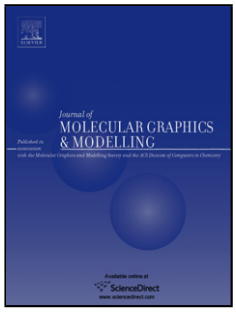


Title: Molecular Dynamics Simulations of Isoleucine-Release Pathway in GAF Domain of N-CodY from *Bacillus Subtilis*

Author: Baoping Ling Min Sun Siwei Bi Zhihong Jing Zhiguo Wang

PII: S1093-3263(13)00120-4
DOI: http://dx.doi.org/doi:10.1016/j.jmgm.2013.07.002
Reference: JMG 6307

To appear in: *Journal of Molecular Graphics and Modelling*

Received date: 7-4-2013
Revised date: 29-6-2013
Accepted date: 1-7-2013

Please cite this article as: B. Ling, M. Sun, S. Bi, Z. Jing, Z. Wang, Molecular Dynamics Simulations of Isoleucine-Release Pathway in GAF Domain of N-CodY from *Bacillus Subtilis*, *Journal of Molecular Graphics and Modelling* (2013), http://dx.doi.org/10.1016/j.jmgm.2013.07.002

Molecular Dynamics Simulations of Isoleucine-Release Pathway in GAF Domain of N-CodY from *Bacillus Subtilis*

Baoping Ling[a*], Min Sun[a], Siwei Bi[a], Zhihong Jing[a], Zhiguo Wang[b]

[a]*Key Laboratory of Phamaceutical Intermediates and Analysis of Natural Medicine, School of Chemistry and Chemical Engineering, Qufu Normal University, Qufu, Shandong, 273165, China*

[b]*Institute of Ageing Research, Hangzhou Normal University, Hangzhou, Zhejiang, 310036, China*

Abstract: The GAF domain located in the N-terminal motifs of CodY (N-CodY) is responsible for increasing the affinity of CodY to its target sites on DNA by its interaction with the branched chain amino acids (BCAAs) involving isoleucine, leucine and valine. The study of the interaction of GAF domain with isoleucine gains much attention in recent years, but the mechanism of isoleucine release still remains unclear. In this paper, a conventional molecular dynamics (MD) and force probe molecular dynamics (FPMD) simulations have been performed with the aim to understand how the isoleucine ligand escapes from the GAF domain of N-CodY from *bacillus subtilis*. The MD results reveal that the ligand release is a gradual process, which is accompanied by the movement of the loop between β3 and β4. During the periods of ligand escaping from the bottom to the top of binding pocket, isoleucine forms hydrogen bonds one after another with series of residues, such as ARG61, THR96, PHE98, VAL100, GLU101 and ASN102, under the mediation of hydrophobic contacts. The FPMD results show that the easiest way to pull ligand out of the cavity is along *x* direction (i.e. the direction is opposite to MET62).

Keywords: GAF domain; N-CodY; *Bacillus subtilis*; molecular dynamics simulations; force probe molecular dynamics

* Corresponding to: Email address: lingbaoping@yahoo.com.cn (B. Ling)

1 Introduction

CodY is a highly conserved protein in the low-G+C Gram-positive bacteria, which was firstly identified as the repressor of the *Bacillus subtilis* dipeptide permease operon [1]. CodY regulates more than 100 *B. subtilis* genes that are typically repressed during rapid exponential growth while induced when the cells encounter nutrient limitation [2], and it plays a critical role in controlling the expression of stationary phase genes and the initiation of sporulation [3]. Most of the CodY-regulated genes are involved in nitrogen and carbon metabolism [4], and they encode extracellar degradative enzymes, transport systems, catabolic enzymes, antibiotic synthesis and so on [2, 5].

The repressor function of CodY is regulated through its interactions with two different effectors, GTP and the branched-chain amino acids (BCAAs), involving leucine and valine, particularly isoleucine, which can independently and additively increase the affinity of CodY to its target sites on DNA [6]. Two key domains of CodY are responsible for the binding of effectors and DNA. One is a GAF domain, locating in the N-terminal motifs of CodY, which interacts with GTP and BCAAs. With the binding stimulations of effectors, a conformational change occurs and triggers DNA binding in the other domain. Simultaneously, a hydrophobic pocket surrounding the bound effector is formed. The other domain is a winged helix-turn-helix (wHTH) motif, locating at the C-terminal region of CodY, which has high affinity for DNA binding and represses the transcription of genes resided in downstream region of the binding site [7].

GAF domains are ubiquitous motifs which are presented in cyclic nucleotide phosphodiesterases, certain adenylyl cyclases and the bacterial transcription factor Fh1A [8]. They are widely distributed in mammalian proteins [9]. The crystal structure in Fig. 1 shows that the GAF domain of the N-terminal CodY spanning residues 1-155 contains a three-layered globular structure [9], i.e., a basal layer of two or more α-helices, a middle layer of a mixed β-pleated sheet formed by four or more strands, and a distal layer of two extended polypeptide segments connecting the strands of the β-sheet [10, 11]. The domain binds a diversity of ligands including linear, cyclic nucleotides, amino acids and porphyrin rings. When the domain binds the effectors, the middle layer of polypeptide segments forms a hydrophobic pocket surrounding the bound ligand and encloses it in a cavity based on β-sheet. Accordingly, the ligand binding and release from the GAF domain are accompanied by a structural rearrangement of distal segment [11].

Levdikov et al. have solved the crystal structures in the absence and presence of ligand [10, 11], which suggested that the core of the protein and its dimerisation interface were almost unchanged. However, a structural change with a displacement of 15 Å and a dramatic twisting were observed in the binding site loops linking β3 and β4 [11]. As a consequence, the ligand-binding cavity no longer exists in the unliganded GAF domain, and the side entrance to the cavity in liganded structure is located on the opposite side of β2-α3 segment in unliganded domain. Moreover, structural analysis also indicated that the high flexibility of the loop linking β3 and β4 other than the segments linking α1-α2 and β4-β5 could play an essential role in ligand binding and release. Thus, the results displayed that accessing to the binding pocket was accompanied by the movement of the loop linking β3 and β4 [11]. However, the detailed accessing/releasing pathways and structural changes accompanying by the ligand uptake and release are still unclear.

In present study, to further identify the binding and releasing pathways of ligand from the binding site of N-terminal GAF domain and the role of the loop linking β3 and β4 in ligand uptake and release process[12], a conventional molecular dynamics (MD) and force probe molecular dynamics (FPMD) simulations were carried out. Understanding the ligand-release pathway and dynamic properties is helpful to gain an insight into the entering mechanism of ligand to the cavity, which is significant for drug delivery to the targets.

2 Methods

2.1 Preparation of structures

The crystal structure of the protein bound ligand in open state (N-CodY_ILE) was taken from the Protein Data Bank (www.rcsb.org) with PDB entry of 2B18 at 1.80 Å resolution [10]. The structure without ligand in closed state (N-CodY_NO) was also downloaded with PDB code 2GX5 at 1.74 Å resolution, whereas Chain C was chosen as the model because four chains were superposed well and only Chain C was intact without the missing residues [11]. In order to identify whether the protein will change its conformation from open state to closed state without ligand, the crystal structure of protein in the absence of ligand in open state (N-CodY_DEL) was obtained from the complex with isoleucine (N-CodY_ILE) by simply deleting ligand.

