The EMBO Journal Vol.18 No.15 pp.4222–4232, 1999

Guanylyl cyclases with the topology of mammalian adenylyl cyclases and an N-terminal P-type ATPase-like domain in *Paramecium, Tetrahymena* and *Plasmodium*

Jürgen U.Linder, Peter Engel, Andreas Reimer, Thomas Krüger, Helmut Plattner[1], Anita Schultz and Joachim E.Schultz[2]

Fakultät für Chemie und Pharmazie, Universität Tübingen, Auf der Morgenstelle 8, D-72076 Tübingen and [1]Fakultät für Biologie, Universität Konstanz, D-78434 Konstanz, Germany

[2]Corresponding author
e-mail: joachim.schultz@uni-tuebingen.de

We cloned a guanylyl cyclase of 280 kDa from the ciliate *Paramecium* which has an N-terminus similar to that of a P-type ATPase and a C-terminus with a topology identical to mammalian adenylyl cyclases. Respective signature sequence motifs are conserved in both domains. The cytosolic catalytic C1a and C2a segments of the cyclase are inverted. Genes coding for topologically identical proteins with substantial sequence similarities have been cloned from *Tetrahymena* and were detected in sequences from *Plasmodium* deposited by the Malaria Genome Project. After 99 point mutations to convert the *Paramecium* TAA/TAG-Gln triplets to CAA/CAG, together with partial gene synthesis, the gene from *Paramecium* was heterologously expressed. In Sf9 cells, the holoenzyme is proteolytically processed into the two domains. Immunocytochemistry demonstrates expression of the protein in *Paramecium* and localizes it to cell surface membranes. The data provide a novel structural link between class III adenylyl and guanylyl cyclases and imply that the protozoan guanylyl cyclases evolved from an ancestral adenylyl cyclase independently of the mammalian guanylyl cyclase isoforms. Further, signal transmission in Ciliophora (*Paramecium*, *Tetrahymena*) and in the most important endoparasitic phylum Apicomplexa (*Plasmodium*) is, quite unexpectedly, closely related.

Keywords: adenylyl cyclase/cGMP/guanylyl cyclase/*Paramecium*/*Plasmodium*

Introduction

cAMP is used universally as a second messenger. Therefore, the existence of at least three phylogenetically separate classes of adenylyl cyclases (ACs) is surprising. The membrane-attached ACs from Gram-negative bacteria (class I ACs) possess no sequence or topological similarity to the class II AC toxins from *Bordetella pertussis* or *Bacillus anthracis*. Both class I and II ACs are equally distant from the class III isozymes, which are found in mammals as well as in slime molds and many prokaryotes. The class III ACs share sequence similarity in their catalytic domains and, to a lesser extent, with the catalytic regions of mammalian soluble and membrane-bound guanylyl cyclases (GCs) (for a review see Barzu and Danchin, 1994). The currently accepted topology of mammalian ACs comprises two hydrophobic membrane cassettes of six α-helical transmembrane spans (TM), designated M1 and M2, and a catalytic center formed by two conserved cytoplasmic domains of ~250 amino acids, C1a and C2a, which are localized downstream of M1 and M2, respectively (see Figure 1B for a schematic representation). C1a and C2a precede poorly conserved cytosolic segments of variable length, designated C1b and C2b. Initially it was speculated that the prominent membrane anchors of the ACs might entertain an independent transport activity because their topology is reminiscent of the multidrug resistance gene, the cystic fibrosis gene, various membrane transporters and voltage-gated channels (Krupinski *et al*., 1989). Yet, all attempts to demonstrate such a secondary function of class III ACs failed. Mammalian ACs have a low basal activity and their hormonal stimulation is mediated predominantly by G-proteins in conjunction with other effectors (for a review see Sunahara *et al*., 1996).

We reported that in the protozoans *Paramecium* and *Tetrahymena*, cAMP and cGMP formation are biochemical correlates of electrophysiological events. Thus, the control of cAMP levels in *Paramecium* is linked to a K^+ outward current because K^+ channel blockers such as tetraethylammonium, Cs^+ and quinine inhibit hyperpolarization-activated cAMP formation, and because a mutant, *restless*, with a defect in the K^+ resting conductance is unable to control cAMP formation (Schultz *et al*., 1992). This raised the possibility that within the AC protein from *Paramecium* the function of an ion pore which is intimately linked to the regulation of enzyme activity may have been preserved. Indeed, the purified AC of 97 kDa, i.e. comparable in size with the prototypical mammalian versions, displayed both an AC activity and cation conductance (Schultz *et al*., 1992).

cGMP levels in *Paramecium* and *Tetrahymena* are regulated by a Ca^{2+} inward current (Schultz *et al*., 1986). *Paramecium* mutants with defects in Ca^{2+} channel regulation such as *pawn* or *pantophobiac* show deficits in cGMP regulation. To understand the molecular properties of *Paramecium* ACs and GCs, we wished to clone and express candidate cDNAs. As a start, we used a 328 bp genomic DNA sequence from *Paramecium* which presumably coded for a partial sequence of a mammalian-type AC (Hinrichsen *et al*., 1995), and expected to clone an AC in which the bulky membrane anchors had retained an ancestral ion conductance. To our surprise, we discovered a GC which is disguised in the topology of mammalian ACs and has an extended N-terminal domain with topological and amino acid sequence similarity to P-type

© European Molecular Biology Organization

ATPase ion pumps. After 99 point mutations for conversion from the ciliate to the universal DNA code and partial resynthesis of the cDNA, the gene was heterologously expressed. Similar genes were identified by cloning in *Tetrahymena* and in *Plasmodium* by searching the data released by the Malaria Genome Project. Our data provide a novel view on the evolution of mammalian ACs and GCs and suggest the existence of a new type of signal transduction in these protozoans which employs P-type ATPase-like structures as receptor.

A

```
   1  MGLLRRLSQIFKSEVPQTNNDHLHENQQYSKKIVLENGESVNVRGNENQF
  51  NLVKTSRYTLTNCLYVLIIKFFKRFLNLYLFVINMIQLDYNYLISLLLTF
 101  TIHIIIELYFDTQKRKRDEIVNSRQATVFADLKKNVQSVASQNLIKLGGR
 151  KSLNQNIVEHENLVNNNNYQKDICMRKVSIDTVTYRQDKQSRIPLFTMKR
 201  WDQLQVGDIIYLKNNEICPADVLILDMGQSVSMASNTVMSGNTNEVRKRA
 251  CPLTTISKEHKIQLLDYRTILNGVIRYDQTDSDQYYKGMVKLKKDPKPLE
 301  INKENIFFREQLLLNDLYLFGVILSVGLDCRCYKSFKHVEKYGYFEKKAN
 351  LYFLIAILSLILLSVIQYTIEYYGEHHNNLEFQLMNYSGLLPLYFYFLID
 401  LLYFTQMLYNNYLFLKEKGNSKHPLKEINPLLDNAHIKLNHLQVNYSPIS
 451  NLSLINHVLIDKTGTLTMPNFKIKWIFIFDSLYKLRYKAFAEKEFQNALK
 501  QTKIRQELNQPIDERLEPESTPMIKCNNNEEVNFKIEDLDDDHPPYDELH
 551  RTQPPRFRPESQLSSQFDQSAEGAQIQKQSLARGSSLIVNSLHESKRQLT
 601  YQRGSQKFHAPRIPSNLEIVSEVEFSGNEITLIRRIQDDEFKPHYFEAMF
 651  ALVVCQNTRSLYQKSDDSFFYEYSSELDKQQLYLCKHYGFKFICRSKQEN
 701  QMRYVIQDGEEIHLIDVMSLHSFNSRKFSIIIKLQEELCTKLGLENEDSQ
 751  YLQYLRDDSLDMIGCLSLEKDQVAKLDLMIKDLQVQGSRPVLFYRSLMTE
 801  SQMQTFNKEANKLLDIQSNKIEFMHEYQFLMHEFEKNCDLVSVIGVQ EKV
 851  NKNVTPTLNYIKQLGLPCWVVSGDNYEKVLPIAYRVQLLNQNDTVVHIQA
 901  KTQDELFLQLKQLLQSLSGQLRSHSNKKELKHSDTLIHKKGSFTCSPRWI
 951  SSGQKNLHIKPFQIVISGESLEQILRDAYLKKHFQFLFQFTSNFIGYRMT
1001  PQQKSILIKLLKDRKLNYKFILSIGDSFSDINLFNHSDFTFQMQSGRQIF
1051  PHDKNNFEDLQHGRVSIKQLDQESQFCQHKQDHGFIEPLYNSDVYIKDFE
1101  LIHKMIALDSRKAALYFERILIFALFRSFGIIYLLCVTQLIRDEIVVITK
1151  DYLKFHSNIFFLLSSSFIISEISEKVENFNHNEFMYFVFKNDQLELNHNK
1201  FGKYIARITFFPLIQALLINLYTQYADLSNQNGIVLTSSELGSGLFLLLL
1251  TVDASKLVQRNYITQKQFVFNIGLTFGLYTLISLQFDSISIIYQWAASVQ
1301  ILFSFVFIVAIHCIFNLVLYHINRPFFMPLAIQTSKEYDEILKIYNKISG
1351  DLAKQIKLHRATIEKISYFARKLFNGDEDMDPIIKQMLSGAFADETENSI
1401  NSITLRFKSQLLQNKFTEDTLQYVVRTYRIIVAITFLFYEVATYIIILLA
1451  FDSFTLSYWTDYYYICMFGVLLLLMLFCWSRYYKLYFFEVNFIILFIRCS
1501  SIIIWLFNQDQSVGGMNLNMLQLSLAVNIYHYQEQPLISQVFSAIYVAIY
1551  LIKKSNEDDVYILINQYVMSIASLFLLVLTFRRIIHIFLEVFLGRINMNK
1601  ENQIMSDILSILLPQFIRDRINKAGQYDIQEDQGMVAVLFCDIIDFDQLI
1651  KNEQSNVVDILDKLFRRFDLLCQQHEVQKIETVGKTYMAAAGLKIHVSQK
1701  SNPVNKVISLALDMKRYVMSNETFQIKIGIHYGNVIAGVIGHHKPQFSLI
1751  GDTINTASRICSTAESWDVAISEQAYRQTNKYELVYVQRDVVAKGKGKLI
1801  TYVVNTKRGGKQRKQTIMIQRPPKQPFRDPKDYQNLDQVHENQLSEIGIG
1851  SNREVQPLVNRNEDPQPVTLNKKQKSEIYSFVTKTQKQISIESPNKSLQQ
1901  QMLFQQSQAQLNQGGGSSKNYASQSIMGGIQSQGYDDRHQKHQDLIQLLE
1951  RKTRCELNVSDYQLNMEYHVEKDKIRALYSVDEDAQGQNLLLTLNPKKLY
2001  LEFNEDATVTLEFYEFLSSLYRPFAVQYLLTIGSCVLISSMSCVSLNGQI
2051  YLANLITGVMVGLLHLLQCTLIHRYESFYNVIYFLILAYLLLGSYALDCF
2101  LITDHYEIQMAKGIIYCLSVLLNRVLRHQNKLVYLVLFLGVTIAAIIVND
2151  WPWSRFYYCAVISIFAFCIQLFMFLKNVNSYNVNKQLIEKSIKYQY LLNY
2201  LLPKNVLEEFFRPNEEKRVLREQADEVTLLFADIAGFTEYSSKVQPEQVV
2251  NMLRNLFTEFDKNSLLHNVFKLYTIGDCYVVMGMVDYGKGIQRNPSQEAV
2301  NVVRMGFAMIDAIRRVRAHINHPTLDMRIGVHTGSIIGGVLGTELVRYDI
2351  YGPDVLIANKMESKGAKGFVQVSQETKDIIEREFPDLFRFEYKQSIEFES
2401  IERKTSGYFVYQ
```

