Modeling and Simulation of Soluble Guanylyl Cyclase

Yu Wang
Department of Computer Science
New Jersey Institute of Technology
University Heights, Newark, NJ 07102

Kentaro Sugino
Department of Computer Science
New Jersey Institute of Technology
University Heights, Newark, NJ 07102

Marc Q. Ma
Department of Computer Science, and
Center of Applied Mathematics and Statistics
New Jersey Institute of Technology
University Heights, Newark, NJ 07102

Annie V. Beuve
Department of Pharmacology and Physiology
New Jersey Medical School
University of Medicine and Dentistry of New Jersey
185 South Orange Avenue, Newark, NJ 07103

CAMS Report 0405-34, Spring 2005
Center for Applied Mathematics and Statistics

Modeling and Simulation of Soluble Guanylyl Cyclase

Yu Wang[1], Kentaro Sugino[1], Marc Q. Ma[1,2,*], Annie V. Beuve[3]
[1]Department of Computer Science, [2]Center for Applied Mathematics and Statistics, New Jersey Institute of Technology, Newark, NJ 07102; [3]Department of Pharmacology and Physiology, New Jersey Medical School, University of Medicine and Dentistry of New Jersey, 185 South Orange Avenue, Newark, New Jersey 07103; [*]Corresponding author
yw6@njit.edu, ks23@njit.edu, qma@njit.edu, Annie.Beuve@umdnj.edu

Abstract

Soluble guanylyl cyclase (sGC) is a heterodimeric enzyme that catalyzes the formation of cGMP from GTP. The mechanisms regulating the catalytic activity of sGC still remain unclear despite extensive experimental studies. We detail the steps used in homology modeling for constructing the 3D structure of sGC's catalytic core region. Many homology-based models are obtained, and their quality is evaluated using various programs including PROCHECK. Based on our homology models, we perform classical molecular dynamics (MD) simulations to study the conformational changes of the protein complex during the allosteric regulation, and a tentative mechanism is established, which can be used to guide the screening and design of new vasodilation drugs.

Keywords: molecular modeling, protein structure prediction and evaluation, molecular dynamics simulations, soluble guanylyl cyclase, YC-1, allosteric activation.

1. Introduction

Soluble guanylyl cyclase (sGC) is an allosterically regulated enzyme, which contains an α and a β subunit to form a heterodimeric structure [1]. Upon activation, the cyclization of the substrate guanosine 5'-triphosphate (GTP) to guanosine 3',5'-cyclic monophosphate (cGMP) takes place in sGC's catalytic core. cGMP is a second messenger that regulates many fundamental biological processes including vasodilatation. Despite many experimental [1-5] and limited theoretical studies [6] on sGC, the mechanism regulating the catalytic activity of sGC still remain unclear. In addition, the experimental (by X-ray crystallography or NMR) structure of sGC currently is not available.

We show methods for the construction and validation of 3D structure of the sGC's catalytic core region using homology-based modeling (HM). We perform all–atom, classical molecular dynamics (MD) simulations on one HM model in complex with a synthetic compound, YC-1 [7], an allosteric activator of sGC [1, 8], to study the conformational changes of the protein complex during the allosteric regulation. The dynamics of the protein, GTP and YC-1 is observed, and a tentative mechanism is established, which could be used to guide the screening and design of new vasodilation drugs.

2. Homology modeling

In theory, protein structure can be obtained by the following methods: (1) Experiments, *e.g.* X-ray crystallography or NMR spectroscopy. (2) *ab initio* prediction, which is a purely theoretical method based on sequence alone, or (3) Homology modeling, which relies on a database of known protein structures [9, 10]. The first two methods cannot be applied to this study due to the characteristics of sGC and issues such as the size of the protein and prohibitive computation time. Homology modeling is the only method of choice.

Ever since the idea of constructing protein structure by homology was first published by Browne and co-workers about 40 years ago [11], there have been lots of algorithms and software packages developed and applied to different stages in homology modeling. Generally, the process contains four sequential steps [12, 13]:

(1) Template search and selection
(2) Target-template alignment
(3) Model building
(4) Model evaluation and refinement.

Figure 1 gives a more detailed flow-chart of this process.

According to several studies on assessment of homology modeling approaches in CASP (Critical Assessment of protein Structure Prediction), EVA (Evaluation of Automatic protein structure prediction) and LiveBench-8, the most important factors in the homology modeling process are the choice of the best template structure(s) and the correctness of the alignment

[14-17]. Using the same template and the same alignment, the accuracy of model predictions by the best few modeling programs are very similar when used optimally [18, 19]. Thus, our modeling procedure will focus on selecting the sequence alignment methods and the model evaluation indicators.