2.2 MD simulations

The above three crystal structures were solvated in a $6.7\times6.7\times6.7$ nm^3 cubic box with periodic boundary condition by adding about 9,000 simple point charge (SPC) water molecules

[13]. 6 water molecules were replaced randomly by 6 sodium ions to keep charge neutrality of each system. As a consequence, each system contains about 29,000 atoms. First, each system was minimized for 10,000 steepest-descent steps. Second, the solvate and ions were equilibrated for 100 ps in constant volume (NVT) and constant pressure (NPT) ensembles, respectively, while the protein and ligand were restrained with harmonic constrains with a force constant of 1000 $kJ \cdot mol^{-1} \cdot nm^{-2}$. Subsequently, all systems were subjected to 100 ns MD simulations in NPT ensemble without any restraints, and the trajectories were stored every 10 ps for further MD analysis. To validate the accuracy of the MD trajectories, another 60 ns MD simulations of the N-CodY_ILE system were carried out repeatedly using method as described above.

All the simulations were performed using program GROMACS version 4.0.4 [14, 15] in GROMOS96 force field [16]. The protein topology was based on the united-atom GROMOS96 force field [16], and ligand topology was generated by the PRODRG sever [17]. The linear constraint solver (LINCS) method [18] was applied to constrain all bond lengths involving hydrogen atoms, allowing integration time step of 2 fs. Long-range electrostatic interactions were calculated using the Particle Mesh Ewald (PME) algorithm [19, 20] with a cutoff of 1.0 nm. Short-range repulsive and attractive interactions were described by Lennard-Jones potential with a cutoff of 1.0 nm. Van der Waals interactions were set at a cutoff of 1.4 nm. The protein, solvate, ligand and ions were separately coupled to a temperature bath at 300 K with a coupling constant of 0.1 ps. The Parrinello-Rahman barostat [21, 22] was used to maintain pressure isotropically at 1.0 bar with a relaxation time constant of 2.0 ps [23, 24].

2.3 Essential dynamics analysis

Essential dynamics (ED) analysis has been conducted to reveal high-amplitude concerted motion using the Cartesian coordinates of the Cα atoms through the eigenvectors of the mass-weighted covariance matrix of the atomic positional fluctuations. Only the eigenvectors with large eigenvalues are important to assess the significant motions of protein, and the detailed principles have been described in elsewhere [25-27]. The root mean square inner product (RMSIP) between the essential subspaces of two different systems is calculated to assess their dynamical similarity [28, 29]. In the present study, the RMSIP value between the first 10 eigenvectors of two different systems is defined as follows:

$$RMSIP = \sqrt{\frac{1}{10}\sum_{i=1}^{10}\sum_{j=1}^{10}(\eta_i^a \nu_j^b)^2}$$

Where η_i^a and ν_j^b are the *i*th and *j*th eigenvectors from the different systems, respectively. The range of RMSIP value from 0 to 1 represents the similarity of two different systems: the value is 1 if the sampled subspaces are identical, in contrast, the value is 0 if they are orthogonal [28].

2.4 FPMD simulations

FPMD simulations can make the system forced to be desired state by pulling a ligand out of the binding site along a pre-defined direction. This mimics experimental measurements of pulling a ligand by atomic force microscopy (AFM) [30, 31]. As starting structures for FPMD simulations, three snapshots were taken from the initial equilibration phase of N-CodY_ILE system in open state. Each starting structure was solvated in a box by adding water molecules and ions, and the system size was increased to ensure ligand out of the cavity and could accommodate the dissociated protein and ligand. Thus, this yields three boxes with dimensions of 8.0×6.0×6.0, 6.0×9.0×6.0, and 6.0×6.5×10.0 nm^3 corresponding to *x, y, z* pulling directions, respectively. Water molecules and ions were equilibrated for 100 ps with positional restraints on the protein and ligand after 50000 steps energy-minimization, subsequently, the mass center of ligand was subjected to a harmonic spring potential, which was moved along the specified pulling direction with a constant velocity of 0.005 $nm \cdot ps^{-1}$, and the force constants were set 500 and 1000 $kJ \cdot mol^{-1} \cdot nm^{-2}$, respectively [32-34]. FPMD simulations along three different directions were performed for 700 ps with software GROMACS version 4.0.4 [14, 15]. All molecular figures in present study were generated with Chimera [35].

3 Results and Discussion

3.1 Equilibrium dynamics and flexibility of CodY

To obtain the dynamic characters of protein in different states, the conventional MD simulations were carried out for N-CodY_ILE, N-CodY_NO and N-CodY_DEL three systems. The root mean square deviations (RMSDs) of the protein and ligand versus the energy-minimized structure were calculated, as shown in Fig. 2. Fig. 2A depicts that the values of RMSDs for N-CodY_ILE system keep stable at 0.3 nm before 50 ns, then they increase distinctly and reach 3.0 nm at 60 ns, afterward they sharply rise to 12.0 nm. These results show that the ligand

gradually escapes and ultimately dissociates from the cavity of protein during the MD simulations.

In order to validate the reliability of trajectories from MD simulations, another 60 ns MD simulations were performed repeatedly for N-CodY_ILE system, and the RMSDs and superimposed snapshots are shown in Fig. 3A and 3B, respectively. The RMSDs of ligand versus the protein stabilize at 0.3 nm before 30 ns, then they increase and reach 6.0 nm at 45 ns, hereafter they have large fluctuations. However, the values of RMSDs of protein versus itself are always stable at 0.3 nm during 60 ns MD simulations. The superimposed snapshots display that ligand gradually moves away from the binding pocket and completely separates from the protein at 50 ns and locates beyond the binding site. The calculated results from repeat simulations also suggest that ligand begins escaping from the cavity at 30 ns and dissociates completely from the protein at 45 ns, further verifying the reliability of MD simulations.

To better observe the releasing pathway of ligand, the RMSDs before 60 ns of N-CodY_ILE system are given in Fig. 2B. It is obvious that the RMSDs of ligand versus the protein are basically maintained at 0.25 nm before 25 ns, after this they increase and maintain at 0.3 nm. While the RMSDs of protein versus itself are always stable at 0.25 nm before 25 ns, hereafter they gradually increase to 1.2 nm at 60 ns. These results indicate that the structure of complex is very stable before 25 ns, whereas ligand slowly escapes from the cavity from 25 ns, which results in the conformational change of protein, finally ligand is completely released from the cavity at 60 ns but the protein structure keeps stable. Meanwhile, the root mean square fluctuations (RMSFs) in Fig. 4A also demonstrate that the residues 90-110 (namely the loop linking β3 and β4) have high flexibility during the MD simulations. Except for C- and N-terminal residues, although the residues 15-25 and 115-125 are more flexible than this loop, they locate beyond the binding site, thus should not be taken into account for contribution to the release of ligand.

For N-CodY_NO system, the RMSDs in Fig. 2C increase gradually before 25 ns and then keep at 0.45 nm, which suggests that the protein structure without ligand is very stable. Likewise, the RMSDs of N-CodY_DEL in Fig. 2D increase before 30 ns, then they have slight fluctuations but basically stabilize at 0.40 nm. The result shows that the protein changes its conformation after removing ligand but keeps stable structure after reaching the equilibrium. Comparison of the average structures from the different MD simulations is shown in Fig. 5A. The significant differences are that helix α3 has moved downwards and generates a large displacement,

meanwhile the loop linking β3 and β4 has twisted and results in dramatic conformational changes, while other structures are almost steady. On the whole, the protein structure tends to gradually convert the conformation to the structure of protein without ligand in closed state during the ligand release.

