Durham Research Online

Deposited in DRO:
06 February 2018

Version of attached file:

Peer-review status of attached file:
Peer-reviewed

Citation for published item:

Further information on publisher’s website:
https://doi.org/10.1017/s0031182017001901

Publisher’s copyright statement:
© Cambridge University Press 2017. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a link is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the full DRO policy for further details.
INTRODUCTION

The phylum of Apicomplexa contains approximately 6000 unicellular, eukaryotic parasites including *Plasmodium* spp., the causative agent of Malaria, and *Toxoplasma gondii*, responsible for toxoplasmosis. The complex life cycle of *T. gondii* is regulated by a unique family of calcium-dependent protein kinases (CDPKs) that have become the target of intensive efforts to develop new therapeutics. In this review, we will summarize structure-based strategies, recent successes and future directions in the pursuit of specific and selective inhibitors of *T. gondii* CDPK1.

Keywords: Calcium-dependent protein kinase, CDPK1, *Toxoplasma gondii*, drug design, protein structure.

SPECIAL ISSUE REVIEW

The calcium-dependent protein kinase 1 from *Toxoplasma gondii* as target for structure-based drug design

EMILY M. CARDEW¹, CHRISTOPHE L. M. J. VERLINDE² and EHMKE POHL¹,³,⁴*

¹ Department of Biosciences, Durham University, Lower Mountjoy Durham DH1 3LE, UK
² Department of Biochemistry, University of Washington, Seattle, Washington, WA 98195, USA
³ Department of Chemistry, Durham University, South Road, Durham DH1 3LE, UK
⁴ Biophysical Sciences Institute, Durham University, Durham DH1 3LE, UK

(Received 31 May 2017; revised 11 August 2017; accepted 25 August 2017)

SUMMARY

The apicomplexan protozoan parasites include the causative agents of animal and human diseases ranging from malaria (*Plasmodium* spp.) to toxoplasmosis (*Toxoplasma gondii*). The complex life cycle of *T. gondii* is regulated by a unique family of calcium-dependent protein kinases (CDPKs) that have become the target of intensive efforts to develop new therapeutics. In this review, we will summarize structure-based strategies, recent successes and future directions in the pursuit of specific and selective inhibitors of *T. gondii* CDPK1.

THE ROLE OF CALCIUM-DEPENDENT PROTEIN KINASES (CDPKS)

In *T. gondii* Ca²⁺-ions play key roles in cell signalling and in pathogen–host interactions including cell invasion, motility of the parasite within the host and differentiation during the parasites complex life cycle (Irvine, 1986; Nagamune et al. 2009; Verlinde et al. 2009; Groftehauge et al. 2015; Hol, 2015; Muller, 2017). CDPKs are a family of serine/threonine kinases that are only found in plants and protists including ciliates and apicomplexan parasites. Importantly, these kinases provide the mechanistic link between calcium signalling and motility, differentiation and invasion (Tzen et al. 2007; Billker et al. 2009). These crucial roles of CDPKs have been proven through a range of knock-out studies in various species and underline the potential of CDPKs as targets for a handful of antimicrobials such as sulphonamides, folic acid derivatives and certain macrolide antibiotics. However, these drugs often show limited efficacy and are associated with significant side effects (Alday and Doggett, 2017). Furthermore, there are no treatments available to target tissue cysts, the persistent form in which the parasite evades the host immune system, and to eradicate persistent *T. gondii* infections (Opsteegh et al. 2015). Therefore, new drug targets and novel therapies are urgently needed. In addition to high-throughput screening approaches (Norcliffe et al. 2014), structure-based methods in close combination with medicinal chemistry and biophysical and biological validation have become powerful tools in the search of new drugs and treatments (Hunter, 2009; Verlinde et al. 2009; Groftehauge et al. 2015; Hol, 2015; Muller, 2017).

* Corresponding author: Department of Biosciences, Durham University, Lower Mountjoy Durham DH1 3LE, UK, E-mail: ehmke.pohl@durham.ac.uk

