1. ABSTRACT

This review explores the origins, diversity and functions of guanylyl cyclases in cellular organisms. In eukaryotes both cGMP and cAMP are produced by the conserved class III cyclase domains, while prokaryotes use five more unrelated catalysts for cyclic nucleotide synthesis. The class III domain is found embedded in proteins with a large variety of membrane topologies and other functional domains, but the vertebrate guanylyl cyclases take only two forms, the receptor guanylyl cyclases with single transmembrane domain and the soluble enzymes with heme binding domain. The invertebrates additionally show a soluble guanylyl cyclase that cannot bind heme, while the more basal metazoans may lack the heme binding enzymes altogether. Fungi, the closest relatives of the metazoans, completely lack guanylyl cyclases, but they appear again in the Dictyostelids, the next relative in line. Remarkably, the two Dictyostelid guanylyl cyclases have little in common with the vertebrate enzymes. There is a soluble guanylyl cyclase, which shows greatest sequence and structural similarity to the vertebrate soluble adenylyl cyclase, and a membrane-bound form with the same configuration as the dodecahelical adenylyl cyclases of vertebrates. There is a difference, the pseudosymmetric C1 and C2 catalytic domains have swapped position in the Dictyostelium enzyme. Unlike the vertebrate guanylyl cyclases, the Dictyostelium enzymes are activated by heterotrimeric G-proteins. Swapped C1 and C2 domains are also found in the structurally similar guanylyl cyclases of ciliates and apicomplexans, but these enzymes additionally harbour an amino-terminal ATPase module with ten transmembrane domains. G-protein regulation could not be demonstrated for these enzymes. Higher plants lack class III cyclase domains, but an unexplored wealth of guanylyl cyclases is present in the green alga Chlamydomonas. Progenitors of all structural variants of the eukaryote guanylyl cyclases are found among the prokaryote adenylyl cyclases. This and the close similarity of many guanylyl cyclases to adenylyl cyclases suggests a paraphyletic origin for the eukaryote enzymes with multiple events of conversion of substrate specificity.

2. INTRODUCTION

Stimulus-induced production of the second messenger cGMP controls a broad spectrum of physiological responses in vertebrates, such as vision, olfaction, smooth muscle relaxation and homeostasis of electrolyte and fluid levels in blood and intestine. The cGMP-mediated functions show only limited overlap with those mediated by the chemically related messenger cAMP. Metazoans have developed sets of proteins with different membrane topologies and domain architectures for synthesis of either cAMP or cGMP. As a consequence, the activation mechanisms of guanylyl cyclases bear little in common with those of adenylyl cyclases. The strict separation into architecturally different forms is quite remarkable in view of the fact that the catalytic domains of adenylyl- and guanylyl cyclases are very similar and can be interconverted by substitution of only a few amino-acids.

cGMP has also been implicated in controlling motility in protists, such as the Alveolates and Dictyostelids. These organisms show a novel repertoire of architectural designs for guanylyl cyclases, accompanied by different modes of regulation. The ongoing and finished sequencing projects of many protist genomes are also uncovering new guanylyl cyclase forms.
Guanylyl cyclase diversity

Figure 1. Alignment of cyclase catalytic domains. The catalytic domains of the mammalian ACV C1 domain and ACII C2 domains, mammalian RetGC, *Dictyostelium discoideum* ACG, *Chlamydomonas reinhardtii* gene genie.777.3 and *Arabidopsis thaliana* AtGC1 were aligned using Clustal W (146). Residues essential for substrate binding and catalysis are highlighted. Green: residues that form hydrophobic pocket; red/yellow: residues that determine guanine/adenine specificity; blue: residues that coordinate the Mg$^{2+}$ ions; amber: interaction with forskolin; magenta: residues that stabilize the transition state; tan: residues that interact with Ppi. Genbank accession: CfACV: AAC32726; RrACII A41541; DdACG: Q03101; HsRetGC: NP_000171; AtGC1: AAM51559. The genie.777.3 sequence can be found at http://genome.jgi-psf.org/chlre1/chlre1.home.html

This review summarizes the evidence for the existence of cGMP signaling in the eight major group of eukaryotes and in the prokaryotes and compares the structure and function of their guanylyl cyclases.

