Delicate conformational changes of a protein in the CRP family lead to dramatic functional changes via binding of an alternate secondary messenger molecule

Shan-Ho Chou1,2
1Institute of Biochemistry; 2National Chung Hsing University Biotechnology Center; National Chung-Hsing University; Taichung, Taiwan Republic of China

CLP (c-AMP receptor protein like protein) from the plant pathogen Xanthomonas campestris (Xc) is a moderately strong c-di-GMP binder. After binding with c-di-GMP, XcCLP is dissociated from DNA, thus changing the expression profile of approximately 300 downstream genes, many of which are involved in the pathogenicity of Xc. To better understand the intricacies of this interesting regulation mechanism, a crystal structure of XcCLP has been determined to a resolution of 2.28 Å. This, in conjunction with an abundance of biochemical studies of mutants based on this structure, has illuminated a better understanding of c-di-GMP receptor.

In this addendum, the following aspects of this unique protein shall be described from a more structural perspective: (1) Why c-di-GMP cannot bind to the c-NMP domain of XcCLP; (2) Why XcCLP can intrinsically bind to cognate DNA in the absence of any ligand; and (3) How c-di-GMP inhibits the DNA-binding capability of XcCLP, thus releasing XcCLP from DNA. Interestingly, the aforementioned questions are related by the subtle side-chain flipping of several key signature amino acids. In particular, through mutant studies and sequence alignments of proteins in the CRP family, four key signature amino acid residues have been identified that may allow one to distinguish whether a member in the CRP-FNR family is a c-AMP or a c-di-GMP responsive protein. In conclusion, a possible evolutionary pathway from EcCRP to XcCLP will be described.

Key words: Xcc, pathogenicity, CRP, CLP, c-di-GMP receptor, quorum sensing

Introduction

Signal transduction is crucial for the adaptation and survival of microorganisms in a dynamic environment. These processes involve not only macromolecules, but also small molecules, especially those of nucleotide origin, such as c-AMP, c-GMP, ppGpp, c-di-GMP and c-di-AMP.1 C-AMP is possibly the most extensively studied secondary messenger molecule in prokaryotes to date.2 It is synthesized by a single adenylyl cyclase (Cya) in response to glucose shortage and is degraded by a single phosphodiesterase (CpdA) when glucose supply is abundant. C-AMP acts by binding to c-AMP-receptor protein (CRP), which induces a conformational change in CRP, thus rendering the holomolecular complex capable of binding to specific DNA sequences.3-5 This signal transduction results in the expression of enzymes responsible for the catabolism of alternative carbon sources. In addition, c-AMP has also been recently found to be directly or indirectly involved in the regulation of a wide variety of other important functions such as flagellum biosynthesis, biofilm formation and virulence.1

C-di-GMP was first discovered over 20 years ago, but has only recently received great attention. It is now considered to be an important and ubiquitous secondary messenger, regulating the bacterial lifestyle transition from the single-cellular motile state to the multi-cellular surfacetached sessile state.7-11 This lifestyle transition is correlated with the quorum sensing phenomenon that can lead to biofilm formation in prokaryotes.12-14
In addition, c-di-GMP has also been shown to substantially affect bacterial virulence, cell differentiation and development etc.\(^1\) It is generally believed that more virulence, cell differentiation and development motifs, and are instead responsible for the EAL or the HD-GYP domain.\(^{15-21}\) However, scientists are just beginning to understand the functionality of c-di-GMP.\(^{20,21}\) To date, several different c-di-GMP effectors have been found to bind to c-di-GMP with reasonable affinities, with \(K_d\) values ranging from \(\mu\)M to sub-\(\mu\)M. The PilZ domain protein is the first such effector predicted by a bioinformatics approach to possess c-di-GMP binding capability,\(^{22}\) which was later experimentally confirmed.\(^{23}\) However, it is important to note that not all PilZ domains contain the RxxxR and D/NxSxxG signature motifs necessary for direct c-di-GMP binding. Some PilZ domain proteins have been found to contain degenerate RxxxR and D/NxSxxG motifs that can only bind to c-di-GMP indirectly via a helper protein.\(^{24,25}\) In addition to PilZ, several other types of c-di-GMP receptors have also been identified, including PleD, PelD, LapD and FleQ.\(^{9,10,26,27}\) However, most of these proteins contain no DNA-binding motif, and are instead responsible for the allosteric regulation of exopolysaccharide (EPS) production. Only FleQ was found to act as a transcription factor, derepressing the gene expression of the EPS synthesis enzymes by binding with c-di-GMP.\(^{26}\)

