Commentary

The ascent of nucleotide cyclases: conservation and evolution of a theme

Cyclic nucleotides are known to regulate a number of metabolic and developmental processes in all living organisms. There are two types of nucleotide cyclases, adenylyl cyclases that produce 3′5′-cyclic adenosine monophosphate (cAMP) and guanylyl cyclases that synthesize the corresponding guanosine (cGMP) analogue. Both enzymes utilize their respective nucleotide triphosphates as substrates, and require a metal co-factor, usually Mg or Mn, for catalysis (Tang and Hurley 1998). With information available from large scale genome sequencing projects, it is apparent that cyclases can be classified into different sub-classes, based on the amino acid sequence similarities, rather than their substrate specificities (see figure 1). Such a classification reveals interesting differences amongst these enzymes as a result of divergence in their structure.

The role of cAMP in the phenomenon of catabolite repression in bacteria is well known (Cases and Lorenzo 1998). In this case, cAMP is produced as a consequence of phosphorylation of an adenylyl cyclase. The phosphorylation of adenylyl cyclase is brought about by the terminal protein in the phosphotransferase system, only in the absence of glucose. Phosphorylation leads to the activation of the enzyme and elevation of intracellular cAMP levels. Cyclic AMP then binds to the cAMP receptor protein which can activate operon(s) for utilization of other sugars. Equilogues of the *Escherichia coli* adenylyl cyclase gene have been found in other Gram-negative organisms such as *Salmonella*, *Pasteurella*, *Haemophilus*, *Vibrio* etc. and these enzymes comprise the class 1 type of adenylyl cyclases and are single copy genes (Danchin 1993).

A second class of adenylyl cyclases is represented by enzymes produced by two pathogenic organisms, *Bordetella pertussis* (Ladant and Ullman 1999) and *Bacillus anthracis* (Baillie and Read 2001), that cause whooping cough and anthrax respectively. These phylogenetically unrelated bacteria share the common feature of producing a virulence factor in the form of a secreted adenylyl cyclase toxin. This adenylyl cyclase, once within the host cell, is activated by the host cellular calmodulin, and the extremely high concentrations of cAMP produced within the host cell results in the disease symptoms of oedema. Both these toxins are similar to each other in amino acid sequence but share no similarity with the adenylyl cyclase in *E. coli*, or enzymes present in the other classes of adenylyl cyclases (Danchin 1993). Whether these toxins can complement each other is not known. The Exo Y toxin of *Pseudomonas aeruginosa* shares partial similarity with these enzymes, but is not stimulated by calmodulin (Yahr *et al* 1998).

The importance of cAMP in higher organisms was first evident in its role as a second messenger in several hormone signalling pathways. It is well known now that the mammalian adenylyl cyclases are coupled to a G-protein signalling cascade and are stimulated by the ***a***-subunit of the stimulatory G-proteins. Cloning of the cDNA and analysis of the predicted amino acid sequence of the mammalian adenylyl cyclases indicated that these were large polypeptides that showed no amino acid sequence similarity to the adenylyl cyclase from *E. coli*, or the toxin adenylyl cyclases, thereby representing a third class of cyclase enzymes (Krupinski *et al* 1989; Danchin 1993). Interestingly, these adenylyl cyclases showed significant amino acid similarity to the mammalian soluble and receptor guanylyl cyclases (Thorpe and Garbers 1989), indicating the common evolution of the two enzymes. Therefore, rather than just their substrate specificity, sequence analysis appears to be a key tool for classification of adenylyl and guanylyl cyclases.

Relatively recently, attempts towards cloning nucleotide cyclases from the eubacteria *Aeromonas* (Sismeiro *et al* 1998) and *Prevotella* (Cotta *et al* 1998) revealed two new cyclase gene sequences

which could complement a cyclase-deficient strain of *E. coli*. However, the putative gene products from these species had no sequence similarity to the bacterial or the mammalian enzymes mentioned earlier, or to each other, therefore representing a fourth and fifth class of adenylyl cyclases. Obviously, when widely differing amino acid sequences, with less than 20% sequence identity, can code for a common function, no peptide sequence can be considered as representing a 'typical' nucleotide cyclase. Perhaps there is no typical cyclase 'structure' either. However, it has now become evident from a series of recent publications that a number of microbial gene products function as cyclases. The new cyclases share significant sequence similarity with the class III nucleotide cyclases, perhaps indicating that it is this class that truly represents a primitive and conserved cyclase catalytic fold.