3. ANIMALS

3.1. The cyclase catalytic domain

In all characterized eukaryote guanylyl- and adenyl cyclases, cyclic nucleotide synthesis is carried out by the conserved class III cyclase domain. The crystal structure of the catalytic core of the G-protein stimulated vertebrate adenylyl cyclase is resolved and this also provides insight in the catalytic mechanism of the guanylyl cyclases. G-protein activated adenylyl cyclases harbour two sets of six transmembrane helices that are separated by the cyclase C1 domain and followed by the cyclase C2 domain. These domains form an antiparallel dimer with ATP binding to the dimer interface. Potentially, two ATP molecules, each associated with two Mg$^{2+}$ ions can be bound. However, the C1-C2 dimer accommodates only one ATP, because the C2 domain lacks two aspartate residues that should coordinate the Mg$^{2+}$ ions, while C1 lacks an arginine and asparagine pair that interacts with the phosphoryl moiety of the same ATP and stabilizes the transition state (1-4). The adenine ring interacts with a conserved lysine, a glutamine and an aspartate residue. Guanylyl cyclases carry a glutamate, an arginine and a cysteine at the equivalent positions (figure 1). Mutagenesis of these residues into their adenylyl cyclase counterpart alters the substrate preference of both soluble- and receptor guanylyl cyclase from GTP to ATP (5, 6). For those adenylyl- and guanylyl cyclases that harbour a single catalytic domain and form homodimers, all essential residues have to be present in the single domain. In these enzymes two symmetric ATP or GTP binding sites are created (7, 8). The C1-C2 dimer accommodates the diterpene forskolin at the position of the degenerate ATP binding site. Forskolin is a well-known activator of mammalian adenylyl cyclases, and it does so by stabilizing the association of the C1 and C2 monomers. It interacts with several hydrophobic residues in a deep hydrophobic cleft, and with a serine residue that forms a hydrogen bond with the 7-acetyl group of forskolin. This serine replaces one of the Mg$^{2+}$ binding aspartates in C2 (1, 2). Gs, the physiological activator, also binds simultaneously to the C1 and C2 monomers and activates catalysis by stabilizing the catalytically competent juxtaposition of the monomers (1).

3.2. Vertebrate cGMP signaling, a brief survey

There are two topologically distinct guanylyl cyclases in mammalian cells, the receptor guanylyl cyclases and the soluble guanylyl cyclases, neither of which is activated by G-proteins. The receptor guanylyl cyclases are integral transmembrane proteins with a single transmembrane helix that separates the external ligand binding domain from a carboxyl-terminal region that contains a protein kinase homology domain (KHD), a coiled-coil (CC) domain and a single guanylyl cyclase (GC) domain (figure 2). The majority of these proteins already form dimers in the inactive state by interactions between their CC and KHD domains (9-11). Binding of the ligand causes juxtamembrane closure of the extracellular domain, which reorients the intracellular KHD domain. This supposedly allows it to bind ATP, which in turn triggers a second conformational change, which results in activation of the cyclase domain (12-14). The ligands are usually peptides, such as the natriuretic peptides for GC-A and GC-B (10, 15, 16) and enterotoxins and guanylin for GC-C (17-19), or unknown odorants for GC-D (20).

The Ca$^{2+}$ regulated retinal guanylyl cyclases (RetGCs) are not (known to be) regulated by ligand; their dimerization state and therefore their activity depends on
Figure 2. Domain architectures of guanylyl- and adenylyl cyclases. The amino-acid sequences of the indicated proteins were analysed with SMART (147) for SMART and PFAM functional domains (148) and by TMHMM (149) for putative transmembrane domains. The domain architecture of all proteins printed in bold is presented at the correct relative scale. The proteins indicated between brackets contain a similar array of domains, but the relative scale may be different. Abbreviations: ANP-R: atrial natriuretic peptide receptor; KHD: kinase homology domain; AAA: ATPase family associated with various cellular activities (76); PP2C: Serine/threonine phosphatases, family 2C; RA: Ras association domain; CHASE: cyclase and histidine kinase associated extracellular sensor; HK: histidine kinase; HA: histidine kinase like ATP-ase; RR: response regulator; C39: peptidase C39 family; VGC: voltage-gated channel; BLUF: sensors of blue-light using FAD; GAF: domain present in phytochromes and cGMP phosphodiesterases. Genbank accession numbers: HsGCA: AAH63304; HsRetGC: NP_000171; S.purGC: P16065; HssolGCα: NP_000847; HssolGCβ: NP_000848; MsGC-I: AAC62238; HsAC1: NP_066939; HssAC1: NP_060887; ScCyr1: P08678; DdsGC: AAK92097; DdGCA: CAB42641; DdACA: AAA33163; DdACG: Q03101; DdAcrA: AAD50121; AtGC1: AAM51559; ZmPsiP: CAC59976; P.tetGC: PTE238859; P.falGC: PFA245435; P.falAC1: AAO64441; P.falAC2: NP_704518; T.bruESAG4: Q26721; T.cruADC1: CAA09919; L.donRacA: Q27675; EgACA: BAB85619; Syn.spp.Cya2: NP_440289; Nostoc CyaC: NP_489003.
Guanylyl cyclase diversity

interaction of each monomer with a Guanylyl Cyclase Activating Protein (GCAP), which dimerizes in the absence, but not in the presence of Ca2+ (21-23). The RetGCs nevertheless harbour the conserved ligand binding, KHD and CC domains of the other receptor guanylyl cyclases (figure 2).