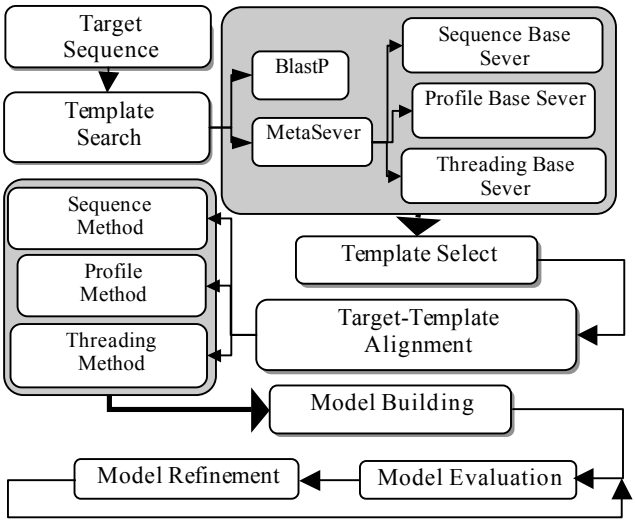

Figure 1. Homology modeling procedure.

In 1997, a homology model of Bovine sGC was created and deposited in the RCSB protein data bank (PDB), 1AWN. Due to the rapid development of the HM prediction methods, now we can make much better predictions than several years ago. We examine this HM model with 3 evaluation programs (Section 2.4). Our evaluation of 1AWN shows that this model can still be improved (in Ramachandran plot, only 71.4% residues are in most favored regions; in Morris *et al* classification, the Chi-1 standard deviation is 7.1 and the H-bond energy standard deviation is 0.80; the G-factor for overall structure is 0.27; ProQ gives LG-score 3.201 and MaxSub 0.23), which motivates us to perform our own predictions using the state-of-the-art technologies.

2.1 Template search and selection.

At the first stage, we submit the amino acid sequences of rat sGC's catalytic core region, α_1V480-L625 (Segment 1), β_1V420-L485 (Segment 2) and β_1H492-E576 (Segment 3), to the NCBI-BlastP sequence analysis sever [20] and the BIOINFO.PL-MetaSever [21]. The MetaSever in BioinfoBank institute is a consensus sever which first takes the amino acid sequence submitted by users, then distributes to all of the participating severs with various fold recognition and local structure prediction methods. The results returned from participating severs are gathered as a list of ranked PDB entries by their scoring function, providing researchers great convenience. The participating severs cover from sequence-based search methods to the profile- and threading-based methods which can recognize common folds even in the absence of any statistically significant sequence similarity. The list of participating severs can be found at http://bioinfo.pl/Meta/servers.html.

From the list of possible homologous proteins and after considering the resolution of each experimental 3D structure (data from RCSB-PDB [22]), their family relative with the target and their biochemical functions and binding ligands (data from Pfam [23], MSD [24]), and the chemostructural restrictions of target protein in prior studies [1, 3, 25, 26], we choose the crystal structure of the Vc1 (chain A) and Iic2 (chain B) of the complex of Gs-alpha with the catalytic domains of mammalian adenylyl cyclase (AC), RCSB-PDB entry: 1AZS, as the modeling template.

The ideal PDB entries and all factors related to template selection for the α_1 subunit of sGC are listed in Table 1 (partial listing). The functional and structural similarities between the AC, which catalyzes the conversion of ATP to cAMP upon allosteric activation by FSK, and sGC are shown in Figure 2 according to previous studies [1].

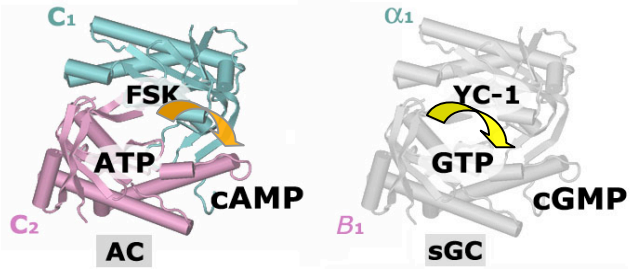

Figure 2. The functional and structural similarity in the catalytic core region of AC and sGC. The cartoon graphics were generated using VMD [27].

The structural features of chain A and chain B of 1AZS are shown in Figure 3 and Figure 4, which also display some sGC patterns such as the active site for GTP. For example, in chain A, there is a guanylate_cyclase_1 pattern, PROSITE pattern #: PS00452, found from GLY495 to GLU518, marked with pink shadow in Figure 3. The location of this GC pattern could be the catalytic domain of α1 subunit in sGC. In chain B, there is a guanylate_cyclase_1 pattern found from GLY1008 to ASP1031, marked with pink shadow in Figure 4. The location of this pattern could be the GTP binding site in Segment 3 of the sGC catalytic center.

2.2 Target-template alignment

Alignment methods can be categorized into four classes including:

(1) Pairwise comparison, *e.g.* BLAST and FASTA

(2) Sequence-profile, *e.g.* PSI-BLAST and SAM-T98

(3) Profile-profile, *e.g.* CLUSTALW and COMPASS
(4) Threading or 3D template structure information methods, *e.g.* THREADER and 3D-PSSM.

