Sains Malaysiana 49(4)(2020): 755-764
http://dx.doi.org/10.17576/jsm-2020-4904-04

Molecular Characterisation of *Eimeria tenella* Porin, a Potential Anticoccidial Drug Target

(Pencirian Molekul Porin *Eimeria tenella*, Sasaran Dadah Antikoksidia yang Berpotensi)

XIN-WEI LEE, SU DATT LAM, MOHD FIRDAUS-RAIH & KIEW-LIAN WAN*

ABSTRACT

Eimeria tenella *is an apicomplexan parasite that causes the economically important disease coccidiosis in chickens. An estimated loss over $3 billion USD per annum has been reported. Control of coccidiosis relies on chemotherapy and vaccination, but drug resistance is common and live vaccines are relatively expensive. Therefore, there is an urgent need to develop new drugs to control* Eimeria *infections. Recent studies have shown that the pore forming structures of porin play important roles in many eukaryotic organisms. In this study, we generated and characterised a putative porin cDNA sequence from* E. tenella *that we have named Etporin. Sequence alignments showed that Etporin is 47 % similar to the putative porin sequence of* Toxoplasma gondii*, while a search against the Conserved Domain Database (CDD) shows that Etporin contains the Porin3 superfamily domain. Multiple sequence alignment with porin sequences from various eukaryotic organisms showed that the conserved VKXKX and GLK/STK motifs are present in Etporin. Analysis of the predicted Etporin 3D structure showed a classic beta barrel structure consisting of 19 beta-strands. Taken together, these results suggested Etporin has the potential to be developed into an anticoccidial drug target.*

Keywords: Coccidiosis; drug target; protein structure

ABSTRAK

Eimeria tenella *adalah parasit apikompleksa yang menyebabkan penyakit koksidiosis pada ayam. Anggaran kerugian ekonomi melebihi USD $3 bilion setahun telah dilaporkan. Pengawalan penyakit ini bergantung kepada kemoterapi dan pemvaksinan, namun kerintangan dadah adalah berleluasa dan vaksin hidup adalah mahal secara relatifnya. Oleh itu, terdapat keperluan yang mendesak untuk membangunkan dadah baru bagi mengawal jangkitan* Eimeria*. Kajian terkini menunjukkan bahawa struktur porin yang terlibat dalam pembentukan liang memainkan peranan penting dalam kebanyakan organisma eukariot. Dalam kajian ini, kami telah menjana dan mencirikan jujukan cDNA porin putatif daripada* E. tenella*, yang telah dinamakan Etporin. Penjajaran jujukan berbilang menunjukkan bahawa Etporin mempunyai 47% keserupaan dengan jujukan porin putatif* Toxoplasma gondii*, sementara pencarian terhadap Pangkalan Data Domain Terpelihara (CDD) menunjukkan bahawa Etporin mengandungi domain superfamili Porin3. Penjajaran jujukan berbilang dengan jujukan porin daripada pelbagai organisma eukariot turut menunjukkan bahawa motif terpelihara VKXKX dan GLK/STK hadir pada Etporin. Analisis struktur ramalan 3D Etporin menunjukkan struktur tong beta klasik yang terdiri daripada 19 bebenang-beta. Secara keseluruhannya, hasil kajian ini mencadangkan potensi Etporin untuk dibangunkan sebagai sasaran dadah antikoksidia.*

Kata kunci: Koksidiosis; sasaran dadah; struktur protein

INTRODUCTION

Eimeria species are intracellular protozoan parasites that cause coccidiosis in animals including chickens, cows, pigs, rabbits, and ducks. Seven different species of *Eimeria*, namely *E. acervulina*, *E. maxima*, *E. tenella*, *E. brunetti*, *E. necatrix*, *E. mitis*, and *E. praecox* have been recognised to infect chickens (Chapman 2014). Of these, *E. tenella* is the most studied and widely used as a standard reference in many laboratories because of its high replication rate, and it can be easily isolated, recovered and handled in the laboratory (Blake et al. 2015; Reid et al. 2014). Each *Eimeria* species has their own specificity where different species will invade different locations and possess different pathogenicity, immunological specificity, pre-patent period and sporulation time. The pathology of coccidiosis is caused by the destruction of epithelial cells and blood capillaries due to the release of merozoites from schizonts that in turn result in debris and blood clots causing blockages that lead to necrosis. In serious cases, necrosis can result in the deaths of infected chickens. It has been estimated that losses due to coccidiosis in the chicken industry are over $3 billion USD per annum (Dalloul & Lillehoj 2006; Williams 1999).

To reduce the occurrences of coccidiosis in chickens, prophylactic anticoccidial drugs and vaccines are being used with considerable success (Blake & Tomley 2014). Prophylactic chemotherapy which involves in-feed anticoccidial drugs has been the main method of controlling

chicken coccidiosis. Generally, anticoccidial drugs such as monensin, lasalocid, salinomycin, narasin and maduramycin are chemicals that act on parasite metabolism by altering ion transport and disrupting osmotic balance (Noack et al. 2019). Vaccination is an alternative way to control coccidiosis in chickens. Protective immunity is induced when chickens are primarily infected with low numbers of *Eimeria* parasites (Chapman & Jeffers 2014). However, the emergence of drug resistance and relatively high production costs of vaccines emphasise the need for the development of more efficient methods of coccidiosis control.

A number of proteins have been previously assessed for their potential as anticoccidial drug targets including glucose-6-phosphate isomerase (Loo et al. 2010) and glycogen synthase kinase-3 (Yao et al. 2016). Recent studies have shown that the porin protein may also be a good candidate as it plays an important role in cell adherence to eukaryotic host cells and pathogen/symbiont recognition (Goo et al. 2006; Hejair et al. 2017; Nyholm et al. 2009). In addition, porin has also been demonstrated as one of the most-heavily lysine acetylated proteins, indicating that lysine deacetylase inhibitors such as apicidin can be used to prevent the acetylation of porin and subsequently attenuate the host-pathogen interaction and the infection process (Jeffers & Sullivan Jr. 2012). Porin has been reported to be conserved across all known apicomplexan parasites and its unique properties as protein transport complexes could be exploited for possible drug targetting (Gajdács 2019; Mather et al. 2006; Pusnik et al. 2008). Here, we report the sequencing and characterisation of a putative porin cDNA sequence for *E. tenella*.

