Article

Sarcocystis neurona Protein Kinases: Computational Identification, Evolutionary Analysis and Putative Inhibitor Docking

Edwin K. Murungi [1] & Henry M. Kariithi [2]

[1] Department of Biochemistry and Molecular Biology, Egerton University, P.O. Box 536, 20115, Njoro, Kenya; eddkimm@gmail.com

[2] Biotechnology Research Institute, Kenya Agricultural and Livestock Research Organization, P.O Box 57811, 00200, Kaptagat Rd, Loresho, Nairobi, Kenya; henry.kariithi@kalro.org

* Correspondence: eddkimm@gmail.com; Tel: (+254)-789-716059

Abstract: The apicomplexan parasite, *Sarcocystis neurona* causes the degenerative neurological equine protozoal myeloencephalitis (EPM) disease of horses. Due to its host range expansion, *S. neurona* is an emerging threat that requires close monitoring. In apicomplexans, protein kinases (PKs) have been implicated in a myriad of critical functions such as host cell invasion, cell cycle progression and host immune responses evasion. Here, we used various bioinformatics methods to define the kinome of *S. neurona* and phylogenetic relatedness of its PKs to other apicomplexans. Further, three-dimensional (3D) homology models for selected *S. neurona* putative PKs were constructed and evaluated for inhibitor docking. We identified 92 putative PKs clustering within the AGC, CAMK, CK1, CMGC, STE, TKL, aPK and OPK groups. Although containing the universally conserved PKA (AGC group), *S. neurona* kinome was devoid of PKB and PKC, but contained the six apicomplexan conserved CDPKs (CAMK group). The OPK group was represented by ROPKs 19A, 27, 30, 33, 35 and 37, but was devoid of the virulence-associated ROPKs 5, 6, 18 and 38. Two out of the three *S. neurona* CK1 enzymes had high sequence similarities to *T. gondii* TgCK1-α and TgCK1-β and the *Plasmodium* PfCK1. Docking of four inhibitors onto homology models of putative ROP27 and PKA indicated that inhibition of *S. neurona* PKs is feasible, but needs to be experimentally tested. The essentiality of apicomplexan PKs makes the elucidation of *S. neurona* kinome a key milestone for development of novel therapeutics for EPM.

Keywords: *Sarcocystis neurona*; EPM; Docking; Apicomplexans; Phylogeny; Homology modeling.

1. Introduction

Equine protozoal myeloencephalitis (EPM) is an infectious, progressive, degenerative neurological disease of horses caused by the apicomplexan parasite, *Sarcocystis neurona* [1]. To complete its life cycle, this heteroxenous parasite requires a reservoir host (i.e. opossums; *Didelphis virginiana, D. albiventris*) and an aberrant (horses) or intermediate host (cats, skunks, raccoons and sea otters) [2]. Opossums become infected upon ingestion of sarcocysts containing hundreds of bradyzoites, which undergo gametogony that sporulate into mature oocysts, which are then shed in the faeces. Upon ingestion by the intermediate or aberrant hosts, the oocysts release the environmentally resistant sporozoites, which chronically parasitize the neural and inflammatory cells of the host's central nervous system (CNS). Clinical EPM symptoms depend on which part of the CNS is parasitized; parasitization generally results in abnormal gait, dysphagia and muscle atrophy in affected horses [3].

The intracellular nature of *S. neurona* and its ability to evade the host's immune surveillance [4] makes EPM treatment expensive, lengthy and challenging. Traditionally, clinical treatment of EPM

involved inhibitors of folate synthesis and metabolism (sulphonamides/pyrimethamine combination) over a prolonged period [5]. More recently, triazines derivatives (diclazuril, ponazuril) that target the parasite's apicoplast [6], nitazoxanide, a pyruvate:ferridoxin oxidoreductase analogue that inhibits the parasite's anaerobic metabolism [7], and anti-inflammatory agents and immune stimulants [8] have been used with variable success in eliminating clinical signs. Despite the availability of these drugs, EPM treatment is complicated by the emergence of drug-resistance (due to intermittent or periodic treatments), cost of therapies and drug toxicity (e.g. diarrhea and/or anemia) and infection relapses due to re-growth of residual parasites after the treatment regimes [2]. As such, discovery and development of novel therapeutics for EPM is imperative.

To successfully invade the host cells, apicomplexans utilize three specialized exocytic organelles (micronemes, rhoptries and dense-granules) [9]. The microneme is used for host cell recognition, binding, penetration and gliding along the cytoskeletal structures. Rhoptry proteins are discharged into the host cell during parasite internalization and are crucial in the formation of the parasitophorous vacuoles (PVs). Developing zoites contain non-pedunculated condensing vesicles that synthesize and package inactive rhoptry proteins, which are proteolytically activated when the rhoptry contents are condensed [10]. The PVs facilitate parasite development by allowing nutrient transport from the host cell and by blocking lysosomal fusion, which would otherwise kill the parasites [11]. Upon internalization, zoites use the dense-granules to remodel the PVs into functionally active organelles.

The virulence of the apicomplexans is highly dependent on protein kinases (PKs) that are involved in the invasion and modification of host cell structure and function. PKs are classified into various groups, comprising of many families and sub-families [12-14]. The major apicomplexans PK groups are the PKs A, G and C (AGCs), which are regulated by cyclic nucleotides and other second messengers; the calmodulin/calcium-dependent PKs (CAMKs); the casein kinase 1 (CK1) and close relatives, the PKs involved in cell-cycle control and signal transduction (containing cyclin-dependent, mitogen-activated and cell division cycle [CDC]-like PKs; or CMGCs); the receptor guanylate cyclase (RGC); the "sterile" serine/threonine kinases (STEs, including several kinases functioning in the MAPK cascades); the tyrosine kinases (TKs); and the tyrosine kinase-like kinases (TKLs) [15-17]. In addition, there is a sizeable proportion of other PKs that do not fit into any of the above-mentioned eight major PK groups (i.e. the "Other" kinases [OPKs]), and another smaller set of atypical PKs (aPKs) that have no clear sequence similarity with the conventional PKs [13, 18, 19].

Each of the PK families have vital roles in parasite's survival. For instance, PfPK-B (AGC family), PfTKL3 (TKL family), and four of the seven CDPKs (CAMK family) are required by *Plasmodium* parasites to complete their asexual cycle [20-22]. In a recent study, deletion of TgCK1α (CK1 family) resulted in defective replication of *Toxoplasma gondii* in vitro [23]. The *P. falciparum* CKL and SRPK1 (CMGC family) complement each other in the regulation of mRNA splicing (via phosphorylation of splicing factors [24]). Since apicomplexans lack typical MAPK cascades, the STEs are not well studied. However, in the parasites that do have the MAPK pathways, STEs are essential for parasite growth (e.g. in the human parasites, *Schistosoma mansoni* [25]). Further, for parasites without the conventional MAPK cascades, aPKs can activate the signaling pathways, for instance the *Plasmodium* Pfnek3 [26]. Finally, some of the notable OPKs include aurora kinases, rhoptry kinases (ROPKs) and parasite-specific eukaryotic initiation factor-2 (elF2) kinases (elF2K), which are important in parasite virulence and differentiation [27-29].

The differences between the parasite's and their host's PK homologs can be used to develop anti-parasitic strategies via specific inhibition of the parasite's PKs [30]. This is feasible due to the essential roles of PKs in the parasite's growth and development. Proof of principle of this approach has been demonstrated by the inhibition of human PKs using chemical ligands to treat cancers and other diseases [31, 32]. Recently, Ojo *et al.*, [33] provided evidence that PKs can be targeted for rationally designed drugs that can potently inhibit growth of *S. neurona*. The technology is available and approved for therapeutic intervention, thus offering a unique prospect of repurposing chemical ligands to manage *S. neurona* infections [34].

Here, we used a genome-wide approach to define the kinome of *S. neurona* and determined the relatedness of the putative PKs to those reported in other apicomplexans. We also tested the docking of some specific inhibitors to a selection of some of the putative PKs. Defining the *S. neurona* kinome is not only important in providing insights into the parasite's biology, but also identification of potential novel drug targets that can be used to clear chronic *S. neurona* infections, reduce parasite survival, augment the host's anti-parasitic immune responses, or as candidates for development of EPM vaccines.

2. Results

2.1 *Sarcocystis neurona* encodes 92 putative kinases

To date, at least 15 apicomplexan genomes (coccidians, gregarines, hemosporidians and piroplasmids) have either been fully sequenced or partially annotated [35]. In the current study, we conducted an exhaustive genome-wide search of the newly sequenced *S. neurona* genome [36], and identified 92 putative PKs (**Table 1**). The identified PKs contained the characteristic PK (IPR000719) or PK-like (IPR011009) domains and three conserved amino acids constituting the catalytic triad (Lys30, Asp125, Asp143). The PKs had sizes ranging between 152 and 6544 amino acids and relative molecular weights of between 15.94 to 671.51 kDa. Majority of the PKs had isoelectric point (pI) greater than 7.0, implying that the PKs have low turnover rates since in general, acidic proteins are thought to be degraded more rapidly than neutral or basic proteins [37].

Assignment of *S. neurona* PK groups was accomplished by sequence clustering using Blast2GO [62], and by BLASTp searches in the Kinase.com database [14]. Out of the eleven known PK groups [12-14], *S. neurona* PKs segregated into the AGC (*n*=9), CAMK (*n*=20), CK1 (*n*=3), CMGC (*n*=19), STE (*n*=2), TKL (*n*=6), aPK (*n*=2) and OPK (*n*=31) (**Table 1**). Apart from the 31 OPKs that do not fit to the major kinase groups, the CAMK and CMGC groups, whose members are essential for the parasite's host cell invasion [38] and differentiation (via cell-cycle regulation) [39], respectively, had the highest number of PKs underlying the importance of these processes in the parasite. Unlike in some parasites (e.g. *P. falciparum*) that lack STEs and TKLs [40], *S. neurona* contains PKs in the two classes.

Table 1. Description of the 92 putative PKs identified in the kinome of *S. neurona*. The putative PKs could be classified into eight groups. The amino acid coordinates of the conserved PK domains in the protein sequences, and the PK homologies to other apicomplexan PKs are shown in columns 7 to 12.

Description of the putative protein kinases (PKs) in the genome of *S. neurona*							Description of protein kinase (PK) homologies (BLASTp)				
Protein ID [a]	**Sequence annotations; description [b]**	**Family; [subfamily] c**	**Length [aa]**	**pI**	**MW [kDa]**	**PK domain coordinates**	**Sequence name; [apicomplexan]**	**Bit Score**	**E-value**	**Identity [%]**	**Accession number**
	1. Kinase Group AGC (Protein kinases A [PKA], G [PKG] and C [PKC] families)										
SRCN_1312	AGC kinase	3-phosphoinositide dependent PK-1 (PDK1)	903	5.52	101.13	137–481	PDPK; [*T. gondii* RUB]	417	6.00E-135	59	KFG61374.1
SRCN_3339	AGC kinase	PKA	1428	8.85	147.52	1102–1417	Putative AGC kinase; [*N. caninum* L]	611	0.0	81	CEL65574.1
SRCN_4249	AGC kinase	Nuclear dbf2-related (NDR)	152	8.91	17.47	6–141	Putative AGC kinase; [*N. caninum* L]	219	2.00E-68	76	XP_003883757.1
SRCN_4518	PK G AGC kinase family member PKG	Ciliate-E2	425	5.83	48.78	97–399	AGC kinase TgPKG1; [*T. gondii* ME49]	830	0.0	92	EPR61116.1
SRCN_3990	cAMP-dependent kinase	CAMKL; [MELK]	1907	9.55	217.55	782–1634	cAMP-dependent protein kinase [*T. gondii* VEG]	75.9	1e-12	30	ESS31194.1
SRCN_5165	cAMP-dependent PK, catalytic chain	PKA	343	8.99	39.36	20–338	AGC kinase; [*T. gondii* ARI]	425	3.00E-150	90	KYF43224.1
SRCN_4913	Putative PK	PKD	2330	6.42	244.64	1093–1737	Putative PK; [*E. tenella*]	107	9.00E-22	60	XP_013228294.1
SRCN_5430	AGC kinase	Ribosomal protein S6 Kinases (RSK; [p70])	1378	5.54	139.28	824–1344	AGC kinase; [*T. gondii* MAS]	223	3.00E-60	59	KFH07588.1
SRCN_5610	cAMP-dependent PK, catalytic chain	PKA	333	9.00	37.96	12–318	cAMP-dependent PK, catalytic subunit; [*T. gondii* ME49]	641	0.0	92	XP_002366464.1
	2. Kinase Group calcium (Ca2+)-/calmodulin-regulated kinases (CAMK)										
SRCN_1071	Ca^{2+}-dependent kinase	CAMK1	1495	9.47	152.66	1085–1401	Putative PK; [*T. gondii* VEG]	188	1.00E-47	67	ESS31884.1
SRCN_2032	Putative PK	Ciliate-C1	297	6.22	33.26	15–297	PK; [*H. hammondi*]	424	1.00E-146	68	XP_008882026.1
SRCN_2165	Ca^{2+}-dependent kinase CDPK2B	CDPK	692	7.33	75.65	101–401	Ca2+-dependent PK CDPK2A; [*T. gondii*	674	0.0	90	KYF44522.1

Description of the putative protein kinases (PKs) in the genome of *S. neurona*							Description of protein kinase (PK) homologies (BLASTp)				
Protein ID [a]	**Sequence annotations; description** [b]	**Family; [subfamily] c**	**Length [aa]**	**pI**	**MW [kDa]**	**PK domain coordinates**	**Sequence name; [apicomplexan]**	**Bit Score**	**E-value**	**Identity [%]**	**Accession number**
							ARI]				
SRCN_2257	Histone kinase	CAMKL; [AMP-activated protein kinase (AMPK)]	1800	8.98	187.71	1159–1448	Putative CAM kinase, SNF1 family; [*E. acervulina*]	377	7.00E-105	64	XP_013252246.1
SRCN_2544	CAM SNF1 AMK1 family	CAMKL;[AMPK-regulated kinase novel kinase (NUAK)]	333	5.71	37.85	62–333	CAM kinase, SNF1/AMK1 family ToxPK1; [*N. caninum* L]	480	9.00E-167	79	XP_003882065.1
SRCN_2937	Ca^{2+}-signalling kinase MARK	CAMKL; [microtubule affinity regulating kinase (MARK)]	278	9.62	31.29	1–250	Putative Ca^{2+} signalling PK MARK; [*T. gondii* GT1]	356	5.00E-124	72	EPR59053.1
SRCN_3011	Calmodulin-dependent PK (CAM) CDPK6	CDPK	1435	9.25	154.44	1238–1435	Cdpk kinase domain; [*T. gondii*]	179	6.00E-49	75	3IS5_A
SRCN_3314	A Chain crystal Structure of TgCDPK1 with inhibitor bound	CDPK1	519	5.99	58.89	37–335	Calmodulin-domain PK 1; [*T. gondii*]	536	0.0	97	3MA6_A
SRCN_3583	Ca^{2+}-dependent kinase CDPK5	CDPK	454	6.09	50.38	35–308	Ca^{2+}-dependent PK CDPK5; [*T. gondii* ARI]	776	0.0	89	KYF43137.1
SRCN_3701	Ca^{2+}-dependent kinase CDPK3	CDPK	560	5.91	62.07	77–362	Ca^{2+}-dependent Kinase; [*T. gondii*]	493	1.00E-172	87	3DXN_A
SRCN_4076	CAM CDPK family	CDPK	1701	5.93	181.42	1079–1663	CAM kinase, CDPK family; [*H. hammondi*]	239	2.00E-62	69	XP_008884897.1
SRCN_4093	PK	CAMKL; [AMP-activated protein kinase (AMPK)]	1155	9.02	118.33	1–261	Putative atypical MEK-related kinase; [*N. caninum* L]	222	1.00E-59	70	XP_003880869.1
SRCN_4390	Ca^{2+}-dependent kinase CDPK2	CDPK	790	6.24	85.47	280–556	Ca^{2+}-dependent PK, related; [*N. caninum* L]	1130	0.0	74	XP_003884321.1
SRCN_4815	Histone kinase (partial)	CAMKL; [AMP-activated protein kinase (AMPK)]	711	6.53	75.98	1–314	SNF1-related PK catalytic-α KIN10, 5 AMP-activated PK; [*N. caninum* L]	568	0.0	48	CEL67550.1
SRCN	CAM CDPK	CDPK	2748	8.98	282.1	796–885	Putative CAM kinase,	117	3.00E-24	59	XP_003881901.

