cGMP that contribute significantly to the opening and closing of the wreath-like hsGC cyclase heterodimer complex. Finally, we discuss the pre- (GTP-bound hsGC cyclase domain) to post-catalysis (cGMP-bound hsGC cyclase domain) structural transition.

## Results

Pairwise sequence alignment (Fig. 1) of the *R. norvegicus* AC with the hsGC cyclase domain revealed that their heterodimeric chains share >30% sequence identity and >50% sequence similarity.

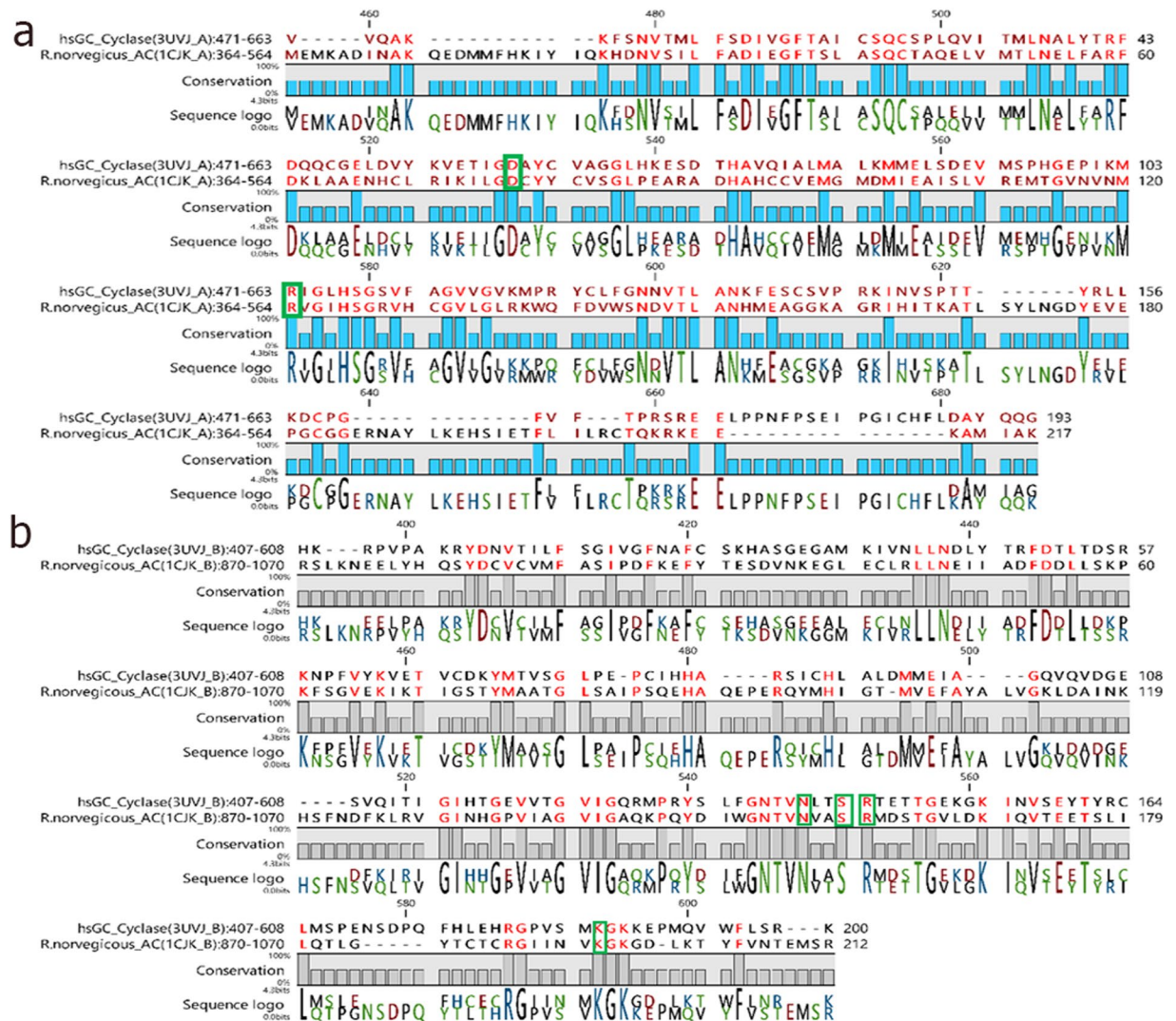
**Molecular docking.** The crystal structure of the human heterodimeric hsGC cyclase domain in the apo-form is available at the Protein Data Bank (PDB ID: 3UVJ). The Glide XP module was employed to dock the GTP and cGMP molecules within the binding pocket of the hsGC cyclase domain. The docking protocol was validated by docking ATP within the *R. norvegicus* sAC catalytic domain and superposing with the already reported ATP-bound complex of sAC (PDB ID: 1CJK)<sup>18</sup> (Supplementary Fig. 1a). Binding pocket residues ( $\alpha$ D396,  $\alpha$ D441,  $\beta$ N1025,  $\beta$ R1029 and  $\beta$ K1065), involved in the binding of nucleotide tri-phosphate (NTP) within the active site of adenylate cyclase, are conserved at positions  $\alpha$ D486,  $\alpha$ D530,  $\beta$ N548,  $\beta$ R552 and  $\beta$ K593 of hsGC (Supplementary Fig. 1b). Conserved catalytic residues highlighted in Fig. 1. The conformation of the Glide-docked ATP was very similar to the experimental one with an RMSD 0.4 Å. The ATP-binding pocket grid coordinates that were extracted from the *R. norvegicus* sAC catalytic domain were used for the docking of GTP and cGMP to the sGC cyclase domain. Secondary structure elements such as helix1 (h1), helix2 (h2),  $\beta$  hairpin loop1 ( $\beta$ 3L1 $\beta$ 4),  $\beta$  hairpin loop2 ( $\beta$ 6L2 $\beta$ 7) of  $\alpha$  and  $\beta$  chains constitute the ligand-binding pocket and the dimeric interface of the hsGC catalytic dimer (Fig. 2a).

The Glide XP dock scores for GTP and cGMP are  $-6.718$  and  $-7.295$  kcal/mol respectively.

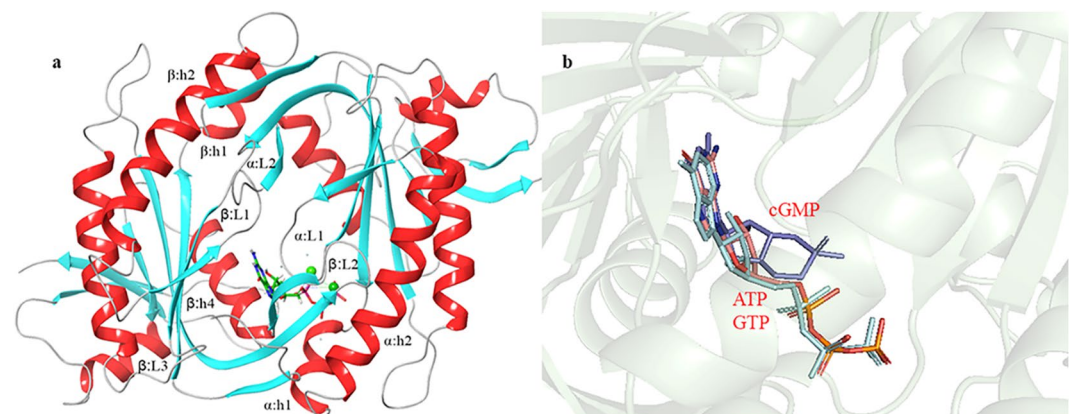
The GTP- and cGMP-docked complexes were again superposed to the ATP-bound *R. norvegicus* sAC catalytic domain to validate the best docked conformations (Fig. 2b). In the GTP-hsGC cyclase domain complex highlighted in Fig. 3a. Residues D486, V488, F490 and D530 from  $\alpha$  chain and N548, S551, R552 and T514 from  $\beta$  chain are encapsulating the GTP molecule within the catalytic pocket through various molecular interactions (Fig. 3b). Of all the GTP-interacting residues, D486 and D530 from the  $\alpha$  chain of the hsGC cyclase showed  $Mg^{2+}$ -mediated interactions with the GTP. Beside these conserved residues, F490 showed hydrogen bond interactions and V488 weak van der Waals interactions with the GTP molecule. All  $\beta$ -chain residues i.e. N548, S551, R552 and T514 formed a dense network of hydrogen bonds with the GTP molecule (Fig. 3b).

Similarly, cGMP adopted a similar orientation within the dimeric pocket of the hsGC cyclase domain (Fig. 3c) but formed interactions with only six binding pocket residues. The  $Mg^{2+}$ -mediated ionic interactions of cGMP with D486 and D530 of the  $\alpha$  chain persisted while only a single hydrogen bond with N548 and three van der Waals interactions with S551, R552 and M443 of the  $\beta$  chain were observed upon binding of cGMP (Fig. 3d).