Information available from the crystal structure of the mammalian adenylyl cyclases (Zhang *et al* 1997; Tesmer *et al* 1997, 1999 and see figure 2) helps us make some generalizations about the overall structure that could be adopted by this class of nucleotide cyclases. The region corresponding to the catalytic domain of the mammalian, membrane-bound adenylyl cyclases is composed of two similar sub-domains termed C1 and C2. Sequence and mutational analysis have shown that the C1 and C2 domains each have specific residues, conserved in most mammalian cyclases, that are critical for enzyme activity and need to be appropriately juxtaposed at the dimer interface. The critical residues are those required for either metal or nucleotide binding, with the C1 domain contributing residues required for metal binding and the C2 domain contributing residues for nucleotide binding. Since the C1 and C2 domains differ in their sequence, the mammalian adenylyl cyclase (forming C1-C2 heterodimers) effectively has a single catalytic center (Tang and Hurley 1998).

Inspection of the several microbial genomes sequences available indicates the presence of one or more genes with coded amino acid sequences similar to those of class III adenylyl cyclases and demonstrates the versatility of this catalytic domain. It appears that the cyclase domain can be incorporated as a separate functional domain in a larger polypeptide, fused to other protein domains with diverse functions. Some of the class III cyclases that have been characterized
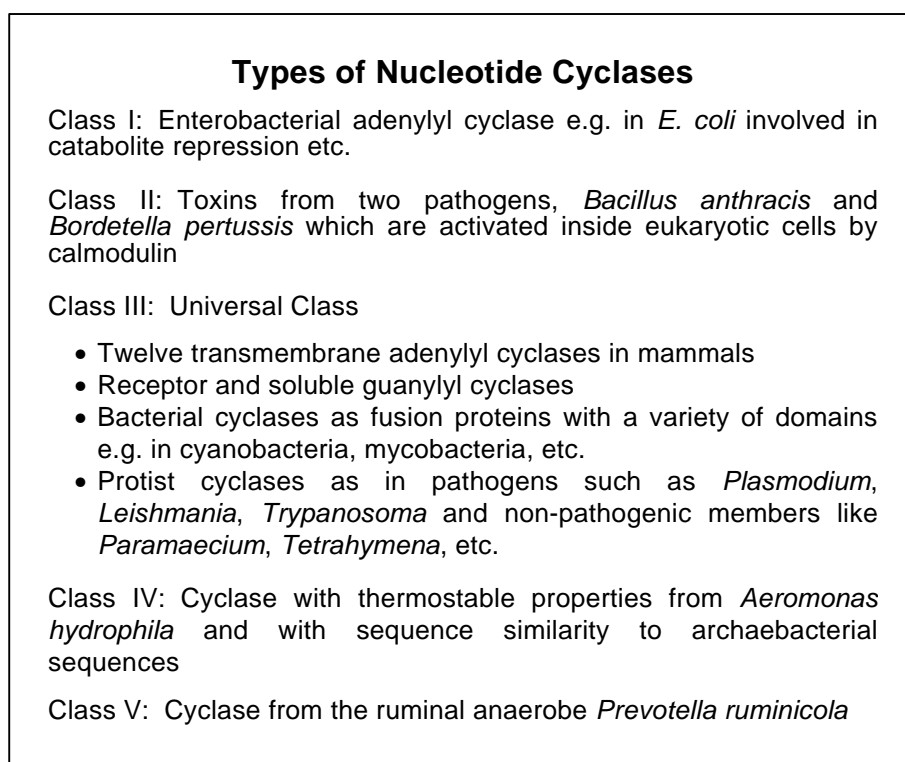

Types of Nucleotide Cyclases

Class I: Enterobacterial adenylyl cyclase e.g. in *E. coli* involved in catabolite repression etc.