Materials and Methods

VECTORS AND CLONES

The recombinant clone used in this study was isolated from a cDNA library of *E. tenella* obtained from the Institute of Medical Sciences, University of Tokyo, Japan. The cDNA library was constructed by cloning *E. tenella* merozoites into pME18S-FL3 vector using an oligo-capping approach (Amiruddin et al. 2012).

BACTERIA GROWTH

Stock cells of *E. coli* carrying pME18S-FL3 vector were streaked onto Luria-Bertani (LB) agar and incubated overnight at 37 ^{0}C. A single colony was inoculated into 5 mL LB Broth medium containing 50 µg/mL ampicillin. The inoculums were grown 18 h at 37 ^{0}C with agitation. These cultures were then used for plasmid extraction.

PLASMID DNA EXTRACTION

The pME18S-FL3 plasmid DNA was extracted with a QIAprep® Miniprep kit (QIAGEN, USA), according to the manufacturer's protocol. The extracted plasmid was eluted out using 50 µL sterilised water.

PLASMID CHARACTERISATION

The quality and quantity of DNA samples obtained were determined by Spectrophotometer ND-1000 (NanoDrop, USA). Agarose gel electrophoresis was carried out to determine the presence of the plasmid. DNA supercoiled ladder (Promega, USA) was used as a DNA size marker. The extracted plasmids were subsequently digested with EcoR1 (Promega, USA) and the digested products were analysed using agarose gel electrophoresis to characterise the Etporin gene.

DNA CYCLE SEQUENCING

Sequencing of extracted DNA plasmids was performed using the ABI PRISM® BigDye® Terminator V3.1 Cycle Sequencing Ready Reaction Kit (Applied Biosystem Inc., USA) according to the manufacturer's protocol. The parameters for cycle sequencing were: heating at 96 ºC for 2 min, denaturation at 96 ºC for 10 s, annealing at 50 ºC for 5 s and elongation at 60 ºC for 4 min. Denaturation, annealing and elongation processes were repeated for 99 cycles. Reaction products were then ethanol precipitated and vacuum dried. Purified DNA samples were sequenced using an ABI PRISM® 3100 Genetic Analyzer (Applied Biosystem Inc., USA).

PRIMER DESIGN

Primers were designed using Primer3 (frodo.wi.mit.edu/primer3/). Oligo cap linker (5' GGA TGT TGC CTT TAC TTC TA 3') and Oligo-dT-adapter (5' TGT GGG AGG TTT TTT CTC TA 3') were used as outer primers of cDNA in cycle sequencing reactions. PwF1 (5' TTT CGA CTT CTC CAC AGA AT 3'), PwF2 (5' CAG CCT GAT GTT CGG GAA GG 3'), PwF3 (5' GAT CGA CGT CGT TTC TTA AC 3'), PwR1 (5' GGG CCT TGC ATG CGT TTT G 3'), PwR2 (5' TCG ATC TTG ATT CCG TAC TT 3') and PwR3 (5' ATT CAA GCT CCC TAT CAC CT 3') were designed as inner primers.

PRIMER DESIGN AND DATABASE SEARCHES

The sequences generated were quality-clipped using Phred (Machado et al. 2011). Vector and host contamination were identified and masked using the sequence comparison program Crossmatch (Gordon et al. 1998). Vector trimming excised the longest non-masked sequence and further trimming removed low quality bases (Phred score less than 30) at both ends of a read. Rearrangement and assembly of the sequences generated were done by Bioedit (Hall 1999). Open reading frames in the sequences obtained were identified by ORF Finder (www.ncbi.nlm.nih.gov/gorf/orfig.cgi) at the National Centre for Biotechnology Information (NCBI) website (Sayers et al. 2018).

Sequence database similarity searches were carried out using BLAST suites (Camacho et al. 2009) for similarity with known porin proteins. Conserved protein

domains were analysed using the Conserved Domain Database (CDD) (Marchler-Bauer et al. 2014), CATH protein domain database (Dawson et al. 2017) and InterPro domain database (Mitchell et al. 2018). Multiple sequence alignments were produced using MAFFT (Katoh & Standley 2013). Protein function prediction was done using ARGOT (Falda et al. 2012), PANDA (Wang et al. 2018) and PANTHER (Mi et al. 2019).

FOLD RECOGNITION AND 3D MODELLING

Secondary structures were predicted using PSIPRED (Buchan & Jones 2019). Different protein structure modelling servers i.e. I-TASSER (Yang et al. 2015), Phyre2 (Kelley et al. 2015), Raptor-X (Peng & Xu 2011) and IntFOLD5 (McGuffin et al. 2019) were utilised to generate 3D models of Etporin. MolProbity (Williams et al. 2018) were used to check the stereochemical quality of the protein structure. ProQ3 (Uziela et al. 2016), a machine learning based model quality assessment method in CASP12 (Kryshtafovych et al. 2018), was also used to determine the model quality.

RESULTS AND DISCUSSION

Eimeria tenella is an apicomplexan parasite that causes coccidiosis, an economically important disease that affects chicken production worldwide. This disease is dominantly controlled by prophylactic chemotherapy but the widespread occurrence of drug resistant strains requires the development of new drug targets. Identification and validation of a potential drug target are the beginning steps in the drug discovery pipeline. Before investing more resources in the target, it is crucial to have collective corroborative evidence that can support the choice of a target to be further explored. A good drug target must be involved in a crucial, ideally an essential biological pathway that has been functionally and structurally characterised as well as being distinguishable from any previously known targets (Bakheet & Doig 2009).

Studies have shown that porin, a transmembrane protein located in the outer membrane of mitochondria, is a major interface between the mitochondrial and the cellular metabolisms (Shoshan-Barmatz et al. 2006). It has been recognised as a key protein in mitochondria-mediated apoptosis, thus making it a potential target for the rational development of new therapeutics. In this study, a putative porin cDNA sequence in *E. tenella* was generated and characterised for further understanding of this protein.

A recombinant plasmid containing the cDNA of interest was sequenced by primer walking and resulted in a contig of 1368 bp in size (Figure 1). Analysis with ORF Finder showed the most probable open reading frame of the sequence is located in the +2 reading frame with a start codon at position 104 encoding methionine and ending with a stop codon at position 982 for a total length of 879 bp and encoding for 292 amino acids.

