약학박사학위논문

인간 유래 Filamin A의 Dimerization Domain 과 대장균 유래 CRP의 Allosteric Conformational Change에 대한 구조적 통찰

Structural Insights into the Dimerization Domain of human Filamin A and Allosteric Conformational Change of CRP from *Escherichia coli*

2014 年 8 月

서울대학교 대학원

약학과 물리약학전공

석승현

Abstract of the Dissertation

Structural Insights into the Dimerization Domain of human Filamin A and Allosteric Conformational Change of CRP from *Escherichia coli*

Seung-Hyeon Seok

Physical Pharmacy, Research Institute of Pharmaceutical Sciences

The Graduate School

Seoul National University

Dissertation director: Professor Bong-Jin Lee

Filamins play important roles in regulating the dynamics of the actin cytoskeleton which plays a central role in many cell functions and plays as integrators of cell mechanics and signalling by interacting with transmembrane receptors and

cytosolic signaling proteins. In humans, three filamin isoforms have been identified: filamin A, filamin B and filamin C. Of these, filamin A is the most abundant and widely expressed. FLNa missense mutations cause familial cardiac valvular dystrophy and putative gain-of-function mutations result in a spectrum of congenital malformations generally characterized by skeletal dysplasias.

Human vertebrate filamins are homodimers of two 280kDa subunits, and each subunit contains an N-terminal actin binding domain consisted of two calponin homology domains followed by 24 tandem repeat domains and two hinge regions. 24^{th} repeat domain was dimerization domain, which results in a V-shaped flexible parallel homodimer. Here, we report the structure of FLNa domain 24 (FLNa24), compare the structure with FLNc24, and discuss how dimerization is formed in FLNa24.

CRP has long served as a typical textbook example describing transcription regulation, DNA-binding motif, and allosteric activation of a protein. CRP plays a critical role in regulating the transcription of more than 200 genes by binding to specific DNA sites and interacting with RNA polymerase. The inactive CRP structure would provide an insight into the structural basis of CRP. In the present work, we succeeded in obtaining higher-resolution (2.2 Å) crystal structure of the wild-type apo-CRP,

which would be the most proper template for precise inspection of inactive conformation and establishes here the mode of conformational transition in atomic detail. In addition, we report the first crystal structure of another inactive form, the cGMP-CRP complex, at also 2.2 Å resolution, which not only support our insight into the conformational allosterism, but also address how CRP discriminates the false ligand cGMP from its authentic effector cAMP. And also the conformational allostery change of CRP by cyclic mucleotide and structural difference between apo-CRP, cAMP-CRP and cGMP-CRP was discussed.

.....

Keywords: Filamin; FLNa24; dimerization; CRP; Crystallography; allostery; cyclic nucleotide

Student number: 2008-30482

Contents

List of Tables

List of Figures

Abbreviations

ADSC	Area Detector Systems Corporation
cAMP	Cyclic Adenosine 3',5'−MonoPhosphate
cGMP	Cyclic Guanosine 3',5'−MonoPhosphate
CRP	Cyclic AMP receptor protein
E. coli.	*Escherichia coli*
IPTG	Isoprophl−β−D−ThioGalactopyranoside
OD	Optical Density
PAGE	Polyacrylamide Gel Electrophoresis
RMSD	Root Mean Square Deviation
SDS	Sodium Dodecyl Sulfate

Chapter 1. Structural insight into the dimerization domain of human filamin A

1.1 Introduction

Filamins play important roles in regulating the dynamics of the actin cytoskeleton through crosslinking actin filaments. Filamins are actin-crosslinking proteins that engender mechanical force to cells by binding to actin filaments and making the filamentous actin form bundles or gel networks[1]. Reorganization of actin cytoskeleton plays a central role in many cell functions such as the maintenance of cell shape, cell division, adhesion, motility, signal transduction and protein sorting[1]. Filamins as integrators of cell mechanics and signalling by interacting with transmembrane receptors and cytosolic signaling proteins and scaffolding signaling molecules of diverse functions[1,2] (Figure 1).

In humans and mice, three filamin isoforms have been identified: filamin A, filamin B and filamin C. Of these, filamin A (FLNa) is the most abundant and widely expressed[1,2]. In mice, FLNa expression is required for cardiac and vascular development[3], FLNb is required for skeletal and microvascular development[4] and FLNc is required for normal myogenesis[5]. In

humans, heterozygous null FLNa alleles result in defective neuronal migration causing periventricular heterotopias and oto palate digital (OPD) spectrum disorders[6] that show symptoms such as cleft palate, facial malformation and bone dysplasia[7]. Mutations in FLNb cause a class of dieases with abnormal vertebral segmentation, joint formation and skeletogenesis[8] and FLNc mutation causes an autosomal dominant myofibrillar myopathy[9]. The diversity in phenotypes associated with different filamin mutations shows that filamins play a variety of essential functions.

There are multiple forms of filamins across species, mainly differing in the number of tandem repeats and the presence of hinge regions[1]. Human vertebrate filamins are homodimers of two 280kDa subunits, and each subunit contains an N-terminal actin binding domain consisted of two calponin homology domains followed by 24 tandem repeat domains (FLNa1-24) that are interrupted by flexible hinge regions between FLNa15 and FLNa16 and FLNa23 and FLNa24[1, 2]. Dimerization through FLNa24 is crucial for the actin-crosslinking function of filamins. Dimerization results in a V-shaped flexible parallel homodimer that can promote high-angle branching of actin filaments and is essential for stabilizing the orthogonal actin networks at the leading edge[10]. In addition to crosslinking actin filaments, filamins act as scaffolds for more than 30 transmembrane

receptors and cytosolic signaling proteins[1]. In general, C-terminal repeat domians of filamin are major regions of these interactions (Figure 2).

Structures of various domains of each isoform have been reported. The actin binding domain, FLNa domain 16-17, 18-19, 23 and 19-21 were determined. And also structures of Ig-like domains of other filamin isoforms were reported. The complex structure of FLNa domain 21 and peptide from β-integrin cytoplasmic tail shows that the peptide binds to the β-strands C and D of filamin domain[11]. Complex structure of FLNa domain 17 (FLNa17) and glycoprotein Ibα (GP Ibα) tail also shows that the binding partner interacts with filamin similarly[12]. Structure of FLNc domain 24 (FLNc24) showed how dimerization is formed in the carboxy terminal of filamin[13]. Speculating from these structures, the face consisting of strands C and D of filamin domains might be the general interaction interfaces with other ligands.

Although structures of various domains have been reported, despite a lot of functional and interactional studies of filamin, relatively little is known about structure of filamin and its domains. So the main purpose of this study is structural determination of FLNa24. Here, we report the structure of FLNa domain 24 (FLNa24), compare the structure with FLNc24, and discuss how dimerization is formed in FLNa24.

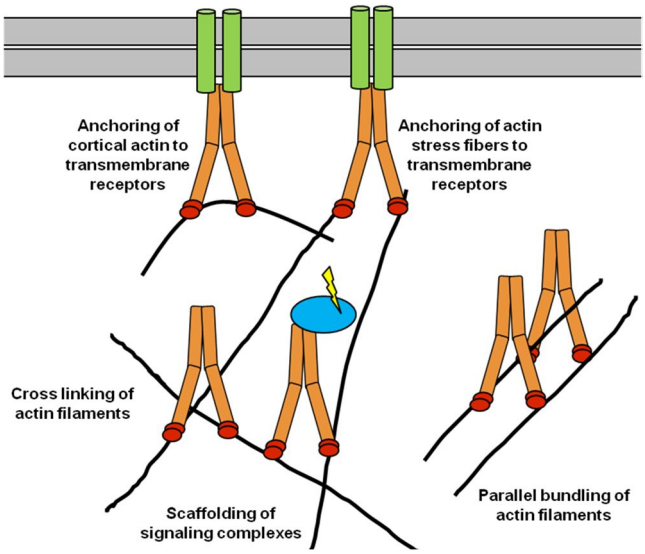

Figure 1. Schematic model of filamin function in F-actin cross-linking, receptor anchoring and cell signaling[2].

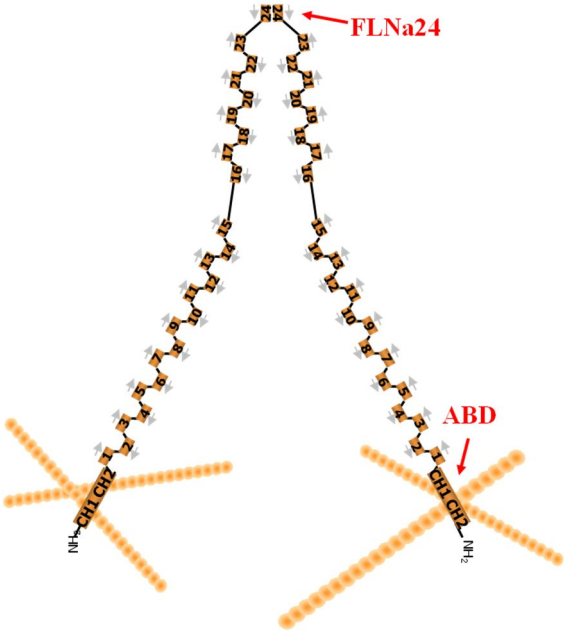

Figure 2. Diagram of a dimer of FLNa. There are two actin binding domains in N-terminal of FLNa and 23th filamin repeats terminating in a 24th – the dimerization domain.

1.2 Experiment Procedure

1.2.1 Protein expression and purification

The open reading frame (ORF) of FLNa24 from *homo sapience* (residues 2559-2647) was cloned into the expression vector pGEX4T-1 and was expressed in the *Escherichia coli* BL21 (DE3) strain. The cloned FLNa24 was expressed as fusion proteins with 26 kDa glutathione S-transferase (GST).

Transformants were grown at 37℃ until an OD_{600} of 0.5 and then induced with 0.5 mM IPTG for 24 h at 20℃. Cells harvested by centrifugation with 8,000 rpm at 4℃ for 15 min. The bacterial cell pellet was suspended in 200 ml of the lysis buffer (1X PBS, 1 mM EDTA with 5% Glycerol). The bacterial lysis was performed by sonication (4 × 30-s pulses) until the lysate became clear. After sonication, the supernatant was applied a GST column, GSTPrep^{TM} FF 16/10 (GE Healthcare Bio-Science) equilibrated with 1X PBS. The proteins was eluted with 80 ml elution buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 10 mM Glutathione reduced) at the flow rate of 1 ml/min. Protein samples were analyzed by SDS-PAGE.

The concentration of proteins was measured by Bradford assay and the thrombin proteases was added (20 units of thrombin per mg of protein) and incubated for 20 h at 20℃ to cleave the GST tag off. After cleavage, the protein was dialyzed into 1X PBS, 1 mM EDTA, 1 mM DDT, 1 mM PMSF, 0.5 mM

sodium azide, and then the GST tag was separated from target protein using GST column, Glutathione Sepharose™ 4 Fast Flow (GE Healthcare Bio-Science). The purified fractions were pooled and dialyzed into 50 mM sodium phosphate, pH 8.0, containing 1 mM EDTA, 1 mM DDT, 1 mM PMSF and 0.5 mM sodium azide. The dialyzed sample was applied to anion exchange column, HiPrep™ 16/10 Q FF (GE Healthcare Bio-Science) and eluted with a linear gradient from 0 to 500 mM NaCl in 50 mM sodium phosphate, pH 8.0. The purified fractions were analyzed by SDS-PAGE and concentrated by ultrafiltration using Amicon® Ultra (Millipore). The concentrated sample was applied to gel filtration on Superdex™ 75 10/300 (GE healthcare Bio-Science) with final buffer (50 mM Tris-HCl, pH 8.0, 100 mM NaCl, 1 mM EDTA) (Figure 3).

After purification of FLNa24, matrix-assisted laser desorption / ionization-time of flight (MALDI-TOF) mass spectroscopy was performed to identify status of samples and denature of samples (Figure 4).

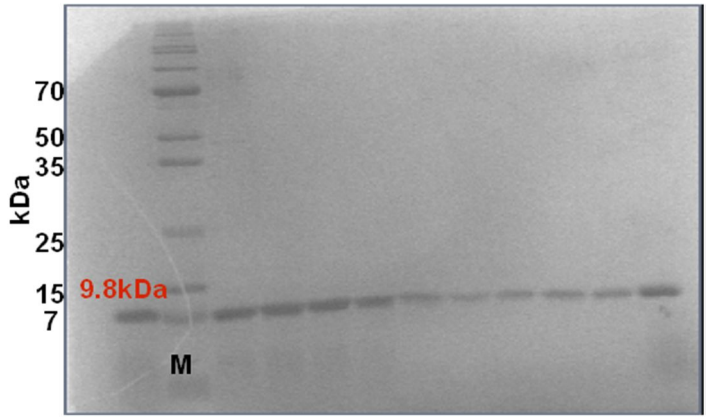

Figure 3. FLNa24 applied size exclusion column. SDS-PAGE shows purification on size exclusion column. Lane M, molecular-mass marker. Other lanes, elution profiles of FLNa24 from size exclusion.

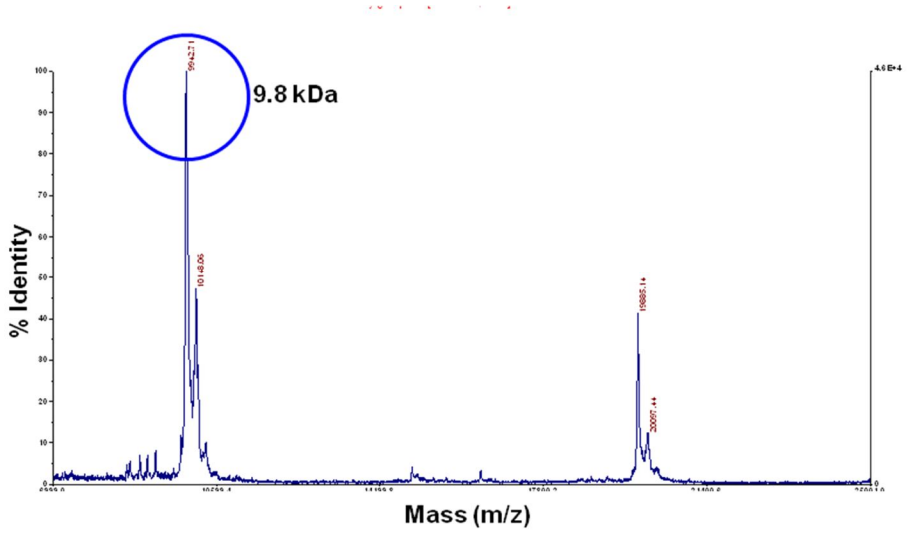

Figure 4. MALDI-TOF mass spectroscopy measurement of purified FLNa24.

1.2.2 Crystallization

Crystals of purified FLNa24 were grown by hanging drop vapor diffusion method at 20℃ after mixing equal volumes (2 ㎕) of protein solution (30 mg/ml in final buffer) and the reservoir solution. Block shaped crystals appeared in 3-5 days in optimized reservoir solution consisting of 10 mM Tris, pH 8.2, 30% PEG 3350 and 200 mM lithium sulfate (Figure 5). For data collection, crystals were equilibrated in 20% glycerol cryo-protective solutions containing reservoir buffer and flash frozen in liquid nitrogen.

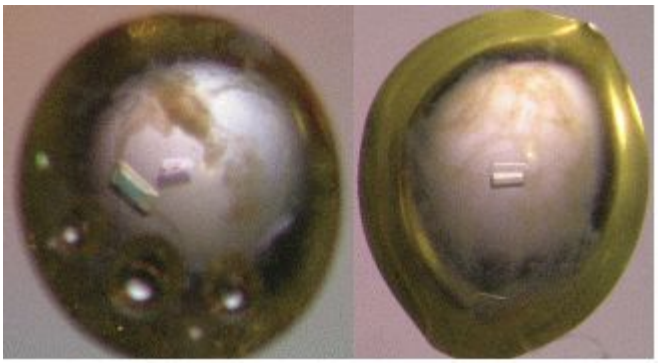

Figure 5. Photomicrograph of FLNa24 crystals.

1.2.3 Data collection and structure determination

Diffraction data were acquired at 100K on the beam-line 4A at the Pohang Light Source (PLS) in Korea. Diffraction images were taken at 18 oscillation for 360° and 180 image frames were used for data processing. The data set was processed using HKL2000[14] and the CCP4 suite of programs[15]. Structure determination was performed by molecular replacement using the previously known human FLNc24 structure as a search model (PDB ID: 1V05)[13] with the program Phaser[16]. CHAINSAW program[17] in the CCP4 was used to mutate the search model into FLNa24 PDB file. Structure refinement and rebuilding were performed by the programs CNS[18], Refmac5[19] in CCP4, and Coot[20]. Refinement was performed using all reflections in 50-1.65Å resolution shell leaving 5% of reflections for R_{free} calculation. Rigid body refinement yielded R_{work} and R_{free} 45.3 and 44.7%, respectively. Refinement of structure was done using simulated annealing, individual B-factor refinement, and minimization implementing bulk solvent correction on all stages. Noncrystallographic symmetry (NCS) was implemented over two monomers with the restraint weight of 300 kcal/mole/Å^2. Model adjustment and refinement were done iteratively, and R_{work} and R_{free} dropped to 30.3 and 31.9%, respectively. At this stage water molecules were picked and model was refined, and R_{work} and R_{free} dropped to 24.5 and

27.3%, respectively. Two sulfate ions (from the crystallization condition) were placed in extra electron density, and model was adjusted and refined. After this stage NCS restraint was released and refinement was done, yielding Rwork and Rfree values of 21.4 and 23.8%, respectively. Model was adjusted and refined further in CCP4 using Refmac5. Finally, Refmac refinement with TLS refinement[21] resulted in R_{work} and R_{free} values of 19.2 and 21.9%, respectively.

The final model had R_{work} and R_{free} of 19.2 and 21.9%, respectively. The data collection and refinement statistics are summarized in Table 1. The final structure was verified using PROCHECK[22] and all the residues were in the most favorable region, additionally allowed region, or generously allowed region of the Ramachandran plot. The atomic structure and structure factor were deposited with the PDB ID 3CNK.

Table 1. Data collection and Refinement Statistics

Data collection	
Beamline	PLS BL4A
Wavelength (Å)	1.0
Space group	$P2_12_12$
Cell dimension	
a, b, c (Å)	57.75, 94.57, 40.95
α, β, γ (°)	90, 90, 90
Resolution (Å)	50−1.65
R_{merge}[a] (%)	5.4 (33.2)[b]
I/σ (I)	31.7 (3.6)[b]
Redundancy[c]	6.2 (5.7)[b]
Completeness (%)	97.5 (99.3)[b]
Unique reflections	27127 (2703)[b]
Refinement	
R_{work}[d] (%)	19.2
R_{free}[e] (%)	21.9
No. atoms	1548
Proteins	1364
Water	184
Ligand	−
B factor (Å^2)	
Protein	16.4
Water	30.8
Sulfate ion	36.4
RMSD[f]	
Bond lengths (Å)	0.010
Bond angles (°)	1.3
Ramachandran plot (%)	
Preferred region	90.4
Allowed region	9.6
Disallowed region	0.0
PDB accession code	3CNK

[a] $\Sigma | I_j - \langle I \rangle | / \Sigma I$

[b] The values in parentheses indicate the highest resolution shell.

