Synthesis and degradation of cAMP in *Giardia lamblia*: Possible role and characterization of a Nucleotidyl Cyclase with a single Cyclase-Homology-Domain

Vanina Saraullo[1]; Nicolas Di Siervi[2], Belen Jerez[1], Carlos Davio[2] and Adolfo Zurita[*1]

[1] Área de Biología Molecular, Departamento de Bioquímica y Ciencias Biológicas, UNSL and Instituto Multidisciplinario de Investigaciones Biológicas de San Luis (IMIBIO-SL-CONICET), San Luis, Argentina.

2 CONICET-Universidad de Buenos Aires. Instituto de Investigaciones Farmacológicas (ININFA), Facultad de Farmacia y Bioquímica, Ciudad Autónoma de Buenos Aires, Argentina.

Corresponding author:
azurita1974@gmail.com (AZ)

Keywords. *G. lamblia;* cAMP; Adenylyl cyclase; Phosphodiesterase.

Abbreviations. NC= Nucleotidyl cyclase; AC= Adenylyl cyclase; GC= Guanylyl cyclase; AGC Adenylyl/ Guanylyl cyclase; PDE= Phosphodiesterase; CD= catalytic domain; CHD= Cyclase Homology Domain; cAMP= cyclic AMP; cGMP= cyclic GMP; ATP= Adenosine triphosphate; GTP = Guanosine triphosphate;

Funding

This study was supported by the Agencia Nacional de Promoción Científica y Tecnológica [PICT 2013-2050 and PICT 2011-0672] and the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) [PIP-2011].

Abstract

Despite its importance in the regulation of growth and differentiation processes of a variety of organisms, the mechanism of synthesis and degradation of cAMP has not yet been described in *Giardia lamblia*.

In this work, we measured significant quantities of cAMP in trophozoites of *G. lamblia*

incubated *in vitro* and later detected how it increases during the first hours of encystation, and how it then returns to basal levels at 24 h.

Through an analysis of the genome of *G. lamblia*, we found sequences of three putative enzymes -one phosphodiesterase (gPDE) and two nucleotydyl cyclases (gNC1 and gNC2)- that should be responsible for the regulation of cAMP in *G. lamblia*. Later, an RT-PCR assay confirmed that these three genes are expressed in trophozoites.

The bioinformatic analysis indicated that gPDE is a transmembrane protein of 154 kDa, with a single catalytic domain in the C-terminal end; gNC1 is predicted to be a transmembrane protein of 74 kDa, with only one class III cyclase-homology-domain (CHD) at the C-terminal end; and gNC2 should be a transmembrane protein of 246 kDa, with two class III CHDs.

Finally, we cloned and enriched the catalytic domain of gNC1 (gNC1cd) from bacteria. After that, we confirmed that gNC1cd has adenylyl cyclase activity. This enzymatic activity depends on the presence of Mn^{2+} and Ca^{2+}, but no significant activity was displayed in the presence of Mg^{2+}.Additionally, the AC activity of gNC1cd is competitively inhibited with GTP, so it highly possible that gNC1 has guanylyl cyclase activity as well.

Introduction

Giardia lamblia (syn. *G. duodenalis* or *G. intestinalis*) is a flagellated, binucleated parasitic protozoan that inhabits in the intestine of humans and domestic animals [1] and its infection (giardiasis) is one of the leading causes of human intestinal diseases worldwide, the most frequent reason of defined waterborne outbreaks of diarrhea in developed countries, as well as a common cause of diarrhea in day-care centers, institutionalized individuals, backpackers and travelers [2,3].

During its life cycle, this parasite exists in two forms: the trophozoite -replicative form-, that grows in the intestine of the host and is responsible for the infection symptoms, and the cyst -infective form- that is relatively inert and allows the parasite to survive outside the host and resist harsh environmental condition [1].

In order to survive, *G. lamblia* shows two adaptive differentiation processes: the encystation/excystation and the antigenic variation of their surface proteins [1,4,5]. With respect to the first process, it is possible to cultivate trophozoites of *G. lamblia in vitro* and to replicate its

whole life cycle in the laboratory by using growth medium supplemented with serum, high concentrations of L-cysteine in an oxygen-poor environment [1,6,7]. However, although the life cycle of *G. lamblia* and the physiological conditions that induce the processes of encystation and excystation have been described in detail, the molecular mechanisms of how this organism is able to sense the environment and respond to a given stimulus it is still unknown.

The second messenger cAMP is present in both eukaryotes and prokaryotes and has a significant role in intracellular signal transduction [8]. It is synthesized from ATP by the enzyme Adenylyl Cyclase (AC) and is negatively regulated by the enzyme cAMP-phosphodiesterase (PDE) [9], which converts the cAMP to 5'AMP. The cAMP mechanism of action in eukaryotes consists mainly of activation of protein kinase A (PKA) and Epac protein (G-protein-activating Rap monomer), which are capable to regulate many cellular processes such as proliferation, differentiation and apoptosis [10–12].

As observed in previous *in vitro* studies [13], the starting process of excystation of *G. lamblia* seems to be regulated by post-translational modifications, since it is a rapid event that occurs within 10 minutes. According to this observation, it has been described that excystation of *G. lamblia* was significantly repressed when the specific inhibitor of PKA, amide 14-22, was used during the first stage of the excystation *in vitro* [14], indicating that cAMP may have an important role in this process. Interestingly, and despite the importance of cAMP as a mediator of numerous intracellular processes in eukaryotic cells, no AC and PDE have been specifically described so far in *G. lamblia,* with the exception of the full meta-analysis of the genome of *G. lamblia* (WB strain clone C6) performed by Morrison et al. (2007) [15], in which 6470 ORFs were automatically identified. So within all these ORFs, we found the sequence of two class III nucleotidyl cyclases (NC) and one PDE, which ,despite their original names, we labeled gNC1, gNC2 and gPDE, respectively.

In this work, we analyzed the expression of ORFs and performed an *in silico* characterization of gNC1, gNC2 and gPDE of *G. lamblia*. Additionally, we cloned and enriched from bacteria the catalytic domain of gNC1 (gNCcd) in order to identify and confirm the biochemical properties of this enzyme. All of this constitute a first step to the characterization of the molecular mechanisms involved in the synthesis and degradation of cAMP in *G. lamblia*.

Experimental

Cell lines and reagents

Trophozoites of *G. lamblia WB* strain were kindly provided by Dr. Maria Carolina Touz laboratory from the Instituto de Investigación Médica Mercedes y Martín Ferreyra, INIMEC - CONICET, Universidad Nacional de Córdoba, Córdoba, Argentina.

ATP, GTP and Anty-poly-Histidine-Peroxidase antibody were purchased from Sigma-Aldrich; $MgCl_2$ and EDTA were purchased from Merck; $CaCl_2$ was purchased from Biopack; $MnCl_2$ was from Cicarelli; TYI-S-33 growth medium was prepared from Fetal adult calfbovine serum (Natocor), BBL Biosate peptone (Becton dickinson and company), L-Ciesteine (Dalton), Ascorbic acid (Cicarelly), Bovine bile (Britania) and ammonium ferric citrate (kindly donated by Dr. Maria Carolina Touz); porcine bile extract porcine for encystation medium was from Sigma-Aldrich.

All the other chemicals and biochemicals were of the highest purity commercially available and used without further purification.

Bioinformatic analysis

The complete sequence of *G. lamblia*'s nucleotidyl cyclases and phosphodiesterase were obtained from the Giardia Data Bank (giardiadb.org/giardiadb/) [15, 16]. The multiple sequence alignment and visualization of sequences was performed by *MAFFT_7* (mafft.cbrc.jp) [17]. The detection of domains in protein sequences was performed with the on line tools *S.M.A.R.T* (smart.embl-heidelberg.de/) [18] and *TMpred* (www.ch.embnet.org/software/TMPRED_form.html). Structural model of thiese domains were performed using the *PHYRE2 Protein Fold Recognition Server* (www.sbg.bio.ic.ac.uk/phyre2/) [19] and visualized with *The PyMOL Molecular Graphics System*, (Version 1.6 Schrödinger, LLC).

Cell Culture

Trophozoites of *G. lamblia WB* strain and competent bacteria Top10 and BL21(DE3)pLysS were used.

Top10 and BL21(DE3)pLysS competent bacterias were cultured in LB (Luria-Bertani) or LB/agar media supplemented with antibiotics (when appropriate) at 37°^{O}C.

G. lamblia trophozoites were cultured in TYI-S-33 growth medium (pH 7.0), supplemented with 10% adult bovine serum and 0.5 mg/ml bovine bile at 37°C in sealed tubes to ensure anaerobic conditions [6]. Trophozoites grow attached to the tube walls and can be collected by placing the tube oin ice for 10 to 15 minutes and inverting several times until the cells detach from the walls of the same.

The encystation of trophozoites was performed according to Boucher and Gillin (1990) [20] with minimal modifications, where the cells were incubated first for 24 h at 37°C with TYI-S-33 growth medium without bovine bile (pre-encystation medium). This was followed by incubation with encystation medium, which consists of pre-encystation medium with pH adjusted to 7.8 and supplemented with 0.25 mg/ml of porcine bile and 5 mM lactic acid.

RT-PCR

The total RNA was isolated from trophozoites incubated in normal medium or encystation medium, was extracted with Trizol (Invitrogen), and quantified with a spectrophotometer and treated with DNAse RQ1 RNAse-free from Promega. complementary DNA (cDNA) was synthesized using Superscript III Reverse Transcriptase (Invitrogen) for 1 hour at 42°C. Controls without Reverse Transcriptase were performed in parallel. The products of the RT-PCR were analyzed in 1% agarose gel. Glutamate dehydrogenase (gGDH) was used as an internal standard.
A set of oligonucleotides were used to detect mRNA for the followed proteins: gGDH (Forward 5'AGCCTGCAGGCTATGGCGCTGTCTACTTCCTG and Reverse 5'TTGCGCCGAGCTGAATGAG); gNC1 (Forward 5'TCCAACAATCAGGCTCAGCAGACT and Reverse 5'ACTAGCCACATTGACTGTGTCGGA); gNC2 (Forward 5'AGACTCATATGGGCAAAAGGAG and Reverse 5'CGCAACCCCAAGTAATCTACC) and gPGE (Forward 5'ACCAGAGGCGGGTAAGGAATGTTT and Reverse 5'TGTGTAGGCCGGGTATTTAAGCCA).

Gene cloning

Because genes of *G. lamblia* lack of introns in theiry sequence, cDNA encoding gNC1 was obtained by PCR with Pfu DNA Polymerase (*Promega*), using total DNA from *WB* trophozoites as template. In order to avoiding problems of expression of heterologous proteins with transmembrane domains in bacteria, only the catalytic domain of gNC1 was amplified. These deleted constructs, called labeled gNC1-268 and gNC1-301, were amplified by using the primers: forward (gNC1_268F 5'GACCGGCTAGCCAAGAGATCGACCGTTCTCTGG, gNC1_301F 5'GACCGGCTAGCGATTTCAATGCACTAGAGCAGC) with an EcoRI restriction site and reverse (gNC1_R 5'TCGATGTCGACAGCTTAAAAAGGTATGTAGACCTGAGG) with a SalI restriction site. These fragments were inserted into the cloning vector pGEM-T-Easy (Promega) and subsequently transferred to the expression vector pET28b (Novagen), which adds the sequence of a 6 histidine epitope (6His) at the 5 ' end of the each insert, in order to be used for purification

purposes and/or immunological detection. These constructions, called pET28-gNC1-268 and pET28b-gNC1-301, were then used to transform the appropriated competent cells.

Expression and enrichment of gNC1cd

BL21(DE3)pLysS competent bacteria transformed with the vector pET28b alone and the constructions pET28-gNC1-268, pET28b-gNC1-301, were incubated overnight at 37ºC in 3 ml of LB/agar medium containing kanamycin (50 µg/ml) and chloramphenicol (40 µg/ml). Subsequently, 400 µl of each culture was transferred to 200 ml of fresh LB medium with the corresponding antibiotics and incubated at 37°^{O}C to an OD_{600} between 0.4 and 0.6. The expression of the desired proteins were induced with 0.5 mM IPTG for 3 hrs. at 37°C. Finally, the cells were harvested by centrifugation at 4000 rpm at 4ºC for 5 minutes in a SIGMA-12256 rotor.

Cells in the pellet were resuspended in 20 ml of ice-cold lysis buffer (20 mM Tris-HCl pH 8; 2 mM 2-mercaptoethanol and 10% v/v glycerol) and lysed by sonication with a SONIFIER-450 (Branson) equipment, by applying four pulses of 30 second (with 20 seconds rest) with 50% of duty cycle and output-control 5. The lysates were centrifuged at 9000 rpm at 4ºC for 1h in a rotor SIGMA-12176-H and the supernatant was applied to 1 ml bed Ni-NTA-agarose column (Qiagen). The column was washed with 10 column volumes of wash buffer (20 mM Tris-HCl pH 8; 200 mM NaCl; 2 mM 2-mercaptoethanol; 40 mM Imidazole and 10% v/v glycerol) and the protein of interest was eluted with 3 column volumes of elution eluate buffer (20 mM Tris-HCl pH 8; 200 mM NaCl; 2 mM 2-mercaptoethanol; 300 mM Imidazole and 10% v/v glycerol). Finally, all the enriched were aliquoted and stored a -20^{O}C.

The protein concentration of the eluted fractions was measured using Bradford's method.

Western Blot

Total protein from cell homogenates and enriched fractions were resolved by SDS/PAGE (12% w/v gel) under reducing conditions (20% v/v 2-mercaptoethanol) and then transferred electrophoretically transferred to nitrocellulose membranes for 1h atto 300 mA. For immunoblotting, non-specific binding sites on the nitrocellulose membrane were blocked with 3% w/v skim milk in 50 mM Tris-HCl pH 7.5; 150 mM NaC and 0.05% v/v Tween-20. Anty-poly-Histidine-Peroxidase antibody (Sigma) was used at a dilution of 1:4000. The specific bands were detected by chemiluminescence assay with SuperSignal West Pico kit (Thermo Scientific) and Kodak Biomax MS films. The molecular masses were calculated on the basis of the calibrated standards using Prestained Broad-Range (Genbiotech) run in every gel.

Adenylyl Cyclase assays

The reactions were performed by incubating 5 µl of the enriched enzymes gNC1 (gNC1-268 or gNC1-301) in a 50 µl reaction mix containing 25 mM Tris-HCl pH 8.0; 4 mM $MnCl_2$ (or other divalent cations) and 100 µM ATP (unless otherwise indicated) during 20 minutes at 32ºC. The reactions were stopped by addition of 1 ml of 100% v/v ethanol and the cAMP formed was quantified.

cAMP-dependent protein kinase (PKA) purification

PKA purification was adapted from the protocol previously described by Gilman (1970) [21]. Fresh bovine muscle (750g) was mechanically homogenized in an ice bath with 1 L of cold phosphate buffer (50mM, EDTA 4mM, pH 7.4). Homogenate was centrifuged at 16000g for 45 min at 4°C in a Sorvall centrifuge, supernant was recovered and precipitated with 30% $(NH_4)_2SO4$ for 20 min and centrifuged in the same conditions. The pellet was resuspended in 100 ml of 5 mM phosphate buffer (pH 7.4, 4 mM EDTA) and dialized in a pore membrane against the same buffer for 2 days at 4°C. The sample was collected and centrifuged in the same conditions as above. Supernatant was tritiated in equilibrium conditions for cAMP binding. Briefly, two-fold serial dilutions of the supernatant were incubated in duplicate for 2 h at 4°C with 2 nM [^{3}H]-cAMP (PerkinElmer, NET1161250UC, 20.7 Ci/mmol) in the absence or presence of 1 mM cAMP, the latter used for nonspecific binding correction. Bound fraction was isolated by dextran-carbon precipitation (15 min, 4°C) followed by centrifugation at 3000 rpm for 15 min at 4°C in an Eppendorf 5810R centrifuge. Supernatants were mixed with Optiphase HiSafe scintilliation cocktail (Perkin Elmer, cat 1200.437) and measured in a Wallac 1410 scintilliation counter. Maximal bound fraction (B0) versus log of protein dilution was fitted to a non-linear dose-response curve for optimal PKA dilution selection.

In all cAMP protein radiobinding experiments, 100 µl of a dilution of PKA in RBP buffer able to bind between 35 and 50% of 2 nM cAMP was used.

cAMP radiobinding-protein (RBP) assay

In all cases, ethanol from each reaction sample was evaporated in a water bath, and residues resuspended in an RBP buffer (50 mM Tris-HCl, 4 mM EDTA, pH 7.4, 0.1% BSA) for further cAMP determination. cAMP content was determined by a competitive radio-binding assay for PKA using [^{3}H]-cAMP as previously described by Davio et al. (1995) [22]. Briefly, tritiated PKA (100 µl) was incubated in equilibrium conditions (2 h, 4°C) with 50 µl of different samples of cAMP

standards (0.1 to 90 pmol) in the presence of 50 µl 8 nM [^{3}H]-cAMP (PerkinElmer, NET1161250UC) in an RBP buffer. Bound fraction was separated by carbon-dextran precipitation, followed by centrifugation (3000 rpm, 15 min, 4°C) and supernatants added to Optiphase HiSafe scintilliation cocktail (Perkin Elmer, cat 1200.437) and quantified in a Wallac 1410 liquid scintilliation counter. Sample cAMP concentrations were determined by interpolating the displacement curves obtained from cAMP standards. Duplicate samples in at least three independent experiments were analyzed.