CLP, in contrast, is a global transcription factor and a newly discovered c-di-GMP receptor\(^{28-31}\) that has been found to control the protein expression of approximately three-hundred downstream genes,\(^{32}\) many of which are involved in the pathogenesis of \textit{Xanthomonas campestris} (\(Xc\)). Thus, it is of considerable interest to be able to understand how c-di-GMP links cell-cell signaling to virulence gene expression via CLP in \(Xc\), which is the main topic of several recent publications.\(^{28-35}\) It is now well established that c-di-GMP binds to \(Xc\)-CLP moderately well, which results in the release of \(Xc\)-CLP from binding to cognate DNA, thus changing the protein expression profile in \(Xc\). However, it is still difficult to understand how \(Xc\)-CLP carries out such crucial functions without a detailed knowledge of its tertiary structure, which was not solved until recently.\(^{29}\) In this addendum, the \(Xc\)-CLP protein shall be described from a more structural perspective in order to illustrate how the subtle structural changes resulting from the few key amino acid substitutions between two highly homologous \(Ec\)-CRP and \(Xc\)-CLP sequences can lead to dramatically different secondary messenger molecule binding proteins and to overall different biological functions.

### Why \(Xc\)-CLP Cannot Bind c-AMP with Strong Affinity

From Figure 3A of our recently published article,\(^{29}\) it is clear that two important substitutions in the c-NMP binding pocket of \(Xc\)-CLP have been made: the substitutions of Ser84 to Glu99 and of Glu130' to Arg150' (residues marked with an apostrophe indicates that they are from an adjacent protomer). The substitution of Ser84 to Glu99 has resulted in the disappearance of five direct or water-mediated H-bonds/salt bridges in this pocket. Also, the bulky negatively charged side chain of Glu99 causes considerable blockage to the c-AMP binding site if its side chain is not rearranged. In addition, substantial negatively charged repulsions between c-AMP and Glu99 are expected to prevent c-AMP from accessing this pocket. Altogether, these factors considerably reduce c-AMP binding in the c-NMP domain of \(Xc\)-CLP. This scenario is very different than that of the c-AMP binding pocket in \(Ec\)-CRP.\(^{28,33}\) When the tertiary structure of CLP was not available, it was suggested that c-AMP and c-di-GMP may bind at the same c-NMP site because addition of a very high level of c-AMP (5 mM) to a CLP sample that has been inhibited by c-di-GMP toward DNA binding reversed that inhibition.\(^{28}\) In another word, c-AMP may competitively displace c-di-GMP from the specific binding site and drive CLP into a conformation suitable for DNA binding. The significance of this finding is not yet clear. From our published \(Xc\)-CLP structure, the c-NMP pocket seems to not be large enough to accommodate a c-di-GMP molecule. Also, no adenylylate cyclase or CRP protein exists in the xanthomonads.\(^{34}\) However, based on our docking study, it is also possible that c-di-GMP first binds at the position between the \(\alpha\)C and \(\alpha\)D helices, and that binding of c-AMP at the c-NMP pocket weakens the binding affinity of CRP toward c-di-GMP, hence replacing c-di-GMP with c-AMP when c-AMP concentration is at the high mM range. This conclusion is consistent with the result that c-AMP only exhibits mM affinity toward \(Xc\)-CLP.\(^{34}\)

