Aus dem Lehrstuhl für physiologische Chemie im Biomedizinischen Centrum (BMC)
der Ludwig-Maximilians-Universität München

Vorstand: Prof. Dr. Andreas Ladurner, PhD

Und

Aus dem Biozentrum, Bereich Genetik der Fakultät für Biologie
der Ludwig-Maximilians-Universität München

Evolution of ligand specificity of protein kinase A isoforms in the phylum Euglenozoa

Dissertation
zum Erwerb des Doktorgrades der Medizin an
der Medizinischen Fakultät der
Ludwig-Maximilians-Universität zu München

vorgelegt von

Veronica Teresa Ober

aus

Starnberg

Jahr

2022

Mit Genehmigung der Medizinischen Fakultät der
Universität München

Berichterstatter:	Prof. Andreas Ladurner, PhD
Mitberichterstatter:	Prof. Dr. F. Ulrich Hartl Prof. Dr. Axel Imhof Prof. Dr. Annette Nicke
Mitbetreuung durch	Prof. Dr. Michael Boshart
Dekan:	Prof. Dr. med. Thomas Gudermann
Tag der mündlichen Prüfung:	17.02.2022

Contents

Abstract vii

Zusammenfassung ix

List of figures xi

List of tables xiii

List of abbreviations xv

1 Introduction 1

1.1 Leishmania - an underestimated pathogen 1

1.1.1 The impact of Leishmaniasis on global health 1

1.1.2 The lifecycle of *L. donovani* 1

1.2 PKA and cAMP signalling 2

1.2.1 PKA - the protein and the activation mechanism 2

1.2.2 cAMP signalling in mammalian cells 4

1.3 PKA and cAMP signalling in trypanosomatids 4

1.4 The phylum Euglenozoa 6

2 Methods 9

2.1 Material 9

2.1.1 Plasmids 9

2.1.2 Oligonucleotides 10

2.1.3 Enzymes 10

2.1.4 Molecular biology kits 10

2.1.5 *E. coli* strains 10

2.1.6 Antibiotics 11

2.1.7 Buffers for protein purification and sample preparation 11

2.1.8 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-Page) 11

2.1.9 Software 12

2.2 Methods . . . 12
2.2.1 DNA cloning protocols . . . 12
2.2.2 Protein purification . . . 15
2.2.3 Nano differential scanning fluorimetry (nDSF) . . . 18
2.2.4 Isothermal titration calorimetry (ITC) . . . 19
2.2.5 Protein crystallization . . . 21

3 Results **25**
3.1 PKAR isoform genes in the phylum Euglenozoa – origin of diversity in ligand specificity . . . 25
3.2 The nucleoside specific EgPKAR2 isoform . . . 29
3.3 Characterization of EgPKAR1 in comparison with mammalian PKAR1α . . 33
3.3.1 Structural determinants of ligand specificity in EgPKAR1 . . . 33
3.3.2 Contribution of key residues to ligand binding affinity in EgPKAR1 . 38
3.4 Evidence for a single nucleoside binding domain in LdPKAR3 . . . 40
3.5 Cell permeable nucleoside analogues bind to LdPKAR1 and LdPKAR3 . . 47

4 Discussion **49**
4.1 Are nucleosides the endogenous ligands of cNMP independent PKA? . . . 50
4.2 LdPKAR3, a novel type of PKAR-like isoform in pathogenic *L. donovani* . . 51
4.2.1 The role of LdPKAR3 in the life cycle of *L. donovani* . . . 51
4.2.2 Structure-derived model of an alternative kinase activation mechanism . . . 53
4.2.3 Comparison of LdPKAR3 to the homology model of EgPKAR2 . . . 56
4.3 Evolution of PKA in distant eukaryotes . . . 57
4.3.1 Distribution of PKAR ligand specificity types within the Euglenozoa . 58
4.3.2 Loss or gain of PKAR ligand specificity types in euglenozoan phylogeny? . . . 59
4.4 Can ligand specificity of CNB domains be predicted from primary sequence? 62

5 Conclusion and perspectives **65**
5.1 Conclusion . . . 65
5.2 Outlook . . . 66

Bibliography **67**

Supplementary material **79**
1 Supplemental material - Chapter 2.2 . . . 79
2 Supplemental material - Chapter 3.1 . . . 81

3 Supplemental material - Chapter 3.3 85

Acknowledgement **87**

Affidavit **89**

Abstract

Protein Kinase A (PKA) is one of the most studied eukaryotic kinases as it plays a key role in regulation of several metabolic pathways and genome expression. It is a main downstream effector of cyclic adenosine monophosphate (cAMP) signalling and its regulation by cAMP was suggested to be conserved among all eukaryotes possessing PKA. Recent findings, however, revealed that PKA in *Trypanosoma brucei* and *cruzi* cannot be activated by cAMP, whereas nucleosides and nucleoside analogues bind and activate the kinase. Based on these findings we aimed to investigate the genome of closely related organisms, *Euglena gracilis* and *Leishmania donovani*, looking for orthologs of regulatory subunits that might also display cAMP independency. We report the existence of 2 regulatory subunits (PKAR) in *E. gracilis* (EgPKAR1 and 2) and also in *L. donovani* (LdPKAR1 and LdPKAR3). Biophysical analysis by Isothermal Titration Calorimetry (ITC) confirmed that 3 of the 4 proteins investigated do not interact with cAMP but with nucleosides. Findings described in the present study highlight cAMP independency in PKA as a phenomenon not unique to trypanosomes, but seen for other representatives of the Euglenozoa. We show that amino acid substitutions found in highly conserved regions called PBC (phosphate binding cassete) within the ligand binding domains alter its affinity for cAMP or rise the affinity for alternative ligands (such as nucleosides). By X-ray analysis of a co-crystal complex between EgPKAR1 and cAMP we show binding of cAMP to an atypical sequence motif that misses a residue previously thought to be essential. Furthermore, characterization of two PKAR isoforms from *L. donovani*, LdPKAR1 and LdPKAR3, reveal nucleoside binding of both, but not necessarily in two binding sites. In a co-crystal complex of LdPKAR3 with guanosine we show that nucleosides only bind to the B site in LdPKAR3 and the A site remains empty. Interestingly, we found site A of LdPKAR3 to be crystallized in a different rotation angle (in relation to site B) when compared to the previous crystal structures of ligand bound PKAR (PDB: 1RGS, 6FTF, 6FLO). Also, the αB/C helix that connects site A and B was found in an unusually stretched conformation that resembles a PKAC-bound state, or the "holoenzyme" conformation, even in the absence of PKAC. Our study suggests the existence of a broad variety of PKAR isoforms in the Euglenozoa, providing a glimpse on the evolution of this ancient signalling protein.

Zusammenfassung

Protein Kinase A (PKA) ist eine der am besten erforschten Proteinkinasen in Eukaryonten. Über das Mitwirken an verschiedenen metabolischen Prozessen sowie der Genexpression, hat PKA eine zentrale Bedeutung für die Zellfunktion von Metazoa und ebenso Protozoa. Im Säuger wird die Aktivität von PKA über das Binden von zyklischem Adenosinmonophosphat (cAMP) reguliert, und man ging bisher davon aus, dass dieser Regulierungsmechanismus für alle Eukaryonten konserviert sei. Neueste Ergebnisse hingegen zeigen, dass PKA in *Trypanosoma brucei*, nicht durch cAMP reguliert wird, sondern dass Nucleoside und Nukleosidanaloga die Kinase binden und aktivieren können. Auf dieser Erkenntnis aufbauend, untersuchten wir vier regulatorische Untereinheiten von PKA in den phylogenetisch verwandten Organismen *Euglena gracilis* und *Leishmania donovani* hinsichtlich ihrer Ligandenspezifität. Unter Verwendung von Isothermaler Titrationskalorimetrie (ITC) konnten wir zeigen, dass drei dieser vier untersuchten regulatorischen Untereinheiten kein cAMP jedoch Nukleoside binden können. Der Liganden Wechsel von cAMP hin zu Nukleosiden in PKA hat somit nicht nur in Trypanosomen stattgefunden, sondern auch in anderen Vertretern der Euglenozoa. Der Liganden Wechsel in PKA korreliert mit Abweichungen in einem hoch konservierten Sequenzmotif, der PBC (phosphate binding cassette). In Nukleosid spezifischer PKA finden wir einen Austausch einzelner Aminosäuren, welche die Bindung von Nukleosiden begünstigen, jedoch eine Bindung von cAMP unmöglich machen. Interessanterweise finden wir aber auch andere Abweichungen in der Aminosäuresequenz, wie beispielsweise in EgPKAR1 aus *E. gracilis*, die lediglich eine Verringerung der Affinität für cAMP bewirken. Mittels Röntgenstrukturanylse von EgPKAR1, konnten wir zeigen, dass trotz des Fehlens von vermeintlich essenziellen Aminosäuren, cAMP in der C-terminalen Bindungstasche binden kann. Des Weiteren zeigt die Charakterisierung zweier Nukleosid-bindender PKARs aus *L. donovani*, LdPKAR1 and LdPKAR3, dass zwar beide Orthologe ausschließlich Nukleoside binden aber nicht unbedingt in beiden Bindungstaschen. Mittels Röntgenstrukturanalyse eines Komplexes aus LdPKAR3 und Guanosine konnten wir zeigen, dass LdPKAR3 im Gegensatz zu LdPKAR1 und anderen Nukleosid-spezifischen PKARs, Nukleoside nur in der C-terminalen Tasche bindet. Die N-terminale Bindungstasche hingegen bleibt leer und ähnelt in ihrer Konformation eher einer Liganden-freien Struktur der Säuger PKARIα. Auch eine die beiden Taschen verbindende αB/C helix zeigt eine außergewöhnlich gestreckte Formation, die im Säugersystem überwiegend bei einer Interaktion mit der katalytischen Untereinheit beobachtet wird. Unsere Ergebnisse implizieren die Existenz

einer großen Vielfalt an PKA Isoformen in Euglenozoa, welche uns einen Einblick in die Entwicklung dieses sehr alten Signalproteins gewähren.

List of figures

1.1 Activation of PKA in higher Eukaryotes 3
1.2 Organisation of the cyclic nucleotide binding domain (CNB) 4

3.1 Consense motifs of cylic nucleotide and nucleoside binding PKA isoforms . 25
3.2 Protein-ligand interactions analyzed by ITC 27
3.3 Ligand binding specificities of investigated PKAR orthologs 29
3.4 Overall organization of TbPKAR, TcPKAR and EgPKAR2 30
3.5 Ligand binding sites in EgPKAR2 31
3.6 Comparison of capping residues in binding site A 32
3.7 Capping motif in EgPKAR2 Bsite 32
3.8 Thermodynamic signature of EgPKAR2 obtained by ITC 33
3.9 Control of protein quality before crystallization of EgPKAR1 34
3.10 Alignment of EgPKAR1 to mammalian RIα 35
3.11 Ball and stick model of EgPKAR1_CNB:A binding cAMP 35
3.12 Structural features of EgPKAR1_CNB:B 36
3.13 Comparison of PBC:B in mammalian RIα and EgPKAR1 37
3.14 Thermodynamic signature of EgPKAR1 obtained by ITC 37
3.15 Substitution of potential key residues in EgPKAR1(255-401) 39
3.16 Control of protein quality before crystallization of LdPKAR3 42
3.17 Arrangements of the two binding domains in LdPKAR3 43
3.18 Empty site A of LdPKAR3 43
3.19 Structural features of LdPKAR3 Bsite 44
3.20 Molecular docking to LdPKAR3 Bsite 45
3.21 Schematic representation of LdPKAR3 46
3.22 Interdomain interaction between site A, site B and C-terminus 47
3.23 Binding of nucleoside analogues to LdPKAR1, LdPKAR3 and TbPKAR *in vitro* 48

4.1 Developmental-stage specific phosphorylation of LdPKAR3 52
4.2 Modelling of an LdPKAR3 holoenzyme by an alignment to a RIα(91-244):C complex 55
4.3 Alignment of EgPKAR2 to LdPKAR3 on site A 57

4.4 Representation of ligand specificity types 58
4.5 Distribution of PKAR subtypes in the phylum Euglenozoa 59

S1 Charts of pUCIDT_AMP_plasmids 79
S2 Charts of pET_Sumo and pET_Duet_vector 79
S3 Protein purification 80
S4 Differing melting temperatures in dependency of protein preparation 81
S5 Multiple sequence alignment of PKAR in representatives of the *Euglenozoa* 82
S6 Differing melting temperatures in dependency of small molecule incubation 83
S7 Negative binding results for EgPKAR1, EgPKAR2, and LdPKAR3 84
S8 Electron density maps for EgPKAR1 CNB:A and B 85

List of tables

2.1 Plasmids 9
2.2 Oligonucleotides 10
2.3 Purchased Enzymes 10
2.4 Purchased kits 10
2.5 *E. coli* strains 10
2.6 Antibiotics 11
2.7 Buffers for protein sample preparation 11
2.8 SDS-PAGE 12
2.9 PCR mastermix 13
2.10 Temperatures for PCR thermocycling 14
2.11 Screening of potential ligands 19
2.12 Constraints for design of receptor grids 23

3.1 Ligand binding properties of PKA orthologs determined by ITC 28
3.2 Comparison of binding data and results from molecular docking 40
3.3 Data collection and refinement statistics for EgPKAR1(109-401) and LdPKAR3(321-647) 41
3.4 Binding of nucleoside analogues to LdPKAR1, LdPKAR3 and TbPKAR . . 48

4.1 Comparison of ligand binding affinities in diffeerent PKA orthologs 50
4.2 Comparison of important residues in RIα and our orthologs 63

S1 Comparison of folded and refolded protein by nDSF 82
S2 Melting temperatures T_m of native proteins in presence of ligand candidates 84

List of abbreviations

APS Ammonium persulfate

cAMP Cyclic adenosine monophosphate

CNB Cyclic nucleotide binding domain

DMSO Dimethylsulfoxide

DTT Dithiothreitol

IPTG Isopropyl-β-D-thiogalactopyranosid

IS Inhibitory sequence

ITC Isothermal titration calorimetry

KCl Potassium chloride

$MgCl_2$ Magnesium chloride

NaCl Natrium chloride

nDSF Nano differential scanning fluorometry

PKAC Protein Kinase A Catalytic subunit

PMSF Phenylmethylsulfonyl fluoride

RBC Ribose binding cassette

RIα Mammalian PKARIα

SDS-Page Sodiumdodecyl Sulfate Poly Acrylamide Gel Electrophoresis

SENP2 Sentrin-specific protease 2

SUMO Small ubiqutin related modifier

TEMED Tetramethylethylendiamin

Chapter 1

Introduction

1.1 Leishmania - an underestimated pathogen

1.1.1 The impact of Leishmaniasis on global health

Leishmania subspecies (spp), belonging to the class of Kinetoplastea, are parasitic protozoa that cause a variety of diseases. These range from single skin lesion in cutaneous leishmaniasis to fulminate mucocutaneous infestation to the severest form, visceral leishmaniasis or kala-Azar [1] [2]. Transmitted by sandflies and endemic in 97 countries , *Leishmania spp.*, have the highest global dispersion among the three human-infective kinetoplastids. These also include *T. brucei* and *T. cruzi*, and all three are highly relevant in terms of public health [3]. Special attention has to be drawn to the severe visceral Leishmaniasis, caused by *Leishmania donovani* and *infantum*. Varying period of incubation and potential manifestation in almost every organ often complicates diagnosis and delays therapy with lethality up to 95% in untreated patients [4]. The high burden of disease becomes even clearer, as visceral Leishmaniasis is endemic in mostly low-income countries with limited access to healthcare and adequate protection against transmission such as repellents or mosquito nets. Treatment with liposomal amphotericin B or other antiparasitic drugs often must be applied intravenous and does not necessarily erase the parasite, so that relapses are common [2].
Thus, research on Leishmaniasis, as a prominent example of neglected tropical diseases, with special regards to potential drug targets and the development of new therapeutics, is of great importance and a high motivation for this study.

1.1.2 The lifecycle of *L. donovani*

L. donovani is a dixenous parasite, i.e. it infects two different hosts during its life cycle, female sand flies and mammals, and has to adapt to both. Thereby, the parasite undergoes major morphological and functional changes. In the gut of the female sand flies, the

parasite develops into a so-called procyclic promastigote, that has an elongated shape with an adjacent flagellum [1]. Developing further into a metacyclic stage, the parasite locates to the anterior part of the digestive system, leading to intense pharyngeal infection in the sand fly. In this stage the parasite becomes infective to mammals and will be transmitted during a blood meal of the fly. Injected into peripheral tissue, the parasite is phagocytosed by mammalian macrophages and starts to round in their lysosoma, forming an amastigote with its flagellum reduced to a remnant [5] [1] [6]. Pursued replication in the macrophages eventually leads to rupture of the cells, allowing the parasite to egress and infect other macrophages. Is an infected mammal bitten by a female sand fly, amastigotes within macrophages are ingested with the blood and the life cycle starts again.
Host specific morphology is induced by environmental cues, especially temperature and acidity, in the mammalian and insect hosts [7]. Whereas in the gut of the female sand fly the parasite is exposed to a rather alkaline milieu and temperatures around 26°C, in mammalian macrophages higher temperatures as well as lower pH has to be tolerated. Mimicking these environmental changes in a host free system revealed that promastigotes start to differentiate to amastigotes when exposed to 37 °C and pH 5.5 [8]. This correlates with conditions of lysosomes in macrophages [7] [8].
Adaption to the host-specific environment is therefore likely to be coordinated by various pathways that transduce changes of outer conditions into inner cell signalling [9]. One approach to elucidate potential signalling pathways that regulate differentiation in *L. donovani* was an analysis of phosphorylation trends during differentiation performed by Tsigankov et al., 2014. In these, the authors came across a putative regulatory subunit of PKA (LdPKAR3) due to its remarkable stage-specific phosphorylation pattern [9]. For a second regulatory subunit of PKA in *L. donovani*, LdPKAR1, a correlation of expression and promastigote maturation has been observed [10]. Furthermore, its involvement into metacyclogenesis, a transformation into highly infective metacyclic promastigotes, has been suggested [11]. Thus, PKA in *L. donovani* is likely playing a role in parasite differentiation, making it an interesting protein to be further investigated.

1.2 PKA and cAMP signalling

1.2.1 PKA - the protein and the activation mechanism

Widespread among eukaryotic organisms and known to play a crucial role in protein regulation by phosphorylation, PKA is one of the best studied eukaryotic protein kinases (EPK) and one of the main effectors of cyclic adenosine monophosphate (cAMP) [12] [13] [14] [15]. In mammals, it is involved in different cellular processes, e.g. metabolism, gene

regulation, and cell differentiation [16] [17] [18] [19].
In higher eukaryotes, PKA is composed of two regulatory (PKAR) and two catalytic (PKAC) subunits. In the absence of cAMP all four subunits are assembled as build an inactive, tetrameric R_2C_2 holoenzyme (Fig. 1.1, left) [16].

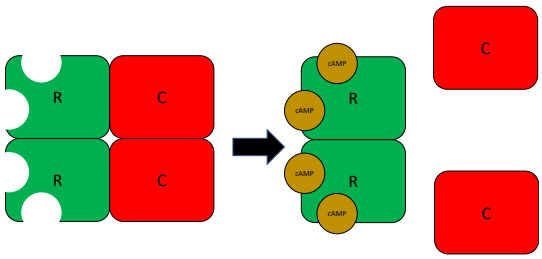


Figure 1.1: PKA activation mechanism of higher eukaryotes is shown as cartoon. Green squares represent PKAR, red squares PKAC. On the left, an inactive holoenzyme is shown as tetrameric structure R_2C_2. On the right, activation of PKA upon cAMP binding (brown circles) is shown.

Upon activation, two cAMP molecules are bound per PKAR molecule in so-called cyclic nucleotide binding domains (CNB). cAMP has to bind to the C-terminal binding site, CNB:B, first so that the N-terminal domain, CNB:A, becomes accessible. When cAMP binds to CNB:A, a conformational change of the R-subunit is induced. This conformational change leads to lower affinity for the catalytic subunit and favours the disruption of an R-C-interface [20] [21]. Consequently, C- and R-subunit dissociate, which enables the activated catalytic subunits to interact with their substrates (Fig. 1.2, right) [16].
Of particular interest are the CNBs, that bind cAMP. These are not unique to PKA but widespread in proteins of bacteria and eukaryotes. In archaea they are only exceptionally present, in plants they are completely absent [22]. CNBs are around 120 amino acids long, arranged in α helices and an eight-stranded β barrel, with the latter harbouring the ligand [23]. As the name implicates, CNBs were thought to exclusively bind cyclic nucleotides, and thereby mainly the ancient second messengers cAMP and cGMP. This selectivity was explained by a highly conserved feature, the so-called phosphate binding cassette (PCB). It is located within a loop, which is followed by a small helix between two beta strands of the barrel, β-strand 6 and β-strand 7. Composed of 14 amino acids of the following sequence:

F_1-G_2-E_3-[LIV]$_4$-A_5-L_6-[LIMV]$_7$-X_8-X_9-X_{10}-[PV]$_{11}$-R_{12}-[ANVQ]$_{13}$-A_{14} [24] [25]

The residues in the PBCs mainly interact with the phosphate moiety of cyclic nucleotides. Especially, arginine in position 12 plays a crucial role upon cyclic nucleotide binding as

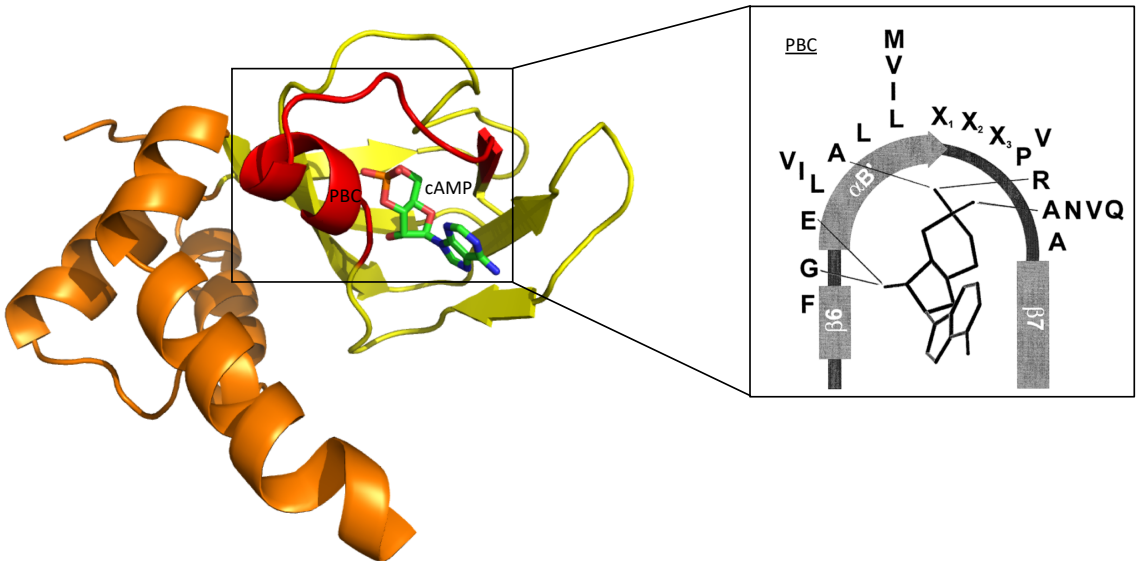


Figure 1.2: Organization of cyclic nucleotide binding domain (CNB). Secondary structure of CNB:A in RIα (PDB: 1RGS) according to the illustration of Figure 1b in the manuscript of Berman et al, 2005 [23]. In the insert, a segment that is defined as PBC is highlighted, taken from Figure 7 Canaves and Taylor, 2002 [24]. Consensus residues are shown in relative position to the cAMP molecule within the PBC.

it interacts with the phosphate moiety and was thought to be indispensable for cyclic nucleotide binding [25] [26] [27].

1.2.2 cAMP signalling in mammalian cells

cAMP is an extensively studied second messenger that plays a crucial role when extracellular stimuli are translated into intracellular response [28]. Such extracellular stimuli like hormones or neurotransmitter bind to a G-protein coupled receptor (GPCR) in the cell membrane. Subsequently, the GPCR associates with a G-protein. A subunit (G_α) separates in the next step and activates adjacent adenylate cyclases (AC). ACs, which produce cAMP, interact with various downstream effectors [16]. Along with cyclic nucleotide gated ion channels and EPACs (exchanging protein directly activated by cAMP), PKA is the main receptor for cAMP in mammalian cells [29]. Activated PKA regulates its substrates via phosphorylation on threonine and serine residues.

1.3 PKA and cAMP signalling in trypanosomatids

To date, there are very few studies of cAMP signalling in kinetoplastids. However, they all find substantial differences in the cAMP signalling pathways compared to mammalian

cells. Although the general machinery of cAMP production through ACs and cAMP degradation by phosphodiesterases (PDEs) is conserved, kinetoplastids do not possess GPCR or G proteins [30] [31].

After the discovery of the first kinetoplastid AC in *T. brucei*, the so-called ESAG 4, which was identified in the expression site of the variant surface glycoprotein [32] [33], similar ACs have also been found in *L. donovani* and *T. cruzi* [34]. Remarkably, the structure of kinetoplastid ACs differs significantly from the structure of mammalian ACs. Instead of two membrane spanning cassettes consisting of six transmembrane domains and two intracellular catalytic domains each, kinetoplastid ACs possess only a single transmembrane domain flanked by a N-terminal extracellular domain and a C-terminal intracellular catalytic domain [35]. Although it has been suggested that the N-terminal domain takes on the role of GPCRs, to date no possible ligand has been identified that would bind to this domain [36]. The importance of ACs for the parasites becomes clear as a lethal phenotype was observed upon knock down of ESAG4 and two further ESAG4 like ACs in *T. brucei* [33], and AC activity was shown to be important for early host infection [37]. In *T. cruzi*, ACs have been suggested to play a role in cell differentiation in the insect host [38].

Together with ACs, PDEs are crucial for cAMP homoeostasis in the cells, as they degrade cAMP. PDEs can be divided into classes I, II and III, according to their individual catalytic domains [39]. In kinetoplastids, we find only class I PDEs and these can further be subdivided into PDE-A, PDE-B, PDE-C and PDE-D, with at least one copy of each PDE present in the genomes of *T. brucei*, *T. cruzi* and *Leishmaina spp.* [39] [40] [41]. High synteny among the three parasites was found for PDE-B of which two paralogs, PDEB1 and 2, exist [41]. In *T. brucei*, PDEB1 and 2 were seen to be essential for the parasite as simultaneous knock-down led to cell death in bloodstream forms [42]. A similar phenotype was observed after inhibition of trypanosmal PDEs by CpdA (tetrahydrophthalazinone compound A) [43] [44]. These findings highlight the importance of cAMP homoeostasis and presumably cAMP signalling for the parasites.

With focus on downstream effectors of cAMP, attention has to be drawn to PKA. In mammals, four functionally non-redundant PKAR isoforms, RIα, β and RIIα, β are present. In all three parasitic trypanosomatids, we find one copy of a regulatory subunit to be conserved. In *Leishmania* and *T. cruzi* a second PKAR ortholog, PKAR3 type, has been identified and interaction with a catalytic subunit C3 has recently been shown for LdPKAR3 from *L. donovani* [Fischer Weinberger, manuscript in preparation]. Compared to the mammalian R subunits, the regulatory subunit in trypanosomatids has a longer N-terminus that lacks a dimerization/docking domain [45] and whether PKA in kinetoplastids can be activated by cAMP has long been subject of research.

In *T. brucei*, PKA could not be activated by cAMP [46] [47] [45]. Instead, cGMP led to

kinase activation, but only at concentration >100μM [47] [48]. Considering the absence of PKG, guanylate cyclases and cGMP in trypanosomatids [49] and the high concentrations of cGMP required, PKA activation by cGMP is unlikely a physiological phenomenon [50].
For *T. cruzi*, the existence of a cAMP responsive PKA (TcPKA) has been reported by Huang et al., 2006 [51]. However, cAMP responsiveness of this protein has to be carefully scrutinized. The catalytic subunit (TcPKAc) has been expressed in mammalian cells and could interact with mammalian R subunits, as earlier reported for PKAC in *T. brucei* [52]. Contamination of the preparation with mammalian R subunits can therefore not be excluded, and it remains unclear whether kinase activation by cAMP was measured for a holoenzyme formed between TcPKAc and mammalian R subunit or TcPKA.
For *Leishmania*, several studies suggest activation of PKA by cAMP: Overexpression of a globin-coupled heme containing adenylate cyclase in *L. major* led to accumulation of cAMP and increase in PKA activity, letting the authors conclude that cAMP would likely induce the kinase [53]. For *L. donovani* it has been reported that recombinantly expressed and purified LdPKAR1 binds cAMP in both binding sites (K_d = 7.15μM and K_d = 5.6μM in CNB:A and B, respectively) [11]. Considering a significant lower binding affinity of mammalian RIα for cAMP (K_d = 28nM [45]), cAMP as a physiological activator of PKA in *L. donovani* has at least to be regarded with caution.
Although cyclic nucleotide dependency of PKA in kinetoplastids has long been questioned, it was only recently that Bachmaier et al. had found an alternative activator of PKA in *T. brucei* [45]. In their study, they show that cell permeable nucleoside analogues are able to bind and activate PKA, *in vivo* and *in vitro* [45]. Further results from our lab show cAMP independency of PKA also for *T. cruzi* and that both, PKA from *T. brucei* and *T. cruzi*, can be activated by nucleosides (George Githure, Yuri Volpato, manuscript in preparation). For other organisms, cAMP independency of PKA has not been reported and ligand specificity of the second regulatory subunit, PKAR3, found in *Leishmania spp.* and *T. cruzi* has not been determined to date.
Characterization of PKAR orthologs in the direct neighbours of trypanosomes, i.e. *Leishmania spp.*, as well as their more distant relatives is of high interest. Here, we want to investigate whether all orthologs in trypanosomatids are cAMP independent and whether other organisms outside the trypanosomatids also possess cAMP independent PKA that can be regulated by nucleosides or other so far unidentified ligands.

1.4 The phylum Euglenozoa

Phylogenetically related but not a direct neighbour to *T. brucei*, perfect candidate for our investigations is the free-living algae *E. gracilis*, found in ponds and other standing water.