A BLASTx search showed the sequence to be similar to only other eukaryotic porin sequences (Table 1). We thus annotated this sequence as a putative porin and named it as Etporin. The database matches included highly significant matches to other apicomplexan sequences from *Toxoplasma gondii, Plasmodium vivax, P. falciparum, P. knowlesi*, and *Cryptosporidium muris*. The Etporin sequence showed the highest similarity with the putative porin protein from *T. gondii* (XP_002365430.1). The Etporin sequence length of 292 aa fitted well into the average sequence length for a single porin unit which averages ~290 aa and therefore indicative of a complete sequence.

The Etporin sequence search against CDD showed that Etporin contains the Porin 3 superfamily domain (E-value: 5.72e-29). The Porin 3 superfamily has two subfamilies - the Voltage Determinant Anion Channel (VDAC) subfamily and the Tom40 subfamily. Etporin is believed to be a VDAC subfamily member because it possesses a sequence with motifs consistent to other VDAC members. Etporin is also mapped to CATH mitochondria outer membrane porin functional family (2.40.160.160.FF9682) and eukaryotic porin family (IPR001925). These families composed of VDAC that behaves as diffusion pores for small hydrophilic molecules.

Multiple sequence alignment with porin sequences from various eukaryotic organisms identified the conservation of two glycine residues (G145 and G285) and two lysine residues (K94 and K283) (Figure 2). The conserved lysine residues may function by funnelling anions toward the binding site and may be required for the formation of the inorganic phosphate-specific binding site of porin (Sukhan & Hancock 1996).

However, a GLK motif found to be present in other porin sequences is absent in Etporin and in its place at the aligned position is an STK sequence (Figure 2). The GLK motif was originally suggested to be an ATP binding site because the replacement of LYS with GLU will lead to the impaired binding of ATP thus resulting in cation selectivity but not anion selectivity (Runke et al. 2000). In Etporin, the GLK motif is not present. Recent studies showed that porin sequences that lack this motif were scattered throughout the phylogenetic tree, and there is no clear relationship between this motif and eukaryotic porin signature motif. However, in plants, there are two clusters of porin sequences where GLK is replaced by a version of STK (Young et al. 2007). According to multiple sequence alignments of Etporin and porin sequences from various eukaryotic organisms, an STK motif is present in Etporin instead of the GLK motif that can be found aligned in other porin sequences. We speculate that evolutionary changes occurred in Etporin or it may be more closely related to porin of plants. Taken together, GLK or STK motif was either derived by a neutral event or it may have been selected due to a specific function.

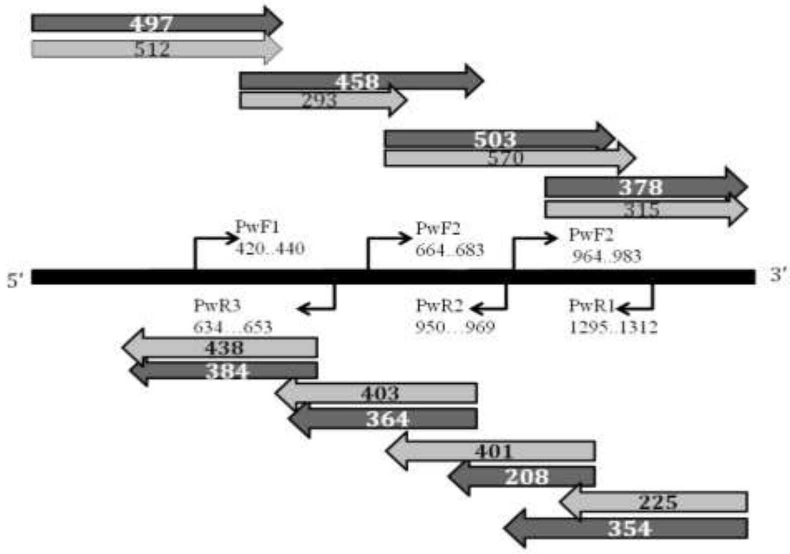
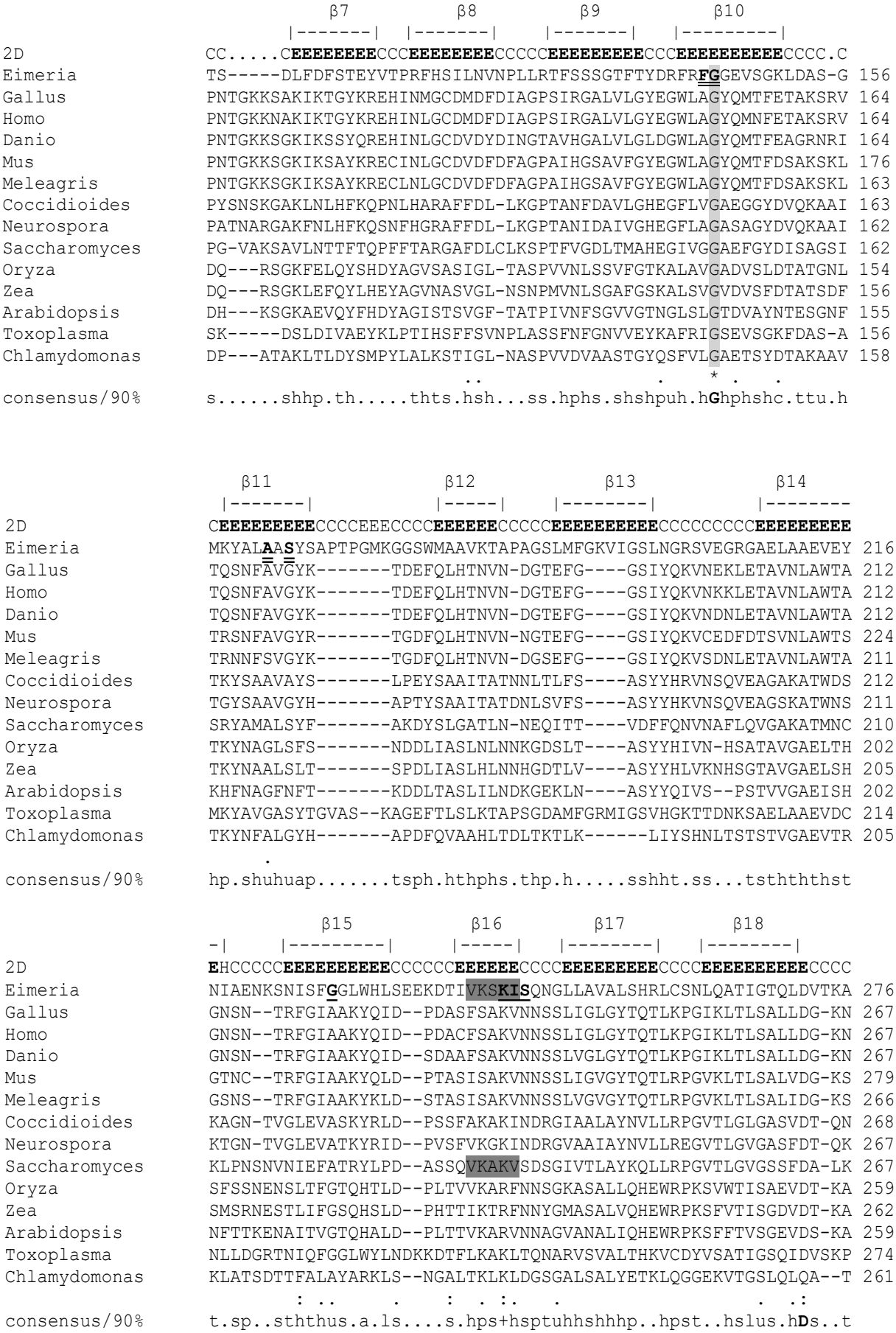