Results

Regulation of cAMP during encystation

For the purpose of detecting the presence of intracellular cAMP in *G. lamblia* and determining if its concentration is affected during encystation, we use trophozoites of *G. lamblia* (WB strain) incubated *in vitro* with normal growth medium and induced to undergo encystation, first by incubation for 24 hrs in pre-encystation medium, and later by incubation with encystation medium during different times [20]. During this last stage, the encysting trophozoites were centrifuged, the pellets were resuspended in 1 ml of 100% (v/v) ethanol, and the total cAMP was measured (as described in the Experimental section). In this experiment, we found a detectable amount of intracellular cAMP, which increases significantly during the first hour of encystation (Fig. 1A), followed by a gradual reduction of the levels of cAMP until the 6th hour of encystation.

Additionally, when we performed the same experiment with longer times of encystation, we found that the levels of cAMP resulted significantly lower than the ones measured at the starting point (Fig 1B). This is consistent with the general decrease in metabolism of *G. lamblia* observed in the final stages of encystation [1]. The encystation of trophozoites was confirmed by checking the expression of the *G. lamblia* cyst wall protein 1 transcripts [23,24] by qRT-PCR, (Supplementary Figure S1).

Identification of adenylyl cyclases and phosphodiesterases in *G. lamblia*

In order to study the molecular mechanisms involved in the regulation of cAMP in *G. lamblia*, we conducted a search in the entire genome of *G. lamblia* for homologous sequences of adenylyl cyclases and phosphodiesterases by using known sequences of mammals, insects, nematodes and protozoans as queries.

As a result, we found three ORFs encoding one phosphodiesterase, accession XM_001706292.1 (automatically labeled *CAMP-specific 3',5'-cyclic phosphodiesterase 4B,*) and two nucleotidyl cyclases, accession XM_001704605.1 (automatically labeled *Adenylate cyclase*) and accession XM_001708713.1 (automatically labeled a *Hypothetical protein,*) [15]. For purposes of simplicity, during the course of this work, these proteins will be labeled: Giardia PDE (gPDE), Giardia nucleotidyl cyclase 1 (gNC1) and Giardia nucleotidyl cyclase 2 (gNC2) respectively (Fig. 2). Additionally, both the genome of the *WB* strain of *G. lamblia* as well as the genome of the strains *DH*, *GS* and *P15* are available at the GiardiaDB site (http://giardiadb.org/giardiadb/), both manually and automatically annotated.

An alignment of the gPDE, gNC1 and gNC2 genes between the WB strain and the DH, GS and P15 strains revealed that the sequences of the corresponding genes show an overall sequence identity ranging from 99% (DH) to 74% (GS), whereas within the catalytic domains, the sequence identities increase from 99% (DH) to 78% (GS) (Supplementary figures S2-S4).

The use of both secondary and tertiary structure prediction and multiple sequence alignment analysis of the translated products of these three ORFs shows the following results:

gPDE would correspond to a transmembrane protein of 1371 residues (154 kDa), with 9 or 10 predicted transmembrane domains (Fig. 2A), and one class I phosphodiesterase catalytic domain (CD) [25] at the C-terminal end (residues 1065 to 1343). The multiple sequence alignment and the tertiary structure analysis predict that this catalytic domain of gPDE has both sequence and structural homology with the mammalian PDE4B, where the amino acids that conform the active site are conserved (Fig. 2B) [26]. No additional or known regulatory elements were found with the prediction methods used.

With respect to the nucleotidyl cyclases, we found that gNC1 could be a transmembrane protein of 642 residues (74 kDa), with two transmembrane domains in the N-terminal end and a single class III CHD at the C-terminal end (residues 325 to 576), while gNC2 would correspond to a transmembrane protein of 2190 residues (246 kDa) with 12 transmembrane domains and two class III CHDs, Ca1, located between residues 288 and 465, and Ca2 located between residues 1810 and 2136 (Fig 3A). The analysis of the sequence of the catalytic domains of both gNC1 and gNC2 indicates that residues required for the formation of the active site are conserved. In both cases, the active site should be formed by the complementary interaction of two CDs [27], therefore, the activation of gNC1 would occur by the formation of homodimers, generating two identical catalytic sites as results of structural symmetry [28,29], while gNC2 should be activated -like all dimeric ACs- by the dimerization of the domains Ca1 and Ca2 of the same protein. In this last case, the sequence analysis of gNC2 indicates that this heterodimer should form only one active site, due to

differential mutations between both domains (Fig. 3B and 3C) [28]. The structure and activation mechanism of gNC2 should be similar to those two-CHD adenylyl cyclases present in mammalian, insect and nematodes [27], which could have been formed by an apparent duplication of a "monovalent" nucleotidyl cyclase domain, similar to gNC1 or AC_Rv1625c from *mycobacterium tuberculosis* [30]. Nevertheless, the multiple sequence alignment analysis shows that the formation of the active dimer in gNC2 is different from those present in higher eukaryotes, where the catalytic domains Ca1 and C2a look inverted with respect to the catalytic domains of the ACs of higher eukaryotes (Fig. 3B and 4).

Expression of gPDE, gNC1 and gNC2 in *G. lamblia*

We confirm the expression of gPDE, gNC1 and gNC2 mRNAs in our system by a RT-PCR assay using total RNA extracted from *G. lamblia* trophozoites incubated *in vitro* with normal growth medium. In all the cases, the specific oligonucleotides were designed to amplify regions between 100 and 150 bp of each mRNA (Fig. 5A, see the Experimental section for details). As shown in Fig. 5B, the trophozoites of *G. lamblia* express the mRNAs encoding the proteins under research. This result is complementary with the regulation of cAMP observed in the previous experiment (Fig. 1) and suggests that proteins gPDE, gNC1 and gNC2 are expressing in this parasite.

Enzymatic characterization of gNC1

The enzymatic characterization of gNC1 was performed through the following steps:

Expression and enrichment of gNC1cd

Because it was not possible to express the full-length gNC1 in bacteria, probably due to the presence of transmembrane domains encoded in its sequence, we only characterize the catalytic domain. To achieve this goal, we generated two deletion constructs by amplifying two sequences containing the catalytic domain of gNC1 without the transmembrane domains using PCR. These partial sequences called gNC1-268 and gNC1-301 (Fig. 6A) -where each number indicates the amount of codons removed from the 5' region of the gNC1 ORF- were inserted into the expression vector pET28b for its expression in bacteria. This vector also adds a 6His-tag at the N-terminal end of each protein to be used for immunology detection and protein purification.

Next, the constructs pET28-gNC1-268 and 301 were used to transform competent BL21 bacteria and the expression of gNC1-268 (44 kDa) and gNC1-301 (40 kDa) was induced by adding 0,5 mM IPTG to the culture and incubated 3 h at 37ºC. As a negative control, we performed the

same procedure using BL21 cells transformed with an empty pET28b vector (mock). After washing the cells with resuspension buffer, these were lysed by sonication and were centrifuged at 9000g for 1 h at 4°C in order to obtain the soluble fraction with the desired recombinant proteins. On the other hand, in order to increase the specific activity of each sample, the proteins of interest were later enriched by a Ni-agarose affinity chromatography -as described in the Experimental section . The protein expressions were demonstrated by western blot with an anti-6His-tag monoclonal antibody (Fig. 6B). The final eluate from the Ni- columns were then used for the experiments described below.

Adenylyl Cyclase activity of gNC1

In order to determine if the catalytic domain of gNC1 has the capacity to synthesize cAMP, we performed an *in vitro* AC assay by incubating 5 µl of the his-tagged gNC1-268 (0.7 µg of total proteins) and gNC1-301 (0.8 µg of total proteins) recombinants eluates,as well as the mock control (0,5 µg of total proteins) in a 50 µl reaction mix, as described in the Experimental section. Fig. 6B shows that gNC1-268 and gNC1-301 eluates have adenylyl cyclase activity, and that they are 7.3 and 34 times higher than controls, respectively. Additionally, as expected for an AC, both gNC1-268 and gNC1-301 showed dependence on Mn^{2+} as a divalent cation in order to be active. Despite the fact that both proteins are expressed in similar quantities (Fig. 6B), the highest adenylyl cyclase activity found in gNC1-301 may be explained because this construction expresses or folds better in bacteria due to its smaller size [32]. However, it can also be postulated that the extra N-terminal region of the gNC2-268 construct has an inhibitory effect on the enzyme by sterically hindering the formation of active dimers. For this reason, we selected gNC1-301 to be used in the following assays, in order to evaluate the catalytic domain of gNC1.

Effect of other divalent cations and optimal pH

In order to be functional, the active site of an AC that is formed by the dimerization of two catalytic domains requires the presence of two divalent cations in two well-established positions between the phosphates of the ATP and two Asp residues [27,28]. Although most of the ACs (including gNC1) respond very well to the presence of Mn^{2+} as a cofactor [28], AC enzymes can be activated in the presence of other cations such as Ca^{2+} and Mg^{2+}. For that purpose, we performed an AC assay of gNC1-301 eluates incubated in the absence of divalent cations (2 mM EDTA) or in the presence of 4 mM of $CaCl_2$, $MgCl_2$, $MnCl_2$ and their combinations. Fig. 7A shows that the AC activity of gNC1-301 is almost null in the absence of divalent cations. We found a significant AC

activity in the presence of individual divalent cations, with a maximum in presence of Mn^{2+}, followed by Ca^{2+} and a very little effect of Mg^{2+}. Additionally, when the divalent cations are combined, we found a synergistic effect in the AC activity when Ca^{2+} and Mn^{2+} were present in the reaction. This effect was not observed when gNC1-301 was incubated in presence of Ca^{2+}/Mg^{2+} and Mn^{2+}/Mg^{2+} combinations.

Finally, in order to determine the optimum pH activity of gNC1-301, we performed an AC assay of gNC1-301 eluates incubated in presence of 100 μM ATP, 4 mM Mn^{2+} in a 25 mM Tris-HCl buffer at different pH. The maximum activity under these conditions was found at pH 9.34 (Fig. 7B).

Inhibitory effect of GTP

As hypothesized above, the absence in gNC1 of a Lys-378 and an Asp-517 suggests that gNC1 lacks specificity for ATP, and it could interact with GTP and use it as a substrate. In order to test this, we evaluated if the nucleotide GTP is able to act as a competitive inhibitor in gNC1 in an AC reaction. For this, we performed three substrate vs. activity AC assays in the presence 0 μM (control), 100 μM and 300 μM of GTP. In each reaction, gNC1-301 eluates were incubated with increasing concentrations of ATP .

As shown in Fig. 8, we found that the best-fit curve of the control reaction shows a hill coefficient of 1.98 with a $K_{0.5}$ of 15.3 +/- 1.4 μM of ATP, and a V_{max} of 6021 +/- 618 pmol of cAMP.min^{-1}.mg of total $proteins^{-1}$. At the same time, the curves in the presence of 100 μM of GTP have a K_M apparent of 125 +/- 30 μM of ATP with a V_{max} of 15808 +/- 2212 pmol of cAMP.min^{-1}.mg of total $proteins^{-1}$, whereas in the presence of 300 μM of GTP, the K_M apparent goes up to 334 +/- 64 μM of ATP with a V_{max} of 14934 +/- 1877 pmol of cAMP.min^{-1}.mg of total $proteins^{-1}$.

These results indicate that the nucleotide GTP should act as a competitive inhibitor of the AC activity of gNC-301 by increasing the value of K_M of the reaction and not changing significantly the V_{max}. Additionally, the hill coefficient of 1.98 shown in the control reaction is consistent with the presence of two catalytic sites in gNC1, as described before.

Discussion

The synthesis and degradation of cAMP is important for the regulation of growth and differentiation processes of a great variety of organisms, from prokaryotes to eukaryotes [8, 28].

In the case of *G. lamblia*, Abel et al. (2001) [14] described the presence of a cAMP-dependent PKA (PKA) and detected significant levels of cAMP in trophozoites incubated *in vitro*.

Later, Gipson et al. (2006) [33] described that inducing encystment *in vitro* of trophozoites, correlates with an increase in the activity of the catalytic subunit of gPKA (gPKAC) and the degradation of its regulatory subunit (gPKAr).

In agreement with these results, we first found the presence of one Phosphodiesterase (gPDE) and two nucleotidyl cyclases (gNC1 and gNC2) genes in *Giardia* genome and also confirmed the presence of mARN of these enzymes in trophozoites of WB strain. . These results also agree with the expression of the gPDE, gNC1 and gNC2 mRNAs, originally shown by Einarsson et al. (2016) [34] in a complete transcriptional analysis of *G. lamblia trophozoites* of the WB strain, subject to encystment at different periods of time (Supplementary Figure S5). Then, we first measured the basal levels of cAMP in trophozoites of *G. lamblia* incubated *in vitro*, and later we detected a significant increase of the cAMP levels after the first hour of encystation, followed by a subsequent decrease of these levels to be below baseline at 24 h of encystation (Fig. 1). All these results describe the existence of mechanisms of synthesis and degradation of cAMP and its possible participation in the processes of adaptive differentiation of *G. lamblia*.

During the bioinformatic analysis, we found that gPDE has a tertiary structure similar to the PDEs of other protozoa such as *Cryptosporidium hominis*, *Plasmodium falciparum* and *Dictyostelium discoideum* (Fig.2C), presenting several transmembrane domains at the N-terminal ends and a class I PDE-catalytic-domain at the C-terminal end (Fig. 2A) [25], in whichthe amino acids involved in the catalysis and substrate recognition are highly conserved (Fig. 2B) [9,25,35].

In the case of the gNC1 and gNC2 nucleotidyl cyclases, these two are presented as enzymes with different characteristics when compared with each other. While both enzymes have a sequence structure of transmembrane proteins, gNC1 has a tertiary structure with a single catalytic domain, similar to protozoan nucleotidyl cyclases like *Plasmodium falciparum*, *Toxoplasma gondii* and *Trypanosoma cruz*as well as, prokaryotes like *Mycobacterium tuberculosis* and even higher eukaryotes guanylyl cyclases (Fig. 3C). On the other hand, gNC2 has a tertiary structure with two catalytic domains, similar to the typical structure of the adenylyl cyclases of mammals, insects, and nematodes [27,36], or to the two-CHD guanylyl cyclases found in *Plasmodium falciparum, Toxoplasma gondii* and *Paramecium tetraurelia*, but without their regulatory domains in the N-terminal ends (Fig. 3C) [31]. Because of this, gNC1 needs to form homodimers in order to be enzymatically active, yielding two active sites, while gNC2 could be activated by the formation of heterodimers between their catalytic domains C1a and C2a (Fig. 4 and 8B) [28].

In the latter case, we made two observations: a) The enzyme gNC2, like the ACs of higher eukaryotes, could only form one active site because of the loss of essential amino acids due to differential mutations of both catalytic domains during the protein evolution; and b) The location of

the two catalytic domains of gNC2 is apparently inverted with respect to the ones of the ACs of higher eukaryotes (Fig. 3D). This last observation could also be made in two-CHD guanylyl cyclases of *Toxoplasma gondii, Plasmodium falciparum* and *Paramecium tetraurelia* (Fig. 3B) [31]. However, differently from these, gNC2 shows the characteristic of a specific AC by presenting a $\beta2^{C1}$-$\beta3^{C1}$ hairpin Lys^{327} and a dimerization arm^{C1} Asp^{432} (Fig. 3A and 8B). Both residues are key for the recognition of ATP over GTP in the active site [28], so gNC2 could be labeled gAC in the future. This β2-β3 hairpin-Lys/dim-arm-Asp feature is lost in GCs, and so they show a significant AC activity as well [28].

One explanation for this inverted configuration of the catalytic domains of gNC2 with respect to two-CHD membrane ACs would be that both gNC2 and the two-CHD ACs of higher eukaryotes arise from two events of domain duplications of an ancestral monovalent nucleotydyl cyclase, similar to gNC1 or rv1625 from *Mycobacterium tuberculosis* [30,37]. Additionally, it is important to highlight that these differences found between gNC2 and others two-CHD ACs could be exploited for the generation of specific therapeutic agents in the future.

Finally, we did not find any detectable regulatory elements neither in the sequence of gPDE nor in gNC1 and gNC2, such as GAF domains [38,39]. Thus, the activities of these enzymes could be regulated either at the level of mRNA expression and/or post-translational modification. In agreement with the first, the early decrease in gPDE mRNA expression found in trophozoites during encystation (Supplementary Figure S5) coincides with the initial increase in cAMP at a similar period of time [34]. Nevertheless, the subsequent increase in the gNC1 mRNA expression is not consistent with the posterior reduction of cAMP observed in our research -considering that gNC2 mRNA expression remains unchanged. The latter would indicate that some kind of regulation of the activity, independent from the expression of proteins, may also be involved.