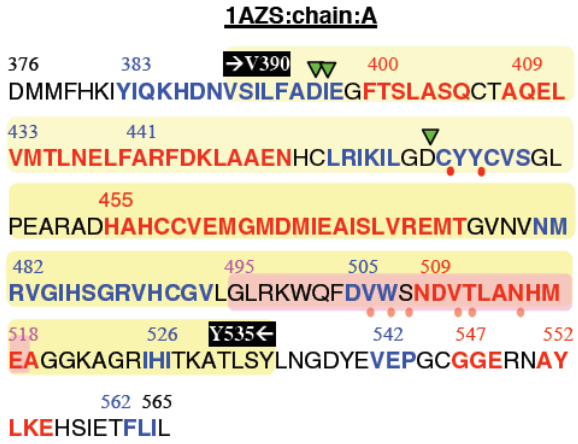

Figure 3. The structural features of chain A (190 amino acids) of 1AZS. The yellow shadow area is the alignment region with our α1 subunit. The blue texts represent the strands. The red texts represent the helices. Green triangles point to the active site of MGB, The red dots mean the residue interaction with ligand. The data is collected from EBI-PDBsum [28].

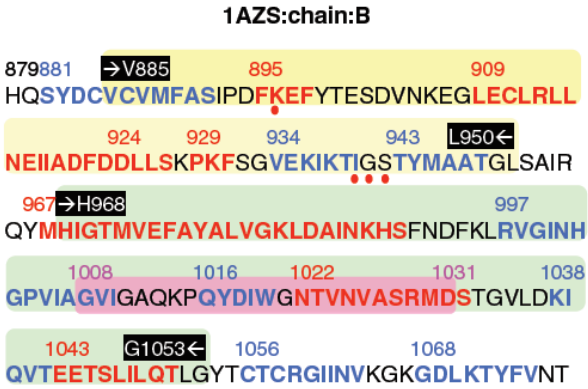

Figure 4. The structural features of chain B (189 amino acids) of 1AZS. The yellow area is the alignment region with Segment 2. The green area is the alignment region with Segment 3. The blue texts represent the strands. The red texts represent the helices. The red dots mean the residue interaction with ligand. The data is collected from EBI-PDBsum [28].

The pairwise comparison methods are best for the situation when target-template sequence identity is higher then 30%. When the sequence identity drops to the "twilight zone", which mean below 25%, the sequence-profile and profile-profile methods are usually capable of increasing the sensitivity and reliability of aligning those distantly homologous proteins. However, the profile-based methods must iteratively generate profiles, which is more computationally costly than the sequence-based methods. The threading or 3D template structure information methods are useful when the target sequence is only related to a very few sequences and thus the alignment can not benefit from the increased sensitivity of the profile-based methods.

In this study, the sequence identity between our target and template are around 30 to 40%, (34%, 30% and 47%, respectively). There is no single alignment method can outperform others when evaluated various group of targets [29]. Moreover, no tools can provide completely confident assessment for the alignment result. Instead of just using a single alignment tool, we collect alignment results from the methods which have been widely used and have higher reliability in different categories, *e.g.* FASTA3 in pairwise, PDB-Blast in profile and 3D-PSSM in threading category. Not all of the programs and data used or generated in this study are listed here. A complete list of sequence alignment, model building, and model evaluation programs used in this study is available at http://web.njit.edu/~yw6/HM_MDS_sGC_SupplementalData.htm.)

2.3 Model building

There are three kinds of model building methods can be used in HM including: (1) Rigid-body assembly. (2) Segment matching or coordinate construction. (3) Satisfaction of spatial restraints. We choose Modeller [30], a spatial restraints model prediction package, which gives excellent results such as those best ranked predictions in CASP5 in 2002 [31]. Besides the accuracy, Modeller has high degree of flexibility and automation. These features are important for case sensibility and model refinement.

There are two steps in our model building process. First, we eliminate those questionable alignment results, *e.g.* with gap in secondary structure elements. Then, we use Modeller to generate models containing backbone, all loops, the N- and C-termini and side-chains from all of the quality alignments that we have obtained at the previous stage. For each alignment, we generate three different candidate models.

2.4 Model evaluation

We submit our three sub-sequences to BlastP and MetaSever, and construct models using Modeller. We set up the "quality indicators" and evaluate these models. There are two types of methods for validation, one focusing on sequence-structure compatibility, which is used by Verify3D, RPOVEN and PEOSAIII, the other examining backbone and side-chain stereochemistry, which is used by PROCHECK, WHATIF and ProQ.

We take two steps to complete the evaluation process. First, when a model is generated by Modeller, an

"objective function" value, which is a combination of the internal restraints, is also calculated to reflect the quality of the predict model. However, this value only reflects the currently selected restraints, models can only be compared if the restrain set is the same. In our case, we use this objective function value as 1st quality indicator using which two thirds of the models are eliminated. For the rest of the models, we choose total nine factors from PROCHECK [32] and PROVEN [33], to cover the criteria from sequence-structure compatibility to side-chain stereochemistry.

Seven factors are chosen from the result of PROCHECK:

(1) Residues in most favored regions
(2) Residues in disallowed regions
(3) Bad contacts numbers
(4) Chi-1 standard deviation (Mean: 18.2, St. dev: 6.2)
(5) H-bond energy standard deviation (Mean: 0.87, St. dev: 0.24)
(6) Average G-factors of main-chain covalent forces (include bond lengths and bond angles)
(7) Average G-factors of dihedral angles (include phi-psi, chi1-chi2 distribution, chi1 only, chi3 & chi4, and omega).