[c] N_{obs}/N_{unique}.

[d] $\Sigma_{hkl} | | F_{obs} | - k | F_{calc} | | / \Sigma_{hkl} | F_{obs} |$

[e] R_{free} was calculated by the same way as R_{work}, but with the 5% of the reflections excluded from the refinement.

[f] Root mean square deviation (RMSD) was calculated with REFMAC.

1.3 Results and Discussion

1.3.1 Crystal structure of FLNa24

Overall structure of the FLNa24 exists as an asymmetric dimer, which is in the asymmetric unit and the C_α carbons superimpose with Root Mean Square Deviation (RMSD) of each subunit is 0.3Å. And the structure of FLNa24 forms immunoglobulin-like (Ig-like) fold, which consists of a β-sandwich of seven or more strands in two sheets with a Greek-key topology. Structure of the FLNa24 is consisting of eight β-strands and an 3_{10}-helix (Figure 6). Strands A, B, E and D form a one β-sheet and strand C, F and G form another β-sheet. The β-strands correspond to residues V2561-K2563, S2576-D2581, L2590-H2595, E2603-G2610, L2613-L2619, G2624-W2632, E2635-H2636 and Y2642-V2646 and an 3_{10}-helix correspond to residues L2565-L2567. Secondary structure assignment was performed by DSSP program[23].

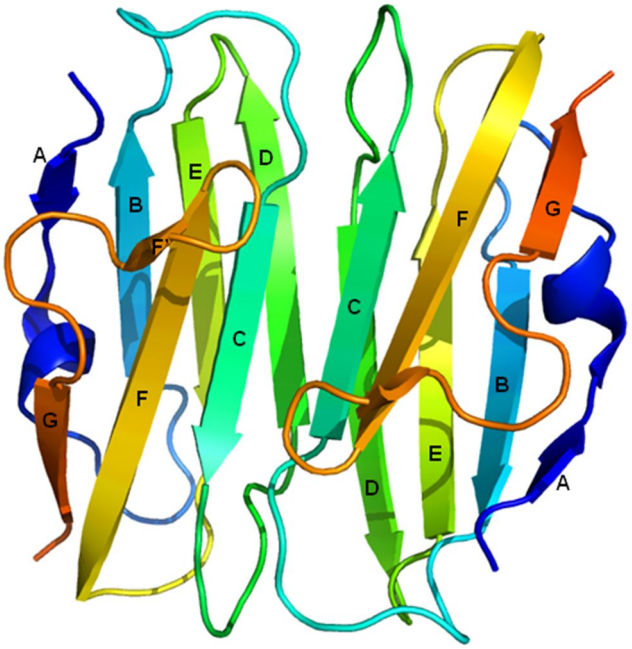

Figure 6. Overall structure of domain 24 of FLNa. Cartoon representation of FLNa24. FLNa24 was consisting of eight strands and a short helix. Figures were generated using PyMOL.

1.3.2 Comparison with other Ig-like domains

As it can be predicted from high sequence with identity other isoforms (69.7 and 67.4% for FLNb and FLNc, respectively), the structure of FLNa24 is very similar to that of FLNb24 (PDB ID: 2EED) and FLNc24 (PDB ID: 1V05)[13] (Figure 7). The RMSD between the dimers of FLNa24 and FLNb24 is 1.2 Å, and FLNc24 is 0.7 Å. The RMSD value between FLNb24 and FLNc24 is 1.2 Å with 70.5%% sequence identity. Because the structure of FLNc24 has the small RMSD with FLNa24 and the paper about the structure of FLNb24 was not published yet, only PDB deposited, the structure of FLNc24 was used to compare with that of FLNc24.

The difference in secondary structure between FLNa24 and FLNc24 lies only in the length of strand D. Strand D of FLNa24 has eight amino acids (E2603-G2610) and its length of amino acid is two longer than that of FLNc24. Considering the small RMSD between the structure of FLNa24 and that of FLNc24, FLNa does not form dimer with other domains of fimamin isoforms while FLNb and FLNc may form heterodimers[24]. In the E strand, that forms a sheet together with the B- and D-strands, FLNb and FLNc show a significantly higher mutual identity than FLNa (Figure 7). Thus, six out of 15 residues that

comprise strand E and its flanking loops are only shared by FLNb and FLNc. This dissimilarity may be causative for the inability of FLNa to heterodimerize with other filamin isoforms.

When the structure of FLNa24 is compared with those of other Ig-like domains of filamin, they show high structural similarity (Table 2). These similarities were derived from not only high sequence identity among them but domains within Ig-like fold-family share the similar structure. All of repeat domains of filamin form Ig-like fold, abbreviated to Ig-like domain. Ig-like domains are one of the most typical protein modules found in animals, existing in a variety of different proteins. These domains are often participated in interactions, commonly with other Ig-like domains via their β-sheets[25,26]. Domains within this fold-family share the similar structure, but can diverge with respect to their sequence. Based on sequence, Ig-like domains can be classified as V-set domains (antibody variable domain-like), C1-set domains (antibody constant domain-like), C2-set domains, and I-set domains (antibody intermediate domain-like). Proteins can contain more than one of these types of Ig-like domains. For example, in the human T-cell receptor antigen CD2, domain 1 (D1) is a V-set domain, while domain 2 (D2) is a C2-set domain, both domains having the same Ig-like fold[27]. Domains with an Ig-like fold can be found in many, diverse proteins in addition to immunoglobulin

molecules. For example, Ig-like domains occur in several different types of receptors (such as various T-cell antigen receptors), several cell adhesion molecules, MHC class I and II antigens, as well as the hemolymph protein hemolin, and the muscle proteins titin, telokin and twitchin.

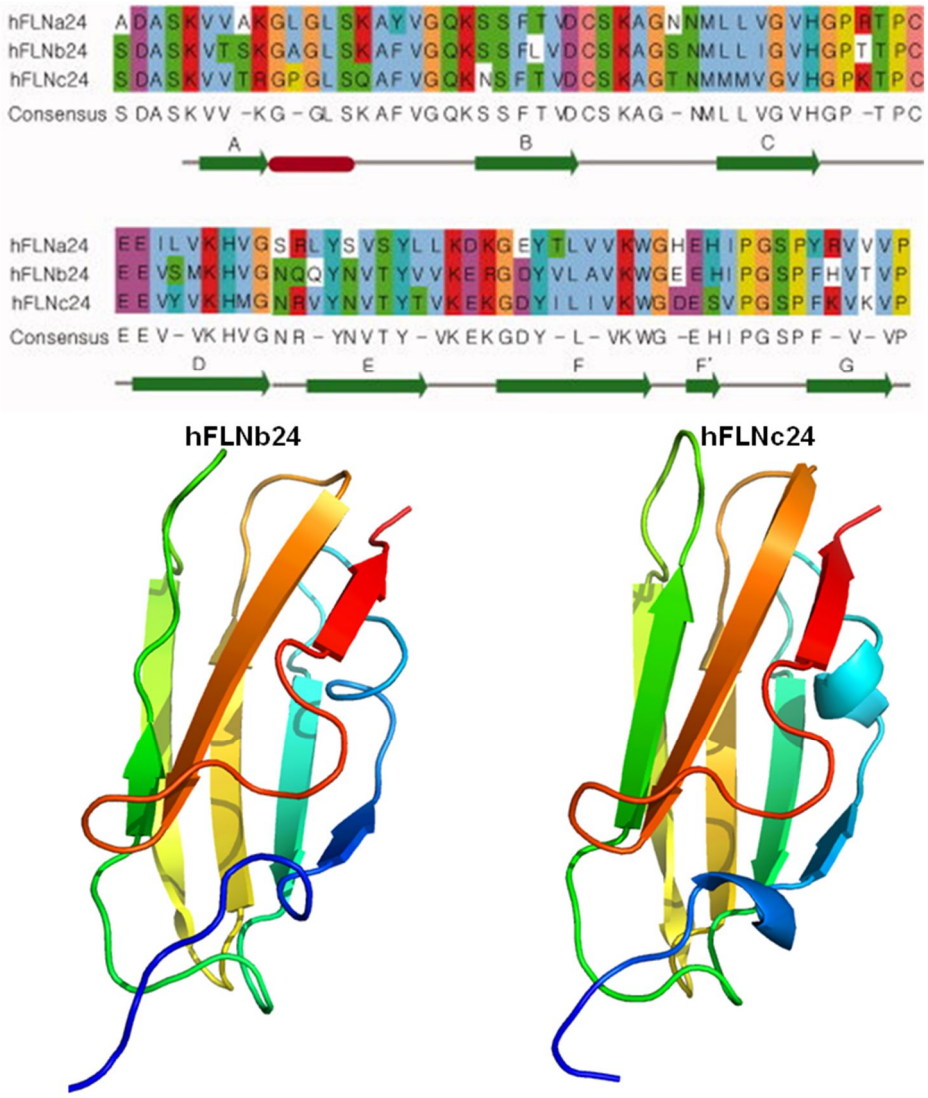

Figure 7. Comparison of dimerization domains of filamins. Multiple sequence alignment of domain 24 of human filamin A, B and C (hFLNa24, hFLNb24 and hFLNc24). Sequence alignment was generated by CLUSTALW and was colored by default according to the ClustalX coloring scheme[28,29] and Overall structure of hFLNb24 and hFLNc24.

Table 2. Root Mean Square Deviation (RMSD) Values for the Structures Superposed on Domain 24 of FLNa

Domain	RMSD (Å)	PDB ID
Gelation factor domain 4	1.3	1WLH
Gelation factor domain 5	1.4	1QFH
Gelation factor domain 6	1.5	1QFH
Filamin A domain 17	1.1	2BP3
Filamin A domain 19	1.1	2J3S
Filamin A domain 21	0.9	2BRQ
Filamin B domain 24	1.2	2EED
Filamin C domain 14	1.7	2D7M
Filamin C domain 16	2.7	2D7N
Filamin C domain 17	2.4	2D7O
Filamin C domain 22	2.1	2D7P
Filamin C domain 23	1.0	2NQC
Filamin C domain 24	0.6	1V05

RMSD values are obtained using secondary-structure matching (SSM) superpose tool in the program Coot[30]. For RMSD comparison, one monomer (chain A) was used if coordinates contain a dimer and in the case of NMR ensembles, the best representing structure was used.

1.3.3 Dimerization interface of FLNa24

The dimerization interface between two monomers in crystal structure of FLNa24 is consisted by mainly strands C and D, as shown in structure of FLNb24 and FLNc24. It is very similar to the dimerization interface of FLNc24 in its size, the characteristics of amino acid residues of these regions, and dimerization pattern[13]. The interface buries 18% (977Å) of the solvent accessible surface of the monomer. The interface (N2857-G2610) contains 42% of polar amino acids and strand D region contains more polar residues than stran C region does. And the dimerization interface is highly conserved in hydrophobic residues, L2590, L2591, V2592 and G2593 of C strands and I2604, L2605, V2606 and V2609 of D strands makes hydrophobic contacts between the two monomers (Figure 8).

Putative hydrogen bond formation between the two C strands and that between the two D strands has distinctive characteristic (Figure 9). Strand Cs does not form direct hydrogen bonding between two monomers but for water molecules are involved in hydrogen bonding network. In stands D, there are six hydrogen bondings between the main chains of each subunit and five indirect hydrogen bondings via water molecules, connect the two D strands. It is similar to that of FLNc24. It is interesting that dimerization of domain 24 is

formed by antiparallel association of each domain. It is conjectured that the hinge region, connecting the FLNa23 and FLNa24, gives room for the domain 24 region to be flexible so that the dimerization domains form antiparallel structure while other domains of whole fimain dimer run parallel along the filamin rod axis.

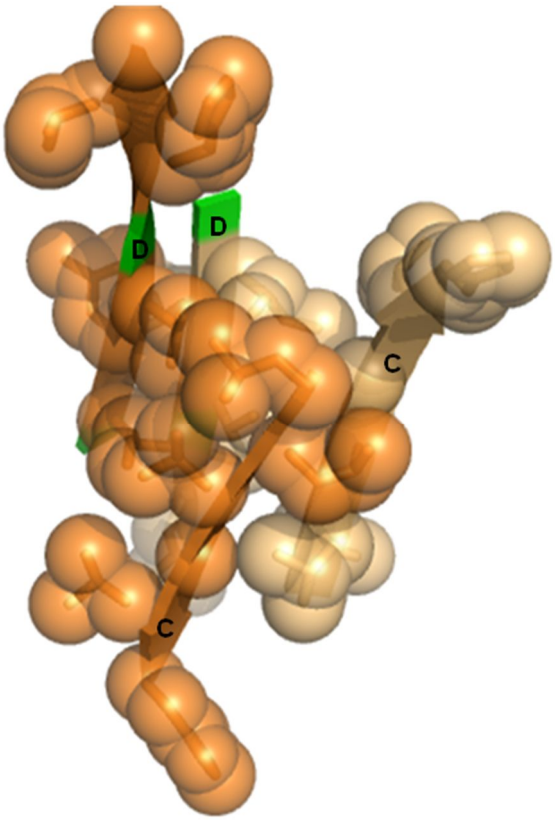

Figure 8. Hydrophobic contacts in dimerization interface. Hydrophobic side chains are depicted as spheres and are colored orange for subunit A and light orange for subunit B.

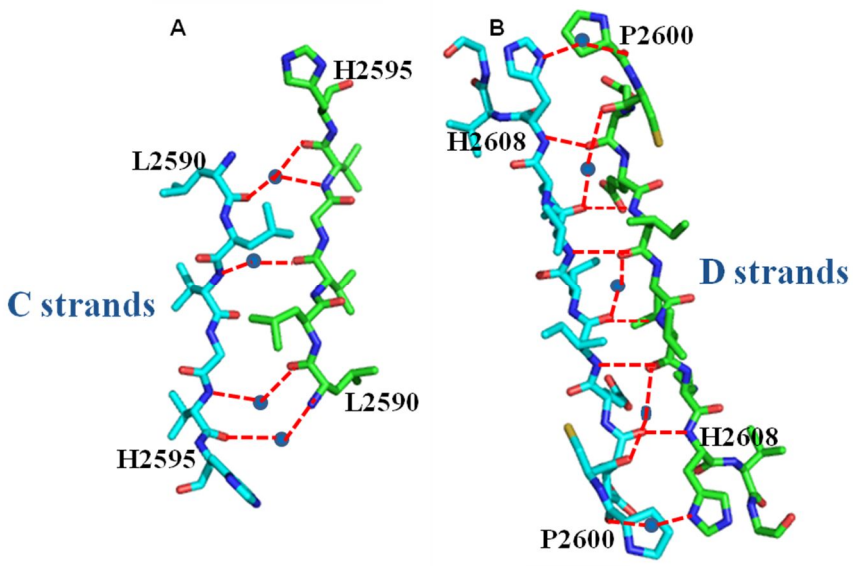

Figure 9. Dimerization interface of domain 24 of FLNa. A: Hydrophobic interaction and putative hydrogen bonding between the two C strands. Four water molecules are involved in the hydrogen bonding network. B: Putative hydrogen bonding network between the two D strands. There are six hydrogen bondings between protein residues, and additionally five water molecules are involved in the hydrogen bonding network.

1.4.4 Possibility of interaction with binding partner

The interaction of filamin with other proteins is regulated by different mechanism such as receptor occupancy, phosphorylation and proteolysis. The potential phosphorylation sites were reported as the conserved three residues, (Ser, Phe)−Pro−(Phe, Tyr, Thr)[31]. The putative phosphorylation site, S640−P2641−Y2642 is also conserved in carboxy terminus of FLNa24. Different binding partners interact with domain 24 or tandem domains including domain 24 of filamin. Domain 21−24 of filamin can interact with various proteins[32]. FLNa24 can interact with the small GTPase RalA[33] and Granzyme B[34]. In the human FLNc, FLNc24 alone is enough and sufficient for dimerization[31]. In the structure of FLNa24, strands C and D are forming dimerization interface as previously shown in that of FLNc24. Interestingly, in the complex structure of FLNa17 and GPIbα, the strands C and D of FLNa17 interact with the GPIbα peptide [12]. These show that domain 24 may have a role in interaction with other binding molecules in addition to the dimerization role.

1.4 Conclusion

We can determine the crystal structure of the dimerization domain of human filamin A. The dimerization of FLNa24 results in the V-shape dimer of whole filamin A, the molecular basis for its actin cross-linking activity. At the view point of structure of FLNa24, FLNa24 has immunoglobulin-like fold, consisting of β-sandwich. The structure of FLNa24 was similar with structure of other Ig-like domains, especially other dimerization domain of filamin isoform, which was the main character of Ig-like fold. Strands C and D were major dimerization interface of FLNa24. Well conserved hydrophobic residues in strands C and D, which compose the compact hydrophobic cluster, and hydrogen bonding networks in strands C and D results in dimer of FLNa24. FLNa24 has putative phophorylation site (S2640-P2641-Y2642) in C-terminus. FLNa24 may have a role in interaction with other binding molecules in addition to the dimerization role.

Chapter 2. Structural insight into the allosteric conformational change of CRP from *Escherichia coli*

2.1 Introduction

Adenosine 3', 5'-Cyclic monophosphate (cAMP) receptor protein (CRP; also referred to as CAP, catabolite gene activator protein; the official gene name is crp) from *Escherichia coli* has long served as a typical textbook example describing transcription regulation, DNA-binding motif, and allosteric activation of a protein[35,36]. CRP plays a critical role in regulating the transcription of more than 200 genes by binding to specific DNA sites and interacting with RNA polymerase[37]. The activated CRP upon binding of the effector molecule cAMP binds to target DNA sites located in or adjacent to the promoter region. The CRP binding to DNA results in the DNA bending and concomitant recruitment of RNA polymerase (RNAP) via the interaction with CRP, to initiate the gene transcription[38-41].

After the isolation of CRP in the early 1970s, more than ten 3-dimentional structures have been determined. The first crystal structure of CRP was determined in complex with cAMP in 1981[42], and afterward was refined to 2.1 Å in 2000[43]. The

structure of cAMP-CRP-DNA complex was determined in 1991[44] and the structure of cAMP-CRP-DNA-RNA Polymerase complex was determined in 2002[45]. In these crystal structures, CRP is a homodimeric protein with each subunit consisting of the C-terminal DNA-binding domain (CDD), the N-terminal nucleotide-binding domain (NND), and a short flexible stretch connecting the two domains, called an interdomain 'hinge' region. Particularly, the C-terminal F-helix (αF) of CRP, which forms a typical helix-turn-helix motif together with the neighboring E-helix (αE), is responsible for the specific DNA recognition[38,44,46,47]. The helix-turn-helix motif is highly conserved in many prokaryotic DNA-binding proteins. The F-helix is oriented to fit into the major groove and the E-helix is oriented to fit into the minor groove of the specific DNA site. However, it doesn't operate in the absence of bound cAMP, of which binding pocket is located apart from the DNA-binding region. Thus, the activation process of CRP by cAMP binding was a classic example of allosteric conformational change. In addition, recent advances in protein dynamics have contributed to elucidation of its dynamic allosterism.