Results

Cloning of a Paramecium cyclase with a mammalian AC topology

The sequence of a 328 bp PCR product from genomic DNA of *Paramecium* which displayed 29% amino acid identity to the catalytic C2a region of the rat type III AC was kindly provided by C.Russel and R.D.Hinrichsen (Hinrichsen *et al.*, 1995). Using specific primers, we amplified a 219 bp piece which was used as a probe to screen a *Paramecium* cDNA library prepared with poly(A)$^+$ mRNA. We obtained two cDNA clones of 4.5 and 7.2 kb, both with a 3′ poly(A) tail. The open reading frame (ORF) was identical in both clones and ended at a TGA stop codon 47 bp upstream of the poly(A) tail. The 7.2 kb clone contained an ORF of 2408 amino acids; an ATG start codon was not identifiable. The 5′ ends of *Paramecium* genes are hardly ever detected in cDNA libraries because their 5′-untranslated regions (UTRs) usually are very short. A genomic DNA (gDNA) library prepared from *Eco*RI-digested DNA and a new probe located at the 5′ end of the clone were used to obtain the ATG start. The genomic sequence unequivocally identified an ATG as start just 12 bp upstream of the 5′ end of the 7.2 kb cDNA clone because of an in-frame TGA stop only a further 24 bp upstream. The ORF coded for a protein of 2412 amino acids with a calculated molecular mass of 282.6 kDa (Figure 1). Hydrophobicity analysis indicated five hydrophobic regions with 22 putative TMs. Similarity searches demonstrated that the protein consisted of two large units: an N-terminal half of 1319 amino acids (155 kDa) with similarity to P-type ATPases and a C-terminal half of 982 amino acids (115 kDa) with a topology identical to that of the prototypical mammalian AC (Figure 1A). Both units were linked by a polypeptide of 111 amino acids. To exclude the possibility that we

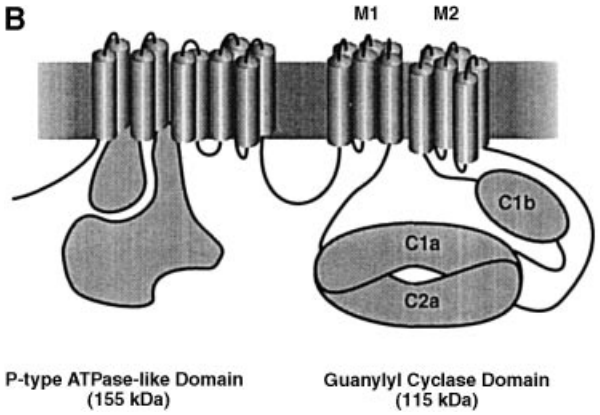

Fig. 1. (**A**) Sequence of the *Paramecium* guanylyl cyclase. Transmembrane regions are shaded; the four consensus blocks in the P-type ATPase-like domain and the two catalytic domains of the guanylyl cyclase region are in bold. (**B**) Suggested topology of the guanylyl cyclase. The 22 putative transmembrane spans are depicted as barrels, and cytosolic portions of the protein are shown as a duck head-like shape for P-type ATPases (Toyoshima *et al.*, 1993; Zhang *et al.*, 1998) and as a doughnut shape for the C1a/C2a heterodimer of the cyclase (Tesmer *et al.*, 1997; Zhang *et al.*, 1997). As in several mammalian adenylyl cyclases, a distinct C2b region is absent in the cloned guanylyl cyclase. For the C-terminal cyclase, the short-hand nomenclature used for mammalian adenylyl cyclases is adopted. M1/2 designate two membrane cassettes of six transmembrane spans, C1a, C1b and C2a domains coalesce as a cytosolic catalytic center.

sol GCα	523	KV	**E**	TIGDAYCVAGGL	589	KMP	**R**	Y	**C**	LFGNNVTLANKF
III C1a	362	RI	**K**	IL**GDCY**YCICGL	428	KRW	**Q**	Y	**D**	VWSTDVTVANKM
VII C1a	321	RI	**K**	IL**GDCY**YCVSGL	387	RKW	**Q**	Y	**D**	VWSHDVSLANRM
VIII C1a	454	RI	**K**	IL**GDCY**YCVSGL	520	RKW	**Q**	F	**D**	VWSWDVDIANKL
PARA C2a	2271	KL	**Y**	TI**GDCY**VVM-GM	2344	ELV	**R**	Y	**D**	IYGPDVLIANKM
PARA C1a	1679	KI	**E**	TVGK**TYMA**AAGL	1743	HKP	**Q**	F	**S**	LIGDTINTASRI
III C2a	973	KI	**K**	TIGS**TYMA**ASGV	1057	RKP	**H**	Y	**D**	IWGNTVNVASRM
VII C2a	926	KI	**K**	TIGS**TYMA**AAGL	1002	RKP	**Q**	Y	**D**	IWGNTVNVASRM
VIII C2a	1029	KI	**K**	TIGS**TYMA**VŠGL	1101	KKP	**Q**	Y	**D**	IWGKTVNLASRM
sol GCβ	471	KV	**E**	TVGDKYMTVSGL	536	RMP	**R**	Y	**C**	LFGNTVNLTSRT

Fig. 2. Local alignment of the catalytic C1a- and C2a-positioned regions of the *Paramecium* GC (PARA) with corresponding mammalian sequences [rat soluble guanylyl cyclase α1- and β1-subunits: sol GCα/β (Nakane *et al.*, 1990); rat type III AC (Bakalyar and Reed, 1990); bovine type VII AC and rat type VIII AC (Cali *et al.*, 1994; Völkel *et al.*, 1996)]. The signature sequences GDCY and TYMA are in bold. Those residues that determine substrate specifity are printed inverted. The boxed amino acids are located in a non-catalytic pocket of the mammalian AC heterodimer and their positions correspond to those which determine the substrate specificity in the catalytic groove (Tesmer *et al.*, 1997; Zhang *et al.*, 1997). The switch of the catalytic domains in the *Paramecium* GC is apparent.

cloned a cDNA which was joined accidentally during library preparation, we cloned a 2.5 kb *Eco*RI fragment from gDNA which was intronless, and we fully covered the intradomain region. Thus we unequivocally established the presence of this gene at the gDNA level.