FIGURE 1. Generation of the full-length sequence of *Etporin* via the primer walking strategy

Dark bar indicates generated full-length sequence. Forward primers 5' 3' are represented by PwF1, PwF2 and PwF3. Reverse primers 3' 5' are represented by PwR1, PwR2 and PwR3. Number below primer name represents the location of each primer. Dark grey arrows indicate nucleotide sequences generated by clone 1 while light grey arrows indicate nucleotide sequences generated by clone 2

```
                                           α1                   β1          β2
                                     |------------|          |------|     |-----
2D                  ...............CCCHHHCCHHHHHHHCCCCCCCCE...EEEEEEECCCCCEEEEEE
Eimeria             ---------------MVLFKDIGKPASDLLGKGFPHEKP---WELEYKYKSKNPQITNVA 42
Gallus              ------------MAVPPAYADLGKSARDVFTKGYGFG-L---IKLDLKTKSENGLEFTSS 44
Homo                ------------MAVPPTYADLGKSARDVFTKGYGFG-L---IKLDLKTKSENGLEFTSS 44
Danio               ------------MAVPPTYVDLGKSARDIFTKGYGFG-L---IKLDLKTRSENGLEFKSS 44
Mus                 MAECCVPVCPRPMCIPPPYADLGKAARDIFNKGFGFG-L---VKLDVKTKSCSGVEFSTS 56
Meleagris           -------------AIPPSYADLGKSARDIFNKGYGFG-L---VKLDVKTKSASGVEFTTS 43
Coccidioides        ------------MAAPAAFGDIAKTVNDLLNKDFYHTSA---ASLEVKSKAPNGVTFNVK 45
Neurospora          ------------MAVP-AFSDIAKSANDLLNKDFYHLAA---GTIEVKSNTPNNVAFKVT 44
Saccharomyces       ------------MSPP-VYSDISRNINDLLNKDFYHATP---AAFDVQTTTANGIKFSLK 44
Oryza               ------------MVGPGLYPEIGKKARDLLYRDYQTD-----HKFTLTTYTSNGVAITAT 43
Zea                 -----------MVVAVGLYTDIGKKTRDLLYKDYNTH-----QKFCLTTSSPNGVAITAA 44
Arabidopsis         -----------MVKGPGLYTEIGKKARDLLYRDYQGD-----QKFSVTTYSSTGVAITTT 44
Toxoplasma          ---------------MVLFKDLNKSCADLLTKGYPHEKA---WDLEYKYKSKNPEIVNTA 42
Chlamydomonas       -------------MPVVSFGDIGKAAKGLLGGDKPTGTFSFDPKLSVSSTTASGVALTAT 47
                                      : :: :   .::  .           :     : .    .
consensus/90%       ..................a.-luKtspDlh.+sa.ht......phphp..o.ss..hpst

                               β3           β4           β5            β6
                    --|    |----------| |-------|   |-------|      |-------|
2D                  EECCCCCEEEEEEEEEEECCEEEEEEEECCCCEEEEEEEEECCCCCCEEEEEEEECCCCC
Eimeria             GVGSHGGFDASSSLVYSAGEAKAEVKVFSAGRSYVDLSYKIPQLPNLIISTKYERKGEKN 102
Gallus              GSANSETSKVSGSLETKYRWVEYGLMFTEKWNTDNTLGTEITLEDQLARGLKLTFDSSFS 104
Homo                GSANTETTKVTGSLETKYRWTEYGLTFTEKWNTDNTLGTEITVEDQLARGLKLTFDSSFS 104
Danio               GSANTETSKVAGTLETKYKWAEHGLTFTEKWNTDNTLGTEITLEDQLTKGLKLTFDSSFS 104
Mus                 GSSNTDTGKVSGTLETKYKWCEYGLTFTEKWNTDNTLGTEIAIEDQICQGLKLTFDTTFS 116
Meleagris           GSSNTDTGKVNGSLETKYKWAEYGLTFTEKWNTDNTLGTEIAIEDQIAKGLKLTFDTTFS 103
Coccidioides        GKS-AHEGPISGSLEAKYVDPPTGLTLTQTWTTGNALDTKLELDNNIAKGLKAEVLTQYL 104
Neurospora          GKS-THDKVTSGALEGKFTDKPNGLTVTQTWNTANALETKVEMADNLAKGLKAEGIFSFL 103
Saccharomyces       AKQPVKDGPLSTNVEAKLNDKQTGLGLTQGWSNTNNLKTKLEFAN-LTPGLKNELITSLT 103
Oryza               STK--KADLIFGEIQSQIKNKN--ITVDVKANSDSNVVTTVTVDEL-TPGLKSILSFAVP 98
Zea                 GTR--KNESIFGELHTQIKNKK--LTVDVKANSESDLLTTITVDEFGTPGLKSIINLVVP 100
Arabidopsis         GTN--KGSLFLGDVATQVKNNN--FTADVKVSTDSSLLTTLTFDEP-APGLKVIVQAKLP 99
Toxoplasma          SVSPAGAFDASSKLKYCVSDVTTEVKMMAMGKSTIDVKYAAPKLKGLTLGAKFDRRGDKT 102
Chlamydomonas       TVQ--KADKVEASLKAAYSTKKYSVDVALSPDNK---VTATASINDVAPGIKLTTSAVLP 102
                                 :          .       .                . *
consensus/90%       u.t..ts..h.uplthpht.....lth..thpst.sl.hth...p..s.GhK..h...h.
```


