Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray analysis of the catalytic domain of the adenylate cyclase GRESAG4.1 from *Trypanosoma brucei*

Boris Bieger and Lars-Oliver Essen*

Max-Planck-Institute for Biochemistry, Department of Membrane Biochemistry, Am Klopferspitz 18a, D-82152 Martinsried bei München, Germany

Correspondence e-mail: essen@biochem.mpg.de

Received 30 October 1999
Accepted 5 January 2000

Adenylate cyclases (ACs) are involved in signal transduction by generating the second messenger, cAMP. In *Trypanosoma brucei*, 3′,5′-cyclic adenosine monophosphate (cAMP) controls the life cycle of this unicellular parasite. cAMP is generated by a class of adenylate cyclases which are either constitutively (GRESAG4.1–4.3) or transiently expressed (ESAG4) during the life cycle. Unlike mammalian ACs, the trypanosomal ACs have a simple topology comprising of a large extracellular region, a transmembrane helix and a cytosolic catalytic region. Two orthorhombic crystal forms of the catalytic AC domain of GRESAG4.1 (residues Ala884–Thr1132) were generated by the hanging-drop vapour-diffusion method. X-ray diffraction data from GRESAG4.1 crystals were collected at 1.9 Å resolution using synchrotron radiation. Furthermore, two heavy-metal derivative data sets were collected from crystal form *A*; heavy-atom sites were subsequently located in difference Patterson maps.

1. Introduction

Adenylate cyclases (ACs) play a crucial role in signal-transduction pathways of eukaryotic organisms by catalyzing the generation of a second messenger, 3′,5′-cyclic adenosine monophosphate (cAMP). There are three eukaryotic classes of adenylate cyclases with distinct architectures (Tang & Gilman, 1992). Class I cyclases occur in higher eukaryotes and comprise a 12-transmembrane helix motif with two cytosolic regions which contain the catalytically active portion. Class II cyclases consist of an extracellular domain with putative receptor function, a transmembrane helix and a cytosolic catalytic domain, while class III cyclases are large solely membrane-associated proteins (Fig. 1).

The transition of *T. brucei*, a bloodstream parasite and causative agent of sleeping sickness in humans, from a long slender form proliferating to the non-dividing short stumpy form is triggered by the activation of class II adenylate cyclases that subsequently raises cytosolic cAMP levels (Vassella *et al.*, 1997). *T. brucei* contains a set of membrane-bound ACs (Alexandre *et al.*, 1996) which are specifically expressed in the mammalian bloodstream (ESAG4; expression site-associated gene) or constitutively during the whole life cycle (GRESAG4.1–4.3; genes related to ESAG). ESAG4 is not encoded by a single gene, but represents a multi-gene family that occurs on the variable surface glycoprotein (VSG) expression sites with at least six copies per cell. Likewise, nine genes were found for the adenylate cyclases GRESAG4.1 which code for gene products that are highly homologous to ESAG4. Consequently, the difficulty of generating *T. brucei* mutants which are devoid of all these multi-gene ACs hampers a concise understanding of the role and regulation of trypanosomal ACs *in vivo*.

Trypanosomal ACs (tACs) are topologically more closely related to membrane-bound guanylate cyclases (pGC) than to mammalian class I ACs, with which they share less than 25% sequence homology over the catalytic domains (220–250 amino-acids in length; Fig. 2). The latter enzymes contain two catalytic domains, C1 and C2. When associated in a soluble C1–C2 heterodimer, these domains exhibit similar catalytic activities as the intact membrane-bound class I enzyme (Whisnant *et al.*, 1996). Low- to medium-resolution studies on a forskolin-linked C2-homodimer (Zhang *et al.*, 1997) or a catalytically competent quarternary complex made of $G_{S\alpha}$, C1 and C2 domains and forskolin (Tesmer *et al.*, 1997) indicated that the active site of class I ACs is located along the C1–C2 interface. Based on this and mutational studies, a common activation mechanism by homo- or heterodimerization was postulated to be valid not only for class I, but also for other adenylate cyclases including tACs and some bacterial ACs (Liu *et al.*, 1997). Various residues which were found to be essential for catalysis in mammalian enzymes are indeed conserved in all nucleotidyl cyclase classes (Fig. 2), thus indicating

that the overall organization of active sites in tACs resembles those of mammalian ACs.