The soluble guanylyl cyclases (SolGCs) form a heterodimer of two isoforms, α and β, which each carry a guanylyl cyclase domain at the C-terminus, and a region at the N-terminus that coordinates a single heme group per dimer (24-26). Two cysteine residues (Cys178 and Cys214) and a histidine (His105) on the β-isoform are particularly important for binding to and coordination of the heme group (27, 28). Similar to the C1 and C2 adenylyl cyclase domains, the α- and β- guanylyl cyclase domains are not identical. Like C2, the β-isoform lacks residues to coordinate the Mg2+ ions, while the α-isoform lacks residues that stabilize the transition state (figure 1). Therefore, also here only one catalytic site can be formed (7). SolGCs are 200-fold activated by the interaction of NO radicals with heme Fe2+ (29, 30). This activation is blocked by a specific inhibitor, 1H-[1,2,4]oxadiazolo[4,3-a]-quinoxalin-1-one (ODQ), which competes with NO for binding to heme and oxidizes the Fe2+ (31, 32). SolGCs play decisive roles in smooth muscle relaxation and are also implicated in synaptic transmission and blood platelet aggregation (26).

cGMP is detected in vertebrates by either cGMP-dependent protein kinases (PKGs) or by cGMP-gated ion channels (33, 34). PKG harbours two intramolecular cGMP binding sites, that are very similar in sequence and structure to the cAMP binding sites on the regulatory subunit of PKA (35). One copy of this canonical cGMP binding domain is present on the cGMP- (and cAMP) gated ion channels. cGMP can also bind to the GAF domain, a family of small molecule binding domains with multiple functions. The cGMP binding GAF domains are present on a number of vertebrate cyclic nucleotide phosphodiesterases, where they mediate allosteric activation of the catalytic activity by cGMP (36). There are 11 families of vertebrate phosphodiesterases, which all share the same highly conserved HD domain for cyclic nucleotide hydrolysis (37). However, the families differ in substrate selectivity, inhibitor sensitivity and the presence of GAF- and other functional domains (38, 39).

3.3. Invertebrate guanylyl cyclases

The receptor guanylyl cyclases were first found in the echinoderms (sea urchins), where they are expressed in sperm cells to detect chemotactice peptides secreted by eggs (40). More recently, they have also been identified in more basal phyla. They are very abundant in the nematode, C.elegans, where they are involved in olfaction (41-43) and in Drosophila where their function has not yet been resolved (44). A Ca2+ inhibited RetGC type receptor guanylyl cyclase was identified in another arthropod, Manduca sexta (45). The NO-stimulated soluble guanylyl cyclases are also functionally conserved in the arthropods (46, 47).

In addition to these “classical” guanylyl cyclases, the invertebrates present us with a few new configurations. The M.sexta enzyme MsGC-I, shares greatest sequence similarity with the mammalian enzyme receptor guanylyl cyclase GC-B in its cyclase and coiled-coil domain, but lacks the extracellular-, transmembrane- and kinase homology domain of GC-B (figure 2). When expressed in COS-7 cells, it is soluble and active as a homodimer (48). Another Manduca enzyme MsGC-β3 is closely related to the SolGC β-isoform. However, it lacks the two cysteine residues that are essential for heme binding (28) and it is active as an NO insensitive homodimer even in the presence of the MsGC-α isoform (49, 50). All predicted C elegans SolGC proteins also lack the essential cysteine residues for heme binding and therefore stimulation by NO (43, 51). In addition, the gene for NO synthase is absent from the C elegans genome (52). Since the Pseudocoelomate Nematoda represents a more basal metazoan form than the Eucelomate Arthropoda, this suggests that NO regulation of guanylyl cyclase activity is an evolutionary novelty of the more derived metazoans. However, this suggestion requires support from fully sequenced genomes of more basal metazoan taxa.

4. FUNGI

The fungi are a large group of non-motile organisms, which feed by adsorption of nutrients. They exist primarily as filamentous networks, hyphae, that infiltrate their substrata. The fungi are the closest relatives of the animals and some unicellular allies, with whom they form the major group of ophisthokonts (figure 3). All fungi share a highly conserved CAMP signaling pathway, that controls differentiation and pathogenicity in response to nutrient status (53, 54). The single fungal adenylyl cyclase, CYR1, is a soluble protein, that is associated to the plasma membrane by a lipid anchor (55). In addition to a single cyclase domain, CYR1 harbours a phosphatase 2C domain, a Ras Association domain and a series of leucine-rich repeats (figure 2). The latter two regions are required for binding to the monomeric G-protein Ras, which activates the enzyme. CYR1 is also activated by Gpa2, the α-subunit of a heterotrimeric G-protein (56-58). Despite, the latter resemblance to regulation of vertebrate adenylyl cyclases, the fungal cyclase domain bears greater sequence similarity to that of the Trypanosomoid receptor adenylyl cyclases than to the vertebrate enzymes (59).

