Dissecting the subunit composition, structure and evolution of the highly diverged mitochondrial F-type ATP synthase complex from *Toxoplasma gondii*.

Thesis Submitted to AcSIR
For the Award of the Degree of
DOCTOR OF PHILOSOPHY
In
BIOLOGICAL SCIENCES

By

Rahul Pandurang Salunke
10BB13J26030

Under the guidance of
Dr. Dhanasekaran Shanmugam

BIOCHEMICAL SCIENCES DIVISION
CSIR-NATIONAL CHEMICAL LABORATORY,
Pune-411008, INDIA
January-2019

DECLARATION

I, Rahul Pandurang Salunke, hereby declare that the work incorporated in the thesis and entitled "**Dissecting the subunit composition, structure and evolution of the highly diverged mitochondrial F-type ATP synthase complex from *Toxoplasma gondii***" submitted for the award of the Degree of **Doctor of Philosophy** in **Biological Sciences** to the **Academy of Scientific & Innovative Research (AcSIR)**, New Delhi, has been carried out by me at Division of Biochemical Sciences, CSIR-National Chemical Laboratory, Pune-411008, India, under the supervision of Dr. Dhanasekaran Shanmugam. The work is original and has not been submitted as a part or full by me for any degree or diploma to this or any other university. I further declare that the material obtained from resources has been duly acknowledged in this thesis.

Date: 7th January 2019

Rahul Pandurang Salunke

(Enrollment No: 10BB13J26030)

सीएसआईआर - राष्ट्रीय रासायनिक प्रयोगशाला
(वैज्ञानिक तथा औद्योगिक अनुसंधान परिषद)
डॉ. होमी भाभा मार्ग, पुणे - 411 008, भारत
CSIR - NATIONAL CHEMICAL LABORATORY
(Council of Scientific & Industrial Research)
Dr. Homi Bhabha Road, Pune - 411 008, India

Certificate

This is to certify that the work incorporated in this Ph.D. thesis entitled "**Dissecting the subunit composition, structure and evolution of the highly diverged mitochondrial F-type ATP synthase complex from *Toxoplasma gondii***" submitted by **Mr. Rahul Pandurang Salunke** to the Academy of Scientific and Innovative Research (AcSIR), New Delhi, India, is in fulfilment of the requirements for the award of the Degree of **Doctor of Philosophy**, and embodies original research work under my supervision and guidance. I further certify that this work has not been submitted to any other University or Institution in part or full for the award of any degree or diploma. Research material obtained from other sources has been duly acknowledged in the thesis. Any text, illustration, table, used in the thesis from other sources, have been duly cited and acknowledged.

It is also certified that this work done by the student, under my supervision, is plagiarism free.

(Student)
Rahul Pandurang Salunke
Biochemical Sciences Division
CSIR-National Chemical Laboratory
Dr. Homi Bhabha Road,
Pune-411008, Maharashtra, India.

Dr. Dhanasekaran Shanmugam
Pr. Scientist
Biochemical Sciences Division,
CSIR-National Chemical Laboratory,
Dr. Homi Bhabha Road,
Pune-411008, Maharashtra, India.

Rahul_7.1.2019

ORIGINALITY REPORT

5%	2%	3%	0%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	An I. Jonckheere. "Mitochondrial ATP synthase: architecture, function and pathology", Journal of Inherited Metabolic Disease, 08/27/2011 Publication	<1%
2	dune.une.edu Internet Source	<1%
3	www.jimmunol.org Internet Source	<1%
4	J. P. Issartel. "The ATP synthase (F0â F1) complex in oxidative phosphorylation", Experientia, 04/1992 Publication	<1%
5	Vannuruswamy, Garikapati, Mashanipalya G. Jagadeeshaprasad, K. Kashinath, Suresh K. Kesavan, Shweta Bhat, Arvind M. Korwar, Ashok D. Chougale, Ramanamurthy Boppana, D. Srinivasa Reddy, and Mahesh J. Kulkarni. "Molecules with O-acetyl group protect protein glycation by acetylating lysine residues", RSC Advances, 2016.	<1%

Dedicated to my family, Ms. Swati and Dr. Anish for their love and never ending support

Acknowledgement

First of all, I would like to express my big appreciation to my mentor Dr. Dhanasekaran Shanmugam for his continuous guidance, support and the fruitful scientific stimulation throughout my PhD. I have learnt a great deal from you.