3.2 Essential dynamics analysis

The concerted motions and dynamic properties of proteins are characterized by ED analysis. Table 1 shows the sum and percentage of eigenvalues for each system. As shown in Table 1, the first 10 eigenvectors contribute to the overall average square deviation for ~70%. Thus they were used to assess the similarities and the motions for three systems. The RMSIP values in Table 2 between N-CodY_ILE and N-CodY_NO, N-CodY_NO and N-CodY_DEL, N-CodY_ILE and N-CodY_DEL are 0.452, 0.455 and 0.599, respectively. These results reveal that the degrees of similarity in the motions of protein for three systems are low and ligand in the presence of the binding pocket affects the motion of protein. However, the similarity between N-CodY_ILE and N-CodY_DEL systems is slightly greater, which shows the motional direction of N-CodY_ILE after releasing ligand is most similar to that of N-CodY_DEL. Meanwhile, traces of covariance matrix of N-CodY_ILE and N-CodY_DEL systems are 10.384 and 6.714 nm^2, respectively, further verifying that ligand restricts the motions of proteins.

To compare regions sampled by the essential motions in different states, two-dimensional (2D) projections of backbone atomic trajectories along with the first two eigenvectors are shown in Fig. 6A. Notable differences in fluctuation were observed among the three states along eigenvectors 1 and 2. It is clear that the conformational space of N-CodY_ILE covers the entire space of N-CodY_DEL system, and three systems have partial overlay of the conformational spaces, but there is significant difference in N-CodY_NO system. These results show that there is a considerable overlap in the cluster of conformations for three systems, especially N-CodY_ILE and N-CodY_DEL, but ligand bound to the protein restricts the motions and conformational space of protein, which is consistent with the above RMSIP values.

3.3 Isoleucine-release pathway from the conventional MD simulations

As indicated in Fig. 7A, ligand is apparently encapsulated in the hydrophobic cavity, and its uptake and release must pass the entrance of the cavity. The experimental result revealed that N-CodY_NO possessed a cavity whose entrance was bounded by the β2-loop-α3 segments. To

verify the accurate isoleucine-release pathway, the trajectories from the conventional MD simulations were analyzed in detail.

From Fig. 7A one can see that the hydrophobic interactions exclude ligand from the binding site. In addition, hydrogen bonds play a critical role for the release of isoleucine. Fig. 7B clearly shows the releasing-pathway and the surrounding residues. According to the trajectories of the N-CodY_ILE system, the distances of heavy atoms between ligand and the surrounding residues are given in Fig. 8A and 8B. It can be seen that before 43 ns the distances between the N atom in ligand and O atom in THR96 and in PHE98 keep at 0.2 nm and 0.3 nm, respectively, and then both of them increase. In contrast, the distances between the other residues and ligand (including VAL100 N⋯⋯O, GLU101 N⋯⋯OXT, GLU101 OE1⋯⋯N, GLU101 OE2⋯⋯N and ASN102 N⋯⋯OXT) fluctuate before 43 ns, but they decrease after 43 ns. Meanwhile, the distances of two O atoms (O and OXT) in ligand interacting with two N atoms (NH1 and NH2) in ARG61 increase at the beginning of MD simulations, and the values fluctuate at 0.8 nm, but they rise at 45 ns. These results indicate that ligand gradually escapes from the bottom to the top of the cavity under the mediation of hydrogen-bonds and hydrophobic interactions, and it successively forms hydrogen bonds interactions with a series of residues, such as ARG61, THR96, PHE98, VAL100, GLU101 and ASN102. Fig. 8C displays that the distance of salt-bridge between ARG61 and GLU101 stabilizes at 0.5 nm before 30 ns, and then it gradually increases. However, the distance of salt-bridge between ARG61 and ASP104 maintains at 1.4 nm before 25 ns, after this it slowly reduces. This verifies that the salt-bridge of ARG61 with GLU101 disappears in N-CodY_ILE and that of ARG61 with ASP104 forms upon ligand escaping, which is in good agreement with the experimental alternative salt-bridge partners [11].

The average structures of N-CodY_ILE system from the first 20 ns and the last 70 ns MD simulations are superimposed in Fig. 5B. It is clear that the conformations of the binding pocket have obvious structural rearrangement. ARG61 is almost perpendicular to the face of the aromatic ring of TYR75 in N-CodY_ILE system, while it has a 2.59 Å displacement and packs against the face of TYR75 after ligand release. Meanwhile, the aromatic face of TYR75 rotates upward about 62° and occupies the position of ligand in the binding site in N-CodY_NO. The carbonyl of THR96 interacting with ligand moves away 7.79 Å to the peripheral surface, and GLU101 moves outward 5.52 Å with ligand escaping. The RMSD is 3.3 Å between the whole loop linking β3 and

β4, further verifying that the loop has high flexibility, which is in consistent with the experimental B factors [11]. The snapshots are superimposed every 10 ns from 0 to 60 ns MD simulations as shown in Fig. 7C and 7D, which clearly show that ligand-release is evidently correlated with the movement of the loop linking β3 and β4, and the loop gradually sweeps into the cavity occupied by ligand. These detail structural rearrangements agree with the experimental results [11].

3.4 Isoleucine-release pathways from FPMD simulations

The conventional MD simulations show that ligand-release is correlated with the loop linking β3 and β4.To explore a variety of releasing pathways of ligand from the binding pocket and avoid bias, FPMD simulations are carried out to study the releasing mechanism of ligand by means of a pulling force in different predefined directions. According to the conformational changes observed from the conventional MD simulations, three pulling directions were chosen as shown solid arrows in Fig. 9A: the mass center of isoleucine was separately pulled from the cavity along the directions opposite to MET62, ILE127 and PRO72, which were correspondingly defined as *x*, *y* and *z* directions.

Fig. 9B and 9C show the comparison of the force profiles obtained from six FPMD simulations with two force constants of 500 and 1000 $kJ \cdot mol^{-1} \cdot nm^{-2}$. It can be seen that at two force constants, *x* pulling direction yields slightly lower rupture forces (199.802 and 237.442 $kJ \cdot mol^{-1} \cdot nm^{-1}$ corresponding to 500 and 1000 $kJ \cdot mol^{-1} \cdot nm^{-2}$, respectively) than other two directions (*y* direction: 253.797 and 259.501 $kJ \cdot mol^{-1} \cdot nm^{-1}$, *z* direction: 213.792 and 311.962 $kJ \cdot mol^{-1} \cdot nm^{-1}$ corresponding to 500 and 1000 $kJ \cdot mol^{-1} \cdot nm^{-2}$, respectively). In addition, although ligand pulled out of the cavity needs a little longer time along *x* direction than other two directions, the pulling force of *x* direction is the lowest among three directions. These results show that it is the easiest way to pull ligand out of the cavity along *x* direction.

Fig. 10A shows the superimposed snapshots along *x* pulling direction written every 50 ps from 0~400 ps FPMD simulations, and the number of hydrogen bond during the pulling process is also given in Fig. 10B. It can be seen that, firstly, ligand forms hydrogen bonds with ARG61, THR96 and VAL100 in the binding pocket. Subsequently, these bonds are broken, and new hydrogen bond is generated between ligand and GLU94 after 80 ps. However, hydrogen bonds of ligand interacting with GLU101 and ASN102 are not observed during the pulling process, because isoleucine was pulled out of the cavity along the specified *x* direction. The above results are

basically in accordance with the pathway observed from the conventional MD simulations.

RMSFs of the CodY protein from three pulling directions under two conditions are also shown in Fig. 4B and 4C. Comparing the RMSFs, one can see that the character of flexibilities of the CodY proteins from FPMD simulations is similar to that of conventional MD simulations, i.e., except for C- and N-terminal residues, the residues 15-25, 90-110 and 115-125 have high flexibility. However, only the residues 90-110 in the binding site play an essential role in the releasing pathway. Moreover, the RMSFs from FPMD are lower than those of conventional MD simulations, and the values of RMSFs along *x* direction are the lowest among the three pulling directions.