Similar indirect interactions of the hsGC cyclase  $\alpha$  chain residues (D486 and D530) with GTP and cGMP through two  $Mg^{2+}$  ions have also been reported in the adenylate cyclase-ATP complex<sup>12,14</sup>. This suggests a conserved role of these residues in the activity of the class III purine nucleotidyl cyclase family of enzymes. In

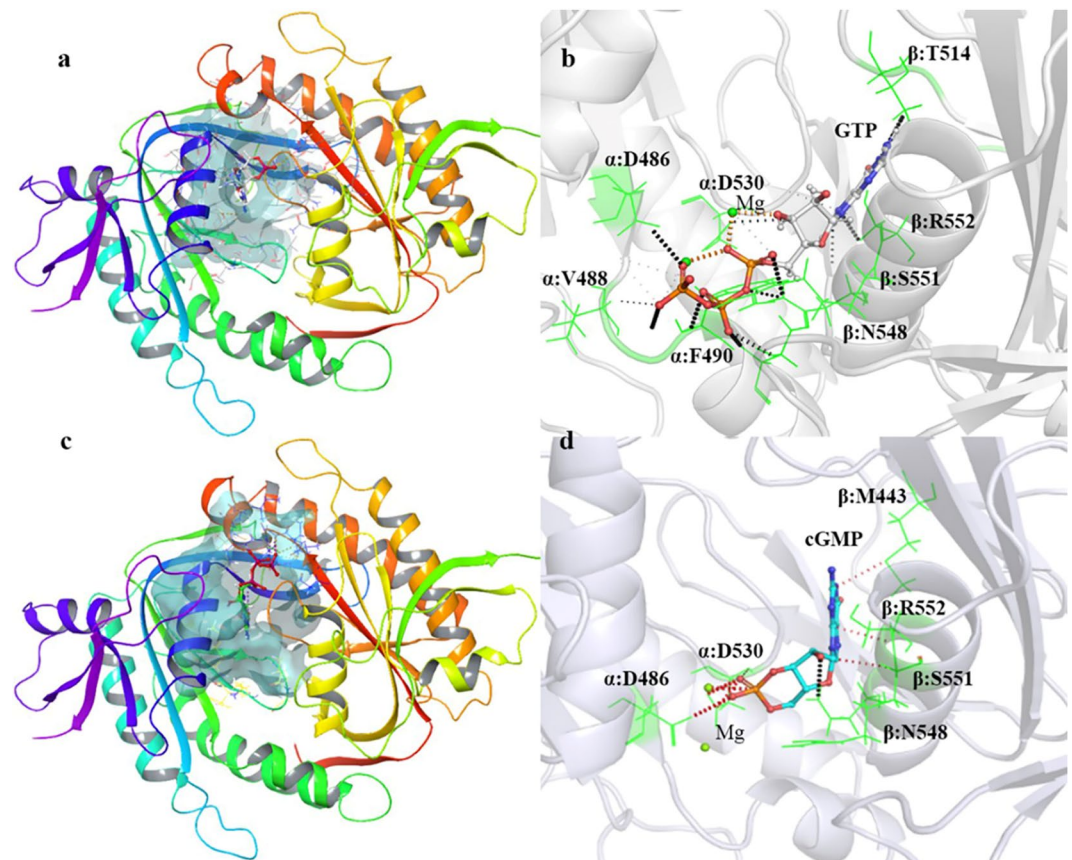


**Figure 1.** Sequence alignment of the  $\alpha$  (a) and  $\beta$  (b) chains of the human soluble guanylate cyclase heterodimer (*hsGC*) with the closely related *R.norvegicus* adenylate cyclase homodimer (AC). Conserved catalytically active residues are highlighted with a green box.



**Figure 2.** (a) The *hsGC* catalytic dimer. Conserved secondary structure elements of the *hsGC* catalytic dimer are shown in red and cyan. (b) docked GTP (pink) and cGMP (blue) illustrate the very similar orientation to the ATP (cyan) in *R.norvegicus* AC.





**Figure 3.** (a) Docked complex of GTP with the *hsGC* catalytic heterodimer. The active site of the catalytic heterodimer is shown in transparent cyan. (b) Interactions of GTP (shown as ball and stick) with the active site residues (shown as green colored lines). Hydrogen bonds are shown as black dotted lines (dense: strong; light: weak) and van der Waals interactions as orange dotted lines. (c) Docked complex of cGMP with the *hsGC* catalytic heterodimer. The active site of the catalytic heterodimer is shown in transparent cyan. (d) Interactions of cGMP (shown as ball and stick) with active site residues (shown as green colored lines). Hydrogen bonds are shown as black dotted lines (dense: strong; light: weak) and van der Waals interactions as red dotted lines.

addition to this, N548, S551, and R552 from the *hsGC* cyclase  $\alpha$  chain are also conserved and involved in the stabilization of the GTP and cGMP molecules in the cyclase domain.

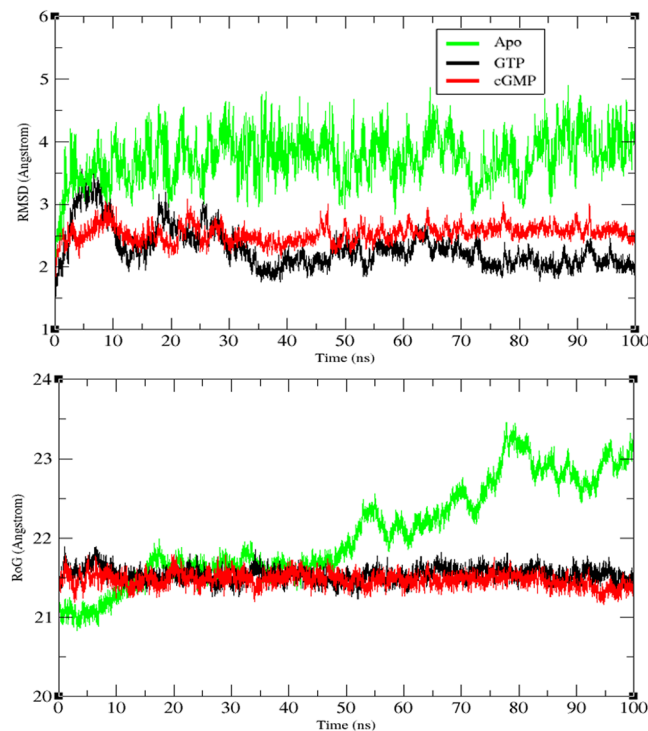
**Stability analysis of MD simulations.** *Root-mean-square deviation.* Stability analysis was carried out for the apo, GTP- and cGMP-bound cyclase  $\alpha\beta$  complexes to identify the changes in trajectories. The root-mean-square deviation (RMSD) and root mean square fluctuation (RMSF) were calculated for the backbone atoms of all systems using the initial apo crystal structure as a reference. Afterwards, the radius of gyration (Rog) was also calculated for all atoms of our systems.

The RMSD analysis showed that the system with the highest structural variation was the apo system with an average RMSD of 3.9 Å and a high fluctuation rate during the simulation (Fig. 4).

The GTP-bound sGC cyclase heterodimer underwent a large structural change (RMSD 3.5 Å) at the beginning of the simulation (5–7 ns) and gradually attained stability at 20–100 ns with an average RMSD of 2.1 Å. The difference of RMSD observed during the first 20 ns is possibly associated with the closed pocket conformation of the cyclase heterodimer. In next 30–100 ns, the closed pocket conformation becomes stable which helps in the gradual decrease in RMSD to 2.1 Å.

Interestingly, the cGMP-bound cyclase- $\alpha\beta$  heterodimer showed an average RMSD of 2.5 Å that remained relatively stable throughout the entire simulation. Of note is the 0.6 Å RMSD difference observed between the GTP- and cGMP-bound systems during the last 20 ns of the simulations that may be attributed to the opening and closing of the wreath-like cyclase heterodimer. The presence of GTP at the substrate binding site stabilizes the dimeric pocket of the *hsGC* cyclase heterodimer with a robust network of interactions that are partially absent in the presence of cGMP where a slight opening of the pocket is observed.

**Radius of gyration (Rog).** The radius of gyration is another stability indicator which represents the mass-weighted RMS distance of a group of atoms from their common center of mass and is used to estimate the global dimension of proteins.



**Figure 4.** Root mean square deviation (RMSD) and radius of gyration (Rog) of the *hsGC* catalytic dimer in the apo form and in complex with GTP and cGMP over the course of 100 ns simulations.

The analysis showed that both substrate-bound complexes were stable during the simulation. The Rog values ranged between 21–21.5 Å, signifying moderate fold movement among GTP/cGMP bound systems (Fig. 4). Results suggest that the GTP- and cGMP-bound dimeric cyclase remains compact over the course of simulation while the apo-cyclase system experiences broader configurational variations up to 23.3 Å. The Rog analysis results are consistent with the RMSD analysis.

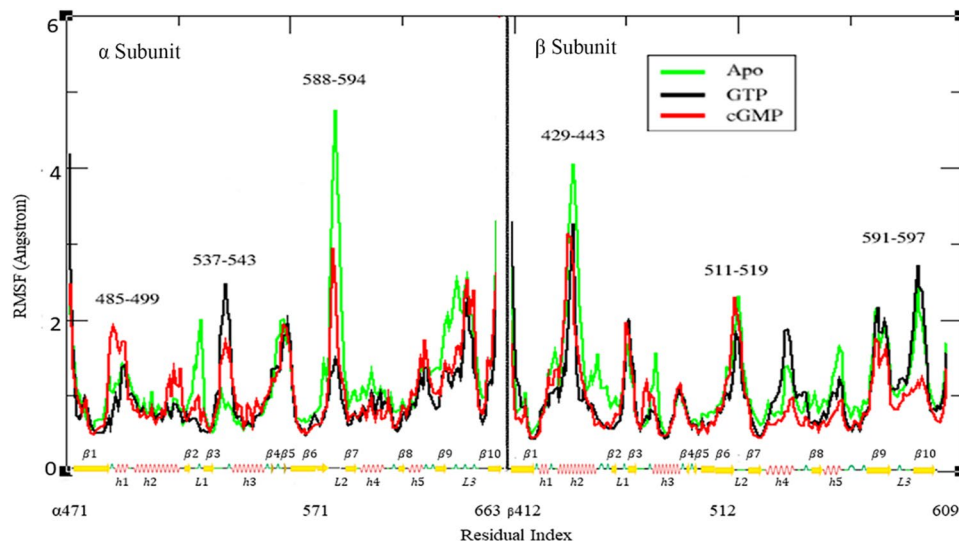
**Root mean square fluctuation.** The root-mean-square fluctuation (RMSF) analysis is indispensable for the characterization of structural fluctuations within local regions of the *hsGC* cyclase dimer after ligand (GTP, cGMP) binding. The substrate-anchoring regions of the *hsGC* catalytic dimer that interact with GTP and cGMP include  $\alpha$ :h1,  $\beta$  hairpin loop  $\alpha$ :L1,  $\beta$  hairpin loop  $\beta$ :L2 and  $\beta$ :h4. During simulations we monitored conformational changes in these important secondary structure elements. RMSF analysis of the *hsGC* cyclase dimer in the absence of ligand revealed 1.4 Å, 4.8 Å and 2.3 Å deviations in  $\alpha$ :h1,  $\alpha$ :L2 and  $\beta$ :L2 regions, respectively, indicating high conformational flexibility (Fig. 5).