### Why \(Xc\)-CLP Can Bind to Cognate DNA with Strong Affinity in the Absence of Ligand Induction

Several \(Ec\)-CRP mutants\(^{35-38}\) and one PrfA mutant\(^{39}\) have been found to be constitutively active in binding with their cognate DNAs without the need of c-AMP induction. Previous structural and functional investigations have uncovered several residues important for intrinsic DNA binding.\(^{35,38,40}\) Again, the substitution of Ser84 to Glu99 and of Glu130' to Arg150' seem to play an important role in conferring \(Xc\)-CLP the ability to bind DNA intrinsically. In the wild-type \(Ec\)-CRP, it is believed that binding of the c-AMP:N6 atom to the side chain oxygen atom of Ser84 in the vicinity of the binding pocket and to that of the ‘Thr149’ located in the \(\alpha\)C helix of an adjacent monomer is responsible for transmitting the conformational change required for DNA binding.\(^{35,36}\) In \(Xc\)-CLP, however, there is no ligand present to induce such a conformational change. In contrast, the substitutions of Ser84 to Glu99 and of Glu130' to Arg150' in the \(\alpha\)C helix, along with the rotations of the Cx-CB torsional angle of the Glu99 and Arg150' residues, have enabled the formation of a novel salt bridge between the two altered residues.\(^{29}\) This alternate direct interaction, as well as an indirect interaction between the residues Glu99 and Arg150' via a bridging water (Sup. Fig. 1 in the article by Chin et al.),\(^{29}\) may mimic the conformational transmission.
as that caused by the binding of cAMP to Ec-CRP. In addition, two other interactions have also been found to be involved in conferring XcCLP the intrinsic ability to bind to cognate DNA: substitutions of residue Gly142 in Ec-CRP to an Asp162 in XcCLP and of residue Ala145 in Ec-CRP to a bulky Val165 in XcCLP.\textsuperscript{37,38} The amino acid residue at position 142 in Ec-CRP is located in the beginning of an αD helix, and through the formation of a salt bridge with the Asp139 at the hinge region, it draws near the αC and αD helices and strengthens their interaction. As a result, the αF helix is released from interacting with the αD helix and is now situated at a more favorable position for interacting with DNA. Similarly, residue 145 in Ec-CRP was proposed to play a key role in pushing the αD helix forward to better interacting with DNA, and its substitution to a bulkier Val group would further enhance this effect.\textsuperscript{37,38} Thus, these two signature amino acid residues control the hinge bending and the helical reorientation of constitutively active Ec-CRP* mutants, thus achieving their c-AMP independency. These two c-AMP-independent substitutions are indeed also found to be present in XcCLP (Asp162 and Val165). Furthermore, it has also been discovered that c-AMP binding induces domain-domain interactions, namely, the formation of a hydrogen bond between residue Tyr63 in the β strand of the c-NMP binding domain and residue Glu171 in the αE helix of the DNA-binding domain.\textsuperscript{38} Such distal interactions have been found to play a role too in the Ec-CRP* stability, thus optimizing the c-AMP independency. This is also the case for XcCLP, in which the Cα-Cβ torsional angle of Glu192 has been found to rotate approximately 180° to allow for hydrogen bond formation between residues Tyr79 and Glu192. However, the nearby Gln175 in Ec-CRP* has been altered to an Arg195, which allows for even more extensive interactions amongst the side chain atoms of residues Arg195, Glu192 and Tyr79. Thus, this interaction between the aforementioned side chain atoms allows for the interconnection of the N-terminal c-NMP domain and the C-terminal DNA-binding domain of XcCLP, leading to a more “closed” and active conformation.

Altogether, these factors may award XcCLP intrinsic DNA binding capability and cause XcCLP to form a more symmetric dimer in its apo-form. In contrast, Ec-CRP does not form a symmetrical dimer either in the apo-form or in the holo-form with a presence of a 2:1 molar ratio of c-AMP over Ec-CRP in the crystal structure.\textsuperscript{3,34,42} It has also been proposed in a molecular dynamics study that one c-AMP-Ec-CRP monomer binds to DNA first to facilitate the binding of another adjacent c-AMP-Ec-CRP monomer, thus forming a symmetrical c-AMP-Ec-CRP-DNA ternary complex.\textsuperscript{4} A solution structure of Ec-CRP was also recently solved by the multiple heteronuclear dimension NMR, and the two monomers were indeed found to be symmetrical.\textsuperscript{6,46} In XcCLP, it already adopts a highly symmetrical dimer, which may lead to the enhancement of the binding rate of CLP toward DNA (Fig. 1).