In addition, three trajectories from FPMD with the force constant of 500 $kJ \cdot mol^{-1} \cdot nm^{-2}$ were also projected onto the first two eigenvectors in Fig. 6B. Comparing with the conventional MD simulations, the three FPMD simulations have the smaller conformational spaces. The space of *z* direction is the largest and covers other two spaces, especially includes the entire space of *x* direction. Although spaces for three directions are overlapped to a great extent, the space of *x* direction is the smallest and its conformation is the most stable. In summary, *x* direction is favorable for pulling ligand from protein, while z direction is unfavorable and consumes more time, further verifying that ligand-release is accompanied with the movement of loop linking β3 and β4.

4 Conclusions

The present work investigates the releasing pathway of the isoleucine ligand from the GAF domain of N-CodY by using conventional MD, ED analysis and FPMD simulations. The results from ED analysis indicate that there is low similarity in motions of protein for N-CodY_ILE, N-CodY_NO and N-CodY_DEL three systems, and ligand in the presence of the binding site restricts the motion of protein. Comparing the average structures from the different conventional MD simulations, it can be found that the protein gradually changes its conformation from open state to closed state upon ligand releasing. Ligand moves away from the binding pocket through sequential hydrogen bonds interactions with a succession of residues, for example, ARG61, THR96, PHE98, VAL100, GLU101 and ASN102. In addition, FPMD simulations show that the easiest way to pull ligand out of the cavity is along *x* direction, which is accordance with the result from conventional MD simulations.

In summary, the escape of ligand from the binding site is accompanied with the movement of loop linking β3 and β4, and in response the protein changes its conformation from open state to closed state after ligand release.

Acknowledgements

This work was sponsored by National Natural Science Foundation of China (21271115, 31100584) and Shandong Province (ZR2010BM010). Special thanks also go to the Scientific Research Startup Funds (20100205) and the School Youth Funds (XJ201211) of Qufu Normal University for their funding.

References

[1] F.J. Slack, J.P. Mueller, A.L. Sonenshein, Mutations that relieve nutritional repression of the Bacillus subtilis dipeptide permease operon, J. Bacteriol. 175(15) (1993) 4605-4614.

[2] A.L. Sonenshein, CodY, a global regulator of stationary phase and virulence in Gram-positive bacteria, Curr. Opin. Microbiol. 8 (2005) 203-207.

[3] M. Ratnayake-Lecamwasam, P. Serror, K.W. Wong, A.L. Sonenshein, Bacillus subtilis CodY represses early-stationary-phase genes by sensing GTP levels, Genes Dev. 15 (2001) 1093-1103.

[4] B.R. Belitsky, A.L.Sonenshein, Contributions of multiple binding sites and effector-independent binding to CodY-mediated regulation in *Bacillus subtilis*, J. Bacteriol. 193 (2011) 473-484.

[5] V. Molle, Y. Nakaura, R.P. Shivers, H. Yamaguchi, R. Losick, Y. Fujita, A.L. Sonenshein, Additional targets of the *Bacillus subtilis* global regulator CodY identified by chromatin immunoprecipitation and genome-wide transcript analysis, J. Bacteriol. 185 (2003) 1911-1922.

[6] R.P. Shvers, A.L. Sonenshein, Activation of the *Bacillus subtilis* global regulator CodY by direct interaction with branched-chain amino acids. Mol. Microbiol. 53 (2004) 599-611.

[7] L. Stenz, P. Francois, K. Whiteson, C. Wolz, P. Linder, J. Schrenzel, The CodY pleiotropic repressor controls virulence in gram-positive pathogens, FEMS immunol. Med. Microbiol. 62 (2011) 123-139.

[8] L. Aravind, C.P. Ponting, The GAF domain: an evolutionary link between diverse phototransducing proteins, Trends Biochem. Sci. 22 (1997) 458-459.

[9] Y.S. J.Ho, L.M. Burden, J.H. Hurley, Structure of the GAF domain, a ubiquitous signaling

motif and a new class of cyclic GMP receptor, EMBO J. 19 (2000) 5288-5299.

[10] V.M. Levdikov, E. Blagova, P. Joseph, A.L. Sonenshein, A.J. Wilkinson, The structure of CodY, a GTP- and isoleucine-responsive regulator of stationary phase and virulence in gram-positive bacteria, J. Biol. Chem. 281 (2006) 11366-11373.

[11] V.M. Levdikov, E.B. Blagova, V.L. Colledge, A.A. Lebedev, D.C. Williamson, A.L. Sonenshein, A.J. Wilkinson, Structural rearrangement accormpanying ligand binding in the GAF domain of CodY from *Bacillus subtilis*, J. Mol. Biol. 390 (2009) 1007-1018.

[12] J.L. Klepeis, K. Lindorff-Larsen, R.O Dror, D.E. Shaw, Long-timescale molecular dynamics simulations of protein structure and function, Curr. Opin. Struct. Biol. 19 (2009) 120-127.

[13] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, J. Hermans, Interaction models for water in relation to protein hydration, Dordrecht, The Netherlands: D. Reidel publishing company, 1981, pp. 331-342.

[14] D. van der Spoel, E. Lindahl, B. Hess, A.R. van Buuren, E. Apol, P.J. Meulenhoff, D.P. Tieleman, A.L.T.M. Sijbers, K.A. Feenstra, R. van Drunen and H.J.C. Berendsen, Gromacs User Manual version 4.0, www.gromacs.org (2005).

[15] D. van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J.C. Berendsen, GROMACS: Fast, Flexible and Free. J. Comp. Chem. 26 (2005) 1701-1718.

[16] C. Oostenbrink, A. Villa, A.E. Mark, W.F. van Gunsteren, A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6, J. Comput. Chem. 25 (2004) 1657-1676.

[17] A.W. Schüttelkopf, D.M.F. van Aalten, PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta* Crystallogr. D 60 (2004) 1355-1363.

[18] B. Hess, H. Bekker, H.J.C. Berendsen, J.G.E.M. Fraaije, LINCS: A linear constraint solver for molecular simulations, J. Comput. Chem. 18 (1997) 1463-1472.

[19] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems, J. Chem. Phys. 98 (1993) 10089-10092.

[20] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, A smooth particle mesh Ewald method, J. Chem. Phys. 103 (1995) 8577-8592.

[21] M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method, J. Appl. Phys., 52 (1981) 7182-7190.

[22] S. Nosé, M.L. Klein, Constant pressure molecular dynamics for molecular systems, Mol. Phys. 50 (1983) 1055-1076.

[23] J.A. Lemkul, D.R. Bevan, Assessing the stability of Alzheimer's amyloid protofibrils using molecular dynamics, J. Phys. Chem. B 114 (2010) 1652-1660.

[24] J.S. Hub, M.B. Kubitzki, B.L. de Groot, Spontaneous quaternary and tertiary T-R transitions of human hemoglobin in molecular dynamics simulation, Plos Comput. Biol. 6 (2010) e1000774.

[25] A. Amadei, A.B.M. Linssen, H.J.C. Berendsen, Essential dynamics of proteins, Proteins 17 (1993) 412-425.

[26] D.M.F. van Aalten, A. Amadei, A.B.M. Linssen, V.G.H. Eijsink, G. Vriend, H.J.C. Berendsen, The essential dynamics of thermolysin: conformation of the hinge-bending motion and comparison of simulations in vacuum and water, Proteins 22 (1995) 45-54.