In contrast to the apo state, the GTP- and cGMP-ligated complexes showed fewer residual transitions. The GTP- and the cGMP-bound *hsGC* cyclase  $\alpha\beta$  systems showed distinct patterns of RMSF peaks different from each other. In the human sGC catalytic dimer complexed with GTP, structurally significant regions, especially  $\alpha$ :h1,  $\alpha$ :L2 and  $\beta$ :L2 showed RMSF peaks of 1.4 Å, 1.5 Å and 1.8 Å, respectively, indicating a robust increase of the stability of the system upon GTP binding compared to the apo-state (Fig. 5). In the presence of cGMP, the  $\alpha$ :h1,  $\alpha$ :L2 and the  $\beta$ :L2 regions showed RMSF peaks at 2 Å, 2.5 Å and 2.3 Å, respectively, indicating a less compact arrangement compared to the GTP-bound system (Fig. 5).

The difference of the RMSF peaks in structurally critical regions of the human cyclase dimer in the presence of GTP indicates a more stable and/ or rigid organization of the secondary structure elements potentially favoring GTP binding and catalysis. The enhanced flexibility of the cGMP-bound system might be associated with the release of cGMP, after GTP catalysis.

**Hydrogen bond analysis of the ligand-bound complexes.** Both cGMP and GTP bound within the cyclase pocket of the *hsGC* catalytic domain were subjected to hydrogen bond analysis to assess the occurrence of hydrogen bonds between the ligand and the side chain atoms of the binding pocket residues over the 50 ns simulations. The occupancy of hydrogen bonds throughout the simulation signifies the contribution of hydrogen bond interactions to the structural stability of the complexes.

The GTP-bound *hsGC* catalytic dimer showed a greater number of hydrogen bond interactions compared to the cGMP system (Table 1). The GTP- and cGMP-mediated hydrogen bond network can be mediated through the heterocyclic guanine (N1, N2, N3, O6, N7), ribose sugar (O3', O2', O5, O3), and phosphate (O1B, O2A, O2B, O1G, O2G, O3G, O7) group atoms (some donor-acceptor pairs in Table 1 have low occupancy but they are included either because the cumulative effect on the same residue is significant or they are residues important in the subsequent analyses).



**Figure 5.** Root mean square fluctuation (RMSF) of the *hsGC* catalytic dimer in the apo form and in complex with GTP and cGMP. The x-axis of the graph is representing the sequence of the chains  $\alpha$  and  $\beta$  with the secondary structure elements highlighted accordingly while the y-axis shows the residual fluctuation in Angstroms ( $\text{\AA}$ ).

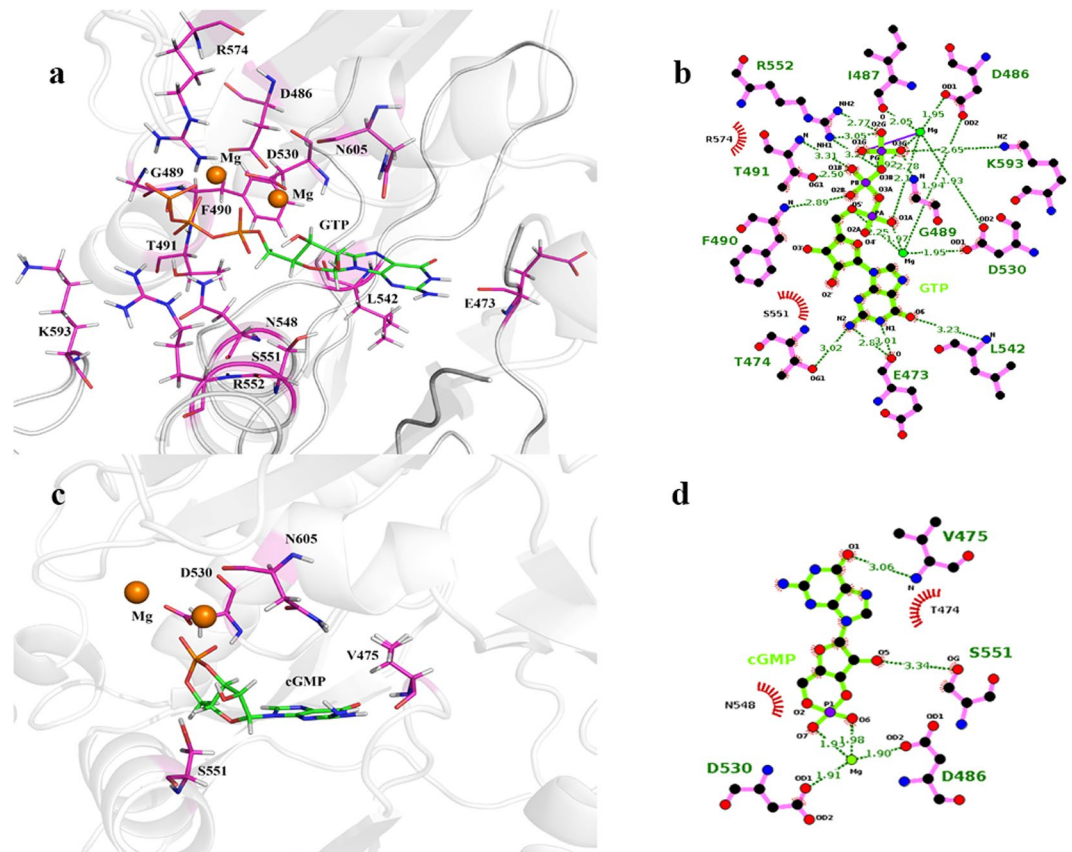
Ligands	Acceptor	Donor	Percentage Occupancy	Average Distance	Average Angle	
<p style="text-align: center;"><b>GTP</b></p>	GTP@O1B	$\alpha$ :T491@O	99.64	2.62	167.51	
	GTP@O3G	$\alpha$ :G489@N	73.90	2.82	160.59	
	$\beta$ :E473@O	GTP@N2	66.78	2.84	155.18	
	GTP@O2B	$\alpha$ :F490@N	59.42	2.86	150.52	
	GTP@O1G	$\alpha$ :R574@N	52.20	2.83	157.66	
	$\beta$ :E473@O	GTP@N1	44.69	2.87	150.11	
	GTP@O2G	$\beta$ :K593@NZ	40.70	2.78	156.20	
	GTP@O2A	$\beta$ :N548@N	30.73	2.81	159.06	
	GTP@O6	$\beta$ :L542@N	26.79	2.88	162.38	
	GTP@O3'	$\beta$ :S551@O	17.69	2.76	161.33	
	GTP@O3G	$\alpha$ :R574@N	14.87	2.86	160.31	
	GTP@O2'	$\beta$ :S551@O	10.84	2.77	158.10	
	GTP@N7	$\beta$ :L542@N	9.15	2.92	157.74	
	GTP@O2A	$\beta$ :R552@N	6.90	2.78	156.71	
	GTP@O3G	$\beta$ :K593@N	6.65	2.77	156.42	
	GTP@O2G	$\alpha$ :R574@N	6.62	2.85	155.48	
	$\beta$ :T474@O	GTP@N2	5.75	2.88	153.39	
	<p style="text-align: center;"><b>cGMP</b></p>	GMP@O1	$\beta$ :V475@N	65.29	2.86	159.68
		GMP@O5	$\beta$ :S551@O	32.67	2.80	162.38
$\beta$ :V475@O		GMP@N1	21.23	2.80	162.09	
GMP@O7		$\alpha$ :D530@N	17.37	2.90	160.61	
$\beta$ :V475@O		GMP@N2	0.20	2.91	149.73	
GMP@O3		$\beta$ :S551@O	0.16	2.75	159.09	
GMP@N3		$\beta$ :S551@O	0.11	2.86	162.96	

**Table 1.** Comparison of the GTP- and cGMP-bound catalytic heterodimers hydrogen bond interactions of the nucleotidyl substrates with the binding pocket residues.

Among the aforementioned groups, the nitrogenous base atoms of GTP interact with E473, L542, while the nitrogenous base of cGMP interacts only with V475. The 5' carbon sugar group atoms established a robust interaction with S551 consistent throughout the 50 ns simulation in both GTP- and cGMP-bound complexes. The phosphate group atoms of GTP fostered a complex hydrogen bond network with T491, G489, F490, R574, K593, N548 and R552 (high to low fractional value), partially absent in case of cGMP because of the release of pyrophosphate during the catalysis of GTP to cGMP. The binding pocket residues of the *hsGC* catalytic dimer that are lying in close proximity to GTP and interact with it through non-covalent interactions are shown in Fig. 6a,b. The electrostatic interactions of the GTP-bound *hsGC* cyclase dimer are consistent with previous findings<sup>7,14</sup>.

In a study reported by Kleinboelting *et al.*, 2014, aspartate residues critical for metal ion binding and indirect interactions of NTPs with the NSR motif (Asn, Ser and Arg) are also conserved in the current study<sup>19</sup>. In addition to the conserved residues, 489 G, 490 F, 491 T at helix 1 ( $h_1$ ), R574 at  $\beta$ -sheet 5 ( $\beta_5$ ) of the  $\alpha$  chain and 473E at  $\beta$ -hairpin loop 1 ( $\beta 3L1\beta 4$ ), 542 L at  $\beta$ -sheet 7 ( $\beta 7$ ) and 593 K at  $\beta$ -hairpin loop 3 ( $\beta 9L3\beta 10$ ) of the  $\beta$  chain are forming a hub that plays a key role in the encapsulation of GTP within the cyclase  $\alpha\beta$  dimer for the cyclization event.