The *Paramecium* cyclase unit contained hydrophobic M1 and M2 regions with six TMs each which were followed by C1a- and C2a-positioned domains with sequence similarity to the catalytic loops of metazoan ACs (Figure 1B). The C1a-positioned domain was followed by a hydrophilic stretch of 221 amino acids reminiscent of the C1b region in mammalian ACs; a distinct C2b region was absent. All mammalian ACs have GDCY as a signature sequence in the C1a domain. In the cloned protozoan cyclase, the GDCY motif is present in the C2a-positioned domain (2276–2279), i.e. close to the C-terminus. Similarly, the motif TYMA is invariant in all C2a regions of mammalian ACs whereas in the ciliate cyclase it was located in the C1a-positioned domain (1686–1689), i.e. toward the 5′ end (Figure 2). Further, mammalian ACs contain a VKGKG motif in their C2a catalytic region. The first lysine presumably binds the γ-phosphate of ATP (Tesmer *et al.*, 1997) and is essential for catalysis. In the *Paramecium* gene, a similar motif, ^{1793}AKGKG^{1797}, was found in the C1a position. Therefore, we conclude that the C1a-positioned and C2a-positioned regions of the *Paramecium* cyclase were inverted compared with the mammalian AC congeners resulting in a (mammalian nomenclature) M1C2aC1bM2C1a architecture. Recently, the amino acids in the heterodimeric catalytic pocket of metazoan ACs and GCs which are important for substrate specificity, K, D and Q in ACs, E, C and R in GCs, have been determined by X-ray crystallography and site-directed mutagenesis (Tesmer *et al.*, 1997; Zhang *et al.*, 1997; Sunahara *et al.*, 1998; Tucker *et al.*, 1998). In the *Paramecium* cyclase, these crucial amino acids are GC-like, with a substitution of C by S1748 (Figure 2). On the other hand, several amino acids which form a second, non-catalytic pocket in mammalian ACs were conserved in the protozoan cyclase, i.e. were like those in ACs (Figure 2). The results imply that the cloned gene represents a novel GC which is very closely related to metazoan ACs by topology and primary structure. Below, the name *Paramecium* GC is adopted for this gene/enzyme.

P-type ion transport ATPases were the only proteins with significant similarity to the N-terminal half of the *Paramecium* GC. It had a predicted membrane topology identical to that of P-type ATPases, i.e. two cassettes of two putative TMs in the N-terminal portion (amino acids 63–110 and 349–407) and one set of six TMs near the C-terminus (amino acids 1121–1319; Figure 1A; Stokes *et al.*, 1994; Zhang *et al.*, 1998). In the cytosolic domains of the P-type ATPase family which has >150 members, several sequence blocks are conserved (Henikoff and Henikoff, 1992). Four of these align to corresponding cytosolic regions of the *Paramecium* GC. DKTGT(L/I)T, which is located within the large cytoplasmic loop, is an invariant signature sequence in all P-type ATPases. For ion transport to occur, the aspartate must be phosphorylated (Allen and Green, 1976). In the *Paramecium* GC, DKTGTLT was conserved at the equivalent position (amino acids 461–467). A conserved GDGXND motif present in the hinge region of P-type ATPases was ^{1025}GDSFSD^{1030} in the *Paramecium* gene, and the (TSND)GE (SNT) block in the transduction domain was retained with a signifcant E to N change at ^{240}SGNT^{243} virtually excluding ATPase function (Clarke *et al.*, 1990a; Fagan and Saier, 1994). Finally, the ATP-binding domain in nearly all eukaryotic P-type ATPases involves an indispensable aspartate in a block of moderately conserved amino acids (Clarke *et al.*, 1990b). The corresponding amino acid in the *Paramecium* P-type ATPase-like domain was Glu848. These decisive deviations of the *Paramecium* sequence from the ATPase consensus imply that this GC domain has adopted another function.

Expression of Paramecium GC

Transfection of *Paramecium* is not yet established. However, heterologous expression of *Paramecium* genes is impossible because it uses the universal stop codons TAA and TAG for glutamine (Preer *et al.*, 1991). Therefore, we painstakingly changed all 99 TAA and TAG codons to CAA/CAG by site-directed mutagenesis (Deng and Nickoloff, 1992). The ‘corrected’ ORF was inserted into the bicistronic expression vector pIRES1neo (Rees *et al.*, 1996). However, upon transfection, no HEK293 cells survived the G418 selection. Using a green fluorescent protein (GFP)-based reporter assay (Trouet *et al.*, 1997) with the cyclase domain, we realized that HEK293 cells did not transcribe a functional mRNA compared with cells transfected with a vector containing a bovine type VII

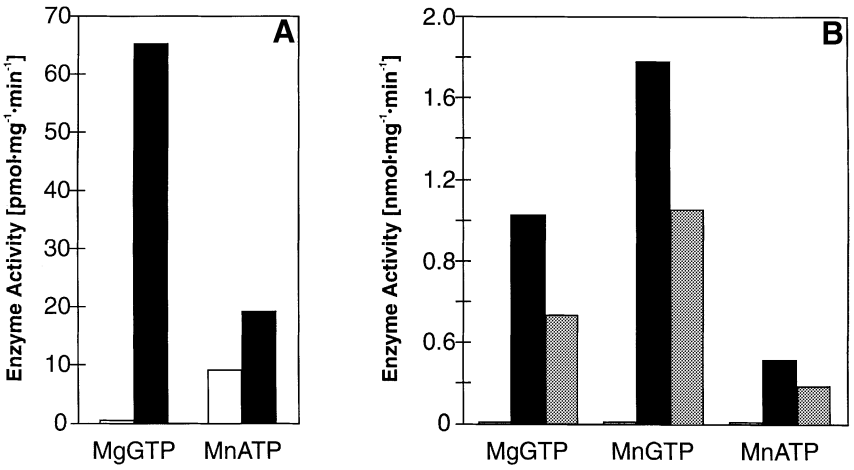

Fig. 3. Cyclase activities of the recombinant guanylyl cyclase domain and of the holoenzyme. (**A**) Guanylyl and adenylyl cyclase activities in cell membranes of HEK293 cells transfected with the GC domain were measured with 75 μM MgGTP or MnATP, respectively. (**B**) Guanylyl and adenylyl cyclase activities in cell membranes from Sf9 cells infected with either the GC domain (black bars) or the GC holoenzyme (shaded bars) were measured with MgGTP or MnATP as indicated. In either expression system, cAMP formation with MgATP as a substrate was barely above background. Open bars are from vector-infected controls. Note the different scales of the ordinates indicating the much higher expression efficiency in Sf9 cells.

AC cDNA (data not shown). Assuming that the problem may be caused by the high AT content of the *Paramecium* GC gene (66%) and in order to localize the problem, we used the bovine AC as a scaffold to test for proper transcription of individual protozoan cDNA segments. We separately inserted six cDNA sequences of the cyclase domain (N-terminal amino acids, 1380–1420; M1, 1421–1589; C1a-positioned, 1590–1806; C1b-positioned, 1807–1996; M2, 1997–2178; and C2a-positioned, 2179–2412) at the corresponding positions of the bovine type VII AC cDNA using appropriately generated restriction sites. No expression of the GFP reporter gene was observed when M1 or M2 of the protozoan cyclase replaced the respective bovine membrane cassettes. Therefore, we resynthesized the ciliate cDNA coding for M1 and M2 using the standard mammalian codon usage. This introduced a 58% G/C content into ~34% of the *Paramecium* GC domain. The partially synthetic gene in pIRESneo was again transfected into HEK293 cells and yielded a G418-selectable cell population. Using an antibody directed against the C1b-positioned domain (K1811–2007), a set of bands at ~95 kDa was stained specifically in Western blots, indicating protein expression of the M1380–Q2412 GC domain (data not shown). Cell homogenates had membrane-bound GC activities of up to 150 pmol/mg/min (Figure 3A). The K_m for MgGTP was 50 μM. The GC activity was unaffected by 50 μM Ca^{2+}, 1 mM EGTA or by *Paramecium* Ca^{2+}/calmodulin. AC activity was not observed with MgATP as a substrate, yet it was detectable using Mn^{2+} as a metal cofactor (Figure 3A). Forskolin did not enhance enzyme activity. These results proved that the cyclase domain of the cloned gene constituted a GC in the disguise of what hitherto had been considered to be a prototypical mammalian AC topology. We were unable to express the holoenzyme cDNA in HEK293 cells. In the hope of obtaining higher expression levels of the GC domain and an expression of the full-length clone, we turned to Sf9 cells and the baculovirus system for expression.

Expression of the *Paramecium* GC domain in Sf9 cells was equally efficient whether the ciliate cDNA (stop codons removed) or the partially synthetic cDNA clone (see above) was used for infection. GC catalytic activities were ~10-fold higher than in HEK293 cells (Figure 3B). The K_m for MgGTP was 32 μM. AC activity was negligible using MgATP as a substrate, yet was substantial with Mn^{2+} as a divalent cation (Figure 3B). In a competition assay with 75 μM GTP and 1 mM ATP in the presence of 5 mM Mg^{2+} as the physiologically relevant divalent cation, GC activity was diminished by only 25%, i.e. the *Paramecium* cyclase domain truly is a GC. The expression of the GC domain was monitored by Western blotting using an antibody directed against the C1b-positioned region (Figure 4, lane 4). The calculated mass of the GC domain was 115 kDa. The slightly lower apparent molecular weight is probably due to the hydrophobic transmembrane regions, as similar effects have been reported for proteins with large transmembrane domains (Hauser *et al.*, 1998; Mons *et al.*, 1998). Variable glycosyl-

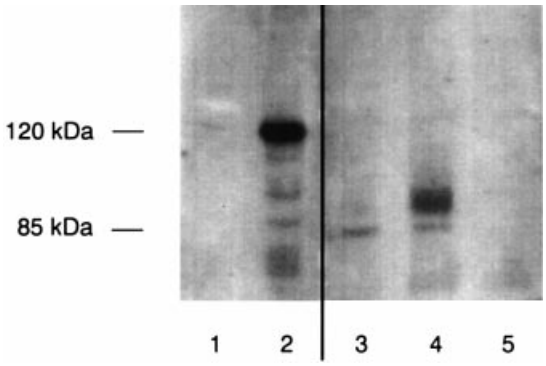

Fig. 4. Proteolytic processing of full-length *Paramecium* guanylyl cyclase expressed in Sf9 cells. Membranes were probed with the anti ATPase antibody (lane 1 and 2) or with the anti-GC antibody (lanes 3–5). Lanes 2 and 3 are from cells infected with the 7.2 kb cDNA coding for the holoenzyme. In lane 4, the expression of a 3 kb cDNA coding only for the GC domain was tested. Lanes 1 and 5 are membranes from control infections. The specificity of the antibodies was ascertained by pre-blocking with the respective antigens and cross-blocking experiments (blots not shown).