The last part of our work consisted of the purification and biochemical characterization of the catalytic domain of gNC1.

In summary, the different assays that we performed show that the catalytic domain of gNC1 (gNC1cd) enriched from bacteria is constitutively active for adenylyl cyclase activity with a K_{50} of 15 μM of ATP (Fig. 6 and 8). This activity is dependent on the presence of Ca^{2+} and Mn^{2+} (activity with Mg^{2+} is actually negligible) as cofactors (Fig. 7A) and has an optimum pH of 9.34 (Fig. 7B).

During the dose-response assay, we found that control curve fits better with a Hill coefficient of 1,98, which is consistent with the presence of the two active sites already predicted from the bioinformatic analysis (Fig. 4) [40]. Additionally, the effect of GTP as a competitive inhibitor shown in the same experiment suggests that gNC1cd can also bind GTP (possibly with similar affinity). This observation was initially predicted by the sequence analysis of gNC1, that

shows that the β2-β3 hairpin-Lys^{378} and the dimerization arm-Asp^{517} are replaced by a Threonine. Interestingly, the dimerization arm-Thr^{517} in gNC1 is topologically equivalent to dimerization arm-Asp in ACs, and so far, the structural model indicates that this Thr^{517} could interact with the adenine of ATP (Fig. 9C), as described for the specific AC of *Ar. platensis* [28]. This feature places gNC1 in an intermediate state between ACs and GCs, so gNC1 should be named gGAC for future references.

Despite the possibility that gNC1 synthesizes cGMP as well as cAMP, probably cGMP is not physiologically relevant as a second messenger in *G. lamblia*. So far, no Transduction signal pathways have been described for cGMP in *G. lamblia* and the *in silico* search for specific agonists of cGMP like cGMP-dependent protein kinase (PKG) did not get positive results [42].

Finally, the synergistic effect found with Ca^{2+} and Mn^{2+} over the AC activity of gNC1cd (Fig. 7A), can be explained by the fact that all the class III Nucleotydyl cyclases have two cation binding sites called site-A and site-B, that can be occupied by the same or different divalent cations. Thus, the results shown in Fig. 7A may indicate that each active site of the gNCcd-gNCcd homodimer is effectively occupied by two divalent cations (Fig. 9C-D) [27,28,43], and in the case of Ca^{2+} with Mn^{2+}, both cations may share an active site favoring a more efficient active conformation than they would have separately.

The adenylyl cyclase activity on gNCsc in the presence of Ca^{2+} alone results interesting, since Ca^{2+} -at the concentrations used for our assays- acts as an inhibitor of all membrane mammalian ACs [43-45], with the exception of mammalian soluble AC10, which can be also activated by Ca^{2+} alone, and it shows a similar synergistic effect in the presence of Ca^{2+}/Mn^{2+} and Ca^{2+}/Mg^{2+} [46,47].

There is experimental evidence over the direct and indirect effect of Ca^{2+} on the activity of mammalian ACs [48]. It has been described, for example, that Ca^{2+} has a direct inhibitory effect on mammalian AC5 and AC6 at submicromolar concentrations of Ca^{2+}, and that this effect could act as a regulatory mechanism in situations in which cAMP induces the opening of Ca^{2+} channels [48]. In these contexts, Reiner et al. [49] described the existence in *G. lamblia* of a thapsigargin-sensitive Ca^{2+} storage compartment. Thus, even if is not clear that Mn^{2+} has a physiological role in the activation of gNC1 in *G. lamblia,* since gNC1 is active in the sole presence of ATP-Ca^{2+}, it is possible that the AC activity of gNC1 be triggered by the release of Ca^{2+} into the cytosol after a certain stimulus.

Additionally, both *mammalian* AC10 and the single-CHD AC of *Spirulina platensis* belong to a group of ACs that, apart from Ca^{2+}, are stimulated by the presence of bicarbonate. So , it is possible that gNC1 behaves similarly to these enzymes. Since we did not find any detectable

regulatory domain in the sequence of gNC1, the hypothesis of AC regulation of gNC1 by Ca^{2+} and/or bicarbonate should be further evaluated. This can be achieved either by directly measuring the AC activity of gNC1 in the presence of different concentrations of Ca^{2+} and bicarbonate *in vitro*, or by indirectly applying an *in vivo* thapsigargin-treatment of trophozoite's culture.

On the other hand, we found that the optimal pH of gNC1cd is 9.34. This measure is higher than the cytosolic pH found in the gastrointestinal tract (6.3-7.0) [50] and also in the *G. lamblia in vitro* culture media. Additionally, Other previously characterized enzymes from *G. lamblia* such as Glutamate Dehydrogenase [51], Fructose-1,6-bisphosphate Aldolase [52] and Triose-phosphate isomerase [53] show optimal pHs of about 7.5. Since the cytosolic pH of *G. lamblia* was measured within 6.7 and 7.1 [50], it is possible that the optimal pH found in gNC1cd may correspond to an intrinsic property of the enzyme (or the catalytic domain alone) and not to the pH to which the enzyme needs to be exposed in order to work in the parasite.

As shown in the literature, and also in this work, cAMP would play an important role in the processes of adaptive differentiation and motility of *G. lamblia*. The present work describes, for the first time, the elements responsible for the synthesis and degradation of this important second messenger in this intestinal pathogen.

In this context, we hope to improve advance our knowledge in this topic in further studies and to beare able to establish the rational bases for future therapeutic treatments.

Acknowledgements

I want to thank to Dr. Maximiliano Ayub for his support and advise during the development of this project. I would also like to thank Dr. María Carolina Touz for the samples, and GAECI (Group of Advice in English Writing) at UNSL.

Author Contributions

Conceived and designed the experiments: VS AZ. Performed the experiments: VS ND BJ CD AZ. Analyzed the data: VS ND CD AZ. Contributed reagents/materials/analysis tools: CD AZ. Wrote the paper: VS ND BJ CD AZ.

References

1 Adam, R. D. R. (2001) Biology of Giardia lamblia. Clin. Microbiol. Rev., Am Soc Microbiol **14**, 447.

2 Feng, Y. and Xiao, L. (2011) Zoonotic potential and molecular epidemiology of Giardia species and giardiasis. Clin. Microbiol. Rev., American Society for Microbiology (ASM) **24**, 110–40.

3 Thompson, R. C. (2000) Giardiasis as a re-emerging infectious disease and its zoonotic potential. Int. J. Parasitol. **30**, 1259–67.

4 Gargantini, P. R., Serradell, M. del C., Ríos, D. N., Tenaglia, A. H. and Luján, H. D. (2016) Antigenic variation in the intestinal parasite Giardia lamblia. Curr. Opin. Microbiol. **32**, 52–58.

5 Lujan, H. D. (2011) Mechanisms of adaptation in the intestinal parasite Giardia lamblia. Essays Biochem **51**, 177–191.

6 Keister, D. B. (1983) Axenic culture of Giardia lamblia in TYI-S-33 medium supplemented with bile. Trans. R. Soc. Trop. Med. Hyg. **77**, 487–8.

7 Diamond, L. S., Harlow, D. R. and Cunnick, C. C. (1978) A new medium for the axenic cultivation of entamoeba histolytica and other entamoeba. Trans. R. Soc. Trop. Med. Hyg. **72**, 431–432.

8 Shemarova, I. V. (2009) cAMP-dependent signal pathways in unicellular eukaryotes. Crit. Rev. Microbiol. **35**, 23–42.

9 Omori, K. and Kotera, J. (2007) Overview of PDEs and their regulation. Circ. Res. **100**, 309–327.

10 Naviglio, S., Caraglia, M., Abbruzzese, A., Chiosi, E., Di Gesto, D., Marra, M., Romano, M., Sorrentino, A., Sorvillo, L., Spina, A., et al. (2009) Protein kinase A as a biological target in cancer therapy. Biochemistry, Informa UK Ltd London, UK 83–92.

11 Stork, P. (2002) Crosstalk between cAMP and MAP kinase signaling in the regulation of cell proliferation. Trends Cell Biol. **12**, 258–266.

12 Wang, Z., Dillon, T., Pokala, V. and Mishra, S. (2006) Rap1-mediated activation of extracellular signal-regulated kinases by cyclic AMP is dependent on the mode of Rap1 activation. Cell. Biol. **26**, 2130–2145.

13 Meng, T. C., Hetsko, M. L. and Gillin, F. D. (1996) Inhibition of Giardia lamblia excystation by antibodies against cyst walls and by wheat germ agglutinin. Infect. Immun. **64**, 2151–7.

14 Abel, E. S., Davids, B. J., Robles, L. D., Loflin, C. E., Gillin, F. D. and Chakrabarti, R. (2001) Possible Roles of Protein Kinase A in Cell Motility and Excystation of the Early Diverging Eukaryote Giardia lamblia. J. Biol. Chem. **276**, 10320–10329.

15 Morrison, H. G. H., McArthur, A. G. A., Gillin, F. F. D., Aley, S. S. B., Adam, R. D., Olsen, G. J. G. J., Best, A. A. A., Cande, W. Z. Z., Chen, F., Cipriano, M. J. M. J., et al. (2007)

Genomic minimalism in the early diverging intestinal parasite Giardia lamblia. Science (80-.)., American Association for the Advancement of Science **317**, 1921.

16 Aurrecoechea, C., Brestelli, J., Brunk, B. P., Carlton, J. M., Dommer, J., Fischer, S., Gajria, B., Gao, X., Gingle, A., Grant, G., et al. (2009) GiardiaDB and TrichDB: integrated genomic resources for the eukaryotic protist pathogens Giardia lamblia and Trichomonas vaginalis. Nucleic Acids Res. **37**, D526-30.

17 Katoh, K. and Standley, D. M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. **30**, 772–80.

18 Letunic, I., Doerks, T. and Bork, P. (2015) SMART: Recent updates, new developments and status in 2015. Nucleic Acids Res. **43**, D257–D260.

19 Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. and Sternberg, M. J. E. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. Nat. Protoc., Nature Research **10**, 845–858.

20 Boucher, S. (1990) Excystation of in vitro-derived Giardia lamblia cysts. Infect. Immun. **58**, 3516–3522.

21 Gilman, A. G. (1970) A protein binding assay for adenosine 3':5'-cyclic monophosphate. Proc. Natl. Acad. Sci. U. S. A. **67**, 305–312.

22 Davio, C. A., Cricco, G. P., Bergoc, R. M. and Rivera, E. S. (1995) H1 and H2 histamine receptors in n-nitroso-N-methylurea (NMU)-induced carcinomas with atypical coupling to signal transducers. Biochem. Pharmacol. **50**, 91–96.

23 Mowatt, M. R., Luján, H. D., Cotten, D. B., Bowers, B., Yee, J., Nash, T. E. and Stibbs, H. H. (1995) Developmentally regulated expression of a Giardia lamblia cyst wall protein gene. Mol. Microbiol. **15**, 955–63.

24 Merino, M. C., Zamponi, N., Vranych, C. V., Touz, M. C., Rópolo, A. S., Adam, R., Birkeland, S., Preheim, S., Davids, B., Cipriano, M., et al. (2014) Identification of Giardia lamblia DHHC Proteins and the Role of Protein S-palmitoylation in the Encystation Process. PLoS Negl. Trop. Dis. (Singer, S. M., ed.), Public Library of Science **8**, e2997.

25 Richter, W. (2002) 3',5'-Cyclic nucleotide phosphodiesterases class III: Members, structure, and catalytic mechanism. Proteins Struct. Funct. Genet. **46**, 278–286.

26 Hengming Ke, Huanchen Wang, and M. Y. (2011) Structural Insight into the Substrate Specificity of Phosphodiesterases. Handb Exp Pharmacol **204**, 121–134.

27 Linder, J. U. (2006) Class III adenylyl cyclases: Molecular mechanisms of catalysis and regulation. Cell. Mol. Life Sci. **63**, 1736–1751.

28 Sinha SC, S. S. (2006) Structures, mechanism, regulation and evolution of class III nucleotidyl cyclases. Rev Physiol Biochem Pharmacol. **159**, 105–140.

29 Steegborn, C., Litvin, T. N., Levin, L. R., Buck, J. and Wu, H. (2005) Bicarbonate activation

of adenylyl cyclase via promotion of catalytic active site closure and metal recruitment. Nat. Struct. Mol. Biol. **12**, 32–7.

30 Ying Lan Guo, Seebacher, T., Kurz, U., Linder, J. U. and Schultz, J. E. (2001) Adenylyl cyclase Rv1625c of Mycobacterium tuberculosis: A progenitor of mammalian adenylyl cyclases. EMBO J. **20**, 3667–3675.

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

32 Finkelstein, A. V., Bogatyreva, N. S. and Garbuzynskiy, S. O. (2013) Restrictions to protein folding determined by the protein size. FEBS Lett., Federation of European Biochemical Societies **587**, 1884–1890.

33 Gibson, C., Schanen, B., Chakrabarti, D. and Chakrabarti, R. (2006) Functional characterisation of the regulatory subunit of cyclic AMP-dependent protein kinase A homologue of Giardia lamblia: Differential expression of the regulatory and catalytic subunits during encystation. Int. J. Parasitol. **36**, 791–799.

34 Einarsson, E., Troell, K., Hoeppner, M. P., Grabherr, M., Ribacke, U. and Sv??rd, S. G. (2016) Coordinated Changes in Gene Expression Throughout Encystation of Giardia intestinalis. PLoS Negl. Trop. Dis. **10**, 1–22.

35 Salter, E. A. and Wierzbicki, A. (2007) The mechanism of cyclic nucleotide hydrolysis in the phosphodiesterase catalytic site. J. Phys. Chem. B **111**, 4547–4552.

36 Kamenetsky, M., Middelhaufe, S., Bank, E. M., Levin, L. R., Buck, J. and Steegborn, C. (2006) Molecular Details of cAMP Generation in Mammalian Cells: A Tale of Two Systems. J. Mol. Biol. **362**, 623–639.

37 Sinha, S. C., Wetterer, M., Sprang, S. R., Schultz, J. E. and Linder, J. U. (2005) Origin of asymmetry in adenylyl cyclases: structures of Mycobacterium tuberculosis Rv1900c. EMBO J. **24**, 663–673.

38 Martinez, S. E., Wu, A. Y., Glavas, N. a, Tang, X.-B., Turley, S., Hol, W. G. J. and Beavo, J. a. (2002) The two GAF domains in phosphodiesterase 2A have distinct roles in dimerization and in cGMP binding. Proc. Natl. Acad. Sci. U. S. A. **99**, 13260–13265.

39 Pandit, J., Forman, M. D., Fennell, K. F., Dillman, K. S. and Menniti, F. S. (2009) Mechanism for the allosteric regulation of phosphodiesterase 2A deduced from the X-ray structure of a near full-length construct. Proc. Natl. Acad. Sci. U. S. A. **106**, 18225–30.

40 Winger, J. A., Derbyshire, E. R., Lamers, M. H., Marletta, M. A. and Kuriyan, J. (2008) The crystal structure of the catalytic domain of a eukaryotic guanylate cyclase. BMC Struct. Biol. **8**, 42.

41 Tesmer, J., Sunahara, R., Johnson, R., Gosselin, G., Gilman, A. and Sprang, S. (1999) Two-metal-Ion catalysis in adenylyl cyclase. Science (80-.). **285**, 756–60.

42 Manning, G., Reiner, D. S., Lauwaet, T., Dacre, M., Smith, A., Zhai, Y., Svard, S. and Gillin, F. D. (2011) The minimal kinome of Giardia lamblia illuminates early kinase evolution and unique parasite biology. Genome Biol. **12**, R66.

43 Mou, T. hung, Masada, N., Cooper, D. M. F. and Sprang, S. R. (2009) Structural Basis For Inhibition Of Mammalian Adenylyl Cyclase By Calcium. Biochemistry. **48**, 3387–3397.

44 Guillou, J. L., Nakata, H. and Cooper, D. M. F. (1999) Inhibition by calcium of mammalian adenylyl cyclases. J. Biol. Chem. **274**, 35539–35545.

45 Hu, B., Nakata, H., Gu, C., De Beer, T. and Cooper, D. M. F. (2002) A critical interplay between Ca2+ inhibition and activation by Mg2+ of AC5 revealed by mutants and chimeric constructs. J. Biol. Chem. **277**, 33139–33147.

46 Litvin, T. N., Kamenetsky, M., Zarifyan, A., Buck, J. and Levin, L. R. (2003) Kinetic properties of "soluble" adenylyl cyclase: Synergism between calcium and bicarbonate. J. Biol. Chem. **278**, 15922–15926.

47 RV, H. and DL, G. (1979) Regulation of Guinea Adenylate Cyclase by Calcium ' The effect of various metal ions on guinea pig sperm adenylate cyclase detergent spermatozoa some appear be studies erythrocyte Dowex-50 Dowex-50. Biol. Reprod. **21**, 1135–1142.

48 Halls, M. L. and Cooper, D. M. F. (2011) Regulation by Ca2+-signaling pathways of adenylyl cyclases. Cold Spring Harb. Perspect. Biol. **3**, 1–22.