It has acquired a chloroplast by endosymbiosis of a green alga, which results in the physiological composition of the Euglena cell in a mixotrophic organism. Hence, it does not rely on a single source of nutrition [54] [55]. It is one of the best studied organism from the euglenids that form, together with kinetoplastids and the two smaller groups, diplonemids and symbiotids, the Euglenozoa.
Euglenoza are a monophyletic group of single celled flagellates belonging to the Excavata [56] [54]. Representatives of these group can be discriminated by the means of their nutrition, free-living or parasitic lifestyle or natural habitat. Whereas most of kinetoplastids and diplonemids show either nutrition by parasitism or predation, feeding via photoautotrophy is exclusively found in some organisms of the euglenids [57][58] [59]. A parasitic lifestyle seems to have evolved several times within kinetoplastids and is not restricted to the three human infective trypanosomatids [60]. Quite the contrary, as we find the mosquites parasitizing *Paratrypanosoma confusum* as well as plants parasitizing *Phytomonas* to be included in this group [61].
Nevertheless, the highest scientific interest is in the human infective species of kinetoplastids, *Leishmania*, *T. brucei* and *T. cruzi* with save and efficient drugs badly needed against these deadly pathogens.
Understanding the evolution of the ligand switch of PKA in Euglenozoa can enable future studies on potential inhibitors of parasitic PKA. Due to the high homology of PKAR in all three human infective trypanosomatids, interfering with the activation mechanism of PKA by blocking the parasite specific PKA can be a promising approach to develop a single drug against all three parasites. Phylogenetic analysis can be of good use as they help us to identify residue substitutions that are correlated to the ligand switch of PKA. Considering these critical residues in the developmental process of a potential PKAR inhibitor might be the key to highly affine binders that do not induce a conformational change of the R-subunit as required for PKA activation.

For this work, we chose two PKAR isoforms from free-living *Euglena gracilis* and parasitic *Leishmania donovani*, respectively, with the following question to be addressed:

1. Do we find cAMP independent PKA orthologs beyond trypanosomes?
2. Can we correlate ligand specificity to sequence and structural aspects of PKA in *E. gracilis* and *L. donovani*?
3. Are we able to determine where and when in evolution a cAMP independent PKA has appeared?
4. Can these insights suggest ways to target the unconventional PKAR of kinetoplastids for drug development?

Chapter 2

Methods

2.1 Material

2.1.1 Plasmids

Table 2.1: Plasmids that were used in this thesis

Protein	Construct	Cloning sites	Selection marker	Origin
EgPKAR1(109-401)	pUCIDT_AMP_EgPKAR1(109-401)	BamHI/ NotI	Ampicillin	IDT Integrated DNA Technologies, Inc., USA
EgPKAR2(183-483)	pUCIDT_AMP_EgPKAR2(183-483)	BamHI/ NotI	Ampicillin	IDT Integrated DNA Technologies, Inc., USA
TbPKAR(199-499)	pET_Duet_6xHisTbR(199-499)	BamHI/ NotI	Ampicillin	Yuri Volpato
LdPKAR3(1-647)	pET_Duet6H/TEV_LdPKARlike	BamHI/ NotI	Ampicillin	Meanie Obermeier, Eleni Polatoglou
LdPKAR1(1-502)	pET_Duet_LdPKAR	BamHI/ NotI	Ampicillin	Eleni Polatoglou
eGFP	pET11_6HIS_Sumo3_eGFP	BamHI/ NotI	Kanamycin	Kristina Malenica
EgPKAR1(109-401)	pETM11_6HIS_Sumo3_EgPKAR1(109-401)	BamHI/ NotI	Kanamycin	Veronica Ober
EgPKAR1(109-292)	pETM11_6HIS_Sumo3_EgPKAR1(109-292)	BamHI/ NotI	Kanamycin	Veronica Ober
EgPKAR1(255-401)	pETM11_6HIS_Sumo3_EgPKAR1(255-401)	BamHI/ NotI	Kanamycin	Veronica Ober
EgPKAR1(255-401)_S347A	pETM11_6HIS_Sumo3_EgPKAR1(255-401)_S347A	BamHI/ NotI	Kanamycin	Veronica Ober
EgPKAR1(255-401)_C354A	pETM11_6HIS_Sumo3_EgPKAR1(255-401)_C354A	BamHI/ NotI	Kanamycin	Veronica Ober
EgPKAR1(255-401)_S347A_C354A	pETM11_6HIS_Sumo3_EgPKAR1(255-401)_C354A_S347A	BamHI/ NotI	Kanamycin	Veronica Ober
EgPKAR2(183-483)	pETM11_6HIS_Sumo3_EgPKAR2(183-483)	BamHI/ NotI	Kanamycin	Veronica Ober
EgPKAR2(183-361)	pETM11_6HIS_Sumo3_EgPKAR2(183-361)	BamHI/ NotI	Kanamycin	Veronica Ober
EgPKAR2(327-483)	pETM11_6HIS_Sumo3_EgPKAR2(327-483)	BamHI/ NotI	Kanamycin	Veronica Ober
LdPKAR3(321-647)	pETM11_6HIS_Sumo3_LdPKAR3(321-647)	BamHI/ NotI	Kanamycin	Veronica Ober
LdPKAR1(200-502)	pETM11_6HIS_Sumo3_LdPKAR1(200-502)	BamHI/ NotI	Kanamycin	Veronica Ober

Charts of pUCIDT_AMP, pETM11_6HIS_Sumo3 and pET_Duet plasmids are shown in Fig. S1 and S2.

2.1.2 Oligonucleotides

Table 2.2: Oligonucleotides used for plasmids listed above

Primer	Sequence	Plasmids
EgPKAR1(109-292)_FW	ACA GAC CGG TGG ATC CAT GTA TGT AGA CCG TCG TAG AGG AGT GA	pETM11_6HIS_Sumo3_EgPKAR1(109-292)
EgPKAR1(109-292)_RV	ATC AGC GGC CGC CTA AAG GGC GTC GGC CAT CGT CAT A	pETM11_6HIS_Sumo3_EgPKAR1(109-292)
EgPKAR1(255-401)_FW	ata ttG GAT CCA TTC TGC TGA CTC AGA CGA TGC GCA AGC GGA GAC AAT ATG AGG ACT TCC TTG CAC AGG TGC CCC	pETM11_6HIS_Sumo3_EgPKAR1(255-401)
EgPKAR1(255-401)_RV	TGC GGC CGC TTA CTC TTG GAC TTC CTC ACG AGG AGC TGA ATA ATT TTG CAT ATT GCG TTG CAG AAT CTC TTC CAT	pETM11_6HIS_Sumo3_EgPKAR1(255-401) pETM11_6HIS_Sumo3_EgPKAR1(255-401)_S347A pETM11_6HIS_Sumo3_EgPKAR1(255-401)_C354A pETM11_6HIS_Sumo3_EgPKAR1(255-401)_S354A_C354A
EgPKAR1(255-401)_S347A_FW	AAC ATT CTG GGC CCC AAA GAT TTC TTC GGC GAG ATG GCA TTA ATG TTC AAT CAG CCC TGT GTG GCA	pETM11_6HIS_Sumo3_EgPKAR1(255-401)_S347A
EgPKAR1(255-401)_C354A_FW	ACA TTC TGG GCC CCA AAG ATT TCT TCG GCG AGA TGA GTT TAA TGT TCA ATC AGC CCG CAG TGG CAA CCG TAG TT	pETM11_6HIS_Sumo3_EgPKAR1(255-401)_C354A
EgPKAR1(255-401)_S347A_C354A_FW	ACA TTC TGG GCC CCA AAG ATT TCT TCG GCG AGA TGG CAT TAA TGT TCA ATC AGC CCG CAG TGG CAA CCG TAG TTA GCG AG	pETM11_6HIS_Sumo3_EgPKAR1(255-401)_S354A_C354A
EgPKAR2(183-361)_FW	AGA CCG GTG GAT CCA TGT CGA GAC GCC GCA CTG	pETM11_6HIS_Sumo3_EgPKAR2(183-361)
EgPKAR2(183-361)_RV	ATG AGC GGC CGC CTA CAG AGC ATC GGC GAT CTG	pETM11_6HIS_Sumo3_EgPKAR2(183-361)
EgPKAR2(327-483)_FW	AGA GGG ATC CGG AGA AAT GAT GAC TCG GCG CAA ACA GT	pETM11_6HIS_Sumo3_EgPKAR2(327-483)
EgPKAR2(327-483)_RV	GTG CGG CCG CCT AGG CTG CGT CTA TTG ATG A	pETM11_6HIS_Sumo3_EgPKAR2(327-483
LdPKAR3(321-647)_FW	TAG TGG ATC CCA GCG CCC AAG CCG CGG	pETM11_6HIS_Sumo3_LdPKAR3(321-647)
LdPKAR3(321-647)_RV	att atg cgg ccG CTT ACT GGA CCG CCG C	pETM11_6HIS_Sumo3_LdPKAR3(321-647)
LdPKAR1(199-592)_FW	TAG TGG ATC CGG CCG GGC TCG TCG CCA GAC AG	pETM11_6HIS_Sumo3_LdPKAR1(199-592)
LdPKAR1(199-592)_RV	att atg cgg ccG CCT ACG CGC TCT CGT G	pETM11_6HIS_Sumo3_LdPKAR1(199-592)

2.1.3 Enzymes

Table 2.3: Purchased enzymes used for cloning

Products	Company
Restriction enzymes: BamHI , NotI, ApaI	New England Bioloabs (NEB), USA
Ligase: T4 DNA Ligase	New England Bioloabs (NEB), USA
DNA Polymerase: Q5 Polymerase	Promega, USA

2.1.4 Molecular biology kits

Table 2.4: Purchased kits for cloning and protein expression

Name	Purpose	Company
Hi Yield® Plasmid Mini DNA Isolationkit	DNA purification	Süd-Laborbedarf GmbH, Germany
PureYield^{TM} Plasmid Miniprep System	DNA purification	Promega, USA
Hi Yield® Gel/PCR DNA Fragment Extraction Kits	DNA recovery from gel / after PCR	Süd-Laborbedarf GmbH, Germany
QuickFold^{TM} Protein Refolding Kit	protein refolding	Athena Environmental Sciences, Inc. ,USA

2.1.5 *E. coli* strains

Table 2.5: *E. coli* strains used for cloning and protein expression

Name	Purpose of use	Resistance	source
E. coli XL10 Gold	Plasmid cloning	Tetracycline and Chloramphenicol	NEB
E. coli Rosetta (DE3)	Protein expression	Chloramphenicol	NEB

2.1.6 Antibiotics

Table 2.6: Antibiotics used for cloning or protein expression

Name	Stock concentration
Chloramphenicol	50 µg/ml Ethanol
Kanamycin	34 µg/ml H_2O
Ampicillin	100 µg/ml H_2O

2.1.7 Buffers for protein purification and sample preparation

Table 2.7: Listed buffers are used for protein purification, protein stability assays and Isothermal titration calorimetry (ITC)

Name	Components
resuspension buffer	50mM Hepes pH 7.5; 50mM NaCl; 30mM Imidazole
elution buffer	50mM Hepes pH 7.5; 50mM NaCl; 250mM Imidazole
minimal buffer	50mM Hepes pH 7.5; 50mM NaCl
buffer A [62]	250mM Citric acid; 375mM Hepes, 500mM CHES
buffer 15	50mM Tris-Cl pH 8.5; 240mM NaCl; 10mM KCl; 2mM $MgCl_2$; 0.4 M sucrose; 10mM DTT
buffer 10	50mM Tris-Cl pH 8.5; 9.6mM NaCl; 0.4mM KCl; 2mM MgCL2; 2mM CaCl2; 0.5M Arginine; 0.75M Guanidine-HCl; 0.4M Sucrose; 10mM DTT
crystallization buffer	100mM NaCl; 10mM Hepes pH 7.0 or 7.5; 1mM DTT
PD10 buffer	8M Urea, 10mM Hepes pH 8.0, 50mM NaCl – sterile filtered

2.1.8 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-Page)

Purity of protein samples is analysed by gel electrophoresis in 12% polyacrylamide gels. Protein samples are mixed 6:1 with Laemmli buffer and boiled for 5 minutes at 95°C before loading onto the gel. Protein samples were run for 60 minutes at approximately 100V. Gels were stained overnight in Colloidal Coomassie staining solution and destained the next day using deionized water.

Table 2.8: Components for SDS-PAGE and gel staining

	Components
Separating gel	2.4ml Acrylamide/Bisacrylamide 37, 5:1 ; 1.5ml separation gel buffer (1.5M Tris-HCl pH 8.8, 0.4% SDS); 2.1ml ddH2O; 20µl 10 % APS; 4µl TEMED
Stacking gel	0.39ml Acrylamide/Bisacrylamide 37, 5:1; 0.75ml stacking buffer (0.5M Tris-HCl pH 6.8; 0.4% SDS); 1.85 ml ddH2O; 15µl 10% APS, 3µl TEMED
Laemmli buffer	1.2g SDS; 6mg Bromophenol blue; 4.7 ml Glycerol; 1.2ml 0.5M Tris pH6.8; 2.1ml ddH_2O; 0.93g DTT
Running buffer (1X)	3g Tris base; 14.4g glycine; 1g SDS; ddH_2O to 1000ml
Staining solution [63]	0.4g CBB G-250; 100g aluminium sulfate-(14-18)-hydrate; 200ml 96% ethanol; 47ml 85%orthophosphoric acid; ddH_2O to 2000ml

2.1.9 Software

(1) NCBI Blast (https://blast.ncbi.nlm.nih.gov/Blast.cgi)
(2) MEGA 7 (https://www.megasoftware.net)
(3) Schrödinger Release 2017-4: Maestro, Schrödinger, LLC, New York, NY, 2017.
(4) PyMol: The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC
(5) Jalview v2.10 (Waterhouse et al. 2009)
(6) Phenix-1.12-2829 (Adams et al, 2010)
(7) Coot v0.8.8 Turtle bay (ccp4) (Emsley et al., 2010)

2.2 Methods

2.2.1 DNA cloning protocols

Origin of plasmids EgPKAR1(109-401) and EgPKAR2(183-483) were de novo synthesized and inserted into a pUCIDT_AMP_vector by IDT (Integrated DNA Technologies, Inc., USA) (Fig. S1). DNA sequence was obtained from Dr. Thankgod Ebenezer from the *Euglena gracilis* genome project (http://euglenadb.org/). Originally, both PKAR subunits were identified as EG_transcript_11236 (corresponding to EgPKAR1) and EG_transcript_8206 (corresponding to EgPKAR2) in the transcriptome from *E. gracilis*. Both constructs were N-terminal truncated, corresponding to a Δ 1-91 deletion mutant of RIα, and codon optimized for expression in *E. coli* [26] (genes designed and optimized by Yuri

Volpato). Synthesized genes were cloned into pETM11_6His_Sumo3_eGFP plasmid (obtained from Kristina Malenica, originally purchased from EMBL, Protein Expression and Purification Core Facility), in which they were N-terminal fused to a His-tagged small ubiquitin related modifier (SUMO).
Truncations and mutants of EgPKAR1 and EgPKAR2 were amplified from pETM11_6His_Sumo3_EgPKAR1(109-401) and pETM11_6HIS_Sumo3_EgPKAR2(183-483) (Tab. 2.1). EgPKAR1(255-401) and EgPKAR2(327-483), for individual expression of binding site B, correspond to the truncation of RIα(G316R, A336T) (aa 234-381) [64]. EgPKAR1(109-255) and EgPKAR2(183-359), for individual expression of binding site A, incorporate binding site A and additionally the αA helix of binding site B. Both proteins were C-terminal truncated after the helix ending sequence motif, A-D-A-L. We thereby preserve a supposed inter-domain interaction between D190 in $\alpha A^{CNB:B}$ and R264 in $\alpha C^{CNB:A}$ seen for other PKAR subunits [65].
LdPKAR1(199-502) (LinJ.13.0160) and LdPKAR3(321-647) (LinJ.34.2680) were amplified from pET_Duet_LdPKAR and pET_Duet_6H/TEV_LdPKARlike (cloned by Eleni Polatoglou and Melanie Obermaier). Both genes were originally amplified from gDNA of the strain Lo8 (subclone from S1). Genomic DNA was obtained from Joachim Clos.
pET_Duet_6xHisTbR(199-499), for expression of TbPKAR(199-499) (Tb927.11.4610, originally amplified from gDNA), was generated by Yuri Volpato.

PCR and insert preparation For generation of protein truncations, DNA was amplified by polymerase chain reaction (PCR) from existing plasmids (Tab. 2.1, upper part). Primers were designed in CLC Main Workbench (QUIAGEN Bioinformatics, Germany) and purchased from Merck KGaA, Germany.

Table 2.9: Components of the PCR mastermix

Component	
5x Q5 PCR reaction buffer (NEB, Germany)	5 µl
10mM dNTPs (Roche, Germany)	0.5 µl
Template DNA	1ng
Forward Primer 2.2	10 µM
Reverse Primer 2.2	10 µM
Q5® High-Fidelity DNA Polymerase (NEB, Germany)	0.25 µl
Nuclease free water	up to 25 µl

All components of the PCR master mix were mixed at room temperature and amplification steps were carried out at the following temperatures:

Table 2.10: Temperatures and duration for each step of the PCR

Process	Temperature [°C]	Time [s]
Initial Denaturation	98	30
Elongation	98 72	10 35-40
Final extension Hold	72 12	120 unlimited

PCR products were digested with BamHI and Not andchecked to be of expected size by agarose gel electrophoresis. For purification, a commercial Gel/ PCR clean up kit was used (Tab. 2.4).

Plasmid generation PCR products were inserted into a pET11_6HIS_Sumo3_eGFP vector, digested with BamHI-HF and NotI-HF. Removal of the eGFP insert was achieved by agarose gel electrophoresis. Ligation of insert and linearized backbone occurred overnight, at 16°C, or for one hour, at room temperature.

Components (for 20µl volume)	volume [µl]
T4 DNA Ligase (NEB, USA)	1 (400 Units per µl)
T4 DNA Ligase buffer (10x) (NEB, USA)	2
Nuclease free water	to a final volume of 20µl
Backbone	50ng
Insert	$\frac{\text{linsert(kb)}}{\text{backbone(kb)}}$ x3x50ng

Transformation For transformation, XL10 competent cells (*E. coli*) were used (New England Biolabs (NEB)). Cells were incubated for 20 min on ice in presence of 10µl ligation preparation. Heat shock was applied at 42°C for 45 seconds, followed by another 2 minutes on ice. For recovery, transformed cells were incubated in 500µl LB at 37°C, shaking at 150 rounds per minute for 35-45 minutes. Recovered cells were plated on agar plates (containing 50µl/ml Kanamycin or 50µl/ml Ampicillin) and incubated overnight at 37°C. Positive clones were picked and incubated in 10ml LB + 10µl Kanamycin (Stock solution: c=50mg/ml) or 10µl Ampicillin (Stock solution: c = 100mg/ml). Cells were shaken overnight, at 37°C, 150 rpm. DNA was purified using a commercial kit according to the attached protocol (Tab. 2.4). Absence of point mutations or other sequence alterations is ensured by Sanger Sequencing carried out by the service unit "sequencing service" of the genetics department at LMU. DNA is stored at -20°C until used for transformation of *E. coli* Rosetta (DE3) for protein expression.

Cloning of EgPKAR1(255-401)_Mutants pET11_6HIS_Sumo3_EgPKAR1(255-401) and pUCIDT_AMP_EgPKAR1(109-401) were digested with BamHI and NotI, and the coding sequence of EgPKAR1(255-401) was placed into the opened pUCIDT_Amp vector (Tab. 2.1, Fig. S1). C-terminal amplifications, including aa 339-401, were generated from pUCIDT_AMP_EgPKAR1(255-401) with the respective mutations, S347A, C345A and the double mutant, S347A;C354A, inserted in the forward primer. As reverse primer, EgPKAR1(255-401)_RV was used (Tab. 2.2). PCR fragments were digested with ApaI and NotI and inserted into pUCIDT_Amp_EgPKAR1(255-401). Linearized plasmids still contained the N-terminus of EgPKAR1(255-401), so that the new PCR products replace the original C-terminus. Thereafter, all three plasmids, were digested with BamHI and NotI and inserts transferred into a pET_Sumo3 vector (Fig. S2). Positive clones were confirmed by sequencing.

2.2.2 Protein purification

Expression and protein recovery from cells For protein expression, *E. coli* strain Rosetta (DE3) was used, which contains an inducible system of T7 RNA polymerase expression. After induction with IPTG (Isopropyl-ß-D-thiogalactopyranosid), T7 RNA polymerase is transcribed. A T7 promoter placed in front of the open reading frame (ORF) in the pET_Duet and pET_Sumo vector is recognized by the polymerase, initiating transcription of the inserted gene. Cells that contain both the inducible T7 RNA polymerase expression system and the plasmids with our gene of interest were selected by antibiotics (Kanamycin or Ampicillin and Chloramphenicol). Chloramphenicol resistance is mediated by a strain specific plasmid. Kanamycin resistance is obtained with the pET_Sumo vector, Ampicillin resistance with the pET_Duet vector.
In order to optimize protein expression, different concentrations of IPTG and temperatures during expression were tested for EgPKAR1(109-401), EgPKAR2(183-483), LdPKAR1(200-502) and LdPKAR3(321-647). For all tested proteins, best expression conditions were: Induction with 0.1mM IPTG and expression over night at 16°C. Under these condition, approximately 50% of the expressed protein was soluble and could be purified in native state.
A preculture was grown from 250µl freshly transformed bacteria suspension or a single colony in 25ml LB (lysogeny broth) medium with 25µl Chloramphenicol and 25µl Kanamycin or Ampicillin (concentration of stock solution: Table 2.6) at 37°C, 150 rpm overnight. 15 ml of preculture are diluted in 800ml LB (+ 800µl Kanamycin or 800µl Ampicillin and 800µl Chloramphenicol). Cells are grown at 37°C, shaking with 150 rpm, until OD_{600} = 0.6–0.9. After induction with 0.1mM IPTG, cells are incubated at 16°C overnight, shaking with 150 rpm. Cells were harvested by centrifugation at 10.000g for 30 minutes.

Obtained cell pellets are resuspended in minimal buffer (Table 2.7). Cells were pelleted again at 3500g for 30 minutes and either stored at -20°C until further use or directly prepared for purification. All further purification steps were carried out on ice to minimize protein degradation.
Cell pellets were resuspended in resuspension buffer (Tab. 2.7), containing 500 µl of PMSF (stock solution: 57mM) (protease inhibitor) and 500 µl DNAse (stock solution: 17 µg/ml). For lysis of cell walls, resuspended cell pellets were passed twice over a french press at maximum pressure of 1500psi. Cell detritus was pelleted at 20000g for 30 minutes. Supernatant was sterile filtered and loaded onto a HisTrap^{Tm} High Performance 5ml (GE healthcare, United Kingdom) at a flow rate of 1ml/min. Remaining, non bound protein within the HisTrap column was removed by washing with resuspension buffer at a flow rate of 4ml/min for 2-3 minutes. The protein was eluted manually with 10ml elution buffer. Protein concentration was determined measuring UV-absorbance at 280nm with a nanoDrop photometer (Thermo Fisher Scientific.Inc, USA). Recombinant expression of each protein led to yields of 20-70mg of soluble, native protein.

Cleavage and removal of Sumo3 Fusion protein Proteins that were transcribed from a pET11_6HIS_Sumo3 vector were N-terminal fused to a SUMO3, which had to be removed before further experiments. Expressed protein were therefore cleaved by sentrin-specific protease 2 (SENP2). Expression and purification of the SENP2 protease was done by Miguel Guirao Ortiz. Aliquots of 2.7mg/ml were stored at –80°C until further use. Cleavage occurred overnight at 4°C, after SENP2 protease was added to the protein at a ratio of 1:500. During cleavage, protein was dialysed in 500-1000ml of minimal buffer to remove remaining imidazole from the elution buffer. SENP2 and SUMO3 were removed using immobilized metal ion chromatography. 2ml of PureCube 100 Ni-NTA Agarose (Cube Biotech, Germany) were washed with 10ml ddH_2O and again with 10ml minimal buffer. Dialyzed and cleaved protein was incubated for 30 minutes in the Ni-NTA agarose. HIS-tagged SENP2 and SUMO3 were captured by the Ni+ ions in the resin and the protein of interest is recollected in the flow through. The column was washed with 4ml of minimal buffer, followed by three fractions of 2ml minimal buffer that contained 20, 40 and 60mM imidazole, respectively. Each fraction was collected and analysed by SDS PAGE. Fractions of washing steps, that still contained protein of interest, were pooled with the flow through and processed in dependency of further use.
For Isothermal titration calorimetry (ITC), collected protein was concentrated or diluted to a final protein concentration of 1-4 mg/ml and stored at -80°C until further denaturing and refolding procedure, as shown in section 2.2.2. For crystallization, collected protein was

concentrated to 1ml and further purified by size exclusion chromatography on a Superdex 200 column.

Gel filtration chromatography For crystallization or isothermal titration calorimetry (ITC), gel filtration chromatography (SEC) was carried out as additional purification step. Gel filtration chromatography was performed on a Superdex 200 Increase 10/30 GL column inserted into an ÄKTA pure (GE healthcare Life sciences, Germany). Gel filtration was carried out at 4°C and an alarm for the pre-column pressure exceeding 2MPa was set. The protein solution was concentrated to 1ml using Amicon Ultra centrifugal filter units Ultra-15, MWCO 10 kDa (Sigma Aldrich, Germany) at 4°C. Protein aggregates were removed by centrifugation at 20000g for 5 min and the protein was loaded manually into a 1ml injection loop on the ÄKTA. Protein was injected at a flow rate of 0.1 to 0.25ml/min, followed by additional 2ml of the respective elution buffer to empty the injection loop completely. The protein sample was eluted at a flow rate of 0.5 to 0.6ml /min in either crystallization buffer for crystallization or in minimal buffer for ITC. Collection was performed automatically in fractions of 1ml or manually according to a peak in UV-absorbance at 280nm, measured by an internal UV detector. Purity of the protein sample was additionally checked by SDS PAGE.
Representative chromatograms and confirming SDS gels of all proteins used in ligand binding studies or for crystallization are shown in Figure S3.

Preparation of APO protein for isothermal titration calorimetry (ITC) For accuracy of ligand binding assays, a correctly folded, ligand-free (Apo) state of he protein had to be guaranteed. Proteins from *L. donovani* and *E. gracilis* were purified via HisTrap and separated from Sumo3 fusion protein as described in paragraph 2.2.2. TbPKAR(199-499) from *T. brucei*, which was fused to a 6HIS_Tag, but not to a Sumo3 protein, was not cleaved before ITC. The remaining preparation steps are equal for all protein.
Protein that was stored in minimal buffer at -80°C after cleavage or HisTrap purification was denatured in 8M urea. Denatured protein was separated from small molecules including metabolites coming from the *E. coli* expression (such as cAMP and nucleosides) by a PD-10 column (GE healthcare, United Kingdom). The protein was eluted in 3ml PD 10 buffer and could either be stored at –80°C for further usage or was directly prepared for refolding. Refolding occurred overnight at 4°C. Therefore, 3ml of sample were put into a dialysis bag and dialysed against 500ml buffer 15 (EgPKAR1, EgPKAR2, LdPKAR1, TbPKAR including truncations and mutants of the respective) or 500ml of buffer 10 (LdPKAR3) (Tab. 2.7). The protein was recovered from the dialysis bag after 12-14 hours of dialysis and centrifuged at 20.000g (5min, 4°C) to pellet aggregated protein. Supernatant

was further purified via SEC in order to guarantee purity of protein and remove remaining aggregates. SEC was performed as described in 2.2.2. Prepared protein had to be used the same day.
Preparation of LdPKAR1 did not require an additional gel filtration (SEC) step after refolding. Refolding occurred nearly to completion and only few aggregates were observed. Once overnight refolding took place against buffer 15, the only required step to prepare the protein for ITC was a buffer exchange. Therefore, the protein was loaded onto a PD10 column (GE healthcare, United Kingdom) and eluted in 3ml minimal buffer. Sample preparations were shown to be pure and of expected size by SDS PAGE.

2.2.3 Nano differential scanning fluorimetry (nDSF)

To monitor protein stability, a valid parameter is the melting Temperature T_m. T_m is defined as temperature, in which half of the protein is denatured while the other half remains folded. For our purposes, to evaluate buffer conditions and potential ligands, we used nano differential scanning fluorimetry (nDSF) on a Prometheus NT.48 (NanoTemper Technologies, Inc., München) that allows a label-free monitoring of T_m. We thereby gain information how certain compositions of protein solution influence protein stability, enabling us to screen for optimal buffer conditions, but also for binding of potential ligand candidates.
Approximately 10µl of protein samples were set into High Sensitivity glass capillaries (Prometheus NT.48 Capillaries nanoDSF Grade High Sensitivity, PR-C006, Nanotemper) and placed on a metallic plate that would heat up according to preset values.
In all our experiments, temperature was increased by 2°C every minute, reaching from 20 to 90°C. During the heating process, an internal detector measures light emission of two wavelengths, 330 and 350nM, mainly send out by tryptophan but also tyrosine residues in the protein. A ratio of both values (F350/F330) is constantly determined and drawn as function of temperature. Whether this ratio shifts towards the red (λ = 350nM) or blue spectrum (λ = 330nM) depends on the position of the tryptophan residues within the protein. In most cases, tryptophan residues are hidden inside the protein when secondary structure is preserved because, as hydrophobic amino acids, they avoid surrounding water molecules and are more likely to form hydrophobic interactions such as Van der Waals interactions or π stacking with other residues in the inside. Due to the loss of secondary structure in the context of temperature increase, more tryptophan residues are exposed at the surface, resulting in a shift of the emitted light towards the red spectrum. The opposite would be the case, if tryptophan residues are buried inside the protein upon temperature increase and a shift of light emission towards the blue spectrum could be observed.
In this experiment we obtain a melting curve and its first derivative is calculated by an in-

built software. T_m is defined as turning point of the sigmoid melting curve and accordingly as vertex of the first derivative. Hence, T_m of all inserted protein samples can be compared directly after the experiment, allowing us to rank buffer conditions or tested ligands according to their influence on protein stabilization.

Screen for ligand candidates Ligand screens were performed for EgPKAR1(109-401), EgPKAR2(183-483), LdPKAR3(321-647) and LdPKAR1(200-501). Protein samples were diluted to 300µg /ml and incubated with potential ligands at a concentration of 1 mM. Both protein and ligand candidate were diluted in the same buffer (minimal buffer). The following ligands, which we know interact with other cNMP independent PKAR subunits, were tested. Results are shown in Figure S6 and Table S2.

Table 2.11: Listed small molecules were tested in regard to ligand protein interactions. Ligand screens were performed for EgPKAR1(109-401), EgPKAR2(183-483), LdPKAR1(200-502) and LdPKAR3(321-647)

Protein	Ligand
EgPKAR1	cAMP; cGMP; Inosine; Guanosine; Adenosine; cAMP + cGMP; cAMP + Inosine
EgPKAR2	cAMP; cGMP; Inosine; Guanosine; Adenosine; cAMP + Inosine
LdPKAR3	cAMP; cGMP; Inosine; Guanosine; Adenosine; Uridine; Thymidine; cAMP + Inosine
LdPKAR1	cAMP; cGMP; Inosine; Adenosine; Guanosine

Screening and success control for protein refolding As accurately refolded protein is essential for ITC binding assays, refolding conditions were screened in 15 different buffers, compiled in a commercial Quickfold^{TM} protein refolding kit (Tab. 2.1.4). Proteins were denatured in 8M urea and diluted in PD10 buffer to a final concentration of 0.5-1.5 mg/ml. 5µl of protein solution were added to 495 µl buffer, which was prepared according to the instructions attached to the kit. All samples were incubated for one hour. Measurements were carried out on a Prometheus NT.48 (NanoTemper Technologies, Inc., Germany). Refolding of protein was successful in buffer 15 for EgPKAR1, EgPKAR2, LdPKAR1, and all their respective mutants or truncations, and buffer 10 for LdPKAR3.