Class II: Toxins from two pathogens, *Bacillus anthracis* and *Bordetella pertussis* which are activated inside eukaryotic cells by calmodulin

Class III: Universal Class

- Twelve transmembrane adenylyl cyclases in mammals
- Receptor and soluble guanylyl cyclases
- Bacterial cyclases as fusion proteins with a variety of domains e.g. in cyanobacteria, mycobacteria, etc.
- Protist cyclases as in pathogens such as *Plasmodium*, *Leishmania*, *Trypanosoma* and non-pathogenic members like *Paramaecium*, *Tetrahymena*, etc.

Class IV: Cyclase with thermostable properties from *Aeromonas hydrophila* and with sequence similarity to archaebacterial sequences

Class V: Cyclase from the ruminal anaerobe *Prevotella ruminicola*

Figure 1. Classification of nucleotide cyclases based on sequence analysis (based on Danchin 1993). Class IV and V were added recently, following the cloning of novel adenylyl cyclases from these bacteria.

from lower organisms appear to function as homodimers. The newest member to be included in this group of organisms is *Mycobacterium tuberculosis* H37Rv (Cole *et al* 1998). Analysis of the genome sequence suggests the existence of at least 14 cyclases with sequence similarity to the class III cyclases as identified by The Institute for Genomics Research (www.tigr.org). One gene product, Rv1625c, has recently been shown to possess adenylyl cyclase activity (Reddy *et al* 2001; Guo *et al* 2001). This is an interesting gene since its sequence suggests the presence of 6 transmembrane helices and one catalytic domain, which represents 'half' of the mammalian membrane bound 12-transmembrane, two-catalytic domain adenylyl cyclases (see figure 3). Further, Rv1625c can homodimerize effectively to form a protein topologically analogous to the mammalian adenylyl cyclases (unpublished observations). Inspection of the *M. tuberculosis* genome sequence does not reveal the presence of any genes that resemble any other class of adenylyl cyclases.
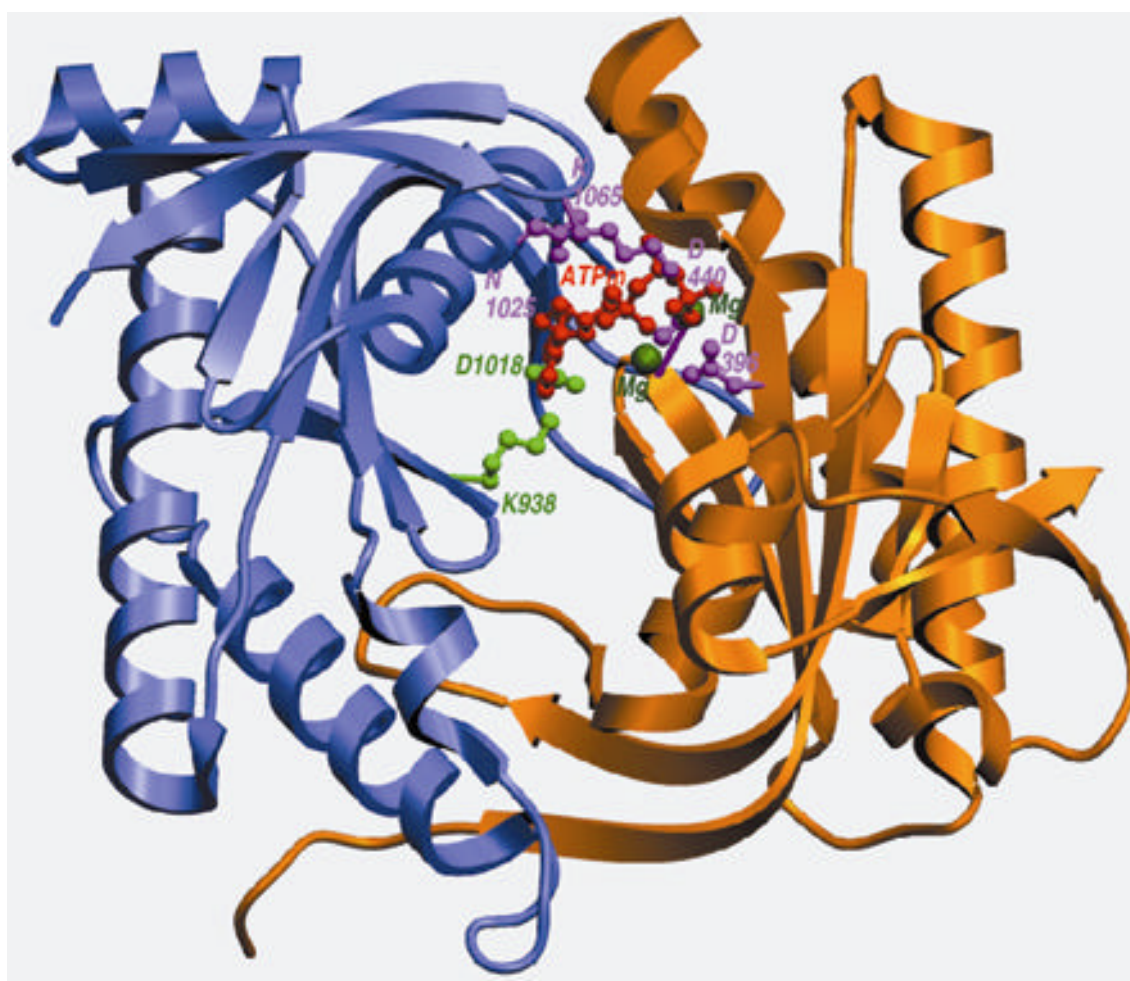