Description of the putative protein kinases (PKs) in the genome of *S. neurona*							Description of protein kinase (PK) homologies (BLASTp)				
Protein ID [a]	**Sequence annotations; description [b]**	**Family; [subfamily] c**	**Length [aa]**	**pI**	**MW [kDa]**	**PK domain coordinates**	**Sequence name; [apicomplexan]**	**Bit Score**	**E-value**	**Identity [%]**	**Accession number**
_5227	CDPK8-like				2		CDPK family; [*N. caninum* L]				1
SRCN_5410	Calmodulin-dependent PK (CAM-SNF1 family)	CAMK1	467	8.94	52.04	168–446	CAM kinase, SNF1 family; [*H. hammondi*]	431	5.00E-134	78	XP_008883430.1
SRCN_5812	Ca^{2+}-dependent kinase CDPK9	CDPK	760	8.37	84.23	254–573	Ca^{2+}-dependent PK CDPK9; [*H. hammondi*]	1139	0.0	81	XP_008889286.1
SRCN_5948	Ca^{2+}-dependent kinase CDPK8	CDPK	3298	7.11	345.85	208–860	EF-hand domain-containing protein; [*T. gondii* ME49]	114	1.00E-23	58	XP_002368547.1
SRCN_6597	Ca^{2+} dependent kinase CDPK7	CAMK1	1374	9.09	138.28	365–623	PK-PH domain-containing protein; [*T. gondii* ME49]	813	0.0	80	XP_002366487.1
SRCN_6606	Ca^{2+}-dependent kinase CDPK4	CDPK	1632	9.42	170.93	813–1236	Ca2+-dependent PK; [*T. gondii*]	731	0.0	58	CAD32376.2
	3. Kinase Group casein kinase 1 (cell kinase 1)										
SRCN_3445	Casein kinase I	CK1-D	323	9.34	37.65	6–290	Casein kinase 1; [*T. gondii* ME49]	603	0.0	94	XP_002366683.1
SRCN_4587	Casein kinase I	CK1-D	137	7.78	15.94	1–137	Casein kinase I; [*H. hammondi*]	226	3.00E-71	81	XP_008883809.1
SRCN_4645	Casein kinase I	CK1-D	229	9.51	25.66	38–229	Casein kinase I; [*T. gondii* GAB2-2007-GAL-DOM2]	259	1.00E-86	79	KFG42638.1
	4. Kinase Group CMGC (including cyclin-dependent kinases, mitogen-activated PKs, glycogen synthase kinases and CDK-like kinases)										
SRCN_1104	Cyclin-dependent kinase family 5	Ca^{2+}-dependent PK-L (CDKL)	372	9.23	42.71	1–318	Cyclin-dependent kinase family 5 protein; [*H. hammondi*]	490	2.00E-173	76	XP_008884207.1
SRCN	Cell-cycle-associated	Serine-arginine rich PK	2911	5.37	302.9	713–1837	PK; [*T. gondii* ME49]	476	9.00E-14	76	XP_002369401.

Description of the putative protein kinases (PKs) in the genome of *S. neurona*							Description of protein kinase (PK) homologies (BLASTp)				
Protein ID [a]	**Sequence annotations; description [b]**	**Family; [subfamily] c**	**Length [aa]**	**pI**	**MW [kDa]**	**PK domain coordinates**	**Sequence name; [apicomplexan]**	**Bit Score**	**E-value**	**Identity [%]**	**Accession number**
_1236	kinase (SRPK)	(SRPK)			0				1		1
SRCN_1479	CMGC Lammer	CLK	748	10.23	79.09	485–748	Cell-cycle-associated PK CLK; [*T. gondii* FOU]	288	2.00E-81	74	KFG33061.1
SRCN_1611	CMGC Dual-specificity tyrosine-regulated kinase (Dyrk)	DYRK; [DyrkP]	1504	5.86	160.34	551–1498	Cell-cycle-associated PK DYRK; [*T. gondii* VEG]	223	2.00E-57	63	ESS33160.1
SRCN_1731	Cell-cycle-associated kinase GSK	Glycogen synthase kinase (GSK)	219	6.59	24.39	1–175	Cell-cycle-associated PK GSK; [*H. hammondi*]	330	2.00E-112	82	XP_008887193.1
SRCN_1732	Cell-cycle-associated kinase GSK	Glycogen synthase kinase (GSK)	203	10.78	20.72	86–203	CMGC kinase, GSK family TgPK3; [*E. brunetti*]	114	1.00E-28	91	CDJ46527.1
SRCN_2759	Cell-cycle-associated kinase partial	Ca^{2+}-dependent PK (CDK); [CRK7]	1122	6.14	118.67	541–1122	Cell-cycle-associated PK CDK; [*T. gondii* VAND]	118	1.00E-25	79	KFH12036.1
SRCN_2845	CMGC DYRK PRP4 kinase	DYRK; [PRP4]	1665	9.76	177.08	1267–1596	Putative PK (CLK3); [*P. malariae*]	330	1.00E-102	69	SBS85334.1
SRCN_3891	CMGC kinase	DYRK; [DYRK2]	674	8.85	73.85	399–674	Putative CMGC kinase; [*T. gondii* ME49]	80.1	6.00E-14	67	EPT25192.1
SRCN_4209	CMGC MAPK family (ERK) MAPK-1	Mitogen-activated PK (MAPK); [ERK])	2361	6.73	247.71	94–754	CMGC, MAPK/ (ERK) TgMAPK-1; [*E. brunetti*]	137	1.00E-30	74	CDJ49492.1
SRCN_4674	Cyclin-dependent kinase	Ca^{2+}-dependent PK (CDK); [CDK7]	138	7.80	15.50	1–138	Cyclin-dependent kinase; [*T. gondii* GT1]	108	7.00E-27	58	EPR60430.1
SRCN_4801	Cell-cycle-associated kinase	Ca^{2+}-dependent PK (CDK); [CDK5]	300	6.08	34.33	1–289	CMGC kinase, CDK family TgPK2; [*N. caninum* L]	576	0.0	91	XP_003885801.1
SRCN_5365	Cell-cycle-associated kinase MAPK	Mitogen-activated PK (MAPK; [ERK])	417	6.77	48.32	7–363	Cell-cycle-associated PK MAPK; [*H. hammondi*]	823	0.0	93	XP_008886907.1
SRCN	Cell-cycle-associated	Ca^{2+}-dependent PK	690	9.55	80.90	208–603	Putative	390	2.00E-12	87	KYF45878.1

Description of the putative protein kinases (PKs) in the genome of *S. neurona*							Description of protein kinase (PK) homologies (BLASTp)				
Protein ID [a]	**Sequence annotations; description [b]**	**Family; [subfamily] c**	**Length [aa]**	**pI**	**MW [kDa]**	**PK domain coordinates**	**Sequence name; [apicomplexan]**	**Bit Score**	**E-value**	**Identity [%]**	**Accession number**
_6346	kinase CDK	(CDK); [CDK5]					cell-cycle-associated PK CDK; [*T. gondii* ARI]		2		
SRCN_6427	CMGC CK2 kinase	Cell Kinase 2 (CK2)	1395	10.29	144.86	885–1356	CMGC kinase, CK2 family; [*T. gondii* MAS]	241	6.00E-73	98	KFH07655.1
SRCN_6472	Cell-cycle-associated kinase ERK7	Mitogen-activated PK (MAPK; [ERK])	983	9.28	104.95	7–317	Cell-cycle-associated PK ERK7; [*T. gondii* ARI]	647	0.0	81	KYF46268.1
SRCN_761	Cell-cycle-associated kinase	Ca^{2+}-dependent PK (CDK); [CDK7]	577	9.34	58.39	144–490	Cell-cycle-associated PK; [*H. hammondi*]	283	3.00E-88	68	XP_008882409.1
SRCN_895	Cell-cycle-associated kinase	Ca^{2+}-dependent PK (CDK); [CDK10]	340	8.93	38.57	1–307	Cell-cycle-associated PK; [*T. gondii* ARI]	234	6.00E-75	76	KYF44017.1
SRCN_977	Cell-cycle-associated kinase CDK	Ca^{2+}-dependent PK (CDK); [PITSLRE/CDK11]	1502	7.38	156.35	1114–1429	Cell-cycle-associated PK CDK; [*T. gondii* p89]	454	3.00E-135	92	KFG28420.1
	5. Kinase Group 'Other' (OPK; i.e. kinases with conventional PK [ePK] domains that do not fit into any of the other major groups of kinases)										
SRCN_108	Unc-51-like autophagy activating kinase 1 (ULK1)	ULK	343	7.13	38.90	1-223	ULK kinase; [*T. gondii* VAND]	376	2.00E-130	75	KFH07419.1
SRCN_1606	eIF2 kinase IF2K-C	PEK; [general control nonderepressible 2 (GCN2)]	4034	8.98	406.57	1235–2178	eIF2 kinase IF2K-C; [*T. gondii* VAND]	259	4.00E-67	35	KFH07289.1
SRCN_2076	Rhoptry kinase family ROP30	Conserved hypothetical protein	1276	9.18	134.73	812–1260	ROP30 [*T. gondii* VEG]	230	3.00E-63	53	CEL76436.1
SRCN_2123	Rhoptry kinase family ROP35	PLK; [PLK-Unclassified]	291	9.30	33.50	53–265	ROP35; [*T. gondii* RUB]	207	2.00E-61	43	KFG59037.1
SRCN_3216	Rhoptry kinase family ROP32	CAMK-Unique	523	7.08	57.00	214–520	Putative PK; [*T. gondii* VAND]	167	1.00E-42	30	KFH00232.1
SRCN_2183	Rhoptry kinase family ROP35	Aurora-like	226	6.36	25.84	1–212	ROP35; [*T. gondii* VEG]	198	8.00E-59	48	ESS33297.1
SRCN_2271	Putative PK (incomplete catalytic	NimA (Never in mitosis gene A)-related Kinase	1463	9.02	157.18	437–1145	Putative PK; [*N. caninum* L]	327	5.00E-90	68	XP_003881849.1

Description of the putative protein kinases (PKs) in the genome of *S. neurona*							Description of protein kinase (PK) homologies (BLASTp)				
Protein ID [a]	**Sequence annotations; description** [b]	**Family; [subfamily] c**	**Length [aa]**	**pI**	**MW [kDa]**	**PK domain coordinates**	**Sequence name; [apicomplexan]**	**Bit Score**	**E-value**	**Identity [%]**	**Accession number**
	triad)	(NEK)									
SRCN_2403	Aurora kinase (incomplete catalytic triad)	PLK; [SAK/Plk4]	778	9.79	79.92	492–778	Putative Aurora kinase; [*N. caninum* L]	127	3.00E-28	44	XP_003880644.1
SRCN_2630	NimA related kinase (NEK) family protein	NEK	351	8.70	38.38	1–336	NEK kinase; [*T. gondii* ME49]	242	8.00E-75	52	XP_018638598.1
SRCN_286	Wee kinase	Inhibitory regulator of the *RAS*-cAMP (IRA1) kinase suppressor (IKS)	1019	6.20	106.67	598–959	Wee kinase; [*H. hammondi*]	445	5.00E-141	58	XP_008882669.1
SRCN_3075	Tyrosine kinase-like (TKL) protein	Numb-associated kinase (NAK)	1571	8.41	164.18	16–500	TKL; [*T. gondii* TgCatPRC2]	138	1.00E-32	73	KYK64203.1
SRCN_3142	PIK3R4 kinase-related	Aurora	997	8.72	106.54	548–899	Putative PIK3R4 kinase-related protein; [*N. caninum* L]	449	2.00E-137	60	XP_003885774.1
SRCN_3151	NimA related kinase (NEK) family protein	NEK	3186	7.96	318.69	352–656	NEK kinase; [*T. gondii* VEG]	468	7.00E-131	73	CEL78174.1
SRCN_3247	Rhoptry kinase family ROP27	Ciliate-D	345	8.94	38.81	23–325	ROP27; [*T. gondii* p89]	163	4.00E-43	31	KFG37427.1
SRCN_3417	Aurora kinase	Aurora	438	7.65	48.49	14–289	Aurora kinase; [*T. gondii* TgCatPRC2]	490	3.00E-155	76	KYK63669.1
SRCN_3444	Unc-51-like Autophagy activating kinase 1 (ULK1)	ULK	406	6.52	44.69	12–406	ULK kinase; [*T. gondii* RUB]	232	3.00E-71	61	KFG59767.1
SRCN_3669	CMGC kinase	ULK	1803	8.41	189.65	736–1200	Putative CMGC kinase; [*N. caninum* L]	624	0.0	62	CEL65030.1
SRCN_4410	Rhoptry kinase family ROP35	PKA-like	204	9.44	23.50	1–166	ROP35; [*H. hammondi*]	107	1.00E-25	39	XP_008885989.1
SRCN_4503	eIF2 kinase IF2K-B	PEK; [general control nonderepressible 2 (GCN2)]	158	5.76	17.59	1–158	eIF2 kinase IF2K-B [*T. gondii* TgCatPRC2]	149	5.00E-41	74	KYK69938.1
SRCN	NimA related kinase	NEK	187	8.20	21.23	1–186	NEK kinase;	177	4.00E-54	64	XP_008885186.

Peer-reviewed version available at *Pathogens* **2017**, *6*, 12; doi:10.3390/pathogens6010012

Description of the putative protein kinases (PKs) in the genome of *S. neurona*							Description of protein kinase (PK) homologies (BLASTp)				
Protein ID [a]	**Sequence annotations; description [b]**	**Family; [subfamily] c**	**Length [aa]**	**pI**	**MW [kDa]**	**PK domain coordinates**	**Sequence name; [apicomplexan]**	**Bit Score**	**E-value**	**Identity [%]**	**Accession number**
_4528	(NEK) family protein						[*H. hammondi*]				1
SRCN_2404	Aurora kinase (incomplete catalytic triad)	Serum and glucocorticoid induced Kinase (SGK)	295	8.81	31.42	1–249	Putative Aurora kinase; [*N. caninum* L]	126	7.00E-31	43	CEL65223.1
SRCN_5653	PEK kinase	Aurora	626	8.27	60.75	513–626	PEK kinase [*T. gondii* TgCatPRC2]	251	1.00E-76	60	KYK62422.1
SRCN_5943	NIMA-related kinase NIMA1	NEK	2842	9.04	295.44	73–383	NIMA-related PK NIMA1; [*T. gondii* MAS]	486	2.00E-140	67	KFH05809.1
SRCN_6157	Unc-51-like autophagy activating kinase 1 (ULK1)	ULK	2420	9.38	250.58	1380–1672	ULK kinase [*T. gondii* ME49]	99.8	3e-21	38	XP_018635814.1
SRCN_6184	Myosin-light-chain kinase	Ciliate-E2-Unclassified	478	5.42	53.65	177–474	ROP19A [*T. gondii ME49*]	127	3.00E-30	27	XP_018637476.1
SRCN_6572	Tyrosine kinase-like (TKL)	ULK	622	6.13	68.80	1–345	TKL; [*T. gondii* VAND]	181	1.00E-46	74	KFH00338.1
SRCN_6812	PK	ULK	199	6.74	22.72	1–183	PK; [*H. hammondi*]	172	3.00E-48	53	XP_008887491.1
SRCN_7083	Rhoptry kinase family ROP35	PKA-like	262	9.62	30.26	1–242	ROP35; [*H. hammondi*]	127	6.00E-32	39	XP_008885989.1
SRCN_4310	Rhoptry kinase family ROP33	Kinase Homologous to SPS1/STE20 (KHS)	1591	9.85	169.63	1265–1578	ROP33; [*H. hammondi*]	306	2.00E-87	39	XP_008887632.1
SRCN_7082	Rhoptry kinase family ROP33	Kinase Homologous to SPS1/STE20 (KHS)	403	9.59	45.92	77–390	ROP33 [*T. gondii* p89]	277	3.00E-89	40	KFG45248.1
SRCN_7084	Rhoptry kinase family ROP37	Ribosomal protein S6 Kinases (RSK; [RSK])	339	5.41	38.09	19–334	ROP37; [*N. caninum* L]	144	1.00E-36	36	CEL64242.1
6.	**Kinase Group "Sterile" serine/threonine kinase, or sterile-phenotype kinases (STE)**										
SRCN_1328	Serine threonine kinase	Conserved hypothetical protein	1461	9.29	158.74	559–671	Hypothetical protein, conserved; [*E. maxima*]	88.6	5.00E-16	68	XP_013335801.1
SRCN_5172	"Sterile" serine/threonine	Mammalian Sterile 20-like (MST)]	6552	6.14	671.51	3410–4122	STE kinase; [*T. gondii* TgCatPRC2]	412	1.00E-114	54	KYK71951.1

Description of the putative protein kinases (PKs) in the genome of *S. neurona*							Description of protein kinase (PK) homologies (BLASTp)				
Protein ID [a]	**Sequence annotations; description** [b]	**Family; [subfamily] c**	**Length [aa]**	**pI**	**MW [kDa]**	**PK domain coordinates**	**Sequence name; [apicomplexan]**	**Bit Score**	**E-value**	**Identity [%]**	**Accession number**
	kinase (STE)										
	7. Kinase Group Tyrosine Kinase-Like (TKL)										
SRCN_1435	Tyrosine kinase-like (TKL)	Mixed lineage kinase (MLK); [Leucine Zipper-bearing Kinase (LZK)]	3064	8.05	306.74	2540–3060	Tyrosine kinase-like (TKL) protein; [*N. caninum* L]	278	3.00E-73	72	CEL64955.1
SRCN_1571	Tyrosine kinase-like (TKL)	Microtubule-associated S/T kinase (MAST)	550	9.76	59.91	135–501	Conserved hypothetical protein; [*E. praecox*]	76.3	6.00E-13	53	CDI87140.1
SRCN_3466	Tyrosine kinase-like (TKL)	TKL-Unique	3002	9.87	320.97	2342–2997	Tyrosine kinase-like (TKL) protein; [*H. hammondi*]	202	3.00E-50	65	XP_008887506.1
SRCN_3928	Tyrosine kinase-like (TKL)	LISK - LIMK (LIM kinase) and TESK (Testicular protein Kinase); [DD1]	5842	8.78	608.80	3639–4268	Tyrosine kinase-like (TKL) protein; [*T. gondii* TgCatPRC2]	216	2.00E-61	79	KYK63216.1
SRCN_4277	Kinase domain-containing protein	TKL-ciliate1	2256	8.43	240.25	1570–2256	Tyrosine kinase-like (TKL) protein; [*N. caninum* L]	204	8.00E-51	61	CEL67693.1
SRCN_811	Tyrosine kinase-like (TKL)	TKL-Unique	1099	9.18	119.26	814–1083	Putative tyrosine kinase-like (TKL) protein; [*E. acervulina*]	403	6.00E-125	59	XP_013252162.1
	8. Kinase Group Atypical (aPKs)										
SRCN_3601	Atypical MEK-related kinase	Muscle-associated kinase TRIO	950	7.20	103.14	381–850	Atypical MEK-related kinase; [*T. gondii* GT1]	171	1.00E-42	32	EPR62774.1
SRCN_5962	Atypical MEK-related kinase	Rho-associated protein kinase (ROCK)-like	805	5.01	87.71	525–805	Atypical MEK-related kinase; [*H. hammondi*]	127	4.00E-29	56	XP_008884362.1

[a] The protein sequences and their corresponding identified were obtained from Toxoplasma Genomics Resource database (Release 28; version May 2016) [41]; [b] The descriptions of the protein sequence is based on BLASTp annotations using Blast2GO (see text for details); [c] The kinase classification is based on BLASTp on the kinase database.