β7 β8 β9 β10
2D
Eimeria
Gallus
Homo
Danio
Mus
Meleagris
Coccidioides
Neurospora
Saccharomyces
Oryza
Zea
Arabidopsis
Toxoplasma
Chlamydomonas
consensus/90%
β11 β12 β13 β14
β15 β16 β17 β18

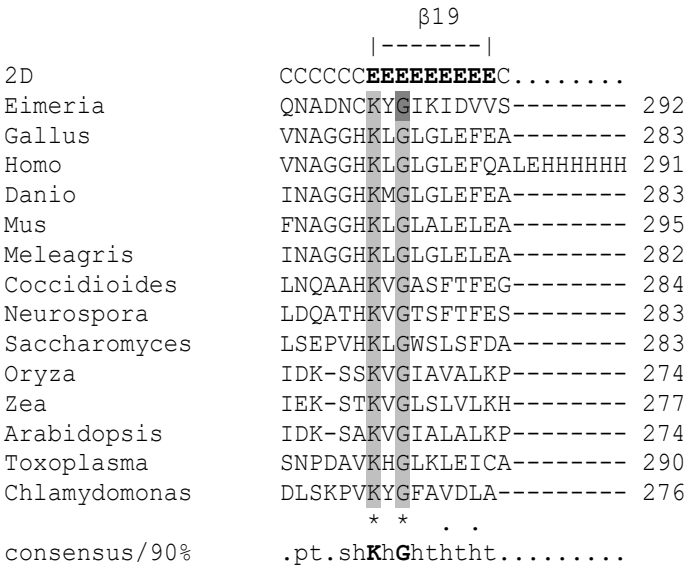


FIGURE 2. Multiple sequence alignment of Etporin and its homologues

Residues with single underline ( ) are predicted site of pocket 1 while residues with double underline ( ) are predicted site of pocket 2. Residues marked (*) are highly conserved residues; (:) are partially conserved residues; and (.) are less conserved residues. For 2D topology analysis, E are beta strands, C are coils and H are helices. Polar residues (KRHEDQNST) are shown as 'p'; turn-like residues (ACDEGKNQRST) are 't'; bulky hydrophobic residues (ACLIVMHYFW) are 'h'; the aliphatic subset of these type residues (LIVM) are 'l'; aromatic residues (FHWY) are 'a'; small residues (ACDGNPSTV) are 's'; and tiny residues (AGS) are shown as 'u'. Dark grey highlighted VKSKI is predicted by JS conservation analysis to be the conserved cluster which may involve in catalytic activity, oxidoreductase activity and oxidation reduction. This motif has been proven to be important in porin assemblye

TABLE 1. Distribution of BLAST hits on the Etporin sequence

Accession	Organism	Protein	Max score	Query coverage	E value	Identity
XP_002365430.1	*Toxoplasma gondii* ME49	Porin, putative	302	63%	8e-80	47%
XP_001616654.1	*Plasmodium vivax* SaI-1	Hypothetical protein	128	62%	2e-27	28%
XP_001348475.1	*Plasmodium falciparum* 3D7	Conserved protein	122	62%	1e-25	28%
XP_002140850.1	*Cryptosporidium muris* RN66	Hypothetical protein	103	64%	4e-20	26%
XP_002260703.1	*Plasmodium knowlesi* strain H	Hypothetical protein	103	58%	4e-20	26%
XP_676234.1	*Plasmodium berghei* strain ANKA	Hypothetical protein	98.2	55%	2e-18	27%
XP_746288.1	*Plasmodium chabaudi chabaudi*	Hypothetical protein	82.4	32%	1e-13	29%
XP_001611333.1	*Babesia bovis* T2Bo	Hypothetical protein	80.9	61%	4e-13	26%
XP_955278.1	*Theileria annulata* strain Ankara	Hypothetical protein	66.2	62%	9e-09	21%
XP_724308.1	*Plasmodium yoelii yoelii*	Hypothetical protein	63.5	28%	6e-08	26%

One signature motif of eukaryotic porins reported by Smith et al. (1995) is that a VKAKV sequence is available in all VDAC sequences that have been studied. Lys-234 and Lys-236 exist in this internal pentapeptide (VKAKV) of yeast porin, and are suggested to have an important role in the membrane insertion of this protein. Besides this, Lys-234 and Lys-236 in yeast have also been shown to be involved in porin assembly. Substitution of these sites by two other polar amino acid residues, glutamate and glutamine, will cause impaired assembly (Smith et al. 1995). In Etporin, we found that VKSKI is present at amino acids 240 to 245 instead of VKAKV, which has been detected in all porin sequences. However, Etporin still possesses the two conserved lysine residues (Lys-241 and Lys-244), which are important for porin assembly. This region VKSKI also has been predicted by Jenson-Shannon Divergence to be involved in catalytic activity, oxidoreductive activity and oxidation activity.

Secondary structure predictions by PSIPRED showed that Etporin consists of an alpha helix at the N-terminal and the expected series of 19 beta-strands that would be necessary for forming the beta barrel arrangement. This is therefore in agreement with existing porin structures. The only region of the sequence predicted to have an alpha-helical secondary structure (4^{th} to 15^{th} aa) is partially conserved between porin sequences from various eukaryotic organisms. This segment is usually located less than 10 residues from the amino-terminus (Young et al. 2007). Most of the beta-strands in Etporin are surrounded by two parallel partners. This is also the major arrangement of beta strands that have been characterised in other porin structures. Although the 16 beta-strands pore structure of porin are commonly seen, porins from different organisms may fold into barrels with different number of strands. The predicted secondary structures for Etporin strongly supports a common structural framework that is associated with porin where the main beta-barrel feature is well-conserved structure across all phyla investigated (Young et al. 2007).