I take this opportunity to sincerely thank my present and former Doctoral Advisory Committee (DAC) members Dr. H.V. Thulasiram, Dr. Kiran Kulkarni, Dr. B.L.V. Prasad, Dr. C.G. Suresh, Dr. Asmita Prabhune and Dr. R.V. Gadre for evaluation of my thesis work, continuous monitoring my progress and their inputs.

I would like to thank Dr. Jomon Joseph from NCCS for his valuable time in evaluating my progress, constructive inputs, sharing reagents and facilitating my extension. Also I would like to thank Dr. Kisan Kodam and Dr. Ram Rup Sarkar for evaluating my progress and facilitating my extensions.

I would like to extend my gratitude for the extended support I have received from Proteomic group, Structural biology group, Neurobiology group and Plant Molecular Biology group. Specifically Dr. Arvind Korwar, Dr. Yogesh Kolekar, Dr.Vishal Dawkar, Dr. Yashwant Kumar, Dr. Kedar Batkulwar, Dr. Jagadeesha M.G, Dr. Arati Deshmukh, Dr. Ijaj, Dr. Amey, Tejas Chirmade, Gouri Patil, Shakutala Bai, Rajeshwari Rathore, Shweta Sonawane, Tushar, Abhishek, Nalinee, Anandsukeerthi, Sneha, Ashwini, Zenia, Gajanan and many more for your scientific input in my initial learning phase of various technique, providing the reagents in urgency and your moral support in troubleshooting of technical problem I have faced during my PhD.

I would take this opportunity to sincerely thank Dr. Mahesh Kulkarni, Dr. Subashchandrabose Chinnathambi, Dr. Narendra Kadoo, Dr. Ashok Giri Dr. Vidya Gupta, Dr. M.V. Deshpande and Dr. Mahesh Dharne for allowing me to use the equipments available in their group.

I would like to extent my gratitude to our collaborator Prof. Arnab Pain, Dr. Tobias Mourier (BESE Division, KAUST) and Dr.Manidipa Banerjee (IIT-Delhi) for their scientific discussion.

I would like to express my deepest appreciation to my labmates Dr. Anurag, Dr. Rupali, Meenakshi, Parag, Tejashri, Sindhuri, Shweta, Dr. Khooshboo, Sayyad Rohil, Milan Mathew for help, support, discussion and suggestion. I heartily thank for all the moments we share, it would be difficult for me to reach my goal without your support. I am also thankful to Dr. Geetha for her encouragement and well wishes, Nithilan for his loving company, Neha and Gayatri for their help and well wishes. Thank you for making the journey easier, happier and memorable.

I would like to thanks my friends Hriushikesh, Swapnil, Hanumant, Mangesh, Kiran, Rajnjeet, Uday, Pradeep, Bharat for constant support in bad or good time, sharing the best memories outside the lab and everlasting friendship.

I would like to thank all my family members for constant support and unconditional love.

My sincerest thanks Head of Department, Biochemical Sciences Division, for their support and lab facilities. I extend my gratitude and thanks to Director CSIR-National Chemical Laboratory for allowing me to carry out my research in this esteemed institute and providing the infrastructure facilities. I sincerely thank SAC office, CSIR-NCL for facilitating all the documentation and Academy of Scientific and Innovative Research (AcSIR), New Delhi for allowing me to submit my thesis for the award of Ph.D. degree. My sincerest thanks to Council of Scientific & Industrial Research (CSIR), New Delhi, Govt. of India for the Ph.D. fellowship.