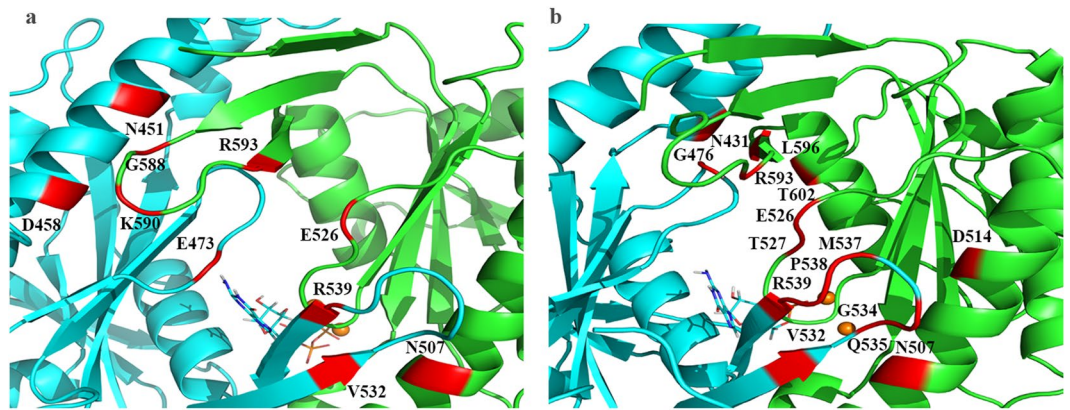
**Figure 6.** (a) Graphical representation of the binding pocket residues (purple) of *hsGC* that interact with GTP throughout the 50 ns MD simulation. (b) LigPlot two-dimensional (2D) ligand interaction diagram of the hydrogen bond interactions between the catalytic pocket residues of the *hsGC* heterodimer and GTP. (c) Graphical representation of the binding pocket residues (purple) of *hsGC* that interact with cGMP throughout the simulation. (d) LigPlot two-dimensional (2D) ligand interaction diagram of the hydrogen bond interactions between the catalytic pocket residues of the *hsGC* heterodimer and cGMP.

The conformation of the binding pocket is determined by four critical structural regions;  $\alpha$ :h1,  $\alpha$ : $\beta$ 5,  $\beta$ :h4,  $\beta$ :L3. These secondary structure elements are forming the dimeric interface of the sGC catalytic heterodimer. The highest hydrogen bond occurrence in the GTP-bound cyclase  $\alpha\beta$  complex might assist these structural components to attain a close pocket conformation. Conversely, in the cGMP-ligated cyclase  $\alpha\beta$  system the reduced number of hydrogen bond interactions at the dimeric interface results in the opening of the cyclase pocket (Fig. 6c,d).

**Substrate-mediated protein-protein interactions in the *hsGC*  $\alpha\beta$  cyclase dimer.** To understand how local changes upon GTP and cGMP binding to the cyclase pocket induce large scale conformational changes in the structure, hydrogen bond analysis of the dimeric interface residues was carried out. In the GTP-bound system, N507 at helix 2 (h2), E526 at loop 1 (L1), G588, K590 at loop 2 (L2), R593 at  $\beta$ -sheet 7 ( $\beta$ 7) of the  $\alpha$  chain are interacting through hydrogen bonds with V532 at  $\beta$ -sheet 6 ( $\beta$ 6), R539 at  $\beta$ -sheet 7 ( $\beta$ 7), E473 at loop 1 (L1), N541 and D548 at helix 2 (h2) of the  $\beta$  chain (Fig. 7). Occupancies, average distances and average angles of the hydrogen bonds between the dimeric interface residues are shown in Table 2. In the GTP-bound *hsGC* cyclase, the abundance of interactions in the dimeric interface may contribute to the tight packing of the cyclase binding pocket and the catalysis of GTP.

The number of inter-residue interactions between the  $\alpha$  and  $\beta$  chains of the cGMP-bound *hsGC* was higher, with only N507 at helix 2 (h2), E526 at loop 1 (L1), and R593 at  $\beta$ -sheet 7 ( $\beta$ 7) of the  $\alpha$  chain interacting with V532 at  $\beta$ -sheet 6 ( $\beta$ 6), R539 at  $\beta$ -sheet 7 ( $\beta$ 7) and G476 at loop 1 (L1) of the  $\beta$  chain (Fig. 7b) common with the GTP-bound complex but also additional interactions between T527 at  $\alpha$ :L1, R593 at  $\alpha$ :L2 of  $\alpha$  subunit with M537, P538, R539 at the  $\beta$ :L2 region of the  $\beta$  subunit (Table 2). A possible reason for this protein-protein interaction rearrangement is the inability of  $\beta$ -hairpin loop 2 ( $\beta$ 6L2 $\beta$ 7) of  $\alpha$  chain to engage with h2 of  $\beta$  chain opening the catalytic pocket which, in turn, results in the opening of the binding pocket.

**Per-residue decomposition of binding free energy.** In order to further delineate the role of key residues that influence the overall binding free energy of the systems; per-residue energy decomposition analysis was performed for the identification of the interaction spectra of the GTP/cGMP-bound *hsGC* catalytic dimer. The last 20 ns snapshots of the 100 ns simulations were considered. The binding affinity of each residue in the ligand-receptor complex depends on the binding free energy ( $\Delta G$ ) contribution, which can be positive or



**Figure 7.** (a) Graphical representation of the GTP-bound *hsGC* catalytic dimer showing a closed-state conformation. Interacting residues of  $\alpha$  (green) and  $\beta$  (cyan) chain are highlighted in red. (b) Graphical representation of the cGMP-bound *hsGC* catalytic dimer showing an open-state conformation. Interacting residues of  $\alpha$  (green) and  $\beta$  (cyan) chain are highlighted in red.

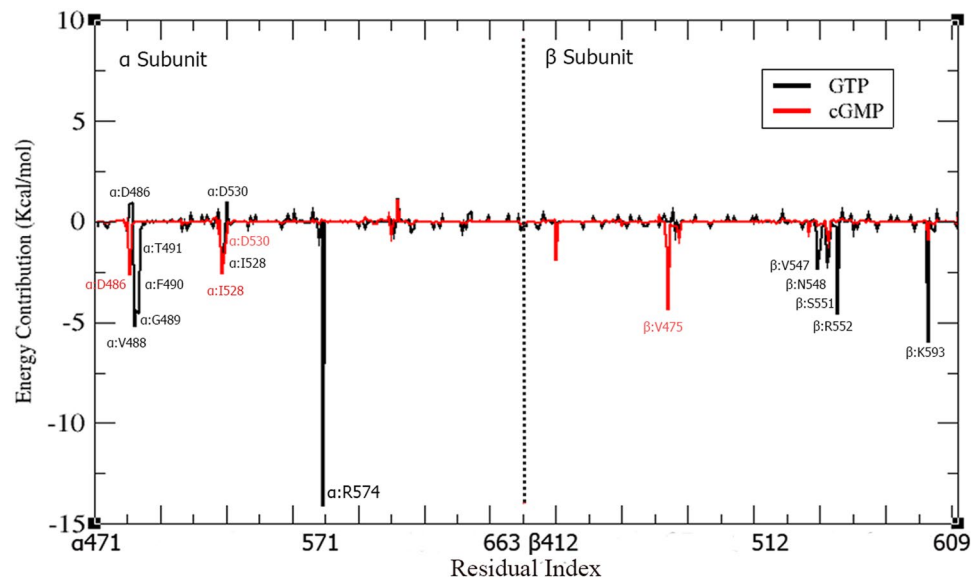
Complexes	Acceptor	Donor	Percent Occupancy	Average Distance	Average Angle
Cyclase-GTP	$\alpha$ :E526@O	$\beta$ :R539@N	87.22	2.80	158.35
	$\beta$ :E473@O	$\alpha$ :R593@N	74.77	2.82	158.39
	$\beta$ :D458@O	$\alpha$ :K590@N	53.40	2.90	165.89
	$\beta$ :N451@O	$\alpha$ :G588@N	43.01	2.90	154.60
	$\beta$ :V532@O	$\alpha$ :N507@N	27.74	2.88	161.13
Cyclase-cGMP	$\alpha$ :D514@O	$\beta$ :Q535@N	81.62	2.83	159.18
	$\beta$ :M537@O	$\alpha$ :T527@O	80.29	2.74	162.54
	$\beta$ :P538@O	$\alpha$ :R593@N	56.53	2.81	152.30
	$\alpha$ :N507@O	$\beta$ :G534@N	51.60	2.86	153.87
	$\alpha$ :R593@O	$\beta$ :G476@N	35.17	2.88	153.52
	$\alpha$ :T527@O	$\beta$ :R539@N	30.07	2.83	150.91
	$\beta$ :N431@O	$\alpha$ :T602@O	25.37	2.75	162.06
	$\beta$ :V532@O	$\alpha$ :N507@N	25.37	2.87	164.44
	$\alpha$ :E526@O	$\beta$ :R539@N	23.88	2.79	155.82

**Table 2.** Comparison of the hydrogen bond interdomain protein-protein interactions of GTP- and cGMP-bound cyclase complexes.

negative. Negative  $\Delta G$  values of a residue indicate higher binding affinity. Conversely, positive  $\Delta G$  values are indicative of a repulsive effect that in turn doesn't facilitate the catalytic activity of the receptor. In GTP-bound *hsGC* catalytic dimer, the most negative contribution to the binding free energy was observed for the residues spanning the following  $\alpha$  and  $\beta$  subdomain regions of cyclase;  $\alpha$ :h1 (V488, G489, F490, T491),  $\alpha$ : $\beta$ 5 (R574),  $\alpha$ : $\beta$ 2L1 $\beta$ 3 (I528), and  $\beta$ :h4 (V547, N548, S551, R552), L3 (K593) (Fig. 8). These residues play a crucial role in the catalytic activity of the *hsGC* cyclase heterodimer. However, in the cGMP-bound *hsGC* catalytic dimer, a different pattern emerged, with  $\alpha$ : $\beta$ 1 (D486),  $\alpha$ : $\beta$ 2L1 $\beta$ 3 (I528, D530), and  $\beta$ : $\beta$ 2L1 $\beta$ 3 (V475) showing the most negative contribution to the binding free energy (Fig. 8). Table 3 depicts the average binding free energy of each residue in the GTP/cGMP-bound *hsGC* cyclase complexes. Interestingly, the D486 and D530 residues of the cyclase  $\alpha$  subdomain showed negative  $\Delta G$  values  $-2.60$  Kcal/mol and  $-0.54$  Kcal/mol, respectively, upon cGMP binding whereas, in case of GTP binding, the same residues showed repulsive effects. The key residues that contribute significantly to the binding of GTP in the GTP-bound *hsGC* catalytic dimer complex;  $\alpha$ :R574 ( $-14.10$  Kcal/mol),  $\alpha$ :V488 ( $-5.18$  Kcal/mol),  $\alpha$ :G489 ( $-4.43$  Kcal/mol),  $\alpha$ :F490 ( $-4.50$  Kcal/mol),  $\beta$ :F543 ( $-2.34$  Kcal/mol),  $\beta$ :V547 ( $-2.03$  Kcal/mol),  $\beta$ :R552 ( $-4.60$  Kcal/mol), and  $\beta$ :K593 ( $-6.00$  Kcal/mol), showed less negative or positive  $\Delta G$  values in the cGMP-bound complex (Table 3). These per-residue energy decompositions findings are consistent with the binding mode analysis, which strengthens the argument that involvement of more residues in the case of the GTP-bound cyclase  $\alpha\beta$  complex contributes to the closed conformation of the binding pocket. The closed pocket structural arrangement is destabilized by the release of the pyrophosphate during the GTP to cGMP catalysis, leading to a decrease of the overall binding free energy and opening of the cyclase pocket.