Table I. Immunogold labeling of *Paramecium* cellular structures with anti-GC antibody or anti-ATPase antibody

Structures analyzed	Au grains per 100 μm membrane length	
	Anti-GC	Anti-ATPase
CM/OASM	37.9	4.8
IASM	34.5	12.0
CiM	17.6	7.2
TM	0.0	0.0
OMM	0.4	0.0

CM/OASM, cell membrane/outer alveolar sac membrane complex; IASM, inner alveolar sac membrane (facing cell center); CiM, ciliary membrane; TM, OMM, trichocyst or outer mitochondrial membranes.

ation may account for the appearance of closely spaced bands.

We were unable to express the full-length *Paramecium* GC gene in Sf9 cells as determined by Western blotting. Again we reasoned that this may have been due to the high A/T content of the gene, in particular to two stretches of seven and nine As in the ATPase-like region. Multiple As in an mRNA may constitute a signal for premature mRNA degradation, and slippage may occur during replication and transcription (Laken *et al.*, 1997). Therefore, we resynthesized four DNA segments of the ATPase-like domain comprising 46% of this gene portion (nucleotides 1–436, 1012–1387, 1973–2363 and 3303–3991). The G/C content of the synthetic DNA was 56%; the amino acid sequence remained unchanged. Nevertheless, for reasons not yet understood, we were unable to express only the ATPase-like domain in Sf9 cells as determined by Western blot analysis. However, upon infection with the partially synthetic gene coding for the holoenzyme, expression in Sf9 cells was achieved (Figure 4, lanes 2 and 3), although we never detected a product corresponding to the full-length protein of 280 kDa. Using affinity-purified polyclonal antibodies directed either against a 20 kDa cytosolic portion of the P-type ATPase-like domain (L549–M718; anti-ATPase antibody), the 23 kDa C1b-positioned segment of the cyclase domain (K1811–A2007; anti-GC antibody) or a peptide antibody directed against the N-terminal hexadecapeptide R5–N20, we labeled specific bands of lower molecular mass in a Western blot. Because Sf9 membrane proteins up to 350 kDa were blotted successfully as visualized by Ponceau S staining, and because inclusion of a cocktail of protease inhibitors during cell lysis had no effect, the result most likely reflected protein processing in the Sf9 cells. The anti-GC antibody recognized proteins with an apparent molecular mass of mainly 85 kDa up to 100 kDa (Figure 4, lane 3). In the same membrane preparations, the anti-ATPase antibody and the N-terminal peptide antibody specifically detected a 120 kDa band (Figure 4, lane 2, and data not shown). Any unspecificity and cross-reactivity of the antibodies was excluded by respective pre-blocking experiments (data not shown). The proteolytic processing in Sf9 cells must have occurred in the area linking the ATPase-like and the GC domains because of the size of the two generated proteins and beause GC activity was obtained (Figure 3B). This conclusion was supported by the isolation of mRNA from infected Sf9 cells and the unequivocal demonstration by RT–PCR that a correct and uninterrupted mRNA existed (data not shown). In Sf9 cells infected with the gene for the holoenzyme, GC activity was ~50% of that obtained with the GC domain expressed alone. All enzymatic properties of this proteolytically processed holoenzyme corresponded to those of the expressed GC domain. This finding indicated that the physical presence of the ATPase-like protein in the membrane did not affect GC activity *per se*.

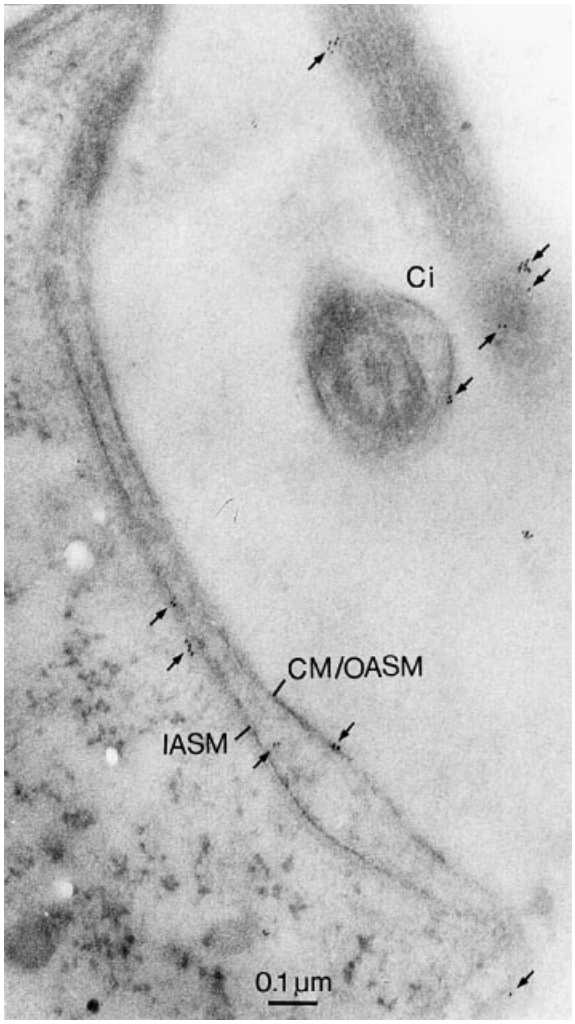

Fig. 5. Immunolocalization of the guanylyl cyclase domain in formaldehyde-fixed *Paramecium* with a rabbit anti-GC antibody followed by incubation with gold-labeled (5 nm) $F(ab)_2$ from goat anti-rabbit antibodies. Note the label on the complex formed by the cell membrane and the outer alveolar sac membrane (CM/OASM), on the inner alveolar sac membrane (IASM), on a cilium cross-section (Ci) and on a 'grazing' section along the membrane of an emerging longitudinally cut cilium (upper right).

Localization of Paramecium GC

Using anti-GC or anti-ATPase antibodies, we detected the GC immunocytochemically in cross-sectioned *Paramecium* using gold conjugates of $F(ab)_2$ as a second antibody (Table I; Figure 5). Both primary antibodies labeled the same membranes, i.e. the GC was present in the ciliary membrane, in the complex formed by the cell membrane and the outer alveolar sac membrane (not resolvable) and in the inner alveolar sac membrane.

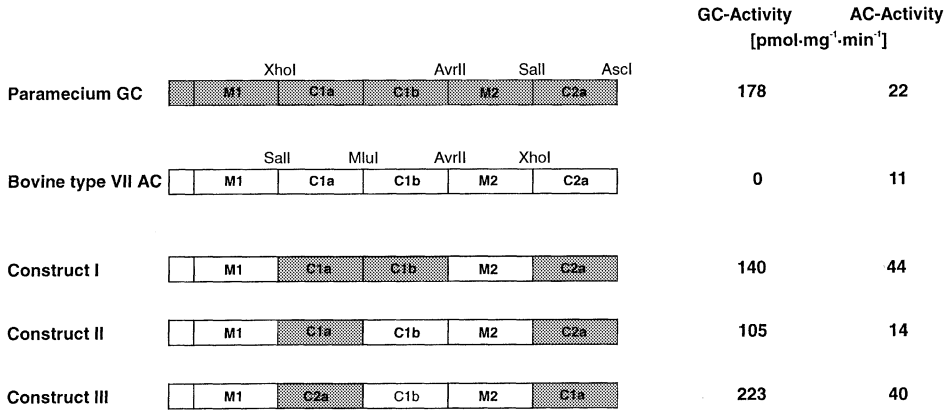

Fig. 6. Guanylyl cyclase and adenylyl cyclase activities of chimeras between the *Paramecium* GC domain and bovine type VII AC. Regions from the *Paramecium* GC are shaded, type VII AC domains are open boxes. Compatible restriction sites (*Mlu*I–*Asc*I and *Sal*I–*Xho*I) were introduced as shown and used for domain shuffling. The locations of the restriction sites in the *Paramecium* GC were: *Xho*I at 4765; *Avr*II at 5986; *Sal*I at 6532; and *Asc*I at the 3′ terminus; and in the bovine type VII AC: *Sal*I at 1183; *Mlu*I at 2032; *Avr*II at 2350; and *Xho*I at 3022. GC activity was assayed with 75 μM GTP and 5 mM Mg^{2+}, AC activity with 75 μM ATP and 2 mM Mn^{2+}. Basal enzyme activities (infections with vector without insert) were subtracted.

Background without the first antibody was negligible (Table I). In semi-quantitative evaluations with both antibodies, we subtracted background labeling determined for probably irrelevant membranes such as outer mitochondrial and trichocyst membranes (Table I). The differences in the labeling intensity were probably caused by differential accessibilities or different affinities of the antigenic sites for the respective primary antibodies.