Last but not least I would like to thanks all friends who help and encourage me to complete the journey.

Thank You.....

Rahul P Salunke

List of figures

ADP	Adenosine diphosphate
ASAPs	ATP synthase associated proteins
ATP	Adenosine triphosphate
BCKDH	Branched chain ketoacid dehydrogenase
BNP	Blue Native PAGE
Cas9	CRISPR-associated 9
CAT	Chloramphenicol Acetyl Transferase
Cd	Cohen's d
cDNA	Complementary DNA
ConF	Confirmatory forward
CRISPR	Clustered regularly interspaced short palindromic repeat
DDM	β-dodecyl maltoside
DEAE	Diethylaminoethanol
DHODH	Dihydroorotate dehydrogenase
DMEM	Dulbecco's Modified Eagle Medium
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
F-ATPase	F-type ATP synthase complex
FBS	Fetal bovine serum
g	Gravitational constant
gDNA	Genomic DNA
gm	Gram
GOI	Gene of interest
Gt1	Glucose transporter 1
HFF	Human foreskin fibroblasts
IP	Immunoprecipitation
KDa	Kilodalton
KEGG	Kyoto Encyclopedia of Genes and Genomes
kV	Kilovolt
LC-MS/MS	Liquid Chromatography tandem Mass Spectrometry
M	Molar
m/z	Mass/charge

mA	Milliampere
MEM	Minimum Essential Medium
mETC	Mitochondrial electron transport chain
$MgCl_2$	Magnesium chloride
$MgSO_4$	Magnesium sulfate
mM	Millimolar
MPI	Max Planck Institute
MQO	Malate:quinone oxidoreductase
MSB	Mitochondria Solubilization Buffer
MWU	Mann-Whitney U *p*-values
NaCl	Sodium chloride
$NaHCO_3$	Sodium bicarbonate
NEB	New England BioLab
ng	Nanogram
nl	Nano liter
OD	Optical density
OSCP	Oligomycin sensitivity conferring protein.
PAGE	Polyacrylamide gel electrophoresis
$Pb(NO_3)_2$	Lead nitrate
PBS	Phosphate-buffered saline
PDH	Pyruvate dehydrogenase
PEPCK	Phosphoenolpyruvate carboxykinase
Pi	Inorganic phosphate
PLGS	Protein Lynx Global Server
Pmf	Proton motive force
PVDF	Polyvinylidene difluoride
RH-wt	Type-I RH [RH-wild type]
RHΔKu80	Type-I RH ΔKu80
SDS	Sodium dodecyl sulfate
SEC	Size exclusion chromatography
TCA	Tricarboxylic acid cycle
TFA	Trifluoroacetic acid

Tg-DHFR	*T. gondii* dihydrofolate reductase-thymidilate synthase
Tg-Δgt1	*T. gondii* glucose transporter 1
USA	United States of America
UTR	Untranslated region
V	Volt
w/v	Weight by volume
YFP-HA	Yellow fluorescent protein plus hemagglutinin
α-HA	Anti- hemagglutinin antibody
µg	Microgram
µl	Microliter

Salunke R, Mourier T, Banerjee M, Pain A, Shanmugam D. Highly diverged novel subunit composition of apicomplexan F-type ATP synthase identified from *Toxoplasma gondii*. PLoS Biol. 2018 Jul 13;16(7):e2006128.