Figure 2. Crystal structure of the mammalian adenylyl cyclase. The crystal structure (PDB code 1cju) showing the C1 domain (orange) and C2 domain (violet) has been depicted using the SETOR program (Evans 1993) and shows the overall structural similarity between the two. The structure reveals their head to tail arrangement (Tesmer *et al* 1999) and the active site at the interface has been occupied by the ATP mimic, *b*-L-2′,3′-dideoxy ATP (ATPm). The two magnesium atoms are shown as spheres. Side chains of a few catalytically important residues are shown. The residues shown in green (K938 and D1018) are involved in binding ATP by interacting with its purine ring. The two metal binding aspartic acid residues (D396 and D440), the lysine (K1065) and asparagine (N1025) interacting with the phosphates and the ribose moiety respectively are shown in pink. Note that the residues interacting with ATP mainly are from the C2 domain and the metal atoms are co-ordinated by the C1 domain.

Equally interesting are the predicted topologies of some recently cloned guanylyl cyclases. Classical mammalian guanylyl cyclases are either single transmembrane receptor cyclases or cytosolic, heterodimeric, nitric oxide-activable enzymes (Barbara and Garbers 1998). However, the microbial guanylyl cyclases reported from protozoa such as *Paramecium*, *Tetrahymena* and *Plasmodium* (Linder *et al* 1999) have 22 transmembrane helices. These proteins have a N-terminal putative ATPase domain followed by the familiar 12-transmembrane spanning domain cyclase cassette seen in mammalian adenylyl cyclases. Moreover, there is a G-protein regulated 12 transmembrane guanylyl cyclase in *Dictyostelium* (Roelofs *et al* 2001a). In these gene products, it appears that a guanylyl cyclase takes on the disguise of a mammalian adenylyl cyclase.