2.2.4 Isothermal titration calorimetry (ITC)

Isothermal titration calorimetry (ITC) experiments were carried out on a Malvern MicroCal PEAQ-ITC (Malvern panalytical, Netherlands and UK). Binding assays with ITC not only proves ligand binding, but also offers information about binding affinity, molar ratio of ligand and protein, as well as changes in enthalpy (ΔH), entropy (–TΔS) and Gibbs Free energy (ΔG).

Small molecules, identified as potential ligand in nDSF experiments, were titrated in defined volumes into a reaction cell, containing the protein of interest. Heat changes occurring within the reaction cell, caused by ligand binding to the protein, were constantly detected by an integrated electrode and compared with a reference cell, filled with water. When a difference in temperature is detected between the reference and measuring cell, a heater readjusts the temperature of both cells and at the same time plots the differential power (DP) needed to perform the temperature adjustment. An integration of DP over time for all the injections matches the total heat released (ΔH) during ligand titration. From the binding curve, enthalpy ΔH and binding affinity and N-value can be derived directly. ΔH is shown as heat released per mole of bound ligand, resulting in a sigmoid binding curve. Binding affinity, expressed in the binding constant K_d, can be derived from the binding curve directly as it corresponds to the slope of this function. Molar ratio of ligand to protein reached at the inflection point of the isotherm is expressed as N-value and represents stoichiometry of ligand to protein under ideal conditions. Gibbs energy ΔG and entropy $-T\Delta S$ were derived indirectly with the following equations:

(1) $\Delta G = RT \ln[\frac{K_d}{c^{\Theta}}]$ (with R ideal gas constant, T Temperature and standard reference concentration $c^{\Theta} = 1 mol/L$)
(2) $\Delta G = \Delta H - T\Delta S$

Calculations were done by an internal software of the Malvern MicroCal PEAQ-ITC.
Proteins were purified and prepared as described above (s. paragraph 2.2.2) and purity was guaranteed by SDS PAGE (Supplements, Fig. S3). Protein was concentrated to 5-15 μM and diluted in minimal buffer for ligand binding assays.

For binding assays with Toyocamycin and 7_cyano_7_Deazainosine (7CN), Dimethylsulfoxide (DMSO) had to be added, as both nucleoside analogues could not be dissolved in minimal buffer. In order to minimize final buffer mismatches, the same concentration of DMSO in the ligand preparation was added to the protein. 50 mM Toyocamycin and 50mM 7CN, respectively, were dissolved in 100% DMSO. To obtain a working concentration of 100 μM, a 500-fold dilution of ligand was needed, meaning that 2 μl stock solution (in 100% DMSO) were added to 998 μl of minimal buffer. In the same manner , 2μl of 100% DMSO were added to 998 μl of protein to adjust final concentrations of DMSO in protein and ligand sample.
All experiments were carried out at 25°C, including either 19 injections of 2 μl or 13 injections of 4μl with 150s equilibration between two injections. Reference power was set to 8-10μcal/s. To distinguish the effect of ligand dilution and buffer mismatch from heat released during binding, ligand to buffer titration were used as control for all measurements.

Reaction cell was filled with 300µl protein solution, the syringe was loaded with 50 µl ligand sample. Between experiments, reaction cell and syringe of the MicroCal PEAQ-ITC were cleaned using the setting "rinse with water".

2.2.5 Protein crystallization

For protein crystallization, high amounts of pure protein were required. Proteins that were purified from *E. coli* were therefore concentrated after removal of the Sumo3 fusion protein and an a size exclusion chromatography was performed as described in paragraph 2.2.2. Afterwards, the protein was again concentrated to 12-20mg/ml with an Amicon Ultra centrifugal filter units Ultra-15, MWCO 10 kDa (Sigma Aldrich, Germany) and purity was confirmed by SDS PAGE (Supplements, Fig.S3).
The method used for initial buffer screenings was the sitting-drop vapour diffusion. For sitting drops screenings, equal volumes (100 nl) of concentrated protein (>10 mg/ml) and reagent solution were mixed and placed over a reservoir of reagent solution in the same chamber. Due to different concentrations of reagent in sample and reservoir solution, water will evaporate from the sample and a vapour equilibrium between both solutions will be formed. Thereby, saturation of protein within the drop increases. The slow removal of water eventually leads to a highly coordinated arranging of protein molecules forming a crystal in the end.
Reagents in the respective solutions depended on chosen in-house screens of the Max Planck institute of biochemistry, Munich, Crystallization Facility, where crystallization was carried out. The following buffer screening kits were used: MAGIC 1, MAGIC 2, IND, and NPG at two temperatures (4°C and 19°C). 0.1µl protein was mixed with 0.1µl screening solution and drops were placed on a 96 well plate using a nanoliter crystallization robot (Phoenix). For monitoring crystallization progress, drops were constantly imaged for 5 weeks and pictures provided online in real time by the facility. Conditions that lead to positive hits in the first crystal screening were then expanded using optimization grids, varying conditions such as pH or concentration of precipitation agents to improve crystal size and geometric arrangement.
Crystallization of EgPKAR1 (c=12mg/ml) was successful in presence of 1mM cAMP in different conditions and at different temperatures. Crystals grew within a week at 19°C in 25% PEG 3000 and 0.1M MES pH 6.5.
Crystals of LdPKAR3 (c=17 mg/ml), grew within 10 days at 19°C in 50mM MES pH 6.0, 14% PEG 8000 and 0.2M Sodium malonate. It is of note that the protein was shown to be co-crystallized with guanosine (as shown by its electron density) even though no ligand was added to the final solution.

Data collection and structure solving Crystal handling and diffraction data collection (both in-house and in the Swiss Light Source beam line PXII) were performed by Dr. Jerome Basquin, PhD (MPI, Biochemistry).
Phases for the co-crystal structure of EgPKAR1 with cAMP were obtained by molecular replacement using the structure of mammalian RIα bound to Rp-cAMPs (PDB: 1NE4) as a search model.
A first structure of LdPKAR3 was solved ab initio by the sulphur SAD method [66] (data not shown). A second structure at higher resolution was then solved by molecular replacement obtaining phases from the first structure of LdPKAR3.

Structure refinement For structure refinement, the software refine phenix-1.12-2829 was used. Model building was done manually in coot-ccp4 - 7.0 and for all structural figures MacPyMOL, Version 1.5, was used (2.1.9). Quality of structure refinement was mainly evaluated by R and R_{free} (Tab 3.3, Sektion 3.4). Constraints due to stereo chemistry were controlled by computation of Ramachandran plot.

Molecular docking For Molecular docking, the program Glide as part of Maestro, Version 2017-4 and provided by Schrödinger GmBH was used [67] [68]. Molecular docking was performed on proteins listed in Table 2.12.
For comparison to ITC data obtained for EgPKAR1(255-401)_S347A, EgPKAR1(255-401)_C354A and EgPKAR1(255-401)_S347A,C354A, molecular docking to EgPKAR2(255-401) was performed. As template the crystal structure of EgPKAR1(109-401) in complex with cAMP (PDB: 6RSX) was used and in silico truncated so that it corresponds to the truncation of EgPKAR1(255-401). Mutations were inserted after energy minimization for each mutant individually.
For docking to LdPKAR3, the crystal structure of LdPKAR3 in complex with guanosine (PDB: 6TFC) was used as a template.
All proteins were prepared using the implemented software Protein Preparation wizard and energetically minimized using an OPLS3 field. Ligands were extracted from the minimized structure and prepared before docking using the LigPrep Tool. Ligands that were not present in the crystal structure were generated using the 3D builder tool before ligand preparation. Receptor grids were built for each mutant individually with the centre of the grid matching the position of the cognate ligand. Constraints, hence, obligatory formed interactions between docked ligand and protein, were chosen as shown in Tab. 2.12

Table 2.12: Receptor grids for molecular docking were defined using listed amino acids as constrains

Protein	Ligand	Site	Constraints	Must match
EgPKAR1(255-401) EgPKAR1(255-401)_S347A EgPKAR1(255-401)_C354A EgPKAR1(255-401)_S347A_C354A	cAMP	B	N335, G344, S347 HOH:46, V355,E345	3
LdPKAR3(321-647)	Guanosine	B	E560 OE1, E560 OE2, A569, I559	2
LdPKAR3(321-647)	Inosine	B	E560 OE1, E560 OE2, A569, I559	2
LdPKAR3(321-647)	Adenosine	B	E560 OE1, E560 OE2, A569, I559	2

Docking strategy demanded at least half of chosen constraints to be fulfilled for all docking experiments. Docked poses were incorporated in the project and values for Emodel and Glide Gscore, which correlates with ligand binding free energy, were calculated. Best poses were selected according to lowest Emodel and Glide Gscore was used to compare best poses of different ligands.

Chapter 3

Results

3.1 PKAR isoform genes in the phylum Euglenozoa – origin of diversity in ligand specificity

Protein sequence analysis indicate the presence of cyclic nucleotide independent PKA in various lineages of the Euglenozoa Recent findings from Bachmaier et al. 2019 [45], and unpublished data from our group show that PKAR isoforms from *T. brucei* and *T. cruzi* are unable to bind cyclic nucleotides. Instead, nucleosides and their analogues bind and activate the kinase [45] [George Githure, PhD Thesis]. Switched ligand specificity has been shown to correlate with sequence deviations from the canonical ligand binding motif in the PBCs, that was defined by Canaves and Taylor, 2002 (Fig.3.1) [24]. In PKA from *T. cruzi* (PDB: 6FTF) and *T. brucei* (PDB: 6FLO), a glutamate in position 5 of the PBC provokes a steric clash with cAMP but interacts with the ribose ring of nucleosides [45]. Also, an arginine in position 12, which binds to the negatively charged oxygen in the phosphate moiety of cAMP, is lost in both orthologs. Thus, in *T. brucei* and *T. cruzi*, a phosphate binding cassette (PBC) has changed into a ribose binding cassette (RBC).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Cyclic nucleotide binding motif	F	G	**E**	LIV	**A**	L	LV IM	X	X	X	P/V	**R**	AQ NV	A
Nucleoside binding motif	X	G	**E**	L	**E**	X	X	X	X	X	X	**X**	X	X

Figure 3.1: Consensus motifs of cyclic nucleotide and nucleoside binding PKA isoforms Consensus motif of cyclic nucleotide binding PBC according to Canaves and Taylor, 2002 (upper part) [24]. Residues that confer selectivity for nucleosides are shown beneath and coloured in blue. Corresponding residues in the cyclic nucleotide binding motif are coloured in red.

As cAMP independency was biochemically verified for two representatives of Kinetoplastida [69] [45] [George Githure, Yuri Volpato, manuscript in preparation], we wondered

whether cAMP independent PKA can also be found in other Kinetoplastida or even outside Kinetoplastida, in more distant protists. We therefore analysed various sequences of PKAR in representatives of the Euglenozoa for deviations from the canonical PBC motif, with particular attention to the mutations found in PKAR from trypanosomes (Fig. 3.1). We found the sequence that deviates from mammalian consensus to be distributed among different lineages of the Euglenozoa. Other putative binding sequences, which differ from both sequences introduced in Fig. 3.1, were also found [24] [45] [George Githure, PhD Thesis]. Based on these survey results, we chose previously uncharacterised PKAR isoforms from *E. gracilis*, EgPKAR1 and 2, and *L. donovani*, LdPKAR1 and LdPKAR3, for biochemical characterization in order to verify ligand specificity and potential ligand candidates.

Binding assays with ligand candidates of PKAR like genes in pathogenic and free-living protozoa Initially, preliminary ligand screens using nano Differentiation Scanning fluorimetry (nDSF) were performed in order to test the binding of candidate ligands for each PKA ortholog (Fig. S6, Table S2).
Isothermal titration calorimetry (ITC) was applied to confirm binding of ligands, which presented a positive thermal shift in the nanoDSF experiment, and to determine the specific binding properties. For accurate determination of ligand binding parameters, an APO, i.e. ligand-free, state of the protein was required. Therefore, the protein was denatured after expression in and purification from *E. coli*, separated from internal bound small molecules and subsequently refolded as described in the method section 2.2.2. Refolded protein was compared to native protein using nDSF. Accurate refolding was concluded when melting temperatures of refolded protein corresponded to those of native protein in presence of potential ligands (Fig. S4). Figure 3.2 shows representative binding data that was derived from titration of the identified ligand candidates to the respective PKA isoform or individual domains there of in ITC. Binding affinity, expressed as dissociation constant K_d, N-value, which approaches stoichiometry of ligand to protein under optimized conditions (accurate determination of protein and ligand concentration; 100% activeness of protein sample), and thermodynamic parameters are calculated as average of three independent measurements and listed in Table 3.1.

Here we show that three out of four tested PKAR subunits bind nucleosides (Fig. 3.2) but no cAMP (Fig.S7).
LdPKAR1 from *L. donovani*, resembles the former characterized nucleoside dependent PKAR isoforms from *T. brucei* and *T. cruzi*. In correlation with specific thermal shifts of each nucleoside (Fig. S6), LdPKAR1 shows the highest affinity for inosine (K_d = 59nM) followed by guanosine(K_d = 173nM) and adenosine (K_d = 1487nM) (Fig. 3.2, G; Tab.

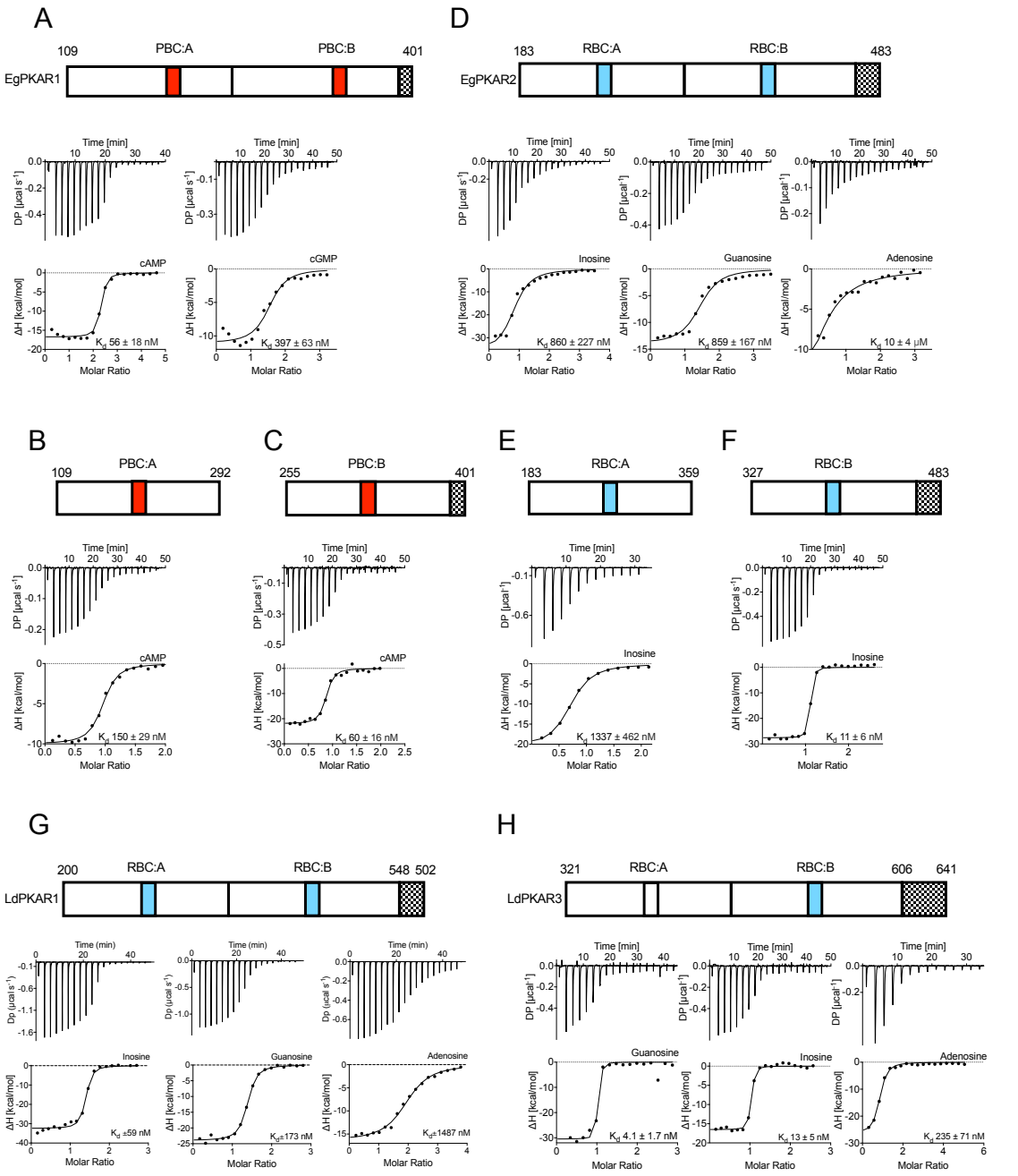


Figure 3.2: Protein-ligand interactions of different PKAR subunits analysed by ITC for (A) EgPKAR1(109-401) and its individual domains (B,C) as well as (D) EgPKAR2(183-483) and its individual domains (E,F); (G) LdPKAR3(321-647) and (H) LdPKAR1(200-502) were obtained from Isothermal titration calorimetry (ITC) experiments. In (A-H) differential power (DP) measured between sample and reference cell is shown as a function of time in the upper plots. Isotherms beneath illustrate the total amount of heat released (corresponding to integrated peak areas in the plots above) per mole of ligand. Cartoons above the binding plots illustrate schematically the respective protein truncation matching the colour code introduced in Figure 3.1. Three ITC measurements were done from at least two independent protein preparations. For separated domains of EgPKAR1 and EgPKAR2 only one protein purification was analysed and binding properties were calculated from a technical triplicate.

Table 3.1: Ligand binding parameters for different PKAR isoforms. Three ITC measurements were done from at least two independent protein preparations. For separated domains of EgPKAR1 and EgPKAR2 only one protein purification was analysed and binding properties were calculated from a technical triplicate

Purified protein	Ligand	$K_d \pm$ SD [nM]	ΔH [kcal/mol]	ΔG [kcal/mol]	$-T\Delta S$ [kcal/mol]	N-value
EgPKAR1(109-401)	cAMP	56 ± 18	-16 ± 1.1	-9.9 ± 0.2	5.8 ± 0.9	2.0 ± 0.3
	cGMP	397 ± 63	-10 ± 0.9	$-8,7 \pm 0.1$	1.6 ± 1.0	1.4 ± 0.1
	Inosine	No binding				
EgPKAR1(109-292)	cAMP	150 ± 29	-10 ± 0.6	-9.3 ± 0.1	1.1 ± 0.7	0.9 ± 0.1
EgPKAR1(255-401)	cAMP	60 ± 16	-22 ± 1.4	-9.9 ± 0.2	12 ± 1.2	0.8 ± 0.1
EgPKAR1(255-401)_S347A	cAMP	157 ± 52	-22 ± 1.3	-9.3 ± 0.2	12.7 ± 1.4	0.6 ± 0
EgPKAR1(255-401)_C354A	cAMP	274 ± 42	-21.0 ± 0.6	-9.0 ± 0.1	11.7 ± 0.7	1.0 ± 0.1
EgPKAR1(255-401)_S347A_C354A	cAMP	405 ± 85	-20 ± 0.8	-8.7 ± 0.1	11.0 ± 0.9	1.1 ± 0.2
EgPKAR2(183-483)	Inosine	860 ± 227	-33 ± 3.7	-8.3 ± 0.2	25 ± 3.8	0.8 ± 0.1
	Guanosine	859 ± 167	-24 ± 8.6	-8.3 ± 0.1	15.3 ± 8.6	1.1 ± 0.3
	Adenosine*	10000 ± 4000	-20 ± 4.1	-6.9 ± 0.2	14 ± 4.0	0.5 ± 0.2
	cAMP	No binding				
EgPKAR2(183-361)	Inosine	1337 ± 463	$-22 \pm 2,3$	-8.0 ± 0.2	$14.3 \pm 2,5$	0.6 ± 0.1
EgPKAR2(327-483)	Inosine	11 ± 6	-28.0 ± 0.2	-10.9 ± 0.3	17 ± 0.3	1.2 ± 0.1
LdPKAR3(321-647)	Inosine	13 ± 5.3	-21 ± 4.2	-11 ± 0.2	11 ± 4.0	0.84 ± 0.1
	Guanosine	4.1 ± 1.7	-28 ± 2.5	-11.5 ± 0.3	16 ± 2.7	1.0 ± 0.1
	Adenosine	235 ± 71	-26.6 ± 0.1	-9.1 ± 0.2	18 ± 0.1	0.7 ± 0.1
	cAMP	No binding				
LdPKAR1(200-502)	Inosine	59 ± 17	-29 ± 7	-9.9 ± 0.2	19 ± 7	1.6 ± 0.4
	Guanosine	173 ± 58	-22 ± 4.3	-9.3 ± 0.2	13.1 ± 4.1	1.5 ± 0.3
	Adenosine	$1200 \pm 300\mu$M	-16 ± 2.6	-8.1 ± 0.1	8.2 ± 2.6	2 ± 0.1
	cAMP	No binding				

$*$ Data from two independent runs only

3.1). With a ligand to molar ratio of N $\geq$ 1.5, the calculated stoichiometry is compatible with the expected number of 2 binding sites. LdPKAR3, a second PKAR isoform from *L. donovani*, differs from LdPKAR1 in that a slightly higher affinity for guanosine (K_d = 4.1nM) than inosine (K_d = 13nM) was detected in ITC and the molar ratio of ligand to protein approximates 1 for all tested ligands. Adenosine has the lowest affinity with K_d = 235nM (Fig. 3.2,G). Preference for guanosine over inosine and adenosine was already indicated in preliminary ligand binding screens using nDSF (Fig. S6, Tab. S2). Although measured stoichiometry of ligand to protein with N $\approx$ 1 differs from expected number of binding sites, ligand binding to only one site and preference for guanosine was also suggested by the crystal structure of LdPKAR3 (Chapter 3.4). Thus, cAMP independency of PKA is also found in other Kinetoplastida.

In *E. gracilis*, we find a nucleoside and cyclic nucleotide specific isoform of PKA to coexist in one organism. EgPKAR2(183-483) binds best to inosine and guanosine (K_d = 860nM), and has very low affinity for adenosine (K_d = 11μM) (Fig. 3.2, D). Furthermore, the measured stoichiometry does not resemble the expected number of binding sites (N $\approx$ 1 for inosine and guanosine). EgPKAR1, on the contrary, displays high affinity for cAMP (K_d = 56nM) with 7-fold lower affinity for cGMP (K_d = 397nM), and binds both ligands in two binding sites (N=2.0 and N=1.4, for cAMP and cGMP, respectively; Fig. 3.2, A). EgPKAR1 resembles a classic cAMP dependent PKAR with slightly lower binding affinity for cAMP compared to RIα of *B. taurus* (K_d = 28 $\pm$ 5nM) [45]). Concluding, cAMP inde-

pendency of PKA is not restricted to Kinetoplastida, but extended to their more distant relatives (Fig. 3.3).

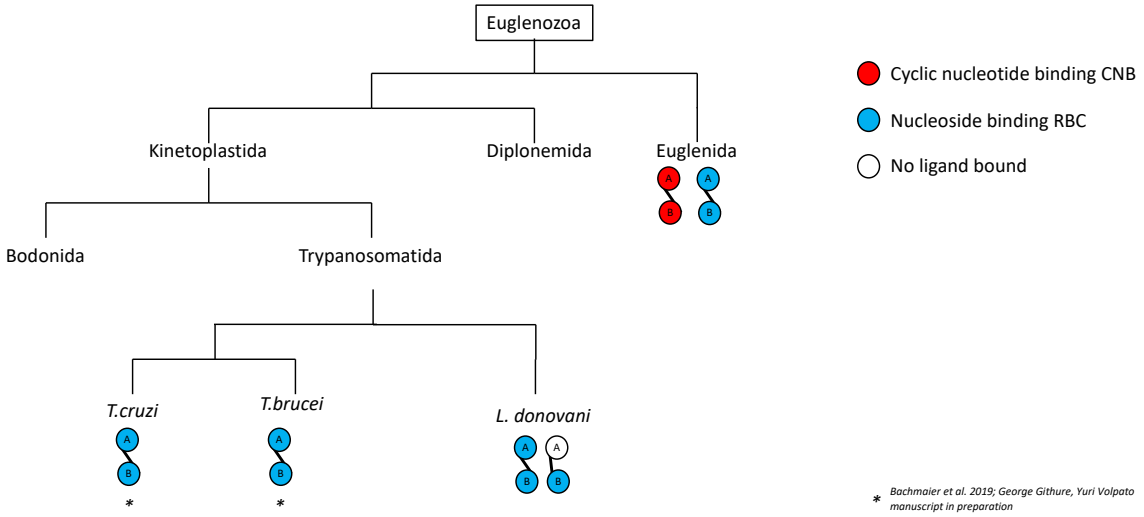


Figure 3.3: Ligand binding specificities of investigated PKAR orthologs from representatives of the Euglenozoa. Ligand specificity is indicated according to the colour code on the right. PKAR isoforms marked with an $*$ were not investigated in this but other studies from our lab [45], [George Githure and Yuri Volpato, manuscript in preparation].

3.2 The nucleoside specific EgPKAR2 isoform

The structure of EgPKAR2 was homology modelled to investigate whether the binding mode of nucleosides is conserved among all nucleoside dependent PKA isoforms. Structural modelling was done by Okke Melse as part of a cooperation with the group of Prof. Dr. Iris Antes and will be shown for comparison in this thesis.

The model was built using the crystal structures of PKAR from *T. brucei* (PDB: 6FLO) and *T. cruzi* (PDB: 6FTF) as templates. Inosine was modelled as ligand into both binding pockets. For refinement, molecular dynamics (MD) simulations were performed on the inosine bound model of EgPKAR2. In these, two different states, an open and a closed state, of the protein were observed. The closed state corresponds to the B-state, i. e. ligand bound state, of PKAR in mammals and kinetoplastids [26] [45] [George Githure, Yuri Volpato, manuscript in preparation], whereas the open state corresponds more to the H-state that was observed for mammalian PKAR in complex with a catalytic subunit [21] [70] [71]. Since the open state was the more stable of both during MD simulations, it was chosen as refined homology model of EgPKAR2.

Overall organization of EgPKAR2 is compatible with *T. brucei* and *T. cruzi* (Fig. 3.4). It

consists of two tandem binding domains, site A and B, which are connected over an elongated αB/C helix. Both sites harbour inosine as potential ligand candidate. Interestingly, the conformational state of EgPKAR2 differs from TbPKAR and TcPKAR. Although inosine was modelled into both domains, EgPKAR2 more closely resembles ligand-free mammalian PKARIα than ligand-bound TbPKAR or TcPKAR [21] [45]. However, the opened conformation of EgPKAR2 explains well the ligand binding behaviour that we observed for EgPKAR2(183-483) (Fig. 3.2 D, E, F; Tab.3.1). Therefore, the structure of both binding sites will be discussed in detail in the following.

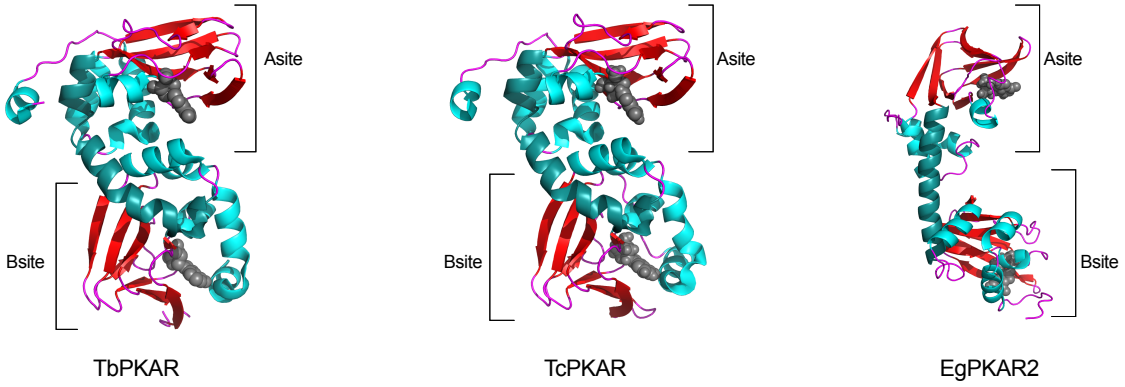


Figure 3.4: Overall organization of TbPKAR from *T. brucei* (PDB: 6FLO), TcPKAR from *T. cruzi* (PDB: 6HYI) and EgPKAR2. Helices are coloured in cyan, loops in purple and β-sheets in red. αB/C helix that connects the two binding sites is coloured in dark cyan. Inosine bound to the binding sites is shown as grey spheres. The model of EgPKAR2 was done by Okke Melse, Group of Prof. Dr. Iris Antes.

Binding site A of EgPKAR2 Inosine in binding site A is not completely buried inside the protein but rather solvent exposed (Fig. 3.5, centre). Similar to ligand binding in PKAR from *T. brucei* and *T. cruzi*, two glutamates, E292 and E294, are the main interactors with the ribose moiety. However, an important part of ligand binding seems to be missing as aromatic capping of the purine base does not occur according to the model. Fig. 3.6 shows an alignment of PKARIα from *B. taurus* (PDB: 1RGS), TbPKAR from *T. brucei* (PDB: 6FLO) and EgPKAR2 on binding site A (aa 113-262 of BtPKARIα). An aromatic capping residue, as it is located in the αA helix$^{CNB:B}$ in the structures of *B. taurus* (green) and *T. brucei* (magenta), is missing in EgPKAR2 (grey). The corresponding residue from sequence alignment, D352 (grey), is not placed in direct proximity to the ligand, but at a greater distance from site A due to a straightened αB/C helix.
These findings correlate with a low binding affinity for inosine to EgPKAR2(183-361), a truncation of EgPKAR2, which includes only the binding site A. K_d for inosine was approximately 1.3μM (Fig. 3.2,E), which is compatible with an overall low ligand binding affinity for

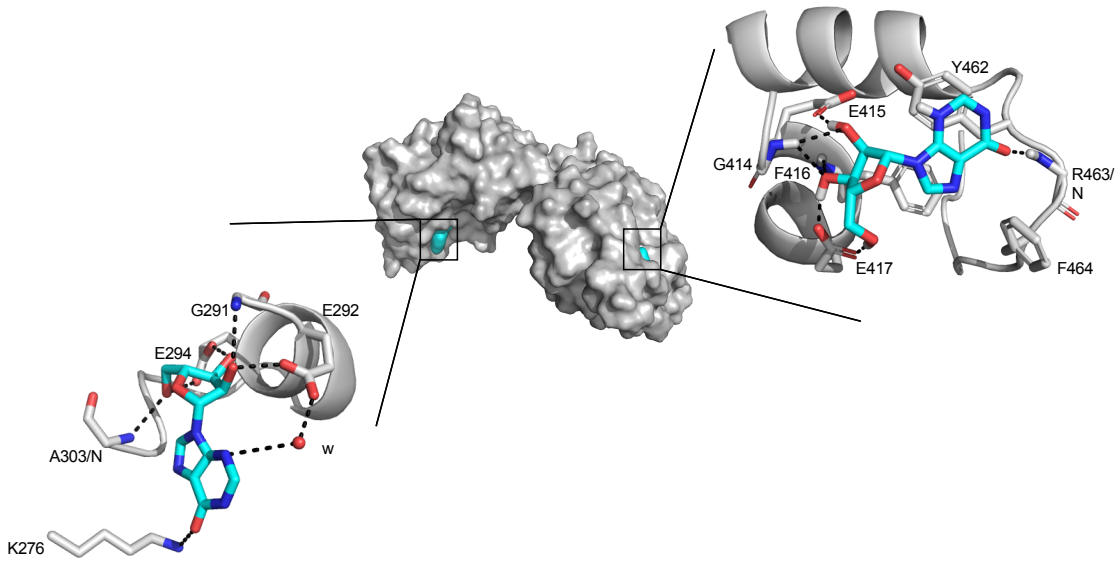


Figure 3.5: Surface representation of EgPKAR2 (grey) and inosine (cyan). Residues that contribute to a hydrogen network around inosine (cyan) in site A (left) and B (right) of EgPKAR2 are shown as sticks. Hydrogen bonds are shown as black dashed lines.

nucleosides of EgPKAR2(183-483) (K_d=860nM for inosine and guanosine, respectively, Fig. 3.2,D, Tab. 3.1).