49 Reiner, D. S., Hetsko, M. L., Meszaros, J. G., Sun, C.-H. H., Morrison, H. G., Brunton, L. L., Gillin, F. D., Reinert, D. S., Hetskot, M. L., Meszaros, J. G., et al. (2003) Calcium signaling in excystation of the early diverging eukaryote, Giardia lamblia. J. Biol. Chem. **278**, 2533–2540.

50 Biagini, G. A., Knodler, L. A., Saliba, K. J., Kirk, K. and Edwards, M. R. (2001) Na+-dependent pH Regulation by the Amitochondriate Protozoan Parasite Giardia intestinalis. J. Biol. Chem. **276**, 29157–29162.

51 Park, J., Schofield, P. J. and Edwards, M. R. (1998) Giardia intestinalis: Characterization of a NADP-Dependent Glutamate Dehydrogenase **138**, 131–138.

52 Galkin, A., Kulakova, L., Melamud, E., Li, L., Wu, C., Mariano, P., Dunaway-Mariano, D., Nash, T. E. and Herzberg, O. (2007) Characterization, kinetics, and crystal structures of fructose-1,6- bisphosphate aldolase from the human parasite, Giardia lamblia. J. Biol. Chem. **282**, 4859–4867.

53 López-Velázquez, G., Molina-Ortiz, D., Cabrera, N., Hernández-Alcántara, G., Peon-Peralta, J., Yépez-Mulia, L., Perez-Montfort, R. and Reyes-Vivas, H. (2004) An unusual triosephosphate isomerase from the early divergent eukaryote Giardia lamblia. Proteins Struct. Funct. Genet. **55**, 824–834.

Figure legends

Fig. 1: **Total cAMP in *G. lamblia***

Determination of the total amount of cAMP measured in trophozoites incubated with encystation medium during the indicated times. The asterisks indicate that there was significant difference compared with T = 0 hrs. (Student's t test: * = p <0.05; ** = p< 0.01).

* Significant difference compared with T = 0 hrs. (Student's t test: * = p <0.05; ** = p< 0.01).

Fig. 2: **Bioinformatic analysis of gPDE**

A) Schematic representation of the domains structure of gPDE. Filled rectangles = transmembrane domains, HDc = phosphohydrolase domain. The remaining domains are labeled according to the Pfan code. The small rectangles below indicate the regions that are shown in the multiple sequence alignment; B) Multiple sequence alignment of the catalytic domains of gPDE compared with other species. The rectangles with green characters indicate the amino acids directly involved in the catalytic activity, while the rectangles with blue characters indicate the amino acids involved in the substrate recognition; C) Schematic representation of the domain structures of the PDEs used in the multiple sequence alignment; D) Model of gPDE interacting with cellular membrane and cAMP as its substrate, two divalent cations are needed for catalysis.

Fig. 3: **Bioinformatic analysis of gNC1 and gNC2**

A) Schematic representation of the domain structures of gNC1 and gNC2. Filled rectangles = transmembrane domains, CHDcys = cyclase homology domain. The remaining domains are labeled according to the Pfan code. The small rectangles below indicate the regions that are shown in the multiple sequence alignment; B) Multiple sequence alignment of the catalytic domains of gNC1 and gNC2 compared with other species. The open rectangles indicate the residues of the active catalytic site. The residues involved in the catalytic activity are shown in green, while the residues involved in the specific ATP base recognition are shown in blue. Amino acids in red indicate mutations with loss of catalytic activity or ATP base recognition; C) Schematic representation of the domain structures of the NCs used in the multiple sequence alignment.

Fig. 4: **Schematic representation of the active sites of ACs formed after dimerization**

The open circle (O) indicates a functional active site, while the X indicates a disabled active site. The residuesinvolved in the catalytic activity and the base recognition are also shown. *Mutation that disables the catalytic site. gNC1 shows two active sites while the two-CHD ACs show only one. Additionally, gNC2 and GCCD from *T. gondii* displays/shows an inverted active site when compared with the two-CHD ACs of higher eukaryotes. This "inversion" on the catalytic domains of the active site location found in gNC2 would be a characteristic of some protozoa two-CHD nucleotidyl cyclases. This property has been also described in two-CHD guanylyl cyclase *Plasmodium falciparum* and *Paramecium tetraurelia* [31] and also in a guanylyl/adenylyl cyclase from *Toxoplasma gondii* (accession XM_002369450.1). However, the analysis of the composition of the residues in the active site of gNC2 shows that it has the residues K-327 and D-432, which are described as essential for the specific ATP-base recognition in specific ACs [28]. This feature is shared only with other ACs in the multiple sequence alignment. As expected, these residues are absent in the enzymes that are already described as guanylyl cyclases (Fig. 3). This observation includes gNC1. Thus, this enzyme would not be a specific AC, but an adenylyl/guanylyl cyclase.

Fig. 5: **Expression of gPDE, gNC1 and gNC2 mRNA in *G. lamblia***

A) Schematic representation of the mRNA analyzed in the RT-PCR assay. The regions amplified are indicated between the opposite arrows; B) Two RT-PCR assays showing the expression of gGDH, gPDE, gNC1 (left) and gNC2 (right) transcripts from *G. lamblia* trophozoites. As a positive control, we used specific primers for *Giardia Glutathione Dehydrogenase* (gGDH). The controls gGDH (RT-) and gNC2 (RT-) indicate no contamination with nuclear DNA.

Fig. 6: **Heterologous expression of the catalytic domain of gNC1**

A) Schematic representation of gNC1 and its deleted constructions gNC1-268 and gNC1-301 expressed in BL21 cells, followed by their predicted size. It was not possible to express the full gNC1 in bacteria. The arrows indicate the residues eliminated from the full-length gNC1. 6H= 6His-tag, TM = transmembrane domain, CHD = cyclase homology domain; B) Adenylyl cyclase activity measured in the indicated eluate containing the recombinant proteins. The eluates were resolved by electrophoresis at 12% SDS–PAGE and the proteins were revealed by immunoblot with monoclonal anti-His antibody (inset).

Fig. 7: **AC activity assay of the catalytic domain of gNC1**

A) AC activity of gNC1-301 incubated in presence of 2 mM EDTA and 4 mM of the indicated divalent cations (P = Student’s t test result). The maximum activity was achieved with the

combination of Ca^{2+} and Mn^{2+}. B) Determination of the optimus pH of the AC activity of gNC1-301.

Fig. 8: **Substrate concentration vs. AC activity of gNC1. Inhibitory effect of GTP**

Three AC assays were performed by increasing the concentration of ATP in presence of 0 (filled squares), 100 (filled spheres) and 300 μM (filled triangles) of GTP. In all the reactions, 5 μl of gNC1-301 were incubated with the indicated concentration of ATP in a reaction mix containing 4 mM MnCl2 in 25 mM Tris-ClH buffer pH 8. The value of K_{50}, K_M^{app} and V_{max} are indicated in the rectangle below. The control curve shows a sigmoid form with a hill coefficient of 1,98. The best-fit curve of each experiment was obtained by the IC50 toolkit (www.ic50.tk) as a graphical interface for the GNUPLOT program (//gnuplot.info/).

Fig. 9: **Structural models of gNC1 and gNC2 catalytic domains**

A) Structural model of a heterodimer of gNC2 catalytic domains (left) based on the crystal structure of the quimeric mammal enzyme adenylyl cyclase (AG) from *Canis lupus* (Ca1) y *Rattus norvegicus* (C2a) (right, PDB reference 1CJK) [41]. The structural model was developed by the online tool Phyre2 (with a coincidence of 100%) and aligned structurally by using the program Pymol. The molecule of Adenosine-5'-Rp-Alpha-Thio-Triphosphate (TAT) and the atoms of Mg^{2+} (green) and Mn^{2+} (violet) shown in the structural model of gNC2 belong to the 1CJK crystal structure; B) Prediction of the catalytic site formed by the dimerization of the structural model of gNC2. The conserved catalytic amino acids of both monomers are shown. The polar interactions were predicted with the Pymol program; C) Structural model of a homodimer of gNC1 catalytic domain (left) based on the crystal structure (1WC6) of the enzyme adenylyl cyclase from *Spirulina platensis* (right, PDB reference 1WC6). The structural model was developed by the on line tool Phyre2 (with a coincidence of 100%) and two identical copies were aligned structurally against the catalytic domains of Spirulina's AC by using the program Pymol. The two monomers in each Figure were colored differently for clarity purposes, despite the fact that both have the same sequence of amino acids. The two molecules of adenosine-5'-rp-alpha-thio-triphosphate (TAT) and the 4 atoms of Mg^{2+} (green) shown in the structural model of gNC1 belong to the 1CW6 crystal structure; D) Prediction of the catalytic site formed by the dimerization of the structural model of gNC1 (left) and a flat graphical representation (right). The conserved catalytic amino acids of both monomers are shown. The polar interactions were predicted with the Pymol program. Besides the catalytic amino acid, there are two divalent cations: one in position "A" that can be seen closest to the nitrogenous base, while "B" is found more proximal to the phosphates β and γ of the ATP.

Supplementary Figures legends

Fig. S1: **The expression of cyst wall protein 1 transcripts during encystation**
qRT-PCR analysis of cwp1 transcript expression after the indicated times of encystation, relative to the expression of gGDH transcripts used as internal standard.

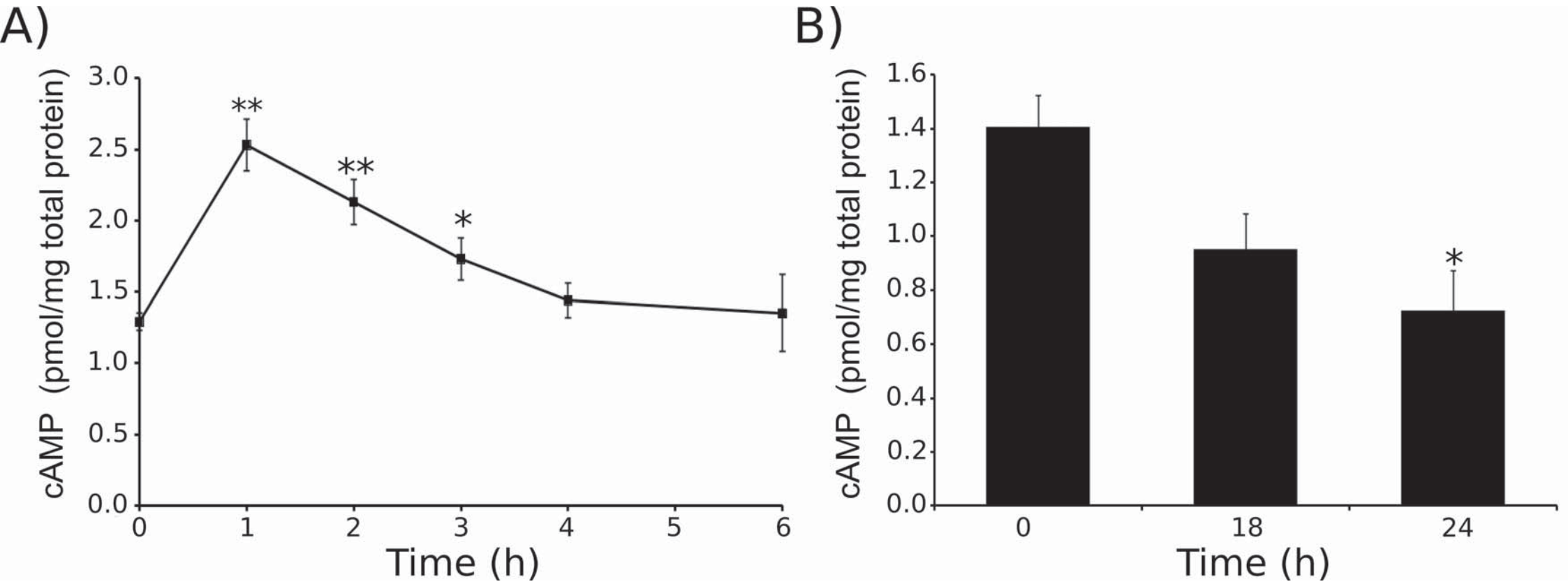
Fig. S2:
A) Multiple sequence alignment of the full-length gPDE of the *WB* strain compared with *DH, GS* and *P15* strains of *G. lamblia*. TM = transmembrane domains, CD = catalytic domain; B) Multiple sequence alignment of the catalytic domain of gPDE of analyzed strains. The residues that should involved in the catalytic activity are shown in green, while the residues that should be involved in the substrate recognition are shown in blue. A summary of the multiple sequence alignments results, of both, the full-length protein and the catalytic domain is shown below; C) Schematic representation of the domain structures of the gPDE sequences analyzed.

Fig. S3:
A) Multiple sequence alignment of the full-length gNC1 of the *WB* strain compared with *DH*, *GS* and *P15* strains of *G. lamblia*. TM = transmembrane domain, CD = catalytic domain; B) Multiple sequence alignment of the catalytic domain of gNC1 of analyzed strains. The residues that should be involved in the catalytic activity are shown in green, while the residues that should be involved in the substrate recognition are shown in blue. The mutation T378A observed in the *GS* strain is indicated in red. A summary of the multiple sequence alignments results, of both, the full-length protein and the catalytic domain is shown below; C) Schematic representation of the domain structures of the gNC1 sequences analyzed.

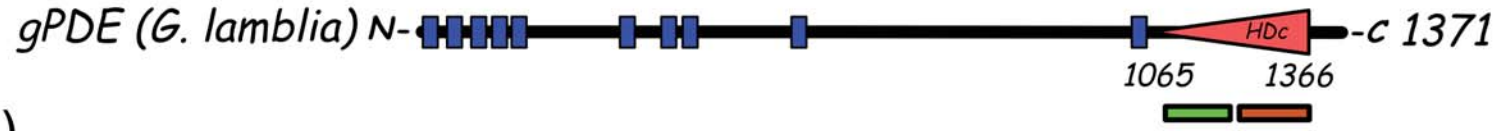
Fig. S4:
A) Multiple sequence alignment of the full-length gNC2 of the *WB* strain compared with *DH*, *GS* and *P15* strains of *G. lamblia*. TM = transmembrane domains, CD = catalytic domain; B) Multiple sequence alignment of the catalytic domain of gNC2 of analyzed strains. The rectangles indicate the residues that belong to the functional active site. The residues that should be involved in the catalytic activity are shown in green, while the residues that should be involved in the substrate recognition are shown in blue. The mutations responsible for the loss of active site function are shown in red. The mutation D292S in the GS strain should not has and additional effect because it is present in the nonfunctional active site. A summary of the result of multiple sequence alignments

of both full-length and catalytic domain is shown below; C) Schematic representation of the domain structures of the gNC2 sequences analyzed.