Various structural modelling methods were employed to model the 3D structures of the Etporin subunit. All the modelling methods predicted the structures to be composed of 19 anti-parallel beta strands arranged to form a barrel shape. This arrangement is consistent with other known porins. The quality of models was assessed using MolProbity and ProQ3. The best structural model was generated using the IntFOLD5 server (Table 2). Ramachandran plot and result showed that 92.8% of total residues located at most favoured regions, 5.2% of them located at additional allowed regions, 1.6% at generously allowed regions and 0.4% residues is in disallowed regions. A good quality of 3D model would be expected to have over 90% residues in the most favoured regions. The model has a good ProQ3 score of 0.514 (the score ranges from 0 to 1, with 1 being the best score).

The best predicted Etporin structural model was compared with the human porin structure (PDB id: 2JK4, Figure 3). In spite of low sequence similarity (14.3%) shown between two sequences, their 3D structures are highly similar (TM-score 0.8533) where they have the same number of beta-strands forming a barrel with an N-terminal alpha-helical region located inside the pore. We speculate that extensive amino acid variations do not bring substantial alterations in the beta-barrel structure of membrane channels and its functions (Shoshan-Barmatz et al. 2010).

The function of Etporin was further investigated using a few protein function prediction web servers (ARGOT, PANDA, PANTHER). Most of the servers provide the most probable annotations for a query sequence in each of the three branches of the Gene Ontology i.e. biological process, molecular function, and cellular component. Its molecular function is described by ion channel activity and voltage-gated ion channel activity, it is thought to be involved in the biological processes associated with transmembrane transport and anion transport, and it can be found in mitochondria outer membrane. Taken together, Etporin has been shown as the porin protein in *E. tenella* and the presence of conserved motifs further suggests that it is an important protein for biological processes.

TABLE 2. Quality of Etporin models generated using various structure modelling methods

Methods	Template used	Percentage of residues located at most favoured region	ProQ3 global model quality score
I-TASSER	2JK4A, 3EMNX, 4BUMX	69.1%	0.448
Phyre2	2K4TA	81.5%	0.399
RaptorX	3EMNX	88%	0.498
IntFOLD5	3EMNX	92.8%	0.514

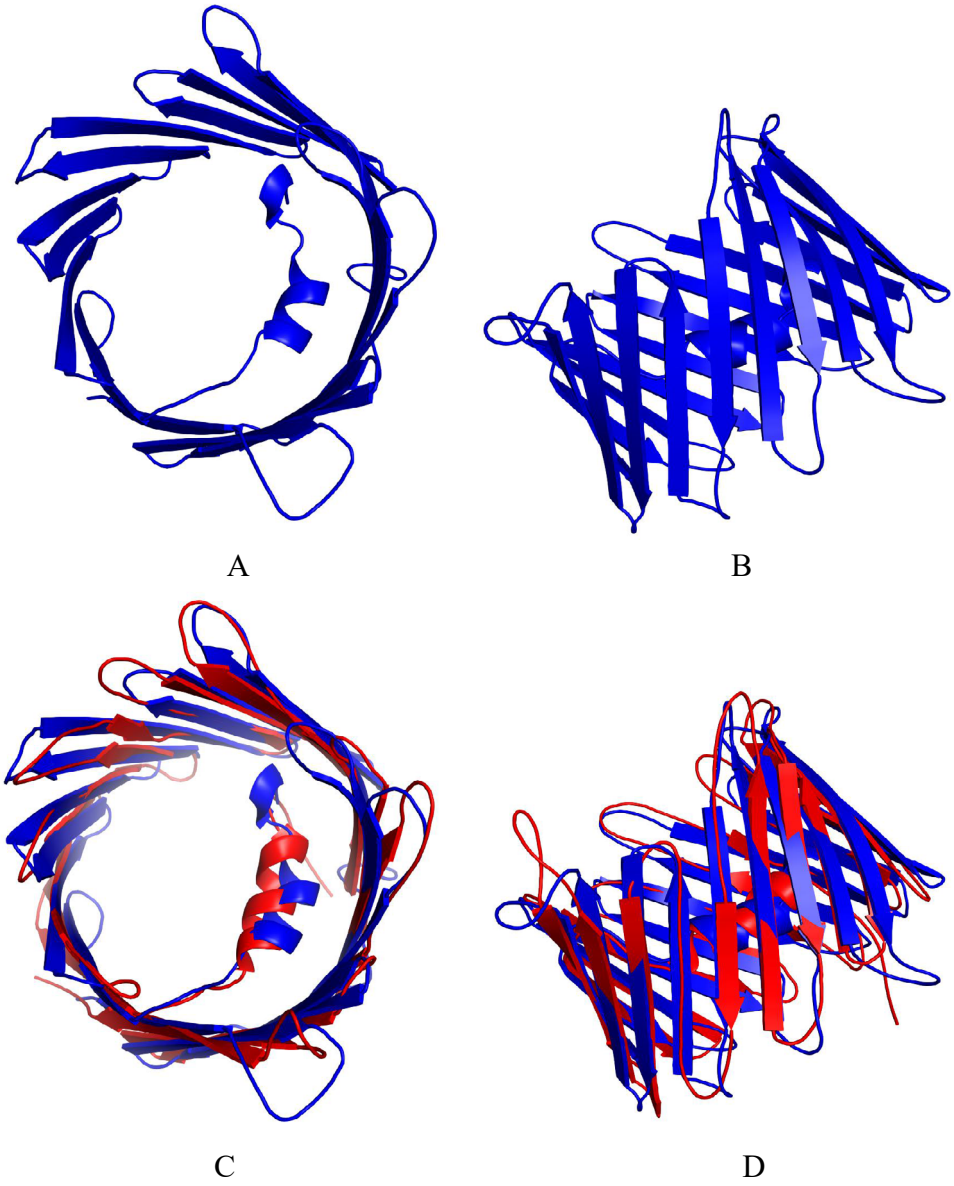


FIGURE 3. 3D structure models of Etporin (A & B) and superposition of Etporin (blue) and Homo sapiens porin (red) (C & D)

(A-C) Top view onto the barrels with the α-helix enclosed by and attached to the barrel wall. (B-D) Side view. Both porin structures are similar but sequence similarity only achieves 14.3%

CONCLUSION

In silico analyses on Etporin has shed light on the putative identity of the protein and showed novel essentials, uncovering its potential as a drug target in *E. tenella*. To increase the value of our findings, additional research, such as gene disruption studies, will be required to confirm the importance of Etporin and to verify the reliability of the data.