The much simpler architecture of trypanosomal class II ACs compared with mammalian ACs makes these proteins very attractive for studies on general structure–function relationships in nucleotidyl cyclases. Furthermore, the structural peculiarities of tACs might point to the physiological roles of this multi-gene family, which are otherwise difficult to study by cell and molecular biological approaches. Consequently, a structural base for the hitherto unknown tAC activation could provide a link to a future pharmacological control of trypanosomal proliferation. In this work, we report the crystallization and preliminary X-ray crystallographic analysis of the catalytic AC domain of GRESAG4.1.

2. Experimental procedures

2.1. Plasmid construction

Gene fragments which code for the cytosolic region of GRESAG4.1 (residues Ala884–Tyr1242) were amplified from genomic DNA of the *T. brucei* strain 927 using Takara Taq polymerase (BioWhittaker Europe) and 30 cycles at 30 s/367 K, 2 min/331 K and 5 min/345 K. For amplification, the primer pairs 5′-CTCG**CTCGAG**CTATGAAGGTGCATGTC and 5′-GCTGTGATCA**GCTAGC**GGTGCGGAACGTCGTAACAAC were used for the GRESAG4.1 gene (restriction sites in bold). The DNA fragment for GRESAG4.1 (1.11 kbp) was digested with *Nhe*I and *Xho*I and ligated into the appropriately cut expression vector pET-28a (Invitrogen) using standard cloning techniques. The resulting construct, pET-gres4.1cyto, was verified by dideoxy sequencing.

2.2. Expression, purification and refolding of the adenylate cyclase fragments

The plasmid pET-gres4.1cyto was transformed into *Escherichia coli* BL21(DE3) cells. *E. coli* cells which harboured the desired construct were grown in LB medium containing 50 µg ml^{-1} kanamycin at 303 K. At an OD_{595} of 0.5–0.6, the cells were induced by the addition of 500 µ*M* IPTG. After 6 h expression at 303 K, the cells were harvested, resuspended in 1/100 th of the culture volume with 1× PBS, 2 m*M* DTT, 0.2 m*M* PMSF, 1 m*M* EDTA and lysed by two passes through a French press at 110 MPa. DNAase was added to the lysate, which was sonified by three continuous pulses of 20 s duration (Branson Sonifier 450). Inclusion bodies were collected by centrifuging the suspension at 25 000*g* for 30 min and resuspended in 6 *M* guanidinium chloride pH 8.0. The suspension was clarified by centrifugation at 25 000*g* for 30 min. The supernatant was loaded onto a 15 ml nickel–nitrilotriacetic acid (Ni^{2+}–NTA) column (Qiagen) that had been pre-equilibrated with washing buffer (8 *M* urea, 20 m*M* imidazole hydrochloride pH 8.0). The column was washed with 40 volumes of washing buffer. The denatured adenylate cyclase fragments were eluted with 8 *M* urea pH 8.0, 250 m*M* imidazole. The eluate fractions were then brought to 6 *M* guanidine hydrochloride, 0.1 *M* dithiothreitol (DTT), 0.4 *M* arginine hydrochloride pH 8.0. The adenylate cyclase fragments were refolded by a preliminary 1:20 dialysis against 20 m*M* Tris–HCl pH 8.0, 0.4 *M* arginine at 277 K for 12 h. The solution was then dialysed 1:20 against 20 m*M* Tris–HCl pH 8.0, 0.5 m*M* DTT at 277 K for 6 h and finally brought to a protein concentration of 3 mg ml^{-1}, before the renatured cyclase fragments were loaded onto a Superdex 200 gel-filtration column (Pharmacia). Gel filtration was carried out at 277 K using 20 m*M* Tris–HCl pH 8.0, 0.6 *M* NaCl, 0.5 m*M* DTT as elution buffer. Protein concentrations were determined using the Bradford reagent (Biorad) and bovine serum albumin as standard.