[27] D.M.F. van Aalten, J.B.C. Findlay, A. Amadei, H.J.C. Berendsen, Essential dynamics of the cellular retinol-binding protein evidence for ligand-induced conformational changes, Prot. Eng. 8 (1995) 1129-1136.

[28] D.M.F. van Aalten, A. Amadei, R. Bywater, J.B.C. Findlay, H.J.C. Berendsen, C. Sander, P.F.W. Stouten, A comparison of structural and dynamic properties of different simulation methods applied to SH3, Biophys. J. 70 (1996) 684-692.

[29] A. Amadei, M.A. Ceruso, A. Di Nola, On the convergence of the conformational coordinates basis set obtained by the essential dynamics analysis of proteins molecular dynamics simulations, Proteins 36 (1999) 419-424.

[30] B. Heymann, H. Grubmüller, Molecular dynamics force probe simulations of antibody/antigen unbinding: entropic control and nonadditivity of unbinding forces, Biophys. J. 81 (2001) 1295-1313.

[31] B. Heymann, H. Grubmüller, ANO2/DNP-hapten unbinding forces studied by molecular dynamics atomic force microscopy simulations, Chem. Phys. Lett. 303 (1999) 1-9.

[32] X. Zhao, S. Wang, X.F. Gao, X.R. Huang, C.C.Sun, Molecular dynamics simulations investigation of neocarzinostatin chromophore-releasing pathways from the holo-NCS protein, J. Struct. Biol., 169 (2010) 14-24.

[33] F. Gräter, B.L. de Groot, H.L. Jiang, H. Grubmüller, Ligand-release pathways in the pheromone-binding protein of *Bombyx mori*, Structure, 14 (2006) 1567-1576.

[34] M. Cecchini, Y. Alexeev, M. Karplus, Pi release from myosin: A simulation analysis of possible pathways, Structure, 18 (2010) 458-470.

[35] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, and T.E. Ferrin, UCSF chimera – A visualization system for exploratory research and analysis, J. Comput. Chem. 25 (2004) 1605–1612.

Figure Captions

Figure 1 Three-dimensional structure (PDB code 2B18) of the GAF domain of N-terminal CodY bound the isoleucine ligand. The protein is represented in ribbon and ligand in ball and stick. The segments with high flexibility are highlighted in green color.

Figure 2 Time evolution of RMSDs from three different systems during 100 ns MD simulations. (A) N-CodY_ILE in open state; (B) N-CodY_ILE in open state before 60 ns MD simulations; (C) N-CodY_NO in closed state; (D) N-CodY_DEL in open state. The black and red lines represent the RMSDs of protein and ligand versus protein in its original state, respectively.

Figure 3 (A) Time evolution of RMSDs and (B) the superimposed snapshots written in every 10 ns from repeat simulation of N-CodY_ILE system during another 60 ns MD simulations. The protein is shown in flat ribbon, and ligand in ball and stick model. The starting structure is colored in blue, and snapshots at 10 ns, 20ns, 30ns, 40 ns, 50 ns in cyan, green, chartreuse, yellow and red, respectively.

Figure 4 RMSFs of the Cα atoms around the energy-minimized structure for the conventional MD and FPMD simulations. (A) RMSF of three different systems during MD simulations. RMSF from FPMD simulations along three different directions at the force constants of 500 $kJ \cdot mol^{-1} \cdot nm^{-2}$ (B) and 1000 $kJ \cdot mol^{-1} \cdot nm^{-2}$ (C).

Figure5 Superimposition of the average structures from different MD simulations. Side view is shown on the left and top view on the right after rotating 90°. The protein is shown in ribbon with different color, and ligand in ball and stick. The distance of heavy atoms is in red dashed line. The segments with conformational changes were highlighted in real line. (A) The energy-minimized structure of N-CodY_ILE from crystal structure is colored yellow, the average structure from N-CodY_ILE after releasing ligand in cyan, that of N-CodY_NO in blue and that of N-CodY_DEL in magenta. (B) The average structure of N-CodY_ILE from the first 20 ns is shown in cyan and that of the last 70 ns in magenta.

Figure 6 2D projections of backbone atomic trajectories along the first two eigenvectors derived from conventional MD (A) and FPMD (B) simulations.

Figure 7 Releasing pathway of ligand from the binding pocket of N-CodY protein and snapshots are superimposed every 10 ns from 0 to 60 ns MD simulations. (A) Ligand bound to the binding pocket which is shown in hydrophobic surface. The hydrophobic surface shows amino acid

hydrophobicity with colors ranging from dodger blue for the most hydrophilic to white at 0.0 to orange red for the most hydrophobic. (B) The residues within 8 Å around ligand in the binding site. The hydrophobic residues are in red, the hydrophilic residues in blue and the residues in between in black. (C) Top view and (D) Side view of the releasing pathway of ligand from the binding site. The proteins are shown in flat ribbon and ligand is in ball and stick. The starting structure is colored in light gray, snapshots at 10 ns, 20 ns, 30ns, 40ns, 50ns and 60ns in magenta, cyan, yellow, red , blue and chartreuse, respectively. The releasing pathway is shown in dashed arrow with magenta color, and the segments with high flexibility correlated with the release of ligand are highlighted in magenta.

Figure 8 Time dependence of distances of heavy atoms and salt-bridge of N-CodY_ILE system during 60 ns MD simulations. (A) The alternative hydrogen bonds arrangements and (B) Distances of heavy atoms between ligand and the surrounding residues. (C) Distances of salt bridge between the alternative partners.

Figure 9 Chosen pulling directions in FPMD (A). The force profiles from *x*, *y*, and *z* directions with the force constants of 500 $kJ \cdot mol^{-1} \cdot nm^{-2}$ (B) and 1000 $kJ \cdot mol^{-1} \cdot nm^{-2}$ (C). The *x*', *y*' and *z*' are smoothed through signal process by adjacent-averaging method every 50 points from *x*, *y* and *z* data.

Figure 10 (A) Superimposed snapshots along *x* direction written every 50 ps from 0~400 ps FPMD simulations. Ligand is represented in ball and stick, and the proteins in flat ribbons. Snapshots at 50, 100, 150, 200, 250, 300, 350, 400 ps are colored in blue, sky blue, cyan, chartreuse, green, yellow, orange and red, respectively. (B) The number of hydrogen bonds of ligand interacting with the surrounding residues from 0~400 ps FPMD simulations.

Table 1 Sum of eigenvalues and the percentage of eigenvalues for each system

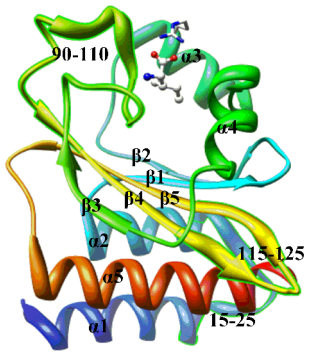
Systems	N-CodY_ILE	N-CodY_NO	N-CodY_DEL
Sum of the first 10 eigenvalues (nm)	8.035	6.371	4.353
Sum of total eigenvalues (nm)	10.384	9.158	6.714
Percentage (%)[a]	77.38	69.57	64.84

[a] Percentage is calculated from sum of the first 10 eigenvalues divided by sum of total eigenvalues.