Peer-reviewed version available at Pathogens 2017, 6, 12; doi:10.3390/pathogens6010012

2.1.1 The AGC group

The numbers of apicomplexan AGCs range from four (in *B. bovis*) to 15 (in *T. gondii*) [15]. Based on our Blast2GO annotations and BLASTp homology searches against the kinome database, five out of the nine *S. neurona* AGCs (SRCN_3339, SRCN_3990, SRCN_5165, SRCN_5610 and SRCN_1312) were homologs to the universally conserved PKAs that are found in *N. caninum* and *T. gondii* (see **Table 1**). The PKAs are essential for completion of schizogony (asexual reproduction) in *Plasmodium* parasites [42]. Further, *S. neurona* contains a putative PKG (SRCN_4518) which shows high homology (92%) to the *T. gondii* TgPKG1 (**Table 1**); PKGs are in essential apicomplexans [43].

2.1.2 The CAMK group

CAMKs form the second-largest apicomplexan PKs (after OPKs). Apicomplexan kinomes constitute varying numbers of CAMKs, which range from seven (in *B. bovis*) to 29 (in *T. gondii*) [15]. The most important CAMK family is the CDPK, which appeared to constitute almost 50% of *S. neurona* putative CAMKs (See **Table 1**). In terms of homologies, the *S. neurona* kinome contained orthologs to the *T. gondii* CDPK1 (SCRN_3314), CDPK2B (SCRN_2165), CDPK3 (SCRN_3701), CDPK4 (SCRN_6606), CDPK5 (SCRN_3583), CDPK6 (SCRN_3011), CDPK7 (SCRN_6597) and CDPK8 (SCRN_5948). Other CDPK orthologs were to the *N. caninum* CDPK2 (SCRN_4390) and *H. hammondi* CDPK9 (SCRN_5812) (**Table 1**). Inhibition of TgCDPK1 has been shown to disrupt the motility, host cell invasion and egress of *T. gondii* [44]. Owing to the absence of mammalian CDPK homologs, the identification of a relatively large number of CDPK homologs in *S. neurona* could be utilized in the rational design of anti-parasitic therapeutics.

2.1.3 The CK1 group

It is notable that *S. neurona* putatively encodes for three CK1 enzymes. Apart from *T. gondii* and some alveolates (e.g. *Cryptosporidium hominis* and *C. parvum*, important causative agents of diarrhea in children), which have three and two CK1 enzymes, respectively, most apicomplexans possess a single CK1 enzyme [15]. Two of the three *S. neurona* putative CK1 (SRCN_3445 and SRCN_4645) showed high sequence similarity (>90%) to the *T. gondii* TME49_040640 (TgCK1-α) and TGME49_089320 (TgCK1-β), respectively (**Table 1**). Inhibition of CK1 showed potential for anti-parasitic interventions in *T. gondii* [45]. CK1 is critical for the asexual proliferation of the *Plasmodium* parasites and is expressed in all the life-cycle stages of the parasite [46]. Three putative *S. neurona* CK1 had significant sequence similarity to the *P. falciparum* PfCK1, i.e. 74% (SRCN_3445), 65% (SRCN_4587) and 56% (SRCN_4645) (data not shown).

2.1.4 The CMGC group

The CMGC is the largest PK group in apicomplexans; CMGC numbers range from 15 in *B. bovis* to 23 in *P. vivax* [15], which is within the range we identified in the *S. neurona* kinome in our study (i.e. 19 CMGCs; See **Table 1**). Notable of these were the two GSK homologs (SRCN_1731 and SRCN_1732). This finding is similar to what has been observed in *Plasmodium* parasites in which two GSK-3 enzymes have been reported, both of which are essential for the parasite [47]. Homology searches showed considerable sequence similarity (51% and 41% for SRCN_1731 and SRCN_1732, respectively) to the PfGSK-3 enzymes (data not shown). Notably, eight of the 19 CMGCs in *S. neurona* were CDKs, including CDK7 (SRCN_4674, SRCN_2759 and SRCN_761), CDK10 (SRCN_895) and CDK11 (SRCN_977). Available data show that CDKs are essential in *P. falciparum* [35]. We also identified two putative MAPK homologs (SRCN_4209 and SRCN_5365), and ERK7 (SRCN_6472) (see **Table 1**), a result which is comparable to the two MAPKs in the kinome of *P. falciparum* [15].

2.1.5 The OPK group

The apicomplexan-specific OPKs are a tight cluster of PKs without clear relation to any of the other major PK groups. Notable of these are ROPKs which have high sequence divergence and have been thought to be largely restricted to *T. gondii* [48], which has a total of 34 members spread in over 40 distinct sub-families [27]. Although their diversification in apicomplexans is poorly understood, some ROPKs are key virulence factors in *T. gondii* [27]. At least nine putative ROPKs could be identified in *S. neurona*, including ROPK19A (SRCN_6184), ROP27 (SRCN_3247), ROP30 (SRCN_2076), ROP33 (SRCN_7082 and SRCN_7086), ROP35 (SRCN_2183, SRCN_2123, SRCN_7083 and SRCN_4410) and ROP37 (SRCN_7084), implying that the ROPKs are not restricted to *T. gondii*. Although largely presumed to be inactive, ROPKs are implicated in the regulation of the host transcription [48], and their presence in *S. neurona* may support the hypothesis that the ROPKs have a unique activation mechanisms in their regulatory functions that facilitate apicomplexan pathogenesis [35, 49]. Other notable OPKs included two parasite-specific eukaryotic initiation factor-2 (eIF2) kinases (eIF2K-C [SRCN_1606] and eIF2K-B [SRCN_4503]), four NEKs (SRCN_4528, SRCN_2630, SRCN_286 and SRCN_3151) and four ULKs (SRCN_3444, SRCN_3669, SRCN_6812 and SRCN_6157) (**Table 1**). The eIF2Ks are conserved in apicomplexans, and are important for the induction of parasite differentiation into the bradyzoites cysts, which are clinically important [29].

2.1.6 The STE group

The STEs are poorly represented in apicomplexans, and although most apicomplexans have one or two STE genes per genome, some parasites such as *C. parvum* are reported to harbor up to six STEs [15, 17]. Our results suggest that *S. neurona* has at least one putative STE (**Table 1**). STEs are thought to function in MAPK pathway cascades despite the fact that this pathway is absent in apicomplexans. The small repertoire of apicomplexan STEs is in contrast to that reported in other parasites such as trypanosomatids, in which these enzymes regulate the length of the flagella [50].

2.1.7 The TKL group

Apicomplexans harbor a maximum of seven TKL-coding genes, which makes it notable that we identified six putative TKLs in *S. neurona* (**Table 1**). In *Plasmodium*, six TKLs are conserved, and are essential in the asexual proliferation of the parasites, hence the TKLs are thought to be ideal drug targets [22]. Two of the six *S. neurona* putative TKLs had considerable sequence similarities to the *Plasmodium* TKLs, including SRCN_3466 (36% similar to *P. malariae* TKL1) and SCRN_1435 (49% similar to *P. ovale* TKL3) (data not shown).

2.1.8 The aPK group

The aPKs share sequence similarity to the OPKs. Although apicomplexan kinomes are generally thought to be devoid of aPKs, at least four genes in *T. gondii* are thought to encode these enzymes, the products of which are hypothesized to be part of the Ovoid Mitochondrial Cytoplasmic (OMC) Complex [51], a composite assembly of organelles observed only in growing tachyzoites of *T. gondii*, but its functions are currently unclear. Our analysis revealed two putative MEK-related aPKs in *S. neurona* (see **Table 1**). A representative of an MEK-related aPK (PfPK7) is reported to be expressed during *Plasmodium* infections in humans [52].

2.2 Evolution of *S. neurona* protein kinases

We phylogenetically deciphered evolutionary relationships among the various PK groups in the *S. neurona* kinomes with the PKs in other apicomplexans, which revealed that the *S. neurona* kinome has slightly fewer AGCs (n=9) compared to the kinomes of *T. gondii* (n=11), *N. caninum* (n=13) and *H. hammondi* (n=15). In general, the phylogenetic clustering of the *S. neurona* AGCs mirrored the homologies of these enzymes to those of the three apicomplexans used in this study (**Figure 1**; compare with **Table 1**).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 January 2017 doi:10.20944/preprints201701.0040.v1

Peer-reviewed version available at Pathogens 2017, 6, 12; doi:10.3390/pathogens6010012

Sequence analysis of *S. neurona* AGCs revealed significant divergence with only ~30% sequence similarity. Two *S. neurona* AGCs (SRCN_5610 and SRCN_3990) clearly cluster with *T. gondii* PKAs TGME49_028420 and TGME49_015670 [53] (**Figure 1**). Moreover, SRCN_5610 shares ~60% full length sequence identity with its ortholog. Further, two other *S. neurona* putative PKAs (SRCN_5165 and SRCN_3339), which also distinctly clustered with their *T. gondii* orthologs (TGME49_026030 and TGME49_086470, respectively) (**Figure 1**), contained the characteristic PKA's '*GxGxxG*' motif harboring a glycine triad that structurally forms a hairpin around the ATP binding pocket [53]. It is also notable that the single putative PKG (SRCN_4518) distinctly clustered with its *T. gondii* PKG ortholog (TGME49_111360) (**Figure 1**).

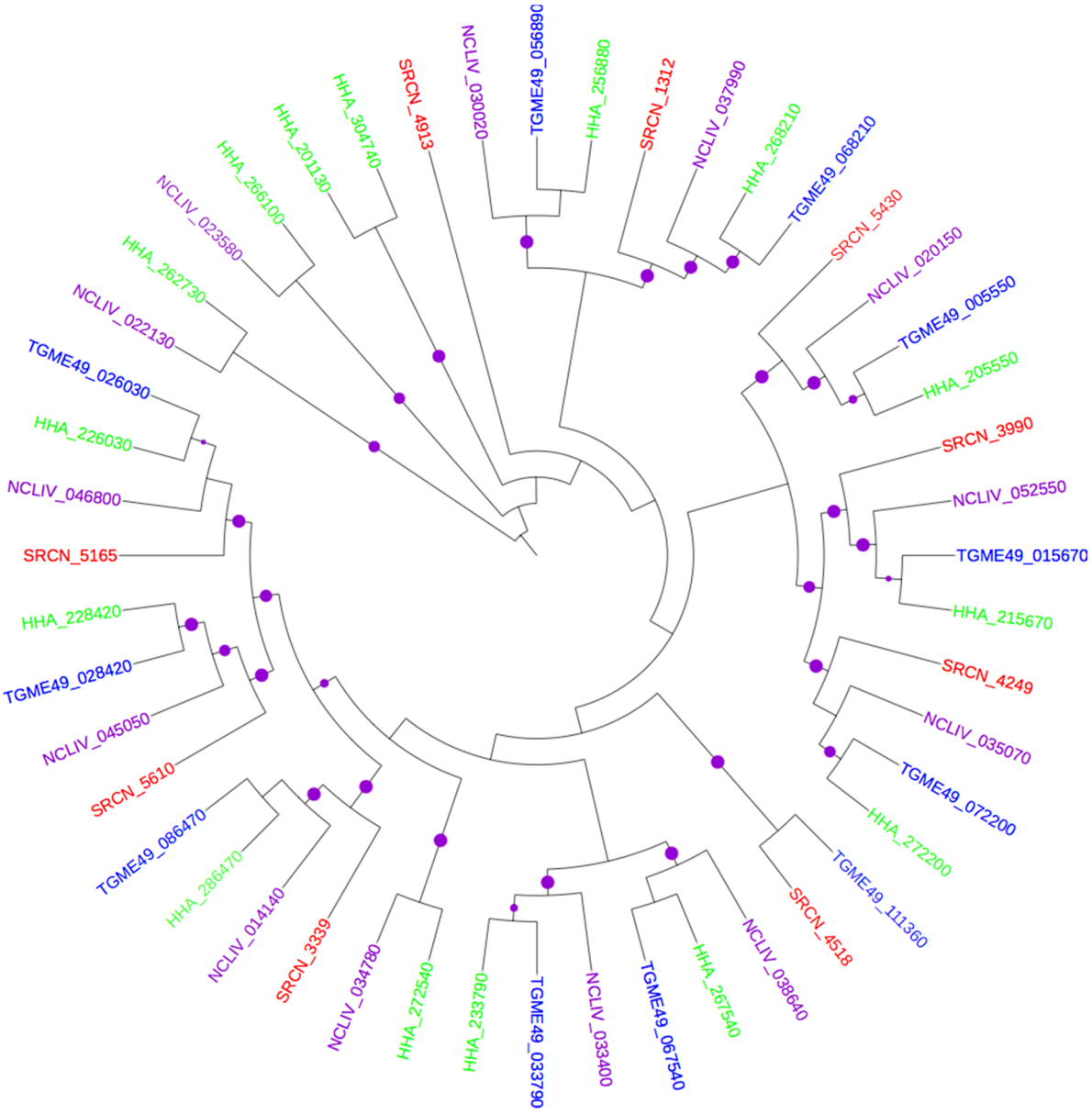

Figure 1. Mid-point rooted ML phylogenetic tree of apicomplexan AGCs. The terminal branches are color-coded for AGCs in the kinomes of *S. neurona* (SRCN; red), *T. gondii*, ME-49 strain (TGME49; blue), *H. hammondi* (HHA; green) and *N. caninum*, Liverpool strain (NCLIV; purple). A solid purple circles on a branch indicates bootstrap support greater than 70. The phylogenetic tree was inferred from a multiple sequence alignment using PhyML with LG amino acid substitution model and gamma model of substitution rate heterogeneity. The tree image was rendered with iTOL.

The SRCN_3339, SRCN_5165, SRCN_5610 and SRCN_4518 possessed an additional functional domain; the AGC-kinase C-terminal domain that contained two of the three conserved

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 January 2017 doi:10.20944/preprints201701.0040.v1

Peer-reviewed version available at Pathogens 2017, 6, 12; doi:10.3390/pathogens6010012

phosphorylation sites in AGCs (data not shown). These conserved sites serve as phosphorylation-regulated switches to control both intra- and inter-molecular interactions [54]. A striking feature in *S. neurona* is that like in *T. gondii*, it lacks PKB and PKC. The putative PDK1 (SRCN_1312) clustered with the *T. gondii* PDPK (TGME49_068210) [53].

Despite the absence of PKC in *S. neurona*, CAMK family members were identified, which perhaps underscores the importance of Ca^{2+} regulation in this apicomplexan. Majority of the identified *S. neurona* CAMKs segregated with their orthologs in *T. gondii*, *N. caninum* and *H. hammondi* in clades with robust bootstraps (**Figure 2**), thus validating the annotation of the CAMKs. Amongst the CAMKs, SRCN_2544 clustered with *T. gondii* PK1 (TGME049_043500) of the AMPK/SNF1 sub-family. There were also three additional SNF1 members in *S. neurona* (SRCN_5410, SRCN_4815 and SRCN_2257), which clustered with *T. gondii* TGME49_115190, TGME49_033900 and TGME49_091050, respectively.