**Principal Component Analysis (PCA).** We have applied principal component analysis to compare the overall conformational pattern of movements of the cyclase  $\alpha\beta$  heterodimer under the influence of GTP and cGMP, with the apo system. The trajectories of all systems (cyclase  $\alpha\beta$ -apo, cyclase  $\alpha\beta$ -GTP, cyclase  $\alpha\beta$ -cGMP) were processed to calculate eigenvectors that correspond to the dominant structural movements during a simulation. An essential dynamics approach was applied to the backbone atoms of all systems to understand their structural transitions. The most notable movements observed in the top two eigenvectors were the motion of  $\alpha$ :h1,  $\beta$ -hairpin loop 2 ( $\alpha$ : $\beta$ 6- $\beta$ 7),  $\beta$ h1,  $\beta$ -hairpin loop 2 ( $\beta$ : $\beta$ 6- $\beta$ 7),  $\beta$ :L3 regions and account for 85% of all the motions.



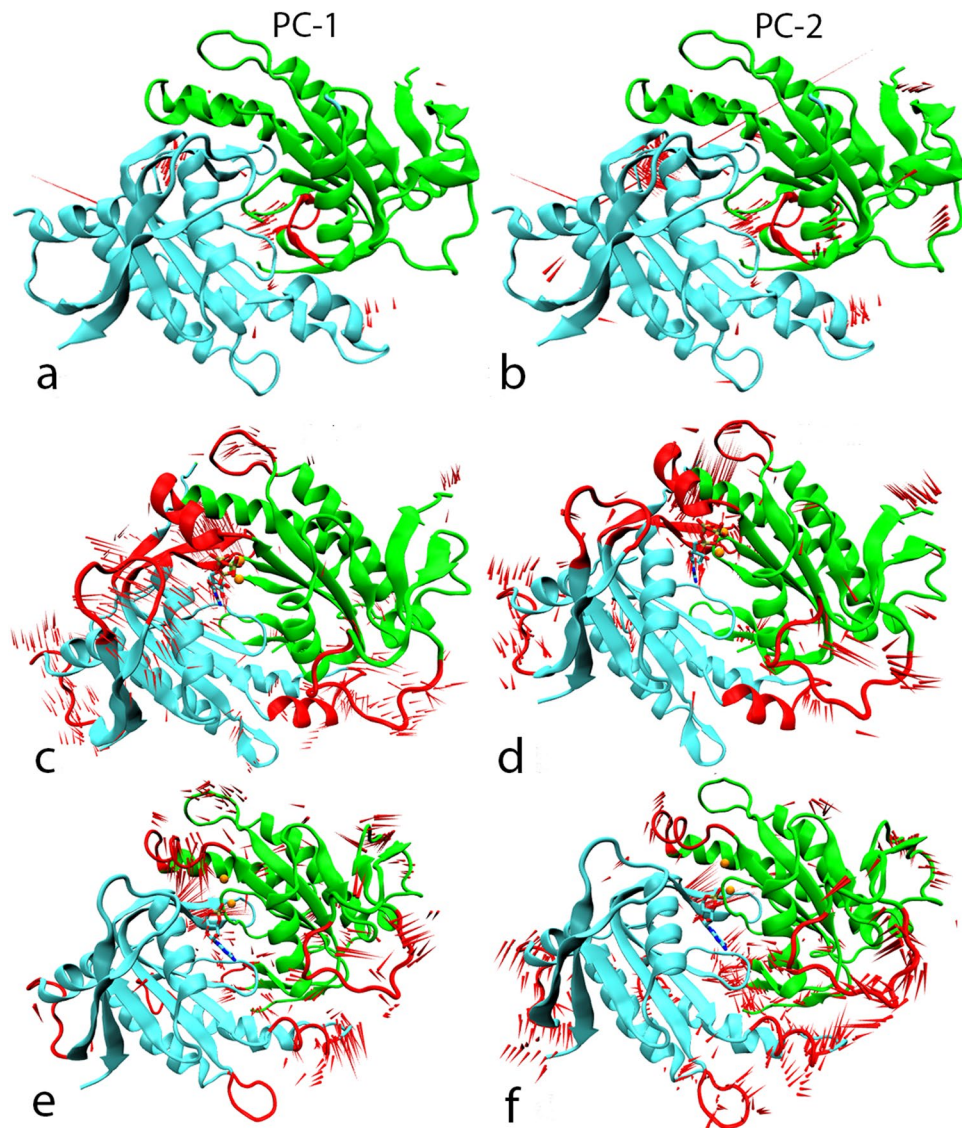


**Figure 8.** Per-residue MM-PBSA binding free energy contribution of the GTP-bound *hsGC* catalytic dimer (black) and cGMP-bound *hsGC* catalytic dimer (red). The residues showing the highest fluctuations are highlighted.

Residues	Cyclase-GTP PRED $\Delta G$ (kcal/mol)	Cyclase-cGMP PRED $\Delta G$ (kcal/mol)
$\alpha$ :D486	-0.90	-2.60
$\alpha$ :V488	-5.18	0.01
$\alpha$ :G489	-4.43	-0.01
$\alpha$ :F490	-4.50	-0.01
$\alpha$ :T491	-0.37	0.01
$\alpha$ :I528	-2.50	-2.60
$\alpha$ :G529	-1.08	-1.60
$\alpha$ :D530	-1.00	-0.54
$\alpha$ :R574	-14.10	0.05
$\beta$ :T474	0.05	-1.24
$\beta$ :V475	-0.04	-4.40
$\beta$ :F543	-2.33	-0.17
$\beta$ :G544	-1.41	-0.10
$\beta$ :N545	-0.22	-0.12
$\beta$ :T546	-0.25	0.04
$\beta$ :V547	-2.03	-0.60
$\beta$ :N548	-1.71	-0.10
$\beta$ :L549	-0.02	0.02
$\beta$ :T550	-0.25	0.13
$\beta$ :S551	-0.25	0.01
$\beta$ :R552	-4.60	0.03
$\beta$ :K593	-6.00	-1.00
$\beta$ :K595	-0.40	-0.10
$\beta$ :K596	-0.45	-0.10

**Table 3.** Per-residue energy decomposition (PRED) of GTP/cGMP-ligated *hsGC* cyclase systems.

Furthermore, the analysis assisted in understanding the conformational effect of cGMP and GTP on the cyclase dimer. The ProDy plugin<sup>20</sup> integrated within VMD<sup>21</sup> was employed to produce porcupine plots associated with the first two normal modes in each system (Fig. 9). The cyclase  $\alpha\beta$ -GTP complex showed substantial movement of the  $\alpha$ :h1,  $\beta$ : $\beta$ 6- $\beta$ 7 and  $\beta$ :L3 regions towards the binding pocket of GTP probably due to the strong electrostatic interaction network generated upon ligand binding. The hydrogen bond analysis of the GTP-*hsGC* complex discussed above also suggests a cooperative effect at the dimerization interface further stabilizing the closed-state conformation upon GTP binding and facilitating the catalysis of GTP to cGMP. The cyclase  $\alpha\beta$ -cGMP system



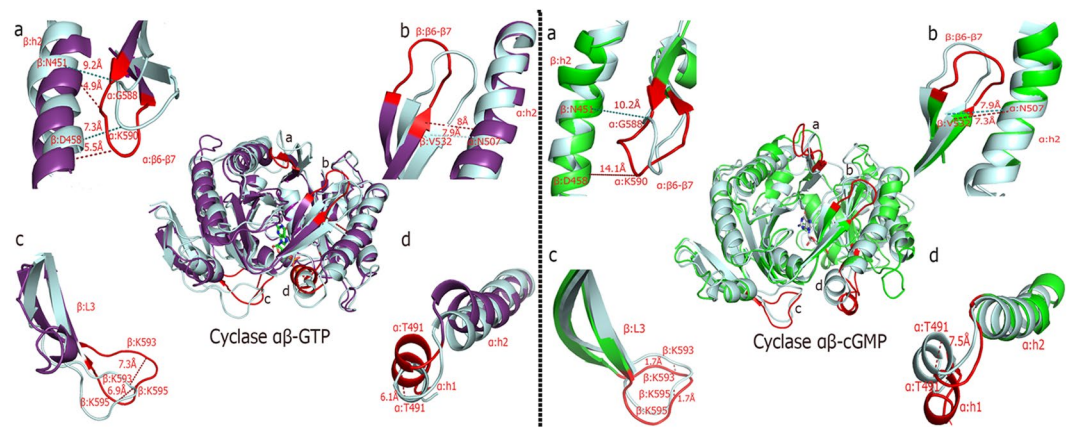
**Figure 9.** Porcupine plots of the first (PC-1) and second (PC-2) modes of motion of the apo (**a,b**) GTP-bound (**c,d**) and cGMP-bound (**e,f**) complexes of the hsGC catalytic dimer. Highly fluctuating regions are highlighted in red. The length of the red “needles” represents the degree of mobility.

first and second principal components show a distinct pattern in which the  $\alpha$ :h1,  $\beta$ :h1 and  $\alpha$ : $\beta$ 6- $\beta$ 7 regions are more distant from the binding pocket, with particularly the  $\beta$ -hairpin loop of  $\alpha$ : $\beta$ 6- $\beta$ 7 moving away from the dimer interface. The structural transition of these elements away from the binding pocket suggests a role in the post-catalysis structural dynamics. These findings are compatible with the stability analysis and earlier experimental studies<sup>7,19,22</sup>.