Table 2 The RMSIP between two systems and the trace of covariance matrix for each system

Systems	N-CodY_ILE	N-CodY_NO	N-CodY_DEL
N-CodY_ILE	-	0.452	0.599
N-CodY_NO	0.452	-	0.455
N-CodY_DEL	0.599	0.455	-
Trace of covariance matrix (nm^2)	10.384	9.156	6.714

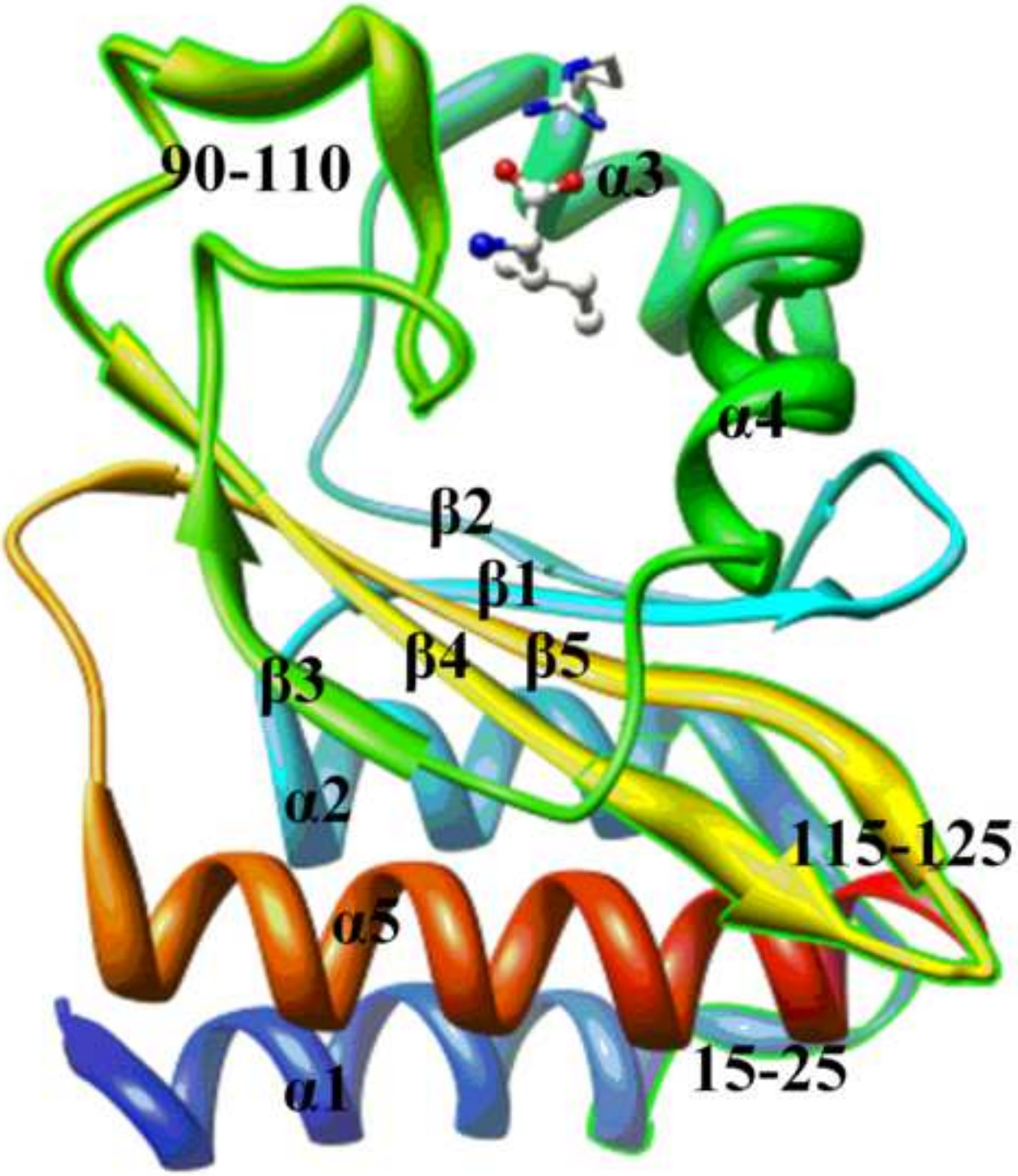
Graphical Abstracts



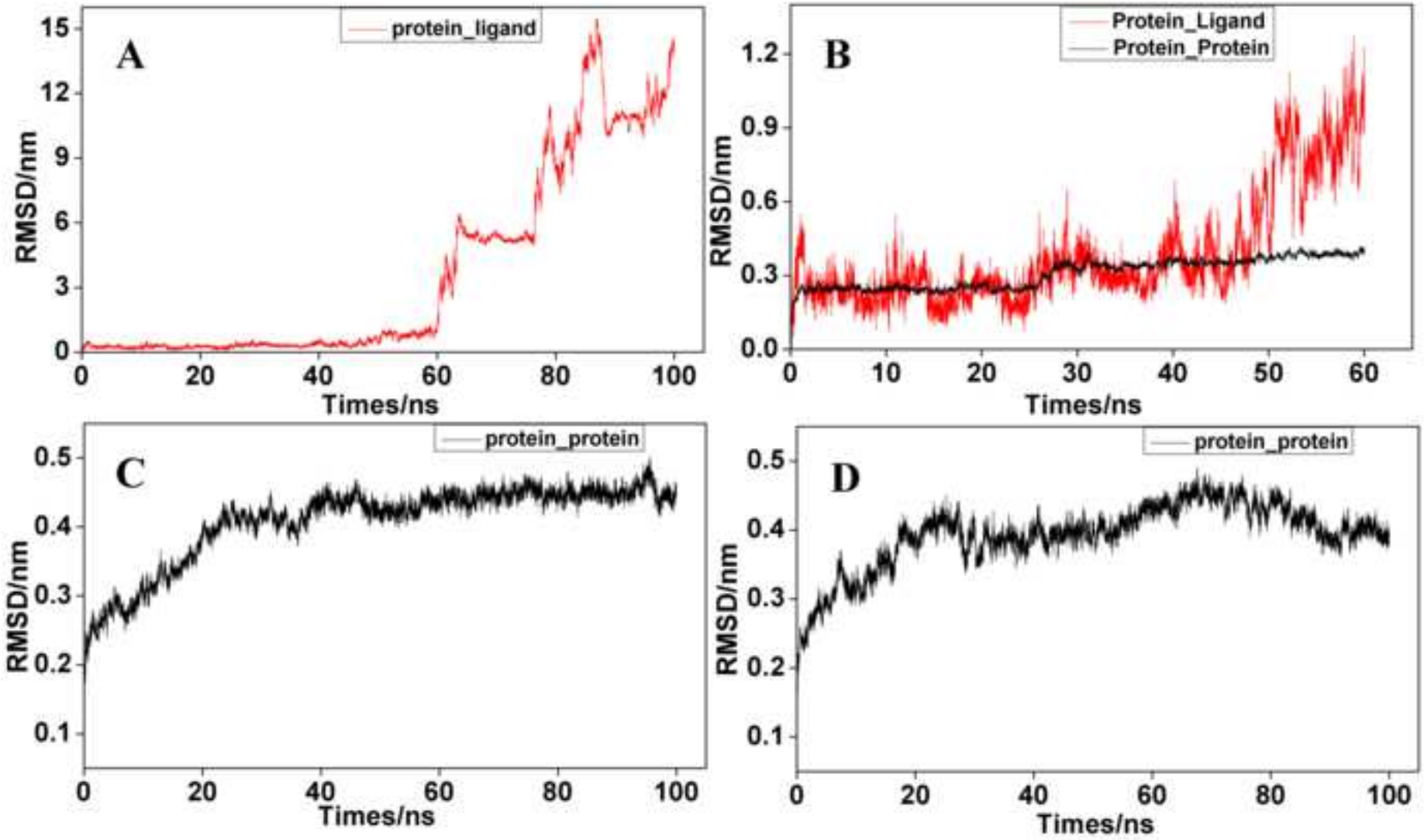
Isoleucine-release pathway from GAF domain in N-CodY is studied by using conventional MD, ED analysis and FPMD simulations. The detail releasing pathway is observed, and the conformational changes during isoleucine release are consistent with the experimental results. These results give useful information for isoleucine entering mechanism to the cavity.