Binding site B of EgPKAR2 In binding site B, we find an extensive hydrogen network and aromatic capping of the ligand that is reflected in high binding affinity to the individual domain, EgPKAR2(327-483) (K_d=11nM; Fig. 3.2, F).
Similar to binding site A, the hydrogen network is mainly formed by two glutamates, E415 and E417, and completed by backbone interactions (Fig. 3.5, right). Main difference, however, is a capping motif (Y462 and F464) that is localized at the C-terminal end of the αC helix. In PKAR from *T. brucei* and *T. cruzi*, the capping motif is at the beginning of an additional αD helix [45] [George Githure and Yuri Volpato, manuscript in preparatio]. An additional structural motif that corresponds to the αD helix is also present in EgPKAR2, although it does not form a helix in the model (Fig. 3.7, left). Nevertheless, Y462 may directly interact with the purine ring via p-shaped π-stacking and F464 engages in a T-shaped π-stacking with the purine base (Fig. 3.7, right).

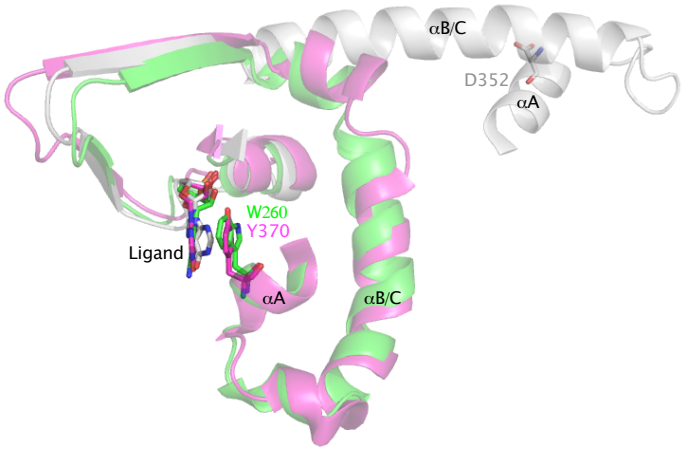


Figure 3.6: Alignment of PKAR subunits from *T. brucei* (PDB: 6FLO, magenta) and *E. gracilis*, EgPKAR2, (grey) to aa 113-262 of PKARIα from *B. taurus* (PDB:1RGS, green). Ligands in binding site A are shown as sticks in the respective colour of the macromolecule. The capping residue of binding site A in PKAR from *T. brucei* (Y370) and *B. taurus* (W260), and the aligning residue in the sequence of EgPKAR2 (D351) are shown as sticks.

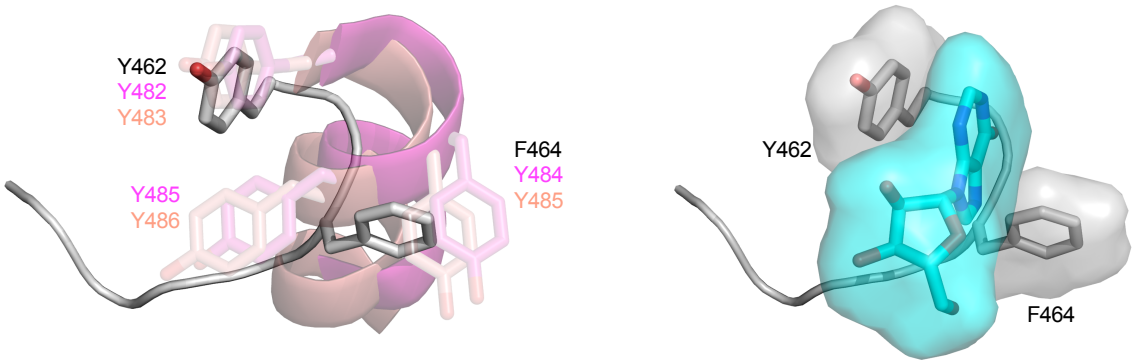


Figure 3.7: αD helices of TbPKAR from *T. brucei* (magenta) and TcPKAR from *T. cruzi* (rose) are compared to the C-terminus of EgPKAR2 (grey) (aa 462-469) in an alignment of TbPKAR and TcPKAR to the B site of EgPKAR2 (aa 349-470) (left). Residues contributing to the capping motif of binding site B are shown as sticks. On the right, the C-terminal capping motif in EgPKAR2 is shown. Surface representation was chosen for the bound ligand (inosine, cyan), the capping residue Y462 and F464.

Thermodynamic signatures of EgPKAR2(183-361) and EgPKAR2(327-483) Differences in binding site A and B are reflected in their thermodynamic signature (Fig. 3.8, Tab. 3.1). In both truncations, ligand binding is enthalpically driven ($\Delta H < 0$), counteracting an unfavourable change of entropy ($T\Delta S > 0$). The change in entropy is higher for EgPKAR2(327-483) ($-T\Delta S$ = 17kcal/mol), site B, than EgPKAR2(183-361), site A ($-T\Delta S$ = 14kcal/mol). Intensive hydrogen bonding in EgPKAR2(327-483), site B, due to contributions of the C-terminal structure motif cause a strong change of enthalpy (ΔH^{Bsite} =

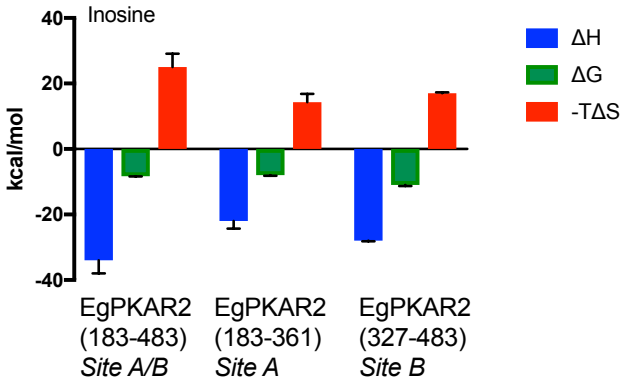


Figure 3.8: Thermodynamic signatures of inosine binding to EgPKAR2(183-483) (left), EgPKAR2(183-361) (centre) and EgPKAR2(327-483) (right), determined by ITC. Bar plots illustrate the thermodynamic signature composed of changes in enthalpy ΔH (blue), entropy –TΔS (red) and Gibbs free energy ΔG (green) in kcal/mol.

–28kcal/mol, ΔH^{Asite} = –22kcal/mol) that exceeds the negative entropic effect. This results in a higher release of Gibbs Free energy ΔG in site B than site A.

3.3 Characterization of EgPKAR1 in comparison with mammalian PKAR1α

EgPKAR1 is a second PKAR subunit encoded in the genome of *E. gracilis*. We could show that unlike EgPKAR2, EgPKAR1 binds cAMP with high affinity and in both binding sites (N=2, Tab. 3.1). While CNB:A of EgPKAR1 resembles a classic cyclic nucleotide binding domain, CNB:B exhibits an unusual cyclic nucleotide binding motif that deviates from the canonical PBC motif defined by Canaves et al. 2002 [24]. To investigate how cAMP can bind to the unusual pocket in CNB:B, crystallization of EgPKAR1 in complex with cAMP was targeted. Due to successful crystallization of N-terminal truncated PKAR subunits, such as mammalian PKARIα(91-379) (PDB:1RGS) [26] and yeast Bc1(168-416) (PDB: 3OF1) [65], we also chose to use an N-terminal truncated protein for crystallization, that we had analysed for cAMP binding already (Fig. 3.2, A-C).

3.3.1 Structural determinants of ligand specificity in EgPKAR1

EgPKAR1(109-401) was expressed in and purified from *E. coli* in native, i.e. non-denatured state as described in section 2.2.2. Purification by size exclusion chromatography led to a symmetric peak that indicated a mono disperse sample and purity of the protein was confirmed by SDS PAGE (Fig. 3.9). Crystallization was carried out at Max Planck institute of

biochemistry, Munich, Crystallization Facility, as described in section 2.2.5. A representative image of the protein crystal that led to the structure is shown in Figure 3.9. Collection of X-ray diffraction data was done by Dr. Jerome Basquin, PhD, from MPI, Crystallization Facility. The structure was solved to a resolution of 1.6 Å. Initial phases were obtained by

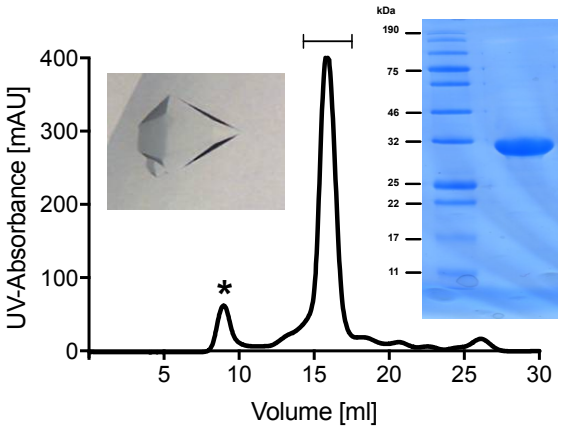


Figure 3.9: Native protein was purified from *E. coli* for crystallization. The line above the higher peak marks the collected fraction that was shown to be pure and of expected size by SDS PAGE (right). The peak marked with an $*$ is caused by aggregated protein and was not collected. A picture of a representative protein crystal of EgPKAR1 is shown on the left.

molecular replacement and multiple cycles of refinement were performed (chapter 2.2.5). Data collection and refinement statistics are provided in Section 3.4, Table 3.3.
The overall structure of EgPKAR1 resembles the structure of mammalian RIα (91-379) and aligns with a root mean square deviation (rmsd) of 1.52 [26] (Fig. 3.10). It is composed of two tandem cAMP binding domains (CNB:A and B) that are connected over an αB/C helix and a cAMP molecule is bound in each domain. Both binding domains consist of an αA helix, followed by a β barrel with a highly conserved phosphate binding cassette (PBC), and two further helices, αB and αC. In CNB:A, αB and αC, are connected, to form a bridge between the two CNB domains.

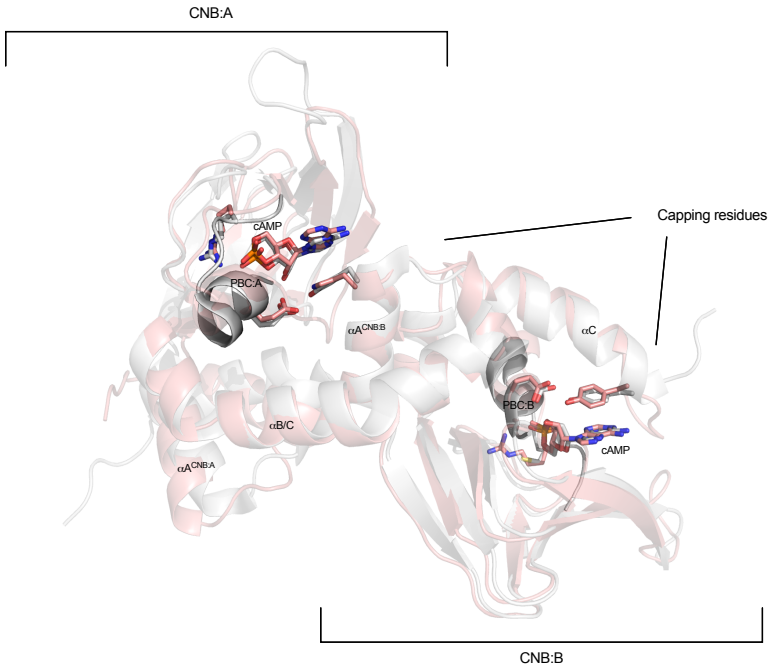


Figure 3.10: Overall structure of EgPKAR1 (grey) aligned to RIα (rose, PDB: 1RGS) is shown. cAMP binding to PBC:A and B of both structures is highlighted with ligand and key residues shown as sticks.

Canonical PBC in binding site A of EgPKAR1 With its overall structure mainly conserved, uniqueness of this protein lays in its binding to cAMP, particularly in the unconventional PBC:B. PBC:A, on the other hand, is a good example for a canonical PBC (Fig. 3.11).

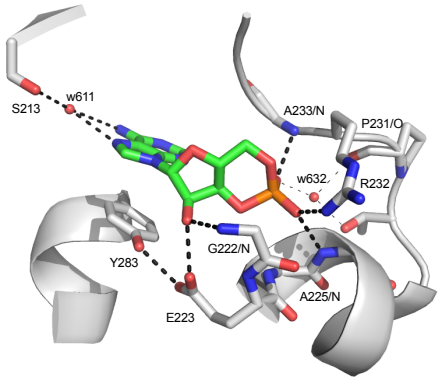


Figure 3.11: Ball and stick model of cAMP (green) binding to CNB:A of EgPKAR1 (grey). Residues, which interact with the ligand are shown as sticks, hydrogen bonds are indicated as black dashed lines.

Here, we find cAMP to be bound in syn conformation with the phosphate moiety pointed towards the PBC and the purine ring towards the solvent surrounding. Similar to mam-

malian RIα, the purine ring is capped by an aromatic residue (Y283). Phosphate moiety and ribose are bound over various hydrogen bonds. Particularly noteworthy are the conserved residues, E223 and R232. The former binds the ribose 2'-OH, the latter the phosphate moiety of cAMP. Further backbone interactions involve A225, G222 and A233. Particular importance was attached to the conserved arginine in the PBC, as its absence would significantly lower the affinity for cAMP or even prevent ligand binding [23] [72] [27] [24].

cAMP binds to CNB:B of EgPKAR1 in a new ligand binding motif cAMP binding to PBC:B, which lacks the conserved arginine described above, is highly interesting. Figure 3.12, (centre), shows the hydrogen bond network (black dashed lines) around cAMP in CNB:B of EgPKAR1. For structural comparison, CNB:B of RIα (PDB: 1RGS) is shown on the left. Superposition of both highlights shared hydrogen bonds (black dashed lines) and those only found in CNB:B of EgPKAR1 (purple dashed lines).

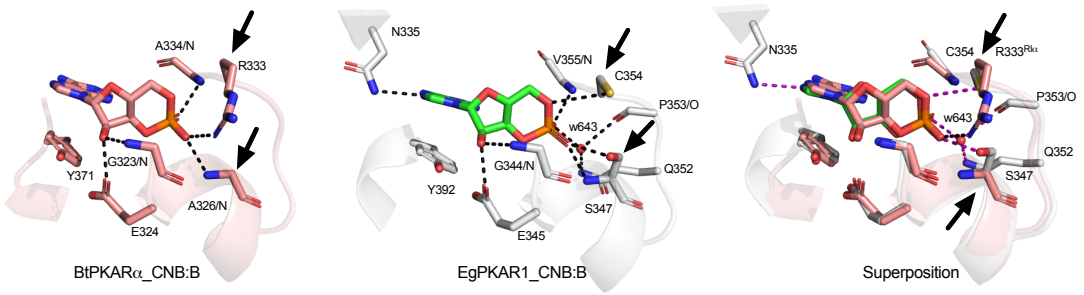


Figure 3.12: On the left, PBC:B of mammalian RIα (amino acids and cAMP in rose) and in the centre, PBC:B of EgPKAR1 (amino acids in grey, cAMP in green) are shown. On the right, PBC:B of mammalian RIα and EgPKAR1 are superimposed. Hydrogen bonds, present in PBC:B of EgPKAR1, but not in RIα, are coloured in magenta instead of black. Black arrows mark residues, critical for binding to RIα, that have been substituted in EgPKAR1.

In both structures, cAMP binds in syn conformation and similar interactions to purine ring and ribose moiety are formed: The purine ring is bound via π-stacking to Y392 (Y371 in RIα) with an additional interaction, absent in RIα, between N335 and N7 of the purine ring. Hydrogen bonds to the ribose moiety are formed with E345 (E324 in RIα) and G344/N (G323/N in RIα). Main difference lies in the binding of the phosphate moiety (Fig. 3.13). In our structure, we find a key residue of mammalian RIα PBC:B, namely R333, to be replaced by a cysteine (C354). In EgPKAR1_CNB:B, the phosphate moiety of cAMP is captured between C354 and S347, which replaces A326 in RIα. Additional hydrogen bonds are formed with the backbone nitrogen of V355 and a highly coordinated water (Fig. 3.13).

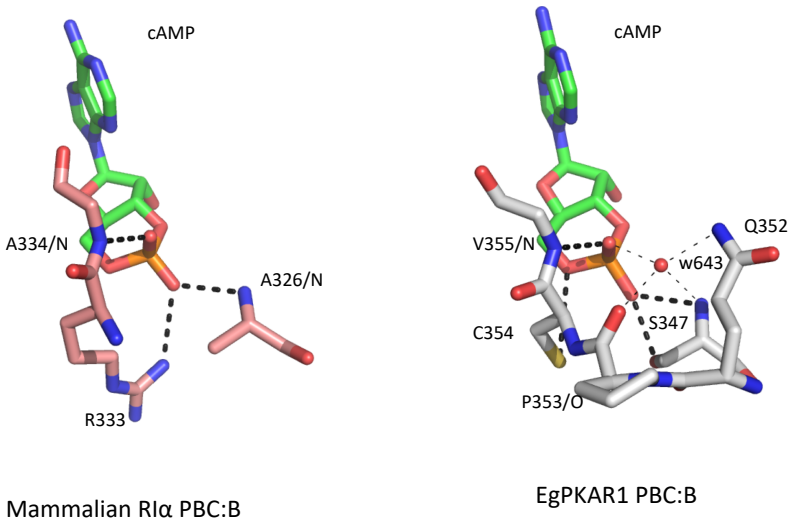


Figure 3.13: Comparison of PBC:B in mammalian RIα and EgPKAR1. Residues localized in the PBC:B of mammalian RIα (left, rose) and EgPKAR1 (right, grey) that interact with the phosphate moiety of cAMP (green) are shown as sticks. Hydrogen bonds are indicated as black dashed lines.

Cyclic nucleotide binding affinities in CNB:A and B To compare binding affinity for cAMP in the canonical and divergent PBC of EgPKAR1, truncations of EgPKAR1, EgPKAR1(109-292) and EgPKAR1(255-401) were generated for expression of individual domains. Ligand binding was investigated for both constructs by ITC.
Obtained K_d values for EgPKAR1(109-292), CNB:A, and EgPKAR1(255-401), CNB:B, are 150nM and 60nM, respectively (Fig. 3.2, B and C; Table 3.1). Thermodynamic signatures of cAMP binding to EgPKAR1(109-292) and EgPKAR1(255-401) reveal an enthalpically driven binding process that has to overcome an unfavourable entropy in both sites (Fig. 3.14, left). The enthalpy of cAMP binding to EgPKAR1(109-292), CNB:A, is reduced

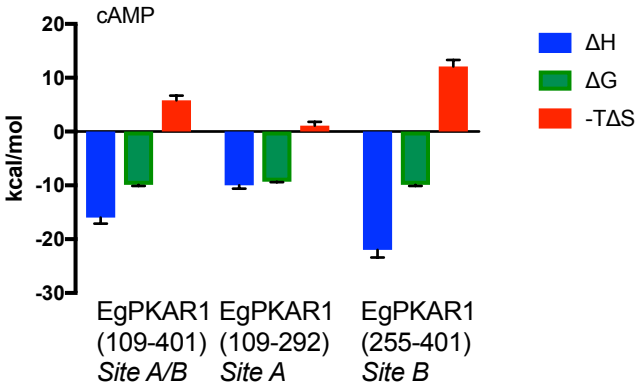


Figure 3.14: Thermodynamic signatures of cAMP binding to EgPKAR1(109-401) (left), EgPKAR1(109-292) (centre) and EgPKAR1(255-401) (right). The representation of ΔH, ΔG and –TΔS corresponds to Figure 3.8.

by half in comparison to EgPKAR1(255-401), CNB:B. Counter directed entropy is lower in CNB:A than CNB:B ($-T\Delta S_{CNB:A}$ = 1.1kcal/mol and $-T\Delta S_{CNB:B}$ = 12kcal/mol). These findings suggest a more flexible interaction between cAMP and EgPKAR1(109-292), CNB:A, and are in good agreement with the literature reporting about a noticeably higher off-rate (5-10 fold) of cAMP from CNB:A than from CNB:B in mammalian PKA [73] [74] [27] [75]. Moreover, higher ligand binding affinity in CNB:B than CNB:A is compatible with experiments performed by Lorenz et al., 2017, on isolated CNB:A (EC_{50} = 151 $\pm$ 39nM) and B (EC_{50} = 3.6nM) of mammalian PKARIα using Surface plasmon resonance (SPR) [64]. However, in comparison to CNB:B in mammalian PKARIα, affinity for cAMP in EgPKAR1(255-401), CNB:B, is reduced by approximately 17 fold ($EC_{50}{}^{RI\alpha CNB:B}$ = 3.6nM and $K_d{}^{EgPKAR(255–401)}$ = 60nM). Thus, we clearly demonstrated that cAMP binds to both CNB domains of EgPKAR1, but with significant lower binding affinity to EgPKAR1(255-401), CNB:B, than to the isolated CNB:B of mammalian RIα due to the divergent ligand binding motif.

3.3.2 Contribution of key residues to ligand binding affinity in EgPKAR1

The surprising extension of the consensus for cAMP binding sites motivated a direct investigation of the role of residues localized in the key positions of PBC. Therefore, we generated mutants with either S347, C354 or both mutated to alanine. Additionally, molecular docking was carried out on the crystal structure of EgPKAR1 with respective mutations inserted in silico as described in the method section 2.2.5.
It was seen that all mutants are still able to bind cAMP with a reduced ligand binding affinity (Tab. 3.2 and Fig. 3.15). Compared to EgPKAR1(255-401) (left panel), we measured 2.5-fold lower binding affinity in the single mutant EgPKAR1(255-401)_S347A and 4.5-fold lower binding affinity in EgPKAR1(255-401)_C354A. Mutating both residues to alanine caused a 6.5-fold decrease in binding affinity. Thus, despite the absence of two residues, whose side chains are involved in ligand binding, cAMP still binds to EgPKAR1(255-401)_S347A_C354A in the nano molar range.
Interestingly, single residue substitution in mammalian PKARIα had significant higher impact on ligand binding affinity. Herberg et al., 1996, described a single site mutant of bovine RIα with R333K in PBC:B, in which the affinity for cAMP decreased 5-fold and the molar ratio of monomeric R-subunit to bound cAMP decreased to 1 (the wild type of PKAR1α was capable of binding 2 mole of ligand per mole of R-subunit) [27]. From these findings, the authors conclude that cAMP cannot bind to the mutated CNB:B in RIα(R333K) [27]. The similar mutant, EgPKAR1(255-401)_S347A_C354A, still binds cAMP and its binding affinity is only reduced 6.5-fold. These differences might be due to the different amino acid in position 12 of the PCB (lysine in RIα(R333K) and alanine

in EgPKAR1(255-401)_S347A_C354A). K333 could have stronger influence on tertiary structure of the binding site than the smaller alanine has. Furthermore, in Herberg et al., 1996, a full size construct of mammalian rRIα was used, whereas in our study CNB:B was expressed individually. It has to be considered that due to cooperative binding, mutations inserted into CNB:B could negatively influence ligand binding in CNB:A. This could lead to a greater loss of binding affinity in EgPKAR1(109-401), where both CNBs are incorporated.

Support for our binding data is gained by molecular docking using Glide enclosed in the program Maestro, Schroedinger suite Version 2017-14. For docking of cAMP to mutated CNB:B of EgPKAR1, the crystal structure was truncated in silico (aa 255-401) and the respective mutations were inserted. Best poses for each mutant were chosen according to lowest emodel and are shown in Fig. 3.15. Glide Gscore (GG) for poses with lowest Emodel is listed in Tab. 3.2. Glide Gscore is a value for estimated energy release upon

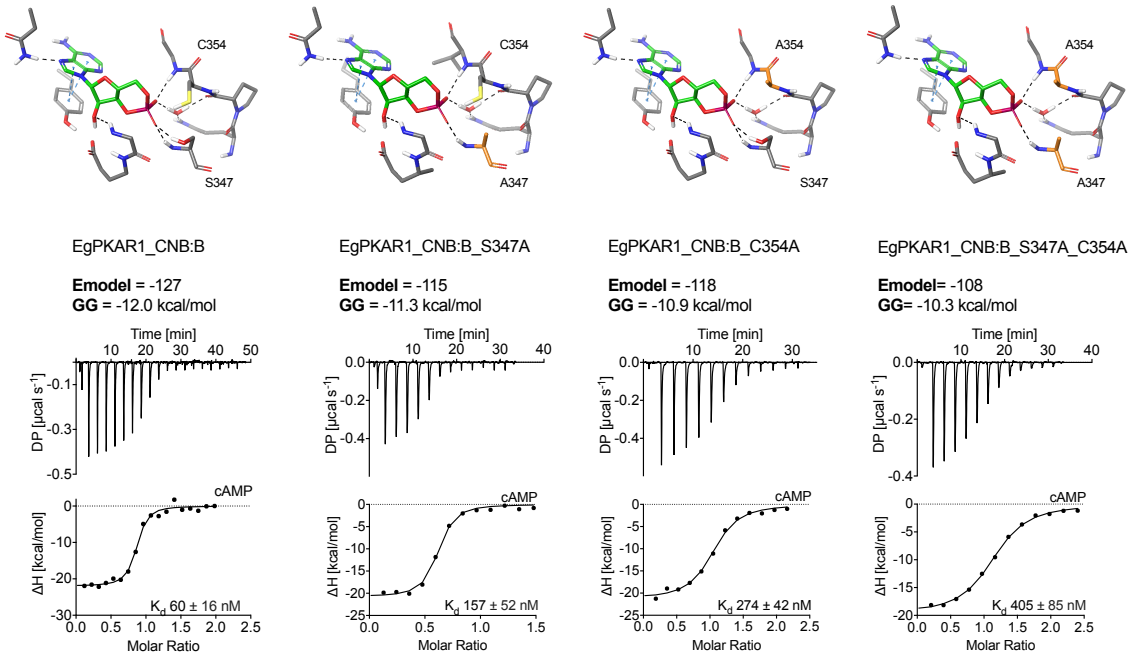


Figure 3.15: cAMP binding to EgPKAR1(255-401), EgPKAR1(255-401)_S347A, EgPKAR1(255-401)_-C354A and the double mutant EgPKAR1(255-401)_S347A_C354A (from left to right) was studied using ITC and molecular docking using Glide incorporated in Maestro (Schroedinger GmBH). Best docking poses were chosen on base of lowest Emodel and are shown above the binding data. Emodel and calculated Gscore (GG) for best poses is shown beneath. Binding data were derived from ITC experiments. Measurements were done in technical triplicates from one protein preparation.

ligand binding and should approximate experimentally determined ΔG. In our case, calculated GG exceeds measured ΔG, but resembles the order of mutants ranked according to their ligand binding affinity. To critically evaluate our docking results, a second experiment, using MM/PBSA, was done by Okke Melse [76]. Values for calculated ΔG_{calc} are

also shown in Tab. 3.2. Values from the MM/PBSA calculations are noticeably higher

Table 3.2: Comparison of experimentally determined ΔG, calculated Gscore obtained by molecular docking using Glide, Maestro, and ΔG_{calc} using MM/PBSA($*$).

Mutant	N	$K_d \pm$ SD [nM]	ΔG [kcal/mol] (ITC)	Gscore (GG) [kcal/mol]	ΔG_{calc} [kcal/mol] ($*$)
EgPKAR1(255-401)	0.8 ± 0.1	60 ± 16	-9.9 ± 0.2	–12.0	-49.8 ± 2.8
EgPKAR1(255-401)_S347A	0.6 ± 0	157 ± 52	-9.3 ± 0.2	–11.3	-49.0 ± 2.5
EgPKAR1(255-401)_C354A	1.0 ± 0.1	247 ± 42	-9.0 ± 0.1	–10.9	-48.4 ± 2.7
EgPKAR1(255-401)_S347A_C354A	1.1 ± 0.2	405 ± 85	-8.7 ± 0.1	–10.3	-46.5 ± 2.5

$*$ Experiments done by Okke Melse

than experimentally determined ΔG, since in calculation of ΔG_{calc} conformational entropy ($-T\Delta S$) was not taken into account. Nevertheless, the method is of good use for comparison of relative binding free energies between similar structures or ligands [76]. Thus, with focus on the ranking of mutants according to ΔG_{calc}, results from MM/PBSA are compatible with experimental data and Glide Gscore from molecular docking.
Concluding, the co-crystal structure of EgPKAR1 reveals a highly coordinated binding of cAMP, also in the unconventional PBC:B. Besides additional binding of the adenine ring, binding of the phosphate moiety is ensured by several backbone interactions that might explain how key residues of the classic cAMP binding domain can be replaced.

3.4 Evidence for a single nucleoside binding domain in LdPKAR3

Exclusive binding of nucleosides to LdPKAR3 from *L. donovani* was found by nDSF and could be confirmed in ligand binding assays using ITC as described in section 3.1. Using the ITC data, we found that the measured ligand-to-protein stoichiometry did not correspond to the two expected binding sites ($N \approx 1$). Therefore, crystallization of N-terminal truncated LdPKAR3(321-647) in complex with guanosine was attempted and the structure was solved to a resolution of 2.0 Å. Crystallographic data and refinement statistics are provided in Table 3.3. After expression in and purification from *E. coli*, size exclusion chromatography (SEC) was performed and the purity of LdPKAR3 was confirmed by SDS PAGE (Fig. 3.16).

Table 3.3: Data collection and refinement statistics for EgPKAR1(109-401) and LdPKAR3(321-647). Values for the best resolution shell are shown in parentheses.