Many biochemical and biophysical studies have shown that allosteric change of CRP was occurred by the binding of cAMP. The binding of cAMP to CRP causes CRP to adopt a

conformation that exhibits a high affinity for specific DNA sequences[48-51]. For example, the cAMP-CRP is easily cleaved by proteases, whereas apo-CRP is resistant to proteolysis[49,52]. In the presence of 5,5' -dithiobis (2-nitrobenzoic acid), the cAMP-CRP forms an intersubunit disulfide bond, whereas apo-CRP does not[48]. Small-angle X-ray scattering and thermodynamic study on CRP also showed that the binding of cAMP to CRP induces the allosteric conformational change[53,54]. The binding of cAMP to the N-terminal domain of CRP allosterically enhances the binding affinity of the C-terminal domain of CRP to the specific DNA sequence, and *vice versa*[55,56]. The cAMP binding induces a small change in the secondary structure elements in either the N- or C- terminal domain of CRP except the length of C- and D-helices, but results in the exposure of F-helix to solvent to facilitate binding of CRP to specific DNA[57-59]. Although the binding affinity of cGMP to CRP is comparable to that of cAMP to CRP, cGMP does not induce high affinity of CRP for specific DNA sequence[60-62]. The maximum two cGMP can bind to CRP while maximum four cAMP can bind to CRP, two in cyclic nucleotide binding domain (CBD) in *anti* form, two in DNA binding domain (DBD) in *syn* from[38,63,64]. In NMR study, the chemical shift perturbations induced by cGMP binding were comparable to but smaller than those by cAMP binding[64]. From many mutational

studies on cAMP and cGMP binding site in CRP, several significant results could be obtained. Mutations on Glu72 and Arg82 residues that interact with phosphate or sugar of cAMP and cGMP, reduce the binding affinity[65-67].

However, structural basis of the transition is essential for thorough understanding of the CRP allostery and it necessarily requires atomic details of both the inactive- and active-state conformations. In this regard, more than 10 three-dimensional coordinates available in the Protein Data Bank provide full pictures of the active-state CRP conformations. Then, the structures of ligand-free (apo-), inactive CRP structure has been independently solved by the X-ray crystallography, albeit at a low resolution (3.6 Å) or with a mutant (D138L) form[68], and by NMR in solution[69]. Unfortunately, however, the NMR and crystal structures were distinct from each other, thereby arguing controversial details of the conformational allosterism.

In the present study, we succeeded in obtaining higher-resolution (2.2 Å) crystal structure of the wild-type apo-CRP, which would be the most proper template for precise inspection of inactive conformation and establishes here the mode of conformational transition in atomic detail. In addition, we report the first crystal structure of another inactive form, the cGMP-CRP complex, at also 2.2 Å resolution, which not only support our insight into the conformational allosterism, but also address

how CRP discriminates the false ligand cGMP from its authentic effector cAMP. And also the conformational allostery change of CRP by cyclic nucleotide and structural difference between apo-CRP, cAMP-CRP and cGMP-CRP will be discussed.

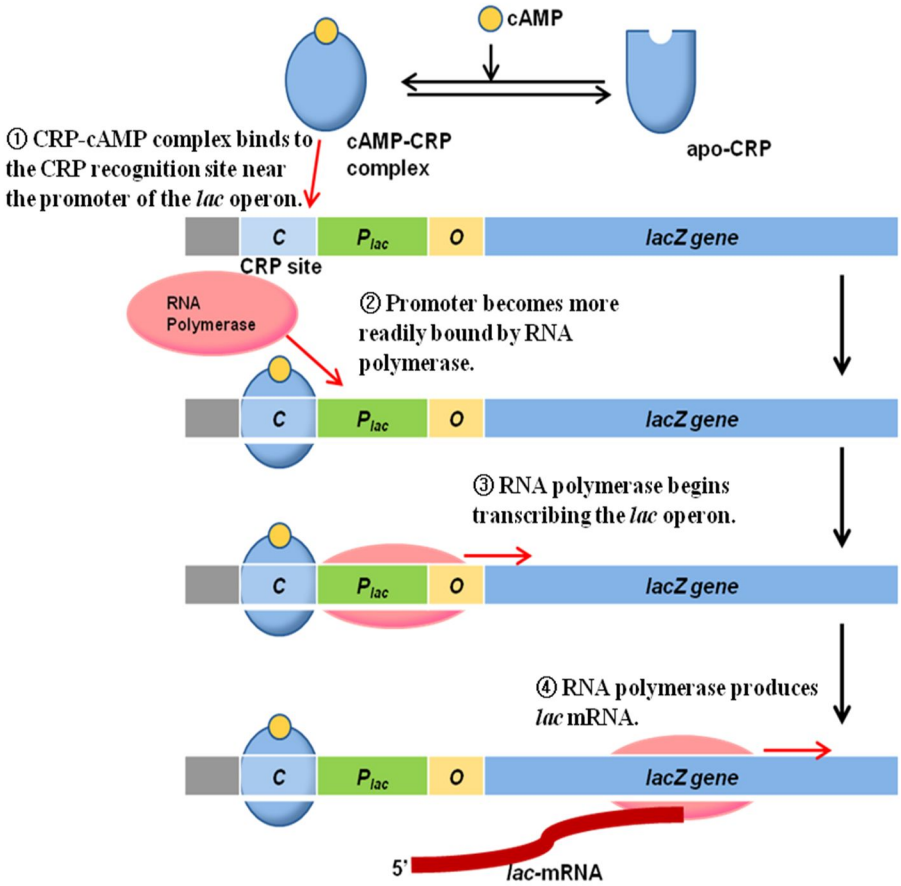

Figure 10. Transcriptional regulatory of CRP. CRP–cAMP complex binds to the CRP recognition site near the promoter of the *lac* open. Then, promoter becomes readily bound by RNA polymerase. RNA polymerase begins transcribing the *lac* operon and produces *lac* mRNA.

2.2 Experiment Procedure

2.2.1 Cloning and protein preparation

The full-length *Escherichia coli* CRP, which was used for the crystallization of the cGMP-CRP complex, was prepared using the recombinant pT7-CRP plasmid.

For the crystallization of the apo-CRP containing an N-terminal eight-residue deletion, DNA fragments encoding P9-R209 were amplified using pT7-CRP as a template and the following pairs of oligonucleotide primers (*Nde*I and *Xho*I restriction sites are shown in bold): the forward 5' -G GAA TTC **CAT ATG** CCG ACT CTC GAA TGG TTC-3' and the reverse 5' -CCG CCG **CTC GAG** TTA ACG AGT GCC GTA AAC GAC-3' primer, which contained a stop codon to produce the protein without artificial histidine tags. The PCR products were purified, digested with *Nde*I and *Xho*I enzymes (NEB) and ligated into an *Nde*I/*Xho*I-digested pET-21a(+) (Novagen) expression vector. The *E. coli* BL21 (DE3) (Novagen) cells that were transformed with the constructed plasmids were grown at 37° C in Luria Broth (LB) media including 100 μg/ml ampicillin. Protein expression was induced at an OD_{600} of approximately 0.5 by adding isopropyl-β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM, and induction was continued for a further 4 h.

To prepare selenomethionine-labeled CRP(9-209), SeMet-CRP(9-209), the cells were grown at 25° C in M9 media containing 100 mg/l Lys, Phe and Thr; 50 mg/l Ile, Leu, and Val; and 60 mg/l selenomethionine until an OD_{600} of 0.5[70]. Protein expression was induced using IPTG, and the cultures were incubated for 12 h.

Cells were harvested by centrifugation and disrupted by sonication. After centrifugation, the supernatant was loaded onto a Bio-Rex 70 (Hercules) cation-exchange column. The protein was eluted with a linear gradient from 0 to 1 M NaCl in 50 mM potassium phosphate buffer containing 1 mM EDTA at pH 6.7. Fractions containing CRP were pooled and concentrated to approximately 2 ml by ultrafiltration using an Amicon® Ultra (Millipore). The concentrated sample was applied to a SuperdexTM 75 (GE Healthcare Bio-Sciences) size-exclusion column that had been equilibrated with the final buffer (20 mM Tris, pH 7.2, 150 mM NaCl, and 1 mM EDTA). The purified protein was concentrated to approximately 10 mg/ml in the final buffer. For the crystallization of cGMP-CRP, the protein solution contained 10 mM cGMP.

2.2.2 Crystallization and data collection

Crystals were grown by the hanging drop vapor diffusion method at 20° C after mixing equal volumes (2 ul) of protein solution (10 mg/ml in the final buffer) and the reservoir solution. Diffraction-quality apo-CRP(9-209) crystals were produced in 3-6 days under the optimal condition of 30% (w/v) polyethylene glycol (PEG) 400 and 100 mM CHES, pH 7.5. Under these conditions, crystals of SeMet-CRP(9-209) were well reproduced. Well-diffracting crystals of the cGMP-CRP complex appeared in 10-12 days under the condition of 22% (w/v) PEG 8000, 15.4% (v/v) glycerol, 100 mM proline, 50 mM HEPES, and 88 mM potassium phosphate, pH 6.4. Prior to data collection, single crystals were equilibrated in the crystallization solution containing additional 5% glycerol and flash-frozen in liquid nitrogen. Diffraction data of the SeMet-CRP(9-209) and cGMP-CRP crystals were collected using the ADSC Quantum 210 CCD detector on the AR-NW12A beamline at the Photon Factory (PF; Tsukuba, Japan). The data sets were processed and scaled using the program HKL2000[14]. Crystals of SeMet-CRP(9-209) and cGMP-CRP belonged to the space groups $P6_1$ and $P2_1$, respectively (Table 3).

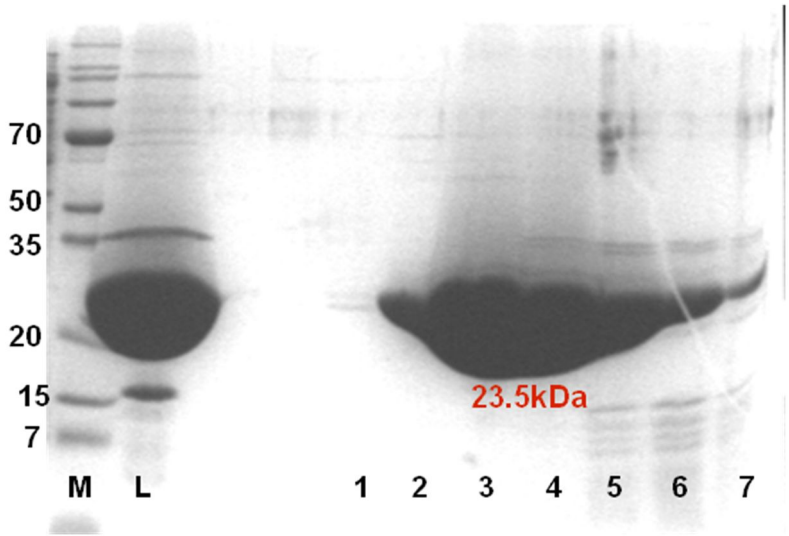

Figure 11. SDS-PAGE showing purification on an size exchange column. Lane L, CRP just before loading onto the size exchage column; lane 1-7, elution profile of CRP, Fractions 2 to 4 were used for protein crystallization trials. Lane *M*, molecular-mass marker.

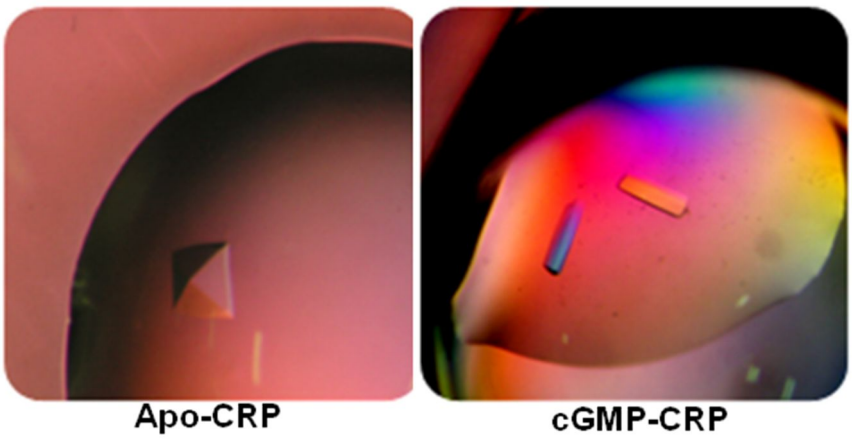

Figure 12. Photomicrograph of CRP crystals. Crystal of apo-CRP has octahedral shape (left) and crystal of cGMP-CRP has block shape (right).

Table 3. Data collection and Refinement Statistics

	Apo-CRP	cGMP-CRP
Data collection		
Beamline	PF-AR NW12A	PF-AR NW12A
Wavelength (Å)	0.97922	1.00000
Space group	$P6_1$	$P2_1$
Cell dimension		
a, b, c (Å)	120.617, 120.617, 60.349	42.568, 56.171, 186.718
α, β, γ (°)	90, 90, 120	90, 90.18, 90
Resolution (Å)	50-2.2	50-2.2
R_{merge}[a] (%)	8.6 (39.4)[b]	6.6 (14.7)[b]
I/σ (I)	94.4 (12.3)[b]	32.8 (12.7)[b]
Redundancy[c]	22.5 (21.9)[b]	6.9 (6.9)[b]
Completeness (%)	99.7 (100)[b]	99.5 (100)[b]
Unique reflections	25544 (1262)[b]	45063 (4489)[b]
Refinement		
R_{work}[d] (%)	20.2	17.0
R_{free}[e] (%)	26.3	20.8
No. atoms	3269	6752
Proteins	3156	6392
Water	113	268
Ligand	-	92
B factor (Å^2)		
Protein	50.26	33.80
Water	52.85	36.78
RMSD[f]		
Bond lengths (Å)	0.019	0.026
Bond angles (°)	2.432	2.116
MolProbity (%)[g]	**2.59**	**2.17**
Ramachandran plot (%)		
Preferred region	95.75	95.50
Allowed region	4.00	4.13
Disallowed region	0.25	0.37
PDB accession code	4N9H	4N9I

[a] $\Sigma|I_j-\langle I\rangle|/\Sigma I$

[b] The values in parentheses indicate the highest resolution shell

[c] N_{obs}/N_{unique}.

[d] $\Sigma_{hkl}||F_{obs}| - k|F_{calc}||/\Sigma_{hkl}|F_{obs}|$

[e] R_{free} was calculated by the same way as R_{work}, but with the 5% of the reflections excluded from the refinement.

[f] Root mean square deviation (RMSD) was calculated with REFMAC.

[g] MolProbity analysis[71]

2.2.3 Structure determination and refinement

The apo-CRP structure was first modeled by MR-SAD (molecular replacement with single anomalous dispersion) using the program PHENIX[72]. The N-terminus domain (residues 9-130) of the cAMP-bound CRP structure (PDB ID: 1G6N)[43] was used as a search model for molecular replacement. Matthews coefficient (V_M) was calculated to be 2.64 $Å^3$ Da^{-1}, and a solvent content 50.6%[73] when two molecules were assumed in asymmetric unit. Together with the solution of molecular replacement, SAD phasing was done by AutoSol of PHENIX using anomalous coefficient f' =8 and f'' =2. The phases were further improved by density modification and model was refined using the program AutoBuild of PHENIX. The Model was improved further by interactive modeling with Coot[20] and refined using REFMAC[19] in CCP4 suite[74] and PHENIX. R_{free} was calculated by randomly setting aside 5% of the data[75]. After iterative refinement R and R_{free} reached 20.2% and 26.3%, respectively (Table 3).

To determine the cGMP-CRP structure, molecular replacement was initially performed using the data set with 75% completeness and the N-terminal domain (residues 8-130) of the low-resolution apo-CRP structure (PDB ID: 3HIF)[68] as the initial search model. The Matthews coefficient was 2.39 $Å^3$ Da^{-1} (48.6% solvent content) with four molecules in the asymmetric

unit. Refinement was performed using Coot and REFMAC in the CCP4 suite, including bulk solvent correction. The bound cGMP was included at the final stage of the refinement, and R and R_{free} reached 19.2 and 25.0%, respectively. Subsequently, the N-terminal domain of the refined structure was alternatively used as a search model for the new cGMP-CRP data set with higher completeness for improved structure determination (Table 3). Model building and refinement was iteratively performed, and the cGMP molecule was included in the positive electron density during the final refinement. The coordinate file for cGMP was obtained from the HIC-Up server[76], whereas the library for REFMAC was generated using the PRODRG server[77]. Finally, the structure was refined to the final R and Rfree of 17.0 and 20.8%, respectively (Table 3).

2.3 Results and Discussion

2.3.1 Crystallization and the Validation of secondary structure of apo-CRP

In all of the known structures of CRP, including NMR data in solution, the N-terminal eight residues (V1-D8) were observed to be highly disordered and flexible[59]. Thus, we could obtain a high-quality crystal with an improved resolution up to 2.2 Å, by deleting the flexible N-terminus and prepared the construct encompassing P9 to R209 of CRP. The X-ray crystal structure of apo-CRP(9-209) was solved using the single-wavelength anomalous dispersion method in the space group $P6_1$ at 2.2 Å resolution. The crystals contained one dimer in the asymmetric unit and except for the Q170 and the C-terminus (T208 and R209), wich were traced only in the backbone, and all the other parts were clearly refined.

The overall fold is similar to the previously published crystal structure of unliganded-D138L mutant (PDB ID: 3FWE)[68]. Each subunit has the larger N-terminal nucleotide-binding domain (NND, residues P10-Q125/E129), formed by eight strands and four helices, the smaller C-terminal DNA-binding domain (CDD, residues L134-R209) containing four strands and three helices, and the hinge region (residues Y126/K130-N133) which connects the C-helix of the NND and the D-helix

of the CDD (Figure 13). For convenience in comparison, the strand numbers and helix letters in Figure 13 were assigned to match the standard nomenclature used in the cAMP-bound CRP structure (PDB ID: 1G6N)[43]. The secondary structure elements of the apo-CRP structure are defined almost identically to those observed n cGMP-CRP. The β2 (residues T28-I30) and β7 (D68-I70), which forms interstrand hydrogen bonds to the antiparallel β7 and β3, respectively, were also regarded as β-strands, similar to the cAMP-CRP, although they are not designeated as regular β-strands in the apo-CRP structure by the standard DSSP criteria[23] (Figure 15). The structure of apo-CRP contains one additional short helix in the N-terminal domain (N, residues E72-E76), which has not been identified in the cAMP-CRP structure, was designated as hN (nascent helix). Our previous NMR analysis has predicted a short helical conformation around the E72 residue in apo-CRP, which directly contact to cAMP in the cAMP-CRP complex. In the present crystal structure, the presence of a tightly ordered 310-helix is evident in both subunits of apo-CRP, with consecutive CO(i)-HN(i+3) hydrogen bonds (Figure 16). A similar helix is well conserved in other cyclic nucleotide binding proteins that share homologous folds of the CRP NND and also observed in the *Mycobacterium tuberculosis* CRP (*Mtb*-CRP). In addition, the ηN has been similarly detected in the crystal

structure of the D138L-CRP and the NMR structure of apo-CRP, but with a relatively loosened conformation due to less number of hydrogen bonds and/or an alternative CO(i)-HN(i+4) interaction (Figure 16). Then, the ηN is not stabilized in the cAMP-CRP complex due to the insufficiency of tight hydrogen bonds (Figure 16), which is attributable to a slight perturbation in the backbone geometry that is caused by cAMP binding.