The mitochondrial F-type ATP synthase (F-ATPase), a multisubunit nanomotor, is critical for maintaining cellular ATP level. In *Toxoplasma gondii* and other apicomplexan parasites many subunits required for proper assembly and functioning of this enzymes are appears to be missing in these parasites. Despite missing crucial subunits, the functionality of F-ATPase is established in apicomplexan parasites, indicating that novel proteins are probably associated with the complex to make it functional. Here, we characterize the subunit composition of mitochondrial F-ATPase enzymes in apicomplexan parasite *T. gondii*. We have partially purified and enriched the F-ATPase enzymes complex from detergent solubilized mitochondrial lysate using three different approaches - Blue Native PAGE, immunoprecipitation and by combination of ion exchange and gel filtration chromatography. Using mass spectrometry based proteomic analysis on partially purified ATP synthase complex, we have identified 20 novel subunit components of *T. gondii* F-ATPase. These studies also confirmed that the enzyme complex is present in monomeric (~600 kDa) and dimeric (>1 MDa) forms, as previously reported in other eukaryotes. Despite the extreme sequence diversification, conserved structure analysis helped in identification of missing key F_O subunits a, b and d, which are necessary for assembly and function of the complex. Notably, orthologs of these novel subunits of *T. gondii* F-ATPase are present in all apicomplexan parasites, except in *Cryptosporidium* species lacking a true mitochondrion. Interestingly, these proteins are also conserved in other alveolate species, such as the chromerids and dinoflagellates, suggesting an ancient origin for these proteins. Further the structural and functional characterization of these highly divergent novel components of F-ATPase will facilitate the fundamental understanding of energy metabolism and development of novel antiparasitic agents against apicomplexan parasites.

Keyword: F-type ATP synthase, Apicomplexa, *T. gondii*, mass spectrometry.

Introduction

Introduction:

The *Alveolata* infrakingdom is a monophyletic group of single celled eukaryotes which inhabit ecologically diverse niches. *Alveolata* is a sister group of stramenopiles which contain diatom algae, kelps and parasitic oomycetes. All alveolates have typical conserved structure, like presence of tubular mitochondria and a single membrane cortical alveolae present below the plasma membrane [1, 2]. This infrakingdom consists five diverse phyletic groups called *Ciliates*, *Dinoflagellates*, *Colpodellidae*, *Chromerida* and *Apicomplexa*. These organisms are adapted to extremely diverse mode of nutrition, from phototrophy (fixing carbon and producing energy *via* photosynthesis) to heterotrophic (either predator or parasite lifestyle for obtaining carbon and energy).

Ciliates are heterotrophic protista that consist of *Tetrahymena* and *Paramecium*, the important members of the microbial food webs. The majority of ciliates are free-living predators. Dinoflagellates are another important subgroup within the alveolate infrakingdom, which are photosynthetic as well as heterotrophic and present in all types of ecosystem [1]. Colpodella are free-living flagellated predators which mainly feed on other smaller free-living protist by penetrating their cell membrane and consuming the prey's cytoplasm [3]. Colpodella are closely related to parasitic apicomplexan and are of phylogenetic importance with respect to the evolution of apicomplexan species. Chromerida are photosynthetic algae and are the closest free-living relatives of apicomplexan parasites [4].

Apicomplexa is the only alveolate phylum in which all known species are exclusively parasitic in nature. These parasites are known to infect a variety of animal hosts, including humans, and thus impose significant medical and economic burden on human and animal welfare across the world. This phylum includes many thousands of protozoan parasites, including a few notorious ones like the *Plasmodium sp*., which are causative agents of malaria,

Eimeria sp., which are important poultry parasites, *Theileria sp*., which cause theileriosis in cattle, *Cryptosporidium sp*., which cause cryptosporidiosis, and the sister parasites *Neospora canium* and *Toxoplasma gondii* which are responsible for neosporosis and toxoplasmosis in humans and animals. The majority of apicomplexan parasites show small host range to complete their life cycle, while the promiscuous parasite *T. gondii* can infect all nucleated cells in all warm blooded animals, and has significant impact on human and animal health [5]. *T. gondii* has been used as the model organism of choice for carrying out this thesis related research work.