Protein	EgPKAR1	LdPKAR3
PDB	6RSX	6TFC
Cell dimension		
a,b,c (Å)	80.586, 80.586, 110.755	211.898, 211.898, 72.41
α, β, γ ($\circ$)	90,90,120	90,90,90
Data collection		
Resolution Range (Å)	43.38 – 1606 (1.664 – 1606)	57.54 – 1963 (2.033 – 1963)
Space group	P61	I41
X-ray wavelength (Å)	0.999910	1000030
Unique reflections	52599 (4906)	112763 (9404)
Reflections used for R-free	2569 (255)	1996(166)
Multiplicity	20.0 (19.6)	13.4 (12.6)
Completeness (%)	99.17 (93.02)	98.23 (82.57)
Mean I/σ	14.33 (0.44)	9.26(0.72)
Copies per ASU	1	4
R_{merge}	0,073 (8.069)	0.127 (2.99)
$R_{pim(withinI+/I–)}$	0.24(2.581)	0.52(1.241)
CC 1/2	1.000 (0.443)	0.996 (0.302)
Refinement statistics		
R_{work} (%)	0.1923 (0.5454)	0.1988 (0.4833)
R_{free}	0.2163 (0.5639)	0.2310 (0.4999)
Average B-factor (A^2)		
Protein atoms	42.25	56.34
Solvent atoms	49.15	57.03
Ramachandran plot (%)		
Favored/Allowed	98.11/1.89	98.45/1.38

Overall organization of LdPKAR3 corresponds to other crystal structures of PKAR (Fig. 3.17). It is composed of two sequential binding domains, A site and B site, of which each contains a β-barrel and a helical subdomain. The arrangement of helices and β-sheets matches mammalian RIα [26] [23], with an additional αD helix in the B site that is also present in PKAR from *T. brucei* and *T. cruzi* [45] [George Githure, Yuri Volpato, manuscript in preparation]. Interestingly, the C-terminus in LdPKAR3 is even longer than in PKA from trypanosomes with additional 13 amino acids after the αD helix.

Another unique feature of the structure of LdPKAR3 is that only the B site binds a ligand. This correlates with an arrangement of the two binding domains that neither resembles the structure of ligand bound PKAR, e.g. TbPKAR (PDB: 6FLO), TcPKAR (PDB:6HYI) or mammalian RIα (PDB: 1RGS), nor the ligand unbound state of PKAR, to date only solved for RIα (PDB: 2QCS) [21]. Alignment of LdPKAR3 to cAMP bound RIα

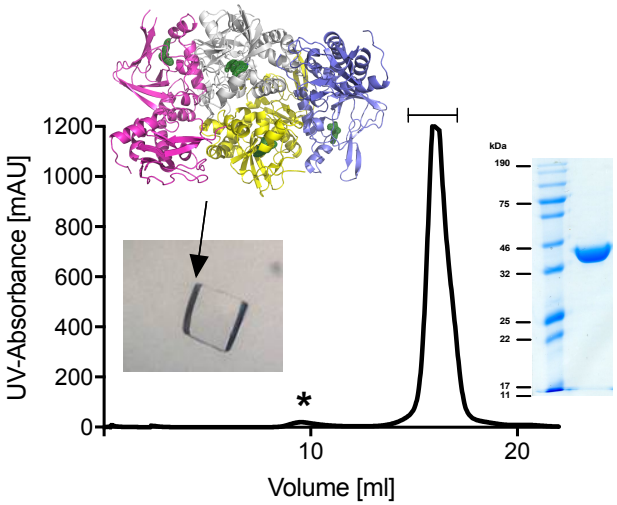


Figure 3.16: Native protein was purified from *E. coli* for crystallization. The line above the higher peak marks the collected fraction that was shown to be pure and of expected size by SDS PAGE (right). The peak marked with an $*$ is caused by aggregate protein and was not collected. A representative protein crystal of LdPKAR3 is shown on the left. The asymmetric unit was built by four molecules of LdPKAR3 and is shown as pymol illustration on the upper left.

(cyan; PDP: 1RGS) and cAMP unbound RIα (magenta, PDB: 2QCS, Chain B) on binding site A (aa 334-475 of LdPKAR3), is shown in Fig. 3.17. It reveals an intermediate state of LdPKAR3 (grey) that does not fully align neither to a ligand bound nor to a ligand free PKAR ortholog. Complete reorganization of mammalian PKARIα upon ligand binding results in an angle of 125° between the αB/C helices of RIα_cAMP bound and unbound. This angle is reduced to 30° between the α B/C helices of LdPKAR3 and RIα_cAMP unbound. Our data suggests that LdPKAR3 does not undergo complete structural reorganization upon ligand binding due to a missing ligand in the A site.

Binding site A appears empty in the crystal structure and in conformational H-state associated with high affinity for PKAC binding In site A, instead of a ligand, we find a structural motif that is known to positively influence the affinity for a catalytic subunit in mammalian PKA [70] [21] [71]. In Figure 3.18, A site from LdPKAR3 is aligned to a deletion mutant of RIα(91-244) in complex with PKAC (PDB: 3FHI). The N-terminal $N3A^{PKAR}$ motif and the $F\text{-helix}^{PKAC}$ are shown. The N-terminal N3A motif is a structural domain that precedes the β-barrel of binding site A and is composed of an αN helix, followed by a 3_{10}loop and the αA helix of the A site [21][71]. In ligand-free mammalian RIα, also called H-state of RIα, the N-terminal N3A moves inward and closer to PBC:A. Thereby, PBC:A and the N3A motif form a hydrophobic docking surface for a catalytic subunit [21] [70][77][78].

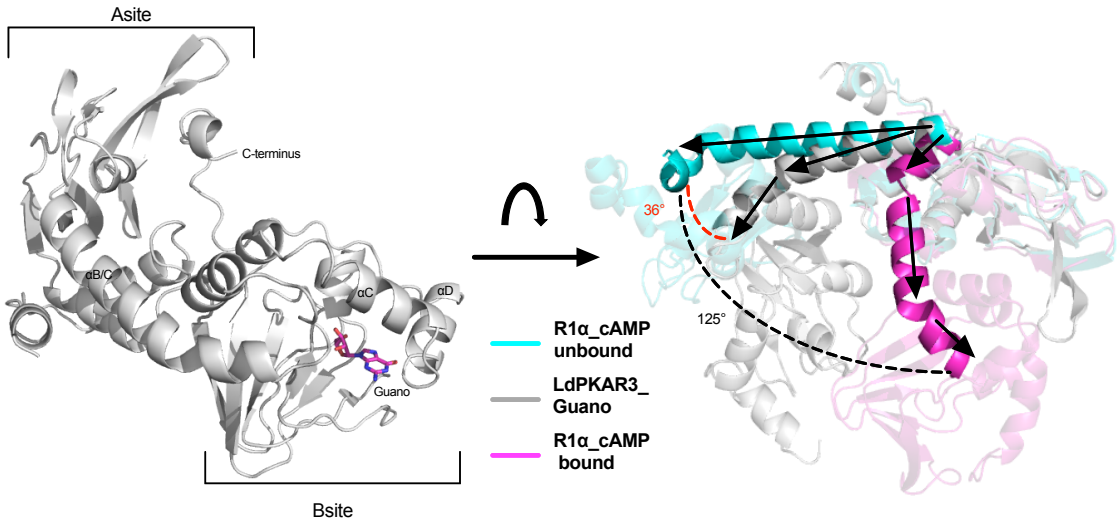


Figure 3.17: Left panel: Overall structure of LdPKAR3 (grey) with guanosine (purple) bound to the B site. Right panel: LdPKAR3 aligned to RIα_cAMP bound (PBD: 1RGS, cyan) and RIα holoenzyme conformation (PDB: 2QCS, Chain B, purple) on CNB:A (aa D334-R461). Dashed lines indicate angles that are formed by the αB/C helices of RIα_cAMP unbound and LdPKAR3 (red) or the αB/C helices of RIα_cAMP unbound and RIα_cAMP bound.

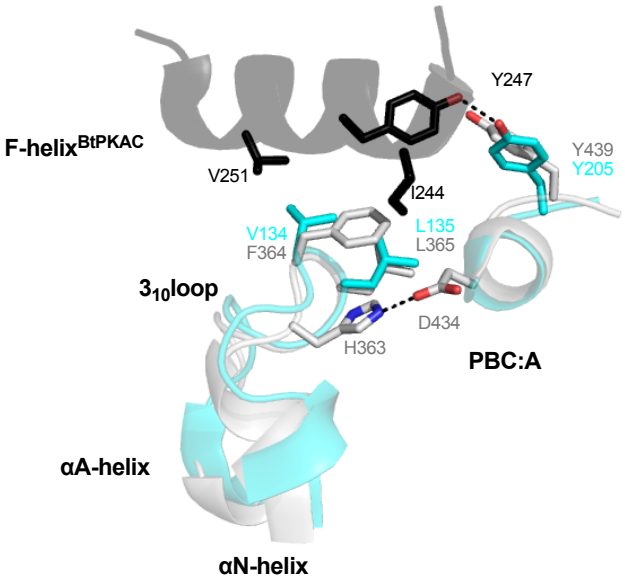


Figure 3.18: Alignment of LdPKAR3 (grey) to RIα(91-244) (cyan) that docks to the F-helix of PKAC (black) (PDB: 3FHI) [70]. Residues that contribute to an interface between RIα(91-244) and the catalytic subunit are shown as sticks in cyan or black. Corresponding amino acids in LdPKAR3 are shown as grey sticks and hydrogen bonds are shown as black dashed lines.

In our structure, we find the N-terminal N3A motif to be directed towards RBC:A, as seen in the ligand-free state of RIα. In this constellation of N3A motif and RBC:A, an interaction between D434$^{LdPKAR3_{PBC:A}}$ and H363$^{LdPKAR3_{N3A}}$ can be formed. This hydrogen bond corresponds to an interaction of E200 and N133 in the ligand-free conformation of RIα (PDB: 2QCS, chain B) in complex with the catalytic subunit [21]. Furthermore, residues that contribute to a hydrophobic docking surface for the catalytic subunit in mammalian RIα (cyan sticks) are conserved in LdPKAR3 and similarly directed (highlighted as grey sticks) [70] [21]. Thus, our data indicates that binding site A of LdPKAR3 does not harbour a ligand but exhibits a structure motif associated with high affinity for a catalytic subunit in mammalian PKARIα.

High affinity binding of guanosine in binding site B Unlike site A, site B harbours a ligand that we could identify as guanosine due to clear electron density in all four chains. Electron densities of ligand and interacting residues in chain A are shown in Fig. 3.19. Since we did not add any ligand before crystallization and best binders according to ITC experiments, inosine and guanosine, do not differ much in their binding affinity, it was highly interesting that we found only guanosine in the crystal structure of LdPKAR3. Nevertheless, the structure of LdPKAR3 explains high affinity for both nucleosides, inosine and guanosine, and reveals a guanosine-specific interaction, which makes a slightly higher affinity for guanosine than for inosine understandable.
Figure 3.19, right, shows the hydrogen network that is formed between protein and ligand.

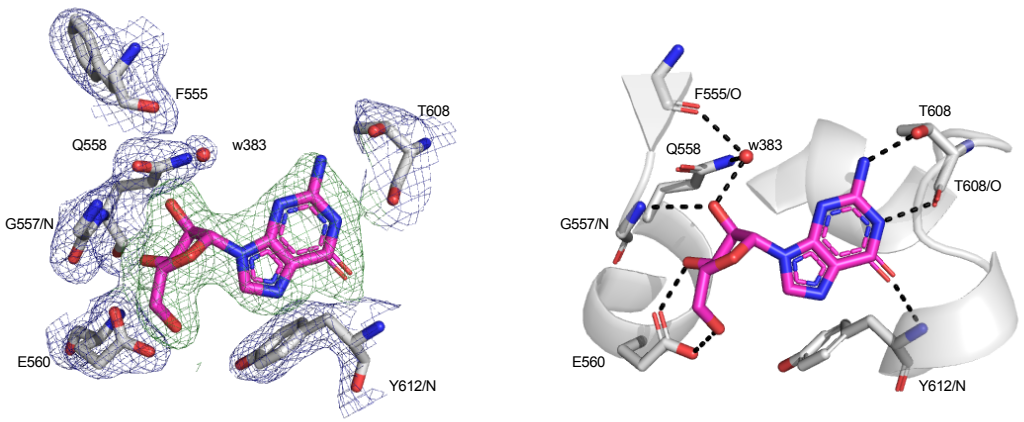
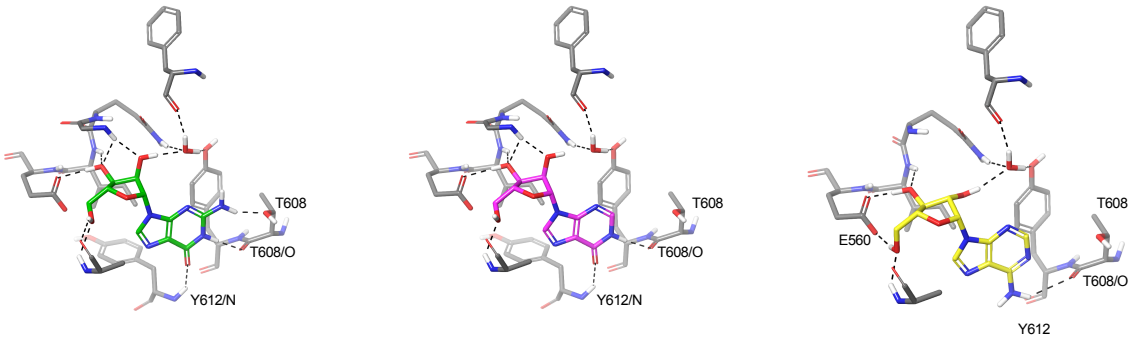


Figure 3.19: On the left, electron density maps were modelled for LdPKAR3 based on the 2Fo-Fc map (1σ blue map) and for the ligand based on Fo-Fc map (3σ green map). On the right, the hydrogen network (indicated by black dashes) surrounding guanosine (magenta) bound to LdPKAR3 in the B site (grey) is displayed. Structural waters are shown as red spheres.

The ribose ring is orientated towards RBC:B, whereas the aromatic ring is placed between αC and αD helix. Comparable to nucleoside binding domains in TbPKAR and TcPKAR, the ribose ring is captured between E560 and Q558 with an additional backbone interaction to G557/N. Q558 replaces a glutamate in position 5 of the RBCs in PKAR from *T. brucei* (PDB: 6FLO) and *T. cruzi* (PDB: 6FTF) and binds guanosine over a water molecule. Unlike PKAR from *T. brucei* and *T. cruzi*, aromatic residues, located in the αD helix, do not interact with the ligand via π-stacking. Instead, the purine ring is bound over hydrogen bonds to Y612/N and T608. Y612 binds to the keto group on the purine ring of guanosine and T608 to the amino group on C2 via the free electron pair of its hydroxyl group.
By molecular docking of guanosine, adenosine and inosine to binding site B of LdPKAR3, we showed that the interaction of Y612/N with the keto group on the purine ring can be formed with guanosine and inosine, whereas the interaction with T608 is guanosine-specific (Fig. 3.20). With adenosine neither of these interactions is formed, explaining its lower binding affinity compared to inosine and guanosine (Tab. 3.1). The missing contribution of Y612/N to the binding of adenosine and T608 to the binding of inosine and adenosine results in a reduction of binding free energy that is expressed in lower Glide Gscore. The lower Glide Gscore is compatible with measured ΔG in ITC experiments (Tab 3.1, Fig. 3.20).
Thus, highest binding affinity for guanosine (K_d = 4nM) is supported by the results of molecular docking. They furthermore explain why the binding of inosine to LdPKAR3 is almost equally good (K_d = 10nM), since only one hydrogen bond cannot be formed with inosine that can be formed with guanosine.



Guanosine

Emodel = -100
GGschore = - 10.2 kcal/mol
ΔG = -11.5 kcal/mol
RMSD = 0.326

Inosine

Emodel = -89.6
GGschore = -9.4 kcal/mol
ΔG = -11 kcal/mol

Adenosine

Emodel = -86.9
GGschore = -7.8 kcal/mol
ΔG = -9.1 kcal/mol

Figure 3.20: Molecular docking of guanosine (left), inosine (centre) and adenosine (right) to bindng site B in LdPKAR3 was performed using Glide incorporated in Maestro(Schroedinger GmBH). Best docking poses were chosen on base of lowest Emodel and calculated Gscore (GG) of best docking pose is shown beneath.

With our co-crystal complex of LdPKAR3 and guanosine we confirmed ITC data, in which ligand binding to only one site was indicated. In thermal denaturation assays using nDSF, guanosine leads to the highest stability of the protein (Fig. S6) and its tight interaction with the B site of LdPKAR3 can be structurally explained.

Interactions of the extended C-terminus with both binding sites in LdPKAR3 In contrast to PKAR subunits of mammals, PKAR subunits from kinetoplastids were shown to have an additional αD helix at their C-terminal end [45] [George Githure, Yuri Volpato, manuscript in preparation]. It forms a closing lid to trap the ligand in binding site B and its closing over the RBC:B was proposed to be an essential part of the activation mechanism [45] [George Githure, Yuri Volpato manuscript in preparation].
Interestingly, we find an even longer C-terminus in LdPKAR3, with additional 13 amino acids that follow the αD helix. It incorporates three proline pairs that stretch the C-terminal fragment and prevent it from forming secondary structure. Figure 3.21 shows a structural (left) and a schematic (right) representation of LdPKAR3 with focus on the C-terminus. In direct proximity to binding site B, the C-terminus contributes to the closing of the αD

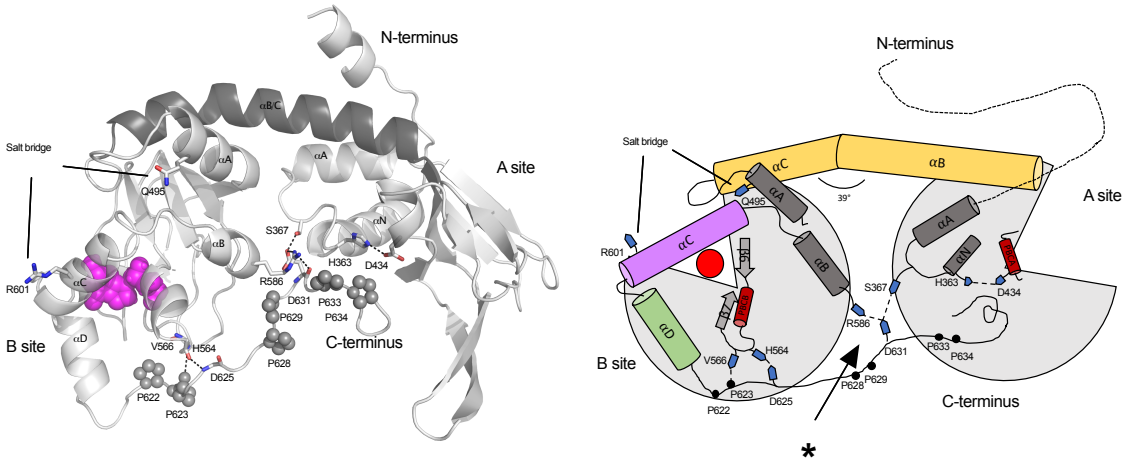


Figure 3.21: On the left, overall structure of LdPKAR3 is shown. Prolines that are localized in the C-terminus are shown as grey spheres. Guanosine in binding site B is shown as spheres in magenta. Residues that undergo interdomain interactions are shown as sticks. Hydrogen bonds are indicated as black dashed lines. On the right, a scheme of LdPKAR3 is shown, corresponding to the pymol illustration on the left. Binding site A and B are indicated as grey pies, secondary structure is partially implied in helices (cylinder) and beta-sheets (arrows). Amino acids that are indicated as sticks in the pymol illustration are shown as small blue pentagon or black dots for prolines. Guanosine in binding site B is indicated as red circle. The $*$ marks an interdomain interaction, which is shown in Figure 3.22.

helix. Unlike in the crystal structures of TbPKAR and TcPKAR, there is no aromatic $\pi-$stacking with the ligand in LdPKAR3. These stacking interactions are indispensable to keep the αD helix closed in TbPKAR and TcPKAR. In the structure of LdPKAR3, we find

the C-terminus to partially overtake this role in LdPKAR3 by forming hydrogen bonds with residues localized in the RBC:B (P623$^{\text{C–terminus}}$ and V566/N$^{\text{RBC:B}}$ and D625/N$^{\text{C–terminus}}$ and H564/O$^{\text{RBC:B}}$).
Furthermore, our data suggests that the C-terminus might hinder a ligand to bind in site A by interacting with the N3A motif in binding site A (Fig. 3.22). D631$^{\text{C–terminus}}$ binds to R586 in the αB helix of binding site B and S367 in the 3_{10}loop of binding site A (Fig. 3.22). This forms an interdomain interaction that stabilizes the inward directed position of the N3A motif and thus the ligand-free state of binding site A.

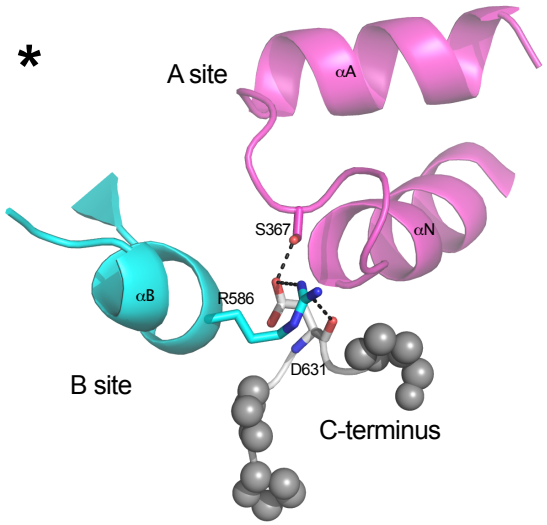


Figure 3.22: Interaction between site A, site B and C-terminus marked with ∗ in Fig. 3.21, right. Site A is coloured in magenta, site B in cyan and the C-terminus in grey. Interacting residues are shown as sticks, hydrogen bonds are indicated as black dashed line, prolines are shown as grey spheres.

3.5 Cell permeable nucleoside analogues bind to LdPKAR1 and LdPKAR3

Two cell permeable compounds, toyocamycin and 7_cyano_7_deazainosine (7CN), were tested on LdPKAR1 and LdPKAR3 in ITC experiments. Both were shown to be high affinity binders and potent activators of TbPKA in *T. brucei* [45]. The aim of this approach was to verify that both compounds bind LdPKAR3 and LdPKAR1 and thus could serve as a tool for kinase activation *in vivo* to investigate PKA function in this pathogen.
For comparison, binding of both nucleoside analogues to TbPKAR(199-499) was repeated (published by Bachmaier et al., 2019). In Figure 3.23, a representative curve for the binding of 7CN or toyocamycin is shown for all three proteins. Table 3.4 summarizes the obtained binding data.

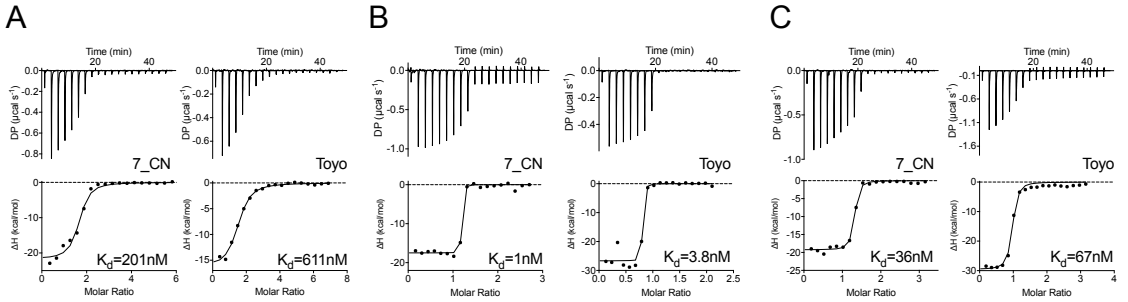


Figure 3.23: Binding of 7CN and toyocamycin to (A) LdPKAR1, (B) LdPKAR3 and (C) TbPKAR was studied by ITC. Illustration of binding data matches the representation in Figure 3.2, Section 3.1.

Table 3.4: Binding properties of 7CN and toyocamycin binding to LdPKAR1, LdPKAR3 and TbPKAR. Values are calculated from at least three independent measurements. Proteins were prepared at least two times.

Organism:PKA	Ligand	$K_d \pm$ SD [nM]	ΔH [kcal/mol]	ΔG [kcal/mol]	$-T\Delta S$ [kcal/mol]	N-value
LdPKAR1(200-502)	7_cyano_7_deazainosine	201 ± 87	-20 ± 2	-9.2 ± 0.3	11 ± 2	1.7 ± 0.4
	Toyocamycin	611 ± 82	$-17 \pm 1^*$	$-8,5 \pm 0.1$	$8.8 \pm 1.3^*$	$1.3 \pm 0.2^*$
LdPKAR3(321-647)	7_cyano_7_deazainosine	1.0 ± 0.6	-18 ± 2	-12.4 ± 0.5	5.3 ± 2.5	1.1 ± 0.0
	Toyocamycin	3.8 ± 2.1	-25 ± 1	-11.5 ± 0.4	14 ± 1.0	0.8 ± 0.1
TbPKAR(199-401)	7_cyano_7_deazainosine	36 ± 12	-19 ± 4	-10.2 ± 0.2	8.3 ± 4.0	1.3 ± 0.2
	Toyocamycin	67 ± 42	-28 ± 2	-9.9 ± 0.5	18 ± 3	0.7 ± 0.1

$*$ Data based on two independent runs only

TbPKAR(199-401) has a 2-fold higher affinity for 7CN than for toyocamycin (K_d = 36nM for 7CN and K_d = 67nM for toyocamycin). These values are in a similar range as published results (K_d(7CN) = 8nM and K_d(toyo) = 58nM) and confirm that both nucleoside analogues bind with high affinity to TbPKAR [45].

LdPKAR1 has a significantly lower binding affinity for both compounds (K_d(7CN) = 201nM and K_d(Toyo) = 611nM). As for TbPKAR, 7CN is the better binder of both and possibly also the better activator, since a 36-fold higher activation potency of 7CN compared to toyocamycin was shown for TbPKAR [45].

LdPKAR3 displays high affinity for both compounds with K_d = 1nM for 7CN and K_d = 3.8nM for toyocamycin. Consistent with ITC data obtained for purine nucleosides, ligand to protein stoichiometry indicates binding in only one site (N $\approx$ 1). Whether a holoenzyme formed between LdPKAR3 and a catalytic subunit can be activated by these compounds remains questionable, as it is not yet clear whether ligand binding to both sites is required for PKA activation in kinetoplastids.

Concluding, we show that 7CN and toyocamycin bind to LdPKAR3 and LdPKAR1 *in vitro* and in all available binding sites. As neither of the compounds is specific for only one of the two isoforms, it might not be possible to discriminate between the two orthologs *in vivo*. Therefore, further preliminary experiments will be required to enable *in vivo* investigations on LdPKAR1 and LdPKAR3 individually.

Chapter 4

Discussion

cAMP independency of PKA in Kinetoplastida has long been indirectly suspected [46] [79], but only recently have *in vivo* activators of PKA in *T. brucei* been identified [45]. Since these activators are nucleoside analogues, it seemed likely that also nucleosides could have an effect on the kinase. The highest potency for kinase activation *in vitro* and also highest binding affinity was observed for purine nucleosides, especially inosine and guanosine [George Githure, Yuri Volpato, manuscript in preparation].
In order to understand the ligand switch on a molecular level, co-crystallization of PKA from *T. brucei* with nucleosides and nucleoside analogues was performed. These crystal structures revealed, how nucleosides can be tightly bound, but cyclic nucleotides would clash with certain residues in the binding pocket [George Githure, Yuri Volpato, manuscript in preparation]. This observation was supported by experiments showing that the ligand change could be reversed by inserting mutations in both binding sites. With the mutation of three residues in RBC:A, E311A, T318R, V319, and RBC:B, E345A, N442R and V443A, respectively, kinase activation could be induced by cAMP (EC_{50} = 33μM) with a 10 fold higher EC_{50} for inosine (EC_{50} = 300μM) [George Githure, Yuri Volpato, manuscript in preparation].
Altered ligand specificity could also be shown for PKAR1 in *T. cruzi* and both paralogs, LdPKAR1 and LdPKAR3, in *L. donovani*, and is likely to be conserved among Trypanosomatida [George Githure, Yuri Volpato, manuscript in preparation] [This work]. To investigate whether the ligand switch also occurred in more distant relatives of trypanosomes, PKA sequences from representatives of the Euglenozoa were screened for deviations from the classic cAMP binding motif [24] (Screens done by Cormac Kinsella). Two PKA subunits from *E. gracilis* were found, that were predicted to differ with regard to their ligand specificity. In this work, we were able to verify the predictions biochemically, using Isothermal titration calorimetry (ITC) and structural analysis.

4.1 Are nucleosides the endogenous ligands of cNMP independent PKA?

With more PKAR orthologs showing cAMP independency but nucleoside binding, the question arises of whether nucleosides are the actual endogenous ligands and whether they are part of an intracellular signalling pathway. Therefore, binding affinities for nucleosides of our investigated orthologs have to be considered (Tab. 4.1).

Table 4.1: Comparison of ligand binding affinities in different PKA orthologs.

PKAR ortholog	Organism	cAMP	K_d Inosine	K_d Guanosine	K_d Adenosine
RIα*	*H. sapiens*	28nM	-	-	-
TbPKAR*	*T. brucei*	-	14nM	78nM	80nM
LdPKAR1	*L. donovani*	-	59nM	173nM	1157nM
LdPKAR3	*L. donovani*	-	10nM	21nM	235nM
EgPKAR2	*E. gracilis*	-	831nM	931nM	11000

* K_d values were taken from Bachmaier et al., 2019 and Yuri Volpato, PhD Thesis

Mammalian RIα, which serves as a reference model in this work, displays high binding affinity for its endogenous ligand cAMP (K_d = 28nM). K_d values that were determined for inosine binding to TbPKAR and LdPKAR1 are comparable to mammalian RIα (K_d = 14nM and K_d = 59nM, respectively). LdPKAR3 shows slightly higher affinity for guanosine (K_d = 4nM) than inosine (K_d = 10nM), but both ligand candidate bind even better to LdPKAR3 than cAMP to mammalian RIα. Hence, high affinity binding of nucleosides to investigated PKA from Kinetoplastida makes them promising ligand candidates.
In contrast, EgPKAR2 from *E. gracilis* has significantly lower affinity for nucleosides than its orthologs in kinetoplastids. In comparison with mammalian RIα, EgPKAR2 has a 30-fold higher K_d for inosine than RIα for cAMP. Since we would expect an endogenous ligand to bind with similar high affinity to EgPKAR2 as cAMP does to mammalian RIα, the role of nucleosides as endogenous ligands for EgPKAR2 must be questioned. Instead, as detailed in section 4.2.3, EgPKAR2 is possibly involved in an alternative activation mechanism, in which activation may not be controlled by ligand binding.
In the other organisms, nucleoside mediated PKA signalling is at least conceivable, although the question arises whether nucleosides serve as second messengers in the same way as cAMP. As kinetoplastids lack the ability of purine synthesis, their nucleoside metabolism is dependent on purine uptake from the host [80]. PKA in trypanosomatids might therefore function as a nucleoside sensor rather than nucleosides as intracellular second messengers. Thus, it could mediate changes in metabolism or differentiation that have been shown to be dependent on extracellular nucleoside supply. For example, it was

shown that metacyclogenesis in *L. donovani* can be triggered by an absence of purines [81]. Starvation-induced metacyclogenesis led to an increase of LdPKAR1, suggesting its involvement in differentiation of metacyclic promastigotes [11].Thus, PKA could act as a nucleoside sensor, mediating the effect of purine absence on metacyclogenesis.
However, it remains elusive to date how PKA is controlled within the parasites. It was shown that activation of TbPKA can be induced by cell permeable nucleoside analogues *in vivo* [45] and that nucleosides display high affinity as well as high activation potency on PKA in *T. brucei* and *L. donovani in vitro*. Whether intracellular changes of nucleoside concentration correlate with PKA activity or whether the actual endogenous ligand is not yet identified remains highly speculative and requires further investigations.