The D-helix (αD) in the NMR structure of apo-CRP determined by Popovych *et al*[69], begins from the T140, which is consistent with the cAMP-bound crystal structures of CRP[43,47]. However, this finding conflicted with our previous NMR analysis of apo-CRP[59], which predicted the L134 residue as the N-terminal end of the αD. The present crystal structure confirms that the αD helix of apo-CRP includes L134 to K152, which is consistent with the D138L-CRP crystal structure[68] and our previous NMR analysis[69]. The length of αD has been suggested to be crucial for the allosteric transition of CRP

The most distinguishing feature of the present apo-CRP structure is the heterogeneity in the lengths observed for the C-helices (αCs), which encompass D111-V126 in one subunit (chain A; open subunit) and further lengthened up to K130 in the other subunit (chain B; closed subunit). The shorter αC is consistent with that determined by NMR in solution[59,69], whereas the longer αC is in good agreement with that

observed in the D138L-CRP crystal structure [68]. However, further examination revealed that the αCs are stabilized as a regular α-helix up to Q125 in the open subunit, whereas they are elongated up to E129 in the closed subunit. The upper regions, *i.e.*, Q125-E129 of the open subunit and E129-V131 of the closed subunit, adopt a helical loop conformation that resembles a 3_{10}-helix (Figure 16). Subsequently, they commonly fold into a portion of a regular α-helix up to F136 in the cAMP-CRP complex. Collectively, the interdomain hinge region of apo-CRP is defined as V126-N133, and it is suggested that the helical conformation in the V126-K130 region of apo-CRP would be transiently adopted or dynamically equilibrated, considering the flexible nature of the hinge[69,78,79,80].

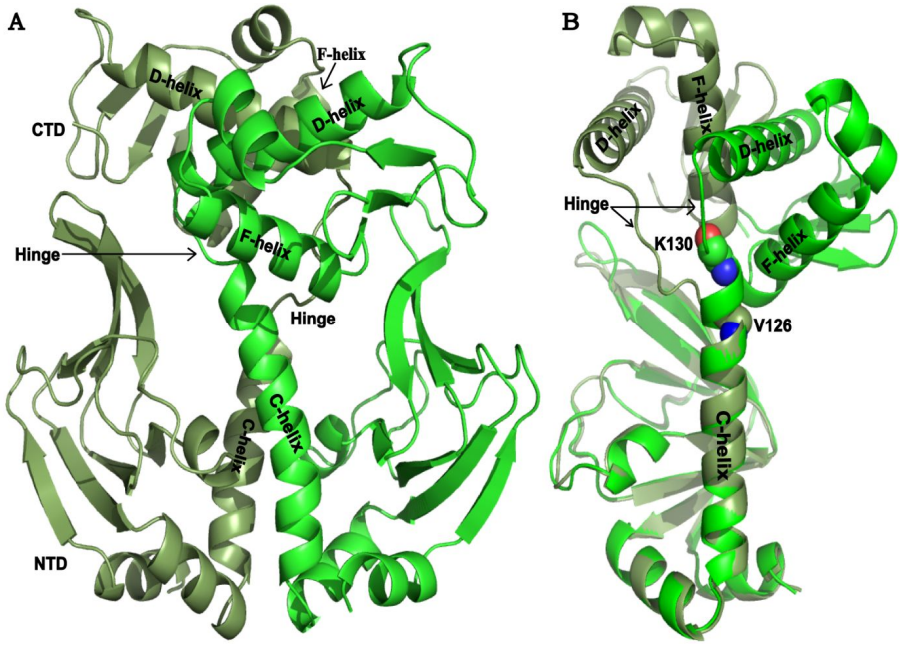

Figure 13. Overall structure of apo-CRP. The structure of apo-CRP is asymmetric dimer consist of N-terminal domain and C-terminal domain (A). The N-terminal domains are very similar between two subunits, while the relative orientation of C-terminal domains and the length of C-helices and hinge regions are different between two subunits (B).

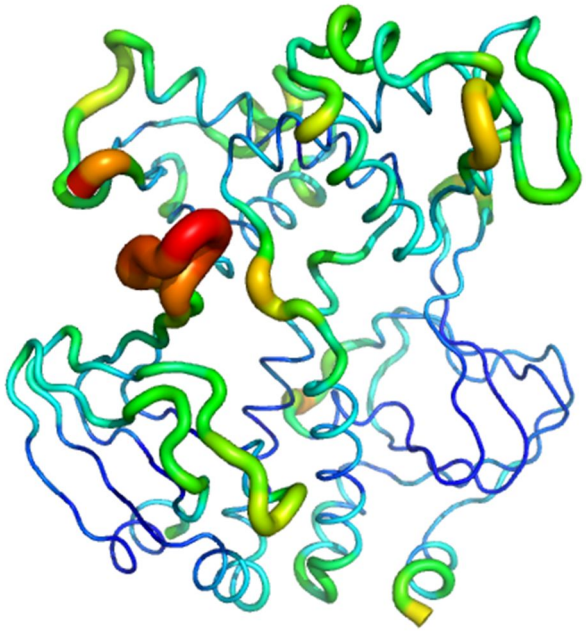

Figure 14. B-factor presentation of apo-CRP.

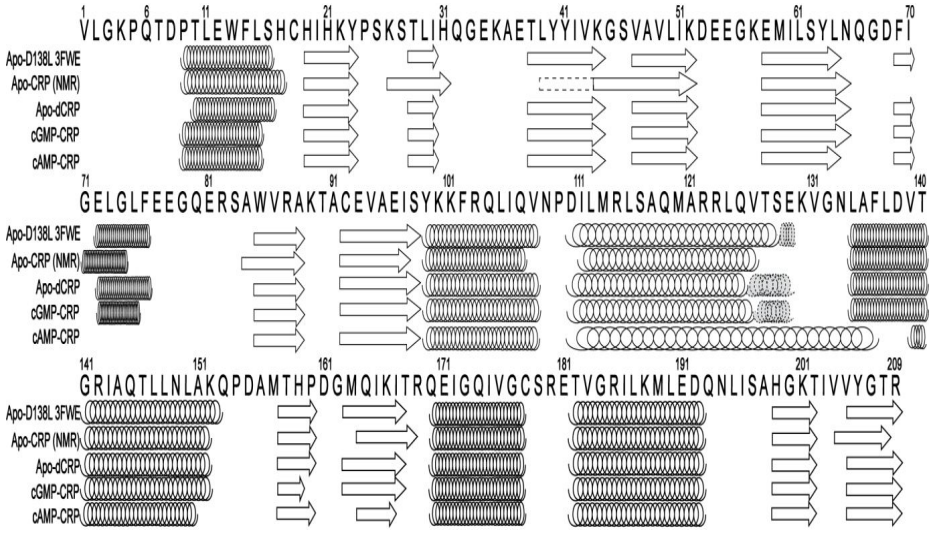

Figure 15. Secondary structure assignment of CRP strcuture.
Secondary structure assignment was performed by DSSP program[23].

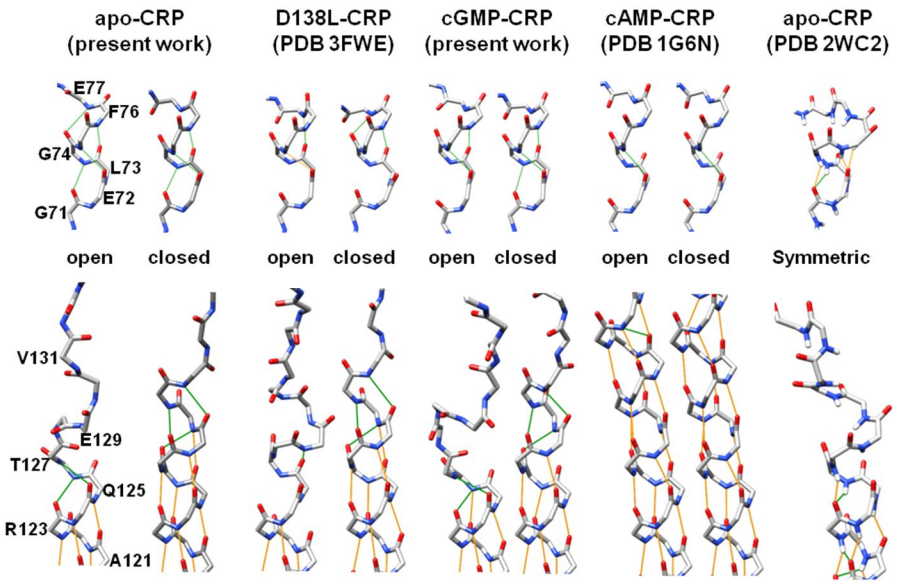

Figure 16. Backbone hydrogen bonding patterns of hN (upper panels) and αC (bottom panels) in individual subunits. The C^{α} positions are labeled with corresponding residues in the most left figures and the others are shown from the same N-terminus (G71 in upper panels and A121 in bottom panels). Hydrogen bonds are depicted as green lines for CO(i)-HN(i+3) and orange lines for CO(i)-HN(i+4) connections.

2.3.2 Dimerization of the apo-CRP structure

Consistent with that previously noted in the D138L-CRP structure, the most distinguished feature of the apo-CRP structure distinct from the cAMP-bound form is the compact dimerization of CDDs (L134-Y206) result in the inward positioning of αFs (Figure 13). The local folds of the two CDDs are almost identical (Figure 17) and they associate in a two-fold symmetry with an antiparallel coiled-coil assembly of αDs providing a fundamental dimer interface. And the strong hydrophobic clusters formed at the ends of the coiled-coil, which stabilize the CDD dimerization (Figure 18A). Each hydrophobic core is constructed as a mutual packing of the two CDDs with the L134 and L137 residues in a subunit and A144, L147, L148, L190, L195, I196, and V205 residues in the opposite subunit. Involvement of the αF residue L190 in this hydrophobic cluster contributes to constraining the inactive orientation of the DNA-binding αF toward the NND. The increased conformational stability of D138L-CRP is also attributable to strengthening this hydrophobic cluster, by supplementing a neighboring hydrophobic sidechain, additionally to the innate L134 and L137 residues (Figure 18A). In particular, the L134 residue occupies the core of the cluster, surrounded by the other eight hydrophobic sidechains. Seizing the L134 into this hydrophobic cluster appears as the structural

basis that retains the long αD of apo-CRP starting from the L134 that forms a backbone hydrogen bond to D138. Upon the cAMP binding to CRP, the L134 sidechains are positioned at the newly generated hydrophobic core at the C-termini of αCs (Figure 18B), with breaking the backbone hydrogen bond to D138 and newly forming the hydrogen bond with K130, leading to the C-terminal elongation of αC and the N-terminal diminution of αD. Similarly, in the known structure of a CRP-family protein CooA[80], the conserved leucine (130 in CooA, 134 in CRP) has been suggested to occupy the central fulcrum for the conformational switching of the hinge conformation. With the L134 ousted, the aforementioned CDD hydrophobic clusters are not integrated and the CDD dimer dissociates. Even at the altered location by cAMP binding, the L134 residues play a central role bridging the two hydrophobic clusters of individual NNDs at their C-termini (Figure 18B). Hence, the functional switch of L134 appears as one of central regulatory machineries in CRP.

Some CRP-family proteins, which resemble the cAMP-bound CRP with dissociated CDD conformation exert an inherent activity with no effector molecules[81]. Conversely, the dimerized CDDs can be a characteristic of inactive CRP species. The CDD dimerization in the inactive, apo-state has been first suggested for the CRP-family protein CooA[80]. Finally, the previous

D138L-CRP and the present *E. coli* apo-CRP and cGMP-CRP structures commonly evidenced the CDD dimerization in the inactive states. From the previously study of *Mtb*-apo-CRP crystal structure[82] and the *Ec*-apo-CRP NMR structure[69], significant sequence variations from *Ec*-CRP was identified. First, the L137 residue in *Ec*-CRP corresponds to T144 in *Mtb*-CRP (Figure 19B). Since the hydrophobic L137, together with L134, plays a critical role in the CDD dimerization of apo-form (Figure 18A), its mutation to the polar amino acid threonine could trigger the CDD dissociation. This possibility is supported by a similar observation of a constitutively active CRP* phenotype by a mutation of the hydrophobic A144 residue at the CDD dimerization interface to threonine. In addition, *Ec*-CRP and *Mtb*-CRP possess significant sequence variations particularly in the flap region (Figure 19B). First, the stabilization of dimerized CDD orientation by the flap residue E54 in *Ec*-apo-CRP, which directly interacts with an αF residue (T182 or R185), is not possible in the *Mtb*-apo-CRP, where the corresponding residue is proline (Figure 19B). In the same way, the interdomain hydrophobic contact between M59 (N67 in *Mtb*-CRP) to the L195 (W203 in *Mtb*-CRP) in *Ec*-apo-CRP is not expected for the polar asparagine at the equivalent position of *Mtb*-CRP. On the contrary, the positive-charged flap residues K54 and K56 of *Mtb*-CRP, which are

replaced respectively by S46 and A48 *E. coli* CRP (Figure 19B), stabilize the dissociated CDDs in *Mtb*-apo-CRP, by contacting to the D174 (K166 in *Ec*-CRP) and E179 (E171 in *Ec*-CRP) residues, respectively. Similary, the *Mtb*-apo-CRP R59 sidechain, which hydrogen bonds to the D140' (N133' in *Ec*-CRP) to position the flap close to the hinge, is mutated to an isoleucine in *Ec*-CRP not to induce that moving. All the take together, the sequence variations between *Ec*- and *Mtb*-CRP appear to have specifically adapted to stabilize the dimerized CDDs in *Ec*-apo-CRP, whereas the dissociated CDDs in *Mtb*-apo-CRP.

Some CRP-family proteins, such as SdrP and TTHB099, which function with no effector molecsles, inherently adopt the CDD-dissociated conformation like the cAMP-bound CRP. Furthermore, the intermediate conformations of CRP-family proteins, observed for CooA, CprK, and PrfA, wherein the αFs are positioned incompatible with DNA binding despite the dissociated CDDs, also support the concept that CRP can adopt an alternatively inactive conformation even with the CDD dissociation, upon a minute structural impact at certain positions.

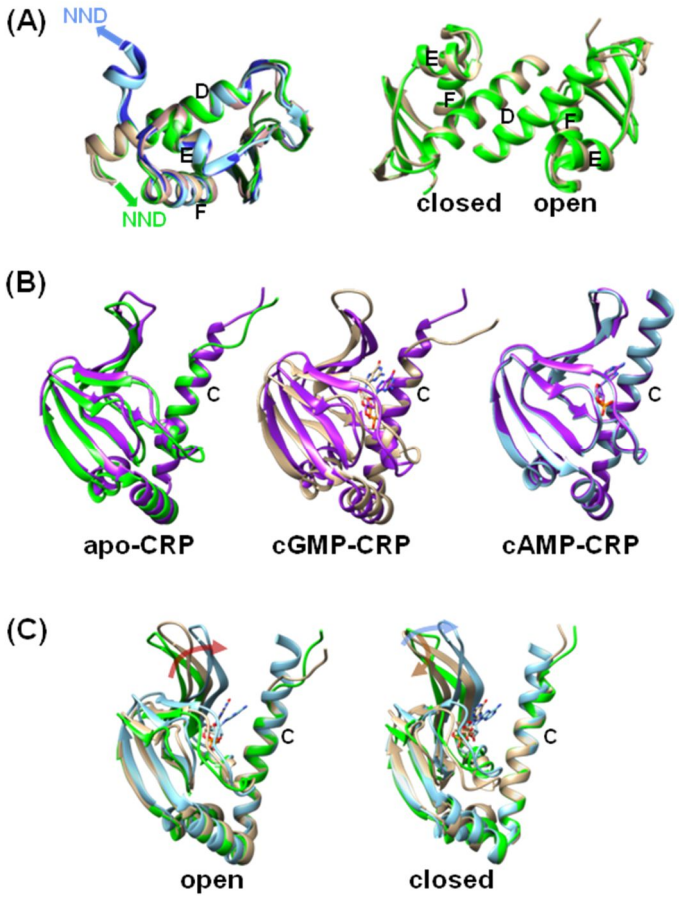

Figure 17. Structural comparison of individual domains. Ribbon colors are basically *green*, *sky blue*, and *tan* for apo−CRP, cAMP−CRP, and cGMP−CRP, respectively. Cyclic nucleotides are shown in a stick presentation. The C−, D−, E−, and F−helix elements are labeled with corresponding letters. (A) In the left panel, all of individual CDDs in the three species of CRP structures are superimposed. CDDs in closed subunits are colored *forest green*, *cornflower blue*, and *orchid* for apo−CRP, cAMP−CRP, and cGMP−CRP, respectively. The N−terminal linkage to NND is depicted by a *green* arrow for apo− and cGMP−CRP and a *sky blue* arrow for cAMP−CRP. In the right panel, CDD dimers of apo− and cGMP−CRP are superimposed by matching C^{a} atoms in the closed−subunit CDDs. (B) Individual NNDs in open and closed (*purple*) subunits are superimposed by matching C^{a} atoms in aCs, for each CRP species. (C) Individual open− and closed−subunit NNDs in the three CRP species are superimposed, respectively, by matching C^{a} atoms in aCs.

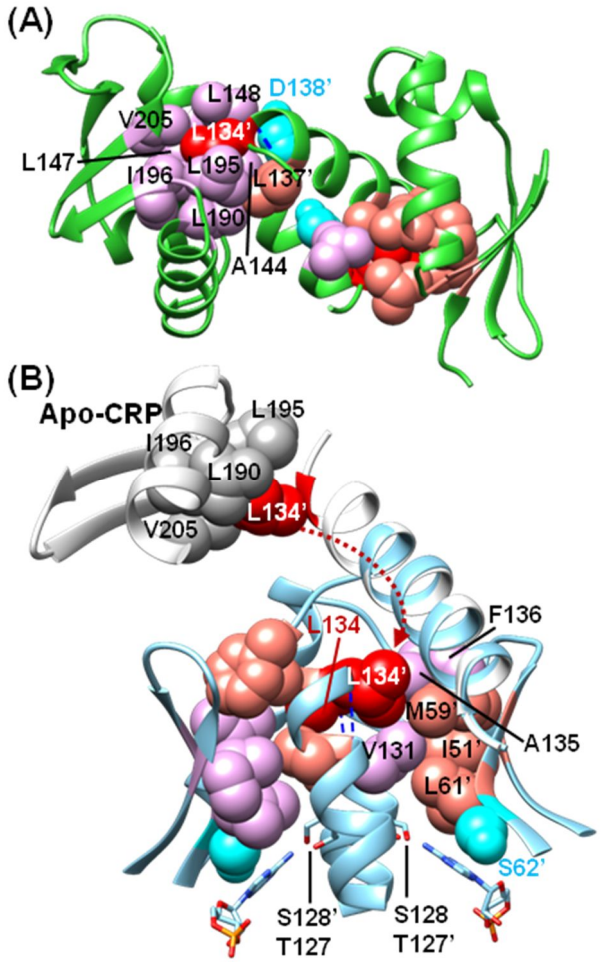

Figure 18. Hydrophobic clusters differentiating CRP conformations. The closed-subunit residues are labeled with apostrophes. Hydrophobic sidechains are presented as spheres and colored *red* for L134 residues, *salmon* or *purple* for open subunit and *orchid* or *magenta* for closed subunit. The L134-mediated backbone hydrogen bonds are shown by *blue*, dashed lines. The L134 residues occupy the center of hydrophobic cluster stabilizing the CDD dimer in apo-CRP. (B) In cAMP-CRP, the L134 residues are captured into the hydrophobic clusters formed at the C-termini of aCs and couple the two NND clusters.