Fig. S5:
Expression levels of gPDE, gNC1 and gNC2 presented as FPKM at the indicated times. Data were obtained from the Supporting Information S1 Table (https://doi.org/10.1371/journal.pntd.0004571.s001) [34] ;


A)
cAMP (pmol/mg total protein)
3.0
2.5
2.0
1.5
1.0
0.5
0.0
**
**
*
0
1
2
3
4
5
6
Time (h)
B)
cAMP (pmol/mg total protein)
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
*
0
18
24
Time (h)

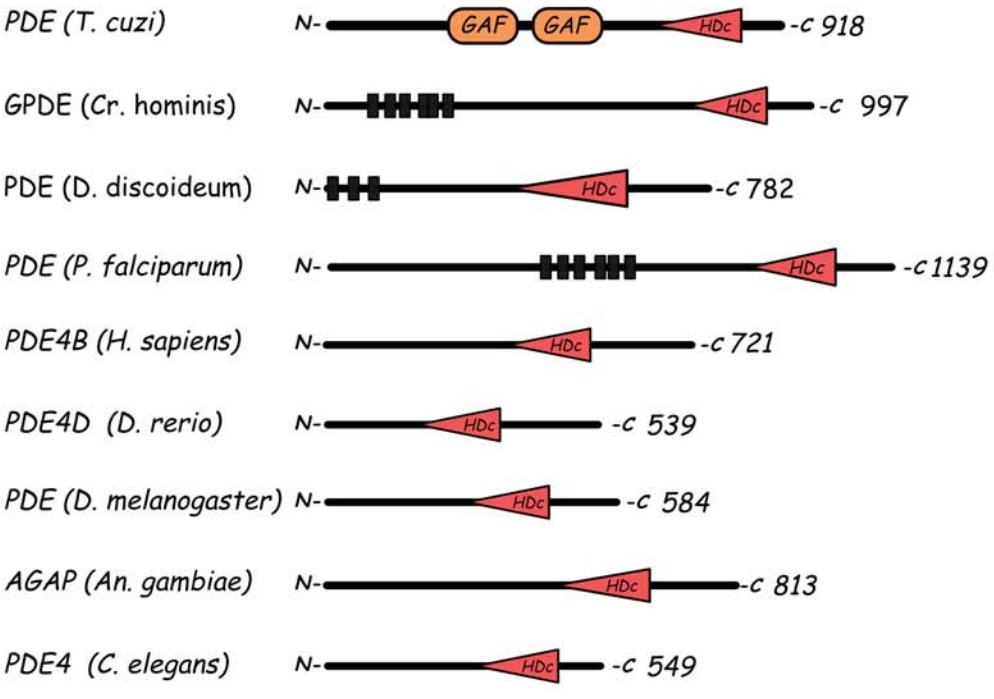
A)



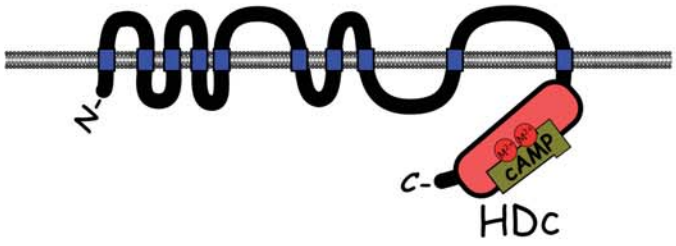
B)

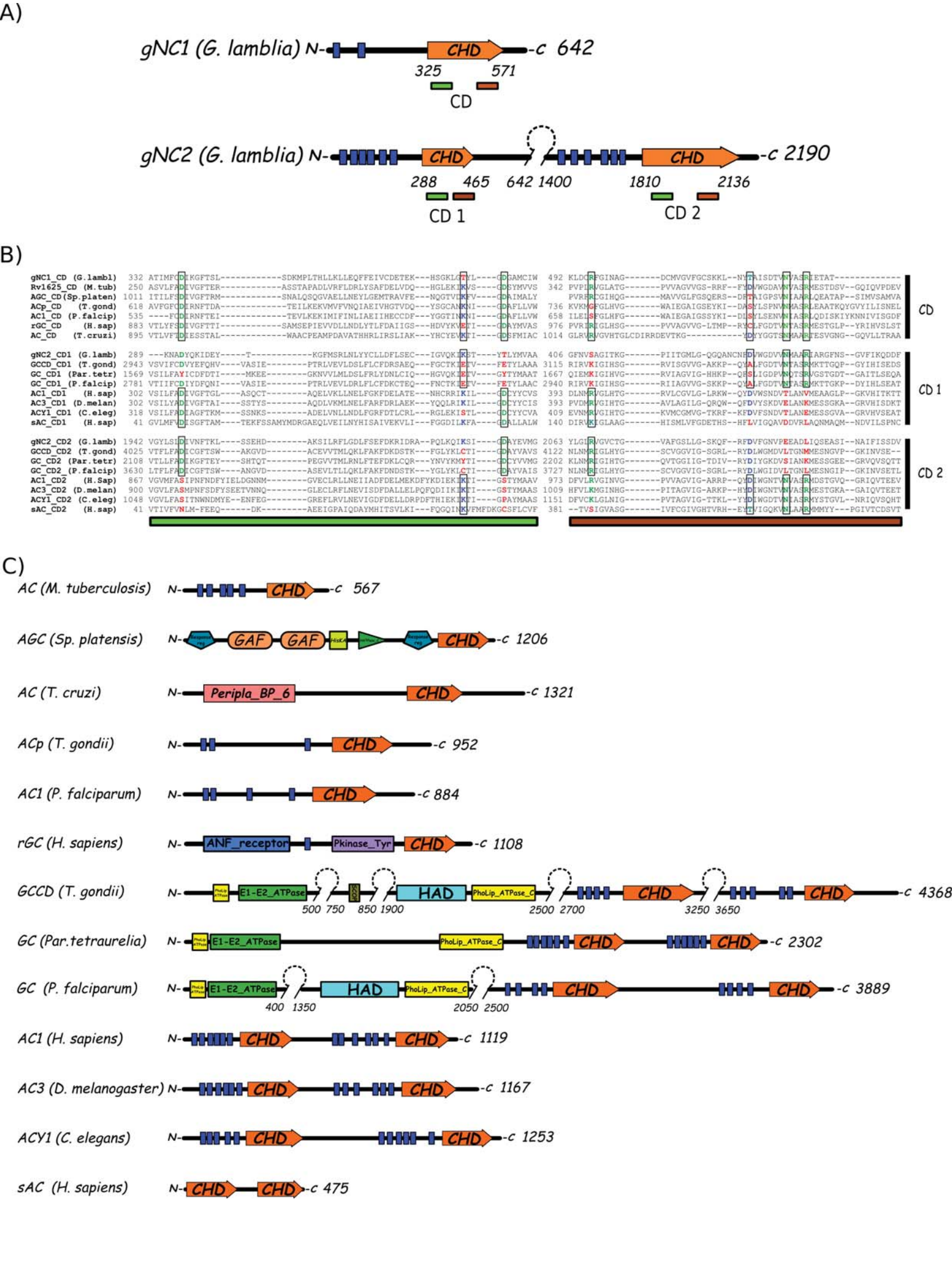
```
gPDE (G. lamblia)    1067 YHNKLHAADVAQMSMYMLSTVYCSLISESPKHPFLCVYKAMRQYKESDYSRLITEQPRQA  LIRPVDFLALLFGSLCHDLGHTGIDNLFCINTENALALLYN---DEAPLEHAHATLSWHI 1183
PDE  (T. cruzi)       667 YHNFYHVVDVCQTIYTFLY----------------------------------KGAVCE  KLTELDCFVLLVTALVHDLLHMGLNNSFYLKTDSPLGILSSASGNASVLEVHHCNLAVEI  762
GPDE (Cr. hominis)    724 YHNELHGTNVCHLAICLSR----------------------------------ATGLWS  HLDTVERLASVIASLGHDVGHIGRTSNFLVNSRHMLAINYN---DRSVLEMFHASLTFRI  805
PDE  (Di. discoid.)   340 YHNSIHATDVLHNLNYFI----------------------------------EKSFGK  FLTDIELFSMILAAIIHDFKHPGVNNHFQINSKSRLALKYN---DKSILENYHLHQAFII  420
NPDE (P. falciparum)  846 YHNTIHATMVTQKFFCLAK----------------------------------KLGIYD  DLEYKIKLVMFISGICHDIGHPGYNNLFFVNSLHPLSIIYN---DISVLENYHASITFKI  927
PDE4B(H. sapiens)      72 YHNSLHAADVAQSTHVLLS----------------------------------TPALDA  VFTDLEILAAIFAAAIHDVDHPGVSNQFLINTNSELALMYN---DESVLENHHLAVGFKL  153
PD4D (D. rerio)       170 YHNSIHAADVVQSTHVLLS----------------------------------TPALEA  VFTDLEILAAMFASAIHDVDHPGVSNQFLINTNSELALMYN---DSSVLENHHLAVGFKL  251
PDE(D. melanogaster)  246 HHNSLHAADVTQSTNVLLN----------------------------------TPALEG  VFTPLEVGGALFAACIHDVDHPGLTNQFLVNSSSELALMYN---DESVLENHHLAVAFKL  327
AGAP (A. gambiae)     455 HHNSLHAADVTQSTHALLH----------------------------------SPALDG  VFTPLEVCAALVAACIHDVDHPGLTNQFLINSSSELALMYN---DESVLENHHLAVAFKL  536
PDE4 (C. elegans)     281 YHNFIHAADVAQSMHVLLM----------------------------------SPVLTE  VFTDLEVLAAIFAGAVHDVDHPGFTNQYLINSNNELAIMYN---DESVLEQHHLAVAFKL  362

gPDE (G. lamblia)    1257 IKFGDLSNPCRPIEISTRYAVALMNEFWSLGDLMLECGLEPDKIKTRPQKGEESLIIANS  QIGFTQSIVKGFWTVVERFWKALAGVEFSD-----LQANLNATVEHWQNVRSEIE----- 1366
PDE  (T. cruzi)       819 LKAGDVSNVTKPFDISRLWAMAVTEEFYRQGDMEKEKGVEVLPMFDRS----KNTELAKG  QIGFIDFVAGPFFKK----IVSALLTGMQW-----TVERVDSNRAEWQRTLDAK------ 918
GPDE (Cr. hominis)    878 LKAADIGHGALKWNQHYKWCRSVVEEFFLQGDEEKALSLPISPICDR-----ESTDVPKS  QVGFLNFVCLPLFQE----LC---YVDVEGDVRR-CIDRILENINNWEDIAEAGIQWDSE  983
PDE  (Di. discoid.)   589 IKCADISNPSKPWNLYTNWSNRVTSEFYKQGDKEKESNMDVSAFMDR-----NKPATTKC  QINFINIFVAPIYEI----WSHH-FPQFKL-----CYQNILSNLSRLEIEQQQQLQLQQQ  693
NPDE (P. falciparum)  996 IKSADISHGSVSWSEHYCWCQRVLSEFYTQGDEELKNKMPLSPLCDRT----KHNEVCKS  QITFLKFVVMPLFEE----LS---HIDNNKFIKSFCLKRLNSNCIMWDTLMKEEKTIEVY 1104
PDE4B(H. sapiens)     227 VHCADLSNPTKSLELYRQWTDRIMEEFFQQGDKERERGMEISPMCDK-----HTASVEKS  QVGFIDYIVHPLWET----WADLVQPDAQD-----ILDTLEDNRNWYQSMI--------- 323
PD4D (D. rerio)       325 VHCADLSNPTKPLELYRQWTDRIMVEFFTQGDRERDKGMEISPMCDK-----HNASIEKS  QVGFIDYIVHPLWET----WADLVHPDAQE-----ILDTLEDNREWYQSMIPHSPTSTPE 430
PDE(D. melanogaster)  401 VHCADLSNPTKPLPLYKRWVALLMEEFFLQGDKERESGMDISPMCDR-----HNATIEKS  QVGFIDYIVHPLWET----WASLVHPDAQD-----ILDTLEENRDYYQSMIPPSPPPSGV 506
AGAP (A. gambiae)     625 VHCADLSNPTKPLPLYRRWVRRLMEEFFRQGDRERAAGMDVSPMCDR-----RSATVESS  QVGFIDYIVHPLWEA----WAELVRPDAQD-----ILDRLEENRAYYQALIPPPSPPSHE 730
PDE4 (C. elegans)     436 IHLADLSNPTKPIELYQQWNQRIMEEYWRQGDKEKELGLEISPMCDR-----GNVTIEKS  QVGFIDYIVHPLYET----WADLVYPDAQN-----ILDQLEENREWYQSRIPEEPDTART 541
```

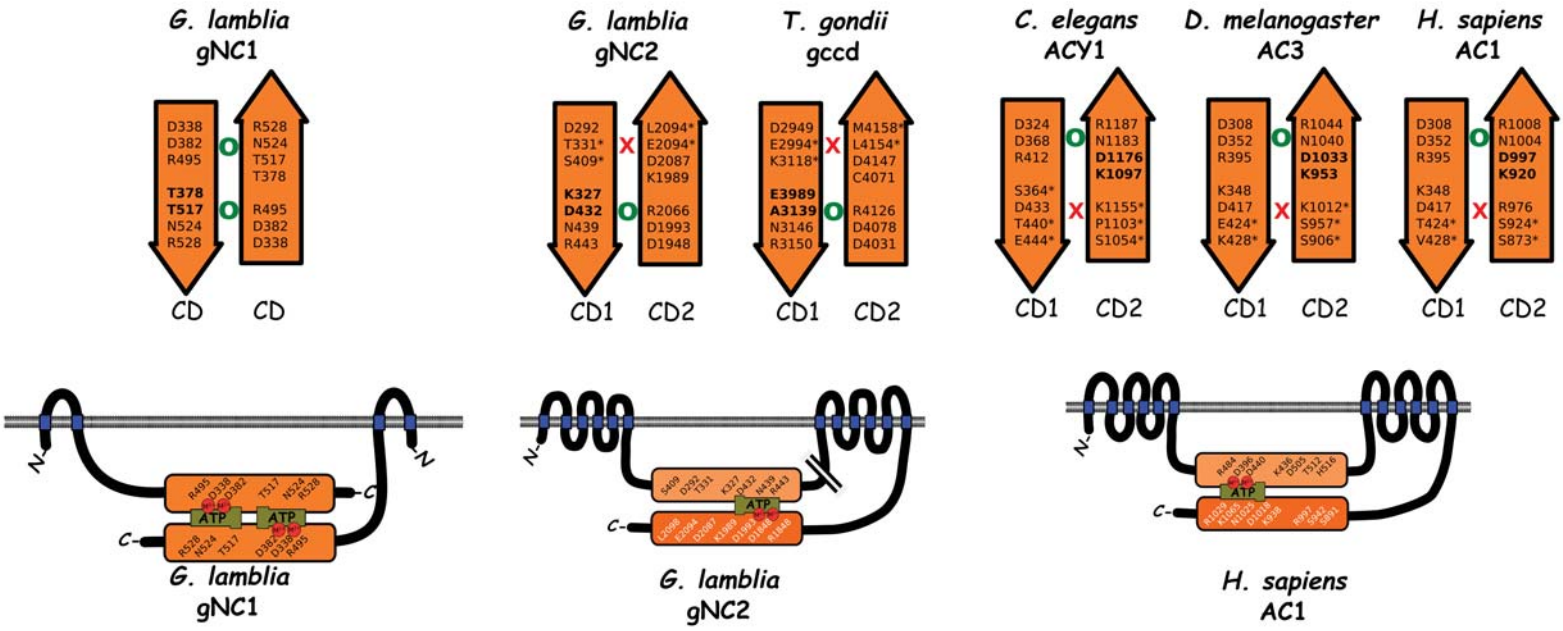
C)

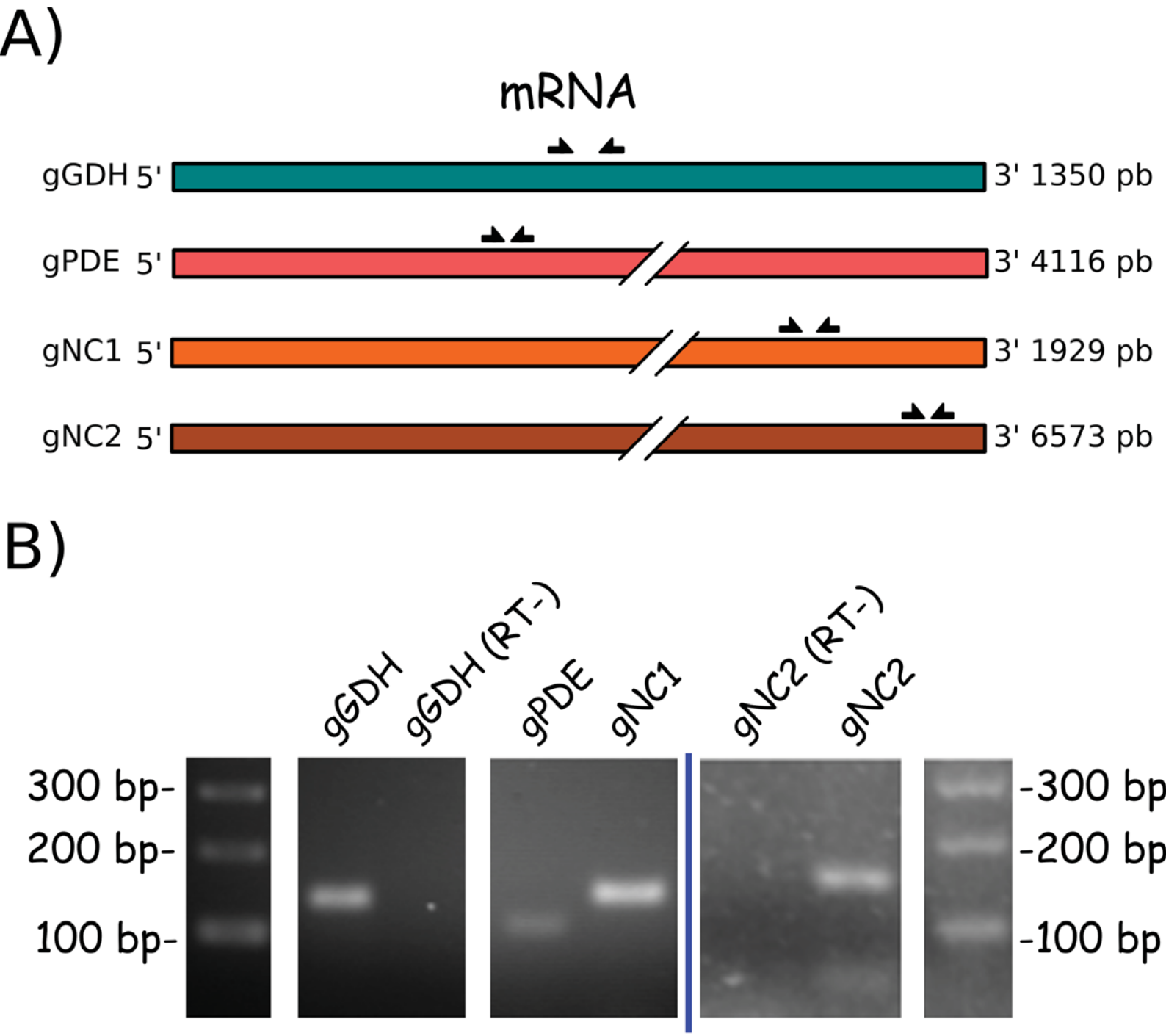


D)