cAMP is detected by a single PKA regulatory subunit, which can interact with up to three catalytic subunits (60, 61). There are two cAMP phosphodiesterases; Pde1 is a low affinity enzyme, which harbours the HxHxDH metal binding motif of the protist type II phosphodiesterases (62), while Pde2 is a high affinity enzyme, with the HD catalytic motif of the vertebrate type I PDEs (37, 63). No genes encoding either guanylyl cyclases, cGMP binding proteins or cGMP-PDEs have been found in the completed genomes of Saccharomyces cerevisiae (64) and Candida albicans (65), nor have they been reported in other fungi. This, and the paucity of fungal adenylyl cyclases contrasts strongly with the abundance of guanylyl- and adenylyl cyclases in the related metazoans and dictyostelids.
Guanylyl cyclase diversity

Figure 3. Consensus phylogeny of eukaryotes. Based on molecular and ultrastructural data, most of the presently known eukaryotes can be assigned to one of eight major groups (150, 151). Courtesy of Dr. Sandra L. Baldauf.

5. DICTYOSTELIDS

The Dictyostelids are the most thoroughly studied representatives of the amoebazoa, the closest sister group of the opisthokonts (figure 3). Most members of the amoebazoa are unicellular predators that feed on bacteria, although some, like the Entamoebas, have adopted an endoparasitic life style. In addition to a unicellular feeding stage, the Dictyostelids aggregate to form a multicellular structure in response to starvation. During this process, extracellular cAMP acts as a chemoattractant to bring starving cells together. cAMP also triggers a transient increase in cytosolic cGMP (66), and mutants with defective cGMP metabolism or detection have implicated cGMP as a regulator of the chemotactic response (67-69). The Dictyostelium adenylyl- and guanylyl cyclases show a marked similarity to the vertebrate cyclases, but structurally similar proteins often have opposite substrate specificities (figure 2). There is both an adenylyl cyclase, ACA, and a guanylyl cyclase, GCA with the same dodecahelical topology as the G-protein regulated mammalian adenylyl cyclases (70, 71). In case of GCA, the order of the C1 and C2 domains is reversed, with C2 N-terminal to C1.

The second Dictyostelium guanylyl cyclase, sGC (72), is both with regard to domain architecture and sequence similarity the closest homologue of the mammalian soluble adenylyl cyclase sAC (73, 74). Both enzymes lack putative transmembrane helices and share highly conserved C1 and C2 domains and an AAA-domain with P-loop motif (75, 76). Quite strikingly, there is no sAC in invertebrates, such as C.elegans and Drosophila (72).

In the eukaryotes, the sAC and sGC cyclase domains are most similar to those of the D.discoideum adenylyl cyclase, AcrA and the alveolate (Plasmodium, Paramecium and Tetrahymena) adenylyl cyclases. However, these enzymes are all transmembrane proteins with a single cyclase domain, and they also lack the AAA domain. As a group, sAC, sGC, AcrA and the Alveolate enzymes show much greater similarity to bacterial adenylyl cyclases, than to any other eukaryote enzyme (72, 73, 77-80).

In contrast to sAC and to all vertebrate guanylyl cyclases, sGC and GCA require a heterotrimeric G-protein, G2, for activity. However, for both proteins, the activation by G2 is indirect and may additionally involve monomeric G-proteins (81, 82). All sGC activity appears to be cytosolic, when measured in the presence of Mn2+ ions. However, with Mg2+ and the G-protein activator, GTPγS, a much lower activity is exclusively detectable in the membrane fraction, suggesting that the G-protein regulated
Guanylyl cyclase diversity

population of sGC molecules is membrane-associated (72).

_Dictyostelium_ cells express a third adenylyl cyclase, ACG, that displays the same topology as the receptor guanylyl cyclases with a large extracellular region, a single transmembrane domain and a cytosolic cyclase domain (15, 70). However, there is no coiled-coil or kinase homology domain in ACG. Furthermore, its extracellular region harbours a CHASE domain instead of the metazoan ANP-R domain (figure 2). The CHASE domain is a predicted sensor domain, that is found on some histidine kinases and bacterial adenylyl cyclases (83). ACG is active as a homodimer and harbours an intramolecular osmosensor that mediates activation of catalytic activity by high osmolarity (84). It is not yet clear whether this osmosensor is identical with the CHASE domain.