The apicomplexan parasite *Toxoplasma gondii*

T. gondii is an obligate intracellular protozoan parasite, one of the most widespread parasite in the world from apicomplexan phylum [6]. It is estimated that approximately one-third of human population is infected by *T. gondii* worldwide [6]. The majority of *T. gondii* infection is chronic and asymptomatic, and recrudescence and fatal pathogenesis can occur in severe disease condition like in immunocompromised individual, or in case of congenitally infected children [6]. To complete its life cycle, *T. gondii* has to pass through a feline host (Figure-1). Despite having the broad mammalian host range, only feline species act as definitive hosts and support the sexual differentiation and development of *T. gondii* [7]. Sexual differentiation happens when the bradyzoite form of the parasite enters the gut environment of an immunologically naïve feline host (i.e., previously uninfected by *T. gondii*), and converts into gametes. The union of male and female gametes happen in the gut epithelial cells, resulting in the formation of diploid oocyst. The unsporulated oocysts are further shed into the environment along with feces, and have the ability to stay for extended time periods in the environment. These unsporulated oocysts then undergo meiosis in the environment to produced infectious sporozoites contained within the sporulated oocysts. Food and water contaminated with sporulated oocysts are major source for *T. gondii* infection in intermediate hosts like

humans and other animals, where the parasites undergo asexual proliferation and differentiation [8]. In addition, direct infection from tachyzoites and bradyzoite tissue cyst stages can occur when infected meat containing the parasite is consumed susceptible host. In humans this can happen when they consume undercooked meat containing tissue cysts.