Parasitology (2018), 145, 210–218. © Cambridge University Press 2017. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

doi:10.1017/S0031182017001901

CrossMark

Downloaded from https://www.cambridge.org/core. University of Cambridge Library, on 06 Feb 2018 at 16:27:14, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0031182017001901
novel therapeutics (Long et al. 2016). CDPKs are members of the Calmodulin/Calcium kinase (CaM) family. They all share an N-terminal kinase domain (KD) linked via a junctional domain to a series of C-terminal Calcium-binding motifs. In T. gondii at least 12 different CDPKs have been putatively identified ranging in size from 507 (CDPK1) to more than 2000 amino acids (CDPK7, CDPK80) (Morlon-Guyot et al. 2014). Although there is probably overlap in functionalities, different sub-cellular locations and varying expression levels during the parasites’ life cycle is likely to lead to different biological functions within the CDPK family (Hui et al. 2015). The shared sequence identities range from 51% (CDPK1 and CDPK3) (Treeck et al. 2014) to lower than 10% (Table 1). These variations in length and sequence support the notion that members of the CDPK family act upon a range of substrates and fulfil different functions in T. gondii biology. Recent knock-out studies using CRISPR-Cas9 indicate that CDPK4, CDPK5, CDPK6, CDPK8 and CDPK9, respectively, have no effect on virulence and on normal growth (Wang et al. 2016), however, knock-down studies have shown that CDPK7 is crucial for survival due to a critical role in parasite division (Morlon-Guyot et al. 2014). More detailed studies have been performed on the smaller family members. CDPK3 with 537 amino acids has been implicated in motility and host cell egress (McCoy et al. 2017). CDPK2 (711 amino acids) has been shown to act as key regulator of amylopectin metabolism (Uboldi et al. 2015). The loss of CDPK2 results in the build-up of amylum with fatal consequences for T. gondii in its chronic stage. Importantly, this family member contains an N-terminal carbohydrate-binding domain that may offer new opportunities for drug design (Uboldi et al. 2015). CDPK1, which is mainly located in the cytosol, has been shown to be required for the microneme secretion at the apical complex and parasite proliferation. The molecular mechanism, however, remains elusive (Lourido et al. 2010; Child et al. 2017). Here we will review strategies and recent results in the discovery, design and potency of inhibitors targeting the KD of CDPK1 from T. gondii (TgCDPK1).

**Activation of TgCDPK1 by Calcium**

The mechanism of activation and inhibition was unravelled in 2010 when the crystal structures of both the auto-inhibited and the Ca\(^{2+}\)-activated forms of TgCDPK1 were published (Ojo et al. 2010; Wernimont et al. 2010). These structures revealed the expected KD in similar overall conformations, however, the Ca\(^{2+}\)-binding domain (also designated CPDK activating domain or CAD) adopted two vastly different conformations and orientations (Fig. 1). In its inactive state the CAD

---

**Table 1. The protein sequence identities between the 12 putative calcium-dependent protein kinases (CDPKs) of T. gondii, rounded to the nearest whole number, derived from a multiple sequence alignment (MSA) generated using Clustal Omega (Sievers et al. 2011).**