Table 1. Possible templates list for sGC-alpha1 subunit (partial)

PDB Entry	Resolution	Sequence Score/E-value/Identities	Protein Families (Pfam ID)	Biochemical Function	Ligands	Metal ions
1AB8	2.60 Å	57.4/1e-09/26% (161 a.a.)	*PF00211 *PF06327	adenylate cyclase activity (A, B) guanylate cyclase activity (A, B)	FOK*2	
1AZS	2.30 Å	97.4/9e-22/34% (143 a.a.)	*PF00211 *PF00503 *PF06327	adenylate cyclase activity (A, B) guanylate cyclase activity (A, B) signal transducer activity (C) GTP binding (C)	GSP, FKP	MG
1CS4	2.50 Å	97.4/9e-22/34% (143 a.a.)	*PF00211 *PF00503 *PF06327	adenylate cyclase activity (A, B) guanylate cyclase activity (A, B) signal transducer activity (C) GTP binding (C)	101, GSP, FOK MES*2, POP	MG*2, CL
1AWN*	Homology model of bovine soluble guanylyl cyclase			intracellular signaling cascade (A) guanylate cyclase activity (A)	GTP	MG

The data of Sequence Score/E-value/Identities is queried from NCBI-BlastP.
The data of Protein Families and Biochemical Function is queried from EMBL-EBI-PDBsum and MSD.
*PF00211: Adenylate and Guanylate cyclase catalytic domain.
*PF00503: G-protein alpha subunit.
*PF06327: Domain of unknown function.
Abbreviations: 101, 2'-deoxy-adenosine 3'-monophosphate; CL, chloride ion; FKP, methylpiperazinoforskolin; FOK, forskolin; GSP, 5'-guanosine-diphosphate-monothiophosphate; GTP, guanosine- 5'- triphosphate; MES, 2-(n-morpholino)-ethanesulfonic acid; MG. magnesiumion; POP, pyrophosphate 2-.
*The homology model, 1AWN, deposited in RCSB-PDB is not used as template in this study since the resolution is not applicable. But, will use to inspect the predicted model's features.

Table 2. Model evaluation table (partial)

Programs	Evaluation Items			Results	
				Blast2	FATSA3
PROCHECK	Ramachandran Plot	residues in most favored regions		91.9 %	94.5 %
		residues in disallowed regions		1.6 %	0.8 %
	Residue properties	Morris et. al. class	Chi-1 st.dev.	7.5	6.5
			H-bond energy st. dev.	0.75	0.69
		Bad contacts		8	5
	G-factors Average	Main-chain covalent forces		-0.35	-0.13
		Dihedral angles		0.05	0.11
PROVEN	Zrms			1.530	1.599
	% Outliers			4.40 %	5.90 %
ProQ	LG-score			3.942	4.749
	MaxSub			0.474	0.533

Another two factors are manipulated by PROVEN: Z-score RMS, percentage of outliers. Additional factors including LG-score and MaxSub from ProQ [19], will be used as quality indicators only when the scores are tied. Model evaluation is shown in Table 2 (partial listing). After obtaining the evaluation table, we can pick the final

winner. The best models that we choose may still have residues that are marked abnormal in the model evaluation reports, which may be improved during energy minimization, heating and equilibration processes using molecular dynamics programs.

We use the "loopmodel" class provided in Modeller to do the loop optimization. Side-chains with bad contacts are rebuilt using SCWRL3 which uses graph theory [34]. The process of energy minimization with restraints also helps in model refinement, which is used in our procedure.

2.5 Homology modeling results

After picking the best models base on the evaluation scores which covers the criterions from sequence-structure compatibility to side-chain stereochemistry, we still need to check its 3D-structure to see if it has the structural features for the biological functions that have been addressed in prior studies. Since there is no experimental structure available for comparison, we compare our HM model with the HM model of Lamothe and co-workers [1] and the HM model deposited in RCSB-PDB (1AWN).

We superimpose these models and calculate the RMS scores, as shown in Table 3. Our HM model performs quite well. The cartoon representation of these models is shown in Figure 5. Another graphic presentation of the structural similarity is shown in Figure 6.

Other important structural features in the catalytic core region of sGC are the pockets for GTP and YC-1 binding. The predict model should contain two pockets for the binding activity taking place. Examination of our model and the possible binding pockets in previous study [1], we can see there exists a pocket in each binding site as shown in Figure 7.

Table 3. Superposition of HM models.

RMS (Å)	Alpha1	Beta1	Beta2
Ours vs. Lamothe's [1]	1.85	1.23	1.37
Ours vs. 1AWN	1.59	1.47	1.41
Lamothe's [1] vs. 1AWN	1.70	1.56	1.48

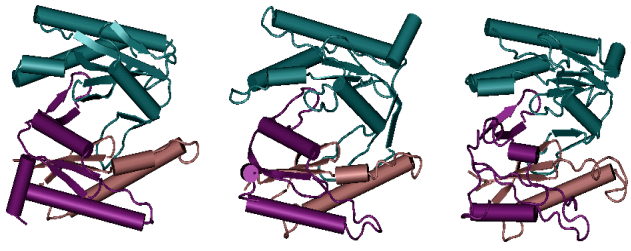

Figure 5. Cartoon representation of three models. From left to right, our predict model, 1AWN, and Lamothe's model. Alpha subunit is colored in cyan. Beta1 subunit is colored in pink and Beta2 subunit is colored with purple. The graphics is generated using VMD.