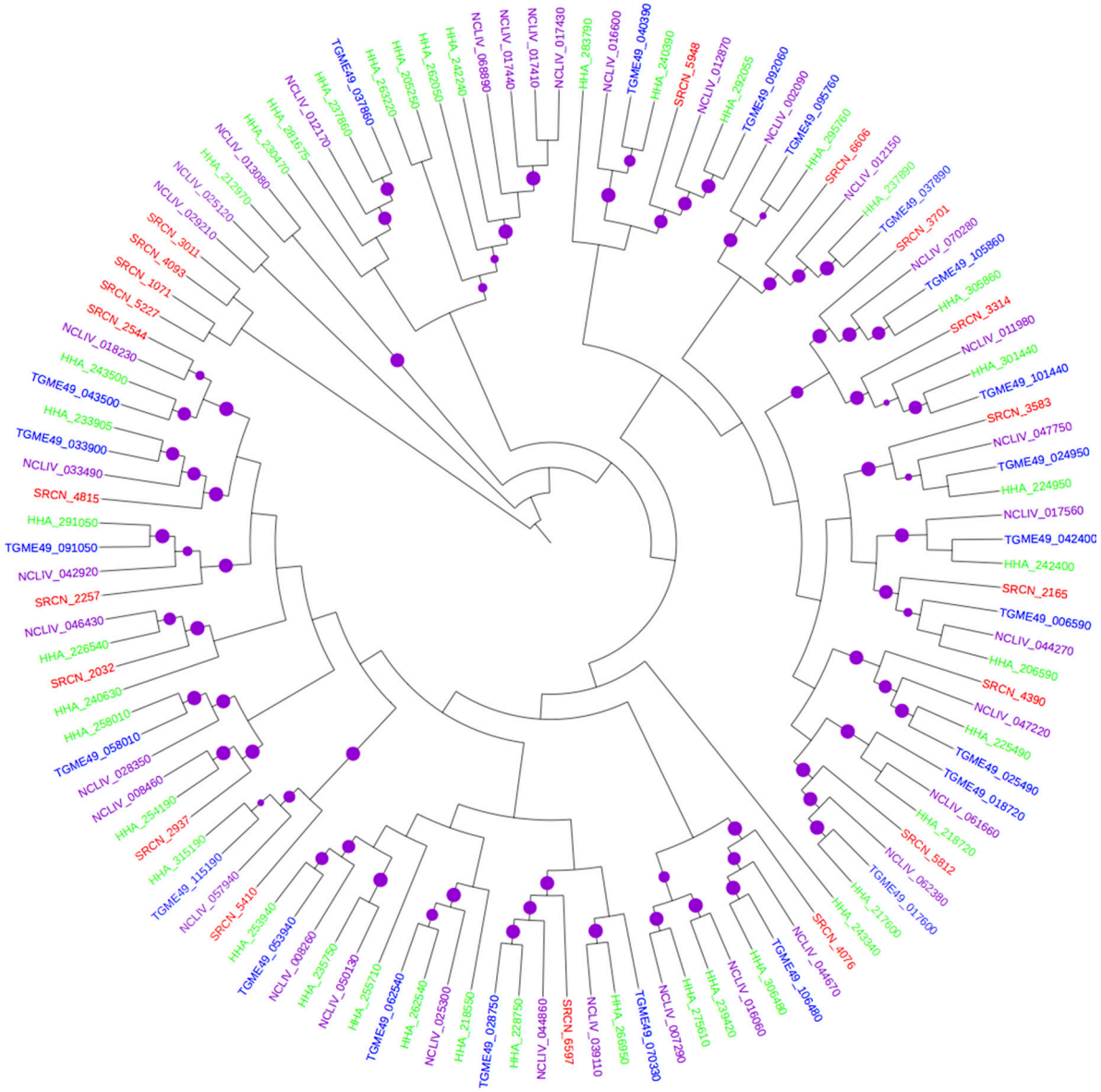

Figure 2. Mid-point rooted ML phylogenetic tree of apicomplexan CAMKs. The terminal branches are color-coded for AGCs in the kinomes of *S. neurona* (SRCN; red), *T. gondii*, ME-49 strain (TGME49; blue), *H. hammondi* (HHA; green) and *N. caninum*, Liverpool strain (NCLIV; purple). A solid purple circles on a branch indicates bootstrap support greater than 70. The phylogenetic tree was inferred from a multiple sequence alignment using PhyML with LG amino acid substitution model and gamma model of substitution rate heterogeneity. The tree image was rendered with iTOL.

Based on the clustering with *T. gondii* CDPK orthologs, 10 *S. neurona* CDPKs, including CDPK1 (SRCN_3314), CDPK2 (SRCN_4390), CDPK2B (SRCN_2165), CDPK3 (SRCN_3701), CDPK4 (SRCN_6606), CDPK5 (SRCN_3583), CDPK6 (SRCN_3011), CDPK7 (SRCN_6597), CDPK8 (SRCN_5948) and CDPK9 (SRCN_5812) were identified (**Figure 2**). This result implies that *S. neurona* has potentially lost at least two CDPKs (compared to the 12 CDPK that have been reported in *T. gondii* [21]). The possible loss notwithstanding, *S. neurona* contained the 6 CDPKs that are expressed and are well-conserved in most apicomplexans (i.e. CDPK1, CDPK3, CDPK4, CDPK5, CDPK6, and CDPK7) [55]. Sequence analysis revealed that, like in other apicomplexans, all identified *S. neurona* CDPKs except CDPK7 (SRCN_6597) contained both a kinase domain and a Ca2+-binding domain known as the EF-hand domain [21]. Like its *T. gondii* ortholog, TGME49_028750 (TgCDPK7), the *S. neurona* CDPK7 (SRCN_6597) contains a pleckstrin-homology (PH) domain just upstream of its PK domain [55]. The domain architecture in CDPKs is such that kinase activity is stimulated upon Ca2+-binding. Finally, the clustering of SRCN_5227, SRCN_1071, SRCN_3011 and SRCN_4093 in Figure 2 imply possible species-specific expansion of the CAMKs in *S. neurona*.

Majority of the putative CMGCs identified in *S. neurona* clustered with robust bootstraps with the conserved CMGCs in *T. gondii*, *N. caninum* and *H. hammondi* (**Figure 3**). Notable were the clustering of the CDKs, including CDK7 (SRCN_4674, SRCN_2759 and SRCN_761), CDK10 (SRCN_895) and CDK11 (SRCN_977). In addition, comparing the phylogenetic clustering of the CDKs in Figure 3 and the annotations presented in Table 1, the *S. neurona* kinome also appeared to have putative CDK5 (SRCN_4801, SRCN_1104 and SRCN_6346), all of which were supported by robust bootstraps. It should be noted that CDKs are amongst the principle molecular switches that regulate cell cycle progression in the apicomplexan parasites [56]. Also notable was the clustering in the same clade of three MAPKs (SRCN_6472, SRCN_4209 and SRCN_5365) and two GSKs (SRCN_1731 and SRCN_1732) (**Figure 3**).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 January 2017 doi:10.20944/preprints201701.0040.v1

Peer-reviewed version available at Pathogens 2017, 6, 12; doi:10.3390/pathogens6010012

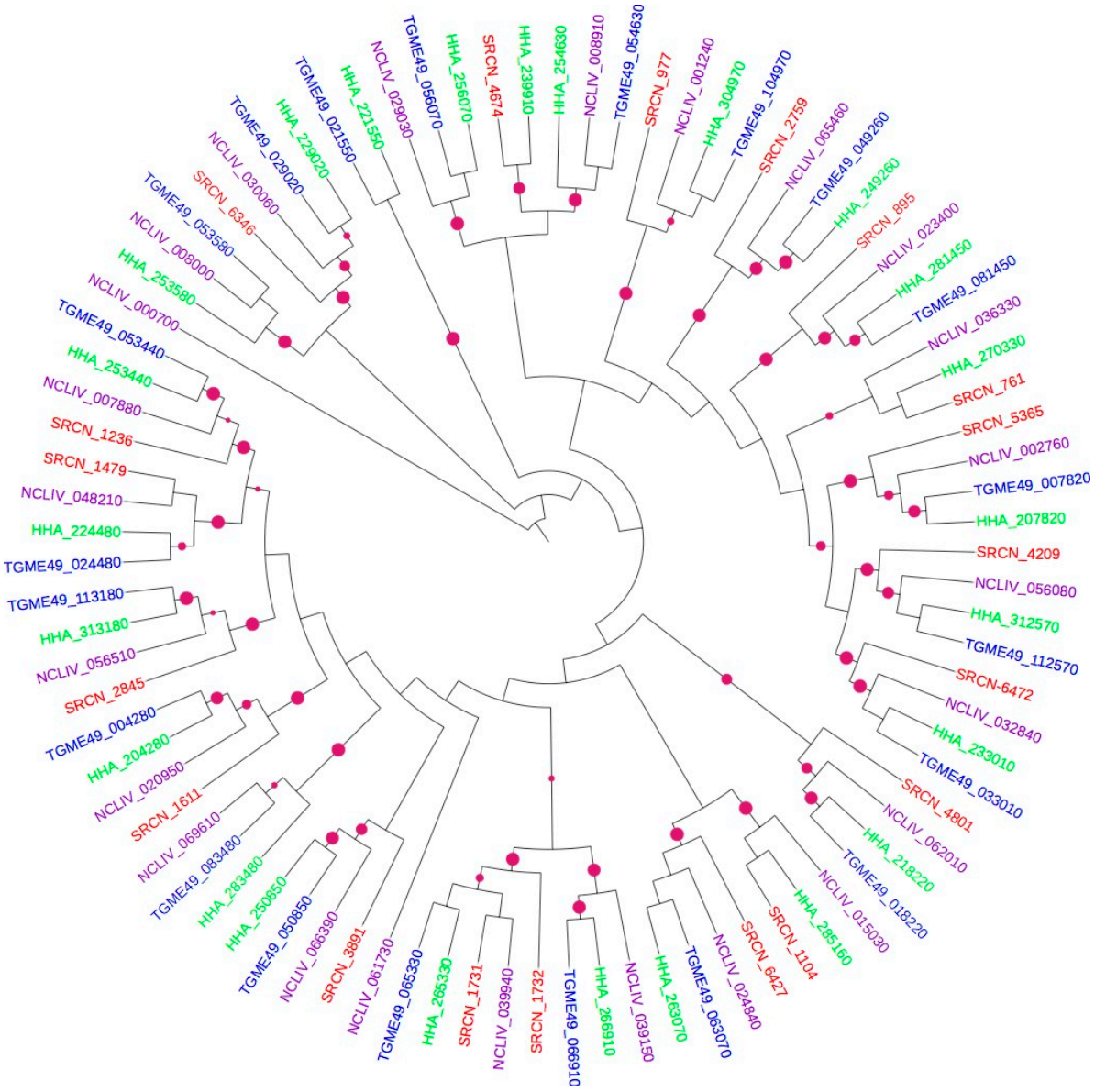

Figure 3. Mid-point rooted ML phylogenetic tree of apicomplexan CMGCs. The terminal branches are color-coded for AGCs in the kinomes of *S. neurona* (SRCN; red), *T. gondii*, ME-49 strain (TGME49; blue), *H. hammondi* (HHA; green) and *N. caninum*, Liverpool strain (NCLIV; purple). A solid purple circles on a branch indicates bootstrap support greater than 70. The phylogenetic tree was inferred from a multiple sequence alignment using PhyML with LG amino acid substitution model and gamma model of substitution rate heterogeneity. The tree image was rendered with iTOL.

In the OPK family, SRCN_4528 and SRCN_2630 are putative NEKs given their clustering with *T. gondii* NEK kinases TGME49_119700 and TGME49_094260, respectively (**Figure 4**). The *S. neurona* SRCN_108 and SRCN_3444 are putative ULKs due to their clustering with their *T. gondii* ULK orthologs (TGME49_035750 and TGME49_040630, respectively). SRCN_3669 clustered with TGME49_066950, a TBC-domain (Tre-2/Bub2/Cdc16) containing kinase. Further, there seemed to be a species-specific expansion of *S. neurona* Aurora kinases, SRCN_1606, SRCN_6157 and SRCN_2403 which cluster with *T. gondii* aurora kinase TGME49_003010; the clade containing these kinases however was not supported by robust bootstrap values (see **Figure 4**). SRCN_286 is a putative Wee kinase given that it clusters with Wee kinases from *T. gondii* (TGME49_073690) as well as from *H. hammondi* (HHA_273690; compare Figure 4 with Table 1).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 January 2017 doi:10.20944/preprints201701.0040.v1

Peer-reviewed version available at Pathogens 2017, 6, 12; doi:10.3390/pathogens6010012

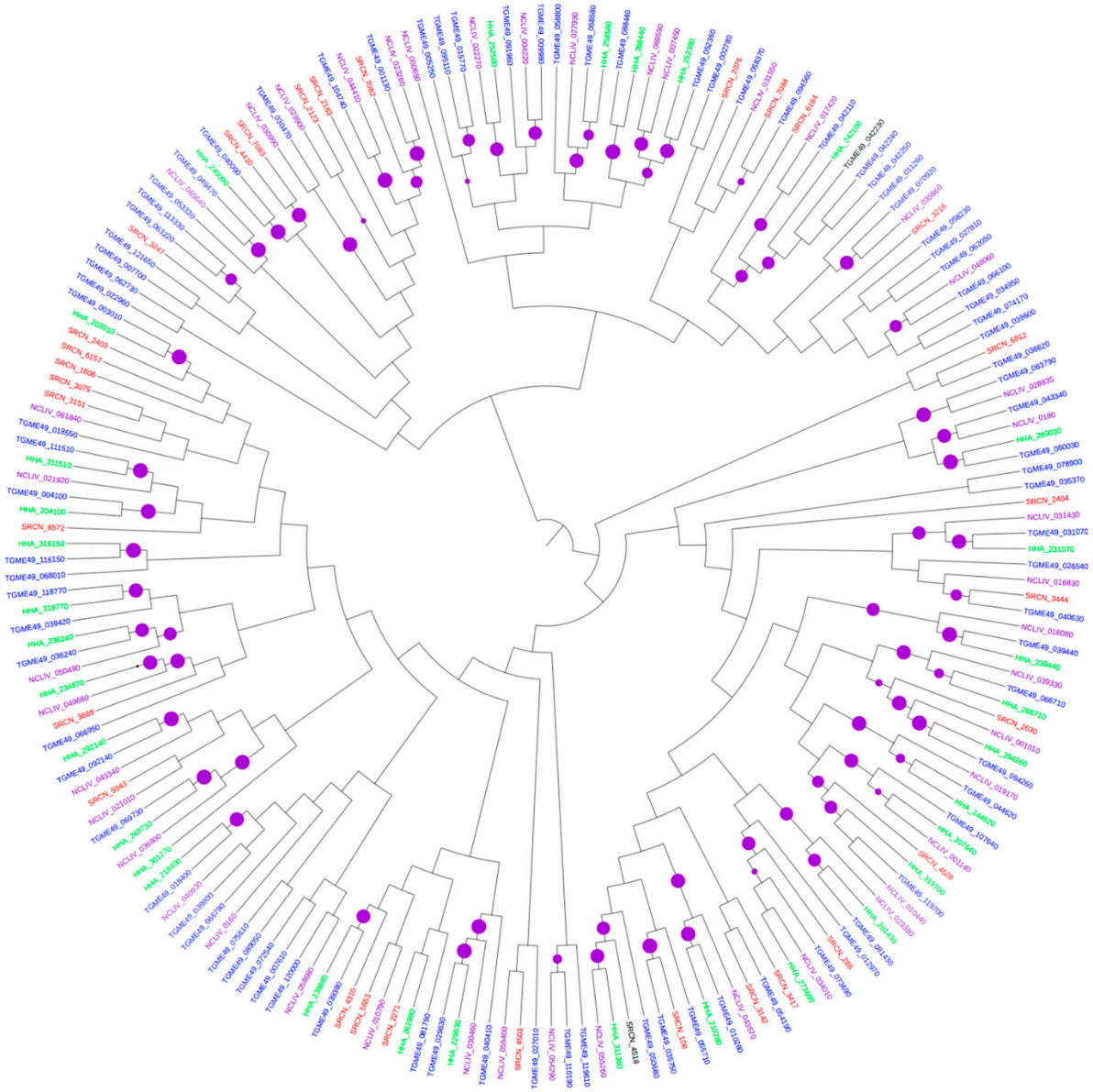

Figure 4. Mid-point rooted ML phylogenetic tree of apicomplexan OPKs. The terminal branches are color-coded for AGCs in the kinomes of *S. neurona* (SRCN; red), *T. gondii*, ME-49 strain (TGME49; blue), *H. hammondi* (HHA; green) and *N. caninum*, Liverpool strain (NCLIV; purple). A solid purple circles on a branch indicates bootstrap support greater than 70. The phylogenetic tree was inferred from a multiple sequence alignment using PhyML with LG amino acid substitution model and gamma model of substitution rate heterogeneity. The tree image was rendered with iTOL.