Pathogens such as *Trypanosoma* and *Leishmania* show the presence of receptor-type adenylyl cyclases where the class III cyclase domain is fused to a putative extracellular domain (Taylor *et al* 1999; Sanchez *et al* 1995). It therefore represents yet another topology for an adenylyl cyclase, and the search for ligands for these receptors, if any, is ongoing. Note that this topology seems reminiscent of the mammalian receptor guanylyl cyclases. A class III adenylyl cyclase in *Stigmatella* is fused to a two-component receiver domain (Coudart-Cavalli *et al* 1997). In addition, a whole spectrum of cyclase catalytic domain fusion proteins are present in the cyanobacteria. Here the nucleotide cyclase domain exists along with other protein domains in the same polypeptide, such as the cGMP-binding GAF domain, the two-component histidine kinase and receiver domains and DNA-binding transcription regulator domains (Katayama and Ohmori 1997). Curiously, it is amongst the cyanobacterial enzymes that we come across the one and only prokaryotic representative of a guanylyl cyclase so far (Ochoa de Alda *et al* 2000). A recently cloned adenylyl cyclase from human sperm has been shown to be a sensor for bicarbonate and is most similar in sequence to the cyanobacterial enzymes (Chen *et al* 2000).
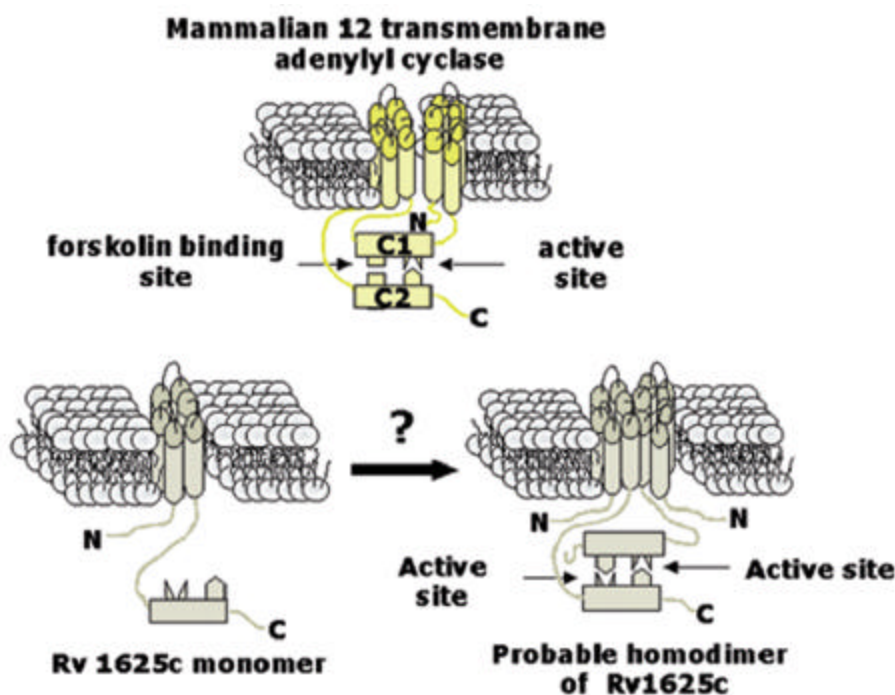

Figure 3. Topology of the mammalian 12-transmembrane adenylyl cyclases and Rv1625c. The mammalian membrane bound adenylyl cyclases have two stretches of 6-transmembrane helices alternating with two cytoplasmic catalytic loops named C1 and C2. The C1–C2 interface gives rise to a single active site and a pseudo-symmetric site that binds forskolin. The Rv1625c gene product has a topology that is half of the mammalian membrane adenylyl cyclases and has been shown to form stable homodimers giving rise to a mammalian-cyclase-like topology with two catalytic centres.

Thus, it is becoming clear that a majority of living organisms have at least one cyclase that is similar to the class III cyclases and the other classes of cyclases are restricted to only a few species (see figure 4). However, our current analysis is based on the available genome sequence from over 80 bacteria and other pathogenic organisms. Does this suggest that the catalytic domain representing the class III cyclases has been retained by these organisms and has evolved to the sequences seen in mammals?

The class III cyclase domain seems extremely flexible, in that it is able to heterodimerize with other similar domains; the added complexity could lead to activation or inactivation of the cyclases depending on the presence or absence of critical catalytic residues. Thus, Rv1625c, shown to possess both C1 and C2 like residues, forms stable and catalytically active homodimers (Guo *et al* 2001 and our unpublished observations). In contrast, the *Leishmania* enzymes are inactivated upon dimerization (Taylor *et al* 1999). This exquisite regulatory potential could have led to the parallel evolution of the catalytic domain of class III cyclases into C1-like or C2-like domains, with the coming together of C1- and C2-like domains resulting in a fruitful union, in the form of an active cyclase, in higher organisms. This is seen in the soluble guanylyl cyclases that possess the C1-like *a*-subunit and the C2-like *b*-subunit that heterodimerize to form the holoenzyme. Perhaps the 12-transmembrane containing adenylyl cyclases have taken this further by fusing the C1- and C2-like catalytic domains in a single polypeptide chain. However, these enzymes are activated only when C1–C2 interaction is stabilized by a G-protein or the diterpene forskolin (Tang and Hurley 1998).