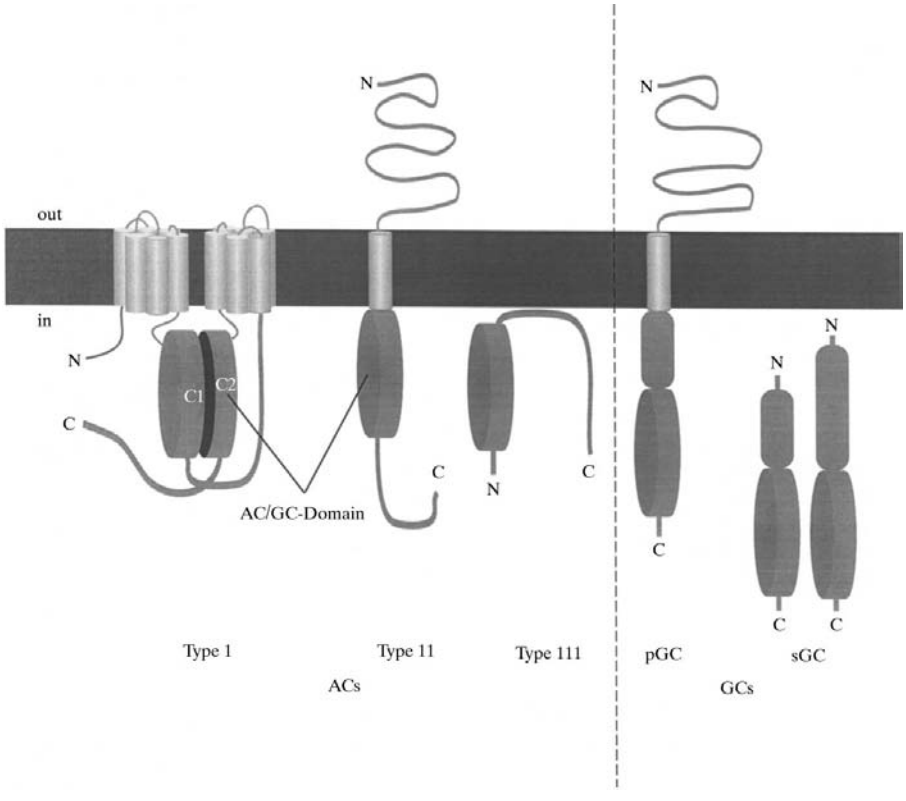


Figure 1
Topology sketch of the different adenylate and guanylate cyclase classes.

2.3. Limited proteolysis and mass spectrometry

To reduce the cytosolic fragment of GRESAG4.1 to the appropriate catalytic domain, the GRESAG4.1-(Ala884–Tyr1242) fragment was incubated at a concentration of 0.5 mg ml^{-1} with 0.05 mg ml^{-1} subtilisin for 30 min at 277 K. The cleavage reaction was carried out in 20 m*M* Tris–HCl pH 8.0 as reaction buffer and stopped by adding 10 m*M* PMSF. The solution was directly loaded onto a Resource Q column (Pharmacia) and the column was washed with five volumes of the reaction buffer. The catalytic domain of GRESAG4.1 was eluted with 60 ml of a linear NaCl gradient (0–600 m*M*). The eluted fractions were concentrated to a final volume of 1.0 ml and loaded onto a Superdex 75 gel-filtration column (Pharmacia) in 20 m*M* Tris–HCl pH 8.0, 0.6 *M* NaCl.

The purified tAC domain was N-terminally sequenced by automated Edman degradation and its molecular mass was analysed on an electrospray mass spectrometer (ABI).

Table 1
Data collection.

	GRESAG4.1 form *A*	Me_3PbOAc form *A*	MeHgOH form *A*	GRESAG4.1 form *B*
X-ray source and wavelength (Å)	BW6, 1.07	BW6, 1.07	BW6, 1.07	Cu $K\alpha$
Space group	$P2_12_12_1$			$P2_12_12$
Unit-cell dimensions (Å)	a = 49.9, b = 60.1, c = 79.7			a = 123.8, b = 35.8, c = 59.3
Resolution (Å)	15–1.9	15–1.6	15–2.3	10–2.1
Observations	64207	97405	38256	40577
Unique reflections	19109	31193	10880	15006
Mosaicity (°)	0.8	0.4	0.5	0.6
Wilson B factor (Å^2)	14.6	16.3	19.3	23.0
$I/\sigma(I)$	15.0 (6.0)	22.3 (11.5)	24.0 (11.1)	21.6 (3.6)
R_{merge}	0.079 (0.268)	0.047 (0.091)	0.079 (0.177)	0.052 (0.309)
Completeness (%)	98.7 (98.9)	96.7 (80.8)	99.3 (99.6)	88.2 (68.8)