4.2 LdPKAR3, a novel type of PKAR-like isoform in pathogenic *L. donovani*

4.2.1 The role of LdPKAR3 in the life cycle of *L. donovani*

In parallel to this work, functional analysis of LdPKAR3 we carried out in collaboration between the groups of Prof. Michael Boshart and Prof. Dan Zilberstein, Technion in Haifa, Israel. Prof. Dan Zilberstein kindly provided us with unpublished data, which will be compared with the results of this work in the following.

LdPKAR3 localizes to the cytoskeleton of the parasite LdPKAR3 was discovered in screenings for post-translationally modified cellular proteins [82]. Together with its paralog, LdPKAR1, it was characterized as a nucleoside binding protein in this work and both are suggested to be involved in the life cycle of *L. donovani* [10] [11][9] [Fischer Weinberger et al., manuscript in preparation].
A knock out of LdPKAR3 in *L. donovani* provokes a strong morphology phenotype, disabling promastigotes to retain their elongated slender form [Fischer Weinberger et al., manuscript in preparation]. This could be correlated with the co-localization of LdPKAR3 with subpellicular microtubules in cell [Fischer Weinberger et al., manuscript in preparation]. In promastigotes, localization of LdPKAR3 is restricted to the cortical region in between nucleus and flagellar pocket, whereas in rounded amastigotes, LdPKAR3 can be detected on the entire cortex [Fischer Weinberger et al., manuscript in preparation]. Interestingly, LdPKAR3 is absent in salivarian trypanosomes that lack the ability of amastigote forming. Stage specific localization as well as stage specific phosphorylation patterns of

LdPKAR3, which have been described earlier, support the hypothesis of its involvement in cell differentiation and will be discussed in the following [9] [83].

N-terminal phosphorylation pattern suggests a role in differentiation In order to simplify expression and work on LdPKAR3, its long and non-conserved N-terminus was truncated. In mammalian RI and RII, the N-terminus carries a dimerization/docking domain, which is essential for homo dimer building and localization of the protein in certain cellular compartments over A-Kinase anchoring proteins (AKAPs) [84]. These mediate the association of various enzymes and substrates in order to facilitate their interaction [78] [85] [86].
The N-terminus in LdPKAR3 differs strongly from the N-termini of mammalian isoforms in terms of size and conserved domains, e.g. it lacks a dimerization/docking domain. However, it has several phosphorylation sites, some of which are (de)phosphorylation according to a stage-specific pattern, as shown in Fig. 4.1.

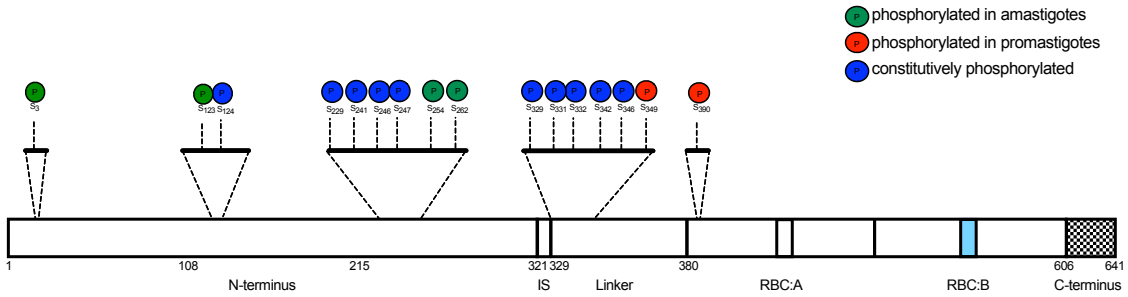


Figure 4.1: Phosphorylation sites that were identified by Tsigankov et al. (2013 and 2014) are shown for the respective domain of LdPKAR3, in which they are localized. Circles above indicate stage specific phosphorylation according to the colour code on the right.

Of in total 16 phosphorylation sites, 10 are localized N-terminal, preceding the inhibitory sequence. The remaining six are distributed in a linker between inhibitory sequence and RBC:A. The majority of those are phosphorylated constitutively. Stage specific phosphorylation sites are S349 and T390, as only found to be phosphorylated in promastigotes, and S3, S123, T254 and T262, only in amastigotes [83]. Augmentation of phosphorylation was shown to be independent of protein abundance and can be considered stage specific.
Of particular note is T262, an amastigote-specific phosphosite that is possibly phosphorylated by PKA itself on a PKA-specific phosphorylation motif (R-X-X-S/T). Phosphorylation of T262 increases strongly during the first five hours after perception of a differentiation signal [9]. Since rounding of cells usually starts 5 to 7 hours after exposure to a differentiation signal, significant changes in phosphorylation abundance during the first five hours

might be correlated with the initiation of amastigote morphogenesis [8] [Fischer Weinberger et al., manuscript in preparation]. Phosphorylation on T262 might therefore be part of an (auto)regulation mechanism of PKA to increase its activity in early differentiation.

4.2.2 Structure-derived model of an alternative kinase activation mechanism

Having successfully crystallized LdPKAR3 in complex with guanosine, we could show that LdPKAR3 binds nucleosides only in bindings site B whereas site A remains empty. Due to this clear deviations from beforehand characterized PKAR subunits, tandem affinity purification of LdPKAR3 with the catalytic subunits C1 and C3 from *Leishmania tarentolae* was performed (experiments done by George Githure). These results prove the capacity of LdPKAR3 to interact with catalytic subunits. Thus, it fulfils the criteria of a regulatory subunit of PKA.
Nevertheless, deviations in sequence and structure let us question whether the activation mechanism is completely conserved for LdPKAR3. Therefore, new features of sequence and structure and their potential influence on an alternative kinase activation mechanism will be discussed for this novel type of PKAR-like isoforms.

Inhibitory sequence in LdPKAR3 Sequence analysis of LdPKAR3 reveals the classic organization of PKAR, consisting of N- and C-terminus, and two binding domains positioned in between. Nevertheless, peculiarities in certain sequence motifs are observed, such as Inhibitory sequence (IS) and RBC:A. IS normally forms the sequence motif R – X – X – **S**/**T**/**G**/**A** – Ψ (with Ψ = hydrophobic residues) and is essential to distinguish between RI and RII isoforms. RII isoforms carry a serine or threonine in the P-site (bold residues in IS motif) of the IS, on which they can be phosphorylated by PKAC. The IS in RI isoforms, on the contrary, is only a pseudo substrate for PKAC, as the P-site is occupied by a glycine or alanine [20] [87]. Additionally, RI isoforms exhibit a serine in position P+2 that can be phosphorylated by PKG and is therefore an essential residue of PKA/ PKG crosstalk [88] [89].
Due to its inhibitory sequence, we conclude that LdPKAR3 resembles an RI like isoform. However, in LdPKAR3, the first arginine of the IS is replaced by a serine (Tab. 4.2), forming a S-R-X-A-Ψ motif. R94 and R95 in a deletion mutant of RIα are important for inhibition of the catalytic subunit because they interact with several residues in the active site cleft [70]. R94 in RIα of *B. taurus* shows intense binding to the C-subunit including ion pairs to $E127^{C}$ in a glycine-rich loop and a hydrogen bond to $T51^{C}$ [70]. These two residues are

conserved in PKAC1 (Accession (TriTryp): LINF_350045100) and C3 (Accession: LINF_-180016200) of *Leishmania infantum*. As R94$^{RI\alpha}$ is replaced by S324 in LdPKAR3, similar intense binding to the glycine-rich loop most likely cannot be established.
However, LdPKAR3 is not the first putative PKAR subunit with an atypical inhibitory sequence, as a T-R-X-S-Ψ motif was found in ParaR from *Paramecium tetraurelia* [24]. Possible consequence of the altered inhibitory sequence can be a looser interaction between R and C subunit, resulting in a more flexible protein complex.

Different roles of binding site A and B in LdPKAR3 It is known that binding site A and B have different functional roles in the activation mechanism of PKA. Therefore, their different conformational states in our crystal structure are of great importance to assess whether a kinase activation might be altered for LdPKAR3.
Ligand binding to CNB:B is described as “gatekeeper” function, necessary to induce a conformational change, that leads to accessibility of CNB:A for ligands in mammals [70]. On the other hand, CNB:B is not necessarily required for R/C interaction, as a truncation of mammalian RIα(91-244) lacking CNB:B could be co-crystallized in complex with a catalyitc subunit [70]. Moreover, it was shown that ligand binding to CNB:A is the critical step to complete the dissociation of R and C [27] [90] [21]. This activation mechanism is likely to be conserved in nucleoside binding regulatory subunits from *T. brucei* and *T. cruzi* [45].
In LdPKAR3, we find a ligand only in binding site B, whereas binding site A crystallized in a ligand-free or H-state, which has a high affinity for a catalytic subunit [77] [70]. Especially, the αB/C helix is nearly extended and outwards orientated, so that it could contribute to an R/C- interface maintaining the interaction between regulatory and catalytic subunit [21] [77]. We therefore consider it rather unlikely that small conformational changes induced by ligand binding to binding site B are effective enough to disrupt the interface between LdPKAR3 and PKAC. It is conceivable that LdPKAR3 is able to bind ligand and catalytic subunit at the same time. This hypothesis is supported by the observation that a holoenzyme complex formed between PKAC1 from *L. donovani* and LdPKAR3(1-647), purified from a *Leishmania tarentolae* expressing system (LEXY), could not be activated by nucleosides (experiments done by George Githure, data not shown).
Furthermore, guanosine that was bound after purification from *E. coli* could not be removed without denaturation of the protein, although a high excess of inosine was added to compete for binding (data not shown). This observation suggests that guanosine is bound constitutively to LdPKAR3, whereas protein activation is overtaken by alternative mechanisms.
Constitutive binding of guanosine in binding site B and inaccessibility of binding site A

for ligands might be favoured by the extended C-terminus of LdPKAR3. The C-terminus closes binding site B by securing the αD helix and helps to maintain the N3A motif in an inward moved position, a key criterion of ligand-free PKARIα (Cf. section 3.4, Fig. 3.22) [77]. In addition, the C-terminus could contribute directly to the R/C interface, as it contains various proline pairs and hydrophobic residues. With their flat, hydrophobic surface, prolines interact with aromatic rings and increase the rigidity of the peptide they are placed in [91].
An alignment of LdPKAR3 to the complex of RIα(91-244):C allows us to imagine a possible holoenzyme, in which guanosine is still bound to the B site of LdPKAR3 (Fig. 4.2). The C-terminus takes a position close to the N3A motif of CNB:A, but also to the F-helix of mammalian PKAC, allowing us to speculate that it directly expands the interface via its prolines and additional hydrophobic amino acids.

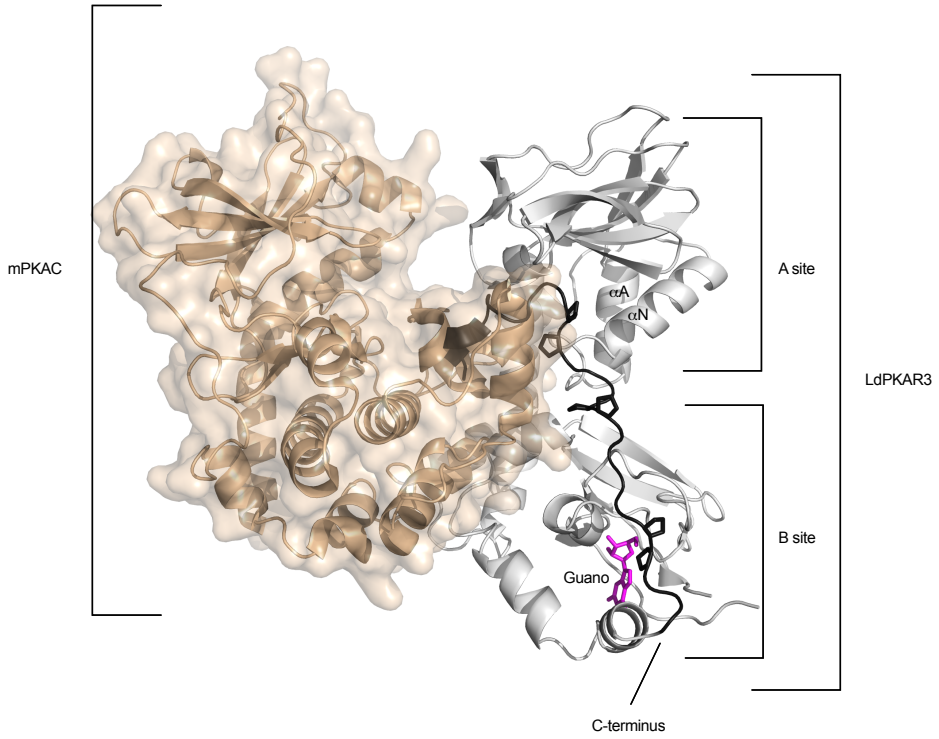


Figure 4.2: LdPKAR3 was aligned to a RIα(91-244):C complex (PDB: 3FHI). Of the RIα(91-244):C complex only the catalytic subunit (gold) is shown. LdPKAR3 is coloured in grey with guanosine (stick representation, magenta) bound to the B site. The C-terminus of LdPKAR3 is coloured in black and proline pairs are shown as sticks.

Possible influence of protein phosphorylation on kinase activation Concluding, we propose a model, in which the C-subunit remains attached to the R-subunit, even in the presence of ligands in the B site. LdPKAR3 and PKAC might be associated in a looser

complex that enables the C-subunit to be active, although it is trapped in a holoenzyme complex. Support for this hypothesis is gained by the postulation of an intermediate state of PKARIα and PKARIIα binding to PKAC [88] [92] [93] [94]. The observation of active PKAC in a holoenzyme complex challenges the previous model, in which holoenzyme activation requires dissociation of the R_2C_2 complex.
Moreover, alternative regulating mechanism with special regards to protein phosphorylation were discussed in these works, completing our knowledge of the PKA activation mechanism. Particular to emphasize is the phosphorylation of S101 by PKG in PKARIα that weakens the interaction between IS of PKARIα and the active site cleft of PKAC. This alteration on S101 in PKARIα, mimicked by the phosphomimetic mutation S101E, does not impede holoenzyme formation but leads to augmented PKAC activity in absence of cAMP [88]. As shown in Figure 4.1, Tsigankov et al. found the corresponding amino acid in LdPKAR3, S329, to be constitutively phosphorylated [83]. It is conceivable, that phosphorylation of S329 has a similar diminishing effect on binding tightness in LdPKAR3, as it has on PKARIα-PKAC.

On base of our own and also aforementioned findings in the literature, we propose an activation model of PKA in *L. donovani* that does not require dissociation of LdPKAR3 and the catalytic subunit. An explanatory model of activation can be, that R and C constitutively stay associated, only switching between a looser and a tighter conformation. This could possibly be provoked by post translational modifications that induce the transition from active to inactive state and back. To address this question in future, a holoenzyme crystallization approach along with kinase activation assays will help to understand the mechanisms underlying this somehow conserved and somehow new regulatory subunit.

4.2.3 Comparison of LdPKAR3 to the homology model of EgPKAR2

A homology model of the structure of EgPKAR2 was built on the base of PKAR from *T. cruzi* (PDB: 6FLO) and *T. cruzi* (PDB: 6FTF) by Okke Melse (Group of Prof. Dr. Iris Antes). Comparison of EgPKAR2 to LdPKAR3 reveals highly similar tertiary structure. Figure 4.3 shows an alignment of EgPKAR2 to site A of LdPKAR3 (aa 334-484). EgPKAR2 thereby resembles LdPKAR3 in organization of binding site A. Despite inosine binding to RBC:A of EgPKAR2, the A site is widely opened to the adjacent solvent surrounding. Moreover, the αB/C helix of EgPKAR2 is stretched, aligning well to the αB/C helix in LdPKAR3.
Regarding the astonishingly low binding affinity in site A of EgPKAR2 (K_d = 1.3μM) and high solvent exposition of the ligand, mostly due to the missing capping residue, we consider the association of ligand and protein in binding site A rather unstable. These obser-

vations make us speculate that, if the alternative kinase activation mechanism is true for LdPKAR3, it might be similar for EgPKAR2. Hence, a kinase activation mechanism that does not require dissociation of C and R subunit may not be restricted to LdPKAR3, but may in fact widespread within the Euglenozoa.

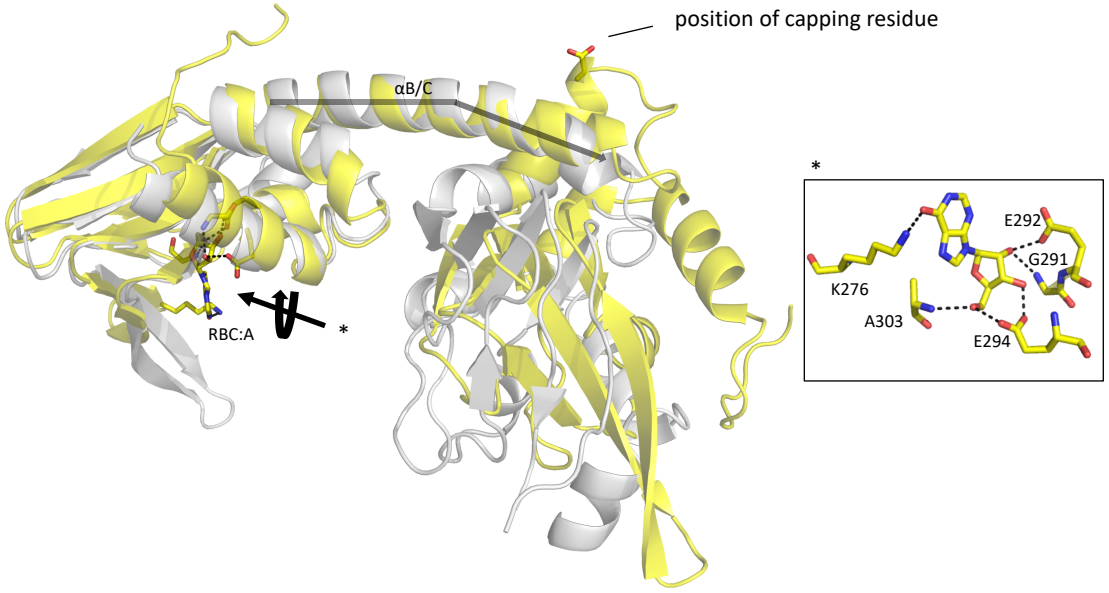


Figure 4.3: Alignment of EgPKAR2 (yellow) to A site (aa 334-484) of LdPKAR3 (grey). Ribose binding cassette A (RBC:A) of EgPKAR2 is shown as ball and stick model on the right. The model of EgPKAR2 was done by Okke Melse (group of Prof. Dr. Iris Antes).

4.3 Evolution of PKA in distant eukaryotes

Many other unicellular eukaryotes, for example *Plasmodium falciparum* or *Yeast*, only possess a single regulatory subunit [95] [65]. Most species in the Euglenozoa have at least two PKAR isoforms. Similar to the development of four distinct PKAR isoforms in mammals, RI(α, β) and RII(α, β), we suggest a subdivision in functionally non-redundant PKAR isoforms has taken place in Euglenozoa [96] [97], but here with respect to ligand specificity. By subdivision into certain ligand specificity types, we can distinguish four different types present in the Euglenozoa (Fig. 4.4): The first, represented by EgPKAR1, binds exclusively cyclic nucleotides. The second, which has not been dealt with in this work, is represented by PKAR from *Diplonema papillatum* and binds nucleosides in the B site and cyclic nucleotides in the A site [Yuri Volpato, PhD Thesis]. The third, represented by EgPKAR2 and LdPKAR1, and fourth, represented by LdPKAR3, bind exclusively nucleosides but the latter only in binding site B.

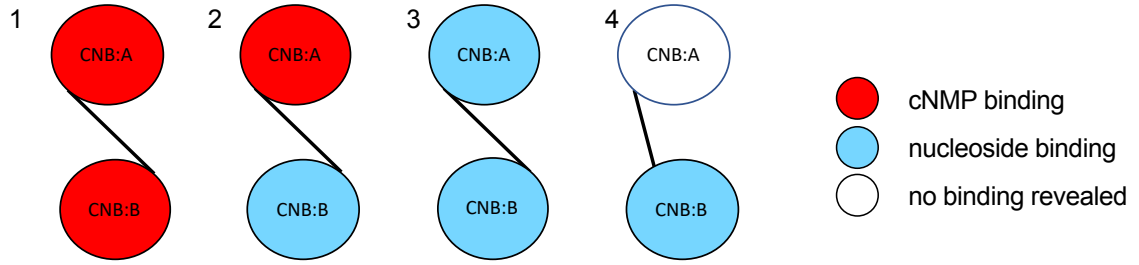


Figure 4.4: Schematic representation of four PKAR subtypes. The two CNB domains (circles) are connected over an αB/C helix (black line). Ligand specificity of the respective sites is indicted according to the colour code on the right. (1) cNMP dependent PKAR (2) Dual specificity type (DpPKAR1) (3) nucleoside type PKAR (4) PKAR3 type.

4.3.1 Distribution of PKAR ligand specificity types within the Euglenozoa

Sequences of different PKAR orthologs from representatives of the Euglenozoa were compared in a multiple sequence alignment using Muscle integrated as web service in Jalview 2.10.4b1 (Fig. S5). Ligand binding specificity was predicted on base of sequences that were found in the PBCs or RBCs, respectively. Were these compatible with the cAMP binding motif defined by Canaves and Taylor, 2002, we considered the ortholog to be cAMP specific [24]. Did the PBC motifs, however, correlate with the derivative sequence that was found in nucleoside dependent PKAR, we predicted the ortholog to be nucleoside specific. The PKAR3 type was assigned to those orthologs whose sequence in RBC:A was neither compatible with a cAMP nor a nucleoside specific motif, but that carried a nucleoside-specific binding motif in RBC:B. A mixed type, as seen for DpPKAR1, was not found in any other organisms. Compared PKAR orthologs are shown in Figure 4.5. Their predicted ligand specificity is indicated according to the colour code on the right.

We find the cAMP dependent type, the nucleoside type and the PKAR3 type in different organisms, with the nucleoside dependent PKAR being the predominant type among the compared species. cAMP dependent PKARs are present in all three classes, hence Kinetoplastida, Euglenida and Diplonemida, but seem to be lost after the branching of Trypanosomatida. Interestingly, in *Diplonema papillatum*, cAMP dependency is integrated in a mixed type, DpPKAR1, that binds both, cyclic nucleotides and nucleosides. Furthermore, it has no clear selectivity for cAMP or cGMP, as both bind equally good to DpPKAR1 (K_d = 3.1nM for cAMP and K_d = 3.5nM for cGMP) (Data not shown, Yuri Volpato, PhD Thesis). The CNB domain in DpPKAR1 could therefore represent an ancient domain that has not yet developed specificity for certain cNMPs.
After the branching of euglenids, a PKAR3 type appears in diplonemids and kinetoplastids. It can be found in both free-living and parasitic organisms, but is absent from try-

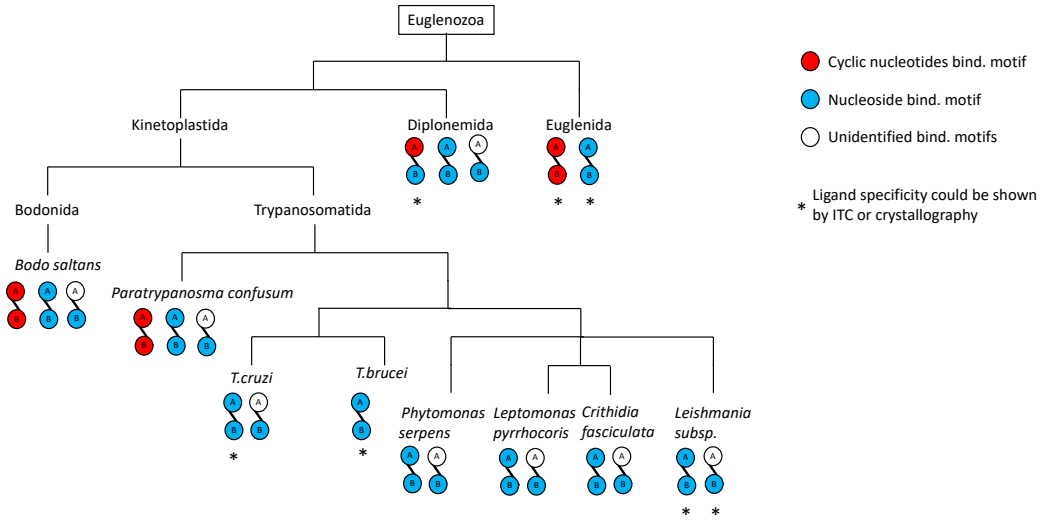


Figure 4.5: Distribution of PKAR subtypes in the phylum Euglenozoa. Ligand specificity of PKAR orthologs was predicted on base of multiple sequence alignment using Muscle integrated as web service in Jalview 2.10.4b1 (Fig. S5) and is indicated according to the colour code on the right. Ligand binding specificity was biochemically proven by ITC or crystallography for orthologs marked with an $*$.

panosomatids, which cannot form an amastigote cell stage [Fischer-Weinberger et al., manuscript in preparation].

4.3.2 Loss or gain of PKAR ligand specificity types in euglenozoan phylogeny?

The distribution of ligand specificity types shown in Figure 4.5 requires a model of gene loss or gain within the Euglenozoa. The appearance of new PKAR isoforms is likely due to gene duplication events with a vertical transfer of duplicated genes to descending lineages. An alternative would be horizontal gene transfer, i.e. between two simultaneously existing species. Therefore, genetic events behind the presence or absence of certain ligand specificity types and possible benefit associated with such events will be discussed in the following.

Possible development of a nucleoside binding type by gene duplication in the last common ancestor of the Euglenozoa Sequence analyses of PKA orthologs in various representatives of the Euglenozoa suggest that a nucleoside specific PKA is the prevailing form of the phylum. A cyclic nucleotide and a nucleoside specific type of PKAR most

likely were present at the time of the separation of the Euglenozoa, as we find them to co-exist in the same organisms, for example *E. gracilis*, *Bodo saltans* or *Paratrypanosoma confusum*.
Distributed among most lineages and found in the earliest eukaryotes, we consider it likely that the cyclic nucleotide dependent type of PKAR is the more ancient of both [98]. Hence, a gene duplication event in the last common ancestor of the Euglenozoa can explain the origin of the second, nucleoside binding type, comparable to the evolution of mammalian isoforms RI and RII [65].
EgPKAR2 could represent an early isoform that has undergone the ligand switch because its affinity for nucleosides is comparatively low. Trypanosomatids later developed the ability to respond to nucleoside binding and activation on a low nano molar level [George Githure, PhD Thesis].
The chimeric DpPKAR1 from *Diplonema papillatum* might also be an early nucleoside binding isoform, since it still binds cAMP in CNB:A. Moreover, co-existence of a cyclic nucleotide specific and a nucleoside specific domain in one protein might indicate domain shuffling of CNBs that have acquired different ligand specificity.

Absence of a cNMP dependent type in various trypanosomatids Shown representatives of the trypanosomatids do not possess a cNMP dependent type of PKA according to our sequence analysis in Figure 4.5.
The absence of of a "classical" PKA can not be correlated with a parasitic lifestyle as we find a cNMP dependent PKA to be present in *P. confusum*, the most basal branch of Trypanosomatida and a parasite of mosquites [99] [61].
Instead, the loss of a cNMP binding PKA might be associated with the appearance of other cAMP downstream effectors in Trypanosomatida. A genome wide RNAi screen in *T. brucei* identified certain genes whose knock-down resulted in resistance to elevated cAMP levels [44]. These were than named cAMP Response Proteins or CARPs [44].
Of these, CARP1 (Tb927.11.16210) has two complete cyclic nucleotide binding domains, and has been shown to bind cAMP (Radek Omelianczk, Eleni Polatoglou, unpublished results). cAMP binding was also shown by affinity chromatography for an ortholog of CARP1 in *T. cruzi* (TcCLB.508523.80) [100]. Interestingly, the protein could be displaced from cAMP coupled beats by cAMP as well as cGMP indicating a low specificity for either of the two cyclic nucleotides. However, conservation of CARP1 also in *Leishmania* [44] suggests that the development of alternative cAMP downstream effectors is a shared feature among the three parasites, and might be paralleled with the loss of a cAMP dependent PKA.
A second reason for the loss of a cAMP dependent PKA in trypanosomatids might be

the role of cAMP in host infection. Elevated cAMP levels produced by adenylate cyclases of lysed *T. brucei* in early host infection are known to activate host PKA. This results in inhibited synthesis of TNFα, a cytokine that is necessary to control parasitemia [33]. *Leishmania spp.* also induce elevated cAMP levels in macrophages to weaken the innate immune response [101]. High levels of cAMP that are produced to fight the hosts immune system could interfere with the parasites own signalling system, making the loss of a high affinity cAMP dependent PKA necessary. However, given the presence of a cAMP machinery for production (ACs) and degradation (PDEs) in trypanosomes, the second argument is less likely, as cAMP signalling certainly occurs in these organisms.

The development of nucleoside dependent PKAR3 in Trypanosomatida A PKAR3 type is found in diplonemids and kinetoplastids, but seems to be absent in euglenids. This observation indicates that a PKAR3 type is either an ancient type that has been lost in euglenids or it developed after euglenids branched from the remaining Euglenozoa.
Nucleoside specificity of binding site B suggests that PKAR3 may have evolved from a nucleoside specific type in which site A was replaced. Whether the non-functional site A of LdPKAR3 is an actual new domain, or whether it developed from either a cAMP binding or a nucleoside binding domain, remains unclear.
Moreover, it is questionable whether all orthologs of the PKAR3 type do not bind ligands in site A. In *Bodo saltans*, we find an ancient relative of LdPKAR3 that misses the prolonged, proline enriched C-terminus, a hallmark of LdPKAR3. As we discussed in section 4.2.2, the C-terminus might favour a ligand-free state of binding site A. Its absence could therefore lead to a restored ligand binding capacity of the A site in this ortholog.
A PKAR3 type can be found in *Leishmania subspecies* as well as most of their monoxenous, insects parasitizing relatives [61]. It is also present in *T. cruzi*, but not in *T. brucei*. *T. brucei*, unlike *Leishmania spp.* and *T. cruzi*, is not an obligate intracellular parasite, but survives in the bloodstream after transmission to mammals [102]. Hence, involvement of PKAR3 in cell morphology could explain why it is such an essential protein in *Leishmania spp.* and *T. cruzi*, but dispensable for *T. brucei* [Fischer Weinberger et al., manuscript in preparation].