<table>
<thead>
<tr>
<th>CDPK1</th>
<th>CDPK2</th>
<th>CDPK3</th>
<th>CDPK2A</th>
<th>CDPK2B</th>
<th>CDPK3</th>
<th>CDPK4</th>
<th>CDPK4A</th>
<th>CDPK5</th>
<th>CDPK6</th>
<th>CDPK7</th>
<th>CDPK8</th>
<th>CDPK9</th>
</tr>
</thead>
<tbody>
<tr>
<td>22%</td>
<td>24%</td>
<td>30%</td>
<td>24%</td>
<td>14%</td>
<td>24%</td>
<td>25%</td>
<td>14%</td>
<td>25%</td>
<td>22%</td>
<td>14%</td>
<td>24%</td>
<td>25%</td>
</tr>
<tr>
<td>23%</td>
<td>28%</td>
<td>32%</td>
<td>25%</td>
<td>17%</td>
<td>25%</td>
<td>28%</td>
<td>17%</td>
<td>27%</td>
<td>23%</td>
<td>17%</td>
<td>26%</td>
<td>27%</td>
</tr>
<tr>
<td>30%</td>
<td>37%</td>
<td>37%</td>
<td>33%</td>
<td>20%</td>
<td>30%</td>
<td>37%</td>
<td>20%</td>
<td>36%</td>
<td>30%</td>
<td>20%</td>
<td>33%</td>
<td>31%</td>
</tr>
<tr>
<td>51%</td>
<td>55%</td>
<td>55%</td>
<td>53%</td>
<td>36%</td>
<td>55%</td>
<td>55%</td>
<td>36%</td>
<td>55%</td>
<td>55%</td>
<td>36%</td>
<td>55%</td>
<td>55%</td>
</tr>
<tr>
<td>14%</td>
<td>14%</td>
<td>14%</td>
<td>14%</td>
<td>14%</td>
<td>14%</td>
<td>14%</td>
<td>14%</td>
<td>14%</td>
<td>14%</td>
<td>14%</td>
<td>14%</td>
<td>14%</td>
</tr>
<tr>
<td>24%</td>
<td>21%</td>
<td>20%</td>
<td>25%</td>
<td>25%</td>
<td>21%</td>
<td>20%</td>
<td>25%</td>
<td>25%</td>
<td>21%</td>
<td>20%</td>
<td>25%</td>
<td>25%</td>
</tr>
<tr>
<td>25%</td>
<td>32%</td>
<td>32%</td>
<td>36%</td>
<td>30%</td>
<td>32%</td>
<td>32%</td>
<td>30%</td>
<td>32%</td>
<td>32%</td>
<td>30%</td>
<td>32%</td>
<td>30%</td>
</tr>
<tr>
<td>22%</td>
<td>26%</td>
<td>24%</td>
<td>26%</td>
<td>25%</td>
<td>24%</td>
<td>26%</td>
<td>25%</td>
<td>24%</td>
<td>26%</td>
<td>25%</td>
<td>24%</td>
<td>26%</td>
</tr>
<tr>
<td>7%</td>
<td>7%</td>
<td>7%</td>
<td>7%</td>
<td>7%</td>
<td>7%</td>
<td>7%</td>
<td>7%</td>
<td>7%</td>
<td>7%</td>
<td>7%</td>
<td>7%</td>
<td>7%</td>
</tr>
<tr>
<td>6%</td>
<td>9%</td>
<td>9%</td>
<td>8%</td>
<td>8%</td>
<td>9%</td>
<td>8%</td>
<td>8%</td>
<td>8%</td>
<td>8%</td>
<td>8%</td>
<td>8%</td>
<td>8%</td>
</tr>
</tbody>
</table>
(shown in raspberry red) adopts an elongated form reminiscent of apo-calmodulin starting with a long helix followed by the first Ca$^{2+}$-binding motifs (EF-hands), which is connected via another long helix to the second pair of C-terminal EF-hands (Fig. 1a). The first long helix has been suggested to be responsible for the auto-inhibitory effect by blocking the substrate binding site and providing a basic lysine residue to bind a cluster of conserved acidic residues. However, this may not be the only mechanism of deactivation as it has more recently been shown that removal of the regulatory domain alone does not lead to an active KD (Ingram et al. 2015). The CAD domain activated by Ca$^{2+}$-binding appears to be required to maintain the KD in its active conformation. Calcium binding leads to a dramatic rearrangement and refolding of the protein chain (Fig. 1b) (Wernimont et al. 2010). The entire regulatory domain is shifted to the other side of the protein hence liberating the active site of the KD as shown in Fig. 1c. In addition, the regulatory calcium-binding domain is collapsed so that the two long helices are no longer arranged in an anti-parallel fashion but are partially unwound and interwoven to form a more globular overall shape. These structural changes are reminiscent to the calcium-bound structure of calmodulin (Kursula, 2014).

Fig. 1. Ribbon representation of the crystal structure of CDPK1 from T. gondii with the kinase domain depicted in cyan, the regulatory domain in raspberry red. (a) CDPK1 in its inactive auto-inhibited state (PDB code: 3KU2) (Wernimont et al. 2010). (b) CDPK1 in its calcium-bound, activated state with the Ca$^{2+}$-ions shown as green spheres and the non-hydrolysable ligand ANP in stick representation (PDB code: 3HX4) (Wernimont et al. 2010). (c) Ribbon diagram of the least-squares superposition of the inactive and active forms of TgCDPK1 with the kinase domains shown in cyan (active) and blue (inactive), the regulatory domain in shades of red, respectively. Only the kinase domain was used to calculate the transformation matrix, which was then applied to the entire protein chain. CDPKs, calcium-dependent protein kinases.
COMPARISON WITH HUMAN KINASES

Historically, characterising (protozoan) kinases as potential drug targets and developing selective inhibitors has been considered challenging due to the fact that the overall protein fold and the active sites of all kinases are structurally well conserved (Scapin, 2002). The structural similarities of the KD are obvious when comparing the crystal structures of the KD of TgCDPK1 with the Calcium/Calmodulin (CaM) dependent-kinase II from H. sapiens (HsCaMKII) (Fig. 2a) (Rellos et al. 2010). These two proteins, which share a sequence identity of approximately 42% over 264 residues of the KD, display the same canonical kinase fold and superimpose with an overall root mean square deviation (rmsd) of approximately 1.5 Å. Note that the loop over the adenosine triphosphate (ATP) binding site adopts a very different conformation presumably due to an induced fit of binding of two very different ligands. TgCDPK1 is bound to the ATP analogue ANP (Fig. 2a) while HsCaMKII is bound to a comparatively small inhibitor. More importantly there are significant differences in the ATP binding site, specifically a residue with no side chain (glycine) close to the adenine binding position. This residue, Gly128 is also termed the gatekeeper residue. Almost all mammalian kinases possess a large residue, a phenylalanine in HsCaMKII for example, in this position. Hence, CDPK1 feature an enlarged ATP binding site with a hydrophobic pocket that can be exploited for structure-based drug design. This key structural difference in the binding pocket is shown in the surface representation where the ATP-analogue is shown as stick representation (Fig. 2b). The additional space at the end of the pocket below the surface of the gatekeeper residue Gly 128 in magenta is clearly visible.