2.4. Adenylate cyclase assays

Adenylate cyclase activities were assayed in 150 µl final volumes for 30 min at 298 K with 0.5 µg of the recombinant proteins in 20 m*M* Tris–HCl pH 8.0. The reaction was started by adding a 3× reaction ready-mix containing 2 m*M* ATP, 40 m*M* DTT, 5 m*M* $MnCl_2$ to the protein solution and stopped by adding 200 µl of a 1:1 mixture of $CHCl_3$/MeOH (Voss & Wallner, 1992). cAMP synthesis was measured with a commercial cAMP assay kit (Amersham).

2.5. Crystallization and data collection

Crystals of the subtilisin-cleaved fragment of GRESAG4.1 were grown in hanging drops using vapour diffusion at 291 K and mixing 1 µl protein solution of GRESAG4.1 (4 mg ml^{-1} in 5 m*M* Tris–HCl pH 8.0) with 1 µl of the corresponding crystallization buffer. Two different crystal forms were obtained for the GRESAG4.1-(Ala884–Thr1132) fragment: crystal form *A* was grown in 1.8–1.9 *M* ammonium sulfate between pH 6.5 and 7.75, while form *B* crystals grew in 1.2 *M* sodium citrate pH 5.4 as crystallization buffer.

Diffraction data were collected on a Rigaku RU-200 Cu $K\alpha$ rotating-anode generator with focusing mirrors (Charles Supper) and a MAR Research imaging plate. Synchrotron data for crystal form *A* were collected from single frozen crystals of GRESAG4.1-(Ala884–Thr1132) at beamline BW6, HASYLAB, Hamburg at 100 K using a MAR CCD detector and a wavelength of 1.07 Å. Data processing was carried out using the *HKL* package (Otwinowski & Minor, 1997). Data statistics are given in Table 1.

3. Results and discussion

As in trypanosomes, the protein-coding genes are devoid of class II introns (Vanhamme & Pays, 1995); the gene fragment coding for the cytosolic region of GRESAG4.1 (residues Ala884–Tyr1242) could be isolated by PCR from genomic DNA of *T. brucei* strain 927 and directly cloned into the bacterial expression vector pET-28a (Invitrogen). In comparison with the corresponding SWISSPROT entry CY41_TRYBB from *T. brucei* strain AnTat1.3, two differences were found for the GRESAG4.1 sequence, which were confirmed by repeated PCR amplifications and sequencing of both strands. The revised GRESAG4.1 sequence misses the codon for Ile1032 and codes for the exchange Tyr1121Ile. The observed amino-acid exchanges were found in conserved regions of the genes and led to an improved sequence homology with other tACs.

Recombinant GRESAG4.1-(Ala884–Tyr1242) was obtained as inclusion bodies after 6 h of expression at 303 K, but could be efficiently refolded by dialysis. Therefore, inclusion bodies were solubilized in the presence of 6 *M* guanidinium chloride and purified by metal-chelating affinity chromatography using Ni^{2+}–NTA columns (Qiagen). After refolding and gel filtration, yields of purified protein were 4–10 mg per litre of culture, with a purity greater than 95%.