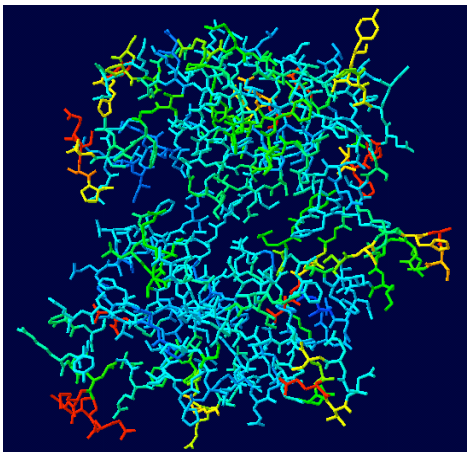

Figure 6. The fit image of our HM model with Lamothe's HM model. Each amino-acid of the active protein is colored accordingly to its RMS backbone deviation from the corresponding amino-acid of the reference protein (Lamothe's model). Dark blue means good superposition whereas red means bad superposition. The image is generated by using DeepView/Swiss-pdb Viewer [35].

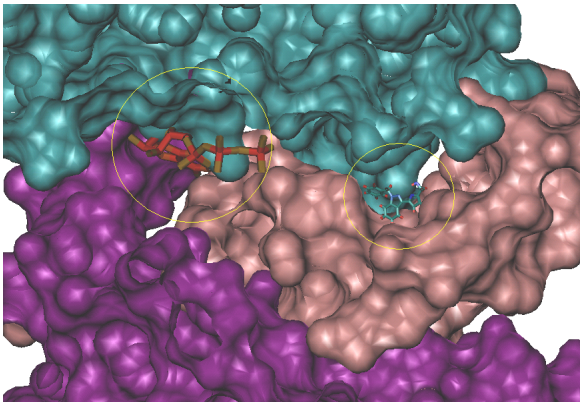

Figure 7. Binding pockets for GTP and YC-1. These two pockets are formed by alpha1, beta1 and beta2 subunits. The GTP is located at the left circle. The YC-1 is binding at the right circle.

From previous assessment and these observations, we have strong confidence with the correctness of our predict model.

3. Molecular dynamic simulation

Molecular Dynamic (MD) is a venerable computer simulation technique in bio-molecular modeling that interfaces mathematics, biology, chemistry, physics and computer science [36]. MD faithfully models the constituent atoms in bio-molecules that continuously interacting with themselves and the environment.

In classical MD, although it does not simulate those of chemical reaction events such as electron donation /acceptation, bond formation/breaking and so on, it is

widely accepted that will defined MD calculation reflects quite reasonably the state of real world phenomenon and theory, *e.g.* classical Newton's dynamics property of given system, and it would be valuable clue to understand targeted molecular complex's mechanism. Therefore, we choose the MD approach for this catalytic mechanism investigation.

From the many MD program packages, NAMD program suite [37] was chosen for this investigation because it has wide scalability, specific features, such as interactive MD which might be used in further analysis, and it accepts CHARMM [38] force field parameters which is well known and accepted as one of the best parameter sets especially for bio molecular simulation.

To understand the catalytic mechanism of sGC, we construct three model systems for MD simulations. In the first system, there is no YC-1 binding. In the second system, YC-1 initially binds with the "Normal" mode, i.e., the hydroxymethyl group of YC-1 initially facing towards the inside of its binding pocket. In the third system, YC-1 binds with the "Flip" mode, i.e., opposite to the orientation in the "Normal" mode. The modeling and simulation process is as follows:

1. Model prediction and refinement
2. Docking YC-1, GTP, 2 magnesium ions
3. Solvating the system
4. Running MD simulations using a force field which is augmented with new parameters for new structures such as YC-1
5. MD trajectory analysis

Solvation of the participative proteins is needed to mimic the biological environment before any production MD run since proteins function in water environment. The TIP3P water model was employed as explicit solvent in this study. We used 3, 700 TIP3P water molecules as the bulk water box (65 Å×54 Å× 54 Å).

To run MD simulation using NAMD, two parameter files must be defined. One is the topology file, which contain the information of atomic mass, whole residue name, atomic charge, atom name, atom type, bond connection, double bond connection, dihedral angle, and improper. The other parameter file is for force constants, which contains force constants of bond / angle / dihedral / improper energy between each atom types, and actual bond length, angle degree. The parameters in the topology files are obtained by using the semi-empirical freely available quantum chemistry calculation software, MOPAC [39]. For the other parameters file, the CHARMM force field is employed.

With these two parameters files and adding one counteracting ion, sodium (Na^+), to make the whole system balance, we can initiate the simulation protocol.