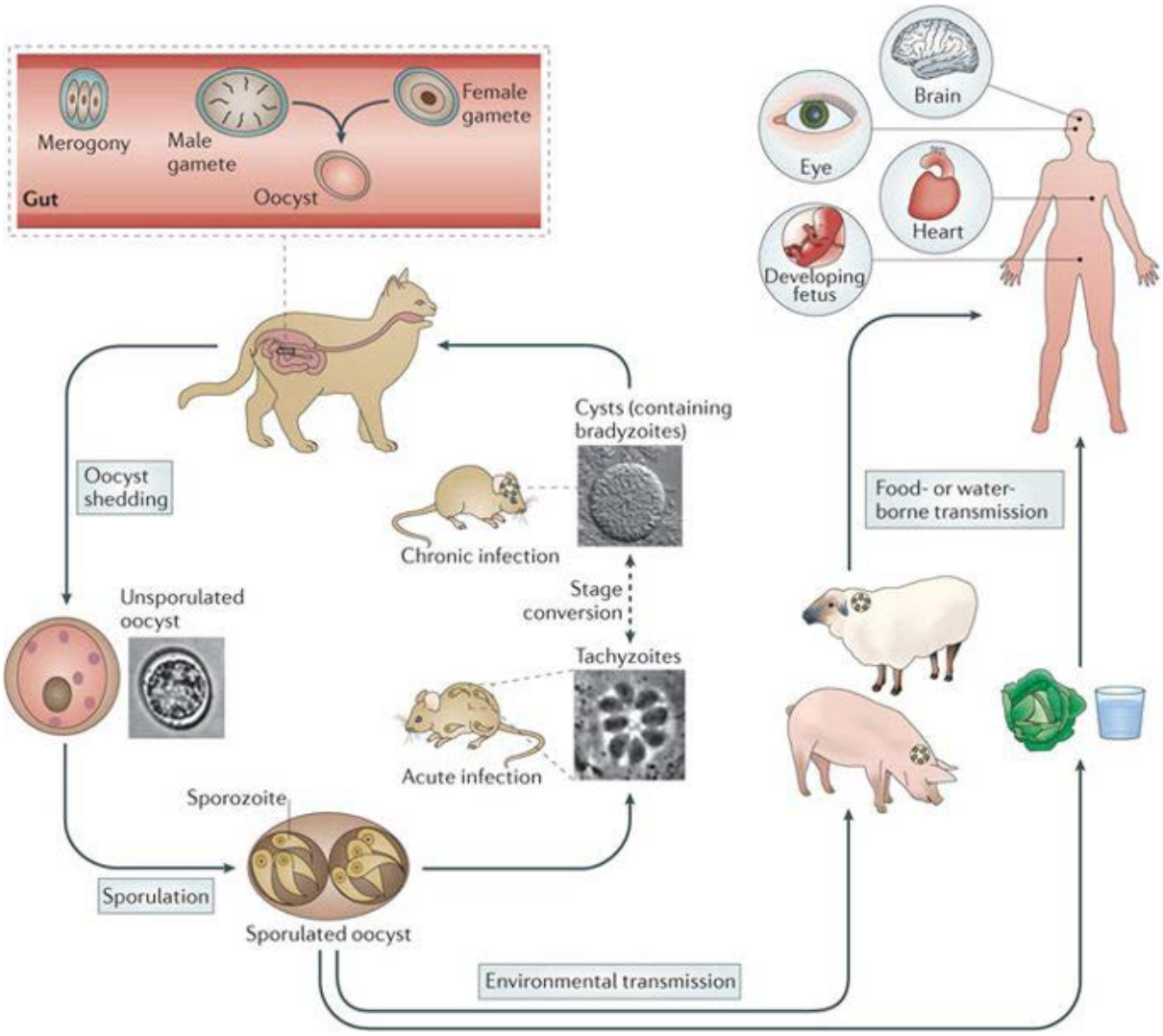


Figure-1: Schematic representation of life cycle of *T. gondii.* The life cycle of *T. gondii* occurs between feline and other intermediate host. The oocysts are shed by definitive host (feline species), which sporulate in the environment and lead to acute infection in the intermediate host before the immune system force the parasite to convert into dormant bradyzoite form. Bradyzoites mainly present in muscle and neuronal cell types. The life cycle is completed when felines feed on infected animal meat containing tissue cysts. Adapted from [8].

In the intermediate hosts, *T. gondii* parasites are present in two different forms called tachyzoite and bradyzoite stages. Tachyzoite stage is a highly replicative stage and is the virulent form of the parasite, which is responsible for acute infection and disease severity [6]. On the other hand bradyzoites are latent and slow growing forms [9], mainly present in muscle and neuronal cells. They have a protective cyst wall and are responsible for chronic infection. The bradyzoite form is reactivated into tachyzoite form in immunocompromised individuals resulting in severe disease condition [6, 7]. Being a highly replicative and virulent form, tachyzoites exhibit robust carbon and energy metabolism, which is probably distinct from the metabolic activity seen in bradyzoites. Being an obligate intracellular parasite, *T. gondii* is dependent on scavenged nutrients from the host for metabolic sustenance.

Carbon acquisition, metabolism and energy production in *T. gondii*

T. gondii and all apicomplexan parasites present inside a protective vacuole inside the host cell called as the parasitophorous vacuole, which serves as a selective barrier and protects the parasite from lysosomal degradation and exposure to host immune system. It also act as molecular sieve, allowing the diffusion of small molecules from host cell to parasite like amino acids, sugars and nucleotides [10]. The hexose sugar glucose is a very important and vital source of carbon and energy for the highly replicative and virulent tachyzoite stage of the parasite. Glucose scavenged by *T. gondii* is catabolized sequentially *via* the glycolytic pathway into pyruvate, which yields 2 moles of ATP per mole of glucose oxidized. Glycolysis supports bulk production of ATP in *T. gondii* and other apicomplexan parasites [11], and the pyruvate produced is mainly converted into lactic acid in the cytosol. Alternatively, pyruvate is also converted into acetyl CoA in the mitochondria, but unlike in most eukaryotes where this conversion is mediated by pyruvate dehydrogenase, in apicomplexan parasites it is mediated by the branched chain ketoacid dehydrogenase (BCKDH) enzyme [12]. The canonical pyruvate dehydrogenase (PDH) enzyme, which converts pyruvate into acetyl CoA, is absent in the