Furthermore, the structural transitions of the open (cGMP-bound) and closed state (GTP-bound) of the hsGC catalytic domain heterodimer are summarily illustrated in Fig. 10 by measuring the deviation of key residues and secondary structure elements from the reference/ initial state structure (docked GTP and cGMP-bound sGC cyclase) at the maximum transition points that were observed during PCA.

**Discussion and Conclusion.** The structural assembly and activation of the soluble guanylate cyclase has been extensively studied by conventional biochemical studies<sup>7,23,24</sup>. Recently a cryo-EM-driven multi-domain quaternary structure of the human cyclase domain has been proposed<sup>1</sup>. Most studies discuss the molecular architecture of sGC in prokaryotes and eukaryotes<sup>13,25</sup> with a few focusing on the molecular mechanism of the NO-based activation of sGC<sup>26</sup> but little is known about the conformational changes within the cyclase domain upon GTP binding and catalysis<sup>24,27</sup>. The current study investigates the structural events that take place within the the hsGC cyclase domain during the binding and the cyclization event of GTP to cGMP.

The soluble guanylate cyclase is a core signaling molecule of the cGMP signaling pathway but biophysical and structural properties driven by GTP and cGMP binding to the cyclase domain remain only partially characterized<sup>24,28</sup>. We characterized the structural behavior of the cyclase domain upon GTP and cGMP binding. We



**Figure 10.** Fluctuations/deviations (Å) of important secondary structure elements a)  $\beta$ :h2,  $\alpha$ : $\beta$ 6- $\beta$ 7 b)  $\beta$ : $\beta$ 6- $\beta$ 7,  $\alpha$ :h2 c)  $\beta$ :L3 d)  $\alpha$ :h1,  $\alpha$ :h2 that govern the dynamics of the closed (GTP-bound, purple) and open (cGMP-bound, green) state of the *hsGC* catalytic heterodimer. Distances were calculated between the initial state (cyan) and maximum transition state of the aforementioned systems.

applied MD simulations to investigate the transition between the open and closed state of the cyclase domain of *hsGC*. Since simulation of the actual transition from the open to the closed state conformation is particularly challenging, a targeted approach was adopted where three systems (apo, GTP- and cGMP-bound cyclase dimers) were studied concurrently and the fluctuations of the secondary structure elements that govern the opening and closing of the binding pocket were closely observed at the nanosecond scale. The structural dynamics of the apo, GTP- and cGMP-bound cyclase dimer complexes are compared with experimental data from the closely related adenylate cyclase catalytic dimer.

The catalytic dimer of adenylate cyclase is sharing a similar catalytic fold with the cyclase dimer of *hsGC* and most of the active site residues are highly conserved<sup>27</sup>. Among the catalytic residues, D440, D530 and E473 are interacting with GTP molecule<sup>22</sup>. A study by Agulló *et al.*, 2016 reported that E473 located at the sGC cyclase  $\beta$  chain is essential for binding to the guanine moiety of the GTP molecule<sup>29</sup>. A role of glutamate residues in the selectivity and binding of GTP versus ATP molecules has also been proposed<sup>30</sup>.

In our simulated GTP-sGC catalytic dimer, E473 lies at a  $\beta$ -hairpin loop ( $\beta$ 2L1 $\beta$ 3) of the  $\beta$  subunit in the catalytic pocket. Additionally, D440 is present at  $\beta$ 2 while D530 at  $\beta$ 2L1 $\beta$ 3 of the  $\alpha$  chain of the sGC catalytic dimer. These secondary structure elements are part of the catalytic pocket assisting in the entrapment of GTP and cGMP. The dynamics of these secondary structure elements involved in the composition of active site have been shown to be substrate-specific<sup>14</sup>. In the current study, the GTP-bound cyclase domain showed a slightly inward movement of  $\alpha$ :h1 which brings 489 G, 490 F, 491 T close to GTP. Strong hydrogen bond interactions with high occupancy were observed between 489 G, 490 F, 491 T and GTP. During the catalysis of GTP to cGMP, threonine residues have been shown to assist in guiding water to the  $\gamma$ -phosphate<sup>31</sup>. These  $\alpha$ :h1 residues are not only binding the phosphate groups of GTP but also stabilizing the closed state of the GTP- sGC catalytic dimer.

Per-residue decomposition analysis complements the hydrogen bond analysis by highlighting the atomistic level energy contributions of active site residues in the ligand-bound states of the *hsGC* cyclase. In the closed state (GTP-bound) cyclase system, more residues interact with the ligand compared to the cGMP-bound state. Significant energy contributions of  $\alpha$ :V488,  $\alpha$ :G489,  $\alpha$ :F490,  $\alpha$ :R574,  $\beta$ :F543,  $\beta$ :V547,  $\beta$ :N548 and  $\beta$ :R552 were observed in the GTP-bound cyclase system, otherwise absent or opposite in the cGMP-bound system. Differences in the protein-protein interactions between intra-domain residues were also witnessed in the two systems under study. The binding free energy analysis results reinforce the molecular docking results proposed in the current study.

Similar structural transitions have also been proposed in other class III nucleotidyl cyclases due to ATP and GTP binding<sup>19,32</sup>. A catalytically active closed state of the GTP binding pocket is induced by the inward movement of h1 of  $\alpha$  chain of the sGC catalytic dimer<sup>14</sup>. In all class III nucleotidyl cyclases members, GTP binding causes rotation of the  $\alpha$  chain at its center translocating h1 close to h4 of the  $\beta$  chain (C2 domain of AC)<sup>32</sup>. In the h4 of the  $\beta$  chain, a conserved NSR motif is present. The serine residue is known to bind to the pentose sugar moiety of the GTP and cGMP<sup>22</sup>. Movement of  $\alpha$ :h1 and  $\beta$ :h4 drives the protein to an active closed conformation which helps all the important catalytic residues to adopt orientations that support the cyclization event. The aforementioned conformational shift is communicated through the entire sGC catalytic dimer inducing  $\beta$ 6L2 $\beta$ 7 of the  $\beta$  subunit to close the active site pocket by interacting with h2 of the adjacent  $\alpha$  subunit. This  $\beta$ 6L2 $\beta$ 7 movement also aids the stabilization of the GTP within the binding pocket through an interaction of L542 with the nitrogenous base of the GTP molecule. Movements of L1 ( $\beta$ 3L1 $\beta$ 4) and  $\beta$ 6 ( $\beta$ 6L2 $\beta$ 7) in  $\beta$  chain are not only holding GTP in the center of the binding pocket through strong hydrogen bond interactions of E473 and L542 with GTP but also generating significant protein-protein interactions through the substrate-induced structural transitions further contributing to the closed state conformation of the sGC catalytic dimer. An inward shift of L3 of the  $\beta$  subunit towards the GTP binding pocket was also observed. L3 of the  $\beta$  subunit harbors K593 which binds with the phosphate group of GTP with significant hydrogen bond interactions and Lys residues in ATPases and GTPases have been shown to stabilize the pentavalent transition state of the phosphate during hydrolysis<sup>33,34</sup>.



In the cGMP-sGC catalytic dimer, due to the cleavage of  $\gamma$ -phosphate from GTP, the sGC catalytic dimer assumes an open conformation. Loss of a few interdomain interactions resulted in the detachment and outward movement of  $\alpha$ :h1 from  $\beta$ :h4 and  $\alpha$ : $\beta$ 6- $\beta$ 7 from the adjacent h2 of the  $\beta$  chain and subsequent opening of the active pocket. Outward movement of  $\alpha$ :h1 and  $\alpha$ : $\beta$ 6- $\beta$ 7 also resulted in the loss of electrostatic interactions of the nucleotide with E473, R574, G489, F490, T491, K593, N548 and R552.

Moreover, residues T491 and L542 located on  $\alpha$ :h1 and  $\beta$ : $\beta$ 6- $\beta$ 7, respectively, establish robust electrostatic interactions with the GTP substrate, possibly facilitating the moderate movement of these regions inward while the absence of interactions in the cGMP-bound system results in the extrication of these regions and the opening of the binding cavity of the cyclase heterodimer. The understanding of the mechanistic basis of the substrate-specific opening and closing of the binding pocket in a more detailed manner was further aided by principal component analysis.

To conclude, we presented the substrate-specific and catalytically restrictive structural dynamics observed in the GTP- and cGMP-bound *hsGC* catalytic heterodimer. All observed conformational changes were validated with prior experimental findings and novel insights have been revealed from a comparative substrate-specific open and closed pocket conformation of the *hsGC* catalytic heterodimer. The findings of this study will be helpful in understanding the molecular mechanism underlying the GTP- and cGMP-induced structural changes and an aid future therapeutic intervention. Although all proposed observations are computationally simulated, they provide a basis for prioritization of future experimental approaches to establish a structural link between the closed and open state of the active pocket upon GTP and cGMP binding.

## Methodology

**Active site prediction of the *hsGC* cyclase domain.** As the active heterodimer of the *hsGC* cyclase domain has not been structurally resolved yet, we applied a knowledge-based computational approach to identify the binding pocket for GTP and cGMP docking. For this purpose, the crystal structure of the heterodimeric *hsGC* cyclase domain in the apo-form solved at 2.08 Å was retrieved from the Protein Data Bank (PDB ID: 3UVJ)<sup>14</sup> Allerston *et al.*, 2013. demonstrated that the catalytic site of the *Rattus norvegicus* sAC catalytic dimer and the *hsGC* cyclase domain share the same structural fold and their catalytic residues are highly conserved<sup>18</sup>. Sequence alignment of the *Rattus norvegicus* sAC catalytic domain and the *hsGC* cyclase domain was performed with the CLC Genomics workbench<sup>35</sup> to assess the highly similar/identical sequence regions. The catalytic dimer of the ATP-bound *Rattus norvegicus* sAC was superposed to the *hsGC* cyclase dimer (PDB ID: 3UVJ) in the Schrödinger software (*Schrödinger Release 2018-3*) to locate the binding pocket of the *hsGC* cyclase domain<sup>36</sup>. The active site residues of the dimeric interface were precisely superimposed. For docking purposes, the catalytic site of the *hsGC* cyclase domain was used as a receptor site for molecular docking.