The clustering of putative *S. neurona* ROPKs (**Figure 4**) was also notable. For instance, SRCN_7082 clustered with *T. gondii* ROP33 (TGME49_001130), implying that it is a ROP33 (**Table 1**). SRCN_6184 clustered with *T. gondii* ROP19A (TGME49_042240) and ROP38 (TGME49_042110). It should however be noted that the sequence similarities with the ROP38 was low, implying that this could be a ROP19A (based on annotations presented in Table 1). Other putative *S. neurona* ROPKs included SRCN_3247, which clustered with ROP27 (TGME49_113330), SRCN_2076 with ROP25 (TGME49_002780). Here, it is notable that the clustering of SRCN_2076 the *T. gondii* ROP25 was supported with low bootstrap value, potentially meaning that based on the annotations (Table 1), SRCN_2076 is an ortholog to ROP30. SRCN_2123 and SRCN_2183 clustered with ROP35 (TGME49_104740), while SRCN_7084 clustered with ROP37 (TGME49_094560). Notably, from the phylogenetic tree in figure 4, SRCN_7083 and SRCN_4410 are probably duplicated forms of ROP34 given their clustering with *T. gondii* TGME49_040090. There was also indication of species-expansion of the *S. neurona* putative ROP35 (SRCN_2123 and SRCN_2183; see **Figure 4**). Although annotations

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 January 2017 doi:10.20944/preprints201701.0040.v1

Peer-reviewed version available at Pathogens 2017, 6, 12; doi:10.3390/pathogens6010012

(Table 1) showed that SRCN_4310 was a putative ROP33 (39% sequence similarity to *H. hammondi* ROP33), this enzyme clustered with a hypothetical protein (TGME49_039380) of *T. gondii* (Figure 4). This discrepancy between annotation and phylogeny could be due to differences in the *T. gondii*. Another apparent discrepancy was the annotation of SRCN_4503 as a putative eIF2K-B (74% sequence similarity to the *T. gondii* TgCatPRC2 [IF2K-B]; see **Table 1**); but phylogeny showed clustering of this protein with a putative *T. gondii* ROP30 (TGME49_027010) (see **Figure 4**). Note that in this figure, the phylogeny bootstraps were low in the segregation of SRCN_4503 with ROP30. Moreover, SRCN_3151 and SRCN_3075 both segregated with TGME49_018550, a PIK3R4 kinase-related protein; from the phylogeny, SRCN_3151 and SRCN_3075 probably represent duplications. SRCN_6572 falls in the same clade with TGME49_004100, an eIF2 kinase IF2K-C, and TGME49_111510, an eIF2 kinase IF2K-B. NEKs are involved in cell cycle regulation, while Aurora kinases play pivotal roles in endodyogeny, duplication rate and parasite virulence [28]. Taken together, the presence of a variety of ROPKs in *S. neurona* is interesting given the fact that in *T. gondii*, ROPKs are key virulence factors [57].

2.3 Structural modeling

Using I-TASSER, we constructed the three-dimensional (3-D) homology models for SRCN_3247 using 4B6L, 4C0T, 1KOB, 4ZHX, 4IC7 and 4CZU as templates, and SRCN_5165 using 1CDK, 1SMK, 4WIH, 4Z84, 3PFQ, 4LQS and 1RDQ as templates. SRCN_3247 and SRCN_5165 are putative ROP27 and PKA respectively. From the predicted models for each of the sequences, the model with the highest C-score, a score that estimates the quality of the models produced, was used. The best SRCN_3247 model had a C-score of -0.46 and an estimated TM-score of 0.65 (**Figure 5A**) while the best SRCN_5165 model had a C-score of 1.6 and a TM-score of 0.94 (**Figure 5B**).

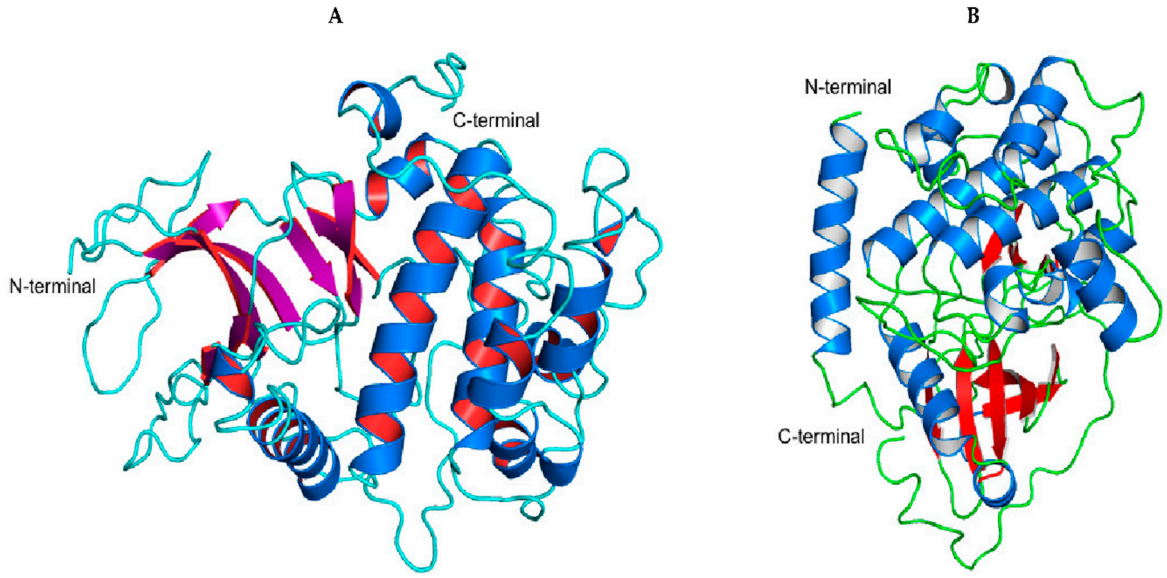

Figure 5. (A) Homology model of *S. neurona* SRCN_3247 (ROP27) generated based on template structures 4B6L, 4C0T, 1KOB, 4ZHX, 4IC7 and 4CZU. (B) Homology model of SRCN_5165 generated based on template structures 1CDK, 1SMK, 4WIH, 4Z84, 3PFQ, 4LQS and 1RDQ.

Evaluation of the models' quality and suitability for docking using PROCHECK revealed that SRCN_3247 and SRCN_5165 had 97.4% and 99.7% of the residues in the favored/allowed regions, respectively (**Supplementary Figure 1A&B**). The SRCN_3247 model contained 9 beta-sheets, 11 alpha and 5 3,10 helices while SRCN_5165 contains 9 beta-sheets, 11 alpha and 6 3,10 helices. Using the improved Protein Block Alignment (iPBA) webserver [58], structural superposition of the models to the top-ranking templates was performed for prediction of pair-wise structural alignment. The structural similarity score was measured using root mean square deviation (RMSD) between both Cα atom of the modeled protein and the corresponding template. The RMSD for the alignment of SRCN_3247 and 4B6L was 0.82 Å for the 276 aligned residues (**Figure 6A**) while that for

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 January 2017 doi:10.20944/preprints201701.0040.v1

Peer-reviewed version available at Pathogens 2017, 6, 12; doi:10.3390/pathogens6010012

SRCN_5165 and 1CDK was 0.46 for 335 aligned residues (**Figure 6B**). The obtained RMSD values indicated that in both cases, the models and the templates had similar folds.

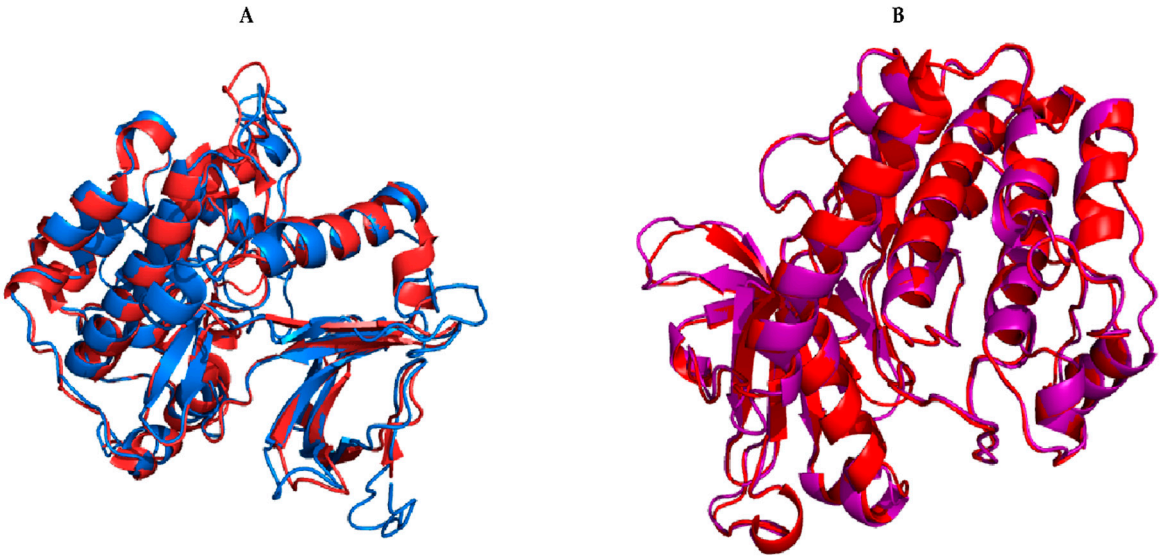

Figure 6. A. Superimposition of SRCN_3247 (marine) and 4B6L template (red). B. Superimposition of SRCN_5165 (purple) and its template, 1CDK (red).

Analysis of putative binding pockets using DoGSiteScorer [59] revealed 14 potential pockets (P0-P13) on the surface of the SRCN_3247 model and 13 potential pockets (P0-P12) on the surface of the SRCN_5165 model. The largest pocket in SRCN_3247 (P0; **Figure 7A**) had a volume of 624.06 $Å^3$ and a drug score of 0.79 while the largest pocket in SRCN_5165 (P0; **Figure 8**) had a volume of 760.51 $Å^3$ and a drug score of 0.77.

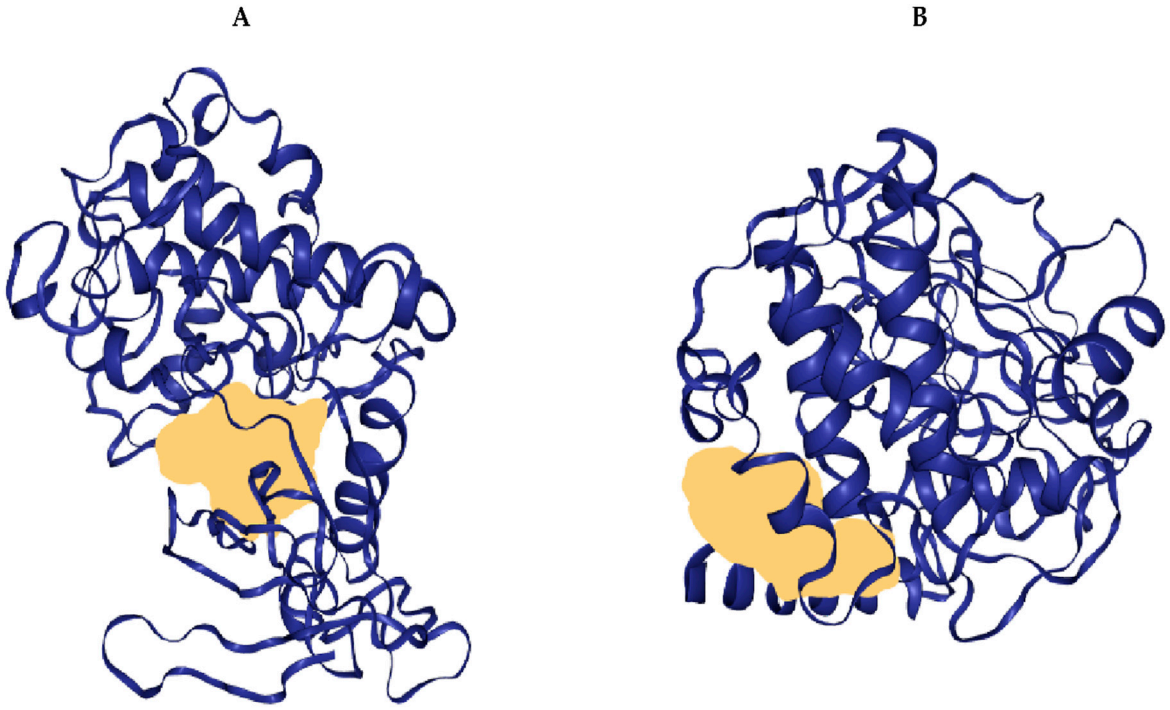

Figure 7. A cartoon of the largest binding pockets of SRCN_3247 (Panel A) and SRCN_5165 (panel B).

Here, it should be noted that the presence of a binding pocket does not necessarily imply that a target protein is druggable. However, the binding volume is one of the important cavity properties

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 January 2017 doi:10.20944/preprints201701.0040.v1

Peer-reviewed version available at Pathogens 2017, 6, 12; doi:10.3390/pathogens6010012

that influence the druggability of a particular target protein, of which most of druggable proteins have volumes of between 500-1000 Å3 [60].

2.4 Molecular docking of SRCN_3247 and SRCN_5165 with PK inhibitors

After the structural modeling, the binding modes of known kinase inhibitors to SRCN_3247 and SRCN_5165 was assessed by molecular docking of the chemical ligands into the predicted active site pockets of the modeled structures. For the docking analyses, the following four inhibitors were used: - BAY61-3606, an ATP-competitive inhibitor of mammalian spleen tyrosine kinase (Srk) [61]; Flavopiridol, a synthetic flavonoid that inhibits a wide range of cyclin-dependent kinases [62]; purvalanolB-B (PurB) is a potent CDK inhibitor [63]; and Piceatannol, a naturally occurring hydroxylated analogue of resveratrol that inhibits Syk kinase [64]. With a binding affinity of -8.4 kcal/mol, the overall binding pose of BAY61-3606 in SRCN_5165 is shown in **Figure 8A**. The binding affinity of flavopiridol to SRCN_5165 was -9.3 kcal/mol (**Figure 8B**). The binding energies for purB and piceatannol to SRCN_5165 were -6.1 kcal/mol (**Figure 8C**), and -6.0 kcal/mol (**Figure 8D**), respectively.

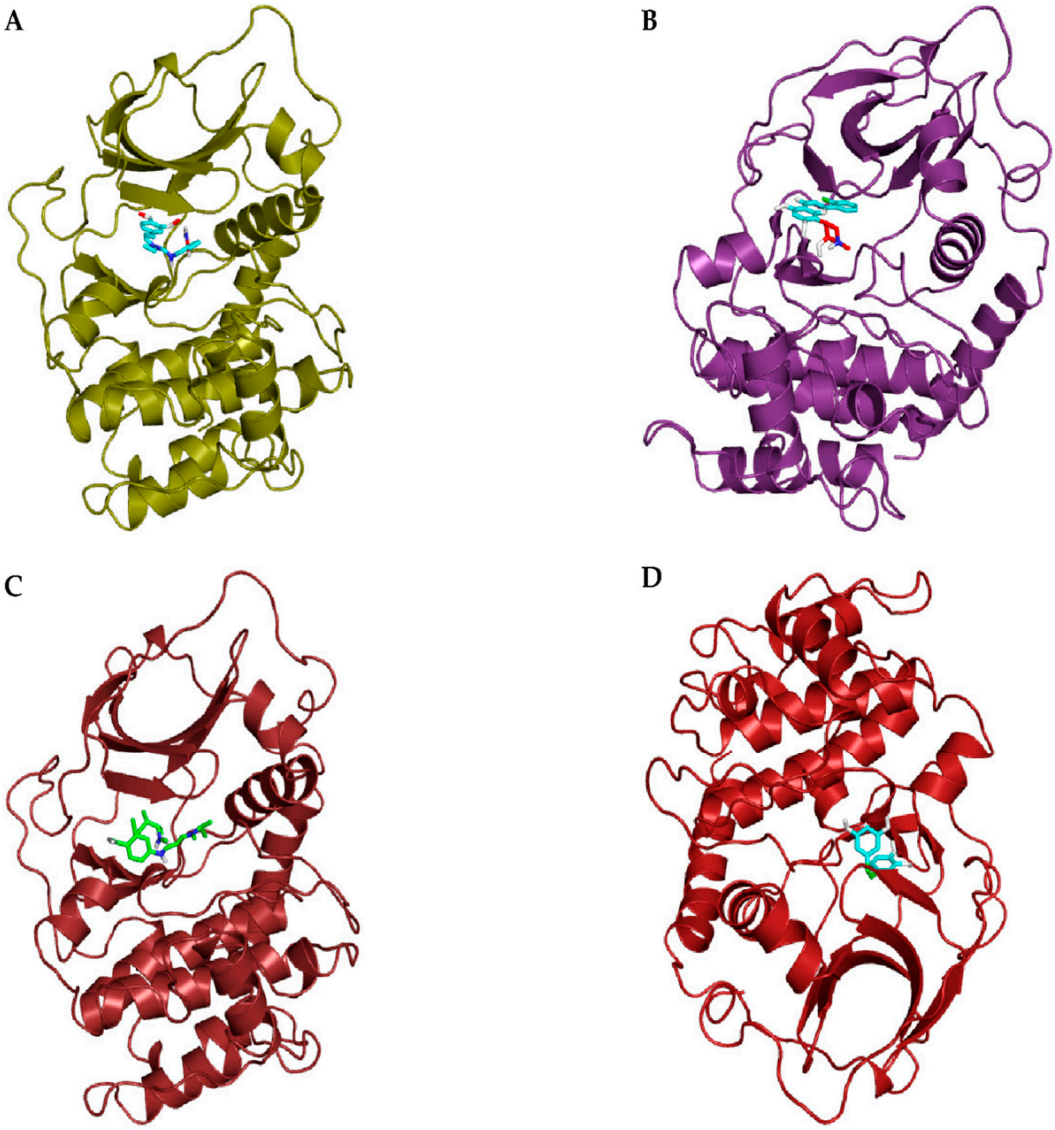

Figure 8. (A) Binding of BAY61-3606 to the active site pocket of SRCN_5165; (B) Binding of flavopiridol to the active site pocket of SRCN_5165; (C) Binding of purvalanolB to the active site pocket of SRCN_5165; and (D) Binding of piceatannol to the active site pocket of SRCN_5165. The inhibitors and structural models are shown in stick and cartoon representations, respectively.