Does a ligand change in PKA correlate with the origin of parasitism in Euglenozoa? The loss of redundant or non-essential proteins is consistent with the so-called streamlining of parasite genomes [103]. Whereas *Bodo saltans* has three ligand specificity types of PKAR, most of the trypanosomatids have lost at least one and in case of *T. brucei* even two ligand specificity types. Assignment of the respective PKAR subtypes to a free-living or parasitic lifestyle is not possible, as all three types were found in parasites as well as

free-living organisms. However, attribution of certain genes to parasitic or non-parasitic lifestyle seems to be quite challenging for the Euglenozoa, as even homologs of virulence associated genes of trypanosomes can also be found in euglenids and diplonemids [104]. Thus, rather than the development of parasitism-specific genes, genes present in free-living and parasitic species could have been adapted and repurposed during evolution to match the requirements of a parasitic lifestyle [104].
These considerations can also be made for PKA. PKA in parasitic trypanosomatids, a high affinity nucleoside binding protein, could be involved in nucleoside sensing. Since kinetoplastids lack the ability of purine synthesis, uptake from host or environment is essential [80] [105]. PKA could be part of a regulation mechanism that controls purine uptake from the host. Such regulation mechanism would not be required for euglenids and diplonemids, which are fully capable of purine synthesis [104]. Therefore, PKA likely has a different, but yet unknown function in these organisms.
Thus, nucleoside specific PKA could have been repurposed from a low affine protein with yet unknown function in free-living Euglenozoa to a high affine nucleoside sensor in parasitic trypanosomatids.

4.4 Can ligand specificity of CNB domains be predicted from primary sequence?

Predicting protein properties on base of amino acid sequence has always been tempting, as it simplifies our work tremendously. However, we clearly experienced the limitations of this method, as we could disprove cAMP dependency of certain PKAR orthologs, although their sequences were altered only marginally [45] [George Githure, Yuri Volpato, manuscript in preparation].
By comparison of our findings with previous knowledge about key residues in PKAR isoforms, we gain a deeper understanding of how single amino acid substitutions alter ligand binding properties significantly. Table 4.2 provides an overview over various amino acids whose functions are known for PKARIα. Critical substitutions in PKAR isoforms from *E. gracilis* and *L. donovani* are listed and highlighted when their function was determined in this work by site directed mutagenesis.
Our data indicates that only a rough assignment to ligand selectivity types is possible. We are able to distinguish with reasonable certainty between cNMP dependent and independent PKAR, as a substitution of an alanine by a glutamate in position 5 of the PBC makes it impossible for cNMPs to bind (Tab. 4.2) [45]. All the nuances that cause preferences within a certain ligand group, as exemplary seen for LdPKAR3 preferentially binding guanosine, were not predicted on base of sequence. Furthermore, the new cAMP binding

Table 4.2: Functional important residues in PKARIα and their substitutions in other PKAR isoforms are summarized. References, in which residue and function have been described, are shown on the right.

Domain	Critical residues in PKARIα from *B. taurus*	Substitution				Function	Reference
		LdPKAR3	LdPKAR1	EgPKAR1	EgPKAR2		
Inhibitory sequence	R94 A97	S324 A227	T206	G115	T188	Inhibiton of PKAC (Pseudo-)substrate site	[70]
Linker region	S99*	S329		S177	S191	P+2, phosphosite for PKG	[94][88]
	A100 E101 V102 Y103 Y106					R/C-Interface	[70]
3_{10}loop	N133	**H363**				Interaction with $E200_{PKARI\alpha}$	[26][21]
	V134 L135					R/C-Interface	[70]
$\alpha A^{CNB:A}$	E143					R/C-Interface	[70]
$PBC^{CNB:A}$	E299	**D434**	E310	E223	E292	Ligand binding	[26][21][45]
	A202	G436	E312	A225	E294		
	Y205	Y439	Y315	Y228	Y297	binds $Y274^{PKAC}$	[65][70][21][106][107]
	R209	N443	A319	R232	A301	cAMP binding	
	A210					ligand binding	
$\alpha B/C^{CNB:A}$	R230 L233 M234 T237 L238					R/C-Interface	[70]
	R241					binds $E200^{R}$, $D267^{R}$ $K213^{C}$	[65]
αACNB : B	W260	H494	Y371	Y283	D352	Capping residue CNB:A	[21] [45]
	E361					Salt bridge	
	D261					Int. with $R241^{PKAR}$	[65]
$PBC^{CNB:B}$	E324	Q588	E434	E345	E315	ligand binding	This work George Githure, PhD thesis
	A326	E560	E436	**S347**	E317		
	R333	A567	T443	**C354**	R424		
$\alpha C^{CNB:B}$	R366	R601	R476	R387	R457	Salt bridge	[21]
$\alpha D^{CNB:B}$		**T608**				LdPKAR3 spec. int. with guano	This work
	Y371	Y609	Y483	Y392	Y462	Capping residue CNB:B	[70][21]
			Y485		F464	Capping motif	[45] [George Githure, Yuri Volpato,
			Y486		Y465	(nucleoside bind. domain)	manuscript in preparation]

∗also referred to as S101 [88]

motif found in CNB:B of EgPKAR1 has not been expected as such on base of sequence. Hence, prediction of ligand selectivity would be oversimplifying and not considering all the aspects we gained from biochemical experiments.

Conversely, sequence helps us to interpret our biochemical data: A missing capping residue in EgPKAR2 explains low binding affinity in binding site A; T608 that specifically interacts with the amine group of the guanine base, explains preference for guanosine over other nucleosides in LdPKAR3 (Tab. 4.2).

The more we compare our biochemical data with variations we find in sequence, the more precise our predictions might get, allowing us to distinguish between more than just cNMP dependent or independent in future.

Chapter 5

Conclusion and perspectives

5.1 Conclusion

In this work, we showed that cNMP independency of PKA is not restricted to trypanosomes, but also found in their more distant relatives. Similar to former characterized PKAR in *T. brucei* and *T. cruzi*, cNMP independent PKAR in *E. gracilis* and *L. donovani* also binds nucleosides instead.

A nucleoside specific PKA seems to be the prevailing form in the phylum Euglenozoa according to sequence comparison. We thereby have to distinguish between a fully nucleoside dependent and a PKAR like type, which binds nucleosides only in one binding site.

Interestingly, not only do we find different ligand specificity types to co-exist in the same organism, but even in the same protein, namely DpPKAR1 from *D. papillatum*. We conclude that ligand selectivity of PKA is not necessarily defined for the whole regulatory subunit, as expected, but only for the single binding site.

In these, we found structural determinants that not only explain selectivity for either nucleosides or cyclic nucleotides, but also preference for certain ligands within one ligand group. Furthermore, investigation of the single binding sites enabled us to identify a new cAMP binding motif of PKA, expanding the canonical motif that was thought to be essential for cAMP binding [24].

We learnt that ligand prediction on base of sequence similarities can be misleading and needs to be confirmed by biochemical analysis. To date, we can distinguish with quite certainty between cNMP and nucleoside dependency of PKA, but further predictions of ligand selectivity are rather vague. Nevertheless, it is clear that cNMP dependency is not universal for PKA and predicted ligand specificity has to be critically scrutinized. Even single amino acid can change ligand specificity radically.

5.2 Outlook

Main focus of future investigations will lay on LdPKAR3, a new PKAR subtype whose activation mechanism might differ noticeably from the so far postulated model.
Therefore, future experiments will give attention to the holoenzyme formation between LdPKAR3 and catalytic subunits of *L. donovani* with a hopefully successful crystallization approach. Thereby, we hope to verify our hypothesis that LdPKAR3 is able to bind guanosine and a catalytic subunit at the same time and dissociation of R and C is not required for activation.
Moreover, ligand binding assays on the isolated site A of LdPKAR3 would be an interesting experiment in the future. With EgPKAR2, we characterized a regulatory subunit for which ligand binding could be proven in both isolated binding sites, but that structurally resembles LdPKAR3. It is therefore conceivable that ligand binding to site A in LdPKAR3 may not be possible for the full length protein, but very well for a truncation that leaves only site A. Furthermore, it has to be taken into consideration that other ligands, which were not tested in this work, are able to bind to site A.
Along with pursued *in vitro* studies, the role of LdPKAR3 in the life cycle of *L. donovani* will be elucidated. Therefore, kinase activation assays *in vivo* will be necessary, which may be realized due to high affinity binding of the cell permeable compounds 7CN and toyocamycin to LdPKAR3.

A second topic will be the extension of phylogenetic analysis of PKA in eukaryotes onto other phyla, as cAMP independency might have evolved more than once.
Furthermore, by means of a statistical coupling analysis (SCA), we hope to identify amino acid positions within protein sequence whose appearance is evolutionary linked. Since these co-evolving residues are often also functionally linked, we hope to identify a cluster of mutations responsible for the ligand switch in PKA that lead us a bit closer to the aim of accurate ligand prediction on base of sequence.

Bibliography

[1] B. L. Herwaldt. Leishmaniasis. *Lancet*, 354(9185):1191–9, 1999.

[2] Herbert Hof and Rüdiger Dörries. *Medizinische Mikrobiologie: Medizinisch relevante Protozoen,Leishmanien*. Thieme, 2017.

[3] Research priorities for Chagas disease, human African trypanosomiasis and leishmaniasis. *World Health Organ Tech Rep Ser*, (975):v–xii, 1–100, 2012.

[4] Leishmaniasis. `https://www.who.int/news-room/fact-sheets/detail/leishmaniasis`. Accessed: 2021-03-26.

[5] Sharon M Gossage, Matthew E Rogers, and Paul A Bates. Two separate growth phases during the development of Leishmania in sand flies: implications for understanding the life cycle. *International Journal for Parasitology*, 33(10):1027 – 1034, 2003.

[6] R.J.V. Pulvertaft and G.F. Hoyle. Stages in the life-cycle of leishmania donovani. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 54(2):191 – 196, 1960.

[7] Dan Zilberstein. Physiological and biochemical aspects of Leishmania development. *Leishmania after the genome*, pages 107–22, 2008.

[8] Efrat Barak, Sigal Amin-Spector, Elena Gerliak, Sophie Goyard, Neta Holland, and Dan Zilberstein. Differentiation of Leishmania donovani in host-free system: analysis of signal perception and response. *Molecular and Biochemical Parasitology*, 141(1):99 – 108, 2005.

[9] Polina Tsigankov, Pier Federico Gherardini, Manuela Helmer-Citterich, Gerald F. Spaeth, Peter J. Myler, and Dan Zilberstein. Regulation dynamics of Leishmania differentiation: deconvoluting signals and identifying phosphorylation trends. *Molecular & Cellular Proteomics*, 2014.

[10] S. Bachmaier, R. Witztum, P. Tsigankov, R. Koren, M. Boshart, and D. Zilberstein. Protein kinase A signaling during bidirectional axenic differentiation in Leishmania. *Int J Parasitol*, 46(2):75–82, 2016.

[11] A. Bhattacharya, A. Biswas, and P. K. Das. Identification of a protein kinase A regulatory subunit from Leishmania having importance in metacyclogenesis through induction of autophagy. *Mol Microbiol*, 83(3):548–64, 2012.

[12] J. F. Kuo and Paul Greengard. Cyclic nucleotide-dependent protein kinases, IV. widespread occurrence of adenosine 3,5-monophosphate-dependent protein kinase in various tissues and phyla of the animal kingdom. *Proceedings of the National Academy of Sciences*, 64(4):1349, 1969.

[13] Jonathan D. Graves and Edwin G. Krebs. Protein Phosphorylation and Signal Transduction. *Pharmacology and Therapeutics*, 82(2):111–121, 1999.

[14] S. S. Taylor, P. Zhang, J. M. Steichen, M. M. Keshwani, and A. P. Kornev. PKA: lessons learned after twenty years. *Biochim Biophys Acta*, 1834(7):1271–8, 2013.

[15] S. H. Francis and J. D. Corbin. Cyclic nucleotide-dependent protein kinases: intracellular receptors for cAMP and cGMP action. *Crit Rev Clin Lab Sci*, 36(4):275–328, 1999.

[16] B. S. Skalhegg and K. Tasken. Specificity in the cAMP/PKA signaling pathway. Differential expression,regulation, and subcellular localization of subunits of PKA. *Front Biosci*, 5:D678–93, 2000.

[17] W. J. Roesler, G. R. Vandenbark, and R. W. Hanson. Cyclic AMP and the induction of eukaryotic gene transcription. *J Biol Chem*, 263(19):9063–6, 1988.

[18] A. Y. Liu. Differentiation-specific increase of cAMP-dependent protein kinase in the 3T3-L1 cells. *J Biol Chem*, 257(1):298–306, 1982.

[19] E. G. Krebs and J. A. Beavo. Phosphorylation-dephosphorylation of enzymes. *Annu Rev Biochem*, 48:923–59, 1979.

[20] Susan S. Taylor and Alexandr P. Kornev. PKA and the Structural Kinome. *Periodicum Biologorum*, 118(4), 2017.

[21] Choel Kim, Cecilia Y. Cheng, S. Adrian Saldanha, and Susan S. Taylor. PKA-I Holoenzyme Structure Reveals a Mechanism for cAMP-Dependent Activation. *Cell*, 130(6):1032–1043, 2007.

[22] C. P. Ponting, L. Aravind, J. Schultz, P. Bork, and E. V. Koonin. Eukaryotic Signalling Domain Homologues in Archaea and Bacteria. Ancient Ancestry and Horizontal Gene Transfer. *Journal of Molecular Biology*, 289(4):729–745, 1999.

[23] Helen M. Berman, Lynn F. Ten Eyck, David S. Goodsell, Nina M. Haste, Alexandr Kornev, and Susan S. Taylor. The cAMP binding domain: An ancient signaling module. *Proceedings of the National Academy of Sciences of the United States of America*, 102(1):45–50, 2005.

[24] J. M. Canaves and S. S. Taylor. Classification and phylogenetic analysis of the cAMP-dependent protein kinase regulatory subunit family. *J Mol Evol*, 54(1):17–29, 2002.

[25] N. Kannan, J. Wu, G. S. Anand, S. Yooseph, A. F. Neuwald, J. C. Venter, and S. S. Taylor. Evolution of allostery in the cyclic nucleotide binding module. *Genome Biol*, 8(12):R264, 2007.

[26] Su, Y. and Dostmann, W. R. and Herberg, F. W. and Durick, K. and Xuong, N. H. and Ten Eyck, L. and Taylor, S. S. and Varughese, K. I. Regulatory subunit of protein kinase a: structure of deletion mutant with camp binding domains. *Science*, 269(5225):807, 1995.

[27] F. W. Herberg, S. S. Taylor, and W. R. Dostmann. Active site mutations define the pathway for the cooperative activation of cAMP-dependent protein kinase. *Biochemistry*, 35(9):2934–42, 1996.

[28] E. W. Sutherland and T. W. Rall. Fractionation and characterization of a cyclic adenine ribonucleotide formed by tissue particles. *J Biol Chem*, 232(2):1077–91, 1958.

[29] J. M. Gancedo. Biological roles of cAMP: variations on a theme in the different kingdoms of life. *Biol Rev Camb Philos Soc*, 88(3):645–68, 2013.

[30] N. M. El-Sayed, P. J. Myler, G. Blandin, M. Berriman, J. Crabtree, G. Aggarwal, E. Caler, H. Renauld, E. A. Worthey, C. Hertz-Fowler, E. Ghedin, C. Peacock, D. C. Bartholomeu, B. J. Haas, A. N. Tran, J. R. Wortman, U. C. Alsmark, S. Angiuoli, A. Anupama, J. Badger, F. Bringaud, E. Cadag, J. M. Carlton, G. C. Cerqueira, T. Creasy, A. L. Delcher, A. Djikeng, T. M. Embley, C. Hauser, A. C. Ivens, S. K. Kummerfeld, J. B. Pereira-Leal, D. Nilsson, J. Peterson, S. L. Salzberg, J. Shallom, J. C. Silva, J. Sundaram, S. Westenberger, O. White, S. E. Melville, J. E. Donelson, B. Andersson, K. D. Stuart, and N. Hall. Comparative genomics of trypanosomatid parasitic protozoa. *Science*, 309(5733):404–9, 2005.

[31] M. Parsons, E. A. Worthey, P. N. Ward, and J. C. Mottram. Comparative analysis of the kinomes of three pathogenic trypanosomatids: Leishmania major, Trypanosoma brucei and Trypanosoma cruzi. *BMC Genomics*, 6:127, 2005.

[32] E. Pays, P. Tebabi, A. Pays, H. Coquelet, P. Revelard, D. Salmon, and M. Steinert. The genes and transcripts of an antigen gene expression site from T. brucei. *Cell*, 57(5):835–45, 1989.

[33] D. Salmon, S. Bachmaier, C. Krumbholz, M. Kador, J. A. Gossmann, P. Uzureau, E. Pays, and M. Boshart. Cytokinesis of Trypanosoma brucei bloodstream forms depends on expression of adenylyl cyclases of the ESAG4 or ESAG4-like subfamily. *Mol Microbiol*, 84(2):225–42, 2012.

[34] M. A. Sanchez, D. Zeoli, E. M. Klamo, M. P. Kavanaugh, and S. M. Landfear. A family of putative receptor-adenylate cyclases from Leishmania donovani. *J Biol Chem*, 270(29):17551–8, 1995.

[35] M. K. Gould and H. P. de Koning. Cyclic-nucleotide signalling in protozoa. *FEMS Microbiol Rev*, 35(3):515–41, 2011.

[36] T. Seebeck, R. Schaub, and A. Johner. cAMP signalling in the kinetoplastid protozoa. *Curr Mol Med*, 4(6):585–99, 2004.

[37] D. Salmon, G. Vanwalleghem, Y. Morias, J. Denoeud, C. Krumbholz, F. Lhomme, S. Bachmaier, M. Kador, J. Gossmann, F. B. Dias, G. De Muylder, P. Uzureau, S. Magez, M. Moser, P. De Baetselier, J. Van Den Abbeele, A. Beschin, M. Boshart, and E. Pays. Adenylate cyclases of Trypanosoma brucei inhibit the innate immune response of the host. *Science*, 337(6093):463–6, 2012.

[38] M. Gonzales-Perdomo, P. Romero, and S. Goldenberg. Cyclic AMP and adenylate cyclase activators stimulate Trypanosoma cruzi differentiation. *Exp Parasitol*, 66(2):205–12, 1988.

[39] J. A. Beavo. Cyclic nucleotide phosphodiesterases: functional implications of multiple isoforms. *Physiol Rev*, 75(4):725–48, 1995.

[40] Y. Shakur, H. P. de Koning, H. Ke, J. Kambayashi, and T. Seebeck. Therapeutic potential of phosphodiesterase inhibitors in parasitic diseases. *Handb Exp Pharmacol*, (204):487–510, 2011.

[41] S. Kunz, J. A. Beavo, M. A. D'Angelo, M. M. Flawia, S. H. Francis, A. Johner, S. Laxman, M. Oberholzer, A. Rascon, Y. Shakur, L. Wentzinger, R. Zoraghi, and T. Seebeck. Cyclic nucleotide specific phosphodiesterases of the kinetoplastida: a unified nomenclature. *Mol Biochem Parasitol*, 145(1):133–5, 2006.

[42] M. Oberholzer, G. Marti, M. Baresic, S. Kunz, A. Hemphill, and T. Seebeck. The Trypanosoma brucei cAMP phosphodiesterases TbrPDEB1 and TbrPDEB2: flagellar enzymes that are essential for parasite virulence. *Faseb j*, 21(3):720–31, 2007.

[43] H. P. de Koning, M. K. Gould, G. J. Sterk, H. Tenor, S. Kunz, E. Luginbuehl, and T. Seebeck. Pharmacological validation of Trypanosoma brucei phosphodiesterases as novel drug targets. *J Infect Dis*, 206(2):229–37, 2012.

[44] M. K. Gould, S. Bachmaier, J. A. Ali, S. Alsford, D. N. Tagoe, J. C. Munday, A. C. Schnaufer, D. Horn, M. Boshart, and H. P. de Koning. Cyclic AMP effectors in African trypanosomes revealed by genome-scale RNA interference library screening for resistance to the phosphodiesterase inhibitor CpdA. *Antimicrob Agents Chemother*, 57(10):4882–93, 2013.

[45] S. Bachmaier, Y. Volpato Santos, S. Kramer, G. B. Githure, T. Klockner, J. Pepperl, C. Baums, R. Schenk, F. Schwede, H. G. Genieser, J. W. Dupuy, I. Forne, A. Imhof, J. Basquin, E. Lorentzen, and M. Boshart. Nucleoside analogue activators of cyclic AMP-independent protein kinase A of Trypanosoma. *Nat Commun*, 10(1):1421, 2019.

[46] R. D. Walter. Multiple protein kinases from Trypanosoma gambiense. *Hoppe Seylers Z Physiol Chem*, 359(5):601–6, 1978.

[47] T. Shalaby, M. Liniger, and T. Seebeck. The regulatory subunit of a cGMP-regulated protein kinase A of Trypanosoma brucei. *Eur J Biochem*, 268(23):6197–206, 2001.

[48] Susanne Kramer. Characterization of a PKA-like kinase from Trypanosoma brucei. April 2005.

[49] S. Laxman and J. A. Beavo. Cyclic nucleotide signaling mechanisms in trypanosomes: possible targets for therapeutic agents. *Mol Interv*, 7(4):203–15, 2007.

[50] Sabine Bachmaier and Michael Boshart. *Kinetoplastid AGC Kinases*, pages 99–122. 2013.

[51] H. Huang, L. M. Weiss, F. Nagajyothi, H. B. Tanowitz, M. Wittner, G. A. Orr, and Y. Bao. Molecular cloning and characterization of the protein kinase A regulatory subunit of Trypanosoma cruzi. *Mol Biochem Parasitol*, 149(2):242–5, 2006.

[52] M. Diskar, H. M. Zenn, A. Kaupisch, M. Kaufholz, S. Brockmeyer, D. Sohmen, M. Berrera, M. Zaccolo, M. Boshart, F. W. Herberg, and A. Prinz. Regulation of cAMP-dependent protein kinases: the human protein kinase X (PrKX) reveals the

role of the catalytic subunit alphaH-alphaI loop. *J Biol Chem*, 285(46):35910–8, 2010.

[53] S. Sen Santara, J. Roy, S. Mukherjee, M. Bose, R. Saha, and S. Adak. Globin-coupled heme containing oxygen sensor soluble adenylate cyclase in Leishmania prevents cell death during hypoxia. *Proc Natl Acad Sci U S A*, 110(42):16790–5, 2013.

[54] T. E. Ebenezer, M. Zoltner, A. Burrell, A. Nenarokova, A. M. G. Novak Vanclova, B. Prasad, P. Soukal, C. Santana-Molina, E. O'Neill, N. N. Nankissoor, N. Vadakedath, V. Daiker, S. Obado, S. Silva-Pereira, A. P. Jackson, D. P. Devos, J. Lukes, M. Lebert, S. Vaughan, V. Hampl, M. Carrington, M. L. Ginger, J. B. Dacks, S. Kelly, and M. C. Field. Transcriptome, proteome and draft genome of Euglena gracilis. *BMC Biol*, 17(1):11, 2019.

[55] Geoffrey Ian McFadden. PRIMARY AND SECONDARY ENDOSYMBIOSIS AND THE ORIGIN OF PLASTIDS. *Journal of Phycology*, 37(6):951–959, 2001.

[56] Patrick J. Keeling, Gertraud Burger, Dion G. Durnford, B. Franz Lang, Robert W. Lee, Ronald E. Pearlman, Andrew J. Roger, and Michael W. Gray. The tree of eukaryotes. *Trends in Ecology and Evolution*, 20(12):670 – 676, 2005.

[57] Julius Lukeš, Brian S. Leander, and Patrick J. Keeling. Cascades of convergent evolution: The corresponding evolutionary histories of euglenozoans and dinoflagellates. 106(Supplement 1):9963–9970, 2009.

[58] SUSANA A. BREGLIA, CLAUDIO H. SLAMOVITS, and BRIAN S. LEANDER. Phylogeny of Phagotrophic Euglenids (Euglenozoa) as Inferred from Hsp90 Gene Sequences. *Journal of Eukaryotic Microbiology*, 54(1):86–92, 2007.

[59] William Marande, Julius Lukeš, and Gertraud Burger. Unique Mitochondrial Genome Structure in Diplonemids, the Sister Group of Kinetoplastids. *Eukaryotic Cell*, 4(6):1137–1146, 2005.

[60] E. Yazaki, S. A. Ishikawa, K. Kume, A. Kumagai, T. Kamaishi, G. Tanifuji, T. Hashimoto, and Y. Inagaki. Global Kinetoplastea phylogeny inferred from a large-scale multigene alignment including parasitic species for better understanding transitions from a free-living to a parasitic lifestyle. *Genes Genet Syst*, 92(1):35–42, 2017.

[61] P. Flegontov, J. Votypka, T. Skalicky, M. D. Logacheva, A. A. Penin, G. Tanifuji, N. T. Onodera, A. S. Kondrashov, P. Volf, J. M. Archibald, and J. Lukes. Paratrypanosoma is a novel early-branching trypanosomatid. *Curr Biol*, 23(18):1787–93, 2013.

[62] J. Newman. Novel buffer systems for macromolecular crystallization. *Acta Crystallogr D Biol Crystallogr*, 60(Pt 3):610–2, 2004.

[63] Nadine Dyballa and Sabine Metzger. Fast and Sensitive Colloidal Coomassie G-250 Staining for Proteins in Polyacrylamide Gels. *Journal of Visualized Experiments*, (30), 2009.

[64] R. Lorenz, E. W. Moon, J. J. Kim, S. H. Schmidt, B. Sankaran, I. V. Pavlidis, C. Kim, and F. W. Herberg. Mutations of PKA cyclic nucleotide-binding domains reveal novel aspects of cyclic nucleotide selectivity. *Biochem J*, 474(14):2389–2403, 2017.

[65] J. Rinaldi, J. Wu, J. Yang, C. Y. Ralston, B. Sankaran, S. Moreno, and S. S. Taylor. Structure of yeast regulatory subunit: a glimpse into the evolution of PKA signaling. *Structure*, 18(11):1471–82, 2010.

[66] James Doutch, Michael A. Hough, S. Samar Hasnain, and Richard W. Strange. Challenges of sulfur SAD phasing as a routine method in macromolecular crystallography. *Journal of Synchrotron Radiation*, 19(1):19–29, Jan 2012.

[67] Thomas A. Halgren, Robert B. Murphy, Richard A. Friesner, Hege S. Beard, Leah L. Frye, W. Thomas Pollard, and Jay L. Banks. Glide : A New Approach for Rapid, Accurate Docking and Scoring 2. Enrichment Factors in Database Screening. *Journal of Medicinal Chemistry*, 47(7):1750–1759, 2004.

[68] Richard A. Friesner, Jay L. Banks, Robert B. Murphy, Thomas A. , Jasna J. Klicic, Daniel T. Mainz, Matthew P. Repasky, Eric H. Knoll, Mee Shelley, Jason K. Perry, David E. Shaw, Perry Francis, and Peter S. Shenkin. Glide : A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *Journal of Medicinal Chemistry*, 47(7):1739–1749, 2004.

[69] José Bubis, Juan Carlos Martínez, Maritza Calabokis, Joilyneth Ferreira, Carlos E. Sanz-Rodríguez, Victoria Navas, José Leonardo Escalona, Yurong Guo, and Susan S. Taylor. The gene product of a Trypanosoma equiperdum ortholog of the cAMP-dependent protein kinase regulatory subunit is a monomeric protein that is not capable of binding cyclic nucleotides. *Biochimie*, 146:166–180, 2018.

[70] C. Kim, N. H. Xuong, and S. S. Taylor. Crystal structure of a complex between the catalytic and regulatory (RIalpha) subunits of PKA. *Science*, 307(5710):690–6, 2005.

[71] A. P. Kornev, S. S. Taylor, and L. F. Ten Eyck. A generalized allosteric mechanism for cis-regulated cyclic nucleotide binding domains. *PLoS Comput Biol*, 4(4):e1000056, 2008.

[72] J. B. Shabb and J. D. Corbin. Cyclic nucleotide-binding domains in proteins having diverse functions. *J Biol Chem*, 267(9):5723–6, 1992.

[73] S. R. Rannels and J. D. Corbin. Two different intrachain cAMP binding sites of cAMP-dependent protein kinases. *J Biol Chem*, 255(15):7085–8, 1980.

[74] J. Wu, S. Brown, N. H. Xuong, and S. S. Taylor. RIalpha subunit of PKA: a cAMP-free structure reveals a hydrophobic capping mechanism for docking cAMP into site B. *Structure*, 12(6):1057–65, 2004.

[75] Dagfinn ØGRIED, Roald EKANGER, Robert H. SUVA, Jon P. MILLER, and Stein Ove DØSKELAND. Comparison of the two classes of binding sites (A and B) of type I and type II cyclic-AMP-dependent protein kinases by using cyclic nucleotide analogs. *European Journal of Biochemistry*, 181(1):19–31, 1989.

[76] E. Wang, H. Sun, J. Wang, Z. Wang, H. Liu, J. Z. H. Zhang, and T. Hou. End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: Strategies and Applications in Drug Design. *Chem Rev*, 119(16):9478–9508, 2019.

[77] M. Akimoto, R. Selvaratnam, E. T. McNicholl, G. Verma, S. S. Taylor, and G. Melacini. Signaling through dynamic linkers as revealed by PKA. *Proc Natl Acad Sci U S A*, 110(35):14231–6, 2013.

[78] J. G. Bruystens, J. Wu, A. Fortezzo, A. P. Kornev, D. K. Blumenthal, and S. S. Taylor. PKA RIalpha homodimer structure reveals an intermolecular interface with implications for cooperative cAMP binding and Carney complex disease. *Structure*, 22(1):59–69, 2014.

[79] C. Banerjee and D. Sarkar. The cAMP-binding proteins of Leishmania are not the regulatory subunits of cAMP-dependent protein kinase. *Comp Biochem Physiol B Biochem Mol Biol*, 130(2):217–26, 2001.

[80] M. H. el Kouni. Potential chemotherapeutic targets in the purine metabolism of parasites. *Pharmacol Ther*, 99(3):283–309, 2003.

[81] T. D. Serafim, A. B. Figueiredo, P. A. Costa, E. A. Marques-da Silva, R. Gonçalves, S. A. de Moura, N. F. Gontijo, S. M. da Silva, M. S. Michalick, J. R. Meyer-Fernandes, R. P. de Carvalho, S. R. Uliana, J. L. Fietto, and L. C. Afonso. Leishmania metacyclogenesis is promoted in the absence of purines. *PLoS Negl Trop Dis*, 6(9):e1833, 2012.