The solvated sGC (before docking GTP, YC-1 and magnesium ions) undergoes a minimization, heating and equilibration process. In the minimization process (100,000 steps), the water molecules are let move while keeping protein's configuration unchanged. Then both protein and water molecules are let move for further minimization. After this process, the system is heated gradually to room temperature using the standard protocols. After the system reaches room temperature (300K), a 300ps equilibration process is performed to bring the system into an equilibrated state. At this stage, the GTP and two magnesium ions and one counter ions are added to the system. Additional 6000 steps of minimization is performed. Then, we dock the YC-1 and perform additional minimization. Minimization is done using conjugate gradient method of NAMD which is used to eliminate illegal close contacts of atoms. After these steps, one 1-nanosecond MD simulation is performed on each of the three systems in NAMD. Verlet-I/r-RESPA/Impulse [40, 41] MTS integrator with 3fs outer step size for good efficiency and stability in solving the governing equations of motion [42], and PME [43] for fast electrostatics calculation are also used.

Key observations from these MD simulations are summarized as follows: A) The distance between the α Phosphorus (Pα) and the 3' hydroxyl (Oxygen O3) of the ribose ring of GTP hovers at a low value (3.6 Å) with the "Normal" binding mode, whereas this distance rises to much higher values (4.9 Å) with the "Flip" binding mode (after 200 ps simulation) or without YC-1 binding (after 700 ps simulation). B) Regardless of the initial binding modes, YC-1 folds to a hairpin structure after initial binding to the protein rather than staying fairly extended, and the distance between the Oxygen O23 in YC-1 and the alpha Carbon in CYS594 in sGC (a residue deep inside the YC-1 binding pocket) converges to a low value (Figures 8 and 9). The folding rate of YC-1 with initial "normal" mode is one order of magnitude faster.

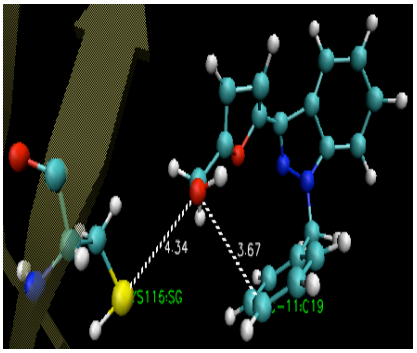

Figure 8. YC-1 folds to form a hairpin structure regardless of initial binding modes. For "Normal" mode, YC-1 forms the hairpin structure within 100ps and stay as a hairpin. For "Flip" mode, it takes much longer for YC-1 to form a stable hairpin structure (after 900ps).

These observations suggest that YC-1 has different capabilities of activating the catalysis in sGC due to different initial binding modes. The "Normal"–mode of

binding allows YC-1 to increase catalysis in sGC, whereas the "Flip"–mode of binding does not. Thus, YC-1 activation of sGC is probably mainly due to the "Normal"– mode of binding. More analysis is needed to fully understand the cascading events — YC-1 folds to hairpin structure and interacts with the surrounding amino acids, which probably causes conformational changes in the catalytic core where GTP binds.

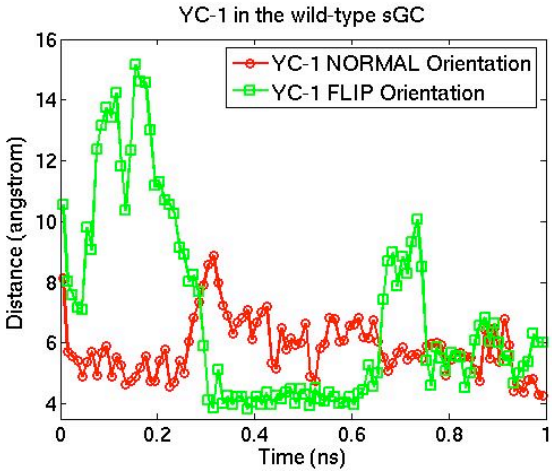

Figure 9. Distance between O23 in YC-1 and the alpha Carbon of CYS594. For "Normal" mode, the Oxygen O23 interacts closely with the CYS594 residue. For "Normal" mode, the Oxygen O23 also interacts closely with the CYS594 residue due to the formation of the hairpin structure.

4. Summary and Future Work

We have successfully constructed homology models of sGC and run molecular dynamic simulations, which has led to the discovery of a tentative mechanism of allosteric regulation sGC. It is possible to gain even deeper and more thorough understanding of this and other complex biological processes, which can be achieved better prediction and more rigorous evaluation of protein structures, more reasonable simulation protocols, and longer and larger simulations using more advanced multiscale numerical schemes.

Current investigation on the mechanisms of allosteric activation in sGC will be further strengthened after more simulations are performed on some experimentally well–studied mutants which will be constructed through *in silico* mutagenesis to study the transient and dynamical behavior of the systems.

Binding affinities indicate how difficult it is to dissociate a ligand from a receptor protein [43]. The binding mode selectivity may be closely related to the difference of binding affinities of YC-1 due to different initial binding modes, which will be studied in the future.