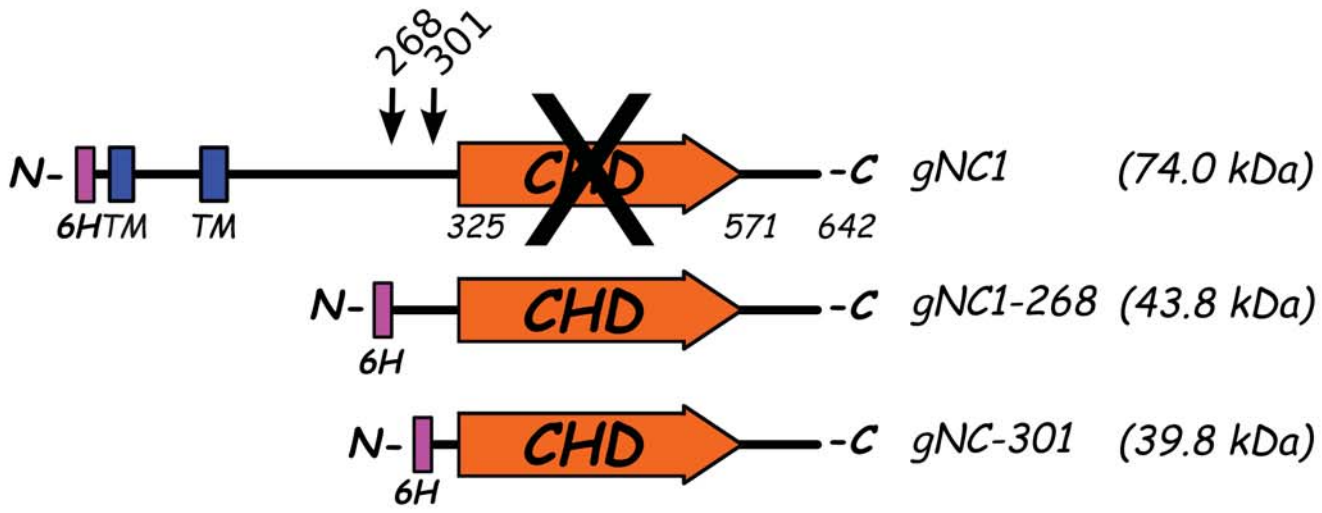



A)
gNC1 (G. lamblia) N-
CHD
-c 642
325
571
CD
gNC2 (G. lamblia) N-
CHD
CHD
-c 2190
288
465
642
1400
1810
2136
CD 1
CD 2
B)
CD
CD 1
CD 2
C)
AC (M. tuberculosis)
CHD
-c 567
AGC (Sp. platensis)
GAF
GAF
HisKA
HATPase
CHD
-c 1206
AC (T. cruzi)
Peripla_BP_6
CHD
-c 1321
ACp (T. gondii)
CHD
-c 952
AC1 (P. falciparum)
CHD
-c 884
rGC (H. sapiens)
ANF_receptor
Pkinase_Tyr
CHD
-c 1108
GCCD (T. gondii)
E1-E2_ATPase
HAD
PhoLip_ATPase_C
CHD
CHD
-c 4368
500
750
850
1900
2500
2700
3250
3650
GC (Par.tetraurelia)
E1-E2_ATPase
PhoLip_ATPase_C
CHD
CHD
-c 2302
GC (P. falciparum)
E1-E2_ATPase
HAD
PhoLip_ATPase_C
CHD
CHD
-c 3889
400
1350
2050
2500
AC1 (H. sapiens)
CHD
CHD
-c 1119
AC3 (D. melanogaster)
CHD
CHD
-c 1167
ACY1 (C. elegans)
CHD
CHD
-c 1253
sAC (H. sapiens)
CHD
CHD
-c 475

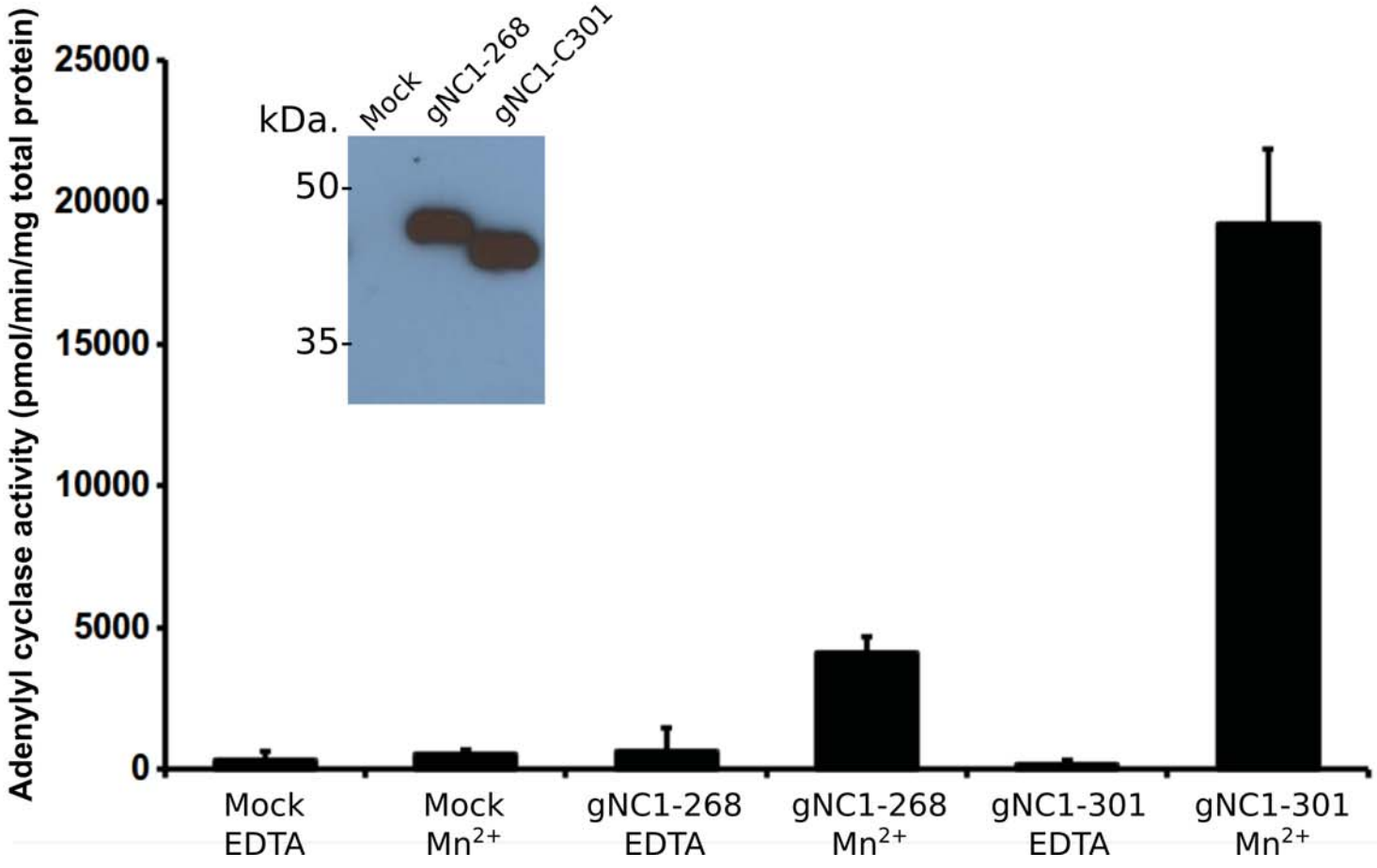

G. lamblia
gNC1
D338
D382
R495
T378
T517
N524
R528
R528
N524
T517
T378
R495
D382
D338
CD
CD
G. lamblia
gNC2
D292
T331*
S409*
K327
D432
N439
R443
L2094*
E2094*
D2087
K1989
R2066
D1993
D1948
CD1
CD2
T. gondii
gccd
D2949
E2994*
K3118*
E3989
A3139
N3146
R3150
M4158*
L4154*
D4147
C4071
R4126
D4078
D4031
CD1
CD2
C. elegans
ACY1
D324
D368
R412
S364*
D433
T440*
E444*
R1187
N1183
D1176
K1097
K1155*
P1103*
S1054*
CD1
CD2
D. melanogaster
AC3
D308
D352
R395
K348
D417
E424*
K428*
R1044
N1040
D1033
K953
K1012*
S957*
S906*
CD1
CD2
H. sapiens
AC1
D308
D352
R395
K348
D417
T424*
V428*
R1008
N1004
D997
K920
R976
S924*
S873*
CD1
CD2
N
C
ATP
ATP
G. lamblia
gNC1
G. lamblia
gNC2
H. sapiens
AC1


A)
mRNA
gGDH 5'
3' 1350 pb
gPDE 5'
3' 4116 pb
gNC1 5'
3' 1929 pb
gNC2 5'
3' 6573 pb
B)
gGDH
gGDH (RT-)
gPDE
gNC1
gNC2 (RT-)
gNC2
300 bp-
200 bp-
100 bp-
-300 bp
-200 bp
-100 bp

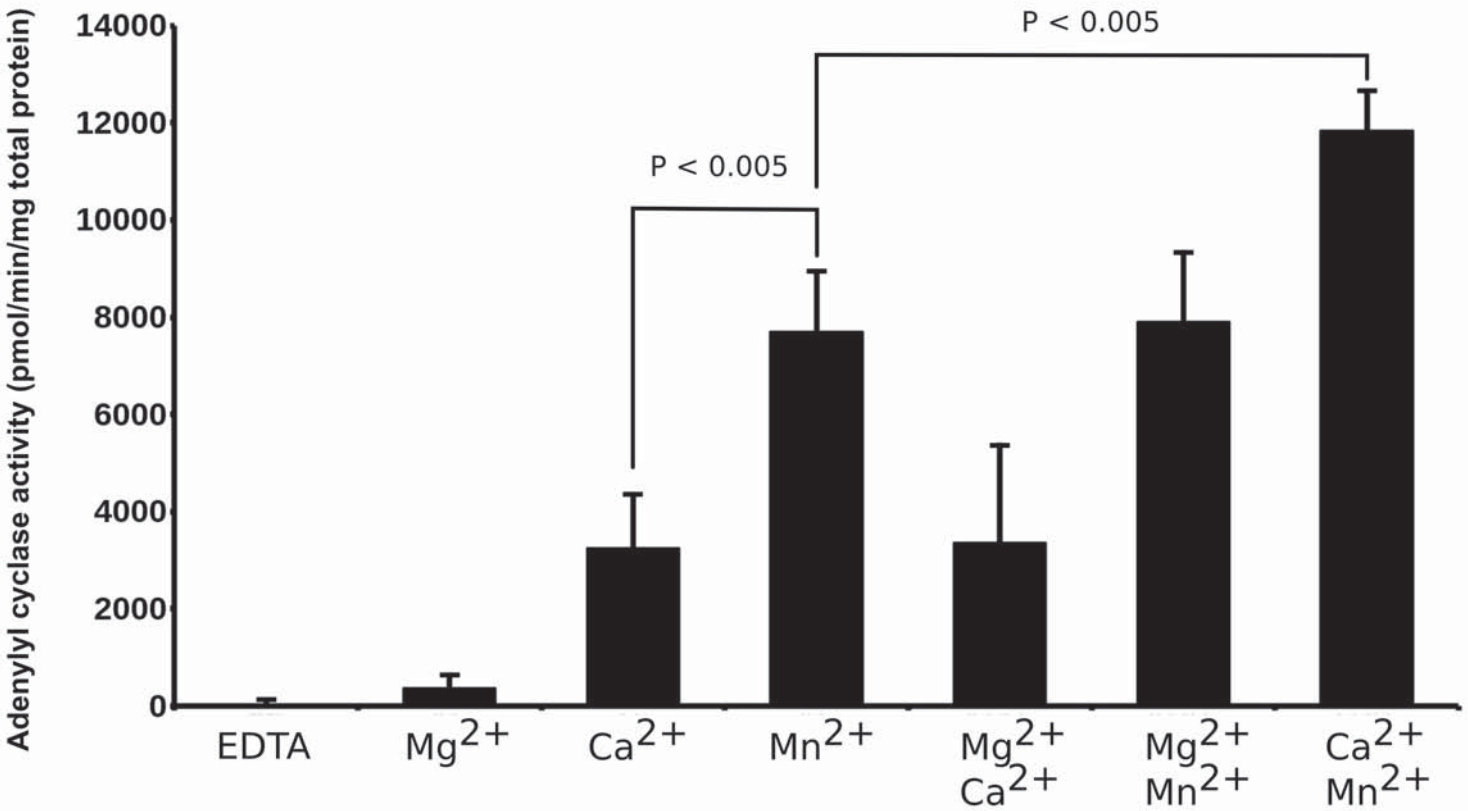
A)


268
301
N-
6H
TM
TM
CHD
-C
325
571
642
gNC1
(74.0 kDa)
N-
6H
CHD
-C
gNC1-268
(43.8 kDa)
N-
6H
CHD
-C
gNC-301
(39.8 kDa)

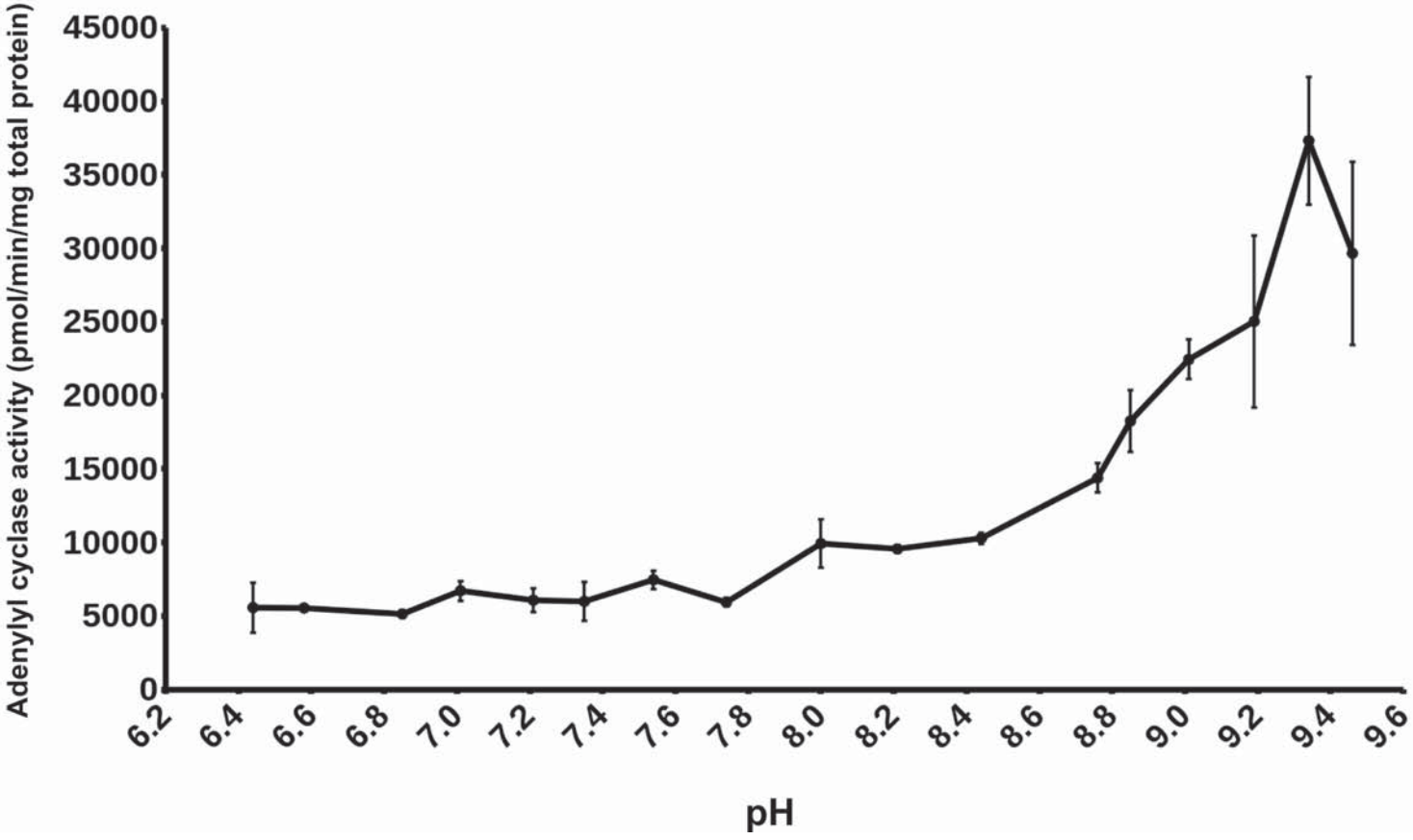
B)