Given the sequence similarity between mammalian adenylyl and guanylyl cyclases, it appears that the same class III fold can adapt to utilize ATP or GTP as substrate. The substrate selection seems to
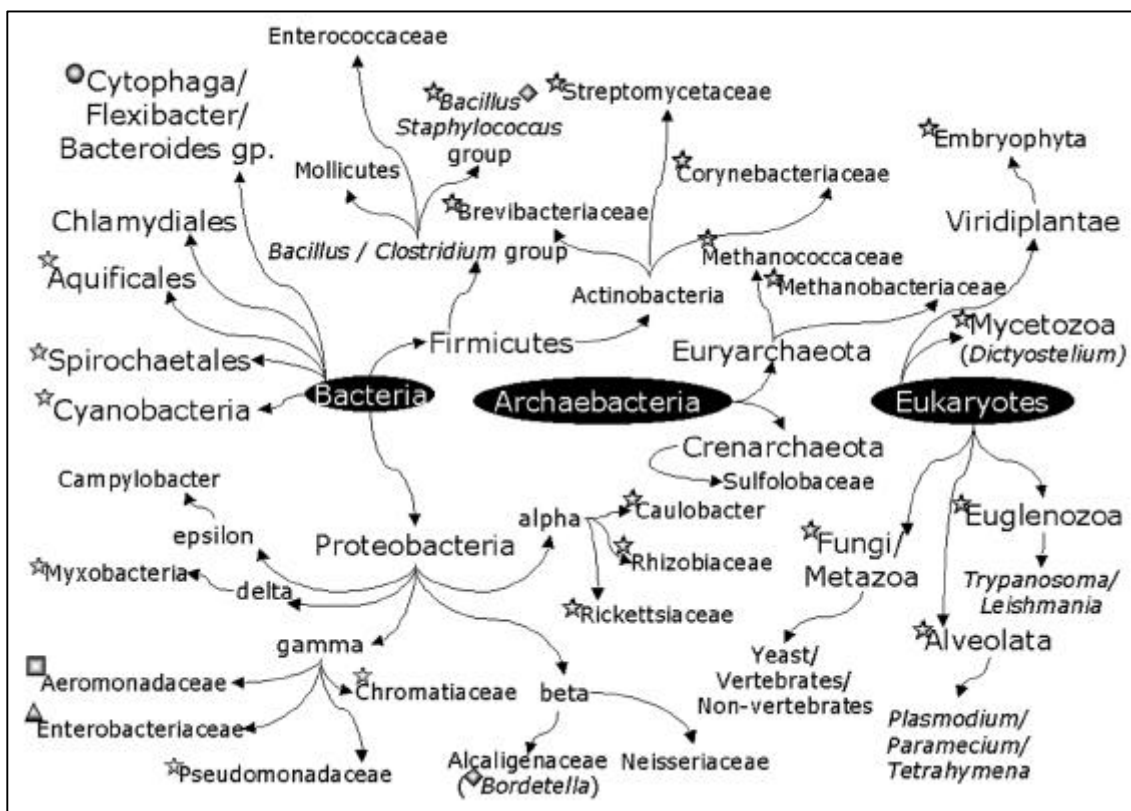

Figure 4. Distribution of nucleotide cyclases in various organisms. The figure is adapted from and based on the Taxonomy information provided by NCBI (http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/genom_tree_cgi and http://www.ncbi.nlm.nih.gov/Taxonomy/taxonomyhome.html/). Symbols are represented as follows: triangle, class I cyclase; diamond, class II cyclase; star, class III cyclase; square, class IV cyclase; circle, class V cyclase. Note the wide distribution of class III cyclases.

depend entirely on two critical residues. Thus, mammalian adenylyl cyclases characterized so far have so far been shown to require a lysine and an aspartate at the catalytic center, while the mammalian guanylyl cyclases possess glutamate and cysteine residues at the corresponding positions. Indeed, it is possible to convert a guanylyl cyclase to an adenylyl cyclase by mutation of 'guanylyl cyclase-like' residues to the corresponding 'adenylyl cyclase-like' residues (Sunahara *et al* 1998; Tucker *et al* 1998). This specific residue requirement does not hold true for microbial cyclases. Some microbial cyclases do not show a clear conservation of these residues and other residues such as histidine (Roelofs *et al* 2001b), threonine, alanine (Roelofs *et al* 2001a), serine (Linder *et al* 1999) are tolerated in place of the aspartate/cysteine residues. It is possible that evolutionary drifting has fixed a cysteine residue in the higher eukaryotic guanylyl cyclases characterized so far.