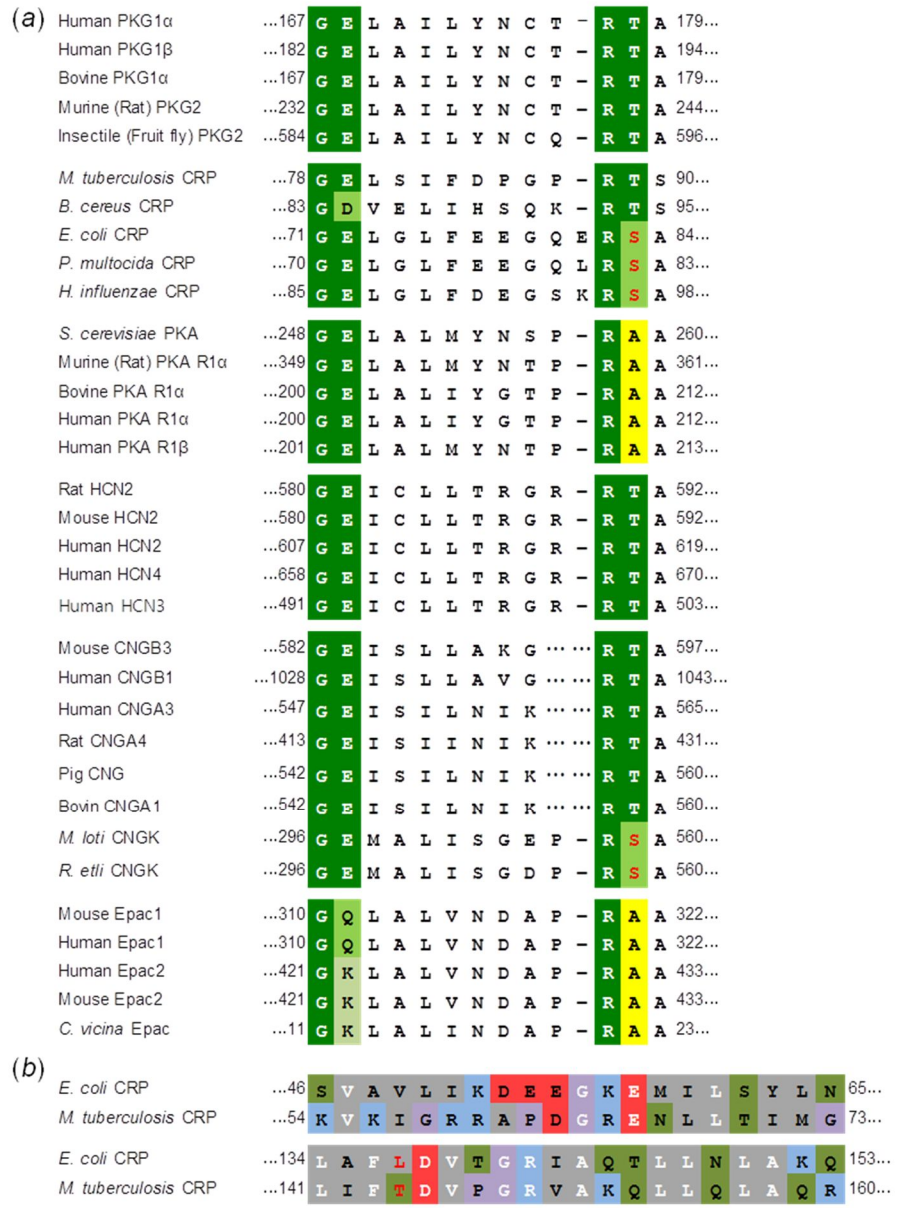

Figure 19. Sequence comparison of CRP with other cNMP-regulated proteins. (A) A part of cNMP-binding b-roll sequence is aligned to show the positional conservation of cNMP-interacting residues, which are color boxed. (B) Amino acid sequences of E. coli CRP and M. tuberculosis CRP are aligned at the flap (upper panel) and aD (bottom panel) regions. Exact matching is indicated by white letter. Box color is differentiated depending on charge, polarity, and hydrophobicity of each amino acid.

2.3.3 Asymmetry and interdomain interaction of the apo-CRP structure

As well as the CDDs, the two NNDs (P9-V126) also dimerize with a comparable fold in a two-fold symmetric manner (Figure 17). Structural aspects of the NND dimerization are apparently the same as observed in the known CRP structures, fundamentally mediated by the coiled-coil assembly of the αCs. Despite the symmetric dimerization in both the NNDs and CDDs, the overall conformation of the present apo-CRP structure is characterized by an overall asymmetry between subunits. One subunit is in an 'open' conformation with relatively splayed orientation between NND and CCD, while the other adopts a 'closed' form relatively stooping (Figure 13). Thus, the appreciable asymmetry is regarded as the one originated primarily from different conformation of the interdomain hinge (V126-N133). Due to the asymmetry in domain orientations, the interdomain interactions are also differently organized between subunits. For example, the electrostatic interaction between the E54 sidechain of the closed subunit (E54' in Figure 20A; henceforth the residues in closed subunit are denoted with primed numbers) and the R185 guanidino group of open subunit is not relevant between the R185' and the E54. The E129 is involved in the intersubunit interaction stabilizing the cyclic nucleotide binding

pocket of the opposite subunit, by contacting with the R123', whereas the E129' provides an interdomain contact to the F188' in its own subunit (Figure 20). In addition, the intersubunit interaction between S128 and K130' is not formed between S128' and K130 (Figure 20). Likewise, the Q193 sidechain from a water-mediated, indirect hydrogen bonding to the backbone of G132', whereas the G132 and Q193' do not interact (Figure 21). The hydrophobic contact between M59 and L195 is not valid either in the closed subunit. Although those interdomain interactions in apo-CRP are not comparable in detail to those observed in the cAMP-bound CRP, it is a common aspect that the interactions are predominantly mediated via the hinge and the so-called flap (the $\beta 4-\beta 5$ hairpin) regions in NNDs. However, their interaction counterparts in CDDs are critically distinguished; *i.e.*, mainly the αF in apo-CRP (Figure 20), instead of αD and αE in cAMP-CRP, mediates the interdomain interactions, albeit not much abundantly. This inspection supports again that the inward retaining of αFs with a dimerized CDD is responsible for the inactiveness of apo-CRP.

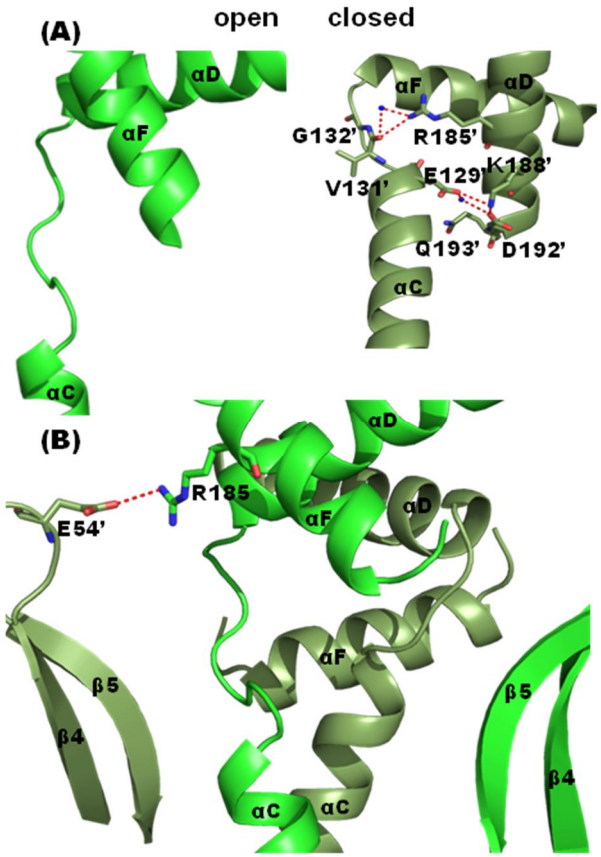

Figure 20. Interdomain interactions between CCD and NND of apo-CRP. Hydrogen bonds and salt bridges are depicted by orange lines. Nitrogen and oxygen atoms are colored in blue and red, respectively. Water molecules and backbone atoms involved in the hydrogen networks are also presented as balls.

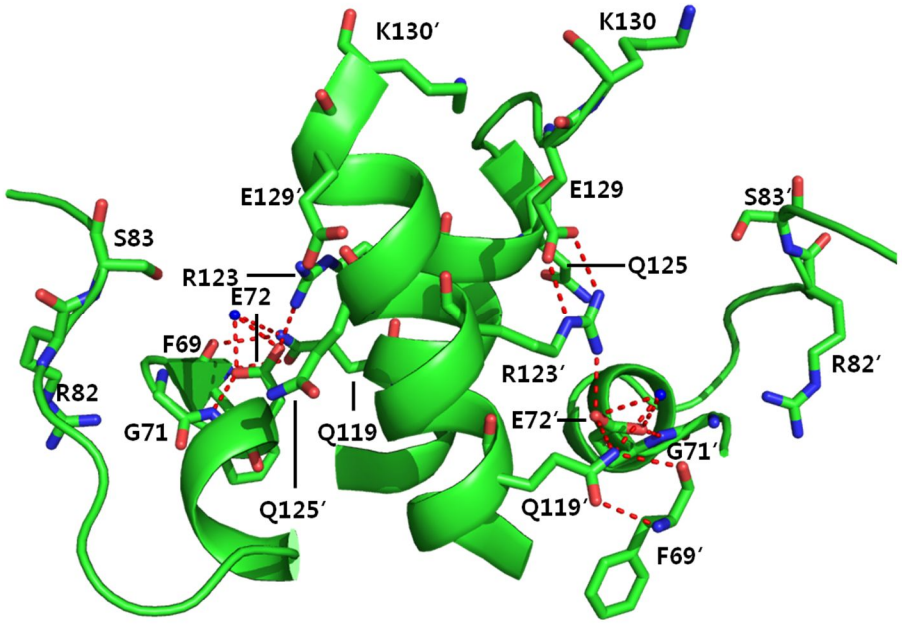

Figure 21. Hydrogen bonding in the cNMP-binding pockets. The residues in closed subunits are denoted with primed numbers. Nitrogen and oxygen atoms are colored in blue and red, respectively. Hydrogen bonds are depicted by red lines. Water molecules and backbone atoms involved in the hydrogen bonding networks are also presented as small balls.

2.3.4 Cystallization and overall conformation of the cGMP-CRP

Although cGMP, which is a homologous structure of cAMP, the binding affinity to CRP is comparable to that of cAMP, cannot induce the activation of CRP[64]. Thus, as a representative cyclic nucleoside monophosphate (cNMP) that cannot activate CRP, cGMP is the most frequently used as a negative control probe for investigating CRP allostery. However, the cGMP binding to some constitutively active CRP mutants (namely CRP*) derives functional regulation of CRP. Thus, the cGMP-CRP complex structure, which has not been available yet, would provide an insight into the structural basis of the inactive state of CRP. In the present work, the crystals of the cGMP-CRP complex were obtained with a full-length protein. The crystal structure of CRP in complex with cGMP (cGMP-CRP) in the space group $P2_1$ was solved using the molecular replacement method at 2.2 Å resolution. Overall structure is shown in Figure 22. The structure of cGMP-CRP complex contains two dimers in asymmetric unit, denoted as chanin A and B and chanin C and D. The conformations of two subunits in the cGMP-CRP structure are different from each other and comparable to apo-CRP structure. The length of C-helix is also different between the two subunits. The r.m.s.d values in the C_α coordinates between subunit A and subunit B is 3.64 Å.

Also, there is the lower r.m.s.d difference between the two N-terminal domains and the two C-terminal domains. The deviation between subunit A of AB dimer and subunit C of the CD dimer is only 0.18 Å (3.62 Å for A versus D) and the deviation between subunit B of the AB dimer and subunit D of the CD dimer is 0.16 Å. The exposed hairpin loops of N-terminal domain (NTD) in subunit B and D (residues G56-I70) show slightly different conformations. It shows some obvious flexibility in NTD hairpin loops. Since these differences do not affect the overall structure and determined conformation, a dimer molecule designated with chains A and B was selected for structural interpretation.

The overall topology of cGMP-CRP is a quite similar with apo-CRP crystal structure. The r.m.s.d difference between one subunit of cGMP-CRP and that of apo-CRP is about 2.0 Å. But the r.m.s.d values in the C_α coordinates between NTD of cGMP-CRP and NTD of apo-dCRP is about 0.79 Å and the r.m.s.d values between CTD of cGMP-CRP and that of apo-CRP is about 0.3 Å.

In the previously NMR study[64], we revealed that a CRP dimer binds two molecules of cGMP as its maxium, whilie four molecules of cAMP. Two cAMPs in *anti*-conformations bound at NNDs are lead to the allosteric activation of CRP and the conformational transition results in two additional binding of

syn-cAMP at CDDs. However, the present structure reveals that the cGMP, in contrast to cAMP, binds in a syn-conformation to the cNMP binding pockets of both subunits, consistent with the previous predictions by NMR[64].

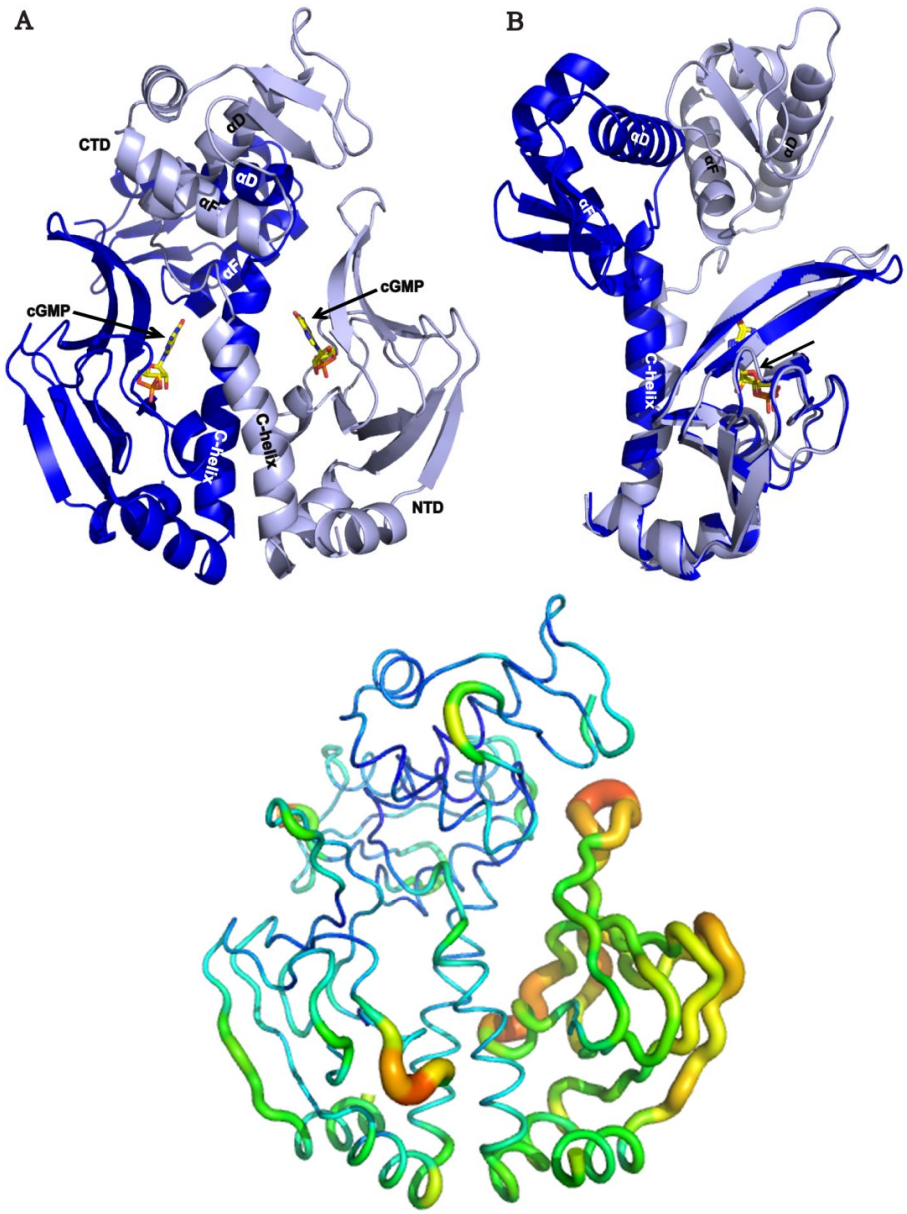

Figure 22. Overall structure of cGMP-CRP and B-factor presentation. A: The overall structure of cGMP-CRP complex. B: comparison of two subunits. B-factor of cGMP was presentation in bottom panels.

2.3.5 A stepwise binding and specific recognition of cGMP

Weber and Steitz have previously supposed that cAMP may first bind to the β-roll residues via the interactions with the phosphoribose moiety, and then form the interaction of adenine moiety[47]. A similar stepwise binding of cAMP has been also suggested from the crystal structures of the cNNP-binding domains in a cAMP-dependent protein kinases (PKAs) and a cAMP-dependent ion channel[83], where the subsequent interactions occur through the phosphate. Interestingly, the present cGMP-CRP structure captured different contacts of the bound cGMP between subunits, which could occur from the stepwise binding. The cGMP binding in the closed subunit is tighter with more direct and indirect contacts to CRP than in the open subunit (Figure 24), in contrast to the cAMP, of which binding interactions were observed almost same in both subunits (Figure 25). Since NMR analyses on the cGMP-bound CRP in solution have evidenced no detectable asymmetry of the complex[64], it is reasonable to suppose that the observed asymmetry reflects an intermediate state of the cGMP-CRP. Then, although the two subunits might be under the fast exchange equilibrium in solution, the crystallographic snapshot of the intermediate state of cGMP-CRP strongly suggests that CRP has two-step manner of cNMP-binding mode, evidencing

the deduction by Weber and Steitz[47].

The minimal binding mode of cGMP observed in the open subunit is assignable to a nascent binding. In detail, first, the cNMPs are taken in the base side of the cNMP-binding pocket by the direct interactions of the common phosphoribose moiety with the conserved β-roll residues G71, E72, R82 and S83 (Figure 24). This initial binding mode is almost identical between the cAMP and cGMP binding and the electrostatic interaction between the sidechains of E72 and R123 are also conserved even in the apo-CRP. However, the purine ring-mediated interactions are distinct different between cAMP and cGMP (Figure 24 and 25). First, the interaction of the hydroxyl group of S83 with the purine ring of a cNMP acts as the structural determinant for discriminating *syn-anti* conformers and specifically stabilizing one of them. As in a *syn*-conformation of cNMP, the N2 amine of the cGMP forms a hydrogen bond to the hydroxyl oxygen of S83. In contrast, the N6 amine of cAMP cannot form the equivalent hydrogen bond, due to long distance away from the S83, even if it adopts a *syn*-conformation. Alternatively, the N3 nitrogen of cAMP, if in a *syn*-conformation, can orient toward the hydroxyl group of S83 in close proximity for hydrogen bonding. However, since the hydroxyl group of S83 acts primarily as a hydrogen donor to the exocyclic phosphate O1 oxygen, the additional hydrogen

bond to the N3 nitrogen cannot be formed. Consequently, the bound cAMP could not be stabilized in a *syn*–conformation by the S83.