Chimeras between Paramecium GC and mammalian AC

The GC domain is topologically identical to mammalian ACs. Therefore, we investigated whether a mammalian membrane anchor would be compatible with the GC catalytic domains derived from *Paramecium*. First, we replaced the mammalian C1a, C1b and C2a regions of a bovine type VII AC (Völkel *et al.*, 1996) by the C1a-positioned, C1a,b-positioned and C2a-positioned regions of the *Paramecium* GC domain, i.e. using the mammalian AC as a scaffold for the *Paramecium* catalyst (Figure 6, constructs I and II). Secondly, we generated a construct with an inverse order of the C1a- and C2a-positioned regions from *Paramecium* because, as outlined above, these two regions seem to have been switched in the *Paramecium* GC domain (construct III). The constructs showed substantial and similar GC activity and a minor AC activity when assayed with MnATP (Figure 6). We conclude that the *Paramecium* GC catalytic center comprised of C1- and C2-positioned loops is expressed and active in the frame of a mammalian AC irrespective of the origin of the C1b region and, surprisingly, of the arrangement of the catalytic subdomains.

Tetrahymena guanylyl cyclase

The heterologous expression of *Paramecium* GC holoenzyme in Sf9 cells did not reveal if and how the GC activity is regulated by the P-type ATPase-like domain. Therefore, we wished to clone a GC from the related ciliate *Tetrahymena* for sequence comparisons. A PCR with primers designed to recognize a protozoan C1a-positioned region and cDNA as a template yielded six different PCR products with a high degree of amino acid identity (Figure 7). One fragment was chosen arbitrarily as a probe to clone the corresponding full-length gene from a *Tetrahymena* gDNA library; 9.7 kb of genomic sequence was obtained by chromosome walking. The 3′ end of the ORF was completed by vectorette PCR (Gonzalez and Chan, 1993) and the introns were identified unambigously by PCR using *Tetrahymena* cDNA as a template. The protein encoded by the 8.4 kb ORF had exactly the same topology as the *Paramecium* GC: an N-terminal P-type ATPase-like domain of 168 kDa, followed by a 146 kDa GC domain with mammalian AC membrane topology. As in the *Paramecium* GC, the C1a- and C2a-positioned regions were inverted compared with mammalian ACs. The amino acids decisive for substrate specificity (E1808 and S1885 in C1a and R2734 in C2a) unequivocally identified the encoded protein as a GC. Interestingly, the P-type ATPase-like domain resembled the general ATPase consensus sequence much less than this region from the *Paramecium* GC. In the *Tetrahymena* sequence, the extremely conserved DKTGTLT phosphorylation site evolved to ^{484}SKSGTLM^{490}. This proves that the protozoan P-type ATPase-like domains lack the capacity for an active ion transport. Further, the consensus sequence GDGXND was ^{1180}GSNISD^{1185} in *Tetrahymena* and the conserved (TSND)GE(SNT) box was changed to ^{234}QGSN^{237}. An attempt to compare the 22 putative transmembrane spans of the *Paramecium* GC with those of the *Tetrahymena* clone revealed a potentially significant conservation in TMs 4 and 5. In TM4, the $YX_4LPX_3YX_3D$ motif, and in TM5, a (F,Y)RSF motif, were conserved between *Paramecium* and *Tetrahymena*. Because no similarity exists to the corresponding transmembrane spans in true P-type ATPases, the functional importance of this remains to be evaluated. We note that amino acids in TM6 which supposedly are crucial for cation binding in P-type ATPases were absent in the ciliate GCs (Zhang *et al.*, 1998). Since *Tetrahymena* also uses the universal stop codons TAA and TAG for glutamine, no attempt was made to express the cloned *Tetrahymena* gene.

```
.............AGLKIHMSQK................SNPVNKVISLA  TetraGC2
.............GGLKDYNISD...............NLNPTERILDTA  TetraGC3
.............SGLDVAKHQK......ARKGNDKLWRENETKKALDAS  TetraGC4
.............GGLKGLPPMK........KYLNNNELIHPVTNCVLMA  TetraGC5
.............AGLQGIKYDT........DLFKRYKNFHPVIRCILMA  TetraGC6
.............AGLKNIVSQN................SANPVLRAIQVA  ParaGC2
.............AGLKNIATQS...............NVNPVLRTIQVA  ParaGC3
.............GGLKDYDVVI............NQKNGNSTSRALETA  ParaGC4
QKIETVGKTYMAAAGLKIHVSQK................SNPVNKVISLA  Paramecium GC
QKIETVGKTYMAAGGLKDAEGKK...........SQFRQINPVARAVIMS  Tetrahymena GC
IKIETVFESYLAASGLSEKKNNALDKIMYDTKCAIKLAIAQLSAKYYISY  Plasmodium GC
   ⇑

LDMKRHVM.....SNE......VFQIKIGIHYGNVIAGVIGHHKPQFSLI  TetraGC2
LQMQESVKTMKYGDNK......AVILKIGIHYGRVIAGVIGAHKPQFSLI  TetraGC3
FAIQKKAQKIYWGSKSQ.....KTILKIGIHYGRVIAGVIGYHKPQFSLI  TetraGC4
YDMYENAKKIQYGSKQT.....KLSVKIGIHYGSVMAGVIGGHKPQFSLI  TetraGC5
NQMFEVSKKITYGKSQE.....KIKIKIGIHYGPVLAGVIGGHKPQFSLI  TetraGC6
FEMRKFASQQEFCGSG......EIIVKIGVHYGNVIAGVIGYHKPQFSLI  ParaGC2
FEMRRFAQQQNFGTSE......EIIVKIGVHYGNVIAGVIGYHKPQFSLI  ParaGC3
IQMMETVKTMKYGDNQ......DVQLKIGIHYGRVIAGVIGVHKPQFSLI  ParaGC4
LDMKRYVM.....SNE......TFQIKIGIHYGNVIAGVIGHHKPQFSLI  Paramecium GC
FEMMEAMKSFTYGDNE......KIKIKIGIHYGRVIAGVIGCHKPQFSLI  Tetrahymena GC
KVLDTREHFSDNSTSYDKYINKNISLKIGIHTGKAISGVIGSVKPQYALF  Plasmodium GC
                                                ⇑
```

Fig. 7. Alignment of C1a-positioned regions of guanylyl cyclases from *Paramecium*, *Tetrahymena* and *Plasmodium*. The sequences of the *Paramecium* and *Tetrahymena* GCs are from the full-length clones (DDBJ/EMBL/GenBank accession Nos AJ238859 and AJ238858, respectively); other *Tetrahymena* (TetraGC) or *Paramecium* (ParaGC) sequences are PCR fragments of the C1a-positioned domains (DDBJ/EMBL/GenBank accession Nos: AJ242491 for ParaGC2; AJ242492 for ParaGC3; AJ242493 for ParaGC4; AJ239077 for TetraGC2; AJ239078 for TetraGC3; AJ239079 for TetraGC4; AJ239080 for TetraGC5; and AJ239081 for TetraGC6); the *Plasmodium* sequence is translated from chromosome 13. Identical amino acids are inverted, the consensus glutamate and serine (alanine in *Plasmodium*) residues which are involved in defining GTP as a substrate are indicated by arrows.

Guanylyl cyclase isoforms from Paramecium and Tetrahymena

The variety of AC isoforms is almost a hallmark of the mammalian cAMP signaling system, and often a single cell expresses several isoforms concomitantly. So far, nine membrane-bound isozymes have been identified in mammals which share extensive sequence identities in their catalytic regions, yet are highly diverged elsewhere. The protozoan GCs have the topology of mammalian ACs. Therefore, we wanted to know whether multiple isoforms of this peculiar GC are expressed in *Paramecium* and *Tetrahymena*. Further, we wished to find out whether protozoan cyclase genes may exist which at the crucial substrate-specifying positions contain amino acids which would endow the protein with a preference for ATP as a substrate. Using several slightly degenerate oligonucleotide primer pairs covering the C1a-positioned region, and genomic and cDNA from *Paramecium* and *Tetrahymena*, respectively, as templates, we identified three further DNA fragments from *Paramecium* and another five from *Tetrahymena*. All fragments coded for 60–75 amino acids of the C1a-positioned domain and had a serine, not an aspartate, as one of the three crucial amino acids which specify GTP as a substrate, i.e. these genes coded for GCs (Figure 7). The extremely conserved regions around the GTP-specifying serine (–20 to +2) were nearly identical in all protozoan variants and highly conserved with respect to the corresponding region in mammalian ACs. The anterior 40 amino acids were diverged (21–93% identity) and had approximately the same extent of diversity as that observed in the corresponding C2a domains of the nine mammalian AC isoforms (24–83% identity; Sunahara *et al.*, 1996). From an evolutionary point of view, it is noteworthy that the fully cloned *Paramecium* GC shared 95% amino acid identity with one of the *Tetrahymena* isoforms (Figure 7, TetraGC2). We conclude that in both ciliates, a family of GC isoforms exists comparable in size with that of mammalian ACs, and that some of the GC isoforms have been highly conserved in these protozoans, indicating a conservation of function.