Since the class III enzymes are present in all three major classes of life forms, the evolution of this protein fold must predate the splitting of the major life forms (figure 4). It is probable that certain organisms lost this class of cyclases during their evolution (e.g. *E. coli*). The crystal structure of the mammalian adenylyl cyclase has shown that they possess a fold similar to DNA polymerases (Tesmer and Sprang 1998), and it is unknown if this unexpected similarity has resulted due to divergent evolution from a common ancestor. Thus, the existence of the class III domain in multiple topologies, and in fusion with other domains, reveals the continuous tinkering this domain has undergone during evolution. Cyclases therefore appear to be ideal model systems to study the evolution of proteins through domain swapping and association with additional motifs, ultimately resulting in complex mechanisms for the regulation of cellular function.

Acknowledgements

References

Baillie L and Read T D 2001 *Bacillus anthracis*, a bug with attitude!; *Curr. Opin. Microbiol.* **4** 78–81

Barbara J W and Garbers D L 1998 Guanylyl cyclases: approaching year thirty; *Trends Endocrinol. Metab.* **9** 213–219

Cases L and Lorenzo V 1998 Expression systems and physiological control of promoter activity in bacteria; *Curr. Opin. Micobiol.* **1** 303–310

Chen Y, Cann M J, Litvin T N, Iourgenko V, Sinclair M L, Levin L R and Buck J 2000 Soluble adenylyl cyclase as an evolutionarily conserved bicarbonate sensor; *Science* **289** 625–628

Cole S T, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon S V, Eiglmeier K, Gas S, Barry C E 3rd, Tekaia F, Badcock K, Basham D, Brown D, Chillingworth T, Connor R, Davies R, Devlin K, Feltwell T, Gentles S, Hamlin N, Holroyd S, Hornsby T, Jagels K, Barrell B G *et al* 1998 Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence; *Nature* (*London*) **393** 537–544

Cotta M A, Whitehead T R and Wheeler M B 1998 Identification of a novel adenylate cyclase in the ruminal anaerobe *Prevotella ruminicola* D13d; *FEMS Microbiol. Lett.* **164** 257–260

Coudart-Cavalli M P, Sismeiro O and Danchin A 1997 Bifunctional structure of two adenylyl cyclases from the myxobacterium *Stigmatella aurantiaca*; *Biochimie* **79** 757–767

Danchin A 1993 Phylogeny of adenylyl cyclases; *Adv. Second Messenger Phosphoprotein Res.* **27** 109–162

Evans S V 1993 SETOR – Hardware lighted three-dimensional solid model representations of macromolecules; *J. Mol. Graph.* **11** 134–138

Guo Y L, Seebacher T, Kurz U, Linder J U and Schultz J E 2001 Adenylyl cyclase Rv1625c of *Mycobacterium tuberculosis*: a progenitor of mammalian adenylyl cyclases; *EMBO J.* **20** 3667–3675

Katayama M and Ohmori M 1997 Isolation and characterization of multiple adenylate cyclase genes from the cyanobacterium *Anabaena* sp. Strain PCC 7120; *J. Bacteriol.* **179** 3588–35931

Krupinski J, Coussen F, Bakalyar H A, Tang W J, Feinstein P G, Orth K, Slaughter C, Reed R R and Gilman A G 1989 Adenylyl cyclase amino acid sequence: possible channel- or transporter-like protein; *Science* **244** 1558–1564

Ladant D and Ullman A 1999 *Bordetella pertussis* adenylate cyclase: a toxin with multiple talents; *Trends Microbiol.* **7** 172–176

Linder J U, Engel P, Reimer A, Kruger T, Plattner H, Schultz A and Schultz J E 1999 Guanylyl cyclases with the topology of mammalian adenylyl cyclases and an N-terminal P-type ATPase-like domain in *Paramecium*, *Tetrahymena* and *Plasmodium*; *EMBO J.* **18** 4222–4232