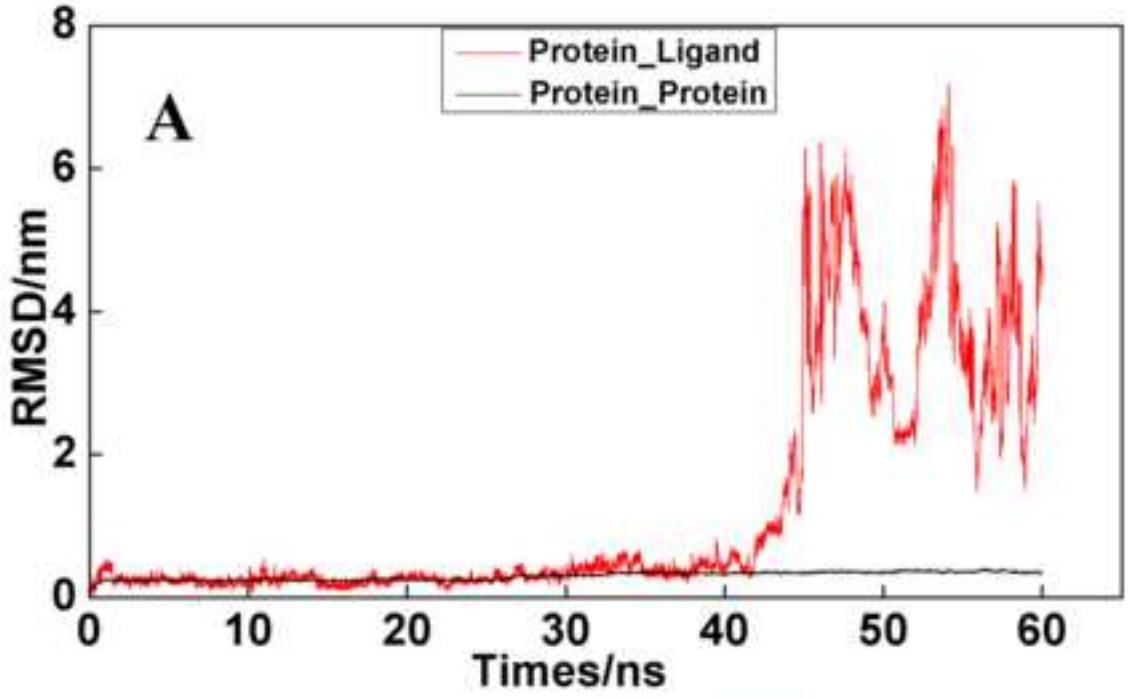
Highlights

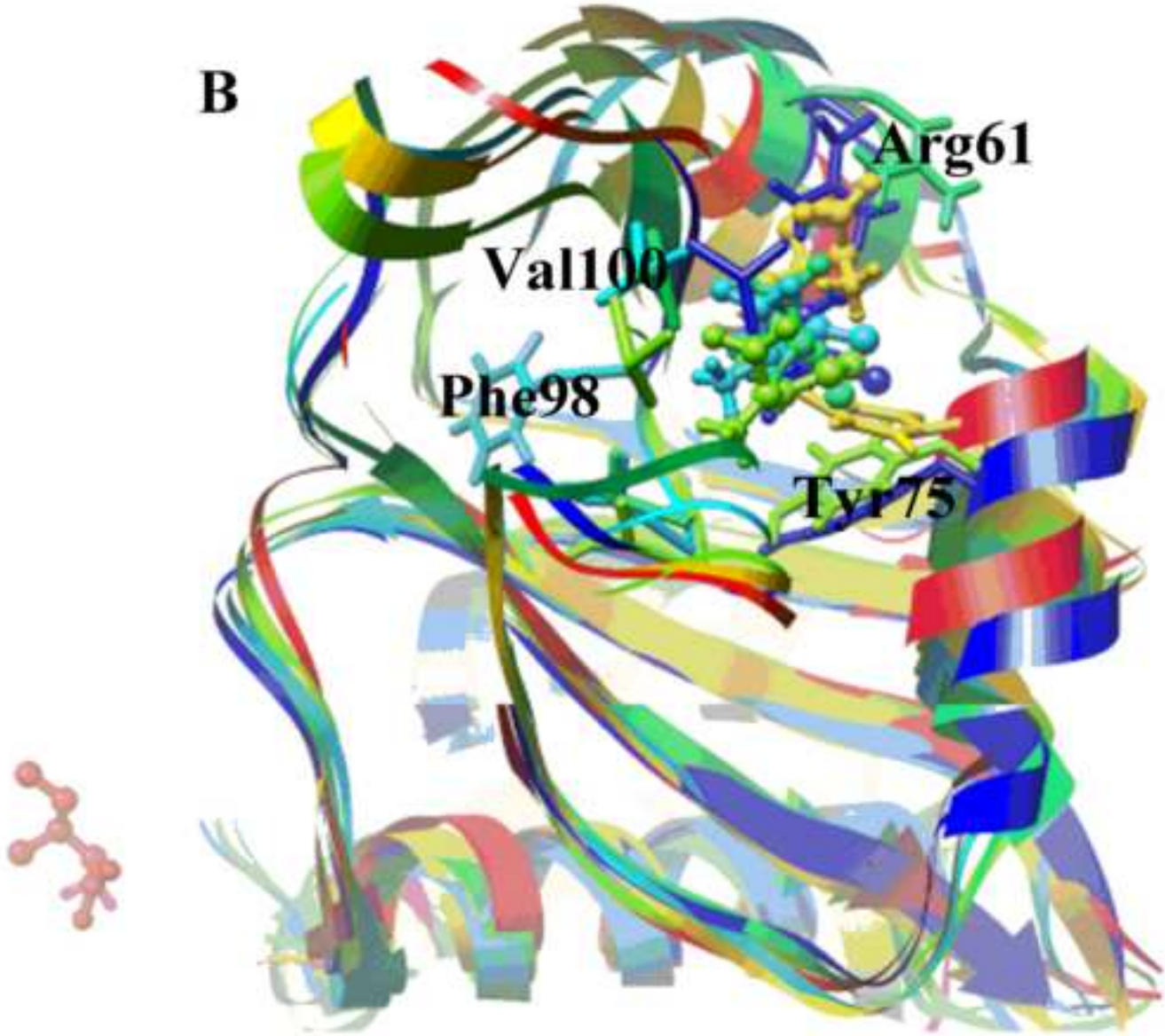
1. Isoleucine-release pathway from GAF domain of N-CodY is proposed based on the MD results.
2. The conformational changes are observed during isoleucine release from the binding pocket.
3. The results from FPMD simulations further determine the ligand releasing pathway.