DEVELOPMENT OF SPECIFIC TgCDPK1 INHIBITORS

Soon after the structural differences between TgCDPK1 and the mammalian homologues were identified, two groups started to develop selective TgCDPK1 inhibitors (Ojo et al. 2010; Wernimont et al. 2010). Initial compounds were based on known inhibitors previously developed for yeast kinases featuring amino acids with small side chains at the gatekeeper position. Importantly, these known kinase inhibitors, termed bumped kinase inhibitors (BKI) have been shown to be inactive against mammalian kinases (Hanke et al. 1996). Generally, BKIs are based on the planar pyrazolo[3,4-d]pyrimidin-4-amine substituted with a bulky hydrophobic group on the C3 position (Bishop et al. 1998). The first example of a BKI with a sub-μmolar IC_{50} is 1-(1-methylethyl)-3-(naphthalen-1-ylmethyl)-1H-pyrazolo[3,4-d] pyrimidin-4-amine. The co-crystal structure of TgCDPK1 shows that the naphtalen-1-ylmethyl moiety fills the hydrophobic pocket.
created by the small gatekeeper residue Gly128 and lined by methionine and leucine residues, and one lysine residue (Fig. 3a and b). The chemically closely related 1-tert-butyl-3-naphthalen-2-yl-1H-pyrazolo[3,4-d]pyrimidin-4-amine (Fig. 3c and d) adopts a similar conformation with the bulky aromatic substituent at the C3 position occupying the space next to the gatekeeper residue. Critically for the subsequent drug development was the fact that these and related BKIs reduced T. gondii proliferation significantly (Ojo et al. 2010; Sugi et al. 2010). These results sparked extensive medicinal chemistry efforts where a large number of compounds based on the BKI scaffold (4-amino-1H-pyrazolo[3,4-d]pyrimidine) were synthesized and tested. A number of compounds exhibited sub- or low-nanomolar IC50 values and high activity in parasite growth models (EC50 in the low- and sub-molar range) while retaining specificity when compared with mammalian kinases (Lourido et al. 2013; Zhang et al. 2014; Moine et al. 2015). In addition to the pyrazolopyrimidine (PP) scaffolds, acylbenzimidazole and 5-amino-pyrazole-4-carboxamide-based compounds have been shown to have similar properties (Zhang et al. 2012; Zhang et al. 2014; Huang et al. 2015). While the initial BKIs showed excellent potency in vitro and in vivo they also exhibited significant hERG (human Ether-Related Gene) inhibition thus posing potential cardiotoxicity (Doggett et al. 2014). Further extensive medicinal chemistry efforts

Fig. 3. Close-up of BKIs bound to TgCDPK1 in the ATP binding site. The gatekeeper residue Gly128 is depicted in magenta, key hydrophobic residue of the binding site are labelled and shown in grey (a) 1-(1-methylethyl)-3-(naphthalen-1-ylmethyl)-1H-pyrazolo[3,4-d]pyrimidin-4-amine shown in ball-and-stick representation (PDB: 3i7b), (b) chemical structure of the ligand, (c) 1-tert-butyl-3-naphthalen-2-yl-1H-pyrazolo[3,4-d]pyrimidin-4-amine (PDB:3i7c) (Ojo et al. 2010) and (d) chemical structure of the ligand. ATP, adenosine triphosphate; CDPKs, calcium-dependent protein kinases.
finally led to the current lead TgCDPK1 inhibitor, (1-{4-amino-3-[2-(cyclopropyloxy)quinolin-6-yl]-1H-pyrazolo[3,4-d]pyrimidin-1-yl}-2-methylpropan-2-ol) that combines high activity and selectivity with favourable pharmacokinetic properties and low hERG activity (Vidadala et al. 2016). Note that the compound is bound to the protein via H-bonds of the pyrimidin ring to the main chain, while the hydrophobic cyclopropyloxy-quinoline moiety forms a large number of hydrophobic interactions (Fig. 5). Taken together, the structure-based approaches of drug development applied to TgCDPK1 has led to three different series of compounds with high inhibitory activity, good pharmacokinetic parameters and promising efficacy in murine models.