### Why c-di-GMP Can Inhibit the Binding of XcCLP from Cognate DNA

Since the crystallization of the XcCLP-c-di-GMP complex is still undergoing, a molecular modeling was used to dock c-di-GMP onto XcCLP, and extensive studies on the binding strength between the c-di-GMP and the XcCLP mutants were made to confirm that the complex model is at least partially correct and is consistent with the available mutant data.\textsuperscript{29} C-di-GMP was predicted to situate in a tight space between the αC and αD helices of XcCLP. The binding of c-di-GMP appears to tilt the αD helix and the DNA binding αF helix upward to a more “open” conformation, thus preventing XcCLP from interacting with DNA, and leading to the release of XcCLP from DNA (Fig. 5C in the article by Chin et al.).\textsuperscript{29} C-di-GMP seems to interact well with residues in the αC and αD helices of XcCLP. Significantly, Arg154 in helix αC was predicted to interact extensively with c-di-GMP, with its guanido group forming salt bridges with the phosphate group of c-di-GMP and with the side chain atoms of Asp170 in the αD helix. The side chain of Arg154 is also fully extended and stacks well upon the guanine base of c-di-GMP. Interestingly, this residue was also found to be highly conserved in the c-di-GMP binding CRP family as revealed in a multiple sequence alignment study.\textsuperscript{30} This result indicates that this amino acid may play a key role in binding c-di-GMP, for substitution of this amino acid reduced the c-di-GMP binding strength eight-fold.\textsuperscript{29}

Thus, we have discovered three key amino acids that are responsible for intrinsic DNA binding (Asp162, Val165 and Glu99) and one amino acid (Arg154) possibly responsible for the c-di-GMP binding capability of the XcCLP protein (Fig. 1). These four signature amino acids may allow one to screen whether a protein of the CRP family is a c-di-GMP binding effector or a c-AMP binding effector. The only available crystal structure of XcCLP\textsuperscript{29} has beautifully illustrated the significance of these amino acids in conferring XcCLP its intrinsic DNA binding capability, as well as in the negative regulation mode executed by c-di-GMP.\textsuperscript{28,30}

### Possible Evolutionary Pathway from CRP to CLP

Currently there is no data to imply whether CRP or CLP evolved first. Since CRP is widely present in almost every living cell except those of the xanthomonads, for the purpose of this discussion, I shall assume that CRP is the more primitive homolog of the two, and that CRP eventually evolved to form a functional CLP. Backtracking to before the evolution of a functional CLP, there may have existed a wild-type CRP molecule with the ability to bind well to upstream promoter DNA in the presence of c-AMP, thus helping RNA polymerase to transcribe downstream genes. However, in the absence of c-AMP, CRP changes conformation and loses its ability to bind to DNA (left of Fig. 1). Through evolution, CRP homolog may have come about that could bind to DNA even in the absence of c-AMP. This hypothesis is backed by the discovery of so-called constitutive CRP*\textsuperscript{36,42}, which can bind well to cognate DNA without the presence of c-AMP. A primary function of CRP is to enhance the expression of enzymes involved in metabolizing alternative carbon sources. However, such
functionality would not be required for a plant pathogen such as Xc, for Xc inhabits the crucifer xylem, which contains plenty of primary carbon sources, i.e., glucose. Without the need to metabolize alternate carbon sources, the reengineering CRP would have been prompted through the alternation of several key signature amino acids (highlighted in orange in Fig. 1), thus forming a new CLP protein whose primary function is now to control the interactions between Xc and its plant host.
This scenario process may have occurred slowly through several stages, via various CRP-like mutants, ultimately evolving into the final active CLP that can bind to DNA constitutively with strong affinity. The presence of the R154 residue allows CLP to bind to c-di-GMP, thus dissociating CLP from DNA. Therefore, two important proteins in the CRP family are defined: CRP, which is able to respond to c-AMP and binds to cognate DNA to regulate downstream genes in a positive mode, and CLP, which is able to bind to DNA intrinsically, but is dissociated from cognate DNA via negative regulation by c-di-GMP when the concentration of c-di-GMP is high.