It is important to note that the S83 to guanine contact is involved in the nascent binding mode of cGMP (Figure 24), which could enable CRP to selectively recognize it at the nascent binding step. The hydroxyl group is well conserved as either a serine or a threonine in CRP orthologues and as a threonine in PKGs[84] (Figure 18), probably for specific recognition of the *syn*–cGMP. Instead, the hydroxyl oxygen of S83 of CRP also likely stabilizes the *anti*–conformation of cAMP, as it makes a water–mediated contact to the N7 nitrogen of the bound *anti*–cAMP (Figure 25). Thus, considering the racemization in solution, the bound cAMP in CRP is expected to be undergoing a conformational equilibrium between *syn*– and *anti*–conformation, albeit preferring *anti*–conformer, at the nascent binding before the final settlement. Then, the settlement of the bound geometry in *anti*–conformation could be accomplished upon the later contacts mediated by the N6 amino group. The *syn*–conformations of cAMPs bound to PKAs are often observed to be stabilized by the interaction via the N6 amine.

The second step of cNMP binding in CRP forms the tighter bound of cNMP by additional indirect contacts via water

molecules, as observed in the closed subunit of the cGMP-CRP structure and both subunits of cAMP-CRP structure (Figure 24 and 25). For example, the water-mediated, indirect hydrogen bonding between the phosphate moiety of cGMP in closed subunit and the sidechain of Q125 in open subunit (Figure 25) is corresponding with that observed in the cAMP-binding mode (Figure 26). This stabilization occurs that the purine rings of cNMPs contact to the hinge regions of CRP for generating specific structural impacts. In particular, the hydrogen-bonding contacts via the 6-keto group of cGMP are contrasted with those of the N6 amino group of cAMP (Figure 26). The direct interactions of the N6 amine of cAMP with the T127 of its own subunit and S128 from the opposite subunit (Figure 25) play a role as critical trigger for the allosteric transition of CRP. On the contrary, the 6-keto group of the bound *syn*-cGMP makes a direct hydrogen bond to the K130 of opposite subunit (Figure 24). Two water-mediated contacts to CRP via the 6-keto group and N1 nitrogen of cGMP that probably stabilize the bound geometry of guanine ring are also observed.

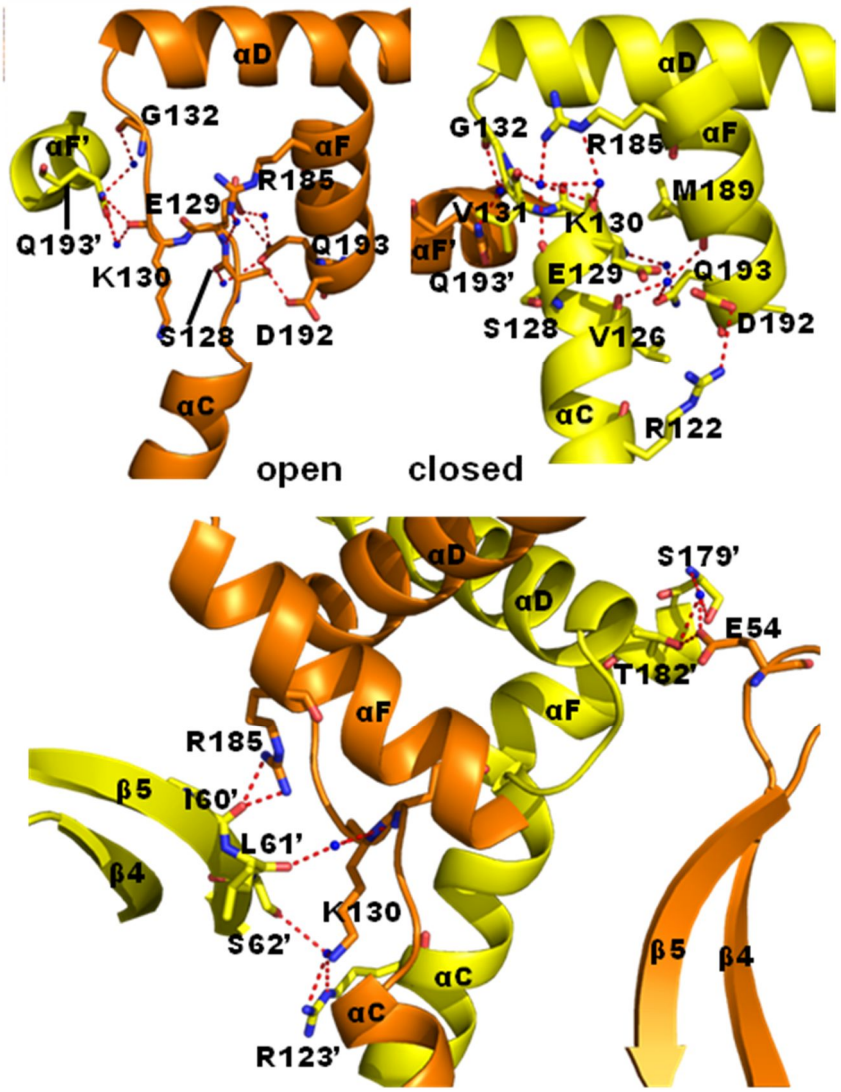

Figure 23. Interdomain interactions between CCD and NND of cGMP–CRP. Hydrogen bonds and salt bridges are depicted by red lines. Nitrogen and oxygen atoms are colored in blue and red, respectively. Water molecules and backbone atoms involved in the hydrogen networks are also presented as balls.

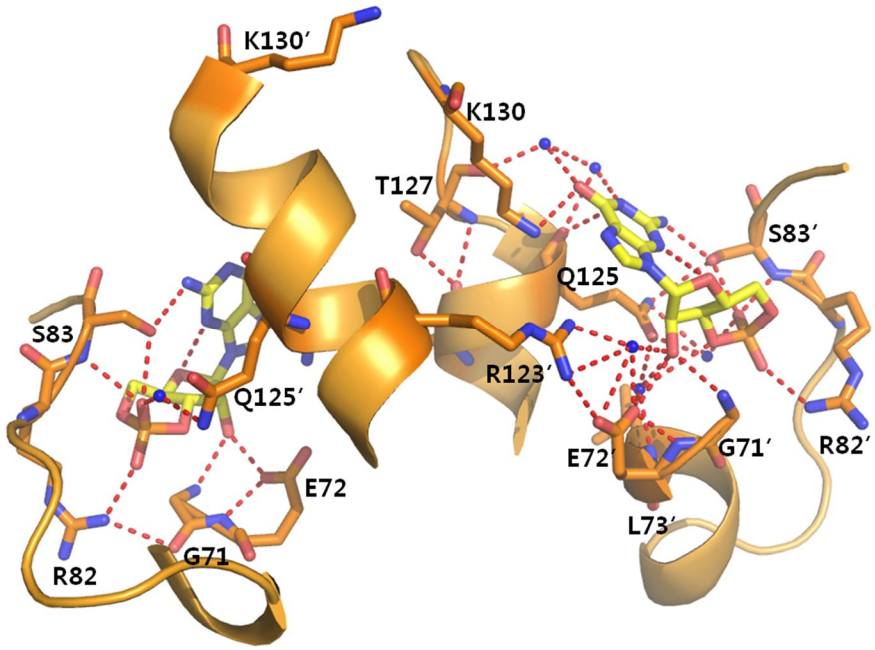

Figure 24. Hydrogen boning in cGMP bindings pocket. The residues in closed subunits are denoted with primed numbers. Nitrogen and oxygen atoms are colored in blue and red, respectively.Hydrogen bonds are depicted by red lines. Water molecules and backbone atoms involved in the hydrogen bonding networks are also presented as small balls

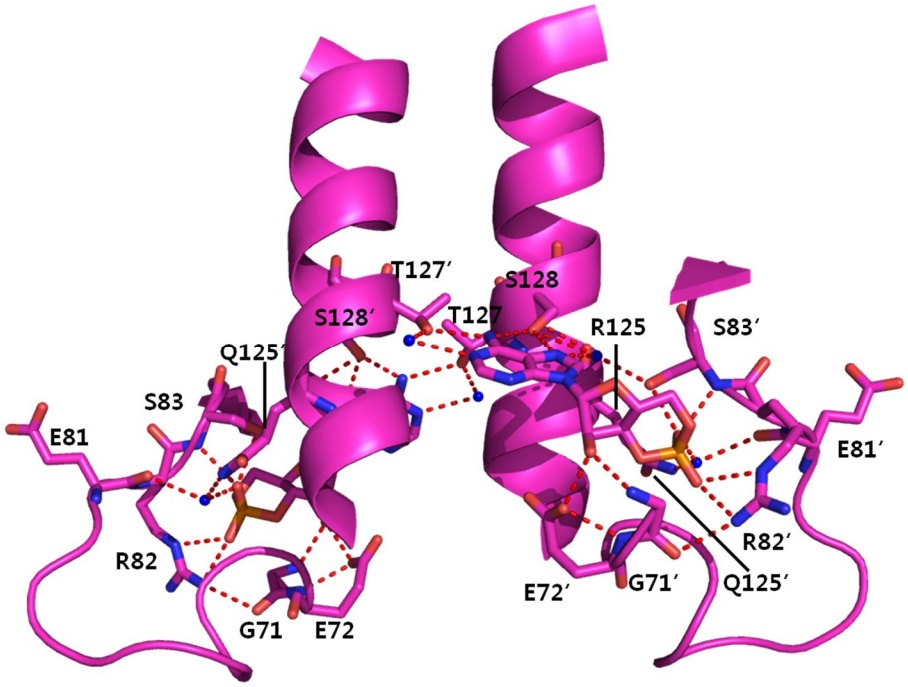

Figure 25. Hydrogen boning in cAMP bindings pocket. The residues in closed subunits are denoted with primed numbers. Nitrogen and oxygen atoms are colored in blue and red, respectively. Hydrogen bonds are depicted by red lines. Water molecules and backbone atoms involved in the hydrogen bonding networks are also presented as small balls

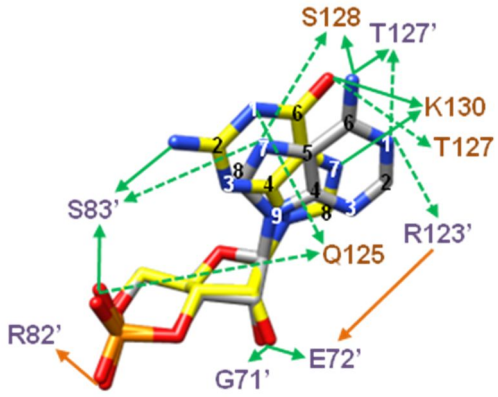

Figure 26. Electrostatic interactions with CRP of the bound cNNP. *anti*−cAMP (*gray*) and *syn*−cGMP (*yellow*) at closed subunit are summarized. Direct electrostatic interactions (hydrogen bond in *green* and salt bridge in *orange*) are illustrated by solid arrows and the water−mediated, indirect interactions by dashed arrows.

2.3.6 Preferences and sequential bindings of cNMPs

Because of the dynamic nature of the hinge region, each subunit of apo-CRP in solution may be equilibrate in conformational exchange between open and closed forms while maintaining the symmetric CDD dimerization. Given that the apo-CRP asymmetry is relevant at the nascent cNMP binding step, the subsequent contact to the K130 of opposite subunit would not be favored for the cGMP bound at the open subunit, due to the reverse direction of the K130 sidechain (Figure 27A, right). On the other hand, the K130 sidechain of open-subunit is located toward the closed subunit of cGMP to readily bind upon a small change of the hinge geometry (Figure 27A, left). The preformed close proximity (*ca.* 4.3 Å) between the 6-ketone of cGMP to the negative-charged E129 sidechain might promote the hinge transformation by inducing a repulsion that flips the E129 sidechain. Therefore, the cNMP-binding pocket in closed subunit is first firmly occupied by cGMP and it fixes the opposite subunit as an open form through the interaction with the K130, while the closed subunit can still equilibrate between open and closed forms. When the first cGMP-bound subunit is in an open state, the second cGMP bound at opposite subunit contacts to its hinge to also stabilize the open state. Finally, both the two subunits are stabilized as an open-state

that is not compatible for DNA binding due to the inward immobilization of αFs.

The sequential binding manner of cAMP is performed likely in a distinct way. Positions of the two apo-CRP hinges relative to the open subunit occupancy by cAMP are already similar to those in cAMP-CRP structure (Figure 27B, right), whereas deviate when compared relatively to the closed subunit occupancy (Figure 27B, left). Thus, the cNMP-binding site in open subunit is likely preferred for the first cAMP binding. In addition, at the nascent binding of the cAMP of open-subunit, the S128 of closed-subunit is expected to be in a especially close proximity (*ca.* 3.9 Å) for preferable binding, whereas it orients apart from the cAMP binding pocket in the other subunit (Figure 21 and 27B). This notice may infer that the S128 contact is also involved in the nascent binding of cAMP to specifically select *anti*-conformation of the inbound cAMP. Then the subsequent occupancy of the cAMP binding is accomplished via the inward moving of flap, which contacts to K130, and a concomitant interaction between the bound cAMP and T127. Thus, assisted by the R123-T127-V131 and V126-K130-L134 hydrogen bonds (Figure 17), the hinge folds into a stable helix up to F136, resulting in the dissociation and rotation of its own CDD (Figure 28). Meanwhile, the interaction of the open-subunit cAMP with the closed-subunit

S128, which was already included in a helical fold, would not be sufficient to induce further helix elongation required for the dissociated CDD rotation (Figure 28). However, the cAMP-free subunit can keep a conformational exchange between open- and closed-flap conformation, which is favored for optimization of the central hydrophobic core, leading to the CDD rotation via αC elongation and αD shortening. Then, the induced state by conformational exchange is almost identical to the final structure occupied by two cAMPs and highly compatible for DNA binding, which is supported by one molar equivalent of bound cAMP per CRP dimer is sufficient for activation of the protein. Meanwhile, this active conformation with one cAMP hinders recruiting of the second cAMP, since the entrance was closed (Figure 29) by the inward reposition of flap. Thus, the dynamic conformational equilibrium at the one cAMP-bound state can operate as an adverse factor for subsequent cAMP binding, which underlies the negative cooperativity that indicates the first cAMP binding interfering with the second cAMP binding.

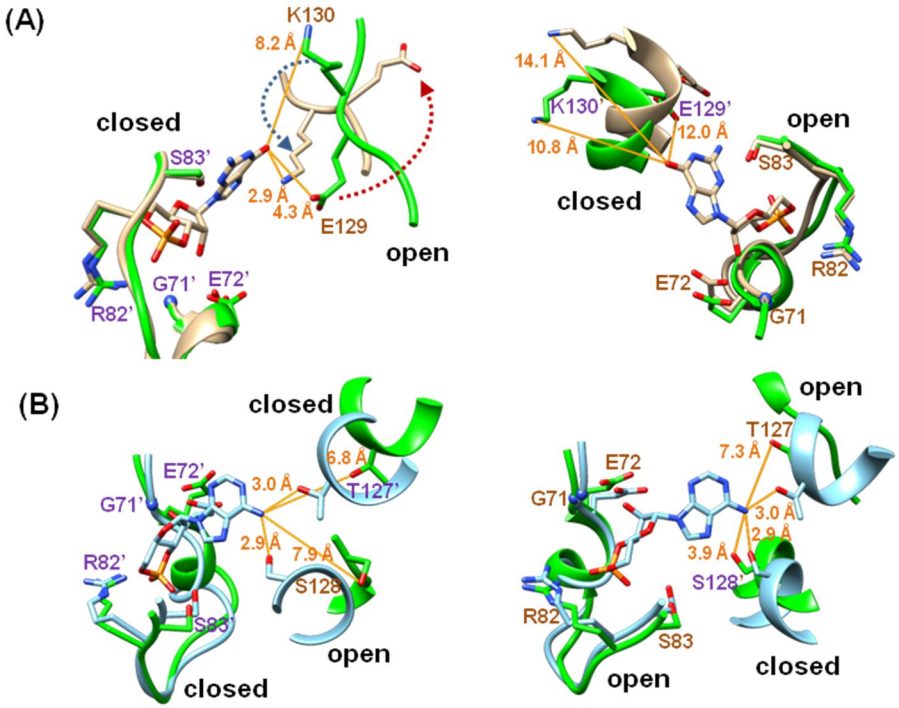

Figure 27. Model for the cNMP binding preferences. The nascent binding of the first inbound cNMP is modeled by superimposing the apo–CRP (*green*) structure with the cGMP–CRP (A; *tan*) and cAMP–CRP (B; *sky blue*) structures, respectively, by matching the G71, E72, R82, and S83 residues in individual subunits. The residues in closed subunits are labeled with apostrophes.

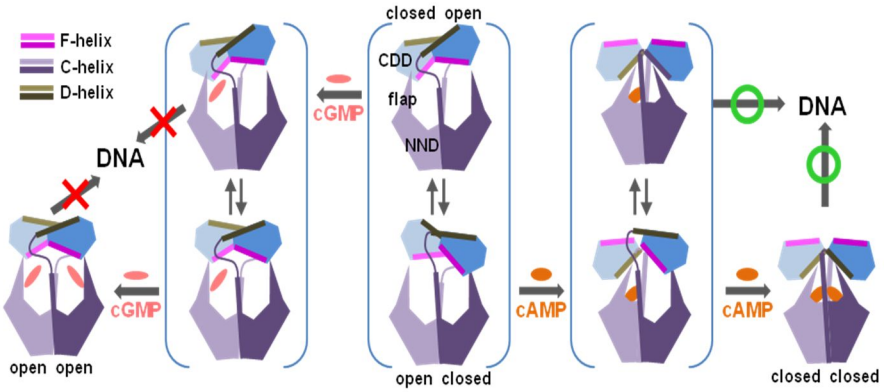

Figure 28. Model for sequential bindings. Conformational exchanges are depicted by antiparallel arrow pairs. Change in aC and aD lengths, inward moving of flaps, closing entrance for cNMP, and CDD rotations are described

2.3.7 Inhibitory conformational change upon the cGMP binding

The cGMP-CRP structure shows that CRP undergoes a certain conformational change upon binding of cGMP. The CDD dimer in cGMP-CRP more tilted and swung compared to that in apo-CRP (Figure 12 and 21). This cGMP-induced CDD rotation is directed, albeit subtly, opposite to that by cAMP (Figure 29) and results in strengthening the interdomain interactions mediated by the αFs, with more compact contacts than in apo-CRP (Figure 20 and 23). However, essential conformational changes to raise the CDD reorientation are observed in NNDs. The contact of 6-keto oxygen of the closed-subunit cGMP to the open-subunit K130 sidechain (Figure 24) forced the altered conformation of the open-subunit hinge (Figure 27A), which primarily contributed to the reorientation of CDD dimer. In addition, due to the altered hinge conformation, the E129 sidechain, which formerly interacted with the R123' in apo-CRP (Figure 21), flipped (Figure 27) to centrally mediate the interusbunit and interdomain interactions in cGMP-CRP (Figure 23).