The _Dictyostelium_ genome has not yielded a mammalian type PKG, but four proteins with PKG-like cNMP binding domains were identified. Two of these proteins, PdeD and PdeE harbour two cNMP binding domains each, and a metallo-β-lactamase domain. This domain is found in a number of Zn2+ -dependent hydrolases (85) and though otherwise dissimilar from the type II cAMP-PDEs of yeast and _Dictyostelium_ (62, 86), it does carry their conserved HxHxDH motif. The metallo-β-lactamase domain of PdeD encodes a highly specific cGMP-phosphodiesterase activity and that of PdeE a less specific cAMP-PDE activity. The cNMP binding domains of PdeD mediate allosteric activation of the cGMP-PDE activity by cGMP, while PdeE is allosterically activated by cAMP (87-89). This differs markedly from the vertebrate phosphodiesterases, where the GAF domain mediates allosteric activation of hydrolytic activity (36).

The other two proteins GbpC and GbpD are also multidomain proteins and most likely the sole effectors for cGMP (90). Both contain two cGMP binding domains, a RAM domain, a RasGEF and a RasGEF-associated domain; GbpC additionally harbours a MAPKKK domain, a Ras domain and a DEP domain. The only known proteins with cNMP binding domains, that share at least some of this plethora of domains are the vertebrate EPACs, in which a RapGEF domain is activated by the associated cAMP binding domain (91). GbpC and GbpD are considered to regulate the phosphorylation status of myosin II in the chemotactic response (69).

6. PLANTS

6.1. Flowering plants

The existence of cyclic nucleotide signaling in higher plants has been controversial for many years. Compared to organisms with well-defined functions for cyclic nucleotides, the levels of both cAMP and cGMP in higher plants are very low and it has been challenging to convincingly demonstrate the presence of nucleotidyl cyclase and phosphodiesterase activities (92). The completed genome sequence of the _Arabidopsis thaliana_ does not contain sequences for the universal class III nucleotidyl cyclases (93). However, two unusual cyclase genes were identified. The _Zea mays_ gene _PsIP_ was identified as a putative adenylyl cyclase on the basis of the homology of its leucine-rich repeat region to that of the fungal ACs (94). PsiP can complement defective sugar fermentation and cAMP production in an _E.coli_ adenylyl cyclase null mutant. cAMP production is stimulated by forskolin, which also stimulates the process of pollen tube outgrowth that is dependent on PsiP (94). In view of the fact that forskolin interacts in a highly structure-specific manner at the interface of the C1 and C2 domains of not even all vertebrate cyclases (1, 2), this stimulation of PsiP, which does not harbour a class III cyclase domain is rather remarkable.

A motif search of the _Arabidopsis_ genome with conserved amino-acids in the guanylyl cyclase catalytic centre yielded a candidate protein, AtGC1, that raises cGMP levels when expressed in _E.coli_ and displays Mg2+-dependent GC activity when purified (95). AtGC1 can also produce cAMP at a three-fold lower rate. Only the posterior third of the GC catalytic domain is conserved at the N-terminus of AtGC1, which means that several residues for GTP and Me2+ binding are lacking (figures 1, 2). However, additional functionality may reside in the C-terminal region of the protein. AtGC1 is a soluble protein that harbours none of the heme- or ligand binding, coiled-coil or kinase homology domains that are common to vertebrate guanylyl cyclases (15, 95, 26).

Quite fascinatingly, plants secrete extracellular peptides, irPNP’s, that bind to antibodies against mammalian atrial natriuretic peptide (ANP) and share a small region of high sequence identity with ANPs (96). Both ANP and irPNP regulate plant solute homeostasis by inducing opening of leaf stomatal pores and ion channels in roots (97-99). These effects can be mimicked by 8Br-cGMP and are in case of the stomatal pores accompanied by an increase in cGMP levels (100).

NO, the activator of the vertebrate SolGC, also regulates important processes in plant physiology and development (101) and increases cGMP levels in spruce and tobacco (102, 103). Inhibitors of vertebrate SolGC, such as ODQ and LY83583 antagonize some NO-induced responses and 8Br-cGMP can reverse this inhibition (103-105) and in one case also mimic the effect of NO (103). If AtGC1 is the only plant guanylyl cyclase, it has to be assumed that both the effects of NO and the natriuretic peptides are mediated by auxiliary sensor proteins that interact with AtGC1. This would represent a truly remarkable case of convergent evolution.

No genes encoding either PKG or the regulatory subunit of PKA have been found in the _Arabidopsis_ genome, but the characteristic cyclic nucleotide (cNMP) binding domain of these proteins is present in a large family of plant ion-channels. These channels show a similar membrane topology as the vertebrate cyclic nucleotide gated ion channels; only the localization of the additional calmodulin binding domain is different. In the
vertebrate channels this domain is located at the N-terminus, while in the plant channels it overlaps with the cNMP binding domain at the C-terminus (106-110). Most, if not all, plant cNMP domains harbour the Ala to Ser/Thr substitution that marks cGMP specificity (35, 107, 111).