Future Perspective

While our XcCLP structure has answered several important questions, there are still many more problems remaining to be explored: (1) The XcCLP-c-di-GMP complex structure needs to be solved in order to confirm the novel c-di-GMP binding mode of XcCLP, which contains the consensus 150RxRXXR motif but only a partial DxxS sequence in the 160D/NxxSxxG motif required for c-di-GMP binding in the PilZ proteins; (2) XcCLP usually binds to cognate DNA asymmetrically. Therefore different CLP-DNA complex structures need to be determined to understand how XcCLP interacts with asymmetric DNA sequences, and to know whether their binding strength is correlated with their downstream gene expression profile; (3) The GGDEF domain is found to fuse with some proteins with a c-NMP binding domain, such as Xc0249 in Xanthomonas campestris. It is likely that c-AMP and c-di-GMP messenger networks are integrated, a topic which has been actively pursued recently. It is of considerable interest to solve the c-NMP-GGDEF protein structure in order to learn whether binding of c-AMP activate or deactivate the diguanylate cyclase activity of the fused GGDEF domain; and (4) XcCLP may serve as a good anti-virulence target for developing novel antibiotics without eliciting drug resistance. Binding screening of small molecule libraries with XcCLP is necessary to achieve this goal.

Acknowledgements

This work is supported by the Academic Excellence Pursuit grant from the Ministry of Education, and by the National Science Council, Taiwan, ROC (grants 97-2113-M005-005-MY3) to S.H. Chou. We appreciate the services of the Structural Genomics Databases provided by the GMBD Bioinformatics Core (www.tbi.org.tw), NRPGM, Taiwan, ROC. We would like to thank the Core Facilities for Protein X-ray Crystallography in the Academia Sinica, Taiwan for their assistance in crystal screening, the National Synchrotron Radiation Research Center (NSRRC) in Taiwan, and the SPring-8 Synchrotron facility in Japan for their assistance in X-ray data collection. The National Synchrotron Radiation Research Center is a user facility supported by the National Science Council, Taiwan, ROC and the Protein Crystallography Facility is supported by the National Research Program for Genomic Medicine, Taiwan, ROC. We also thank Bonnie Levitt for her help in reviewing this manuscript.

References

1. Pesavento C, Hengge R. Bacterial nucleotide-based second messengers. Curr Opin Microbiol 2009; 12:170-6.
2. Bondford J, Harman J. Cyclic AMP in prokaryotes. Microbiol Rev 1992; 56:100-22.
3. Passner JM, Schultz SC, Steitz TA. Modeling the cAMP-induced allosteric transition using the crystal structure of CAP-CAMP at 2.1 Å resolution. J Mol Biol 2000; 304:847-59.
4. Popowycz N, Talbot SR, Tondicelli M, Ebeling RH, Kalodmos CG. Structural basis for cAMP-mediated allosteric control of the catabolite activator protein. Proc Natl Acad Sci USA 2009; 106:6927-32.
5. Sharma H, Yu S, Kong J, Wang J, Steitz TA. Structure of apo-CAP reveals that large conformational changes are necessary for DNA binding. Proc Natl Acad Sci USA 2009; 106:16604-9.
6. Tseng SR, Kalodimos CG. Dynamic activation of an allosteric regulatory protein. Nature 2002; 462:368-72.
7. Ryan R, Fouhy Y, Lucey JF, Ryan RP. The HD-GYP domain and c-di-GMP binding protein. Proteins 2007; 66:282-90.
8. Jenal U, Malone J. Mechanisms of cyclic-di-GMP signaling in bacteria. Curr Opin Microbiol 2005; 8:629-39.
9. Tesfaye B, Radmilovic S, He YW, et al. Cell-cell signaling in Xanthomonas campestris adopts a PilZ domain-like structure without a c-di-GMP switch. Proc Natl Acad Sci USA 2009; 106:6927-32.
10. Chen HC, Lee YC, Tu ZL, Chen CH, Tseng YH, Yang JM, Fouhy Y, Lucey JF, Ryan RP. The HD-GYP domain of the global regulator Clp defines a new class of cyclic-di-GMP responsive transcription factor. Mol Microbiol 2008; 69:376-89.
11. Schirmer T, Jenal U. Structural and mechanistic determinants of c-di-GMP signalling. Nat Rev Microbiol 2009; 7:724-5.
12. Leduc JL, Roberts GP. Cyclic-di-GMP allosterically inhibits the CRP-like protein (Cpl) of Xanthomonas axonopodis pv. citri. J Bacteriol 2009; 191:7121-2.
13. Chin KH, Lee YC, Tu ZL, Chen CH, Tseng YH, Yang JM, et al. The c-AMP receptor-like protein Cpl is a novel c-di-GMP receptor linking cell-cell signalling to virulence gene expression in Xanthomonas campestris. J Mol Biol 2010; 396:646-62.
14. Zhao F, He YW, Wu DH, Swarup S, Zhang LH. The cyclic nucleotide monophosphate domain of Xanthomonas campestris global regulator Cpl defines a new class of cyclic-di-GMP effectors. J Bacteriol 2010; 192:1020-9.
15. Zhang LH. A novel c-di-GMP effector linking intracellular virulence regulon to quorum sensing and hypo/anaerobiosis. Virulence 2010; 1:391-4.
32. He YW, Ng AYJ, Xu M, Lin K, Wang LH, Dong YH, et al. *Xanthomonas campestris* cell-cell communication involves a putative nucleotide receptor protein Clp and a hierarchical signalling network. Mol Microbiol 2007; 64:281-92.