Guanylyl cyclase isoforms in Plasmodium falciparum

We searched databases to determine whether GCs with this architecture were present in other organisms, particularly in other protozoans. In DDBJ/EMBL/GenBank database, we found a 2.797 bp fragment designated as an AC from chromosome 11 of *Plasmodium falciparum* (deposited by D.J. Carucci and D.C.Warhurst; PFAC, accession No. U33118). The ORF codes for six TMs (M2) and a C2a-positioned domain. At nucleotide 1966, a substrate-specifying arginine is encoded, which strongly indicates that this gene, contrary to the given definition, codes for a GC with mammalian AC topology. A C1a-positioned domain is absent in this clone. Further, searching the data which have been released, but not yet edited, from the Malaria Genome Project, we discovered an unequivocal GC sequence on chromosome 13 of *P.falciparum*. It was possible to assemble a complete putative ORF which exactly matched the *Paramecium* GC topology, i.e. it contained a P-type ATPase-like domain fused with a GC of mammalian AC topology with the C1a and C2a positions inverted. The alignment of the C1a-positioned region with the *Paramecium* and *Tetrahymena* GCs clearly identified GTP as the substrate (Figure 7).

Discussion

We have cloned a GC from *Paramecium* which has an N-terminal P-type ATPase-like domain and a C-terminal cyclase domain. So far, mammalian GC catalysts are either homodimers fused to an extracellular receptor by a single TM, or soluble heterodimers linked to a nitric oxide-binding site (Schulz *et al.*, 1991; Yu *et al.*, 1996; Tang and Hurley, 1998; Koesling and Friebe, 1999). The ciliate GC domain differs. It has the unmistakable topology of a

mammalian AC complete with catalytic C1a and C2a regions and two prominent membrane anchors, M1 and M2, with six TMs each. Therefore, the ciliate GC is not related to metazoan GCs but is a close relative of metazoan ACs. A striking difference between the ciliate GCs and their mammalian AC congeners was the switch of the C1a and C2a positions. The preference for GTP could be rationalized by looking at three amino acids which determine the nucleotide specifity in the catalytic pocket. In metazoan ACs, glutamine, lysine and aspartate define ATP as a substrate, whereas in GCs the respective positions are taken by arginine, glutamate and cysteine and define GTP as a substrate (Sunahara *et al.*, 1998; Tucker *et al.*, 1998). In the protozoan GCs, we have arginine, glutamate and a serine instead of a cysteine residue. According to molecular modeling of mammalian GCs, the cysteine SH group interacts with the O^6 of the guanine ring via a hydrogen bridge (Tucker *et al.*, 1998). A serine at this position has a similar space requirement and potential for interaction.

Another remarkable feature of the ciliate GC domain was the successful expression of chimeras with the bovine type VII AC. The catalytic activity of the *Paramecium* GC did not depend on a specific order of the C1a- and C2a-positioned segments, and was completely unaffected by the origin of the C1b region or the M1 and M2 membrane anchors and suggests a considerable mobility of the membrane-anchored loops to form an active heterodimer. Such a cross-compatibility has not yet been reported for mammalian ACs.

In the two fully sequenced ciliate GCs, the C1b and M1/M2 regions displayed no sequence similarity. This was not surprising since the catalytic loops of ciliate GCs were quite insensitive to the origin of the membrane anchors or C1b segments, as demonstrated by the active chimeras. This does not exclude the existence of regulatory functions which depend on C1b or even on M1/M2, but so far those must have escaped detection. The existence of a couple of almost identical GCs in *Paramecium* and *Tetrahymena* leads to two conclusions. First, this isoform must have evolved before *Paramecium* and *Tetrahymena* segregated. Secondly, the existence of several GC isozymes is not an indication of redundancy, rather it defines functional peculiarities which were retained during ciliate evolution. This is not unlike the situation concerning the mammalian AC isoform variety.

The biochemical properties of the recombinant *Paramecium* GC differed from those reported earlier (Schultz and Klumpp, 1982). We were unable to demonstrate an effect of Ca^{2+}. The reasons for this discrepancy may be: (i) the expressed GC is a minor isoform in late-log *Paramecium* cells; (ii) the calcium dependence is conferred by a protein which has no homolog in HEK293 or Sf9 cells or by a post-translational modification which occurs in *Paramecium* cells only; or (iii) a Ca^{2+} dependence may only be demonstrable in the holoenzyme which we consistently failed to express in Sf9 cells due to the inevitable proteolytic processing. The latter may also be responsible for the failure to demonstrate a function for the 155 kDa N-terminal domain with its topology identical to the family of P-type ATPases. Sequence similarity was restricted to four conserved cytosolic regions. We have reasons to believe that these domains do not operate as transporters because decisive amino acid deviations preclude an ATPase function. The most conclusive proof comes from an aspartate invariant in all P-type ATPases. Mutation of this Asp to Glu or Asn completely abrogates enzyme activity (Clarke *et al.*, 1990b). In the *Paramecium* and *Tetrahymena* GCs, the equivalent positions (848 and 1005, respectively) are taken by Glu, thus excluding ATPase activity provided the structure–function relationships which were established for conventional P-type ATPases can be extrapolated to the N-terminal domain of the ciliate GCs. In addition, two genes coding for canonical P-type ATPases have been cloned from *Paramecium* which possess all the structural hallmarks required for an ion transport activity (Elwess and van Houten, 1997; Hauser *et al.*, 1998).

With regard to a potential function of the ATPase-like domains in the GCs, we may get a hint by examining the pore and cation-binding helices TMs4, 5 and 6. TM4 and TM5 are significantly conserved in the ciliates, yet differ substantially from the corresponding helices of ion transporters, and the ciliate TM6 lacks amino acids which supposedly are involved in cation binding (Zhang *et al.*, 1998). We tentatively propose that the protozoan 'pore region' does not bind inorganic ions. Perhaps it associates with organic molecules as does another family of P-type ATPases, the putative aminophospholipid transporters. The protozoan P-type ATPase-like domains would then constitute a receptor-like entity reminiscent of the atrial natriuretic receptor–guanylyl cyclase couple in mammals. However, at present this must remain purely speculative.

Finally, the data allow us to discuss the evolutionary steps which most likely have occurred during the development of these cyclases. The fact that cyclases exist with C1a and C2a arranged in both ways strongly supports the hypothesis that initially membrane-anchored monomers formed a homodimeric AC. This structure is still in existence in several bacteria and may actually represent a common ancestor of mammalian and ciliate nucleoside triphosphate cyclases (Coudart-Cavalli *et al*., 1997; Tang and Hurley, 1997). Upon gene duplication, independent evolution led to the formation of a heterodimeric AC in eukaryotes before the monomers were fused in a single peptide chain. Evidently, the order of monomer linkage has occurred in either way. Only after this event did the ciliate GC evolve by a few point mutations in the purine-binding pocket from an ancestor AC module. All 10 GC isoforms detected in *Paramecium* and *Tetrahymena* by PCR are most likely to have evolved from a common ancestor because the sequences surrounding the substrate-determining serine are almost identical. On the other hand, the soluble heterodimeric GCs in mammals must have diverged from an ancestral cyclase homodimer even before the bulky membrane anchors were added and before heterodimers evolved, because the GCs have retained identical amino acids at positions which determine substrate specifity and at equivalent positions of a non-catalytic pocket (see Figure 2; Tesmer *et al*., 1997). In summary then, the ciliate GCs evolved from a membrane-anchored, heterodimeric AC with inverted catalytic centers, whereas mammalian GCs most probaby descended from a primordial cytosolic cyclase homodimer. The aquisition by the protozoan GCs of an ATPase-like domain

may have occurred before or after the change in substrate specificity.

Recently, the ciliates *Paramecium* and *Tetrahymena* were systematically reclassified as belonging to the parvkingdom 2, Alveolata, which is one out of seven protozoan parvkingdoms. The alveolates comprise the phyla Ciliophora, Dinozoa and Apicomplexa (Cavalier-Smith, 1993). The Apicomplexa contain many dreaded parasites such as *Plasmodium*, *Toxoplasma*, *Theileria* and *Babesia*. Currently, the *P.falciparum* genome with an A/T content of ~80% is being sequenced. The database search identified GCs with mammalian AC topology on chromosomes 11 and 13. Viewed together, our data strikingly complement the novel systematic classification of protozoans which is based on entirely different parameters, and support the notion of a common ancestor for *Paramecium*, *Tetrahymena* and the parasites of the phylum Apicomplexa, e.g. *Plasmodium*.

Materials and methods

Cloning of the guanylyl cyclase from Paramecium

A gDNA library in the λ ZAPII(*Eco*RI/CIAP) vector was generated from 100 ng of *Eco*RI-digested total DNA of *Paramecium tetraurelia 51s*. A cDNA library was established with 8.8 μg of poly(A)$^+$ RNA isolated from stationary phase *Paramecium* using the Messenger RNA Isolation Kit, the ZAP cDNA synthesis Kit and the ZAP Express *Eco*RI–*Xho*I Vector Cloning Kit (Stratagene). Using 500 ng of gDNA as a template, a 219 bp fragment suspected to code for a *Paramecium* AC (Hinrichsen *et al.*, 1995) was amplified by PCR using specific primers (2.5 mM Mg^{2+}, 40 cycles, annealing temperature 47°C; a ramp of 6°C/min from 47 to 72°C was used in the initial five rounds). The product was labeled with [α-^{32}P]dCTP (Rediprime DNA Labeling Kit, Amersham) and used as a probe for filter screening. Nine plates with 50 000 p.f.u./plate of the cDNA library were extracted individually with SM buffer (50 mM Tris–HCl pH 7.5, 15 mM $MgSO_4$, 100 mM NaCl). The 219 bp fragment was identified in two fractions by PCR. Further subfractionation and filter screening of the first fraction yielded a 4.5 kb cDNA and of the second positive fraction a 7.2 kb cDNA clone. To obtain the ATP start codon, we screened the gDNA library with a 290 bp probe located at the 5′ end of the 7.2 kb cDNA clone and a 2.4 kb gDNA clone was isolated (DDBJ/EMBL/GenBank accession No. AJ238860).