It is also well established that the NND flap (β4-β5 hairpin region in Figure 12 and 21) stretches crucially mediate the allosteric conformational change of CRP. Upon the cAMP binding, the flaps move towards hinge totally by more than 10

Å (Fig. 16C). The inward moving of hinge is regarded as a resultant of the overall contraction of NNDs, which has been biochemically evidenced. This shortened distance between flaps is not relevant in the cGMP-CRP structure. However, superposition of NNDs indicates that a subtle inward moving of flap occurred in the open subunit, whereas it laterally out-swung in the closed subunit (Figure 29 and 17). Since the flap residues are involved in the interdomain interaction with the CDD αFs (Figure 21 and 23), the subtle reorientation of the CDD dimer in cGMP-CRP is attributable to the flap reposition in concert with the hinge alteration.

Another significant change upon the cGMP binding is verified in the cNMP binding site occupancy. In the cAMP-CRP structure, the purine and furanose rings are compactly stacked into nonpolar sidechains of CRP through maximized van der Waals interactions (Figure 30). The top face of cAMP is attached to L61 and V49 sidechains via the purine ring and to V86 and I30 sidechains via the furanose ring. Those residues preformed the β-roll hydrophobic cluster in apo-CRP and their hydrophobic packing onto cAMP is also relevant in the cGMP binding at both subunits. However, the bottom face packing of cNMPs onto the preformed small hydrophobic cluster composed of L73 and L124 pairs is quite distinguished. The furanose ring of the open-subunit cGMP packed onto the L73,

which is consistent with the cAMP occupancy and responsible for the slight inward moving of the flap (Figure 29). However, the purine ring packing onto the opposite subunit L124, shown in the cAMP binding, didn't occur in the cGMP binding. Considering the open-subunit cGMP interactions as a nascent binding of a cNMP, the L124 contact to cAMP can be regarded as the second step interaction. In contrast, the bound cGMP in the closed subunit was compelled to contact neither the L124 nor the L73 sidechains. Furthermore, disruption of the preformed L73-L124 hydrophobic cluster is observed in the closed subunit occupancy by cGMP. In particular, the L124 has been identified as one of critical residues in communicating cAMP binding site occupancy to the CRP hinge; *i.e.*, a long-range signal transmission from L124 bypassing the αC reaches the hinge residue K130 through cAMP and L61 (Figure 30 and 24). Then, the K130 in the cAMP-bound CRP drags the L134 out of the CDD dimer hydrophobic core by forming a backbone hydrogen bond (Figure 17) and plays a central role in stabilizing the rearranged flap and CDD positions (Figure 25). Thus, conversely, the atomic contact of cGMP to the K130 sidechain could be responsible for the prevention of the L124 contact to cGMP, as it laterally out-swung the β-roll (Figure 16 and 28), leading to the dissociation of L73 from the opposite L124.

In summary, cGMP binding to CRP renders the protein inert by inducing an opposite conformational change to that by cAMP. It is also significant that the K130, which is crucial for both the signal translation of cAMP binding and conformational transition of CRP, was specifically grabbed by cGMP to avoid the conformational activation. Finally, the conformational effect by cGMP appears to be inhibitory to CRP, rather than null, as it is accomplished by specific and stepwise interactions with the S83 and K130 residues. This inhibitory potential of cGMP to CRP would be also relevant to the complete deactivation of a dynamically activated CRP* mutant recently elucidated.

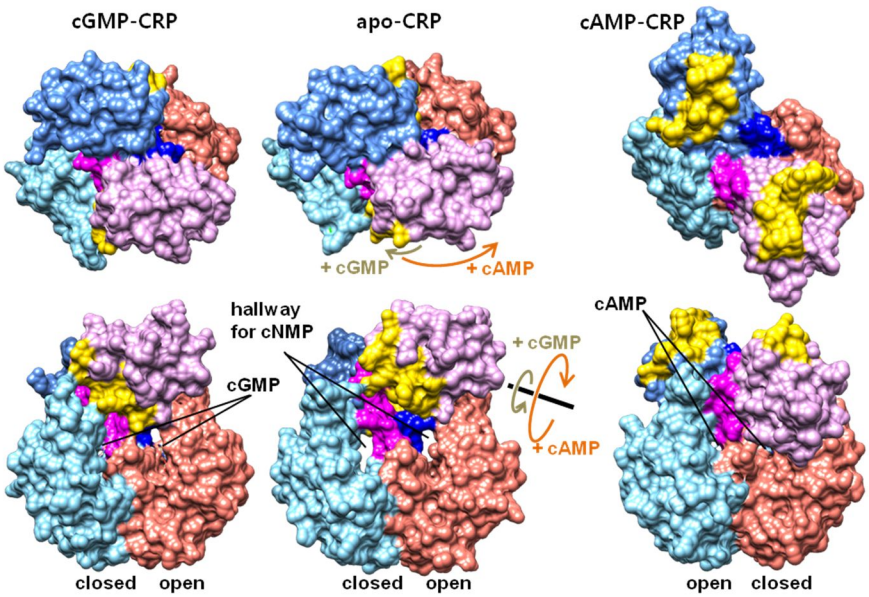

Figure 29. Comparison of domain orientations of CRP structures. Upper and bottom panels are top view and front view, respectively, of each molecule. Surfaces of NND, hinge, and CDD are colored salmon, magenta, and orchid, respectively, in one subunit and sky blue, blue, cornflower blue, respectively, in the other subunit. aFs in both subunits are indicated in *gold*. Directions of the cNMP–induced CDD rotations are depicted by arrows on the apo–CRP structure.

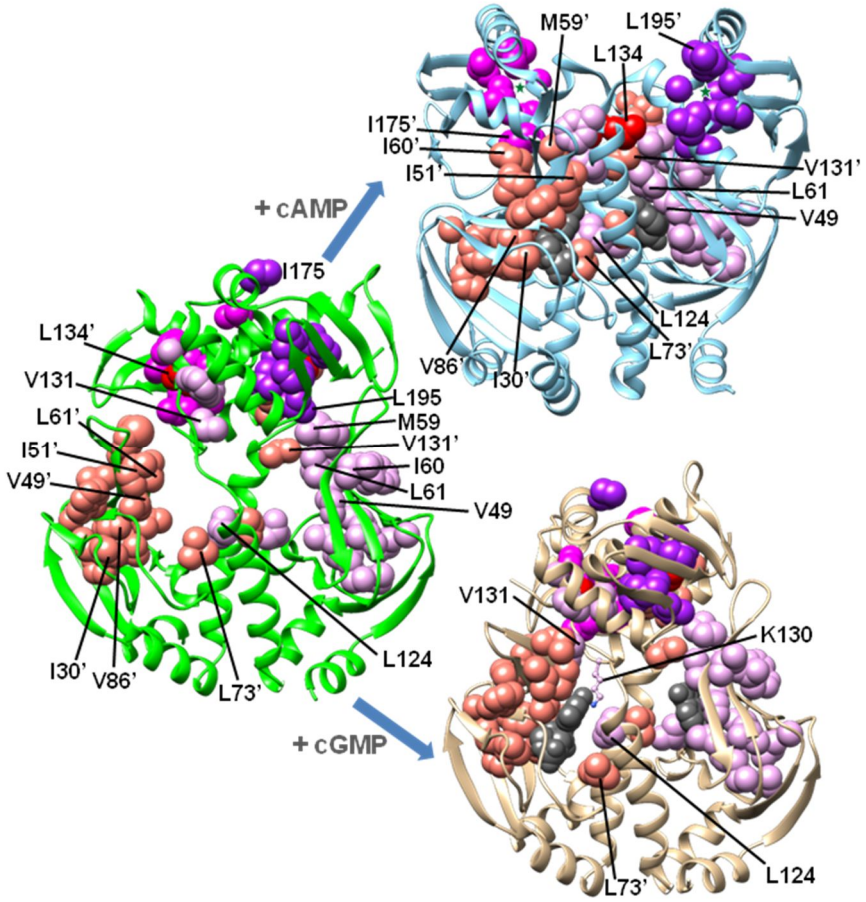

Figure 30. Hydrophobic clusters differentiating CRP conformations. Connections of individual hydrophobic clusters in CDD, flap, and C-helix are changed upon the cNMP binding to apo-CRP. In the cAMP-bound structure, the positions formerly occupied by L134 in apo-state are indicated by *green* star symbols.

2.3.8 Conformational allosterism and driving forces

Owing to paucity of 3D-structural information of the inactive CRP structure, a definite elucidation of the CRP allostery was precluded. The previous D138L-CRP and the present our apo-CRP and cGMP-CRP structures commonly corroborates that the structural determinant of inactiveness is the inward positioning of αFs by the CDD dimerization and the αF interaction with NNDs, while the conformational transition by cAMP is accomplished apparently through a rigid-body rotation of the two CDDs with a concomitant dissociation and without an alteration in their innate folds (Figure 29 and 30). The fact that substitution of a hydrophobic residue in the CDD dimer interface, such as A144, L148 or L195 in Figer 12 and 24, with a polar amino acid falls into a constitutively active CRP* phenotype also indicates that the CDD dissociation would be the most critical requirement for CRP activation.

The present structures provide both the macroscopic and microscopic insights into the global transition of CRP conformation. Overall, the cAMP-induced conformational change is illustrated by a central convergence of hydrophobic clusters that was formerly distributed to αCs, flaps, and CDDs (Figure 30). The classical theory underlying fundamental forces driving protein folding in solution is the construction of hydrophobic core through an internal convergence of

hydrophobic sidechains with neutralizing mainchain polarity by forming secondary structures with hydrogen bonds. The allosteric conformational change of CRP also appears to follow those folding forces to form an optimum hydrophobic core, which in apo-form is hardly driven due to the vacancy of cNMP-binding pockets. Our insights into atomic details how cAMP contributes to driving conformational allosterism are depicted as follows. As a nascent binding, the inbound cAMPs are fitted into the preformed binding pockets of β-rolls, via electrostatic interactions with G71, E72, R82 and S83 residues (Figure 21 and 23) and hydrophobic stacking onto I30, V86, V49, L61 and L73 residues (Figure 30). Then, the L124 of αC, which were communicating with the L73 residues, tug the bound cAMPs by a hydrophobic interaction force. Thus, the β-rolls adhered to cAMPs are also forced to move inward, resulting in the contraction of NNDs. Since the L73 is a fulcrum of that NDD contraction, the inward moving of the distal region flap is more remarkable and contacts to the hinge residue K130 with the cAMP-attached L61. In addition, the bound cAMPs, now shifted toward αCs, accomplish their hydrogen bondings to the hinge residues T127 and S128. As the hinge in apo-CRP is dynamic in nature and had an inherent helical propensity (Figure 15), the quenched flexibility by immobilizing T127, S128, and K130 results in a firm stabilization of helical

conformation through a reconstitution of their backbone hydrogen bonds. As the hydrophobic clusters at flaps are repositioned close to the hinge, the helix elongation at hinge is further forced to neutralize the mainchain polarity and to gather the nearby small hydrophobic clusters composed of the L134-A135-F136-L147 sidechains. Thus, the elongated hydrogen bonding networks starting from the N-proximal region of the hinge finally reach the L137 amide proton and break at between L137 and D138. Then, the concerted hydrogen bonding forces at hinge are translated to a helical torque that triggers a rigid-body rotation of the CDDs. Concomitantly, the CDD dimer, deprived of L134 and L137 at the critical dimer interface, dissociates into monomers and the L137-D138 linkage with a broken network of backbone hydrogen bonds forms a new short hinge connecting the CDDs and NNDs. Finally, the altered orientations of CDDs are stabilized via αD and αE interacting with the new hinge and the moved flap (Figure 25). All those atomic momentums driving the global transition of CRP could operate in a concerted manner and be efficiently blocked by the K130 contact to cGMP. This alternative contact rivets the hinge as an extended stretch not to fold into a helix and pushes out the bound cGMP. Accordingly, the lateral outward swing of β-roll (Figure 16) disrupted the L73-L124 communication (Figure 30) that otherwise functions as the starting point of

active transition. Additionally, the altered conformation of hinge and shifted position of flap detains more strongly the αFs than in apo-CRP (Figure 24), which is relevant to a further deactivation of CRP.

2.4 Conclusion

The purpose of this study is determining the inactive structure of CRP and eventually understanding the mode of allostery conformational change of CRP. In the present study, we can obtain two inactive structure of CRP from *Escherichia coli* by using X-ray crystallography. A higer-resolution (2.2Å) crystal structure of the wild-type apo-CRP, which would be the most proper template for the precise inspection of the inactive conformation was determined. In addition, we succeeded in obtaining the first crystal structure of another inactive form, the cGMP-CRP complex, also at 2.2 Å resolution. These inactive structures underpin our insight into conformational allosterism of CRP and the mode of cyclic nucleotide binding. Inactive structures of CRP are asymmetric dimer, consisting of open subunit and closed subunit and formed the compact dimerization of C-terminal DNA-binding domains and this dimerization result in the inward positioning of F-helix, DNA recognition site. The functional switch of L134 appears as one of central regulatory machineries in CRP by relocation in two hydrophobic clusters, which locate at D-helix in inactive state and hinge region in active state. The binding of cyclic nucleotide to binding pocket of CRP has two stepwise modes and S83 discriminates the false ligand cGMP from its *bona fide*

effector cAMP at initial binding step. In contrast, the binding of cAMP results in activate the CRP, the binding of cGMP cause inhibitory conformational change.

Summary

We could determine the structure of 24th Ig-like domain of filamin from *Homo sapience* and two inactive structure of CRP from *Escherichia coli* using X-ray crystallography.

The structure of FLNa24 compare the structure with other Ig-like domain of filamin and how dimerization is formed in FLNa24. FLNa24 belongs to Ig-like domain which consists of a β-sandwich of seven or more strands in two sheets and FLNa24 play role as dimerization domain of FLNa. Major dimerization interace of FLNa24 located in strands C and D, which has highly conserved hydrophobic residues and forms putative hydrogen bonding network.

CRP has long served as a typical textbook example describing transcription regulation, DNA-binding motif, and allosteric activation of a protein. We succeeded in obtaining higher-resolution crystal structure of the wild-type apo-CRP and another inactive form, the cGMP-CRP complex. These inactive structures not only support our insight into the conformational allosterism, but also address how CRP discriminates the false ligand cGMP from its authentic effector cAMP. We discussed structural difference between apo-CRP, cAMP-CRP and cGMP-CRP. And also the conformational allostery change of

CRP by cyclic mucleotide and inhibitory conformational change by binding of cGMP was discussed.

References

1 Stossel, T. P. *et al.* Filamins as integrators of cell mechanics and signalling. *Nature reviews. Molecular cell biology* 2, 138-145, doi:10.1038/35052082 (2001).

2 van der Flier, A. & Sonnenberg, A. Structural and functional aspects of filamins. *Biochimica et biophysica acta* 1538, 99-117 (2001).

3 Feng, Y. *et al.* Filamin A (FLNA) is required for cell-cell contact in vascular development and cardiac morphogenesis. *Proceedings of the National Academy of Sciences of the United States of America* 103, 19836-19841, doi:10.1073/pnas.0609628104 (2006).

4 Zhou, X. *et al.* Filamin B deficiency in mice results in skeletal malformations and impaired microvascular development. *Proceedings of the National Academy of Sciences of the United States of America* 104, 3919-3924, doi:10.1073/pnas.0608360104 (2007).

5 Dalkilic, I., Schienda, J., Thompson, T. G. & Kunkel, L. M. Loss of FilaminC (FLNc) results in severe defects in myogenesis and myotube structure. *Molecular and cellular biology* 26, 6522-6534, doi:10.1128/MCB.00243-06 (2006).

6 Robertson, S. P. Otopalatodigital syndrome spectrum disorders: otopalatodigital syndrome types 1 and 2, frontometaphyseal dysplasia and Melnick-Needles syndrome. *European journal of human genetics : EJHG* 15, 3-9, doi:10.1038/sj.ejhg.5201654 (2007).

7 Fox, J. W. *et al.* Mutations in filamin 1 prevent migration of cerebral cortical neurons in human periventricular heterotopia. *Neuron* 21, 1315-1325 (1998).

8 Krakow, D. *et al.* Mutations in the gene encoding filamin B disrupt vertebral segmentation, joint formation and skeletogenesis. *Nature genetics* 36, 405-410, doi:10.1038/ng1319 (2004).

9 Vorgerd, M. *et al.* A mutation in the dimerization domain of filamin c causes a novel type of autosomal dominant myofibrillar myopathy. *American journal of human genetics* 77, 297-304, doi:10.1086/431959 (2005).

10 Flanagan, L. A. *et al.* Filamin A, the Arp2/3 complex, and the morphology and function of cortical actin filaments in human melanoma cells. *The Journal of cell biology* 155, 511-517, doi:10.1083/jcb.200105148 (2001).

11 Kiema, T. *et al.* The molecular basis of filamin binding to integrins and competition with talin. *Molecular cell* 21, 337-347, doi:10.1016/j.molcel.2006.01.011 (2006).
12 Nakamura, F. *et al.* The structure of the GPIb-filamin A complex. *Blood* 107, 1925-1932, doi:10.1182/blood-2005-10-3964 (2006).
13 Pudas, R., Kiema, T. R., Butler, P. J., Stewart, M. & Ylanne, J. Structural basis for vertebrate filamin dimerization. *Structure* 13, 111-119, doi:10.1016/j.str.2004.10.014 (2005).
14 Otwinowski, Z. & Minor, W. Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods in enzymology* 276, 307-326 (1997).
15 The CCP4 suite: programs for protein crystallography. *Acta crystallographica. Section D, Biological crystallography* 50, 760-763, doi:10.1107/S0907444994003112 (1994).
16 McCoy, A. J. *et al.* Phaser crystallographic software. *Journal of applied crystallography* 40, 658-674, doi:10.1107/S0021889807021206 (2007).
17 Stein, N. CHAINSAW: a program for mutating pdb files used as templates in molecular replacement. *Journal of applied crystallography* 41, 641-643, doi:Doi 10.1107/S0021889808006985 (2008).
18 Brunger, A. T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D* 54, 905-921, doi:Doi 10.1107/S0907444998003254 (1998).
19 Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta crystallographica. Section D, Biological crystallography* 53, 240-255, doi:10.1107/S0907444996012255 (1997).
20 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta crystallographica. Section D, Biological crystallography* 60, 2126-2132, doi:10.1107/S0907444904019158 (2004).
21 Winn, M. D., Murshudov, G. N. & Papiz, M. Z. Macromolecular TLS refinement in REFMAC at moderate resolutions. *Methods in enzymology* 374, 300-321, doi:10.1016/S0076-6879(03)74014-2 (2003).
22 Laskowski, R. A., Macarthur, M. W., Moss, D. S. & Thornton, J. M. Procheck - a Program to Check the Stereochemical Quality of Protein Structures. *Journal of applied crystallography* 26, 283-291, doi:Doi 10.1107/S0021889892009944 (1993).
23 Joosten, R. P. *et al.* A series of PDB related databases for everyday needs. *Nucleic acids research* 39, D411-419, doi:10.1093/nar/gkq1105 (2011).

24 Himmel, M., Van Der Ven, P. F., Stocklein, W. & Furst, D. O. The limits of promiscuity: isoform-specific dimerization of filamins. *Biochemistry* 42, 430-439, doi:10.1021/bi026501+ (2003).

25 Halaby, D. M., Poupon, A. & Mornon, J. The immunoglobulin fold family: sequence analysis and 3D structure comparisons. *Protein engineering* 12, 563-571 (1999).