Peer-reviewed version available at Pathogens 2017, 6, 12; doi:10.3390/pathogens6010012

For SRCN_3247, BAY-61-3606 bound to the active site with an affinity of -7.2 kcal/mol (**Figure 9A**), while the affinities for flavopiridol, purB and piceatannol were -7.9 kcal/mol (**Figure 9B**), -6.9 kcal/mol (**Figure 9C**), and -7.3 kcal/mol (**Figure 9D**), respectively.

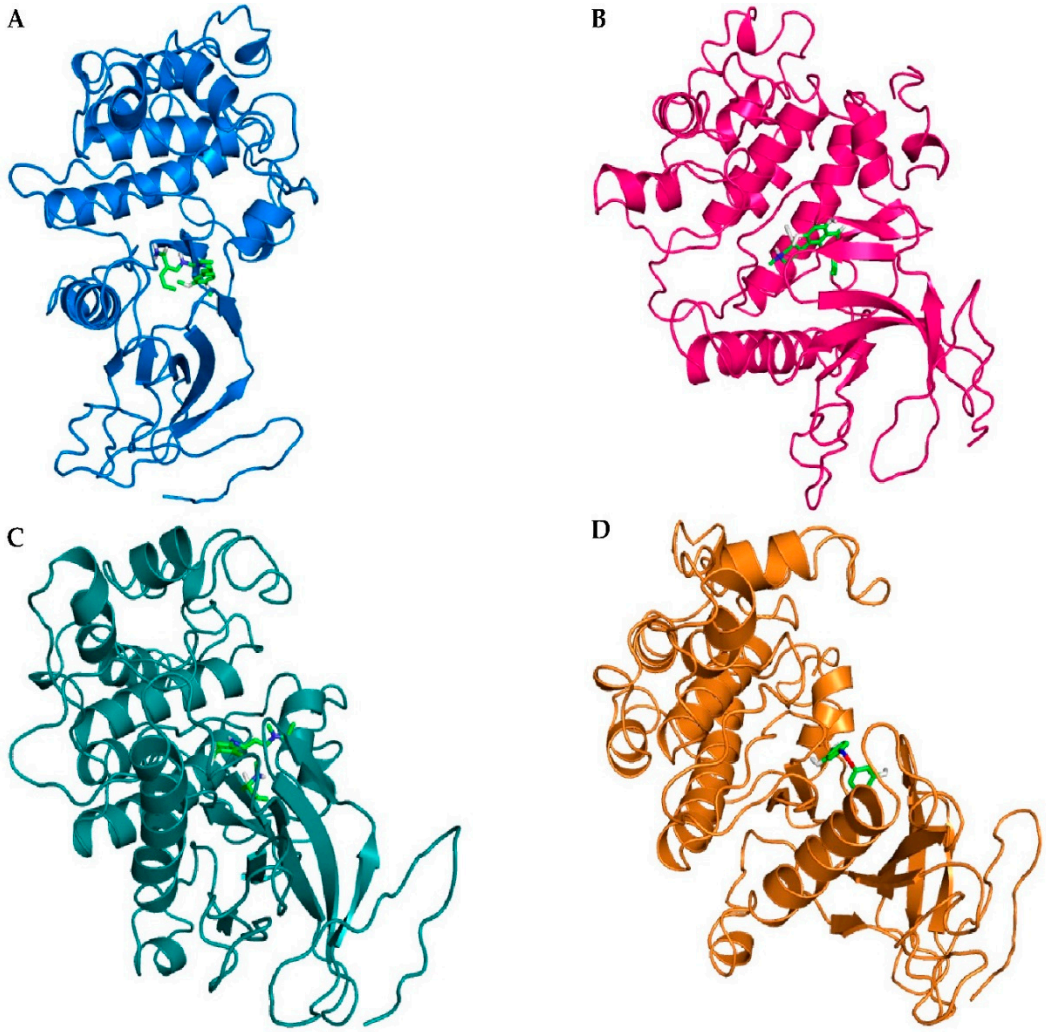

Figure 9. (A) Binding of BAY61-3606 to the active site pocket of SRCN_3247; (B) Binding of flavopiridol to the active site pocket of SRCN_3247; (C) Binding of purvalanolB to the active site pocket of SRCN_3247; and (D) Binding of piceatannol to the active site pocket of SRCN_3247. The inhibitors and structural models are shown in stick and cartoon representations, respectively.

Taken together, by structural modeling and docking analysis, the above results indicate that the putative *S. neurona* PKA and ROP27 could be considered as drug targets. The data from the modeling and docking could pave way for further laboratory experiments to design potential drugs against EMP.

3. Discussion

We used different bioinformatics approaches to identify and classify the PKs of *S. neurona*, and docked four well-known inhibitors into modelled structures of two PK representatives. The critical roles performed by PKs make them attractive targets for novel therapeutics, and as such, the current study provides a rich source of potential drug targets, vaccine candidates or diagnostic proteins for developing new treatment and diagnostics strategies. PKs are considered viable drug targets because their catalysis mechanism and overall structure are conserved. Because small molecules can bind to the PKs' catalytic clefts [65], many kinase inhibitors (e.g. imatinib, trastuzumab and lapatinib) have been developed to treat various human diseases [66].

The kinomes of apicomplexans range from 35 PKs (in *Babesia bovis*) to 135 PKs (in *T. gondii*) [35]. We identified a total of 92 putative PKs in the kinome of *S. neurona*, compared to the PKs reported in the kinomes of *T. gondii* (*n*=135), *N. caninum* (*n*=130) and *H. hammondi* (*n*=124) [15]. Although the total number of *S. neurona* PKs appeared markedly reduced compared to that of its close coccidian relatives (*T. gondii* and *N. caninum* [27]), taken as a percentage of total genome size, the proportion of *S. neurona* PKs is comparable to the 2% observed in humans [13] and other coccidians [27]. The contraction of the *S. neurona* kinome could be attributed to genome compaction, which occasionally offsets lineage-specific expansions of specific gene families. Notably, genome contraction is a common mode of genomic evolution in intracellular parasites, including apicomplexans [67, 68]. As such, the evolution of PKs may be in tandem to the overall genomic adaptive strategies of these parasites.

Using a hierarchical scheme based on the major PK groups, the *S. neurona* kinases could be classified and phylogenetically clustered into the various PK families. A complement of nine putative AGC kinases was identified in *S. neurona*, which is reduced compared with that of *T. gondii*, *N. caninum* and *H. hammondi*. Despite this potential gene loss, seven of the nine AGCs (SRCN_5165, SRCN_5610, SRCN_3339, SRCN_4249, SRCN_3990, SRCN_5430 and SRCN_1312) had orthologs in *T. gondii*, *N. caninum* and *H. hammondi*. In agreement with the observation that PKA is conserved in apicomplexans [27], two PKAs (SRCN_5610 and SRCN_3990) were identified in *S. neurona.* In eukaryotes, PKA signaling regulate various cellular responses such as DNA replication [69], cell growth and metabolism [70] and gene transcription by phosphorylating transcription factors such as the cAMP Response Element-Binding Protein (CREB) [71]. In *T. gondii*, increases in cytosolic cAMP levels activate PKA to trigger the developmental switch from the rapidly proliferating tachyzoites to the quiescent bradyzoites [72]. Additionally, two other *S. neurona* AGCs, (SRCN_5165 and SRCN_3339), were putative PKAs given that they contained the characteristic *GxGxxG* motif found in PKA [53]. The PKAs are attractive targets for the disruption of *S. neurona* growth. Notably, based on orthology, *S. neurona* contains a single putative PKG (SRCN_4518) that distinctly clustered with *T. gondii* PKG (TGME49_111360).

In apicomplexans, CAMKs modulate the intracellular Ca^{2+} concentration, which in turn regulates vital processes such as host-cell invasion, protein secretion and parasite differentiation. We identified four AMPK/SNF1 family members (SRCN_2544, SRCN_5410, SRCN_4815 and SRCN_2257). The AMP-activated PK cascade acts as a metabolic sensor that monitors cellular AMP and ATP levels and is activated by an elevation of the AMP:ATP ratio. In yeast, scarcity of nutrients and energy results in the activation of SNF1 [73]. When activated, AMPK restores energy balance by switching off ATP-consuming anabolic pathways and switching on ATP generating catabolic pathways such as fatty acid oxidation [74]. Further, we identified 10 putative CDPKs in the *S. neurona*, including CDPK1 (SRCN_3314), CDPK2 (SRCN_4390), CDPK2A (SRCN_2165), CDPK3 (SRCN_3701), CDPK4 (SRCN_6606), CDPK5 (SRCN_3583), CDPK6 (SRCN_3011), CDPK7 (SRCN_6597), CDPK8 (SRCN_5948) and CDPK9 (SRCN_5812). Compared to the 12 CDPKs reported in *T. gondii* [21], it appears that *S. neurona* had all the six well-conserved apicomplexan CDPKs (CDPK1, CDPK3, CDPK4, CDPK5, CDPK6 and CDPK7), which provide a link between Ca^{2+} signaling and parasite differentiation, motility, invasion and egress [55]. In *T. gondii*, down-regulation of CDPK1 interfered with parasite motility, host cell invasion and egress [44] while disruption of CDPK3 caused defective parasite egress [75]. Further, the essentiality of CDPK6 and CDPK7 in *T. gondii* have recently been demonstrated [76]. Indeed, TgCDPK1 has been targeted for the development of new drugs for toxoplasmosis [77]. Sequence analysis revealed that, similar to other apicomplexans, all identified *S. neurona* CDPKs except CDPK7 (SRCN_6597) contain both a PK domain and a EF-hand (Ca^{2+}-binding) domain [21]. Similar to its *T. gondii* ortholog, TGME49_028750 (TgCDPK7), the *S. neurona* CDPK7 (SRCN_6597) contains a pleckstrin-homology (PH) domain just upstream of its PK domain [55]; the domain architecture is such that kinase activity is stimulated upon Ca^{2+} binding. Moreover, our phylogeny provided clues of possible gene duplications giving rise to SRCN_3990 and SRCN_3011 and well as SRCN_4093 and SRCN_1071). Interestingly, based on phylogenetic analysis, *S. neurona* probably contains four (SRCN_5227, SRCN_1071, SRCN_4093 and SRCN_3011) species- specific CAMKs.

The CMGCs, comprising of CDKs, MAPKs, GSKs, and CLKs coordinate a wide range of cellular functions in different species. For instance, members of the CDK subfamily are major coordinators of cell division in both mitosis and meiosis while the MAPKs subfamily is crucial in regulating cell proliferation, cell differentiation and cell death in eukaryotes [56, 78]. By both annotations and phylogenetic analyses, we identified four putative CDKs sub-family members; CDK5 (SRCN_4801 and SRCN_6346), CDK7 (SRCN_2759, SRCN_4674 and SRCN_761), CDK10 (SRCN_895) and CDK11 (SRCN_977). The finding of CDKs in *S. neurona* suggests that this parasite's cell cycle regulation could be CDK-dependent and perhaps similar to that of higher eukaryotes [79]. The identification of three putative MAPKs (SRCN_4209, SRCN_6472 and SRCN_5365) in *S. neurona* points to the existence of MAPK regulated transduction pathway(s) in this pathogen. Similar to its *T. gondii* ortholog (TgMAPK1), which is a p38α MAPK homolog [80], SRCN_4209 may be potentially involved in parasite proliferation/stage differentiation, stress response and manipulation of the host immunity to enhance virulence. On the other hand, SRCN_6472 and SRCN_5365 may augment the roles of SRCN_4209 in the parasite. In *T. gondii* MAPK1/ERK7 is involved in intracellular proliferation [81]. We also identified two putative GSKs (SRCN_1731 and SRCN_1732). A genome-wide gene knockout approach in *P. falciparum* demonstrated PfGSK-3 to be critical for schizogony of the parasite [82]. Other *S. neurona* CMGC kinases identified include CLK (SRCN_1479), PRP4 (SRCN_2845), DYRK (SRCN_1611) CK2 (SRCN_6427) and SRPK (SRCN_1236).

ROPKs are secreted by *T. gondii* into the host cell and play roles in adhesion, motility and manipulation of immune responses [83]. We identified at 10 putative ROPK sub-family members in the *S. neurona* kinome; - ROP38 (SRCN_6184), ROP33 (SRCN_7082), ROP27 (SRCN_3247), ROP25 (SRCN_2076), ROP37 (SRCN_7084), ROP34 (SRCN_4410 and SRCN_7083), ROP35 (SRCN_2183), SRCN_2123 and ROP23 (SRCN_6812). It has been recently demonstrated that *T. gondii* ROP21 and ROP27 play a role in a constitutive pathway based on their localization in the PV and cyst matrix [84]. Moreover, *T. gondii* ROP35 has been shown to play a crucial role in chronic infection [85]. Although the *S. neurona* genome is more than twice the size of other coccidians whose genomes sequenced so far (e.g. *Toxoplasma* and *Neospora*), it has a considerably reduced number of ROPKs that nevertheless may have vital roles in the parasite's virulence. Specifically, *S. neurona* is devoid of ROP5, ROP16, ROP18 and ROP38 that have been shown to confer virulence and alter the host's cellular signaling pathways [81]. Putatively therefore, *S. neurona* ROPKs may have multiple roles in the survival of the parasite. In search of drug targets against *S. neurona*, the reduced ROPKs with possible multiple roles and absent in the vertebrate host are thus attractive candidates.

We explored the possibility of repositioning existing PK inhibitors as lead compounds in the development of new drugs for EPM, an approach that would substantially lower the cost of drug development. The lead compounds found via repositioning could be modified to increase efficacy. Molecular docking studies were therefore conducted to explore the potential binding of four inhibitors to the structural models of SRCN_3247 and SRCN_3314, representatives of putative *S. neurona* ROP27 and PKA, respectively. Probing the *S. neurona* kinome with small-molecule has many potential advantages. For instance, given that most kinase inhibitors in therapeutic use target multiple kinases with similar binding pockets [65], the promiscuity of such agents lead to polypharmacology [86] that would help minimize drug resistance. However, our modeling and docking analysis are only predictive, meaning that experimental validations are required.

4. Conclusion

The kinome of *S. neurona* contain members of the major classes of PKs, including AGC, CMGC, GSK, CAMK, CK, TKL, aPKs and several PKs in the OPK family. Similar to other apicomplexans, *S. neurona* kinome is devoid of PKC and the conventional TKs, but appears to have reduced MAPK family members. *S. neurona* kinome did not also have some of the ROPKs that have been implicated in the virulence of *T. gondii*. Known kinase inhibitors can be used as starting scaffolds in the search for drugs acting against EPM.

5. Materials and Methods

5.1. Genome-wide identification of putative *S. neurona* PKs

The predicted *S. neurona* proteome was downloaded from Toxoplasma Genomics Resource database (Release 28; version May 2016) [41]. A hidden Markov model (HMM) profile of signature PK domains obtained from the Kinomer database v 1.0 [87] was used to search for *S. neurona* kinases using HMMER v 3.1b2 [88]. The sequences having PK domain (IPR011009) or PK-like domain (IPR000719) were considered as putative kinases. Annotation of the putative kinase sequences was performed by blast search against the non-redundant (nr)-NCBI protein and UNIPROT/SWISSPROT databases at an e-value of 0.0001. Gene ontology (GO) mapping was performed using Blast2GO v 3.3 [89]. The molecular weight (Mw) and isoelectric point (pI) were obtained using ExPASy compute pI/Mw tool [90]. Motifs analysis was performed with the MEME Suite v 4.11.2 [91]. The parameters were as follows: number of repetitions, any; maximum numbers of motifs, 30; and the optimum motif widths, between 6 and 200 residues.

5.2. Phylogenetic analysis

Phylogenetic trees were constructed to decipher the orthologous and paralogous relationships of *S. neurona* kinases. Sequences of *S. neurona* kinase subfamilies obtained using Blast2Go were aligned with their homologs in *T. gondii* [41], *H. hammondi* and *N. caninum* using MUSCLE [92]. The alignments were subsequently manually edited in Jalview [93]. Phylogenetic reconstruction was undertaken using the maximum likelihood program PhyML 3.0 [94] and RAxML v 8.0 [95] and the Bayesian inference program MrBAYES v 3.2 [96]. For PhyML, the LG substitution model was selected assuming an estimated proportion of invariant sites and 4 gamma-distributed rate categories to account for rate heterogeneity across sites. The gamma shape parameter was estimated directly from the data. Robustness of internal branches was evaluated using 100 bootstraps. MrBayes was run for 5, 000,000 generations with two runs and four chains in parallel and a burn-in of 25%. Obtained trees were rendered with the Interactive Tree of Life server (iTOL) [97].