Acknowledgment

This research is funded by NSF (MRI-DMS 0420590), NIH (RO1-GM 067640-01) and New Jersey Institute of Technology (NJIT). Simulations are performed on a SunFire 6900 system at the University of Medicine and Dentistry of New Jersey and the Hydra cluster at NJIT.

References

[1] M. Lamothe, F.-J. Chang, N. Balashova, R. Shirokov, and A. Beuve, "Functional characterization of nitric oxide and YC-1 activation of soluble guanylyl cyclase: structural implication for the YC-1 binding site?" *Biochem.*, vol. 43, pp. 3039-3048, 2004.

[2] P. Condorelli and S. C. George, "In Vivo Control of soluble guanylate cyclase activation by nitric oxide: a kinetic analysis," *Biophysical Journal*, vol. 80, pp. 2110-2119, 2001.

[3] J. W. Denninger and M. A. Marletta, "Guanylate cyclase and the NO/cGMP signaling pathway," *Biochim. Biophys. Acta.*, vol. 1411, pp. 334-350, 1999.

[4] D. Koesling, "Studying the structure and regulation of soluble guanylyl cyclase," *Methods*, vol. 19, pp. 485-493, 1999.

[5] M. Koglin, J. P. Stasch, and S. Behrends, "BAY 41-2272 activates two isoforms of nitric oxide-sensitive guanylyl cyclase," *Biochem. Biophys. Res. Commun.*, vol. 292, pp. 1057-1062, 2002.

[6] M. Q. Ma, K. Sugino, Y. Wang, N. Gehani, and A. V. Beuve, "Modeling and simulation based approaches for investigating allosteric regulation in enzymes," in *Macromolecular Simulation Algorithms*, C. Chipot, R. Elber, A. Laaksonen, B. Leimkuhler, A. Mark, T. Schlick, C. Schuette, and R. D. Skeel, Eds. Berlin, New York: Springer-Verlag, 2005 (in press).

[7] J. Sopkova-de Oliveira Santos, V. Collot, I. Bureau, and S. Rault, "YC-1, an activation inductor of soluble guanylyl cyclase," *Acta Crystallographica Section C*, vol. 56, pp. 1035-1036, 2000.

[8] E. Martin, Y.-C. Lee, and F. Murad, "YC-1 activation of human soluble guanylyl cyclase has both heme-dependent and independent components," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 98, pp. 12938-12942, 2001.

[9] D. Baker and A. Sali, "Protein structure prediction and structural genomics," *Science*, vol. 294, pp. 93-96, 2001.

[10] A. Fiser, M. Feig, C. L. Brooks III, and A. Sali, "Evolution and physics in comparative protein structure modeling," *Acc. Chem. Res.*, vol. 35, pp. 413–421, 2002.

[11] W. J. Browne, A. C. T. North, D. C. Phillips, K. Brew, T. C. Vanaman, and R. C. Hill, "A possible three-dimensional structure of bovine lactalbumin based on that of hen's egg-white lysosyme," *J. Mol. Biol.*, vol. 42, pp. 65-86, 1969.

[12] M. A. Marti-Renom, A. Fiser, R. Sanchez, F. Melo, and A. Sali, "Comparative protein structure modeling of genes and genomes," *Annu. Rev. Biophys. Biomol. Struct.*, vol. 29, pp. 291-325, 2000.

[13] R. Sanchez and A. Sali, "Advances in comparative protein structure modeling," *Curr. Opin. Struct. Biol.*, vol.