ACKNOWLEDGEMENTS

This work was financially supported by the Research University Grant from Universiti Kebangsaan Malaysia (DIP-2012-21).

REFERENCES

Amiruddin, N., Lee, X.W., Blake, D.P., Suzuki, Y., Tay, Y.L., Lim, L.S., Tomley, F.M., Watanabe, J., Sugimoto, C. & Wan, K.L. 2012. Characterisation of full-length cDNA sequences provides insights into the *Eimeria tenella* transcriptome. *BMC Genomics* 13: 21.

Bakheet, T.M. & Doig, A.J. 2009. Properties and identification of human protein drug targets. *Bioinformatics* 25(4): 451-457.

Blake, D.P. & Tomley, F.M. 2014. Securing poultry production from the ever-present Eimeria challenge. *Trends in Parasitology* 30(1): 12-19.

Blake, D.P., Clark, E.L., Macdonald, S.E., Thenmozhi, V., Kundu, K., Garg, R., Jatau, I.D., Ayoade, S., Kawahara, F. &

Moftah, A. 2015. Population, genetic, and antigenic diversity of the apicomplexan Eimeria tenella and their relevance to vaccine development. *Proceedings of the National Academy of Sciences U.S.A* 112(38): E5343-E5350.

Buchan, D.W.A. & Jones, D.T. 2019. The PSIPRED protein analysis workbench: 20 years on. *Nucleic Acids Research* 47(W1): W402-W407.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K. & Madden, T.L. 2009. BLAST+: Architecture and applications. *BMC Bioinformatics* 10(1): 421.

Chapman, H.D. 2014. Milestones in avian coccidiosis research: A review. *Poultry Science* 93(3): 501-511.

Chapman, H.D. & Jeffers, T.K. 2014. Vaccination of chickens against coccidiosis ameliorates drug resistance in commercial poultry production. *Internation Journal for Parasitology: Drugs and Drug Resistance* 4: 214-217.

Dalloul, R.A. & Lillehoj, H.S. 2006. Poultry coccidiosis: Recent advancements in control measures and vaccine development. *Expert Review of Vaccines* 5(1): 143-163.

Dawson, N.L., Sillitoe, I., Lees, J.G., Lam, S.D. & Orengo, C.A. 2017. CATH-Gene3D: Generation of the resource and its use in obtaining structural and functional annotations for protein sequences. *Methods in Molecular Biology* 1558: 79-110.

Falda, M., Toppo, S., Pescarolo, A., Lavezzo, E., Di Camillo, B., Facchinetti, A., Cilia, E., Velasco, R. & Fontana, P. 2012. Argot2: A large scale function prediction tool relying on semantic similarity of weighted Gene Ontology terms. *BMC Bioinformatics* 13(14): S14.

Gajdács, M. 2019. The concept of an ideal antibiotic: Implications for drug design. *Molecules* 24(5): 892.

Goo, S.Y., Lee, H.J., Kim, W.H., Han, K.L., Park, D.K., Lee, H.J., Kim, S.M., Kim, K.S., Lee, K.H. & Park, S.J. 2006. Identification of OmpU of *Vibrio vulnificus* as a fibronectin-binding protein and its role in bacterial pathogenesis. *Infection and Immunity* 74: 5586-5594.

Gordon, D., Abajian, C. & Green, P. 1998. Consed: A graphical tool for sequence finishing. *Genome Research* 8: 195-202.

Hall, T.A. 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 41: 95-98.

Hejair, H.M.A., Zhu, Y., Ma, J., Zhang, Y., Pan, Z., Zhang, W. & Yao, H. 2017. Functional role of ompF and ompC porins in pathogenesis of avian pathogenic Escherichia coli. *Microbial Pathogenesis* 107: 29-37.

Jeffers, V. & Sullivan Jr., W.J. 2012. Lysine acetylation is widespread on proteins of diverse function and localization in the protozoan parasite Toxoplasma gondii. *Eukaryotic Cell* 11(6): 735-742.

Katoh, K. & Standley, D.M. 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution* 30(4): 772-780.

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N. & Sternberg, M.J.E. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols* 10(6): 845-858.

Kryshtafovych, A., Monastyrskyy, B., Fidelis, K., Schwede, T. & Tramontano, A. 2018. Assessment of model accuracy estimations in CASP12. *Proteins: Structure, Function, and Bioinformatics* 86: 345-360.

Loo, S.S., Blake, D.P., Mohd-Adnan, A., Mohamed, R. & Wan, K.L. 2010. Eimeria tenella glucose-6-phosphate isomerase: Molecular characterization and assessment as a target for anti-coccidial control. *Parasitology* 137: 1169-1177.

Machado, M., Magalhães, W.C.S., Sene, A., Araújo, B., Faria-Campos, A.C., Chanock, S.J., Scott, L., Oliveira, G., Tarazona-Santos, E. & Rodrigues, M.R. 2011. Phred-Phrap package to analyses tools: A pipeline to facilitate population genetics re-sequencing studies. *Investigative Genetics* 2(1): 3. doi: 10.1186/2041-2223-2-3.

Marchler-Bauer, A., Derbyshire, M.K., Gonzales, N.R., Lu, S., Chitsaz, F., Geer, L.Y., Geer, R.C., He, J., Gwadz, M. & Hurwitz, D.I. 2014. CDD: NCBI's conserved domain database. *Nucleic Acids Research* 43(D1): D222-D226.

Mather, M., Henry, K. & Vaidya, A. 2006. Mitochondrial drug targets in apicomplexan parasites. *Current Drug Targets* 8: 49-60.

McGuffin, L.J., Adiyaman, R., Maghrabi, A.H.A., Shuid, A.N., Brackenridge, D.A., Nealon, J.O. & Philomina, L.S. 2019. IntFOLD: An integrated web resource for high performance protein structure and function prediction. *Nucleic Acids Research* 47(W1): W408-W413.

Mi, H., Muruganujan, A., Ebert, D., Huang, X. & Thomas, P.D. 2019. PANTHER version 14: More genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Research* 47(D1): D419-D426.