mitochondrion of apicomplexan parasites, and instead is present in the apicoplast organelle [13-15]. Glycolysis is important for the lytic cycle of *T. gondii* [16], and reallocation of the glycolytic enzymes to the periphery of the cell happens when the parasites are extracellular, presumably to localize ATP production [17]. Despite the importance of glycolysis as a quick source of ATP, glucose is not an essential nutrient for *T. gondii*. For example the glucose transporter 1 (*Tg*-Δgt1) which facilitates glucose transport into the parasite is dispensable [18], similarly the hexokinase enzyme, which catalyzes the first step in glycolysis is also dispensable [19]. Another glycolytic enzyme, aldolase, is also dispensable in *T. gondii* [20]. While in case of *Plasmodium sp*., the causative agent of malaria, the hexose sugar transporter is essential for parasites survival [21]. It was also observed that in the absence of glycolysis, *T. gondii* tachyzoites are able to maintain their cellular ATP homeostasis, likely *via* mitochondrial oxidative phosphorylation. The reducing equivalents produced in the parasite mitochondrion from the oxidation of pyruvate and glutamate *via* the tricarboxylic acid cycle (TCA) cycle are used to power the mitochondrial respiration, which is in turn coupled to ATP synthesis by complex V, also called the F-type ATP synthase.

Mitochondrial metabolism in apicomplexan parasites

Mitochondria are primary endosymbiotic organelles derived from α-proteobacteria and are present in almost all eukaryotes [22]. They are called the power house of the cell and are the hub for various metabolic processes in eukaryotic cells, such as ATP synthesis, various oxidative metabolic functions, biogenesis of iron sulfur cluster, amino acid metabolism, and heme biosynthesis, coenzyme A biosynthesis etc... They are primarily involved in catabolism of amino acids, carbohydrates, fatty acids and intracellular ion regulation [23]. The TCA cycle is the conserved pathway present in the mitochondrion *via* which the metabolites derived from catabolism of amino acid, carbohydrates and lipids are oxidized resulting in the formation of $NADH,H^+$ and $FADH_2$ as a reducing equivalents. It also provide anabolic precursor for

biosynthesis of various molecules [24-25]. Comparative genomics studies of various apicomplexan parasites, including *T. gondii*, and other alveolate organisms reveals the presence of a complete set of TCA cycle enzyme across the alveolate infrakingdom, except in case of *Cryptosporidium sp*., which lack a true mitochondrion [26-27].