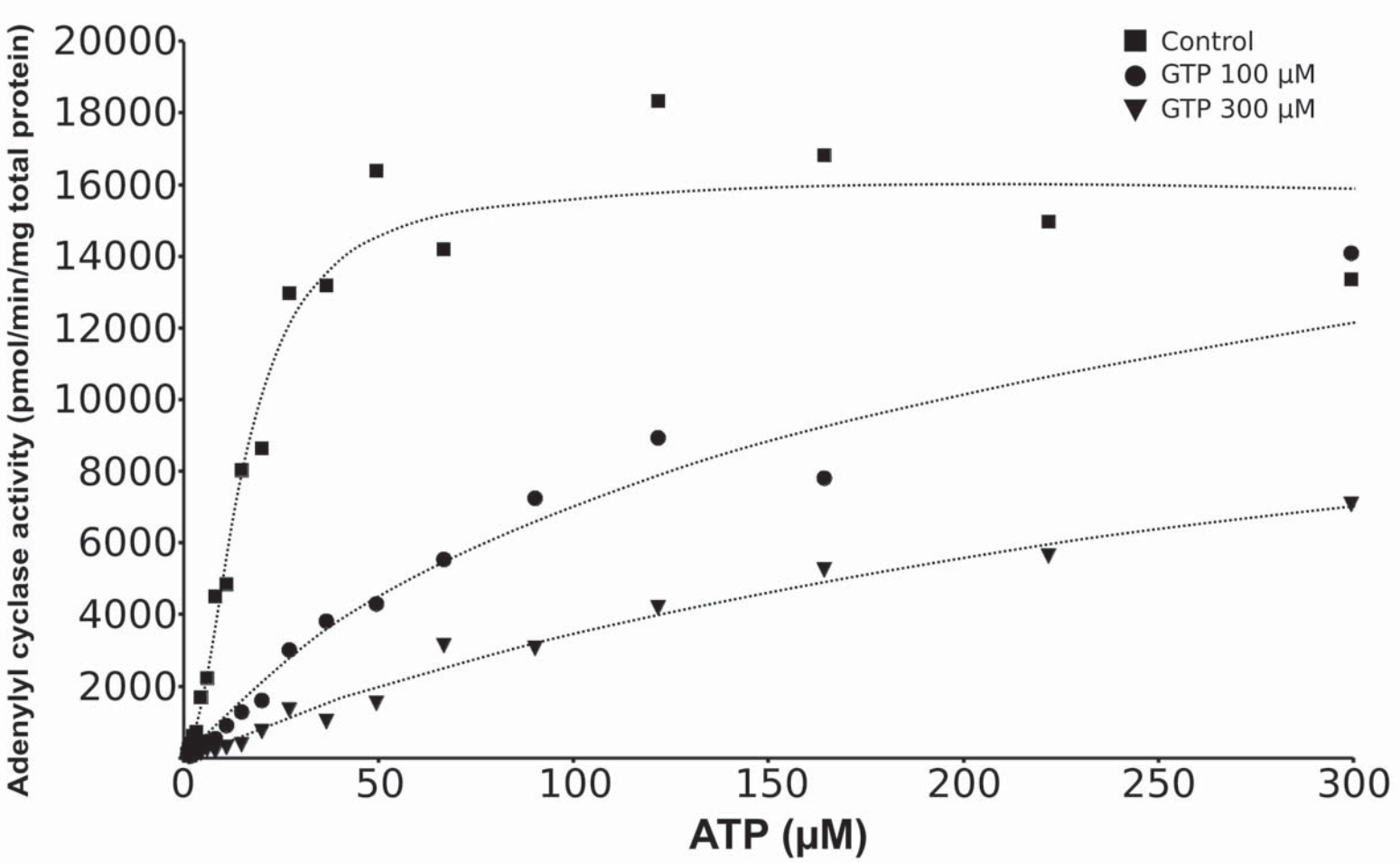

Adenylyl cyclase activity (pmol/min/mg total protein)
25000
20000
15000
10000
5000
0
kDa.
Mock
gNC1-268
gNC1-C301
50-
35-
Mock EDTA
Mock Mn2+
gNC1-268 EDTA
gNC1-268 Mn2+
gNC1-301 EDTA
gNC1-301 Mn2+

A)


Adenylyl cyclase activity (pmol/min/mg total protein)
14000
12000
10000
8000
6000
4000
2000
0
P < 0.005
P < 0.005
EDTA
Mg2+
Ca2+
Mn2+
Mg2+ Ca2+
Mg2+ Mn2+
Ca2+ Mn2+

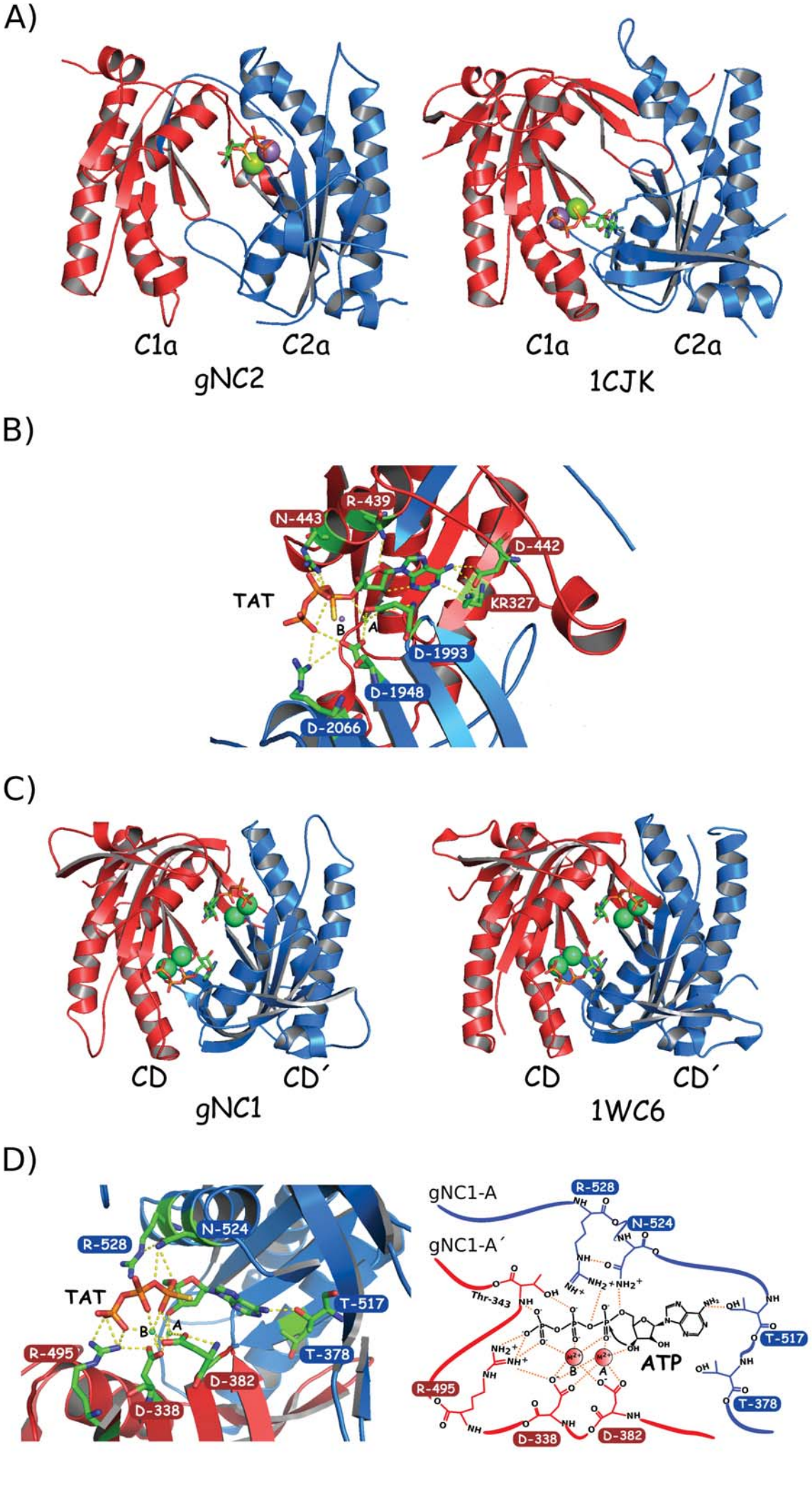
B)

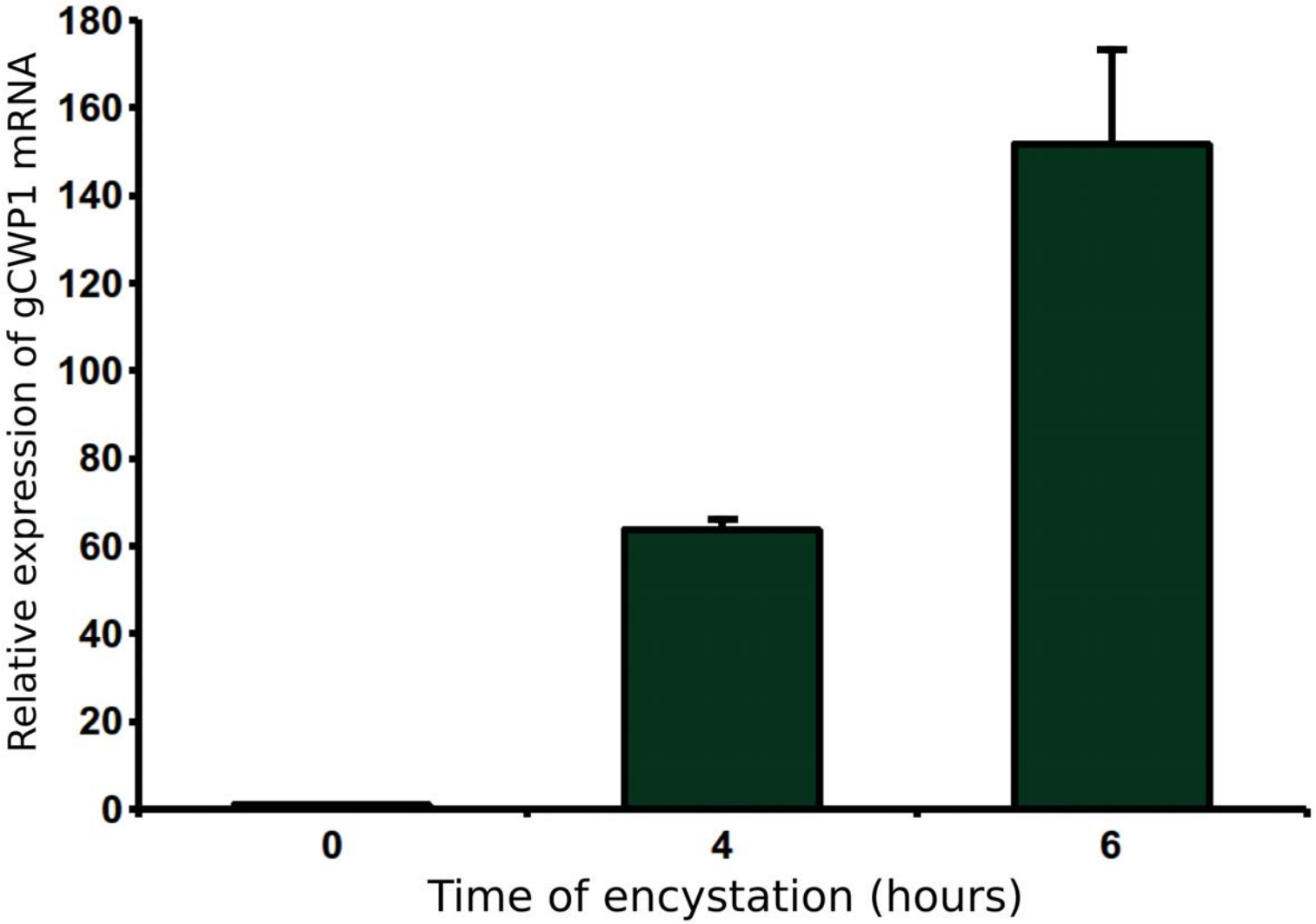

Adenylyl cyclase activity (pmol/min/mg total protein)
45000
40000
35000
30000
25000
20000
15000
10000
5000
0
6.2
6.4
6.6
6.8
7.0
7.2
7.4
7.6
7.8
8.0
8.2
8.4
8.6
8.8
9.0
9.2
9.4
9.6
pH



Control	$K_{0.5}$ = 15 µM +/- 1.4 µM	V_{MAX} =16021 +/- 618 *	Hill coeff. 1.98 +/- 0.30
GTP 100 µM	K_M^{app}= 125 µM +/- 30 µM	V_{MAX} =15808 +/- 2212 *	
GTP 300 µM	k_M^{app}= 334 µM +/- 64 µM	V_{MAX} =14934 +/- 1877 *	

* Adenylyl cyclase activity (pmol/min/mg total protein)


A)
C1a
C2a
gNC2
C1a
C2a
1CJK
B)
R-439
N-443
D-442
TAT
KR327
B
A
D-1993
D-1948
D-2066
C)
CD
CD´
gNC1
CD
CD´
1WC6
D)
N-524
R-528
TAT
T-517
B
A
R-495
T-378
D-382
D-338
gNC1-A
R-528
gNC1-A´
N-524
Thr-343
T-517
ATP
R-495
T-378
D-338
D-382


Relative expression of gCWP1 mRNA
180
160
140
120
100
80
60
40
20
0
0
4
6
Time of encystation (hours)

A)

B)

```
gPDE_CD_WB    1065 TLYHNKLHAADVAQMSMYMLSTVYCSLISESPKHPFLCVYKAMRQYKESDYSRLITEQPR
gPDE_CD_DH    1065 TLYHNKLHAADVAQMSMYMLSTVYCSLISESPKHPFLCVYKAMRQYKESDYSRLITEQPR
gPDE_CD_GS    1061 TLYHNKLHAADVAHMSMYMLSTVYCSLLDESPKHPFLYVYKAMQQYNENDYDRLVEEQPK
gPDE_CD_P15   1065 TLYHNKLHAADVAHMSMYMLSTVYCSLLDESPKHPFLYVYKAMQQYNENDYDRLVEEQPK
                   ************:************:.******** *****:**:*.**.**: ***:

gPDE_CD_WB    1125 QALIRPVDFLALLFGSLCHDLGHTGIDNLFCINTENALALLYNDEAPLEHAHATLSWHII
gPDE_CD_DH    1125 QALIRPVDFLALLFGSLCHDLGHTGIDNLFCINTENALALLYNDEAPLEHAHATLSWHII
gPDE_CD_GS    1121 RALIRPVDFLALLFGSLCHDLGHTGIDNLFCINTENALALLYNDEAPLEHAHATLSWHII
gPDE_CD_P15   1125 RALIRPVDFLALLFGSLCHDLGHTGIDNLFCINTENALALLYNDEAPLEHAHATLSWHII
                   :**************************************************************

gPDE_CD_WB    1185 TQMAVYFKHFTPCQYREFRALFLEIILATDMSTHFNFLRRLESLDEDLIIKILEHHNDSE
gPDE_CD_DH    1185 TQMAVYFKHFTPCQYREFRALFLEIILATDMSTHFNFLRRLESLDEDLIIKILEHHNDSE
gPDE_CD_GS    1181 TQMAIYFKHFTSSQYREFRALFLEIILATDMSTHFSFLRRLESFDKSFISKIFERHNTSE
gPDE_CD_P15   1185 TQMAIYFKHFTSSQYREFRALFLEIILATDMSTHFSFLRRLESFDKSFISKIFERHNTSE
                   ****:******..**********************.*******:*:.:* **:*:** **

gPDE_CD_WB    1245 IALLRWYILKVCIKFGDLSNPCRPIEISTRYAVALMNEFWSLGDLMLECGLEPDKIKTRP
gPDE_CD_DH    1245 IALLRWYILKVCIKFGDLSNPCRPIEISTRYAVALMNEFWSLGDLMLECGLEPDKIKTRP
gPDE_CD_GS    1241 ISLLRWYILKVCIKFGDLSNPCRPIEISTRYAVALMNEFWSLGDLMLECGLEPDKIKTRP
gPDE_CD_P15   1245 ISLLRWYILKVCIKFGDLSNPCRPIEISTRYAVALMNEFWSLGDLMLECGLEPDKIKTRP
                   *:**************************************************************

gPDE_CD_WB    1305 QKGEESLIIANSQIGFTQSIVKGFWTVVERFWKALAGVEFSDLQANLNATVEHWQNVRSE
gPDE_CD_DH    1305 QKGEESLIIANSQIGFTQSIVKGFWTVVERFWKALAGIEFSDLQANLNATVEHWQNVRSE
gPDE_CD_GS    1305 QKGEESLIIANSQIGFTQSIVKGFLIVVESFWKALADVEFQDLQANLSATVEHWQNIRAE
gPDE_CD_P15   1305 QKGEESLIIANSQIGFTQSIVKGFLIVVESFWKALADVEFQDLQANLNATVEHWQHVRSE
                   ************************  *** ******.:**.*** **.*******::*:*
```