33. Gomelsky M. C-di-GMP-binding CRP-like protein: a spectacular new role for a veteran signal transduction actor. J Bacteriol 2009; 181:6785-7.

34. de Grece-Lagard V, Glaser P, Lejeune P, Simeiro O, Barber CE, Daniels MJ, et al. A *Xanthomonas campestris* pv. *Campestris* protein similar to catalytically active factor is involved in regulation of phytopathogenicity. J Bacteriol 1990; 172:5877-83.

35. Chu SY, Toedova M, Gilliland GL, Gorshkova I, Shi Y, Wang S, et al. The structure of the T127L/S128A mutant of cAMP receptor protein facilitates promoter site binding. J Biol Chem 2001; 276:11230-6.

36. Harman JG, McKenney K, Peterkofsky A. Structure-function analysis of three cAMP-independent forms of the cAMP receptor protein. J Biol Chem 1986; 261:16352-8.

37. Kolb A, Bushy S, Bue H, Gargas S, Adhya S. Transcriptional regulation by cAMP and its receptor protein. Annu Rev Biochem 1993; 62.

38. Kim J, Adhya S, Gargas S. Allosteric changes in the cAMP receptor protein of *Escherichia coli*: Hinge reorientation. Proc Natl Acad Sci USA 1992; 89:9700-4.

39. Eiting M, Hagelukien G, Schubert WD, Heinz DW. The mutation G145S in PrfA, a key virulence regulator of *Listeria monocytogenes*, increases DNA-binding affinity by stabilizing the HTH motif. Mol Microbiol 2005; 56:433-46.

40. Harman JG. Allosteric regulation of the cAMP receptor protein. Biochim Biophys Acta 2001; 1547:1-17.

41. McKay DB, Weber IT, Stetza TA. Structure of catabolite gene activator protein at 2.9 Å resolution. J Biol Chem 1982; 257:9518-24.

42. Gallagher DT, Smith N, Kim SK, Robinson H, Reddy PT. Profound asymmetry in the structure of the cAMP-free cAMP receptor protein (CRP) from *Mycobacterium tuberculosis*. J Biol Chem 2009; 284:8228-32.

43. Li L, Uversky VN, Dunker AK, Menoueh SO. A computational investigation of allostery in the catabolite activator protein. J Am Chem Soc 2007; 129:15668-76.

44. Kreueger S, Gorshkova I, Brown J, Hoekin J, McKenney KH, Schwarz FP. Determination of the conformations of cAMP receptor protein and its T127L/S128A mutant with and without cAMP from small angle neutron scattering measurements. J Biol Chem 1998; 273:20001-6.

45. Hsiao YM, Liao HY, Lee MC, Yang TC, Tseng YH. Clp upregulates transcription engA gene encoding a virulence factor in *Xanthomonas campestris* by direct binding to the upstream tandem Clp sites. FEBS Lett 2005; 579:3525-33.

46. Ryan RP, Fouhy Y, Lucey JF, Jiang BL, He YQ, Feng JX, et al. Cyclic di-GMP signalling in the virulence and environmental adaptation of *Xanthomonas campestris*. Mol Microbiol 2007; 63:429-42.