7, pp. 206-214, 1997.
[14] M. A. Marti-Renom, A. Fiser, M. S. Madhusudhan, B. John, A. Stuart, N. Eswar, U. Pieper, M.-Y. Shen, and A. Sali, "Modeling protein structure from its sequence," in *Current Protocols in Bioinformatics*, A. D. Baxevanis, D. B. Davison, R. D. M. Page, G. A. Petsko, L. D. Stein, and G. D. Stormo, Eds.: John Wiley & Sons, 2003, pp. Section 5.1 (CD-ROM).
[15] B. Rost and V. A. Eyrich, "EVA: Large-scale analysis of secondary structure prediction," *Proteins Suppl.*, vol. 5, pp. 192-199, 2001.
[16] L. Rychlewski and D. Fischer, "LiveBench-8: The large-scale, continuous assessment of automated protein structure prediction," *Protein Sci.*, vol. 14, pp. 240-245, 2005.
[17] B. Wallner and A. Elofsson, "All are not equal: A benchmark of different homology modeling," *Protein Sci.*, vol. 14, pp. 1315-1327, 2005.
[18] S. E. Brenner, C. Chothia, and T. J. Hubbard, "Assessing sequence comparison methods with reliable structurally identified distant evolutionary relationships," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 95, pp. 6073–6078, 1998.
[19] B. Wallner, H. Fang, and A. Elofsson, "Automatic consensus-based fold recognition using Pcons, ProQ, and Pmodeller," *Proteins: Structure, Function and Genetics*, vol. 53, pp. 534-541, 2003.
[20] S. F. Altschul, T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman, "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs," *Nucleic Acids Res.*, vol. 25, pp. 3389-3402, 1997.
[21] K. Ginalski, A. Elofsson, D. Fischer, and L. Rychlewski, "3D-Jury: a simple approach to improve protein structure predictions," *Bioinformatics*, vol. 19, pp. 1015-1018, 2003.
[22] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne, "The Protein Data Bank," *Nucleic Acids Res.*, vol. 28, pp. 235-242, 2000.
[23] A. Bateman, L. Coin, R. Durbin, R. D. Finn, V. Hollich, S. Griffiths-Jones, A. Khanna, M. Marshall, S. Moxon, E. L. L. Sonnhammer, D. J. Studholme, C. Yeats, and S. R. Eddy, "The Pfam Protein Families Database," *Nucleic Acids Res. Database Issue*, pp. D138-D141, 2004.
[24] A. Golovin, D. Dimitropoulos, T. Oldfield, A. Rachedi, and K. Henrick, "MSDsite: A database search and retrieval system for the analysis and viewing of bound ligands and active sites," *PROTEINS: Structure, Function, and Bioinformatics*, vol. 58, pp. 190-199, 2005.
[25] J. H. Hurley, "The adenylyl and guanylyl cyclase super family," *Curr. Opin. Struct. Biol.*, vol. 8, pp. 770-777, 1998.
[26] J. J. Tesmer, R. K. Sunahara, A. G. Gilman, and S. R. Sprang, "Crystal structure of the catalytic domains of adenylyl cyclase in a complex with $G_{s}\alpha$GTPS," *Science*, vol. 278, pp. 1907-1916, 1997.
[27] W. Humphrey, A. Dalke, and K. Schulten, "VMD - Visual Molecular Dynamics," *J. Molec. Graphics*, vol. 14, pp. 33-38, 1996.
[28] R. A. Laskowski, V. V. Chistyakov, and J. M. Thornton, "PDBsum more: new summaries and analyses of the known 3D structures of proteins and nucleic acids," *Nucleic Acids Res.*, vol. 33, pp. 266-268, 2005.
[29] D. Fischer, "Hybrid fold recognition: Combining sequence derived properties with evolutionary information," in *Pacific Symp. Biocomputing*. Hawaii, 2000, pp. 119-130.
[30] A. Sali, A. Fiser, R. Sanchez, M. A. Marti-Renom, B. Jerkovic, A. Badretdinov, F. Melo, J. Overington, and E. Feyfant, "MODELLER, A protein structure modeling program, release 8v0," 2005.
[31] A. Tramontano and V. Morea, "Assessment of homology-based predictions in CASP5," *Proteins: Structure, Function and Genetics*, vol. 53, pp. 352-368, 2003.
[32] R. A. Laskowski, M. W. MacArthur, D. S. Moss, and J. M. Thornton, "PROCHECK: a program to check the stereochemical quality of protein structures," *J. Appl. Cryst.*, vol. 26, pp. 283-291, 1993.
[33] J. Pontius, J. Richelle, and S. J. Wodak, "Quality assessment of protein 3D structures using standard atomic volumes," *J. Mol. Biol.*, vol. 264, pp. 121-136, 1996.
[34] A. A. Canutescu, A. A. Shelenkov, and R. L. Dunbrack Jr., "A graph theory algorithm for protein side-chain prediction," *Protein Sci.*, vol. 12, pp. 2001-2014, 2003.
[35] N. Guex and M. C. Peitsch, "SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling," *Electrophoresis*, vol. 18, pp. 2714-2723, 1997.
[36] T. Schlick, *Molecular modeling and simulation: An interdisciplinary guide*, vol. 21: Springer-Verlag New York, Inc., 2002.
[37] L. Kalé, R. Skeel, M. Bhandarkar, R. Brunner, A. Gursoy, N. Krawetz, J. Phillips, A. Shinozaki, K. Varadarajan, and K. Schulten, "NAMD2: Greater scalability for parallel molecular dynamics," *Journal of Computational Physics*, vol. 151, pp. 283-312, 1999.
[38] N. Foloppe and A. D. MacKerell Jr., "All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data," *J. Comp. Chem.*, vol. 21, pp. 86-104, 2000.
[39] J. J. P. Stewart, "MOPAC: A semiempirical molecular orbital program," *J. Comput. Aided Mol. Design*, vol. 4, pp. 1-105, 1990.
[40] H. GrubmÜller, H. Heller, A. Windemuth, and K. Schulten, "Generalized Verlet algorithm for efficient molecular dynamics simulations with long-range interactions," *Molecular Simulation*, vol. 6, pp. 121–142, 1991.
[41] M. Tuckerman, B. J. Berne, and G. J. Martyna, "Reversible multiple time scale molecular dynamics," *J. Chem. Phys.*, vol. 97, pp. 1990-2001, 1992.
[42] Q. Ma, J. A. Izaguirre, and R. D. Skeel, "Verlet-I/r-RESPA/Impulse is limited by nonlinear instability," *SIAM J. Sci. Comput.*, vol. 24, pp. 1951-1973, 2003.
[43] T. A. Darden, D. M. York, and L. G. Pedersen, "Particle mesh Ewald: A method for Ewald sums in large systems," *Journal of Chemical Physics*, vol. 98, pp. 10089–10092, 1993.