5.3. Structural modeling

The amino acid sequences of the target kinases, a putative ROP27 (SRCN_3247) and a PKA (SRCN_5165) (see Table 1), were retrieved ToxoDB [30]. Structural modelling was carried out using two approaches. The first approach was modelling by satisfaction of spatial restraints as implemented in MODELLER [98]. For this approach, template structures were obtained by threading target sequences in the PHYRE database [99] and selecting top ranked template structures for model building. The second modelling approach was the iterative threading assembly refinement implemented in I-TASSER [100]. The template structures for SRCN_3247 were 4B6L, 4C0T, 1KOB, 4ZHX, 4IC7 and 4CZU, while those for SRCN_5165 were 1CDK, 1SMK, 4WIH, 4Z84, 3PFQ, 4LQS and 1RDQ. The quality of the models was further evaluated using PROCHECK [101] to ensure suitability for carrying out docking studies. PROCHECK outputs a Ramachandran plot that shows the number (%) of residues mapping to the favored, allowed, generally allowed and disallowed regions of the plot. Structural superpositioning was performed using the improved Protein Block Alignment (iPBA) webserver [58]Results were visualized in the PyMOL Molecular Graphics System [102].

5.4 Molecular docking

Docking calculations were carried out using the AutoDock Suite [103]. To compute the free energy of binding on each protein model, essential hydrogen atoms and solvation parameters were added. Dockings were performed with grid boxes sufficient to cover the entire protein of interest using parameters tabulated in Table 2. Four PK inhibitors, BAY61-3606, Flavopiridol, Purvalanol B

Peer-reviewed version available at Pathogens 2017, 6, 12; doi:10.3390/pathogens6010012

and Piceatannol were docked onto SRCN_3247 and SRCN_5165 and their overall binding poses evaluated in terms of binding energies.

Table 2. Parameters used for molecular docking of kinase inhibitors with SRCN_3247 and SRCN_5165, which are representatives of *S. neurona* putative ROP27 and PKA, respectively.

Proteins	Ligands	Center-X	Center-Y	Center-Z	Size-X	Size-Y	Size-Z
SRCN_5165	BAY61-3606	65	63	60	52	62	42
	Flavopiridol	63.929	62.41	65.345	56	58	70
	Purvalanol B	62.0	60.5	63.0	50	55	63
	Piceatannol	58	61.23	63	48	50	58
SRCN_3247	BAY61-3606	62.665	64.359	65.712	62	48	40
	Flavopiridol	60.157	64.359	65.52	60	55	40
	Purvalanol B	62.34	60.15	63.45	56	55	43
	Piceatannol	63.23	62.43	60.15	60	56	45

All grid boxes were formed at spacing size of 1.0 Angstrom. Following the docking procedure, the resulting PDBQT output files were visualized using PyMOL [102] and all ligand conformations were analyzed for their binding affinities.

Supplementary Materials

Supplementary Figure S1: Ramachandran plots generated using PROCHECK to evaluate the robustness three-dimensional models SRCN_5165 (Panel A) and SRCN_3247 (Panel B), which are putative PKA and ROP27, respectively, in *S. neurona* kinome. Here, the red, yellow, light yellow, white regions indicate the favored, allowed, generously allowed and disallowed regions, respectively. The Phi and Psi are dihedral angles that determine torsion of peptide bonds.

Acknowledgments

The authors thank Dr. Peter Waweru for useful discussions and Egerton University for funds to cover the costs to publish in open access.

Author Contributions

E.M conceived and designed the experiments; E.M and H.K performed the experiments, analyzed the data and wrote the paper. The authors read an approved the final version of the manuscript.

Conflicts of Interest

The authors declare that there is no conflict of interest in this work.

References

1. Dubey, J. P.; Lindsay, D. S.; Saville, W. J.; Reed, S. M.; Granstrom, D. E.; Speer, C. A. A review of *Sarcocystis neurona* and equine protozoal myeloencephalitis (EPM). *Vet. Parasitol.*, **2001**, *95* (2-4), 89-131. doi: 10.1016/S0304-4017(00)00384-8.

2. Reed, S. M.; Furr, M.; Howe, D. K.; Johnson, A. L.; MacKay, R. J.; Morrow, J. K.; Pusterla, N.; Witonsky, S. Equine protozoal myeloencephalitis: an updated consensus statement with a focus on parasite biology, diagnosis, treatment, and prevention. *J. Vet. Intern. Med.*, **2016**, *30* (2), 491-502. doi: 10.1111/jvim.13834.

3. Dubey, J. P.; Howe, D. K.; Furr, M.; Saville, W. J.; Marsh, A. E.; Reed, S. M.; Grigg, M. E. An update on *Sarcocystis neurona* infections in animals and equine protozoal myeloencephalitis (EPM). *Vet. Parasitol.*, **2015**, *209* (1-2), 1-42. doi: 10.1016/j.vetpar.2015.01.026.

4. Howe, D. K.; MacKay, R. J.; Reed, S. M. Equine protozoal myeloencephalitis. *Vet. Clin. North Am. Equine Pract.*, **2014**, *30* (3), 659-675. doi: 10.1016/j.cveq.2014.08.012.

5. Colahan, P. T.; Bailey, J. E.; Cheeksa, J. P.; Jones, G. L.; Yangc, M. Effect of sulfadiazine and pyrimethamine on selected physiologic and performance parameters in athletically conditioned thoroughbred horses during an incremental exercise stress test. *Vet. Therap.*, **2002**, *3* (1), 49-63. PMID:12050828.

6. McClure, S. R.; Palma, K. G. Treatment of equine protozoal myeloencephalitis with nitazoxanide. *J. Equine Vet. Sci.*, **1999**, *19* (10), 639-641. doi: 10.1016/S0737-0806(06)82197-0.

7. Bernard, W. V.; Beech, J. Neurological examination and neurological conditions causing gait deficits. In *Diagnosis and Management of Lameness in the Horse*, 2 ed.; Ross, M. W., Dyson, S. J., Eds.; Elsevier Saunders: 2003; pp 135-145.

8. Warschauer, B. A.; Sondhof, A. Equine protozoal myeloencephalitis. *Iowa State Univ. Vet.*, **1998**, *60* (2), Article 10.

9. Dubremetz, J. F.; Garcia-Reguet, N.; Conseil, V.; Fourmaux, M. N. Apical organelles and host-cell invasion by Apicomplexa. *Int. J. Parasitol.*, **1998**, *28* (7), 1007-1013. doi: 10.1016/S0020-7519(98)00076-9.

10. Sadak, A.; Taghy, Z.; Fortier, B.; Dubremetz, J. F. Characterization of a family of rhoptry proteins of *Toxoplasma gondii*. *Mol. Biochem. Parasitol.*, **1988**, *29* (2-3), 203-211. doi: 10.1016/0166-6851(88)90075-8.

11. Cesbron-Delauw, M. F.; Gendrin, C.; Travier, L.; Ruffiot, P.; Mercier, C. Apicomplexa in mammalian cells: trafficking to the parasitophorous vacuole. *Traffic.*, **2008**, *9* (5), 657-664. doi: 10.1111/j.1600-0854.2008.00728.x.

12. Hanks, S. K.; Hunter, T. Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J.*, **1995**, *9* (8), 576-596. PMID:7768349.

13. Manning, G.; Whyte, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S. The protein kinase complement of the human genome. *Science*, **2002**, *298* (5600), 1912-1934. doi: 10.1126/science.1075762.

14. Manning, G.; Plowman, G. D.; Hunter, T.; Sudarsanam, S. Evolution of protein kinase signaling from yeast to man. *Trends Biochem. Sci.*, **2002**, *27* (10), 514-520. doi: 10.1016/S0968-0004(02)02179-5.

15. Miranda-Saavedra, D.; Gabaldón, T.; Barton, G. J.; Langsley, G.; Doerig, C. The kinomes of apicomplexan parasites. *Microb. Infect.*, **2012**, *14* (10), 796-810. doi: 10.1016/j.micinf.2012.04.007.

16. Talevich, E.; Mirza, A.; Kannan, N. Structural and evolutionary divergence of eukaryotic protein kinases in Apicomplexa. *BMC Evol. Biol.*, **2011**, *11* 321. doi: 10.1186/1471-2148-11-321.

17. Talevich, E.; Tobin, A. B.; Kannan, N.; Doerig, C. An evolutionary perspective on the kinome of malaria parasites. *Phil. Trans. R. Soc. B,* **2012**, *367* (1602), 2607-2618. doi: 10.1098/rstb.2012.0014.

18. Kannan, N.; Taylor, S. S.; Zhai, Y.; Venter, J. C.; Manning, G. Structural and functional diversity of the microbial kinome. *PLoS Biol.,* **2007**, *5* (3), e17. doi: 10.1371/journal.pbio.0050017.

19. Miranda-Saavedra, D.; Barton, G. J. Classification and functional annotation of eukaryotic protein kinases. *Proteins: Struct. Funct. Bioinf.,* **2007**, *68* (4), 893-914. doi: 10.1002/prot.21444.

20. Kumar, A.; Vaid, A.; Syin, C.; Sharma, P. PfPKB, a novel protein kinase B-like enzyme from *Plasmodium falciparum* I. Identification, characterization, and possible role in parasite development. *J. Biol. Chem.,* **2004**, *279* (23), 24255-24264. doi: 10.1074/jbc.M312855200.

21. Billker, O.; Lourido, S.; Sibley, L. D. Calcium-dependent signaling and kinases in apicomplexan parasites. *Cell Host Microbe,* **2009**, *5* (6), 612-622. doi: 10.1016/j.chom.2009.05.017.

22. Abdi, A.; Eschenlauer, S.; Reininger, L.; Doerig, C. SAM domain-dependent activity of PfTKL3, an essential tyrosine kinase-like kinase of the human malaria parasite *Plasmodium falciparum*. *Cell. Mol. Life Sci.,* **2010**, *67* (19), 3355-3369. doi: 10.1007/s00018-010-0434-3.

23. Wang, Z.; Wang, S.; Wang, W.; Gu, Y.; Liu, H.; Wei, F.; Liu, Q. Targeted disruption of CK1a in *Toxoplasma gondii* increases acute virulence in mice. *Eur. J. Protistol.,* **2016**, *56* 90-101. doi: 10.1016/j.ejop.2016.07.006.

24. Agarwal, S.; Kern, S.; Halbert, J.; Przyborski, J. M.; Baumeister, S.; Dandekar, T.; Doerig, C.; Pradel, G. Two nucleus-localized CDK-like kinases with crucial roles for malaria parasite erythrocytic replication are involved in phosphorylation of splicing factor. *J. Cell. Biochem.,* **2011**, *112* (5), 1295-1310. doi: 10.1002/jcb.23034.

25. Andrade, L. F.; Nahum, L. A.; Avelar, L. G.; Silva, L. L.; Zerlotini, A.; Ruiz, J. C.; Oliveira, G. Eukaryotic protein kinases (ePKs) of the helminth parasite *Schistosoma mansoni*. *BMC Genomics,* **2011**, *12* 215. doi: 10.1186/1471-2164-12-215.

26. Low, H.; Lye, Y. M.; Sim, T. S. Pfnek3 functions as an atypical MAPKK in *Plasmodium falciparum*. *Biochem. Biophys. Res. Commun.,* **2007**, *361* (2), 439-444. doi: 10.1016/j.bbrc.2007.07.047.

27. Talevich, E.; Kannan, N. Structural and evolutionary adaptation of rhoptry kinases and pseudokinases, a family of coccidian virulence factors. *BMC Evol. Biol.,* **2013**, *13* (1), 117. doi: 10.1186/1471-2148-13-117.

28. Berry, L.; Chen, C.-T.; Reininger, L.; Carvalho, T. G.; El Hajj, H.; Morlon-Guyot, J.; Bordat, Y.; Lebrun, M.; Gubbels, M.-J.; Doerig, C. The conserved apicomplexan Aurora kinase TgArk3 is involved in endodyogeny, duplication rate and parasite virulence. *Cell. Microbiol.,* **2016**, *18* (8), 1106-1120. doi: 10.1111/cmi.12571.

29. Sullivan, W. J.; Narasimhan, J.; Bhatti, M. M. Parasite-specific eIF2 (eukaryotic initiation factor-2) kinase required for stress-induced translation control. *Biochem. J.*, **2004**, *380* (2), 523-531. doi: 10.1042/BJ20040262.

30. Gajria, B.; Bahl, A.; Brestelli, J.; Dommer, J.; Fischer, S.; Gao, X.; Heiges, M.; Iodice, J.; Kissinger, J. C.; Mackey, A. J. ToxoDB: an integrated *Toxoplasma gondii* database resource. *Nucleic Acids Res.*, **2008**, *36* (Suppl 1), D553-D556. doi: 10.1093/nar/gkm981.

31. Cohen, P. The regulation of protein function by multisite phosphorylation - a 25 year update. *Trends Biochem. Sci.*, **2000**, *25* (12), 596-601. doi: 10.1016/S0968-0004(00)01712-6.

32. Eglen, R. M.; Reisine, T. The current status of drug discovery against the human kinome. *Assay Drug Dev. Technol.*, **2009**, *7* (1), 22-43. doi: 10.1089/adt.2008.164.

33. Ojo, K. K.; Dangoudoubiyam, S.; Verma, S. K.; Scheele, S.; DeRocher, A. E.; Yeargan, M.; Choi, R.; Smith, T. R.; Rivas, K. L.; Hulverson, M. A. Selective inhibition of *Sarcocystis neurona* calcium-dependent protein kinase 1 for equine protozoal myeloencephalitis therapy. *Int. J. Parasitol.*, **2016**, *46* (13-14), 871-880. doi: 10.1016/j.ijpara.2016.08.003.

34. Dissous, C.; Grevelding, C. G. Piggy-backing the concept of cancer drugs for schistosomiasis treatment: a tangible perspective? *Trends Parasitol.*, **2011**, *27* (2), 59-66. doi: 10.1016/j.pt.2010.09.001.

35. Talevich, E.; Kannan, N.; Miranda-Saavedra, D. Computational analysis of apicomplexan kinomes. In *Protein Phosphorylation in Parasites: Novel Targets for Antiparasitic Intervention*, 5 ed.; Doerig, C., Späth, G., Wiese, M., Eds.; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, 2014; pp 1-36.

36. Blazejewski, T.; Nursimulu, N.; Pszenny, V.; Dangoudoubiyam, S.; Namasivayam, S.; Chiasson, M. A.; Chessman, K.; Tonkin, M.; Swapna, L. S.; Hung, S. S. *et al*., Systems-based analysis of the *Sarcocystis neurona* genome identifies pathways that contribute to a heteroxenous life cycle. *MBio.*, **2015**, *6* (1), e02445-14. doi: 10.1128/mBio.02445-14.

37. Dice, J. F.; Goldberg, A. L. Relationship between in vivo degradative rates and isoelectric points of proteins. *Proc. Natl. Acad. Sci. U. S. A.*, **1975**, *72* (10), 3893-3897. doi: 10.1073/pnas.72.10.3893.

38. Kato, K.; Sugi, T.; Takemae, H.; Takano, R.; Gong, H.; Ishiwa, A.; Horimoto, T.; Akashi, H. Characterization of a *Toxoplasma gondii* calcium calmodulin-dependent protein kinase homolog. *Parasit. Vectors*, **2016**, *9* (1), 405. doi: 10.1186/s13071-016-1676-1.

39. Khan, F.; Tang, J.; Qin, C.; Kim, K. Cyclin–dependent kinase TPK2 is a critical cell cycle regulator in *Toxoplasma gondii*. *Mol. Microbiol.*, **2002**, *45* (2), 321-332. doi: 10.1046/j.1365-2958.2002.03026.x.

40. Ward, P.; Equinet, L.; Packer, J.; Doerig, C. Protein kinases of the human malaria parasite *Plasmodium falciparum*: the kinome of a divergent eukaryote. *BMC Genomics*, **2004**, *5* (1), 79. doi: 10.1186/1471-2164-5-79.

41. Kissinger, J. C.; Gajria, B.; Li, L.; Paulsen, I. T.; Roos, D. S. ToxoDB: accessing the *Toxoplasma gondii* genome. *Nucleic Acids Res.*, **2003**, *31* (1), 234-236. doi: 10.1093/nar/gkg072.

42. Solyakov, L.; Halbert, J.; Alam, M. M.; Semblat, J. P.; Dorin-Semblat, D.; Reininger, L.; Bottrill, A. R.; Mistry, S.; Abdi, A.; Fennell, C. Global kinomic and phospho-proteomic analyses of the human malaria parasite *Plasmodium falciparum*. *Nat. Commun.*, **2011**, *2* 565. doi: 10.1038/ncomms1558.

43. Gurnett, A. M.; Liberator, P. A.; Dulski, P. M.; Salowe, S. P.; Donald, R. G.; Anderson, J. W.; Wiltsie, J.; Diaz, C. A.; Harris, G.; Chang, B. Purification and molecular characterization of cGMP-dependent protein kinase from apicomplexan parasites a novel chemotherapeutic target. *J. Biol. Chem.*, **2002**, *277* (18), 15913-15922. doi: 10.1074/jbc.M108393200.

44. Lourido, S.; Shuman, J.; Zhang, C.; Shokat, K. M.; Hui, R.; Sibley, L. D. Calcium-dependent protein kinase 1 is an essential regulator of exocytosis in *Toxoplasma*. *Nature*, **2010**, *465* (7296), 359-362. doi: 10.1038/nature09022.

45. Donald, R. G.; Zhong, T.; Meijer, L.; Liberator, P. A. Characterization of two *T. gondii* CK1 isoforms. *Mol. Biochem. Parasitol.*, **2005**, *141* (1), 15-27. doi: 10.1016/j.molbiopara.2005.01.011.