Mitchell, A.L., Attwood, T.K., Babbitt, P.C., Blum, M., Bork, P., Bridge, A., Brown, S.D., Chang, H.Y., El-Gebali, S. & Fraser, M.I. 2018. InterPro in 2019: Improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Research* 47(D1): D351-D360.

Noack, S., Chapman, H.D. & Selzer, P.M. 2019. Anticoccidial drugs of the livestock industry. *Parasitology Research* 118(7): 2009-2026.

Nyholm, S.V., Stewart, J.J., Ruby, E.G. & McFall-Ngai, M.J. 2009. Recognition between symbiotic Vibrio fischeri and the haemocytes of Euprymna scolopes. *Environmental Microbiology* 11: 483-493.

Peng, J. & Xu, J. 2011. RaptorX: Exploiting structure information for protein alignments by statistical inference. *Proteins* 79(S10): 161-171.

Pusnik, M., Charrière, F., Mäser, P., Waller, R.F., Dagley, M.J., Lithgow, T. & Schneider, A. 2008. The single mitochondrial porin of Trypanosoma brucei is the main metabolite transporter in the outer mitochondrial membrane. *Molecular Biology and Evolution* 26(3): 671-680.

Reid, A.J., Blake, D.P., Ansari, H.R., Billington, K., Browne, H.P., Bryant, J., Dunn, M., Hung, S.S., Kawahara, F., Miranda-Saavedra, D., Malas, T.B., Mourier, T., Naghra, H., Nair, M., Otto, T.D., Rawlings, N.D., Rivailler, P., Sanchez-Flores, A., Sanders, M., Subramaniam, C., Tay, Y.L., Woo, Y., Wu, X., Barrell, B., Dear, P.H., Doerig, C., Gruber, A., Ivens, A.C., Parkinson, J., Rajandream, M.A.,

Shirley, M.W., Wan, K.L., Berriman, M., Tomley, F.M. & Pain, A. 2014. Genomic analysis of the causative agents of coccidiosis in domestic chickens. *Genome Research* 24: 1676-1685.

Runke, G., Maier, E., O'Neil, J.D. & Benz, R. 2000. Functional characterization of the conserved 'GLK' motif in mitochondrial porin from *Neurospora crassa*. *Journal of Bioenergetics and Biomembranes* 32(6): 563-570.

Sayers, E.W., Agarwala, R., Bolton, E.E., Brister, J.R., Canese, K., Clark, K., Connor, R., Fiorini, N., Funk, K. & Hefferon, T. 2018. Database resources of the national center for biotechnology information. *Nucleic Acids Research* 47(D1): D23-D28.

Shoshan-Barmatz, V., Israelson, A., Brdiczka, D. & Sheu, S.S. 2006. The voltage-dependent anion channel (VDAC): Function in intracellular signalling, cell life and cell death. *Current Pharmaceutical Design* 12(18): 2249-2270.

Shoshan-Barmatz, V., De Pinto, V., Zweckstetter, M., Raviv, Z., Keinan, N. & Arbel, N. 2010. VDAC, a multi-functional mitochondrial protein regulating cell life and death. *Molecular Aspects of Medicine* 31(3): 227-285.

Smith, M.D., Petrak, M., Boucher, P.D., Barton, K.N., Carter, L., Reddy, G., Blachly-Dyson, E., Forte, M., Price, J. & Verner, K. 1995. Lysine residues at positions 234 and 236 in yeast porin are involved in its assembly into the mitochondrial outer membrane. *Journal of Biological Chemistry* 270(47): 28331-28336.

Sukhan, A. & Hancock, R.E.W. 1996. The role of specific lysine residues in the passage of anions through the Pseudomonas aeruginosa porin OprP. *Journal of Biological Chemistry* 271(35): 21239-21242.

Uziela, K., Shu, N., Wallner, B. & Elofsson, A. 2016. ProQ3: Improved model quality assessments using Rosetta energy terms. *Scientific Reports* 6: 33509.

Wang, Z., Zhao, C., Wang, Y., Sun, Z. & Wang, N. 2018. PANDA: Protein function prediction using domain architecture and affinity propagation. *Scientific Reports* 8(1): 34-84.

Williams, C.J., Headd, J.J., Moriarty, N.W., Prisant, M.G., Videau, L.L., Deis, L.N., Verma, V., Keedy, D.A., Hintze, B.J., Chen, V.B., Jain, S., Lewis, S.M., Arendall, W.B., Snoeyink, J., Adams, P.D., Lovell, S.C., Richardson, J.S. & Richardson, D.C. 2018. MolProbity: More and better reference data for improved all-atom structure validation. *Protein Science* 27(1): 293-315.

Williams, R. 1999. A compartmentalised model for the estimation of the cost of coccidiosis to the world's chicken production industry. *International Journal for Parasitology* 29: 1209-1229.

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J. & Zhang, Y. 2015. The I-TASSER suite: Protein structure and function prediction. *Nature Methods* 12(1): 7-8.

Yao, P.P., Firdaus-Raih, M., Sidek, H.M., Embi, N. & Wan, K.L. 2016. Molecular characterisation of glycogen synthase kinase-3 from *Eimeria tenella*. *Sains Malaysiana* 45(12): 1947-1957.

Young, M.J., Bay, D.C. & Hausner, G. 2007. The evolutionary history of mitochondrial porins. *BMC Evolutionary Biology* 7(1): 31.

Xin-Wei Lee, Su Datt Lam, Mohd Firdaus-Raih & Kiew-Lian Wan*
School of Biosciences and Biotechnology
Faculty of Science and Technology
Universiti Kebangsaan Malaysia
43600 UKM Bangi, Selangor Darul Ehsan
Malaysia

Su Datt Lam & Mohd Firdaus-Raih
Department of Applied Physics
Faculty of Science and Technology
Universiti Kebangsaan Malaysia
43600 UKM Bangi, Selangor Darul Ehsan
Malaysia

Kiew-Lian Wan*
Department of Biological Sciences and Biotechnolgy
Faculty of Science and Technology
Universiti Kebangsaan Malaysia
43600 UKM Bangi, Selangor Darul Ehsan
Malaysia

*Corresponding author; email: klwan@ukm.edu.my

Received: 25 September 2019
Accepted: 20 December 2019