26 Bork, P., Holm, L. & Sander, C. The immunoglobulin fold. Structural classification, sequence patterns and common core. *Journal of molecular biology* 242, 309-320 (1994).

27 Bodian, D. L., Jones, E. Y., Harlos, K., Stuart, D. I. & Davis, S. J. Crystal structure of the extracellular region of the human cell adhesion molecule CD2 at 2.5 A resolution. *Structure (London, England : 1993)* 2, 755-766 (1994).

28 Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic acids research* 25, 4876-4882 (1997).

29 Thompson, J. D., Higgins, D. G. & Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic acids research* 22, 4673-4680 (1994).

30 Krissinel, E. & Henrick, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta crystallographica. Section D, Biological crystallography* 60, 2256-2268, doi:10.1107/S0907444904026460 (2004).

31 Sjekloca, L. *et al.* Crystal structure of human filamin C domain 23 and small angle scattering model for filamin C 23-24 dimer. *Journal of molecular biology* 368, 1011-1023, doi:10.1016/j.jmb.2007.02.018 (2007).

32 Feng, Y. & Walsh, C. A. The many faces of filamin: a versatile molecular scaffold for cell motility and signalling. *Nature cell biology* 6, 1034-1038, doi:10.1038/ncb1104-1034 (2004).

33 Ohta, Y., Suzuki, N., Nakamura, S., Hartwig, J. H. & Stossel, T. P. The small GTPase RalA targets filamin to induce filopodia. *Proceedings of the National Academy of Sciences of the United States of America* 96, 2122-2128 (1999).

34 Browne, K. A., Johnstone, R. W., Jans, D. A. & Trapani, J. A. Filamin (280-kDa actin-binding protein) is a caspase substrate and is also cleaved directly by the cytotoxic T lymphocyte protease granzyme B during apoptosis. *The Journal of biological chemistry* 275, 39262-39266, doi:10.1074/jbc.C000622200 (2000).

35 Won, H. S., Lee, Y. S., Lee, S. H. & Lee, B. J. Structural overview on the allosteric activation of cyclic AMP receptor protein. *Biochimica et biophysica acta* 1794, 1299-1308, doi:10.1016/j.bbapap.2009.04.015 (2009).

36 Harman, J. G. Allosteric regulation of the cAMP receptor protein. *Biochimica et biophysica acta* 1547, 1-17 (2001).

37 Hollands, K., Busby, S. J. & Lloyd, G. S. New targets for the cyclic AMP receptor protein in the Escherichia coli K-12 genome. *FEMS microbiology letters* 274, 89-94, doi:10.1111/j.1574-6968.2007.00826.x (2007).

38 Passner, J. M. & Steitz, T. A. The structure of a CAP-DNA complex having two cAMP molecules bound to each monomer. *Proceedings of the National Academy of Sciences of the United States of America* 94, 2843-2847 (1997).

39 Kolb, A., Busby, S., Buc, H., Garges, S. & Adhya, S. Transcriptional regulation by cAMP and its receptor protein. *Annual review of biochemistry* 62, 749-795, doi:10.1146/annurev.bi.62.070193.003533 (1993).

40 Botsford, J. L. & Harman, J. G. Cyclic AMP in prokaryotes. *Microbiological reviews* 56, 100-122 (1992).

41 Aiba, H., Fujimoto, S. & Ozaki, N. Molecular cloning and nucleotide sequencing of the gene for E. coli cAMP receptor protein. *Nucleic acids research* 10, 1345-1361 (1982).

42 McKay, D. B. & Steitz, T. A. Structure of catabolite gene activator protein at 2.9 A resolution suggests binding to left-handed B-DNA. *Nature* 290, 744-749 (1981).

43 Passner, J. M., Schultz, S. C. & Steitz, T. A. Modeling the cAMP-induced allosteric transition using the crystal structure of CAP-cAMP at 2.1 Å resolution. *Journal of molecular biology* 304, 847-859, doi:10.1006/jmbi.2000.4231 (2000).

44 Schultz, S. C., Shields, G. C. & Steitz, T. A. Crystal structure of a CAP-DNA complex: the DNA is bent by 90 degrees. *Science* 253, 1001-1007 (1991).

45 Benoff, B. *et al.* Structural basis of transcription activation: the CAP-alpha CTD-DNA complex. *Science* 297, 1562-1566, doi:10.1126/science.1076376 (2002).

46 Parkinson, G. *et al.* Structure of the CAP-DNA complex at 2.5 angstroms resolution: a complete picture of the protein-DNA interface. *Journal of molecular biology* 260, 395-408 (1996).

47 Weber, I. T. & Steitz, T. A. Structure of a complex of catabolite gene activator protein and cyclic AMP refined at 2.5 A resolution. *Journal of molecular biology* 198, 311-326 (1987).

48 Eilen, E. & Krakow, J. S. Effects of cyclic nucleotides on the conformational states of the alpha core of the cyclic AMP

receptor protein. *Biochimica et biophysica acta* 493, 115-121 (1977).

49 Krakow, J. S. & Pastan, I. Cyclic adenosine monophosphate receptor: loss of cAMP-dependent DNA binding activity after proteolysis in the presence of cyclic adenosine monophosphate. *Proceedings of the National Academy of Sciences of the United States of America* 70, 2529-2533 (1973).

50 Lee, B. J., Lee, S. J., Hayashi, F., Aiba, H. & Kyogoku, Y. A nuclear magnetic resonance study of the cyclic AMP receptor protein (CRP): assignments of the NH protons of histidine and tryptophan residues and the effect of binding of cAMP to CRP. *Journal of biochemistry* 107, 304-309 (1990).

51 Sixl, F., King, R. W., Bracken, M. & Feeney, J. 19F-n.m.r. studies of ligand binding to 5-fluorotryptophan- and 3-fluorotyrosine-containing cyclic AMP receptor protein from Escherichia coli. *The Biochemical journal* 266, 545-552 (1990).

52 Eilen, E., Pampeno, C. & Krakow, J. S. Production and properties of the alpha core derived from the cyclic adenosine monophosphate receptor protein of Escherichia coli. *Biochemistry* 17, 2469-2473 (1978).

53 Ghosaini, L. R., Brown, A. M. & Sturtevant, J. M. Scanning calorimetric study of the thermal unfolding of catabolite activator protein from Escherichia coli in the absence and presence of cyclic mononucleotides. *Biochemistry* 27, 5257-5261 (1988).

54 Kumar, S. A., Murthy, N. S. & Krakow, J. S. Ligand-induced change in the radius of gyration of cAMP receptor protein from Escherichia coli. *FEBS letters* 109, 121-124 (1980).

55 Heyduk, T. & Lee, J. C. Escherichia coli cAMP receptor protein: evidence for three protein conformational states with different promoter binding affinities. *Biochemistry* 28, 6914-6924 (1989).

56 Takahashi, M., Blazy, B., Baudras, A. & Hillen, W. Ligand-modulated binding of a gene regulatory protein to DNA. Quantitative analysis of cyclic-AMP induced binding of CRP from Escherichia coli to non-specific and specific DNA targets. *Journal of molecular biology* 207, 783-796 (1989).

57 Baichoo, N. & Heyduk, T. Mapping conformational changes in a protein: application of a protein footprinting technique to cAMP-induced conformational changes in cAMP receptor protein. *Biochemistry* 36, 10830-10836, doi:10.1021/bi970714v (1997).

58 Dong, A., Malecki, J. M., Lee, L., Carpenter, J. F. & Lee, J. C. Ligand-induced conformational and structural dynamics changes in Escherichia coli cyclic AMP receptor protein. *Biochemistry* 41, 6660-6667 (2002).

59 Won, H. S. *et al.* Structural understanding of the allosteric conformational change of cyclic AMP receptor protein by cyclic AMP binding. *Biochemistry* 39, 13953-13962 (2000).

60 Ebright, R. H., Le Grice, S. F., Miller, J. P. & Krakow, J. S. Analogs of cyclic AMP that elicit the biochemically defined conformational change in catabolite gene activator protein (CAP) but do not stimulate binding to DNA. *Journal of molecular biology* 182, 91-107 (1985).

61 Ren, Y. L., Garges, S., Adhya, S. & Krakow, J. S. Characterization of the binding of cAMP and cGMP to the CRP*598 mutant of the E. coli cAMP receptor protein. *Nucleic acids research* 18, 5127-5132 (1990).

62 Takahashi, M., Blazy, B. & Baudras, A. An equilibrium study of the cooperative binding of adenosine cyclic 3',5'-monophosphate and guanosine cyclic 3',5'-monophosphate to the adenosine cyclic 3',5'-monophosphate receptor protein from Escherichia coli. *Biochemistry* 19, 5124-5130 (1980).

63 Malecki, J., Polit, A. & Wasylewski, Z. Kinetic studies of cAMP-induced allosteric changes in cyclic AMP receptor protein from Escherichia coli. *The Journal of biological chemistry* 275, 8480-8486 (2000).

64 Won, H. S., Lee, T. W., Park, S. H. & Lee, B. J. Stoichiometry and structural effect of the cyclic nucleotide binding to cyclic AMP receptor protein. *The Journal of biological chemistry* 277, 11450-11455, doi:10.1074/jbc.M112411200 (2002).

65 Belduz, A. O., Lee, E. J. & Harman, J. G. Mutagenesis of the cyclic AMP receptor protein of Escherichia coli: targeting positions 72 and 82 of the cyclic nucleotide binding pocket. *Nucleic acids research* 21, 1827-1835 (1993).

66 Gronenborn, A. M., Sandulache, R., Gartner, S. & Clore, G. M. Mutations in the cyclic AMP binding site of the cyclic AMP receptor protein of Escherichia coli. *The Biochemical journal* 253, 801-807 (1988).

67 Moore, J., Kantorow, M., Vanderzwaag, D. & McKenney, K. Escherichia coli cyclic AMP receptor protein mutants provide evidence for ligand contacts important in activation. *Journal of bacteriology* 174, 8030-8035 (1992).

68 Sharma, H., Yu, S., Kong, J., Wang, J. & Steitz, T. A. Structure of apo-CAP reveals that large conformational changes are necessary for DNA binding. *Proceedings of the National Academy of Sciences of the United States of America* 106, 16604-16609, doi:10.1073/pnas.0908380106 (2009).

69 Popovych, N., Tzeng, S. R., Tonelli, M., Ebright, R. H. & Kalodimos, C. G. Structural basis for cAMP-mediated allosteric control of the catabolite activator protein. *Proceedings of the*

National Academy of Sciences of the United States of America 106, 6927-6932, doi:10.1073/pnas.0900595106 (2009).

70 Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L. & Clardy, J. Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *Journal of molecular biology* 229, 105-124, doi:10.1006/jmbi.1993.1012 (1993).

71 Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta crystallographica. Section D, Biological crystallography* 66, 12-21, doi:10.1107/S0907444909042073 (2010).

72 Adams, P. D. *et al.* PHENIX: building new software for automated crystallographic structure determination. *Acta crystallographica. Section D, Biological crystallography* 58, 1948-1954 (2002).

73 Matthews, B. W. Solvent content of protein crystals. *Journal of molecular biology* 33, 491-497 (1968).

74 Potterton, L. *et al.* Developments in the CCP4 molecular-graphics project. *Acta crystallographica. Section D, Biological crystallography* 60, 2288-2294, doi:10.1107/S0907444904023716 (2004).

75 Brünger, A. T. Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* 355, 472-475 (1992).

76 Kleywegt, G. J. & Jones, T. A. Databases in protein crystallography. *Acta crystallographica. Section D, Biological crystallography* 54, 1119-1131 (1998).

77 Schüttelkopf, A. W. & van Aalten, D. M. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta crystallographica. Section D, Biological crystallography* 60, 1355-1363, doi:10.1107/S0907444904011679 (2004).

78 Popovych, N., Sun, S., Ebright, R. H. & Kalodimos, C. G. Dynamically driven protein allostery. *Nature structural & molecular biology* 13, 831-838, doi:10.1038/nsmb1132 (2006).

79 Tomlinson, S. R., Tutar, Y. & Harman, J. G. Interaction of CRP L124 with cAMP affects CRP cAMP binding constants, cAMP binding cooperativity, and CRP allostery. *Biochemistry* 42, 3759-3765, doi:10.1021/bi027126o (2003).

80 Lanzilotta, W. N. *et al.* Structure of the CO sensing transcription activator CooA. *Nature structural biology* 7, 876-880, doi:10.1038/82820 (2000).

81 Chan, M. K. CooA, CAP and allostery. *Nature structural biology* 7, 822-824, doi:10.1038/79559 (2000).

82 Kumar, P. *et al.* Mapping conformational transitions in cyclic AMP receptor protein: crystal structure and normal-mode analysis of Mycobacterium tuberculosis apo-cAMP receptor

protein. *Biophysical journal* 98, 305-314, doi:10.1016/j.bpj.2009.10.016 (2010).

83 Rehmann, H., Wittinghofer, A. & Bos, J. L. Capturing cyclic nucleotides in action: snapshots from crystallographic studies. *Nature reviews. Molecular cell biology* 8, 63-73, doi:10.1038/nrm2082 (2007).

84 Kim, J. J. *et al.* Co-crystal structures of PKG Ibeta (92-227) with cGMP and cAMP reveal the molecular details of cyclic-nucleotide binding. *PloS one* 6, e18413, doi:10.1371/journal.pone.0018413 (2011).

국문초록

Filamin 은 actin filaments 에 교차 결합함으로 actin 세포골격을 조절하는 역할을 수행한다. Actin 세포골격의 재구성은 세포의 모양유지, 세포 분화, 세포 이동과 식세포작용과 같은 다양한 세포의 기능에 중요한 역할을 한다. 또한 Filamin 은 β-integrins 와 같은 transmembrane receptors 와 cytosolic signaling proteins 를 고정시키는 역할을 하여 다양한 cellular process 들에 관여한다.

인간에게 있는 세 종류의 isoform 중 filamin A 가 가장 흔하고 널리 발현되는 Filamin A 는 각 subunit 당 280kDa 의 분자량을 갖는 V-형의 dimeric 단백질이고, 한 개의 Actin binding domain 과, 24 개의 repeating homologous rod domain 들과 두 개의 hinge regions 으로 구성되어 있다. 각 rod domain 은 샌드위치 모양의 β-sheets 로 구성된 immunoglobulin-like folding 을 갖고 있다. 그 중 24 번 rod domain 을 Crystallography 를 이용하여 3 차 구조를 규명하고, 이를 바탕으로 dimerization 양상을 관찰하였다. Strands C 와 D 에서 쌍방의 수소결합과 Hydrophobic 결합을 통하여 강하게 dimerization 을 이루고 있으며, filamin 전체를 dimerization 시켜 유연한 V-형의 homodier 로 만들고, 이를 통해 actin networks 를 안정화 시킨다. 또한 Ig-like 단백질들과 다른 filamin isoform 의 dimerization domain 과 구조를 비교하여 공통점과 차이점을 분석하였다.

E. coli 유래의 Adenosine 3', 5'-Cyclic monophosphate (cAMP) receptor protein (CRP)은 transcription regulation,

DNA-binding motif, 단백질의 allosteric activation 에 대한 예시로 오래 전부터 교과서에 실린 널리 알려진 단백질이다. 특정 DNA sites 와 결합하고 RNA polymerase 와 상호작용을 통해 200 여 개 이상의 유전자의 전사를 조절하는데 중요한 역할을 한다.

CRP 는 각 subunit 이 209 개의 아미노산으로 구성된, 24kDa 의 분자량을 갖는 homodier 이다. 각 subunit 은 DNA 와 결합하는 C-terminal domain (CDD)과 cyclic nucleotide 와 결합하는 N-terminal domain (NND), 두 domain 을 연결하는 짧고 flexible 한 hinge region 을 갖는다. 특히 C-terminal 의 F-helix 는 이웃한 E-helix 와 전형적인 helix-turn-helix 를 구성하여 특정 DNA 를 인식하는 역할을 한다. 하지만 이 DNA 인식작용은 cAMP 의 부재시에는 일어나지 않으며, cAMP 결합에 의한 활성화에 의해 allosteric conformation 변화에 의해서 일어난다. 또한 최근 단백질 dynamics 에 대한 연구를 통해 CRP 의 dynamic allosterism 에 대해서도 밝혀지고 있다.

1980 년대 이래로 CRP 의 3 차 구조는 10 여개 이상 밝혀져 있지만, 이들 대부분이 CRP- cAMP, CRP-cAMP-DNA, CRP-cAMP-DNA-RNAP 등의 활성화 상태의 구조들이다. 근래에 inactive 상태의 CRP 구조가 X-ray crystallography 와 NMR 을 통해서 밝혀졌지만, 이것들은 너무 낮은 resolution 을 갖고 있거나, 변종들의 구조인데다가, 결정 구조와 NMR 구조가 서로 다르기에 CRP 의 conformational allosterism 에 대해서는 명확히 결론 지을 수 없는 상황이다.

cGMP 는 purine 부분을 갖는 cAMP 유사체이고, CRP 와의 결합 상수도 비슷하지만, CRP 가 특정 DNA sequences 와의 결합을 활성화 시키지 못한다. 또한 cAMP 는 최대 4 개의 분자가 CRP

dimer 와 결합하는데, cGMP 는 최대 2 개의 분자가 결합한다. 많은 cAMP 와 cGMP 의 CRP 결합에 관한 변종 실험들에서도 cGMP 는 CRP 의 allostery 의 control probe 로 사용되었다. 하지만 아직까지 cGMP-CRP complex 의 구조는 규명되지 않고 있다.

본 연구에서는, 2.2Å이라는 높은 resolution 의 wild-type 의 apo-CRP 의 구조를 규명하는데 성공하여, 이전의 inactive 구조에서 보다 향상된 원자 수준의 분석이 가능하게 되었다. 게다가 2.2Å의 높은 resolution 의 cGMP-CRP 구조도 최초로 규명하였다. 이를 통해 cAMP 의 유사체이지만 비 유효 분자인 cGMP 의 작용 기전에 대한 연구를 수행하였으며, 이를 통하여 CRP 의 conformational allosterism 에 대한 연구를 진행하였다.

CRP 의 allosteric conformational change 는 도메인의 local fold 에는 변화가 없지만 NTD 와 CTD 를 연결하는 hinge 의 재구성을 통하여 방향성의 변화에 의하여 일어나며, 이는 L134 를 스위치로 하는 hydrophobic cluster 의 재구성에 의해 강한 힘을 갖고 있으며, hinge region 의 helical 특성의 변화도 분석된다.
또한 cGMP 와의 complex 에서는 cyclic nucleotide 들이 2 단계로 결합하는 단계적 결합에 대한 증거를 확인하였고, cGMP 와의 결합은 CRP 를 단순히 활성화시키지 못하는 것뿐만 아니라, conformational change 가 cAMP 와는 반대적으로 일어나고, 내부에 파묻혀있는 DNA 인식 부위인 F-helix 를 더 강하게 붙잡고 있는 것으로 보아 CRP 의 비활성 상태를 더욱 안정화 시키는 것을 확인하였다. 이러한 분석 결과를 토대로 CRP 의 allosteric conformational change 에 대한 model 을 제시하였다.

.....

주요어: Filamin; FLNa24; dimerization; CRP; Crystallography; allostery; cyclic nucleotide

학번: 2008-30482