Only a few members of the plant cyclic nucleotide gated channel (CNGC) family have been investigated. AtCNGC2/DND1 mediates pathogen-induced apoptosis and displays inwardly rectified transport of K⁺, Ca²⁺ and other cations except Na⁺. Channel opening is activated equally well by 8Br-cAMP and 8Br-cGMP (112-115). AtCNGC4/HLM1 is also involved in defence to pathogens, but is permeable to both K⁺ and Na⁺. Here, cGMP triggers channel opening more efficiently than cAMP (116). The presence of cyclic nucleotide gated ion-channels is thus far the strongest genetic evidence that flowering plants use cyclic nucleotides and perhaps particularly cGMP as signaling molecules.

6.2. Chlorophyte green algae

The chlorophyte green algae are basal in the line of descent that gave rise to the flowering plants (figure 3) and the unicellular alga *Chlamydomonas* is used as a model system for plant signaling and cell biology. cAMP signaling plays a well-documented role in the mating process of *Chlamydomonas*, where it is produced when cells of opposite mating type agglutinate. cAMP is produced by an adenylyl cyclase activity, that is insensitive to regulators of vertebrate adenylyl cyclases, such as forskolin and GTPγS. The cells also display a cAMP phosphodiesterase activity, that is inhibited by IBMX, an inhibitor of vertebrate phosphodiesterase (117-119). However, no genes encoding the two activities were isolated.

A genome sequencing project for *Chlamydomonas* is underway and a simple query of the annotated gene models at http://genome.jgi-psf.org/cgi-bin/searchGM2.cgi?id=chlrl with “cyclase” yields 92 hits annotated as guanylyl cyclase. I analysed the bottom twelve of this list for sequence similarity with adenylyl- or guanylyl cyclases domains. Three genes harboured all the essential residues for metal binding and catalysis and these genes were all putative soluble guanylyl cyclases. The sequence of the catalytic domain of one of the genes, genic777.3, is presented in figure 1 and the domain architecture of all three is outlined in figure 2. ESTs encoding guanylyl cyclases and cGMP phosphodiesterases were also recognized in other chlorophyte algae (120). Although there is no information on the role of cGMP signaling in algae, the projected large number of guanylyl cyclases in these organisms suggests that it could be a major one.

7. ALVEOLATES

For three branches of the eukaryote tree, the Cercozoa, Heterokonts and Excavates (figure 3) there is no information on cGMP or cAMP signaling. The remaining two, the Alveolates and Discicristates contain major human pathogens and have received much closer scrutiny. The free-living ciliates and parasitic apicomplexans are both members of the alveolates. While the ciliates are a group of fairly innocuous fresh-water protists, the apicomplexans are a family of obligatory parasites, with the malaria parasite *Plasmodium falciparum* as its most (in)famous member. Other unsavory relatives are *Toxoplasma gondii*, a major cause of congenital birth defects and *Eimeria tenella*, the causative agent of the wide-spread poultry disease, coccidiosis.

In ciliates like *Paramecium* and *Tetrahymena*, both adenylyl- and guanylyl cyclase are regulated by ion currents across the plasmamembrane. In case of adenylyl cyclase, activation is caused by an outward K⁺ current and purified adenylyl cyclase protein exhibits both catalytic activity and cation conductance (121). The *Paramecium* adenylyl cyclase gene harbour both a voltage-gated cation channel with six transmembrane helices and a cytosolic adenylyl cyclase domain. There are multiple isoforms of the gene in *Paramecium* and *Tetrahymena* (79). Less well conserved orthologues are also present in several *Plasmodium* species, *Toxoplasma gondii* and *Cryptosporidium parvum* (78, 79). The *P. falciparum* genome harbours a second putative adenylyl cyclase, PfAC2, which encodes a soluble protein with a single cyclase domain (78). The cyclase domains of all alveolate adenylyl cyclases are more similar to *Dictyostelium* sGC and mammalian sAC than to the dodecahelical adenylyl cyclases (59, 80).

The *Paramecium* and *Tetrahymena* guanylyl cyclases are Ca²⁺/calmodulin activated enzymes, that are regulated by voltage-gated Ca²⁺ influx (122-124). A single guanylyl cyclase gene was identified in the two ciliates and two orthologous genes were found in *P. falciparum*. All genes encode large proteins with in total 22 transmembrane helices. The N-terminal set of 10 helices conforms to the P-type cation transport ATPases (125-127). However, several permutations of essential residues make it doubtful whether the ciliate ATP-ase domain can carry out active ion transport (125). The C-terminal region shows the characteristic topology of the mammalian adenylyl cyclases with two sets of six transmembrane helices interspersed and followed by a catalytic domain. However, as is the case for *Dictyostelium* GCA, the order of the C2 and C1 cyclase domains is reversed. The identity of the enzymes as a guanylyl cyclases was borne out by heterologous expression of the holoenzyme and the catalytic domains (125, 126, 128). In ciliates, both cAMP and cGMP have distinct roles in controlling cell motility (129), while in *P. falciparum*, cGMP has been implicated in exflagellation. This crucial step in the differentiation of male gametocytes is induced by xanthurenic acid (130-132). Recent work shows that xanthurenic acid increases the activity of membrane associated guanylyl cyclase, which provides an important clue for the function of the *Plasmodium* guanylyl cyclases (133). Despite its structural similarity to the G-protein regulated mammalian adenylyl cyclases, the *Paramecium* guanylyl cyclase was not activated by forskolin or by GTPγS (128). The *P. falciparum* guanylyl cyclases cannot be regulated by heterotrimeric G-proteins *in vivo*, because heterotrimeric G-proteins are absent from...
Guanylyl cyclase diversity

the P. falciparum genome (80).