Ochoa de Alda J A G, Ajlani G and Houmard J 2000 *Synechocystis* strain PCC 6803 *cya2*, a prokaryotic gene that encodes a guanylyl cyclase; *J. Bacteriol.* **B182** 3839–3842

Reddy S K, Kamireddi M, Dhanireddy K, Young L, Davis A and Reddy P T 2001 Eukaryotic like adenylyl cyclases in *Mycobacterium tuberculosis* H37Rv: cloning and characterization; *J. Biol. Chem.* **276** 35141–35149

Roelofs J, Snippe H, Kleineidam R G and Van Haastert P J 2001a Guanylate cyclase in *Dictyostelium discoideum* with the topology of mammalian adenylate cyclase; *Biochem. J.* **15** 354697–354706

Roelofs J, Meima M, Schaap P and Van Haastert P J M 2001b The *Dictyostelium* homologue of mammalian adenylyl cyclase encodes a guanylyl cyclase; *EMBO J.* **20** 4341–4348

Sanchez M A, Zeoli D, Klamo E M, Kavanaugh M P and Landfear S M 1995 A family of putative receptor-adenylate cyclases from *Leishmania donovani*; *J. Biol. Chem.* **270** 17551–17558

Sismeiro O, Trotot P, Biville F, Vivares C and Danchin A 1998 *Aeromonas hydrophila* adenylyl cyclase 2: a new class of adenylyl cyclases with thermophilic properties and sequence similarities to proteins from hyper-thermophilic archaebacteria; *J. Bacteriol.* **180** 3334–3349

Sunahara R K, Beuve A, Tesmer J J G, Sprang S R, Garbers D L and Gilman A G 1998 Exchange of substrate and inhibitor specificities between adenylyl and guanylyl cyclases; *J. Bio. Chem.* **273** 16332–16338

Tang W J and Hurley J H 1998 Catalytic mechanism and regulation of mammalian adenylyl cyclases; *Mol. Pharmacol.* **54** 231–240

Taylor M C, Muhia D K, Baker D A, Mondragon A, Schaap P B and Kelly J M 1999 *Trypanosoma cruzi* adenylyl cyclase is encoded by a complex multigene family; *Mol. Biochem. Parasitol.* **104** 205–217

Tesmer J J G and Sprang S R 1998 The structure, catalytic mechanism and regulation of adenylyl cyclase; *Curr. Opin Struct. Biol.* **8** 713–719

Tesmer J J G, Sunahara R K, Gilman A G and Sprang S R 1997 Crystal structure of the catalytic domains of adenylyl cyclase in a complex with Gs$_a$.GTPγS; *Science* **278** 1907–1916

Tesmer J J G, Sunahara R K, Johnson R A, Gosselin G, Gilman A G and Sprang S R 1999 Two-metal-ion catalysis in adenylyl cyclase; *Science* **285** 756–760

Thorpe D S and Garbers D L 1989 The membrane form of guanylate cyclase. Homology with a subunit of the cytoplasmic form of the enzyme; *J. Biol. Chem.* **264** 6545–6549

Tucker C L, Hurley J H, Miller T R and Hurley J B 1998 Two amino acid substitutions convert a guanylyl cyclase, RetGC-1 into an adenylyl cyclase; *Proc. Natl. Acad. Sci. USA* **95** 5993–5997

Yahr T L, Vallis A J, Hancock H K, Barbieri J T and Frank D W 1998 Exo Y, an adenylate cyclase secreted by the *Pseudomonas aeruginosa* type III system; *Proc. Natl. Acad. Sci. USA* **95** 13899–13904

Zhang G, Liu Y, Ruoho A E and Hurley J H 1997 Structure of the adenylyl cyclase catalytic core; *Nature (London)* **386** 247–253

AVINASH R SHENOY
N SRINIVASAN[†]
SANDHYA S VISWESWARIAH*
Department of Molecular Reproduction,
Development and Genetics,
and
[†]*Molecular Biophysics Unit,*
Indian Institute of Science,
Bangalore 560 012, India

(Email, sandhya@mrdg.iisc.ernet.in)