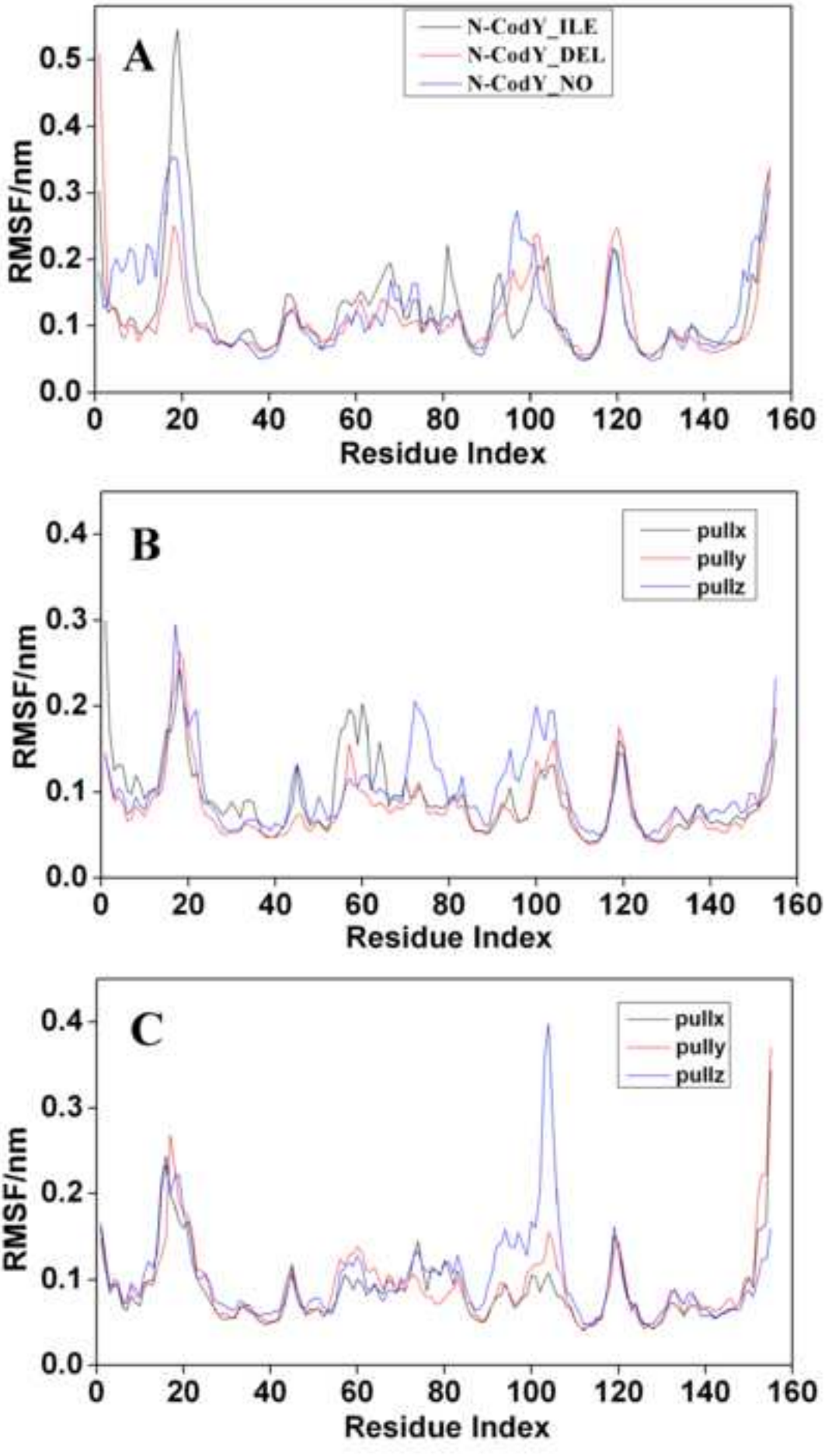


Figure(s)

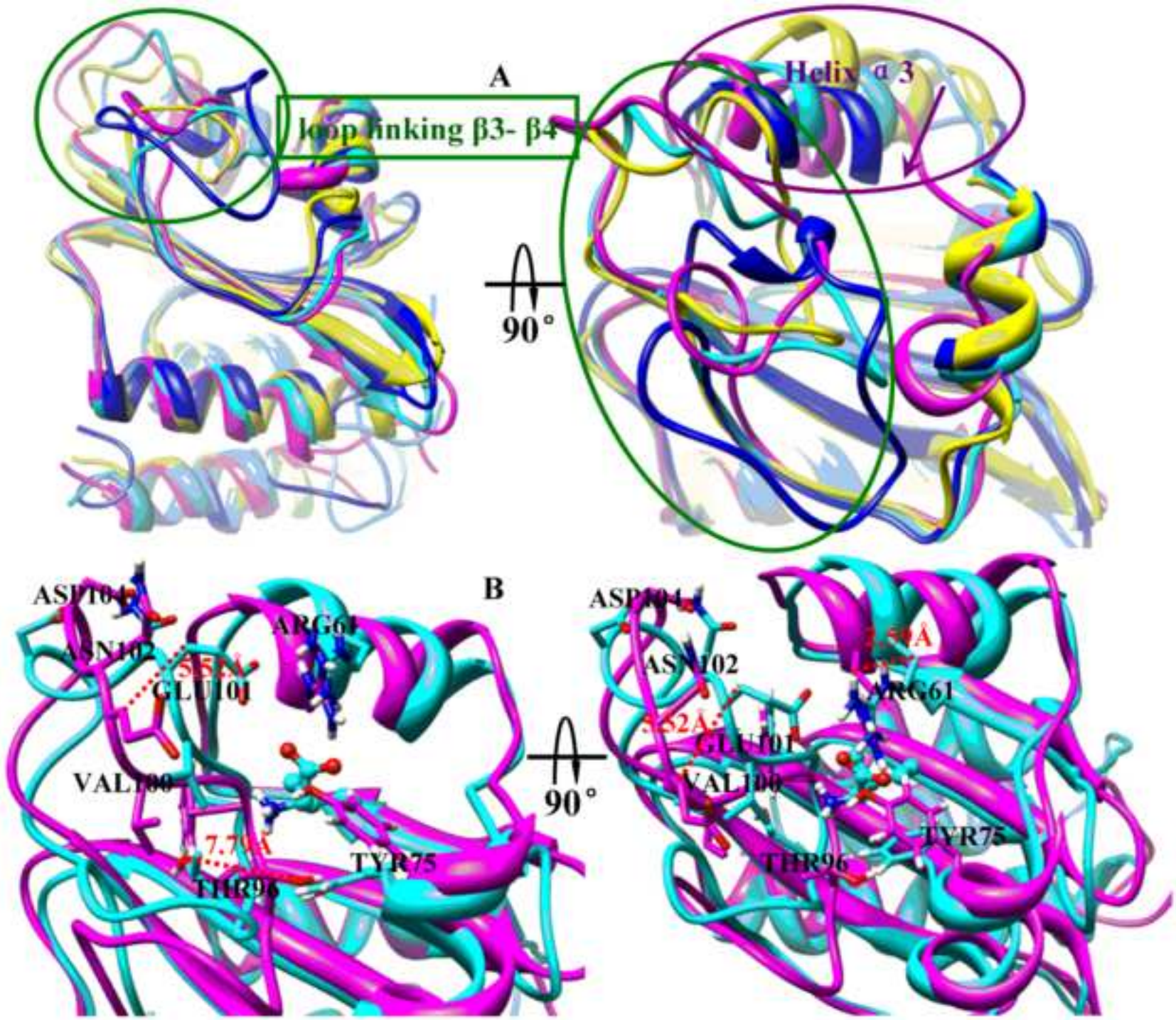

A
protein_ligand
RMSD/nm
Times/ns
B
Protein_Ligand
Protein_Protein
RMSD/nm
Times/ns
C
protein_protein
RMSD/nm
Times/ns
D
protein_protein
RMSD/nm
Times/ns







Figure(s)


A
loop linking β3- β4
Helix α 3
90°
B
ASP104
ASN102
ARG61
5.55Å
GLU101
VAL100
7.79Å
TYR75
THR96
90°
ASP104
ASN102
5.59Å
ARG61
5.52Å
GLU101
VAL100
TYR75
THR96