8. DISCICRISTATES

The common feature of the discicristates (figure 3) are their discoid shaped mitochondrial cristae, but other aspects of their morphology are entirely different. There are amoeboid forms, of which some, the acrasids, show a primitive form of multicellularity. The euglenoids are photosynthetic freshwater protists which use a single flagellum for movement, and the trypanosomatids are obligate parasites, that retained the flagellum and the plastid, but lost the photosynthetic enzymes of the euglenoids. cGMP signaling seems to be more prevalent in the whole group than cGMP signaling. Two homologous light-activated adenylyl cyclases were isolated from the parflagellar body of the photosynthetic flagellate Euglena gracilis. These remarkable enzymes are soluble proteins with two flavin binding domains that are interspersed with two cyclase catalytic domains (figure 2). Cyclase activity is 80-fold stimulated by blue light that is captured by the flavin moiety. The subsequent production of cAMP mediates the photoavoidance response of this organism (134, 135). The adenylyl cyclases that are present in the trypanosomatids T.cruzi, T.brucei and Leishmania donovani do not resemble the Euglena enzyme. They all share the same overall structure with a large extracellular domain, a single transmembrane helix and a single intracellular cyclase domain (136-139). Although their membrane topology resembles that of the receptor guanylyl cyclases, no ligand binding and kinase homology domains are present. In this aspect they resemble Dictyostelium ACG. All trypanosomatid cyclases are part of multigene families, but their number is particularly large in T.brucei. Here, at least 6 genes are part of polycistronic units that also encode the variant surface glycoproteins (VSG), which the parasite uses to escape the host immune system. These ESAG (Expression Site Associated Genes) adenylyl cyclases are only expressed when the corresponding VSG expression site is activated. In addition, several multi-membered gene families of cyclases called GRESAG (Genes Related to ESAGs) are constitutively expressed (136, 140). Only a few of the parasite genes have been functionally characterized as adenylyl cyclases and there may well be guanylyl cyclases among the remaining multitude.

Soluble NO synthase and guanylyl cyclase activities were detected in T. cruzi epimastigotes. The NO donor, nitroprusside stimulates the guanylyl cyclase activity 2.5-fold (141), and both nitroprusside and 8Br-cGMP stimulate epimastigote motility (142). Although the NO-induced stimulation of the parasite guanylyl cyclase is insignificant compared to the 200-fold stimulation of mammalian sGC by NO (29, 30), the data suggest that cGMP is present in trypanosomes. In T. brucei, this is substantiated by the presence of a highly unusual cGMP binding protein. The TbRSU gene encodes a protein, that has all the properties of PKA-R and also interacts with a PKA-type catalytic activity. However, both cNMP binding sites of TbRSU specifically bind cGMP with a Kd ~ 10 µM, while cAMP up to 100-fold higher concentrations does not bind at all. In addition, only cGMP can activate the associated PKA activity (143). In view of these data, there are good reasons to believe that guanylyl cyclases will also be present and their sequences may be revealed once the ongoing sequencing projects of the T.brucei and T.cruzi genomes have reached completion.

9. PROKARYOTES

The prokaryotes outrank all other forms of life in the large number and diversity of nucleotidyl cyclases. There are 5 classes of adenylyl cyclases (I, II, IV, V, VI) with structurally unique catalytic domains that are absent from eukaryotes. In addition, they display an unequalled variety of the universal class III adenylyl cyclases. A recent survey of 129 bacterial genomes identified 193 class III cyclase genes in only 29 species (144). These enzymes share conserved catalytic domains with each other and with all eukaryote adenylyl- and guanylyl cyclases. However, the bacterial cyclases additionally show an unrivalled range of combinations with other functional domains and a variety of membrane topologies. The majority has a single cyclase catalytic domain, but there are also enzymes with the characteristic C1 and C2 catalytic domains of the mammalian dodecahelical adenylyl cyclases (144). A selection of cyclases with similar domain architectures as the eukaryote cyclases are presented in figure 3.