[82] D. Rosenzweig, D. Smith, P. J. Myler, R. W. Olafson, and D. Zilberstein. Post-translational modification of cellular proteins during Leishmania donovani differentiation. *Proteomics*, 8(9):1843–50, 2008.

[83] P. Tsigankov, P. F. Gherardini, M. Helmer-Citterich, G. F. Spath, and D. Zilberstein. Phosphoproteomic analysis of differentiating Leishmania parasites reveals a unique stage-specific phosphorylation motif. *J Proteome Res*, 12(7):3405–12, 2013.

[84] P. Banky, L. J. Huang, and S. S. Taylor. Dimerization/docking domain of the type Ialpha regulatory subunit of cAMP-dependent protein kinase. Requirements for dimerization and docking are distinct but overlapping. *J Biol Chem*, 273(52):35048–55, 1998.

[85] M. Colledge and J. D. Scott. AKAPs: from structure to function. *Trends Cell Biol*, 9(6):216–21, 1999.

[86] L. K. Langeberg and J. D. Scott. A-kinase-anchoring proteins. *J Cell Sci*, 118(Pt 15):3217–20, 2005.

[87] S. S. Taylor, C. Kim, C. Y. Cheng, S. H. Brown, J. Wu, and N. Kannan. Signaling through cAMP and cAMP-dependent protein kinase: diverse strategies for drug design. *Biochim Biophys Acta*, 1784(1):16–26, 2008.

[88] K. J. Haushalter, D. E. Casteel, A. Raffeiner, E. Stefan, H. H. Patel, and S. S. Taylor. Phosphorylation of protein kinase A (PKA) regulatory subunit RIalpha by protein kinase G (PKG) primes PKA for catalytic activity in cells. *J Biol Chem*, 293(12):4411–4421, 2018.

[89] Nina M. Haste, Hana Talabani, Alex Doo, Anais Merckx, and Susan S. Langsley, Gordon and. Exploring the Plasmodium falciparum cyclic-adenosine monophosphate (cAMP)-dependent protein kinase (PfPKA) as a therapeutic target. *Microbes and Infection*, 14(10):838–850, 2012.

[90] M. Zorn, K. E. Fladmark, D. Ogreid, B. Jastorff, S. O. Doskeland, and W. R. Dostmann. Ala335 is essential for high-affinity cAMP-binding of both sites A and B of cAMP-dependent protein kinase type I. *FEBS Lett*, 362(3):291–4, 1995.

[91] M. P. Williamson. The structure and function of proline-rich regions in proteins. *Biochemical Journal*, 297(Pt 2):249–260, 1994.

[92] Jörg Isensee, Melanie Kaufholz, Matthias J. Knape, Jan Hasenauer, Hanna Hammerich, Humberto Gonczarowska-Jorge, René P. Zahedi, Frank Schwede, Friedrich W. Herberg, and Tim Hucho. PKA-RII subunit phosphorylation precedes

activation by cAMP and regulates activity termination. *The Journal of Cell Biology*, 217(6):2167–2184, 2018.

[93] M. G. Gold. Swimming regulations for protein kinase A catalytic subunit. *Biochem Soc Trans*, 47(5):1355–1366, 2019.

[94] Robert L. Geahlen, David F. Carmichael, Eikichi Hashimoto, and Edwin G. Krebs. Phosphorylation of cAMP-dependent protein kinase subunits. *Advances in Enzyme Regulation*, 20:195 – 209, 1982. Advances in Enzyme Regulation.

[95] D. R. Littler, H. E. Bullen, K. L. Harvey, T. Beddoe, B. S. Crabb, J. Rossjohn, and P. R. Gilson. Disrupting the Allosteric Interaction between the Plasmodium falciparum cAMP-dependent Kinase and Its Regulatory Subunit. *J Biol Chem*, 291(49):25375–25386, 2016.

[96] E. M. Reimann, D. A. Walsh, and E. G. Krebs. Purification and properties of rabbit skeletal muscle adenosine 3',5'-monophosphate-dependent protein kinases. *J Biol Chem*, 246(7):1986–95, 1971.

[97] Mandy Diskar, Hans-Michael Zenn, Alexandra Kaupisch, Anke Prinz, and Friedrich W. Herberg. Molecular basis for isoform-specific autoregulation of protein kinase A. *Cellular Signalling*, 19(10):2024 – 2034, 2007.

[98] Mao Peng, Thin Thin Aye, Berend Snel, Bas van Breukelen, Arjen Scholten, and Albert J. R. Heck. Spatial Organization in Protein Kinase A Signaling Emerged at the Base of Animal Evolution. *Journal of Proteome Research*, 14(7):2976–2987, 2015.

[99] Tomáš Skalický, Eva Dobáková, Richard J. Wheeler, Martina Tesařová, Pavel Flegontov, Dagmar Jirsová, Jan Votýpka, Vyacheslav Yurchenko, Francisco J. Ayala, and Julius Lukeš. Extensive flagellar remodeling during the complex life cycle of Paratrypanosoma, an early-branching trypanosomatid. *Proceedings of the National Academy of Sciences of the United States of America*, 114(44):11757–11762, 2017.

[100] A. V. Jäger, J. G. De Gaudenzi, J. G. Mild, B. Mc Cormack, S. Pantano, D. L. Altschuler, and M. M. Edreira. Identification of novel cyclic nucleotide binding proteins in Trypanosoma cruzi. *Mol Biochem Parasitol*, 198(2):104–12, 2014.

[101] Amrita Saha, Moumita Basu, and Anindita Ukil. Recent advances in understanding Leishmania donovani infection: The importance of diverse host regulatory pathways. *IUBMB Life*, 70(7):593–601, 2018.

[102] Keith R. Matthews, James R. Ellis, and Athina Paterou. Molecular regulation of the life cycle of African trypanosomes. *Trends in Parasitology*, 20(1):40–47, 2004.

[103] Andrew P Jackson, Thomas D Otto, Martin Aslett, Stuart D Armstrong, Frederic Bringaud, Alexander Schlacht, Catherine Hartley, Mandy Sanders, Jonathan M Wastling, Joel B Dacks, Alvaro Acosta-Serrano, Mark C Field, Michael L Ginger, and Matthew Berriman. Kinetoplastid Phylogenomics Reveals the Evolutionary Innovations Associated with the Origins of Parasitism. *Current Biology*, 26(2):161–172, 2016.

[104] A. Butenko, M. Hammond, M. C. Field, M. L. Ginger, V. Yurchenko, and J. Lukes. Reductionist Pathways for Parasitism in Euglenozoans? Expanded Datasets Provide New Insights. *Trends Parasitol*, 37(2):100–116, 2021.

[105] Anzhelika Butenko, Fred R. Opperdoes, Olga Flegontova, Aleš Horák, Vladimír Hampl, Patrick Keeling, Ryan M. R. Gawryluk, Denis Tikhonenkov, Pavel Flegontov, and Julius Lukeš. Evolution of metabolic capabilities and molecular features of diplonemids, kinetoplastids, and euglenids. *BMC Biology*, 18(1):23, 2020.

[106] Jian Wu, Simon H. J. Brown, Sventja von Daake, and Susan S. Taylor. PKA Type IIa Holoenzyme Reveals a Combinatorial Strategy for Isoform Diversity. *Science*, 318(5848):274, 2007.

[107] S. H. Brown, J. Wu, C. Kim, K. Alberto, and S. S. Taylor. Novel isoform-specific interfaces revealed by PKA RIIbeta holoenzyme structures. *J Mol Biol*, 393(5):1070–82, 2009.

[108] R. C. Edgar. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, 5:113, 2004.

[109] Andrew M. Waterhouse, James B. Procter, David M. A. Martin, Michèle Clamp, and Geoffrey J. Barton. Jalview Version 2–a multiple sequence alignment editor and analysis workbench. *Bioinformatics (Oxford, England)*, 25(9):1189–1191, 2009.

Supplementary material

1 Supplemental material - Chapter 2.2

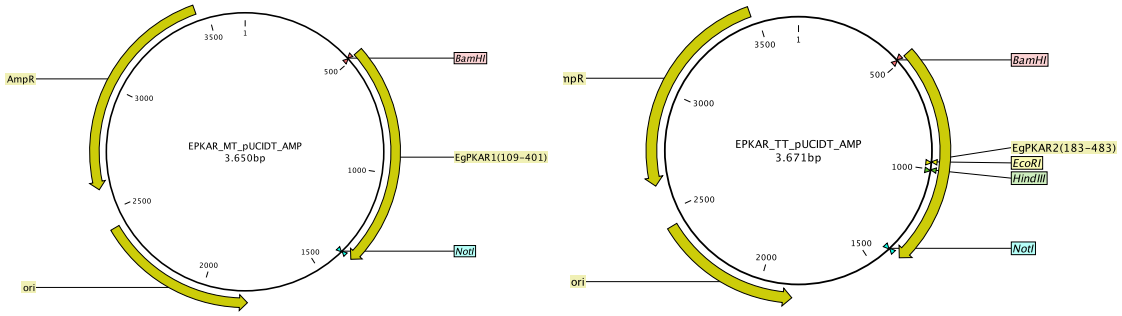


Figure S1: Charts of pUCIDT_AMP_EgPKAR1(109-401) (left) and pUCIDT_AMP_EgPKAR2(183-483) (right)

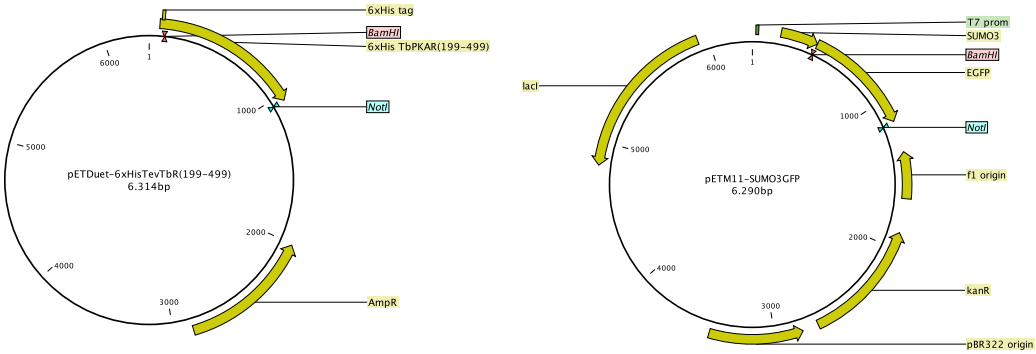


Figure S2: Charts of pET11_6His_Sumo3_eGFP (left) and pET_Duet_TbPKAR(199-499) (right). Other constructs in this work have been inserted into a pET_Sumo vector

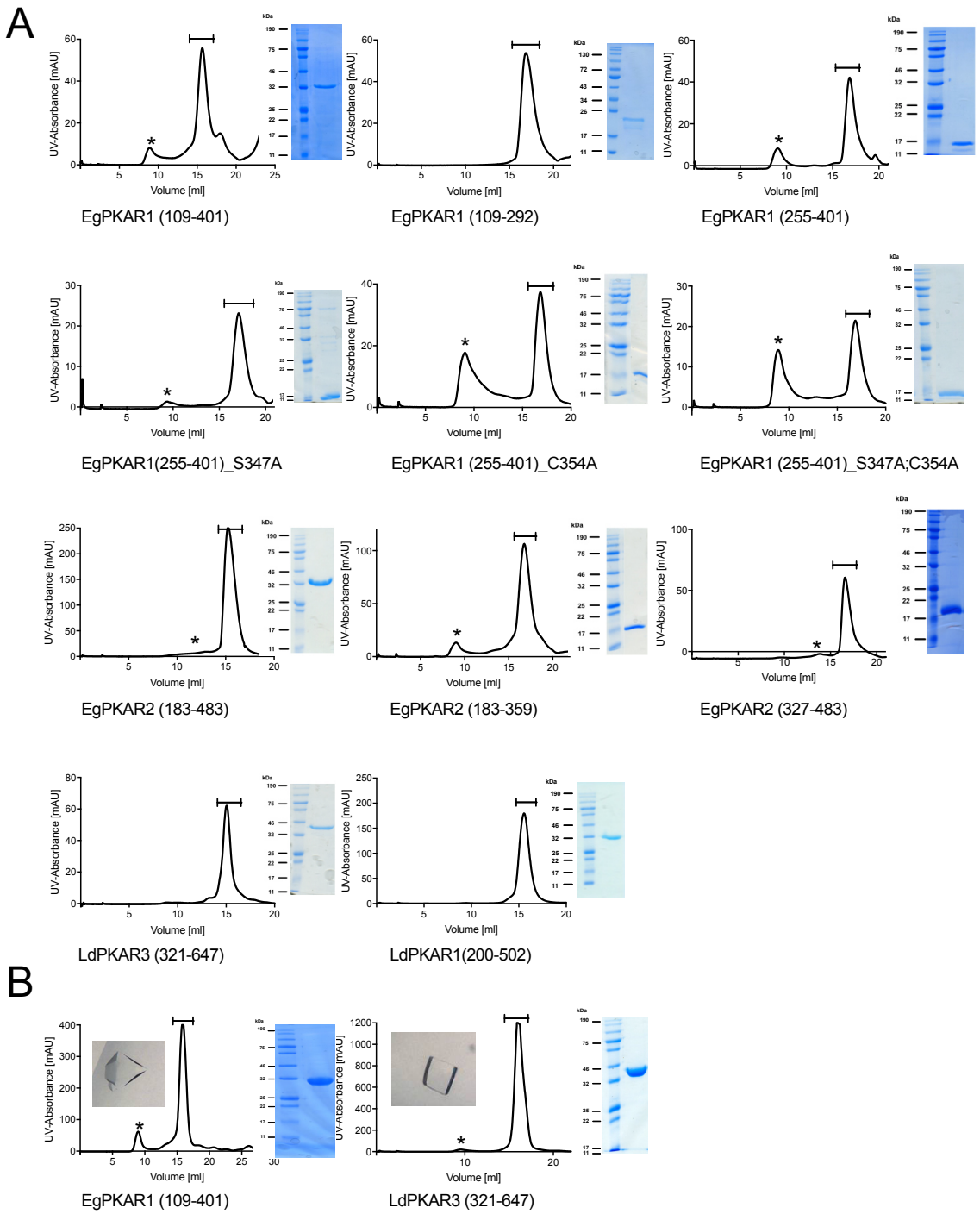


Figure S3: Representative curves of (A) denatured and subsequently refolded proteins for ITC-experiments and (B) native protein for crystallization were purified using gel filtration chromatography. Lines above the respective peaks mark collected fractions, which were confirmed to be pure by SDS PAGE (right). (*)-labelled peaks are caused by aggregates of unfolded protein. Pictures of crystallized EgPKAR1 and LdPKAR3 are shown as inserts in the respective chromatogram

2 Supplemental material - Chapter 3.1

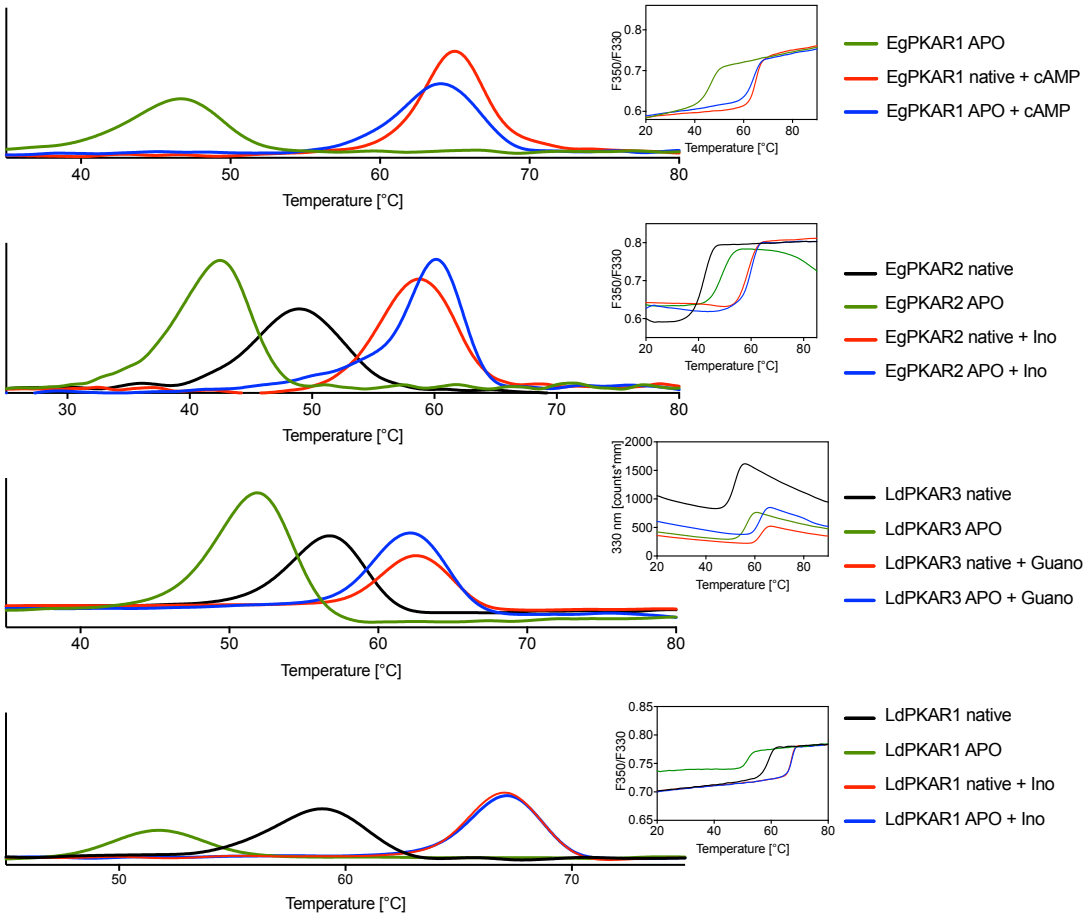


Figure S4: Denatured and subsequently refolded protein (APO) was compared to corresponding native, folded, protein (black) using nDSF. Incubation with ligand is indicated according to the colour code on the right. Corresponding melting temperatures, T_m, of each protein sample is listed in Table S1. For LdPKAR3 (3rd plot), fluorescentic emission at λ = 330nm was monitored instead of the ratio of fluorescentic emission at 350 and 330 nm

Table S1: Melting temperatures, T_m, of native and refolded protein in pre- and absence of ligands are compared by nDSF. In addition of 1mM ligand, similar T_m for native and refolded protein was obtained, indicating accurate refolding of the protein

Protein	State and Ligand	T_m [°C]
EgPKAR1	Apo, ligand free	46.6
	APO+cAMP	64.1
	Native+cAMP	65.0
EgPKAR2	Apo, ligand free	42.2
	Native	48.9
	APO+Inosine	60.0
	Native+Inosine	58.8
LdPKAR3	APO, Ligand free	52.1
	Native	56.6
	APO+Guanosine	62.2
	Native+Guanosine	62.5
LdPKAR1	APO, ligand free	51.8
	Native	58.8
	APO+Inosine	67
	Native+Inosine	66.9

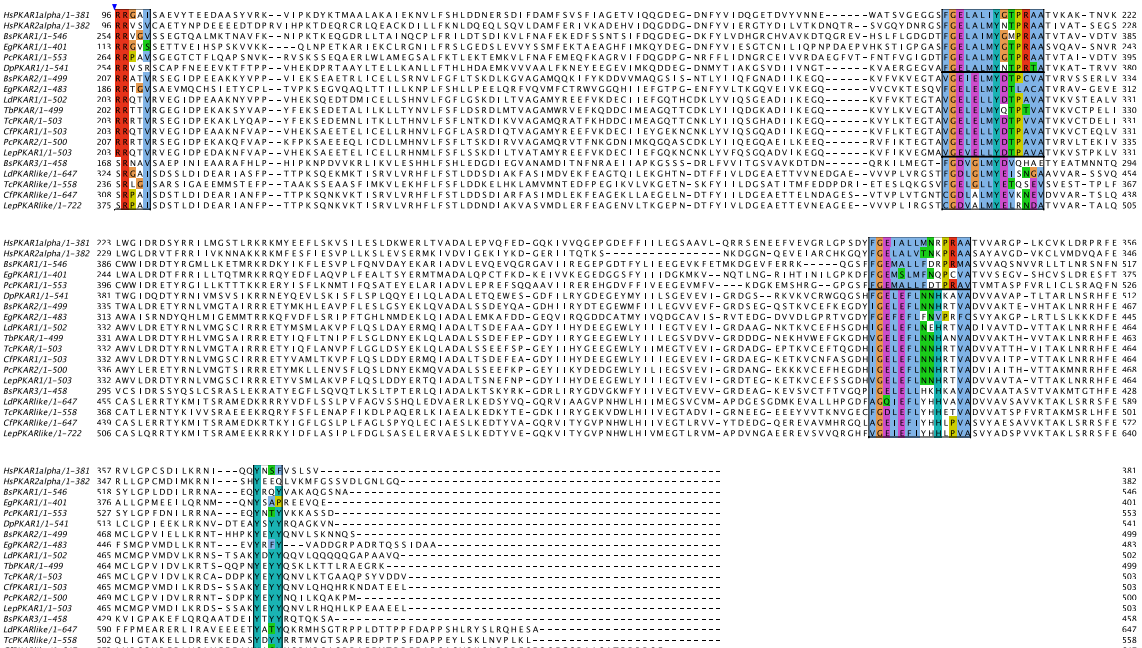


Figure S5: Multiple sequence alignment of PKAR in representatives of the *Euglenozoa* using MUSCLE integrated as webservice in Jalview2.10.4b1 [108] [109]. Highlighted segments, from N- to C-terminus, are Inhibitory sequence, PBC:A, PBC:B, and C-terminal capping motif. Positively charged amino acids are coloured in red, negatively charged in purple. Aliphatic amino acids are shown in blue, besides glycines that are coloured in orange. Polar, but uncharged side chains are coloured in light green except for prolines (yellow). Aromatic residues and histidine are highlighted in dark cyan.

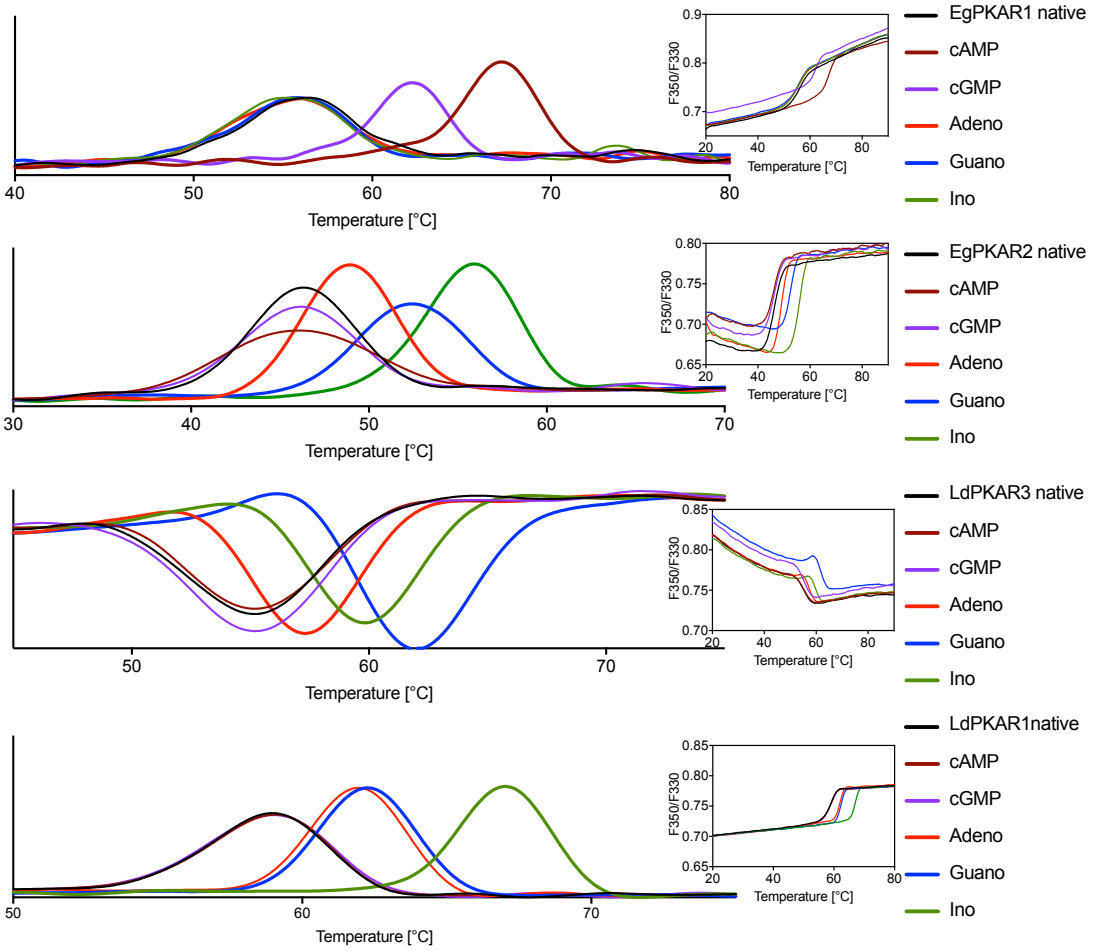


Figure S6: Melting curves were obtained for EgPKAR1(109-401), EgPKAR2(183-483), LdPKAR3(321-647) and LdPKAR1(200-502) from nDSF experiments. Graphs in the inset display the fluorescent emission of the protein as a function of temperature. Main plots show the first derivative of the corresponding graphs in the inset. Minima or maxima of the derivatives define the melting temperature T_m. Protein were fused to a SUMO3-Tag and purified one time. Addition of potential ligand candidates (c=1mM) is indicated on the right

Table S2: Melting temperatures, T_m, of native protein incubated with 1mM of listed ligands were obtained from nDSF experiments. For ΔT_m [°C], the difference of T_m of native protein and T_m of the same protein in addition of small molecules was calculated

Protein	Ligand	T_m [°C]	ΔT_m [°C]
EgPKAR1	**Native**	**56.1**	**-**
	cAMP	67.2	11.1
	cGMP	62.2	6.1
	Inosine	55.4	-0.7
	Guanosine	55.8	-0.3
	Adenosine	55.6	-0.5
EgPKAR2	**Native**	**46.3**	**-**
	cAMP	46.1	-0.2
	cGMP	46.2	-0.1
	Inosine	55.9	9.6
	Guanosine	52.4	6.1
	Adenosine	48.9	2.6
LdPKAR3	**Native**	**55.1**	**-**
	cAMP	55.1	0
	cGMP	55.1	0
	Inosine	59.8	4.7
	Guanosine	62.1	7.0
	Adenosine	57.3	2.2
LdPKAR1	**Native**	**58.8**	**-**
	cAMP	58.8	0
	cGMP	58.9	0.1
	Inosine	66.9	12.1
	Guanosine	62.2	3.4
	Adenosine	61.8	3.0

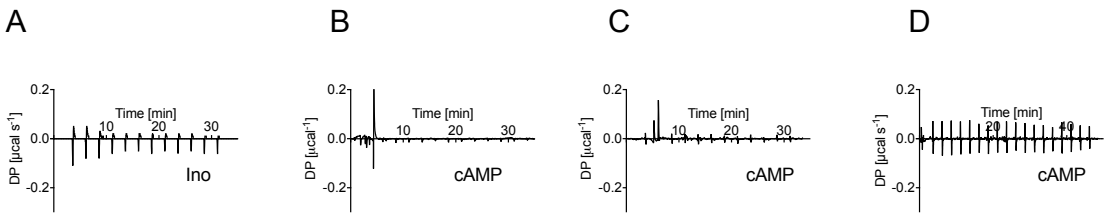


Figure S7: Negative binding results for (A) EgPKAR1(109-401), (B) EgPKAR2(183-483), (C) LdPKAR3(321-647) and (D) LdPKAR1(200-502) were obtained from ITC experiments.Measurements were done in triplicates from at least two independent protein preparations

3 Supplemental material - Chapter 3.3

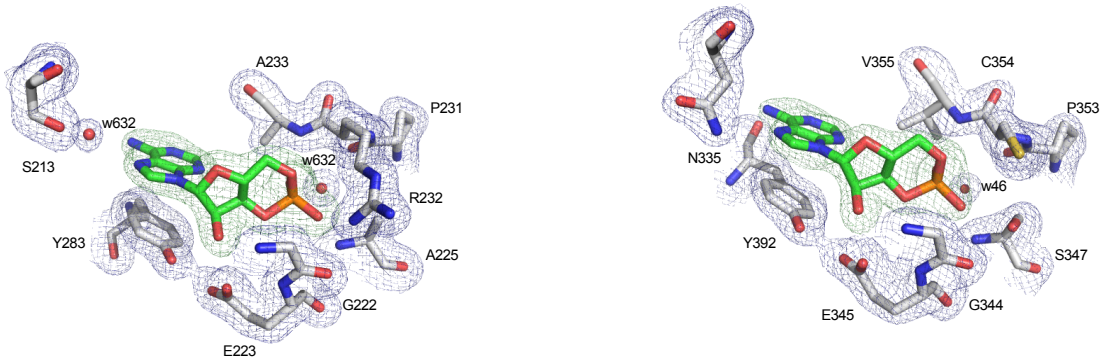


Figure S8: Electron density maps for EgPKAR1 CNB:A and B. The molecular model of the protein was built inside the 2Fo-Fc map (1σ, blue) electron density map whereas the ligand was built into the omit map (Fo-Fc, 3σ, green). Waters are indicated as red spheres

Acknowledgement

At this point I would like to thank every college, friend and member of my family, who supported me during the time of my MD thesis.
In particular, I would like to give a special thanks to

Prof. Dr. Michael Boshart for giving me the opportunity to write me thesis in such an interesting research field, for his constant support and the many productive discussions,

Prof. Andreas Ladurner, PhD, for the kind support and his valuable suggestions regarding troubleshooting of ITC experiments

Yuri Volpato Santos who supervised my work in the lab, taught me all his methods and answered every question patiently,

our collaborators, Okke Melse and Prof. Dr. Iris Antes, for modelling the structures of EgPKAR2 and DpPKAR1 and the helpful support on various questions in regard to our structural data,

Prof. Dr. Michaela Smolle for giving me the opportunity to do ITC experiments at Biophysics core facility of the department for physiological chemistry at BMC and

Prof. Dr. Dan Zilberstein, Technion Israel, Haifa, for kindly sharing unpublished information with us.

I am very grateful to my parents, as well as Tina, Verena and Magnus for accompanying and supporting me the whole time,

and to all members of the AG Boshart who not only contributed to this thesis with productive discussions but also by generating a great working environment that made it fun to come here every day.

Affidavit

Eidesstattliche Versicherung

Ober, Veronica Teresa

Name, Vorname

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Titel:

Evolution of ligand specificity of protein kinase A isoforms in the phylum Euglenozoa

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

München, 25.02.2022	Veronica Teresa Ober
Ort, Datum	Unterschrift Doktorandin bzw. Doktorand