**Protein preparation.** Before docking, the retrieved crystal structure of the heterodimeric *hsGC* cyclase domain in the apo-form from the Protein Data Bank (PDB ID: 3UVJ) was prepared in Schrödinger's (*Schrödinger Release 2018-3*) multi-step Protein Preparation Wizard<sup>36</sup>. Protein preparation involved addition of all missing hydrogen atoms and removal of water molecules and any heteroatoms. In addition to this, bond orders were assigned and protonation states were then adjusted. To minimize steric hindrance between the atoms, restrained energy minimization was performed by applying the Optimized Potential for Liquid Simulations 2005 (OPLS 2005) force field<sup>37</sup>.

**Molecular docking.** The Glide XP (extra precision) module of Maestro<sup>38,39</sup> was selected for the docking of GTP and cGMP to the receptor protein. To validate the docking algorithm, ATP was re-docked to the binding cavity of the *R. norvegicus* sAC catalytic domain. Based on the centroid of the already bound ATP within the active site of *R. norvegicus* sAC catalytic domain, a grid box of specific dimensions was generated.

The GTP and cGMP molecules were prepared using the Schrödinger 3D builder of (*Schrödinger Release 2018-3*) and minimized using the OPLS3e force field<sup>40</sup>. To dock GTP and cGMP within the dimeric active site of *hsGC*, the grid box dimensions (grid size  $\leq 16$  Å) of the *R. norvegicus* catalytic domain used for the ATP docking were applied. Two-dimensional (2D) ligand interaction diagrams were generated with Intermezzo (Ochoa-Montaña B, Blundell TL – unpublished).

**Substrate parametrization.** Accurate force fields are vital for generating the dynamic and conformational response of condensed-phase systems. The force field parameters for GTP were taken from Amber force field database<sup>41</sup> while the parameterization of cGMP was done by applying quantum mechanics calculations through Gaussian using the Hartree Fock method and the 6-31 G\* basis set<sup>42</sup>. The restrained electrostatic potential (RESP) charges of non-standard residues were calculated with the antechamber17 module<sup>43</sup>.

**MD system preparation.** The cyclase- $\alpha_1\beta_1$  (apo), cyclase- $\alpha_1\beta_1$  GTP/cGMP-bound systems were subjected to all-atom MD simulations in explicit solvent employing the GPU version of the PMEMD engine provided with Amber14<sup>44</sup>. The receptor protein was parameterized with the ff14SB force field<sup>45</sup>. The tleap program was applied to add missing hydrogen atoms and counter ions for neutralizing the systems. All three complexes were immersed in an orthogonal box with a TIP3P water model and periodic boundary conditions<sup>46</sup>. The long-range electrostatic interactions were calculated with the Particle Mesh Ewald (PME) method with a vdW cut-off of 10 Å<sup>47</sup>. After the completion of parametrization, the geometry of all prepared systems was minimized (20000 steps) to remove steric hindrances. The complexes were heated from 0 to 300 K at constant volume with a Langevin thermostat (200 ps). Subsequently, productive MD simulations at constant pressure (NPT) were conducted for each complex for, at least, 100 ns. During the entire MD run, the covalent bonds bearing hydrogens were restrained applying the SHAKE algorithm<sup>48</sup>. Trajectory snapshots were acquired every 2 fs.

**Stability analysis.** Trajectories taken from all MD simulations were interpreted with the help CPPTRAJ v17.00 module<sup>49</sup>. The Root Mean Square Deviation (RMSD) was calculated for all backbone atoms with the starting structure as a reference frame. The Root Mean Square Fluctuation (RMSF) of the backbone atoms over the entire simulation was also calculated. The overall compactness of the structures was assessed with the radius of gyration (Rog) calculation. Stability analysis plots were generated using Xmgrace<sup>50</sup>.

**Hydrogen bond occurrence.** Hydrogen bonding plays a critical role in folding and the wreath-like quaternary packing of the hsGC catalytic subdomains. Therefore, changes in the hydrogen bond occupancy pattern possibly contribute to different aggregation propensity and folding of the cyclase  $\alpha\beta$  heterodimer. Hydrogen bonds of the different conformational arrangements of the cyclase  $\alpha\beta$  heterodimer in the presence of GTP and cGMP substrates were analyzed by considering the distance between the heavy atoms and the angle between the acceptor and donor atoms. The occurrence of hydrogen bonds was visualized with the Ligplot<sup>+</sup> tool<sup>51</sup> which depicts the most persistent hydrogen bond interactions over the course of the 50 ns MD simulations.

**MM-PB/GBSA calculations.** The Molecular Mechanics with Poisson Boltzmann/Generalized Born and Surface Area solvation MM-PB/GBSA protocol was used to estimate the nucleotidyl (GTP/cGMP) binding free energy ( $\Delta G_{\text{bind}}$ ) against the cyclase- $\alpha\beta$  heterodimer by considering the interacting molecules solvation energy along with molecular mechanics (MM) energies. The amber14 MMPBSA.py module<sup>52</sup> was implemented for the estimation of MM-PB/GBSA. The program calculates multiple energy terms such as electrostatic energy, van der Waals interaction energy, polar and non-polar solvation free energy. In addition, the binding free energy of each residue was also calculated in the aforesaid energy terms. This approach assists in identifying critical residues and helps in reproducing their relative binding affinities to a given set of ligands. Overall, binding free energy ( $\Delta G_{\text{bind}}$ ) of a ligand-receptor complex can be estimated by assessing the energy of three reactants with these equations<sup>53</sup> Eq. 1

$$\Delta G_{\text{bind}} = \text{avg G complex} - \text{avg G receptor} - \text{avg G ligand} \quad (1)$$

where avg indicates an average over an ensemble of trajectory snapshots, and G the free energy contributions of the complex, the receptor and the ligand. Therefore, the binding energy of the reactant is calculated with Eq. 2.

$$G = \text{MM} + G_{\text{solvent}} - \text{TS solute} \quad (2)$$

where MM term corresponds to the molecular mechanics contribution in vacuum, comprising of the aggregate of internal, van der Waals and electrostatic energies; G solvent corresponds to the solvation free energy delineated in terms of total polar and non-polar solvation free energy; whereas TS solute explores the temperature and solute entropy respectively<sup>54</sup>. In order to investigate the per-residue binding energy of GTP/cGMP bound cyclase- $\alpha\beta$  complexes, snapshots of the last 20 ns were extracted from the trajectories of 100 ns, and the average binding free energies over the collection of conformers was calculated. The GB and PB methods for the estimation of polar solvation energies were applied with default settings. The non-polar solvation free energy was determined by measuring the solvent-accessible surface area<sup>55</sup>.

**Essential dynamics.** Principal component analysis (PCA) was carried out to identify and distinguish the significant structural transition modes in the cyclase- $\alpha\beta$  apo-form, GTP- and cGMP-bound complexes by utilizing MD trajectory data. This method works by constructing a covariance matrix of the backbone structural transitions ignoring the overall rotational and translational movements<sup>56</sup>. The diagonalization of the covariance matrix results in the generation of structurally uncorrelated variables corresponding to certain eigenvectors (eigenvalues). These eigenvalues are arranged in decreasing order where the first principal component (PC-1) is associated with the largest structural transitions of the proteins. PCs of all three systems were obtained with the CPPTRAJ module of Amber14. In all three systems, the first three PCs contributed the majority of structural transitions over the entire simulation. The Visual Molecular Dynamics (VMD) tool was employed to comprehensively investigate the uncorrelated system transitions and the Normal Mode Wizard (NMWiz) was used to create porcupine plots<sup>20</sup>. Graphical representations were constructed with PyMol and VMD<sup>57</sup>.

Received: 4 October 2019; Accepted: 4 May 2020;

Published online: 11 June 2020

## References

- Khalid, R. R., Maryam, A., Fadoulglou, V. E., Siddiqi, A. R. & Zhang, Y. Cryo-EM density map fitting driven in-silico structure of human soluble guanylate cyclase (hsGC) reveals functional aspects of inter-domain cross talk upon NO binding. *Journal of Molecular Graphics and Modelling* **90**, 109–119 (2019).
- Loscalzo, J. & Welch, G. Nitric oxide and its role in the cardiovascular system. *Progress in cardiovascular diseases* **38**, 87–104 (1995).
- Warner, T. D., Mitchell, J. A., Sheng, H. & Murad, F. In *Advances in pharmacology* Vol. 26 171–194 (Elsevier, 1994).
- Maryam, A. *et al.* The Molecular Organization of Human cGMP Specific Phosphodiesterase 6 (PDE6): Structural Implications of Somatic Mutations in Cancer and Retinitis Pigmentosa. *Computational and structural biotechnology journal* **17**, 378–389 (2019).
- Friebe, A., Mergia, E., Dangel, O., Lange, A. & Koesling, D. Fatal gastrointestinal obstruction and hypertension in mice lacking nitric oxide-sensitive guanylyl cyclase. *Proceedings of the National Academy of Sciences* **104**, 7699–7704 (2007).
- Koesling, D., Russwurm, M., Mergia, E., Mullershausen, F. & Friebe, A. Nitric oxide-sensitive guanylyl cyclase: structure and regulation. *Neurochemistry international* **45**, 813–819 (2004).
- Pan, J. *et al.* Probing the Molecular Mechanism of Human Soluble Guanylate Cyclase Activation by NO *in vitro* and *in vivo*. *Scientific reports* **7**, 43112 (2017).
- Ma, X., Sayed, N., Baskaran, P. & Beuve, A. & Van Den Akker, F. PAS-mediated dimerization of soluble guanylyl cyclase revealed by signal transduction histidine kinase domain crystal structure. *Journal of Biological Chemistry* **283**, 1167–1178 (2008).