The abundance of cAMP-producing enzymes forms a stark contrast with the presence of only a single guanylyl cyclase. This enzyme, Cya2, from the cyanobacterium Synechocystis harbours a single cyclase domain, 4 putative transmembrane domains and a periplasmic region with a CHASE domain. This domain is also found in Dictyostelium ACG, in other bacterial adenylyl cyclases and in sensor histidine kinases and is hypothesized to detect small molecules (83). Targeted inactivation of Cya2 causes a 60% reduction of cGMP in the bacterium without reducing cAMP (145). However, a direct demonstration of cGMP production by the enzyme has not yet been presented.

10. SUMMARY AND PERSPECTIVE

Guanylyl- and adenylyl cyclases exist in a broad variety of structural forms. The distribution of these forms across the major groups of eukaryote organisms seems more or less random and except between closely related taxa, no patterns of inheritance are evident. In the prokaryotes, the class III cyclases show even greater structural variety and many structural forms can be recognized as prototypes for the eukaryote enzymes. For instance, the hexahelical bacterial enzymes with a single catalytic domain may, by modification of some of the helices, have given rise to the K+ channel of the alveolate adenylyl cyclases, or by gene duplication to the dodecahelical form of the vertebrate adenylyl cyclases. The response regulator and histidine kinase domains of the Nostoc enzyme CyaC are also found in Dictyostelium AcrA (figure 2). The AAA-type ATP binding domain is common to the soluble eukaryote enzymes sAC and sGC, but is also
found in soluble bacterial cyclases. An extracellular CHASE domain is a common feature of Dictyostelium ACG and a number of bacterial cyclases.

As opposed to a scenario where all eukaryote cyclases diverged from a single ancestral eukaryote form, many of the bacterial enzymes seemed to have found their way into the eukaryotes in parallel. Some cyclases could have come with the archeabacterial ancestor of the first eukaryotes. Others could have been acquired from the ancestral bacteria and cyanobacteria that gave rise to eukaryote organelles of endosymbiotic origin, such as the mitochondria and chloroplasts. Thereafter, the emerging eukaryote groups seemed to have elaborated in different ways on some forms and have lost the others. In some organisms this loss is particularly dramatic. For example, class III guanylyl cyclases appear to be abundant in the chlorophyte algae, but have vanished from their relatives, the flowering plants. A residual function for cGMP in ion channel regulation has been retained, but when and why were the guanylyl cyclases lost? The fungi present a similar case. Phylogenetically close to both animals and dictyostelids, which share related types of adenylyl- and guanylyl cyclases, the fungi have only a single adenylyl cyclase, with little resemblance to either its animal- and dictyostelid counterparts. When were the other forms lost and what did they regulate? The obvious common property of higher plants and fungi, that is not shared by any other group, is their sedentary life style and the complete loss of motility in all their cells including the gametes. Because cyclic nucleotide signaling is so often implicated in control of cell movement, it was possibly the loss of motility during plant and fungal evolution that led to the demise of their guanylyl- and adenylyl cyclases.

We are only at the beginning of the post-genomic era and as the number of completely sequenced genomes increases, the immense gaps in our knowledge of the evolution of cyclic nucleotide signaling will gradually be filled. Only then can we begin to understand the diversity and versatility of this most universal and most ancient system for signal transduction.

11. ACKNOWLEDGEMENT

This work is supported by BBSRC grant 94/COD16760 and Wellcome Trust University Award Grant 057137.

12. REFERENCES


Guanylyl cyclase diversity

on guanylate cyclase activity, cyclic GMP concentration, and ion transport in small intestine. *Proc Natl Acad Sci USA* 75, 2800-2804 (1978)


41. L'Etoile, N. D. and C. I. Bargmann: Olfaction and odor
Guanylyl cyclase diversity


Guanylyl cyclase diversity


88. Meima, M. E., R. M. Biondi and P. Schaap:
Guanylyl cyclase diversity


Guanylyl cyclase diversity

nucleotidegated channel from *Arabidopsis* indicates its involvement in programmed cell death. *Planta* 213, 327-332 (2001)


120. Becker, B., N. Feja and M. Melkonian: Analysis of expressed sequence tags (ESTs) from the scaly green flagellate *Scherffelia dubia* Pascher emend. Melkonian et Preisig. *Protist* 152, 139-147 (2001)


136. Alexandre, S., P. Paindavoine, P. Tebabi, A. Pays, S.
Guanylyl cyclase diversity


Key Words: Guanylyl cyclase, cGMP phosphodiesterase, Cyclic nucleotide gated channel, PKG, Adenylyl cyclase, Arthropod, Nematode, Fungi, Dictyostelium, Arabidopsis, Zea mays, Chlamydomonas, Ciliate, Plasmodium falciparum, Trypanosoma, Review

Send correspondence to: Professor Pauline Schaap, School of Life Sciences, University of Dundee, MSI/WTB complex, Dow Street, Dundee DD1 5EH, UK, Tel: 44-1382-348078, Fax: 44-1283-345386, E-mail: p.schaap@dundee.ac.uk

http://www.bioscience.org/current/vol10.htm