CDPK1 INHIBITORS FOR RELATED PARASITES

Based on the success of developing specific TgCDPK1 inhibitors, recent work has branched out towards related apicomplexan parasites. For example, Neospora caninum, a cyst-forming parasite closely related to T. gondii represents the leading cause of abortion in cattle. This parasite expresses a CDPK1 with 96% sequence identity to TgCDPK1 where all residues in the active side are conserved, bar one conservative variation from phenylalanine to tyrosine (Ojo et al. 2014). Consequently, the crystal structures of TgCDPK1 and NcCDPK1 show very similar overall structures (root mean square deviations (rmsd) on C-alpha atoms 0.5 Å) and the same binding mode of a BKI. Importantly, a number of BKIs display comparable in-vivo activities (Ojo et al. 2014; Winzer et al. 2015). Members of the Cryptosporidium genus are the causative agent of cryptosporidiosis in immune-compromised patients and malnourished children (Shoultz et al. 2016). CDPK1 from C. parvum Iowa II (CpCDPK1) shares a sequence identity of approximately 41% with TgCDPK1, however, the active site residues including the gatekeeper residue are highly conserved. The screening of BKI libraries resulted in highly active CpCDPK1 inhibitors based on the 5-aminopyrazole-4-carboxamide scaffold with clear potential for drug development (Castellanos-Gonzalez et al. 2016). High-throughput screening for Plasmodium falciparum CDPK1 (PfCDPK1) inhibitors resulted in five chemical series, including the PP scaffold (Fig. 4a). PfCDPK1 shares a sequence identity of approximately 47% with TgCDPK1 and the gatekeeper residue threonine harbours a slightly larger side chain compared to glycine. However, this side chain is still relatively small and mutational studies clearly indicated that these inhibitors bind at the same site (Ansell et al. 2014). More recent studies with PfCDPK1 inhibitors based on the chemically very similar Imidazopyridazine scaffold appear to show that these compounds also target cyclic GMP dependent kinases as well as Heat Shock Protein 90. These findings question the prospect of PfCDPK1 inhibitors for further drug development (Green et al. 2015). Taken together, these recent results show the potential of BKIs for future drug
development in Toxoplasma and related parasites but they also illustrate the limitations of transferring detailed structural data to more distantly related proteins.

**FUTURE CHALLENGES**

Over the last 5 years there has been significant progress in the development of selective inhibitors of one of the key CDPKs from *T. gondii* achieved by taking advantage of a series of high-resolution crystal structures. While most of the previous research has focused on *T. gondii*, further efforts are currently underway to investigate inhibitors of CDPK1 from Cryptosporidium and *Plasmodium* spp. (Gaji et al. 2014; Green et al. 2015; Crowther et al. 2016). However, further research is required to unravel the biological roles of *P*CDPKs and their potential as future drug targets (Kumar et al. 2017).

Although the most promising *Tg*CDPK1 inhibitors show high efficacy in murine models, future research is required to increase solubility and bioavailability in order to proceed to clinical trials. Furthermore, current lead compounds only target the ATP binding site of *Tg*CDPK1. However, allosteric kinase inhibitors and modulators have shown enormous potential to target specific kinases and could be further exploited (Fang et al. 2013). Additional binding sites in less conserved regions such as the carboxylic binding site recently discovered in *Tg*CDPK2 can serve as starting points for the development of new inhibitors (Ubaldi et al. 2015). Clearly, more works needs to be done to understand the role of the other members of the Apicomplexan CDPK family. In this regard, the recent development of CRISPR/Cas9 technology to modify the genes of members of the Apicomplexan family (Shen et al. 2014; Vinayak et al. 2015) will greatly facilitate the detailed analysis of the biological function of CDPK family members (Long et al. 2016; Wang et al. 2016).

**FINANCIAL SUPPORT**

This work was generously supported by the BBSRC grant BB/M024156/1. In addition, we would like to thank the Biophysical Sciences Institute (BSI) for seed corn funding and the Institute for Advanced Studies (IAS) for a Senior Research Fellowship at Durham University to CLMV. EMC is grateful for a studentship by the Newcastle-Liverpool-Durham BBSRC Doctoral Training Partnership.

**REFERENCES**


Biochemical and antiparasitic properties of inhibitors of the *Plasmodium falciparum* calcium-dependent protein kinase *P*CDPK1. Antimicrobial Agents and Chemotherapy 58, 6032–6043.