46. Dorin-Semblat, D.; marta-Gatsi, C.; Hamelin, R.; Armand, F.; Carvalho, T. G.; Moniatte, M.; Doerig, C. Malaria parasite-Infected erythrocytes secrete PfCK1, the *Plasmodium* homologue of the pleiotropic protein kinase casein kinase 1. *PLoS One*, **2015**, *10* (12), e0139591. doi: 10.1371/journal.pone.0139591.

47. Masch, A.; Kunick, C. Selective inhibitors of *Plasmodium falciparum* glycogen synthase-3 (PfGSK-3): new antimalarial agents? *Biochim. Biophys. Acta, Proteins Proteomics*, **2015**, *1854* (10), 1644-1649. doi: 10.1016/j.bbapap.2015.03.013.

48. Peixoto, L.; Chen, F.; Harb, O. S.; Davis, P. H.; Beiting, D. P.; Brownback, C. S.; Ouloguem, D.; Roos, D. S. Integrative genomic approaches highlight a family of parasite-specific kinases that regulate host responses. *Cell Host. Microbe*, **2010**, *8* (2), 208-218. doi: 10.1016/j.chom.2010.07.004.

49. Sibley, L. D.; Qiu, W.; Fentress, S.; Taylor, S. J.; Khan, A.; Hui, R. Forward genetics in *Toxoplasma gondii* reveals a family of rhoptry kinases that mediates pathogenesis. *Eukaryot. Cell*, **2009**, *8* (8), 1085-1093. doi: 10.1128/EC.00107-09.

50. Erdmann, M.; Scholz, A.; Melzer, I. M.; Schmetz, C.; Wiese, M. Interacting protein kinases involved in the regulation of flagellar length. *Mol. Biol. Cell*, **2006**, *17* (4), 2035-2045. doi: 10.1091/mbc.E05-10-0976.

51. Kõhler, S. Multi-membrane-bound structures of Apicomplexa: II. the ovoid mitochondrial cytoplasmic (OMC) complex of *Toxoplasma gondii* tachyzoites. *Parasitol. Res.*, **2006**, *98* (4), 355-369. doi: 10.1007/s00436-005-0066-y.

52. Dorin, D.; Semblat, J.; Poullet, P.; Alano, P.; Goldring, J. P.; Whittle, C.; Patterson, S.; Chakrabarti, D.; Doerig, C. PfPK7, an atypical MEK-related protein kinase, reflects the absence of classical three-component MAPK pathways in the human malaria parasite *Plasmodium falciparum*. *Mol. Microbiol.*, **2005**, *55* (1), 184-186. doi: 10.1111/j.1365-2958.2004.04393.x.

53. Artz, J. D.; Wernimont, A. K.; lali-Hassani, A.; Zhao, Y.; Amani, M.; Lin, Y. H.; Senisterra, G.; Wasney, G. A.; Fedorov, O.; King, O. The *Cryptosporidium parvum* kinome. *BMC Genomics*, **2011**, *12* 478. doi: 10.1186/1471-2164-12-478.

54. Pearce, L. R.; Komander, D.; Alessi, D. R. The nuts and bolts of AGC protein kinases. *Nat. Rev. Mol. Cell Biol.*, **2010**, *11* (1), 9-22. doi: 10.1038/nrm2822.

55. Morlon-Guyot, J.; Berry, L.; Chen, C.–T.; Gubbels, M.-J.; Lebrun, M.; Daher, W. The *Toxoplasma gondii* calcium-dependent protein kinase 7 is involved in early steps of parasite division and is crucial for parasite survival. *Cell. Microbiol.*, **2014**, *16* (1), 95-114. doi: 10.1111/cmi.12186.

56. Iwanaga, T.; Sugi, T.; Kobayashi, K.; Takemae, H.; Gong, H.; Ishiwa, A.; Murakoshi, F.; Recuenco, F. C.; Horimoto, T.; Akashi, H. Characterization of *Plasmodium falciparum* cdc2-related kinase and the effects of a CDK inhibitor on the parasites in erythrocytic schizogony. *Parasitol. Int.*, **2013**, *62* (5), 423-430. doi: 10.1016/j.parint.2013.05.003.

57. Fox, B. A.; Rommereim, L. M.; Guevara, R. B.; Falla, A.; Triana, M. A. H.; Sun, Y.; Bzik, D. J. The *Toxoplasma gondii* rhoptry kinome is essential for chronic infection. *MBio.*, **2016**, *7* (3), e00193-16. doi: 10.1128/mBio.00193-16.

58. Gelly, J. C.; Joseph, A. P.; Srinivasan, N.; de Brevern, A. G. iPBA: a tool for protein structure comparison using sequence alignment strategies. *Nucleic. Acids Res.*, **2011**, *39* (Suppl. 2), W18-W23. doi: 10.1093/nar/gkr333.

59. Volkamer, A.; Kuhn, D.; Rippmann, F.; Rarey, M. DoGSiteScorer: a web server for automatic binding site prediction, analysis and druggability assessment. *Bioinformatics*, **2012**, *28* (15), 2074-2075. doi: 10.1093/bioinformatics/bts310.

60. Lahti, J. L.; Tang, G. W.; Capriotti, E.; Liu, T.; Altman, R. B. Bioinformatics and variability in drug response: a protein structural perspective. *J. R. Soc. Interface*, **2012**, *9* (72), 1409-1437. doi: 10.1098/rsif.2011.0843.

61. Yamamoto, N.; Takeshita, K.; Shichijo, M.; Kokubo, T.; Sato, M.; Nakashima, K.; Ishimori, M.; Nagai, H.; Li, Y. F.; Yura, T. The orally available spleen tyrosine kinase inhibitor 2-[7-(3, 4-dimethoxyphenyl)-imidazo [1, 2-*c*] pyrimidin-5-ylamino] nicotinamide dihydrochloride (BAY 61-3606) blocks antigen-induced airway inflammation in rodents. *J. Pharm. Exp. Ther.*, **2003**, *306* (3), 1174-1181. doi: 10.1124/jpet.103.052316.

62. Wirger, A.; Perabo, F. G. E.; Burgemeister, S.; Haase, L.; Schmidt, D. H.; Doehn, C.; Mueller, S. C.; Jocham, D. Flavopiridol, an inhibitor of cyclin-dependent kinases, induces growth inhibition and apoptosis in bladder cancer cells in vitro and in vivo. *Anticancer Res.*, **2005**, *25* (6B), 4341-4347. PMID: 16309238.

63. Gray, N. S.; Wodicka, L.; Thunnissen, A. M. W.; Norman, T. C.; Kwon, S.; Espinoza, F. H.; Morgan, D. O.; Barnes, G.; LeClerc, S.; Meijer, L. Exploiting chemical libraries, structure, and genomics in the search for kinase inhibitors. *Science*, **1998**, *281* (5376), 533-538. doi: 10.1126/science.281.5376.533.

64. Piotrowska, H.; Kucinska, M.; Murias, M. Biological activity of piceatannol: leaving the shadow of resveratrol. *Mutat. Res.*, **2012**, *750* (1), 60-82. doi: 10.1016/j.mrrev.2011.11.001.

65. Zhang, J.; Yang, P. L.; Gray, N. S. Targeting cancer with small molecule kinase inhibitors. *Nat. Rev. Cancer,* **2009**, *9* (1), 28-39. doi: 10.1038/nrc2559.

66. Boyle, S. N.; Koleske, A. J. Dissecting kinase signaling pathways. *Drug Discov. Today,* **2007**, *12* (17-18), 717-724. doi: 10.1016/j.drudis.2007.07.019.

67. Lawrence, J. G.; Hendrix, R. W.; Casjens, S. Where are the pseudogenes in bacterial genomes? *Trends Microbiol.,* **2001**, *9* (11), 535-540. doi: 10.1016/S0966-842X(01)02198-9.

68. Templeton, T. J.; Iyer, L. M.; Anantharaman, V.; Enomoto, S.; Abrahante, J. E.; Subramanian, G. M.; Hoffman, S. L.; Abrahamsen, M. S.; Aravind, L. Comparative analysis of apicomplexa and genomic diversity in eukaryotes. *Genome Res.,* **2004**, *14* (9), 1686-1695. doi: 10.1101/gr.2615304.

69. Costanzo, V.; Avvedimento, E. V.; Gottesman, M. E.; Gautier, J.; Grieco, D. Protein kinase A is required for chromosomal DNA replication. *Curr. Biol.,* **1999**, *9* (16), 903-906. doi: 10.1016/S0960-9822(99)80395-9.

70. Smith, A.; Ward, M. P.; Garrett, S. Yeast PKA represses Msn2p/Msn4p–dependent gene expression to regulate growth, stress response and glycogen accumulation. *EMBO J.,* **1998**, *17* (13), 3556-3564. doi: 10.1093/emboj/17.13.3556.

71. Mayr, B.; Montminy, M. Transcriptional regulation by the phosphorylation-dependent factor CREB. *Nat. Rev. Mol. Cell Biol.,* **2001**, *2* (8), 599-609. doidoi: 10.1038/35085068.

72. Sugi, T.; Ma, Y. F.; Tomita, T.; Murakoshi, F.; Eaton, M. S.; Yakubu, R.; Han, B.; Tu, V.; Kato, K.; Kawazu, S. I. *Toxoplasma gondii* cyclic AMP-dependent protein kinase subunit 3 is involved in the switch from tachyzoite to bradyzoite development. *MBio.,* **2016**, *7* (3), e00755-16. doi: 10.1128/mBio.00755-16.

73. McCartney, R. R.; Garnar-Wortzel, L.; Chandrashekarappa, D. G.; Schmidt, M. C. Activation and inhibition of Snf1 kinase activity by phosphorylation within the activation loop. *Biochim. Biophys. Acta, Proteins Proteomics,* **2016**, *1864* (11), 1518-1528. doi: 10.1016/j.bbapap.2016.08.007.

74. Hardie, D. G.; Schaffer, B. E.; Brunet, A. AMPK: an energy-sensing pathway with multiple inputs and outputs. *Trends Cell Biol.,* **2016**, *26* (3), 190-201. doi: 10.1016/j.tcb.2015.10.013.

75. Gaji, R. Y.; Johnson, D. E.; Treeck, M.; Wang, M.; Hudmon, A.; Arrizabalaga, G. Phosphorylation of a myosin motor by TgCDPK3 facilitates rapid Initiation of motility during *Toxoplasma gondii* egress. *PLoS Pathog.,* **2015**, *11* (11), e1005268. doi: 10.1371/journal.ppat.1005268.

76. Long, S.; Wang, Q.; Sibley, L. D. Analysis of noncanonical calcium-dependent protein kinases in *Toxoplasma gondii* by targeted gene deletion using CRISPR/Cas9. *Infect. Immun.,* **2016**, *84* (5), 1262-1273. doi: 10.1128/IAI.01173-15.

77. Ojo, K. K.; Larson, E. T.; Keyloun, K. R.; Castaneda, L. J.; DeRocher, A. E.; Inampudi, K. K.; Kim, J. E.; Arakaki, T. L.; Murphy, R. C.; Zhang, L. *Toxoplasma gondii* calcium-dependent protein kinase 1 is a

target for selective kinase inhibitors. *Nat. Struct. Mol. Biol.*, **2010**, *17* (5), 602-607. doi: 10.1038/nsmb.1818.

78. Morrison, D. K. MAP Kinase Pathways. *Cold Spring Harb. Perspect. Biol.*, **2012**, *4* (11), a011254. doi: 10.1101/cshperspect.a011254.

79. Malumbres, M.; Barbacid, M. Cell cycle, CDKs and cancer: a changing paradigm. *Nat. Rev. Cancer*, **2009**, *9* (3), 153-166. doi: 10.1038/nrc2602.

80. Cao, L.; Wang, Z.; Wang, S.; Li, J.; Wang, X.; Wei, F.; Liu, Q. Deletion of mitogen-activated protein kinase 1 inhibits development and growth of *Toxoplasma gondii*. *Parasitol. Res.*, **2016**, *115* (2), 797-805. doi: 10.1007/s00436-015-4807-2.

81. Li, Z.-Y.; Wang, Z.-D.; Huang, S.-Y.; Zhu, X.-Q.; Liu, Q. *TgERK7* is involved in the intracellular proliferation of *Toxoplasma gondii*. *Parasitol. Res.*, **2016**, *115* (9), 3419-3424. doi: 10.1007/s00436-016-5103-5.

82. Srinivasan, N.; Krupa, A. A genomic perspective of protein kinases in *Plasmodium falciparum*. *Proteins: Struct. Funct. Bioinf.*, **2005**, *58* (1), 180-189. doi: 10.1002/prot.20278.

83. Behnke, M. S.; Fentress, S. J.; Mashayekhi, M.; Li, L. X.; Taylor, G. A.; Sibley, L. D. The polymorphic pseudokinase ROP5 controls virulence in *Toxoplasma gondii* by regulating the active kinase ROP18. *PLoS Pathog.*, **2012**, *8* (11), e1002992. doi: 10.1371/journal.ppat.1002992.

84. Jones, N. G.; Wang, Q.; Sibley, L. D. Secreted protein kinases regulate cyst burden during chronic toxoplasmosis. *Cell. Microbiol.*, **2016**. doi: 10.1111/cmi.12651.

85. Knoll, L. J. Functional analysis of the rhoptry kinome during chronic *Toxoplasma gondii* infection. *MBio.*, **2016**, *7* (3), e00842-16. doi: 10.1128/mBio.00842-16.

86. Knight, Z. A.; Lin, H.; Shokat, K. M. Targeting the cancer kinome through polypharmacology. *Nat. Rev. Cancer*, **2010**, *10* (2), 130-137. doi: 10.1038/nrc2787.

87. Martin, D. M.; Miranda-Saavedra, D.; Barton, G. J. Kinomer v. 1.0: a database of systematically classified eukaryotic protein kinases. *Nucleic Acids Res.*, **2009**, *37* (suppl 1), D244-D250. doi: 10.1093/nar/gkn834.

88. Eddy, S. R. Profile hidden Markov models. *Bioinformatics*, **1998**, *14* (9), 755-763. doi: 10.1093/bioinformatics/14.9.755.

89. Conesa, A.; Gotz, S.; Garcia-Gomez, J. M.; Terol, J.; Talon, M.; Robles, M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, **2005**, *21* (18), 3674-3676. doi: 10.1093/bioinformatics/bti610.

90. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M. R.; Appel, R. D.; Bairoch, A. Protein identification and analysis tools on the ExPASy server. In *The Proteomics Protocols Handbook*, 1 ed.; Walker, J. M., Ed.; Humana Press: Totowa, NJ, 2005; pp 571-607.

91. Bailey, T. L.; Bodén, M.; Buske, F. A.; Frith, M.; Grant, C. E.; Clementi, L.; Ren, J.; Li, W. W.; Noble, W. S. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.*, **2009**, *37* (Web Server Issue), W202-W208. doi: 10.1093/nar/gkp335.

92. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic. Acids Res.*, **2004**, *32* (5), 1792-1797. doi: 10.1093/nar/gkh340.

93. Waterhouse, A. M.; Procter, J. B.; Martin, D. M.; Clamp, M.; Barton, G. J. Jalview Version 2: a multiple sequence alignment editor and analysis workbench. *Bioinformatics.*, **2009**, *25* (9), 1189-1191. doi: 10.1093/bioinformatics/btp033.

94. Guindon, S.; Dufayard, J. F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.*, **2010**, *59* (3), 307-321. doi: 10.1093/sysbio/syq010.

95. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, **2014**, *30* (9), 1312-1313. doi: 10.1093/bioinformatics/btu033.

96. Ronquist, F.; Teslenko, M.; van der, M. P.; Ayres, D. L.; Darling, A.; Hohna, S.; Larget, B.; Liu, L.; Suchard, M. A.; Huelsenbeck, J. P. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.*, **2012**, *61* (3), 539-542. doi: 10.1093/sysbio/sys029.

97. Letunic, I.; Bork, P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.*, **2016**, *44* (W1), W242-W245. doi: 10.1093/nar/gkw290.

98. Webb, B.; Sali, A. Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinformatics*, **2016**, *54* 5.6.1-5.6.37. doi: 10.1002/cpbi.3.

99. Kelley, L. A.; Mezulis, S.; Yates, C. M.; Wass, M. N.; Sternberg, M. J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.*, **2015**, *10* (6), 845-858. doi: 10.1038/nprot.2015.053.

100. Yang, J.; Yan, R.; Roy, A.; Xu, D.; Poisson, J.; Zhang, Y. The I-TASSER Suite: protein structure and function prediction. *Nat. Methods*, **2015**, *12* (1), 7-8. doi: 10.1038/nmeth.3213.

101. Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.*, **1993**, *26* (2), 283-291. doi: 10.1107/S0021889892009944.

102. DeLano, W. L. The PyMOL molecular graphics system. *http://www. pymol. org* **2002**.

103. Forli, S.; Huey, R.; Pique, M. E.; Sanner, M. F.; Goodsell, D. S.; Olson, A. J. Computational protein-ligand docking and virtual drug screening with the AutoDock suite. *Nat. Protoc.*, **2016**, *11* (5), 905-919. doi: 10.1038/nprot.2016.051.