GC isoforms in Paramecium and Tetrahymena

The sense primer coding for KTYMAA and the antisense primers for G(D,N)T(V,I)NVA or GPTVN(V/M)A were used with gDNA as a template. Three additional GC isoforms were identified. Similar experiments were carried out using gDNA and cDNA from *Tetrahymena* as templates; sense primers coded for KTYMAA or KTYMAC and antisense primers for GDTINTA or G(D,N)T(V,I)NVA. Six further *Tetrahymena* GC isoforms were detected.

Cloning of a guanylyl cyclase from Tetrahymena

gDNA libraries from *Tetrahymena pyriformis* were created in λ ZAPII (Stratagene) using *Eco*RI- or *Mfe*I-restricted total DNA. The *Eco*RI gDNA library was screened with a 249 bp fragment as a probe obtained by PCR (see above). A 5 kb fragment coding for a *Tetrahymena* GC isoform was cloned. The remainder of the gDNA sequence of 9.7 kb was obtained by chromosome walking using the *Mfe*I-restricted gDNA library (DDBJ/EMBL/GenBank accession No. AJ242506). To reveal the exon–intron structure of the gene, overlapping cDNA fragments were produced by PCR using specific primers, and cDNA as a template. Thereby the full ORF of a *Tetrahymena* GC was reconstructed.

Computational analysis

BLASTP (Altschul *et al.*, 1997) searches were carried out online at the NCBI. Alignments were done by the CLUSTAL method (Higgins and Sharp, 1988) and refined by manual editing. Transmembrane spans were identified by the method of Goldman (Engelmann *et al.*, 1986) and by using Kyte–Doolittle plots.

Expression of Paramecium GC in HEK293 cells

Site-directed mutagenesis to convert 99 protozoan TAA/TAG-Gln codons to the universal CAA/CAG-Gln triplets was according to Deng and Nickoloff (1982). The 7.2 kb cDNA clone was cut into fragments of 0.2–1.2 kb, ligated into pBluescriptII SK(–) and mutated separately. A 0.6 kb fragment at the 5′ end was amplified by PCR using specific primers which introduced a 5′ *Ehe*I restriction site, a Kozak sequence and the bases for the four N-terminal amino acids not found in the 7.2 kb cDNA clone. Similarly, a 0.2 kb fragment of the 3′ end was amplified to eliminate the 3′-UTR region and generate a 3′ *Apa*I restriction site. The ORF was reconstituted in pBluescriptII SK(–). A list of the primers used in this study is available from J.E.S. on request.

Five DNA stretches (nucleotides 1–436, 1012–1387, 1973–2363, 3303–3991, 4261–4764 and 6001–6537) were redesigned for chemical synthesis by reverse translation of the corresponding amino acid sequence using the most common mammalian codon usage as defined by the Lasergene software package. The sequences were edited manually to remove potential strong secondary structures and RNA splice sites. Restriction sites were designed as needed for ligating the synthetic fragments to neighboring sequences. Synthesis was carried out by annealing, elongation and amplification with 0.1 μM template oligonucleotides and 0.5 μM amplification primers (Chio *et al.*, 1994). Twenty-one PCR cycles were run with an annealing temperature of 56°C. Synthetic stretches of the P-type ATPase-like domain were generated with 0.5 μM template primers without amplification primers. Eleven PCR cycles were run with an annealing temperature of 54°C. The synthetic DNA fragments were joined to the TAA/TAG-corrected parts of the ORF. The DDBJ/EMBL/GenBank accession No. of the synthetic gene is AJ239047.

The GC domain (nucleotides 4141–7239) was cloned into pBluescriptII SK(–) via the 5′ *Bam*HI site at position 4141 and the 3′ *Apa*I site. Using the oligonucleotide GCTAGCCGCCACCATGGATCC, we concomitantly created a start ATG preceded by a Kozak sequence and a 5′ *Nhe*I site. The ORF was excised from pBluescriptII SK(–) by *Nhe*I and *Apa*I and blunt ligated into the *Eco*RV site of pIRES1neo (Clontech).

HEK293 cells were grown at 37°C in minimum essential medium (Gibco-BRL #31095-029) containing 10% fetal calf serum (FCS) and 20 μg/ml gentamycin. For transfection, the calcium phosphate method was used with 10 μg of plasmid DNA/100 mm dish (Kingston *et al.*, 1992). Cells were selected with 1 mg/ml G418 starting 3 days after transfection. After 2 weeks, cells were rinsed once with phosphate-buffered saline (PBS; 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.4), suspended in 5 ml of PBS and collected. Cells were lysed by freeze–thaw in 20% (v/v) glycerol, 1% (v/v) thioglycerol, 20 mM KCl, 20 mM Tris–HCl, pH 7.5. The homogenate was used for cyclase assays.

Sf9 cell culture and recombinant baculoviruses

Sf9 cells were grown at 27°C in TC-100 medium with 10% FCS, 200 U/l penicillin and 200 μg/l streptomycin as a monolayer or in suspension (90 r.p.m., 0.1% pluronic F68). The ORF of the *Paramecium* GC holoenzyme, consisting of the partially resynthesized ATPase-like domain and the TAA/TAG-corrected cyclase domain, and the GC domain alone were inserted into pFastBac1, respectively. The BAC-TO-BAC baculovirus expression system from Gibco-BRL was used to produce recombinant virus.

Membranes were prepared from Sf9 cells 48 h after infection (m.o.i. 0.1). Cells were lysed in 3 mM EDTA, 3 mM EGTA, 12 mM HEPES pH 8.0 containing a protease inhibitor cocktail (1 tablet/10 ml Complete Mini from Roche Diagnostics) by two freeze–thaw cycles and repeated passages through a 0.45 mm needle. After removal of nuclei, membranes were pelleted, washed and resuspended in 1 mM EDTA, 16 mM HEPES pH 8.0, 5.7% sucrose and protease inhibitors.

Enzyme assays

AC activity was determined for 10 min at 37°C with 75 μM [α-^{32}P]ATP (Salomon *et al.*, 1974). The reactions contained 22% glycerol, 50 mM Tris–HCl, pH 7.5, 3 mM creatine phosphate, 0.23 mg of creatine kinase and 2 mM [2,8-^{3}H]cAMP (1.7×10^6 Bq/mmol). GC assay conditions were identical, except that cAMP and ATP were replaced by [8-^{3}H]cGMP and [α-^{32}P]GTP (Klumpp and Schultz, 1982).

Antibodies and Western blot analysis

Rabbit antibodies were generated against a synthetic N-terminal peptide (R5–N20), a cytosolic region of the ATPase-like region (L549–M718) and the C1b-positioned region (K1811–A2007) of the GC domain. Corresponding cDNAs were cloned in pQE8 (Qiagen) and expressed in

Escherichia coli M15[pRep4]. Peptides were affinity purified on Ni-NTA and by SDS–PAGE and used for immunization. Antibodies were affinity purified with antigen linked to NHS–agarose (Bio-Rad). A 10–20 μg aliquot of cell membrane protein was incubated in sample buffer for 2 h at 37°C prior to SDS–PAGE. Proteins were blotted onto nitrocellulose, transiently stained with Ponceau-S, incubated with the respective antibodies and detected by chemiluminescence.

Construction of functional Paramecium GC/type VII AC chimeras

In the remodeled ORF of the *Paramecium* GC domain, restriction sites were created by site-directed mutagenesis: *Spe*I at position 4258 with an L1421S change, *Xho*I at position 4765, *Avr*II at position 5986 with a K1997R change, *Sal*I at position 6532 with an N2179D change and a 3′-terminal *Asc*I site by extension of the ORF with the sequence GAGGCGCGCCAA. For constructs II and III (Figure 6), an *Mlu*I site was generated at position 5446 with an I817R change. Compatible sites were introduced into bovine type VII AC: *Spe*I at position 664, *Sal*I at 1183 with a Q220V change, *Mlu*I at 2032, *Avr*II at 2350 and *Xho*I at 3022 with I833L/D834E changes. Chimeras were generated by assembling the appropriate restriction fragments. Chimeric ORFs were cloned into the *Eco*RV site of pIRES1neo and expressed in HEK293 cells.

Immunocytochemical localization of guanylyl cylase in Paramecium

Cells were fixed in 4–8% formaldehyde, 0.1% glutaraldehyde and 1 mM $CaCl_2$. After ethanol dehydration and cooling, cells were embedded in hydrophilic methacrylates (Unicryl, British BioCell, or LR Gold, London Resin Co., London, UK) and UV polymerized at –20 or –35°C, respectively. Identical results were obtained with both methods and data were pooled. Ultrathin sections were incubated with affinity-purified anti-GC or anti-ATPase antibodies followed by incubation with gold conjugates of $F(ab)_2$ (5 nm, Au_5) from goat anti-rabbit antibodies (Aurion, Wageningen, The Netherlands). For further details, see Grothe *et al*. (1998). On electron micrographs of uranyl acetate-stained sections, gold grains per unit membrane length were determined by the hit point method (Plattner and Zingsheim, 1983).

Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft and by the Fonds der Chemischen Industrie. Sequence data for *P.falciparum* chromosome 13 were obtained from The Sanger Centre website at http://www.sanger.ac.uk/Projects/P_falciparum/. Sequencing of *P.falciparum* chromosome 13 was accomplished as part of the Malaria Genome Project with support by The Wellcome Trust.

References

Allen,G. and Green,N.M. (1976) A 31-residue peptide from the active site of the [Ca^{2+}]-transporting adenosine triphosphatase of rabbit sarcoplasmic reticulum. *FEBS Lett.*, **63**, 188–191.

Altschul,S.F., Madden,T.L., Schaffer,A.A., Zhang,J., Zhang,Z., Miller,W. and Lipman,D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**, 3389–3402.

Bakalyar,H.A. and Reed,R.R. (1990) Identification of a specialized adenylyl cyclase that may mediate odorant detection. *Science*, **250**, 1403–1406.

Barzu,O. and Danchin,A. (1994) Adenylyl cyclases: a heterogenous class of ATP-utilizing enzymes. *Progr. Nucleic Acid Res. Mol. Biol.*, **49**, 241–283.

Cali,J.J., Zwaagstra,J.C., Mons,N., Cooper,D.M.F. and Krupinski,J. (1994) Type VIII adenylyl cyclase. A Ca^{2+}/calmodulin-stimulated enzyme expressed in discrete regions of rat brain. *J. Biol. Chem.*, **269**, 12190–12195.

Cavalier-Smith,T. (1993) Kingdom protozoa and its 18 phyla. *Microbiol. Rev.*, **57**, 953–994.

Chio,C.L., Drong,R.F., Riley,D.T., Gill,G.S., Slightom,J.L. and Huff,R.M. (1994) D4 dopamine receptor-mediated signaling events determined in transfected Chinese hamster ovary cells. *J. Biol. Chem.*, **269**, 11813–11819.

Clarke,D.M., Loo,T.W. and MacLennan,D.H. (1990a) Functional consequences of mutations of conserved amino acids in the beta-strand domain of the Ca^{2+}-ATPase of sarcoplasmic reticulum. *J. Biol. Chem.*, **265**, 14088–14092.

Clarke,D.M., Loo,T.W. and MacLennan,D.H. (1990b) Functional consequences of alterations to amino acids located in the nucleotide binding domain of the Ca2(+)-ATPase of sarcoplasmic reticulum. *J. Biol. Chem.*, **265**, 22223–22227.

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.

Deng,W.P. and Nickoloff,J.A. (1992) Site-directed mutagenesis of virtually any plasmid by eliminating a unique site. *Anal. Biochem.*, **200**, 81–88.

Elwess,N.L. and van Houten,J.L. (1997) Cloning and molecular analysis of the plasma membrane Ca(2+)-ATPase gene in *Paramecium tetraurelia*. *J. Eukaryot. Microbiol.*, **44**, 250–257.

Engelmann,D.M., Steitz,T.A. and Goldman,A. (1986) Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. *Annu. Rev. Biophys. Biophys. Chem.*, **15**, 321–353.

Fagan,M.J. and Saier,M.J.,Jr (1994) P-type ATPases of eukaryotes and bacteria: sequence analyses and construction of phylogenetic trees. *J. Mol. Evol.*, **38**, 57–99.

Gonzalez,D.H. and Chan,R.L. (1993) Screening cDNA-libraries by PCR using λ-sequencing primers and degenerate oligonucleotides. *Trends Genet.*, **9**, 231–232.

Grothe,K., Hanke,C., Momayezi,M., Kissmehl,R., Plattner,H. and Schultz,J.E. (1998) Functional characterization and localization of protein phosphatase type 2C from *Paramecium*. *J. Biol. Chem.*, **273**, 19167–19172.

Hauser,K., Pavlovic,N., Kissmehl,R. and Plattner,H. (1998) Molecular characterization of a sarco(endo)plasmic reticulum Ca^{2+}-ATPase gene from *Paramecium tetraurelia* and localization of its gene product to sub-plasmalemmal calcium stores. *Biochem. J.*, **334**, 31–38.

Henikoff,S. and Henikoff,J.G. (1992) Amino acid substitution matrices from protein blocks. *Proc. Natl Acad. Sci. USA*, **89**, 10915–10919.

Higgins,D.J. and Sharp,P.M. (1988) CLUSTAL: a package for performing multiple sequence alignments on a microcomputer. *Gene*, **73**, 237–244.

Hinrichsen,R.D., Fraga,D. and Russell,C. (1995) The regulation of calcium in *Paramecium*. *Adv. Second Messenger Phosphoprotein Res.*, **30**, 311–338.

Kingston,R.E., Chen,C.A. and Okayama,H. (1992) Calcium phosphate transfection. In Ausubel,F.M., Brant,A., Kingston,R.E., Moore,P., Sudman,J.G., Smith,J.A. and Struhl,K. (eds), *Short Protocols in Molecular Biology*. John Wiley and Sons, New York, pp. 9.3–9.7.

Klumpp,S. and Schultz,J.E. (1982) Characterization of a Ca^{2+}-dependent guanylate cyclase in the excitable ciliary membrane from *Paramecium*. *Eur. J. Biochem.*, **124**, 317–324.

Koesling,D. and Friebe,A. (1999) Soluble guanylyl cyclase: structure and regulation. *Rev. Physiol. Biochem. Pharmacol.*, **135**, 41–65.

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

Laken,S.J. *et al*. (1997) Familial colorectal cancer in Ashkenazim due to a hypermutable tract in APC. *Nature Genet.*, **17**, 79–83.

Mons,N., Yoshimura,M., Ikeda,H., Hoffman,P.L. and Tabakoff,B. (1998) Immunological assessment of the distribution of type VII adenylyl cyclase in brain. *Brain Res.*, **788**, 251–261.

Nakane,M., Arai,K., Saheki,S., Kuno,T., Buechler,W. and Murad,F. (1990) Molecular cloning and expression of cDNAs coding for soluble guanylate cyclase from rat lung. *J. Biol. Chem.*, **265**, 16841–16845.

Plattner,H. and Zingsheim,H.P. (1983) Electron microscopic methods in cellular and molecular biology. *Subcell. Biochem.*, **9**, 1–236

Preer,J.R.,Jr, Preer,L.B., Rudman,B.M. and Barnett,A.J. (1991) Deviation from the universal code shown by the gene for surface protein 51A in *Paramecium*. *Nature*, **314**, 188–190.

Rees,S., Coote,J., Stables,J., Goodson,S., Harris,S. and Lee,M.-G. (1996) Bicistronic vector for the creation of stable mammalian cell lines that predisposes all antibiotic-resistant cells to express recombinant protein. *Biotechniques*, **20**, 102–110.

Salomon,Y., Londos,C. and Rodbell,M. (1974) A highly sensitive adenylate cyclase assay. *Anal. Biochem.*, **58**, 541–548.

Schultz,J.E. and Klumpp,S. (1982) Characterization of a Ca^{2+}-dependent guanylate cyclase in the excitable ciliary membrane from *Paramecium*. *Eur. J. Biochem.*, **124**, 317–324.

Schultz,J.E., Pohl,T. and Klumpp,S. (1986) Voltage-gated Ca^{2+} entry into *Paramecium* linked to intraciliary increase in cyclic GMP. *Nature*, **322**, 271–273.

Schultz,J.E., Klumpp,S., Benz,R., Schürhoff-Goeters,W.J.CH. and Schmid,A. (1992) Regulation of adenylyl cyclase from *Paramecium* by an intrinsic ion conductance. *Science*, **255**, 600–603.

Schulz,S., Yuen,P.S. and Garbers,D.L. (1991) The expanding family of guanylyl cyclases. *Trends. Pharmacol. Sci.*, **12**, 116–120.

Stokes,D.L., Taylor,W.R. and Green,N.M. (1994) Structure, transmembrane topology and helix packing of P-type ion pumps. *FEBS Lett.*, **346**, 32–38.

Sunahara,K.S., Dessauer,C.W. and Gilman,A.G. (1996) Complexity and diversity of mammalian adenylyl cyclases. *Annu. Rev. Pharmacol. Toxicol.*, **36**, 461–480.

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. Biol. 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.

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 $G_{s\alpha}$GTPγS. *Science*, **278**, 1907–1916.

Toyoshima,C., Sasabe,H. and Stokes,D.L. (1993) Three-dimensional cryo-electron microscopy of the calcium ion pump in the sarcoplasmic reticulum membrane. *Nature*, **362**, 469–471.

Trouet,D., Nilius,B., Voets,T., Droogmans,G. and Eggermont,J. (1997) Use of a bicistronic GFP-expression vector to characterise ion channels after transfection in mammalian cells. *Pflügers Arch.*, **434**, 632–638.

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.

Völkel,H., Beitz,E., Klumpp,S. and Schultz,J.E. (1996) Cloning and expression of a bovine adenylyl cyclase type VII specific to the retinal pigment epithelium. *FEBS Lett.*, **378**, 245–249.

Yu,S., Avery,L., Baude,E. and Garbers,D.L. (1997) Guanylyl cyclase expression in specific sensory neurons: a new family of chemosensory receptors. *Proc. Natl Acad. Sci. USA*, **94**, 3384–3387.

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

Zhang,P.J., Toyoshima,C., Yonekura,K., Green,N.M. and Stokes,D.L. (1998) Structure of the calcium pump from sarcoplasmic reticulum at 8 Å resolution. *Nature*, **392**, 835–839.

Received May 31, 1999; revised and accepted June 15, 1999