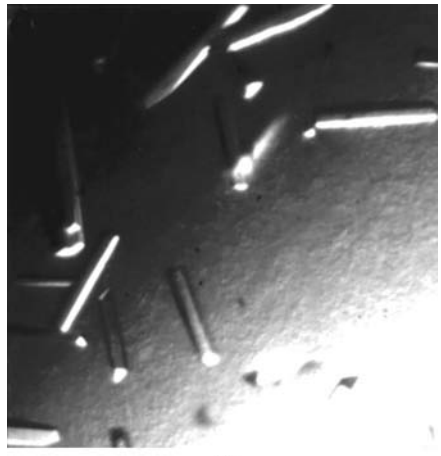
```
                              ** ***                                          * *
cyg1_hum   412 VPAKRYDNVT ILFSGIVGFN AFCSKHGAMK IVNLLNDLYT RFDTLTDSRK NPFVYKVETV  461
cya22_rat  877 LYHQSYDCVC VMFASIPDFK EFYTESEGLE CLRLLNEIIA DFDDLLSKPK FSGVEKIKTI  926
cya51_rat  461 IYIQKHDNVS ILFADIEGFT SLASQCTAQE LVMTLNELFA RFDKLAAENH ...CLRIKIL  527
cy41_try   897 APKEPTDPVT LIFTDIESST ALWAAHPDLM PDAVAAHHRM VRSLIGRYKC ....YEVKTV  942
```

(*a*)

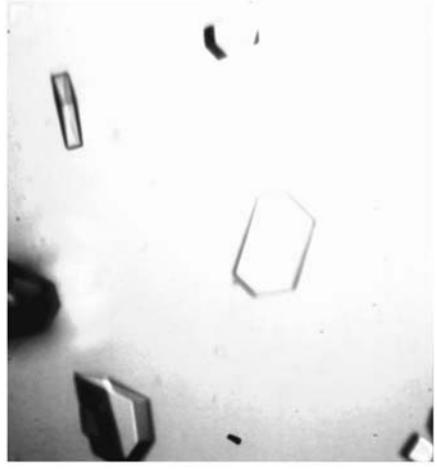
```
                     *                          * *     *    *
cyg1_hum   535 .GES.VQITI GIHTGEVVTG VIGQRMPRYC LFGNTVNLTS RTETTGEKGK INVSEYTYRC  593
cya22_rat  990 SFNDF.KLRV GINHGPVIAG VIGAQKPQYD IWGNTVNVAS RMDSTGVLDK IQVTEETSLI 1049
cya51_rat  571 ..GVNVNMRV GIHSGRVHCG VLGLRKWQFD VWSNDVTLAN HMEAGGKAGR IHITKATLNY  629
cy41_try  1010 LWNG.LRVRV GIHTGLCDIR .HDEVTKGYD YYGRTPNMAA RTESVANGGQ VLMTHAAYMS 1068
```

(*b*)

Figure 2
Sequence alignment for (*a*) the N-terminal and (*b*) the C-terminal regions of the AC/GC domains from human soluble GC (sGC; SWISSPROT entry CYG1_HUM, β1 subunit), rat C1 domain (CYA51_RAT), rat C2 domain (CYA22_RAT) and GRESAG4.1 (CY41_TRYBB). Asterisks mark the positions of conserved residues which were found to be essential for catalysis in class I enzymes.



(*a*)



(*b*)

Figure 3
Crystals of the catalytic domain of the trypanosomal adenylate cyclase GRESAG4.1 in the orthorhombic crystal forms *A* (*a*) and *B* (*b*). Typical crystal dimensions are 0.20 × 0.04 × 0.04 mm for form *A* and 0.12 × 0.07 × 0.07 mm for form *B*.

Extensive screening of crystallization conditions for GRESAG4.1-(Ala884–Tyr1242) were largely unsuccessful owing to a pronounced tendency towards aggregation. Consequently, limited proteolytic digestion was performed to generate truncated products with improved crystallization characteristics. The cytosolic region of GRESAG4.1 (molecular weight 40.2 kDa) could be proteolytically truncated to an AC domain that showed an apparent molecular weight of 30 kDa on 15% SDS–PAGE. The specific activities of the AC domain were in the range 3–4 nmol min^{-1} mg^{-1}, like the non-digested GRESAG4.1-(Ala884–Tyr1242) protein. Analysis of the GRESAG4.1 cleavage product by Edman degradation and ESI-MS indicated that the C-terminus ended at Thr1132, while the N-terminal His-tag was clipped off with the protein starting at Ala884. Consequently, the last 110 C-terminal residues of GRESAG4.1 are dispensable for catalytic activity. This region is unique to tACs and might merely serve a regulatory function.