9. Ma, X., Sayed, N., Beuve, A. & Van Den Akker, F. NO and CO differentially activate soluble guanylyl cyclase via a heme pivot-bend mechanism. *The EMBO Journal* **26**, 578–588 (2007).
10. Pellicena, P., Karow, D. S., Boon, E. M., Marletta, M. A. & Kuriyan, J. Crystal structure of an oxygen-binding heme domain related to soluble guanylate cyclases. *Proceedings of the National Academy of Sciences* **101**, 12854–12859 (2004).
11. Ma, X., Beuve, A. & Van den Akker, F. Crystal structure of the signaling helix coiled-coil domain of the  $\beta 1$  subunit of the soluble guanylyl cyclase. *BMC structural biology* **10**, 2 (2010).
12. Derbyshire, E. R. & Marletta, M. A. Structure and regulation of soluble guanylate cyclase. *Annual review of biochemistry* **81**, 533–559 (2012).
13. Khalid, R. R., Siddiqi, A. R., Mylonas, E., Maryam, A. & Kokkinidis, M. Dynamic Characterization of the Human Heme Nitric Oxide/Oxygen (HNOX) Domain under the Influence of Diatomic Gaseous Ligands. *International journal of molecular sciences* **20**, 698 (2019).
14. Allerston, C. K., von Delft, F. & Gileadi, O. Crystal structures of the catalytic domain of human soluble guanylate cyclase. *PLoS one* **8**, e57644 (2013).
15. Karow, D. S. *et al.* Spectroscopic characterization of the soluble guanylate cyclase-like heme domains from *Vibrio cholerae* and *Thermoanaerobacter tengcongensis*. *Biochemistry* **43**, 10203–10211 (2004).
16. Beste, K. Y., Burhenne, H., Kaever, V., Stasch, J.-P. & Seifert, R. Nucleotidyl cyclase activity of soluble guanylyl cyclase  $\alpha 1\beta 1$ . *Biochemistry* **51**, 194–204 (2011).
17. Underbakke, E. S., Iavarone, A. T. & Marletta, M. A. Higher-order interactions bridge the nitric oxide receptor and catalytic domains of soluble guanylate cyclase. *Proceedings of the National Academy of Sciences* **110**, 6777–6782 (2013).
18. Tesmer, J. J. *et al.* Two-metal-ion catalysis in adenylyl cyclase. *Science* **285**, 756–760 (1999).
19. Kleinboelting, S., van den Heuvel, J. & Steegborn, C. Structural analysis of human soluble adenylyl cyclase and crystal structures of its nucleotide complexes—implications for cyclase catalysis and evolution. *The FEBS journal* **281**, 4151–4164 (2014).
20. Bakan, A., Meireles, L. M. & Bahar, I. ProDy: protein dynamics inferred from theory and experiments. *Bioinformatics* **27**, 1575–1577 (2011).
21. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *Journal of molecular graphics* **14**, 33–38 (1996).
22. Hahn, D. K., Tusell, J. R., Sprang, S. R. & Chu, X. Catalytic mechanism of mammalian adenylyl cyclase: a computational investigation. *Biochemistry* **54**, 6252–6262 (2015).
23. Friebe, A. & Koesling, D. Regulation of nitric oxide-sensitive guanylyl cyclase. *Circulation research* **93**, 96–105 (2003).
24. Campbell, M. G. *et al.* reveals the higher-order domain architecture of soluble guanylate cyclase. *Proceedings of the National Academy of Sciences* **111**, 2960–2965 (2014).
25. Winger, J. A. & Marletta, M. A. Expression and characterization of the catalytic domains of soluble guanylate cyclase: interaction with the heme domain. *Biochemistry* **44**, 4083–4090 (2005).
26. Winger, J. A., Derbyshire, E. R. & Marletta, M. A. Dissociation of nitric oxide from soluble guanylate cyclase and heme-nitric oxide/oxygen binding domain constructs. *Journal of Biological Chemistry* **282**, 897–907 (2007).
27. Vercellino, I. *et al.* Role of the nucleotidyl cyclase helical domain in catalytically active dimer formation. *Proceedings of the National Academy of Sciences* **114**, E9821–E9828 (2017).
28. Denninger, J. W. & Marletta, M. A. Guanylate cyclase and the NO/cGMP signaling pathway. *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **1411**, 334–350 (1999).
29. Agulló, L., Buch, I., Gutiérrez-de-Terán, H., García-Dorado, D. & Villà-Freixa, J. Computational exploration of the binding mode of heme-dependent stimulators into the active catalytic domain of soluble guanylate cyclase. *Proteins: Structure, Function, and Bioinformatics* **84**, 1534–1548 (2016).
30. Huang, G. Y. *et al.* Structural basis for cyclic-nucleotide selectivity and cGMP-selective activation of PKG I. *Structure* **22**, 116–124 (2014).
31. Kityk, R., Kopp, J., Sinning, I. & Mayer, M. P. Structure and dynamics of the ATP-bound open conformation of Hsp70 chaperones. *Molecular cell* **48**, 863–874 (2012).
32. Tesmer, J. J., Sunahara, R. K., Gilman, A. G. & Sprang, S. R. Crystal structure of the catalytic domains of adenylyl cyclase in a complex with G $\alpha$ -GTP $\gamma$ S. *Science* **278**, 1907–1916 (1997).
33. Flaherty, K. M., Wilbanks, S. M., DeLuca-Flaherty, C. & McKay, D. B. Structural basis of the 70-kilodalton heat shock cognate protein ATP hydrolytic activity. II. Structure of the active site with ADP or ATP bound to wild type and mutant ATPase fragment. *Journal of Biological Chemistry* **269**, 12899–12907 (1994).
34. O'Brien, M. C., Flaherty, K. M. & McKay, D. B. Lysine 71 of the chaperone protein Hsc70 is essential for ATP hydrolysis. *Journal of Biological Chemistry* **271**, 15874–15878 (1996).
35. Sequencing, H. (Workbench, 2011).
36. Sastry, G. M., Adzhigirey, M., Day, T., Annabhimoju, R. & Sherman, W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *Journal of computer-aided molecular design* **27**, 221–234 (2013).
37. Jorgensen, W. L., Maxwell, D. S. & Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society* **118**, 11225–11236 (1996).
38. Friesner, R. A. *et al.* Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein–ligand complexes. *Journal of medicinal chemistry* **49**, 6177–6196 (2006).
39. Friesner, R. A. *et al.* Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *Journal of medicinal chemistry* **47**, 1739–1749 (2004).
40. Roos, K. *et al.* OPLS3e: Extending force field coverage for drug-like small molecules. *Journal of chemical theory and computation* **15**, 1863–1874 (2019).
41. Hornak, V. *et al.* Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Structure, Function, and Bioinformatics* **65**, 712–725 (2006).
42. Martinez, J. P. The Hartree-Fock method: from self-consistency to correct symmetry. *Annalen der Physik* **529**, 1600328 (2017).
43. W Sousa da Silva, A. & Vranken, W. F. ACPYPE-AnteChamber PYthon Parser interface. *BMC Research Notes* **5** (2012).
44. Cerutti, D. *et al.* (University of California, San Francisco, 2016).
45. Maier, J. A. *et al.* ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *Journal of chemical theory and computation* **11**, 3696–3713 (2015).
46. Mark, P. & Nilsson, L. Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *The Journal of Physical Chemistry A* **105**, 9954–9960 (2001).
47. 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**, R55–R60 (1999).
48. Gonnet, P. P-SHAKE: a quadratically convergent SHAKE in O( $n^2$ ). *Journal of Computational Physics* **220**, 740–750 (2007).
49. Roe, D. R. & Cheatham, T. E. III. PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *Journal of chemical theory and computation* **9**, 3084–3095 (2013).
50. Turner, P. XMGRACE, Version 5.1. 19. Center for Coastal and Land-Margin Research, Oregon Graduate Institute of Science and Technology, Beaverton, OR (2005).
51. Laskowski, R. A. & Swindells, M. B. (ACS Publications, 2011).



52. Miller, B. R. III *et al.* MMPBSA. py: an efficient program for end-state free energy calculations. *Journal of chemical theory and computation* **8**, 3314–3321 (2012).
53. Srinivasan, J., Cheatham, T. E., Cieplak, P., Kollman, P. A. & Case, D. A. Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate– DNA helices. *Journal of the American Chemical Society* **120**, 9401–9409 (1998).
54. Genheden, S. & Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert opinion on drug discovery* **10**, 449–461 (2015).
55. Massova, I. & Kollman, P. A. Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding. *Perspectives in drug discovery and design* **18**, 113–135 (2000).
56. Van Aalten, D., Findlay, J., Amadei, A. & Berendsen, H. Essential dynamics of the cellular retinol-binding protein evidence for ligand-induced conformational changes. *Protein Engineering, Design and Selection* **8**, 1129–1135 (1995).
57. DeLano, W. L. The PyMOL molecular graphics system. <http://www.pymol.org> (2002).

## Acknowledgements

Support by the High Performance Computing facility of the University of Crete Data Center and Sezerman Lab are acknowledged. This work was co-financed by Greece and the European Union in the framework of the project “NATIONAL RESEARCH INFRASTRUCTURES ON INTEGRATED STRUCTURAL BIOLOGY, DRUG SCREENING EFFORTS AND DRUG TARGET FUNCTIONAL CHARACTERIZATION–INSPIRED” from the EPAnEK 2014–2020 action, operational programme “Competitiveness-Entrepreneurship-Innovation”. The authors would also like to thank the National Research Program for Universities (NRPU-4050) by Higher Education Commission (HEC) Pakistan. R.R.K., and A.M. thanks 2216 Research Fellowship Program for International Researchers by TUBITAK for their funding of the computational work presented in the current study.

## Author contributions

M.K. and A.R.S. conceived and designed the experiments; R.R.K. performed the experiments and wrote the paper; R.R.K., A.M. and E.M. analyzed the data. M.K., E.M., O.U.S. and A.R.S. revised and improved the paper.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-66310-4>.

**Correspondence** and requests for materials should be addressed to A.R.S. or M.K.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020