***G. lamblia_WB* as Query (full-length)**

gPDE_WB: Identities = 1371/1371 (100%), Positives = 1371/1371 (100%), Gaps = 0/1371 (0%)

gPDE_DH: Identities = 1359/1371 (99%), Positives = 1363/1371 (99%), Gaps = 0/1371 (0%)

gPDE_GS: Identities = 1069/1371 (78%), Positives = 1184/1371 (86%), Gaps = 4/1371 (0%)

gPDE_P15: Identities = 1228/1371 (90%), Positives = 1291/1371 (94%), Gaps = 0/1371 (0%)

***G. lamblia_WB* as Query (cyc catalytic domains)**

gPDECD_WB: Identities = 191/191 (100%), Positives = 191/191 (100%), Gaps = 0/191 (0%)

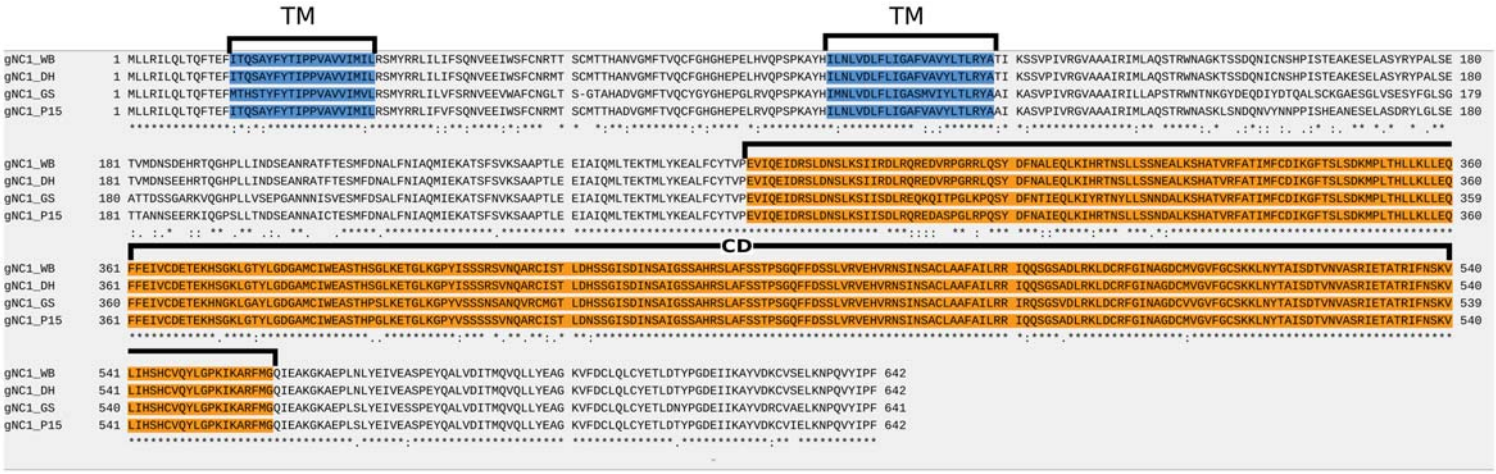
gPDECD_DH: Identities = 191/191 (100%), Positives = 191/191 (100%), Gaps = 0/191 (0%)

gPDECD_GS: Identities = 166/191 (87%), Positives = 176/191 (92%), Gaps = 0/191 (0%)

gPDECD_P15: Identities = 177/191 (93%), Positives = 186/191 (97%), Gaps = 0/191 (0%)

C)

A)



G. lamblia_WB as Query (full-length)

gNC1_WB: Identities = 642/642 (100%), Positives = 642/642 (100%), Gaps = 0/642 (0%)

gNC1_DH: Identities = 639/642 (99%), Positives = 640/642 (99%), Gaps = 0/642 (0%)

gNC1_GS: Identities = 536/642 (83%), Positives = 581/642 (90%), Gaps = 1/642 (0%)

gNC1_P15: Identities = 597/642 (93%), Positives = 611/642 (95%), Gaps = 0/642 (0%)

***G. lamblia_WB* as Query (cyc catalytic domains)**

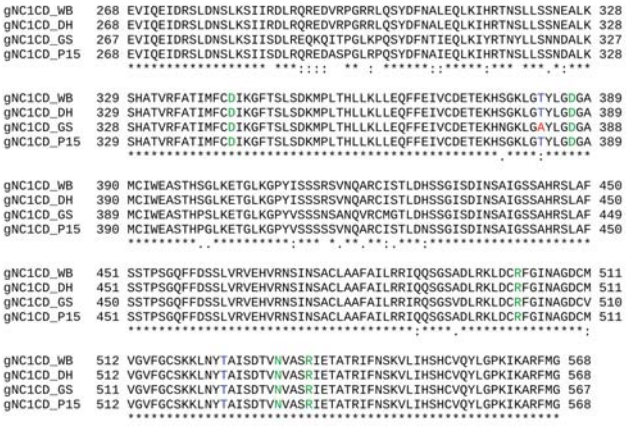
gNC1CD_WB: Identities = 296/296 (100%), Positives = 296/296 (100%), Gaps = 0/296 (0%)

gNC1CD_DH: Identities = 296/296 (100%), Positives = 296/296 (100%), Gaps = 0/296 (0%)

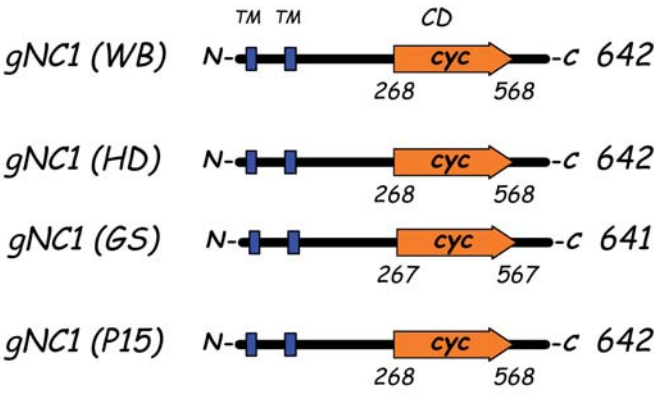
gNC1CD_GS: Identities = 267/296 (90%), Positives = 281/296 (95%), Gaps = 0/296 (0%)

gNC1CD_P15: Identities = 285/296 (96%), Positives = 289/296 (98%), Gaps = 0/296 (0%)

B)



C)



A)

B)

```
gNC2CD1_WB      288 AKNADYQKIDEYTKGFMSRLNLYYCLLDFLSECIGVQKIKSTYTLYMVAAGLESILDINY
gNC2CD1_DH      317 AKNADYQKIDEYTKGFMSRLNLYYCLLDFLSECIGVQKIKSTYTLYMVAAGLESILDINY
gNC2CD1_GS      288 VEGTSNQKIDDGTKGFMSRLNLYYCLLDFLSECIGVQKVKSTYTLYMVATGLESILDISY
gNC2CD1_P15     288 AKNADYQKIDECTKGFMSRLNLYYCLLDFLSECIGVQKIKSTYTLYMVAAGLESILDINY
                    .:.:. ****: **************************:*********:********.*

gNC2CD1_WB      349 KGPVSGYLLGQDDDQEPQAEQLLLQTDPLAYKLRAEKLKFKQIATFCLAARIAGESLGFN
gNC2CD1_DH      378 KGPVSGYLLRQDDDQGPQAEQLLLQTDPLAYKLRAEKLKFKQIATFCLAARIAGESLGFN
gNC2CD1_GS      349 KDTDNDFLLGTDRDQSIQ------QGGSLGCKAGAERLKFKQIATFCLAARIAGESLGFN
gNC2CD1_P15     349 KSPVNDYLLRQDGEQGQQAEQLLLQTDPLSYKLRIEKLKFKQISTFCLAARIAGESLGFS
                    *.. ..:**  * :*  *      * ..*. *   *:******:***************.

gNC2CD1_WB      410 VSAGITKGPIITGMLGQGQANCNFDVWGDVVNMAARIARGFNSGVFIKQDDFYSQLDD
gNC2CD1_DH      439 VSAGITKGPIITGMLGQGQANCNFDVWGDVVNMAARIARGFNSGVFIKQDDFYSQLDD
gNC2CD1_GS      404 VSAGITKGPIITGMLGQGQANCNFDVWGDVVNMAARIARGFNSGVFIKQDDFYTHVDE
gNC2CD1_P15     410 VSAGITKGPIITGMLGQGQANCNFDVWGDVVNMAARIARGFSSGVFIKQDDFYSQLDE
                    ****************************************.**********:::*:

gNC2CD2_WB     1930 LSYDRPIVYFARVGYLSLDIVNFTKLSSEHDAKSILRFLGDLFSEFDKRIDRAPQLKQIK
gNC2CD2_DH     1959 LSYDRPIVYFARVGYLSLDIVNFTKLSSEHDAKSILRFLGDLFSEFDKRIDRAPQLKQIK
gNC2CD2_GS     1921 LPYDRPIVYFGRVGYLSLDIVNFTKLSSEHDAKSILRFLGDLFSEFDKRIDRAPQLKQIK
gNC2CD2_P15    1927 LSYDRPIVYFARVGYLSLDIVNFTKLSSEHDAKSILRFLGDLFSEFDKRIDRAPQLKQIK
                    *.********.**************************************************

gNC2CD2_WB     1991 SIGDAYEVMGLPGILKKYGNAGKRRNKDSYGQKEDDRILQLYDGLLLLVMAGFGFIEDCI
gNC2CD2_DH     2020 SIGDAYEVMGLPGILKKYGNAGKRRNKDSYGQKEDDRILQLYDGLLLLVMAGFGFIEDCI
gNC2CD2_GS     1982 SIGDAYEVMGLPGMLKKYGGAEKRRNKHSYGQKEDDRILQLYDGLLLLVMAGFGFIEDCI
gNC2CD2_P15    1988 SIGDAYEVMGLPGMLKKYGNAGKRRNKDSYGQKEDDRLLQLYDGLLLLVMAGFGFIEDCI
                    ************:*****.* *****.*********:**********************

gNC2CD2_WB     2052 DVSRKHFSDKVDYLGLRVGVCTGVAFGSLLGSKQFRFDVFGNVPEEADLIQSEASINAIF
gNC2CD2_DH     2026 DVSRKHFSDKVDYLGLRVGVCTGVAFGSLLGSKQFRFDVFGNVPEEADLIQSEASINAIF
gNC2CD2_GS     2043 DINRKHFSDKVDYLGLRVGVCTGIAFGSLLGSKQFRFDVFGNVPEEADSIQSEASINTIF
gNC2CD2_P15    2049 DVSRKHFSDKVDYLGLRVGVCTGVAFGSLLGSKQFRFDVFGNVPEEADLVQSEASINAIF
                    *:.********************:********************** :*******:**

gNC2CD2_WB     2113 ISSDVKKILEEGESKWKSRFHEYQVEQ
gNC2CD2_DH     2087 ISSDVKKILEEGESKWKSRFHEYQVEQ
gNC2CD2_GS     2104 VSSDVKRILEEGENKWRSRFHEYQVEQ
gNC2CD2_P15    2110 ISADVKKILEEGESKWKSRFHEYQVEQ
                    :*:***:******.**:**********
```

***G. lamblia_WB* as Query (full-length)**

gNC2_WB: Identities = 2190/2190 (100%), Positives = 2190/2190 (100%), Gaps = 0/2190 (0%)

gNC2_DH: Identities = 2166/2190 (99%), Positives = 2174/2190 (99%), Gaps = 0/2190 (0%)

gNC2_GS: Identities = 1631/2180 (75%), Positives = 1853/2180 (85%), Gaps = 10/2193 (0.5%)

gNC2_P15: Identities = 1934/2187 (88%), Positives = 2050/2187 (94%), Gaps = 3/2187 (0%)

***G. lamblia_WB* as Query (cyc catalytic domains)**

gNC2CD1_WB: Identities = 178/178 (100%), Positives = 178/178 (100%), Gaps = 0/178 (0%)

gNC2CD1_DH: Identities = 176/178 (99%), Positives = 176/178 (99%), Gaps = 0/178 (0%)

gNC2CD1_GS: Identities = 139/172 (81%), Positives = 148/172 (86%), Gaps = 6/172 (3%)

gNC2CD1_P15: Identities = 163/178 (92%), Positives = 170/178 (96%), Gaps = 0/178 (0%)

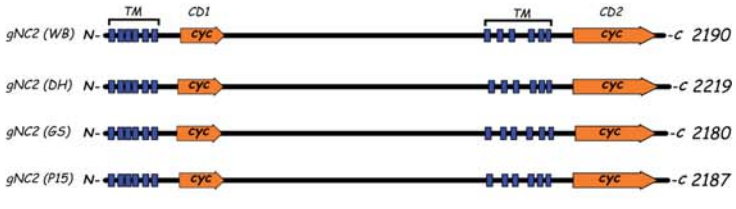
gNC2CD2_WB: Identities = 327/327 (100%), Positives = 327/327 (100%), Gaps = 0/327 (0%)

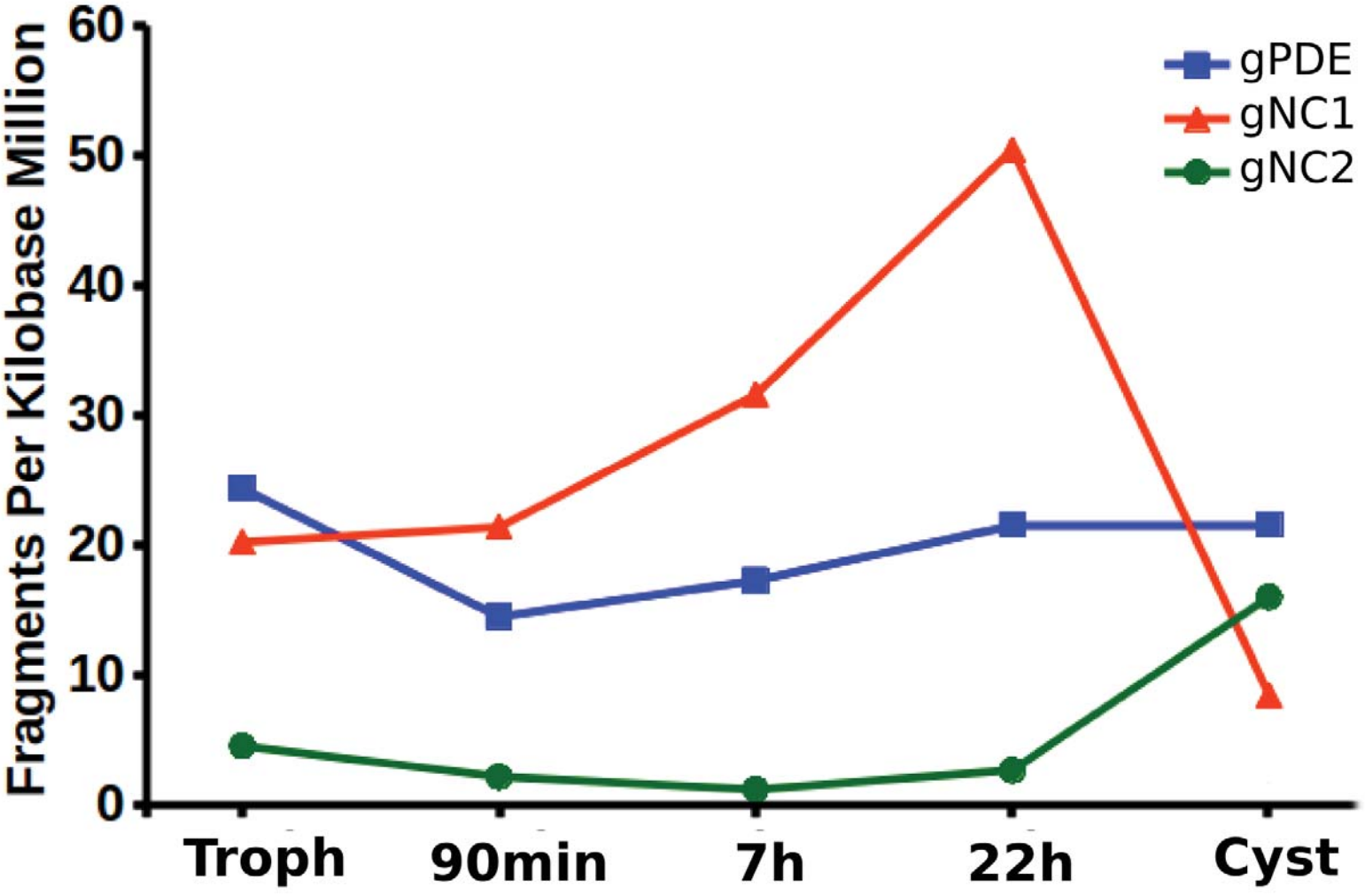
gNC2CD2_DH: Identities = 325/327 (99%), Positives = 325/327 (99%), Gaps = 0/327 (0%)

gNC2CD2_GS: Identities = 274/327 (84%), Positives = 300/327 (92%), Gaps = 0/327 (0%)

gNC2CD2_P15: Identities = 307/327 (94%), Positives = 320/327 (98%), Gaps = 0/327 (0%)

C)




Fragments Per Kilobase Million
60
50
40
30
20
10
0
gPDE
gNC1
gNC2
Troph
90min
7h
22h
Cyst