The AC domain of GRESAG4.1 behaved much better in crystallization screens than the whole cytosolic region. Two orthorhombic crystal forms could be generated using the hanging-drop vapour-diffusion method (Fig. 3). Crystal form *A* of GRESAG4.1-(Ala884–Thr1132) was indexed in space group $P2_12_12_1$, with unit-cell parameters $a = 49.9$, $b = 60.1$, $c = 79.7$ Å. Native crystals diffracted to a resolution limit of 1.9 Å with a mosaicity of 0.4–0.8°. The crystals have one monomer per asymmetric unit, with 30% solvent content as calculated by the Matthews equation (Matthews, 1968). Crystal form *B* appeared when citrate was used as precipitant and belongs to space group $P2_12_12$, with unit-cell parameters $a = 121.4$, $b = 35.8$, $c = 59.3$ Å and one monomer per asymmetric unit. These crystals diffract to 2.1 Å resolution using Cu $K\alpha$ radiation.

For structure determination, X-ray diffraction data were collected from single frozen crystals of form *A* at 100 K using synchrotron radiation and 2.0 *M* ammonium sulfate pH 7.0, 20%(*v*/*v*) glycerol as cryo-protectant. Initial attempts to solve the structure of this tAC by molecular replacement using the previously determined structures of mammalian C1 and C2 domains (Tesmer *et al.*, 1997) failed. Subsequently, heavy-metal derivatives of crystal form A of GRESAG4.1-(Ala884–Thr1132) were prepared by incubating the crystals in cryo-protection buffer for 4–8 h in the presence of 10 m*M* of the corresponding heavy-metal salt.

The major heavy-atom sites for each of the Me_3PbOAc and the MeHgOH derivatives were determined by interpretation of the Harker section in isomorphous difference Patterson maps (Collaborative Computational Project, Number 4, 1994). Two sites were found for the mercury derivative at $x = 0.77$, $y = 0.39$, $z = 0.92$ and $x = 0.04$, $y = 0.00$, $z = 0.11$ and two sites for the trimethyllead derivative at $x = 0.32$, $y = 0.59$, $z = 0.04$ and $x = 0.17$, $y = 0.25$, $z = 0.21$. Additional minor sites were determined by difference Fourier analyses (one for Me_3Pb, two for the MeHg derivative).

Current efforts focus on collecting further derivatives, as phase refinement by solvent-flattening is ineffective owing to the low solvent content of the GRESAG4.1 crystals. The high-resolution structures of catalytically active trypanosomal ACs promise to yield insights into how the family of nucleotidyl cyclases diversified in their function and regulation during evolution.

This work is supported by a predoctoral fellowship (BB) of the Fonds der Chemischen Industrie. The authors are very grateful for the support by H. Bartunik at synchrotron beamline BW6, MPG-ASMB, Hamburg and J. Kellermann for protein sequencing.

References

Alexandre, S., Paindavoine, P., Hanocq-Quertier, J., Paturiaux-Hanocq, F., Tebabi, P. & Pays, E. (1996). *Mol. Biochem. Parasitol.* **77**, 173–182.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Liu, Y., Ruoho, A. E., Rao, V. D. & Hurley, J. H. (1997). *Proc. Natl Acad. Sci. USA*, **94**, 13414–13419.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Tang, W. J. & Gilman, A. G. (1992). *Cell*, **70**, 869–872.
Tesmer, J. J., Sunahara, R. K., Gilman, A. G. & Sprang, S. R. (1997). *Science*, **278**, 1907–1916.
Vanhamme, L. & Pays, E. (1995). *Microbiol. Rev.* **59**, 223–240.
Vassella, E., Reuner, B., Yutzy, B. & Boshart, M. (1997). *J. Cell Sci.* **110**, 2661–2671.
Voss, T. & Wallner, E. (1992). *Anal. Biochem.* **207**, 40–43.
Whisnant, R. E., Gilman, A. G. & Dessauer, C. W. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 6621–6625.
Zhang, G., Liu, Y., Qin, J., Vo, B., Tang, W. J., Ruoho, A. E. & Hurley, J. H. (1997). *Protein Sci.* **6**, 903–908.