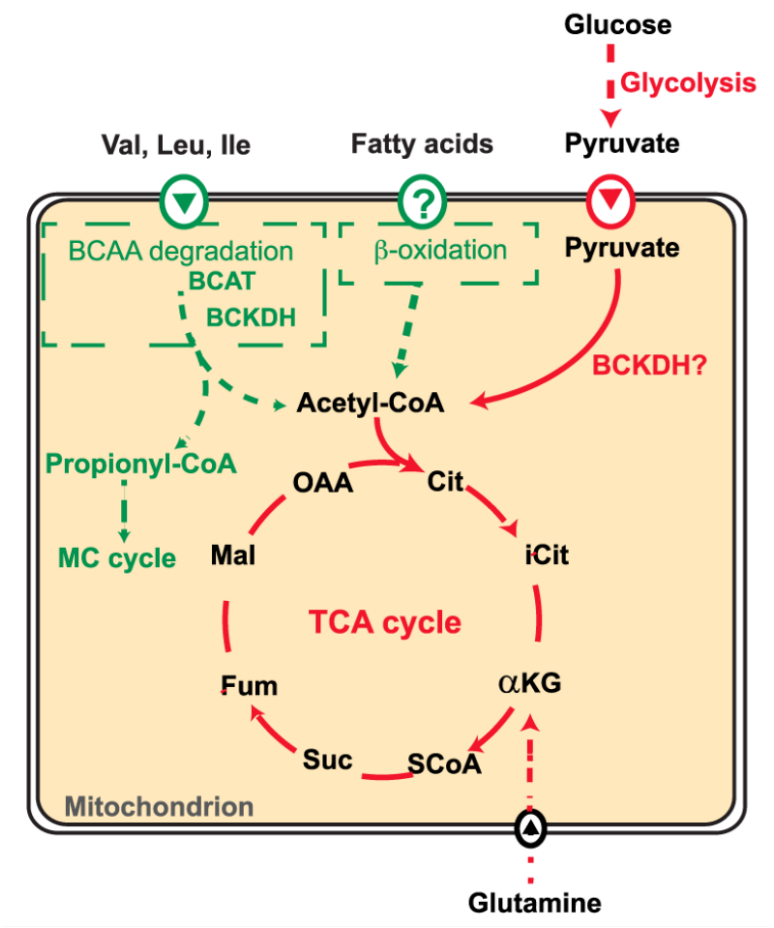


Figure-2: Schematic representation of pathways to produce acetyl-CoA in the mitochondrion. In green are pathways specific to *T. gondii* and in red pathways common to *T. gondii* and *Plasmodium* spp. Adapted from [12]

Despite the absence of a canonical PDH in the mitochondria of apicomplexan parasites, the fully functional TCA cycle has been observed in *T. gondii* and *Plasmodium sp* [28, 29]. The mitochondrial BCKDH, a multi-subunits enzyme complex, was shown to convert pyruvate into acetyl CoA in *T. gondii* and *Plasmodium* species (Figure -2) [12]. In the absence of glucose as carbon source in *T. gondii*, glutaminolysis act as carbon and ATP source [28]. The glycolytic

intermediate required for parasite in the absence of glucose is produced by gluconeogenesis. The gluconeogenesis enzymes phosphoenolpyruvate carboxykinase (PEPCK) present in the mitochondria, converts oxaloacetate, a TCA cycle intermediate, into phosphoenolpyruvate, a glycolytic intermediate [30]. PEPCK is not essential for *T. gondii* in presence of glucose, while it is essential only in glucose-deprived condition where mitochondrial glutaminolysis is the only source of carbon and energy [19, 31]. Surprisingly despite the operation of TCA cycle in asexual blood stage *P. falciparum*, many TCA cycle enzymes are dispensable for normal growth of the parasite. However, the parasites with TCA cycle knockout could not progress into sexual development within mosquitos, indicating the important of TCA cycle for life cycle completion [29, 32]. In summary, the functionality of TCA cycle is important for survival of apicomplexan parasites. During the operation of TCA cycle, reduced cofactors such as NADH, H^+ and $FADH_2$ are produced, which are oxidized in mitochondria by the component of electron transport chain and contribute to the maintenance of the proton gradient across the inner mitochondrial membrane. This proton gradient is then used for production of ATP by complex V (F-type ATP synthase) by a process called as oxidative phosphorylation.

Mitochondrial electron transport chain (mETC) in apicomplexan parasites

The reduced cofactors NADH, H^+ and $FADH_2$ produced by TCA cycle in *T. gondii*, are oxidized by the mitochondrial electron transport chain (mtETC), which results in the sequential transfer of electron through a series of multi-subunit protein complexes to molecular oxygen. Typically, the mtETC is comprised of four different complexes – I, II-III & IV, each of which sequentially accepts and transfers the electrons. Complex I or complex II split hydrogen donated by reduced cofactors into H^+ and electron. The H^+ is pumped across the inner membrane and into the inter-membrane space to establish the proton gradient, while the electron is passed on to ubiquinone, a lipophilic molecule. The reduced ubiquinone donates the electron to complex III or cytochrome c reductase. The cytochrome c protein facilitates the

transfer of electron from complex III to complex IV (cytochrome c oxidase), subsequently leading to reduction of molecular oxygen to H_2O. Complex IV also pumps protons across the intermembrane space, thus contributing to establishing the proton motive force (pmf) across the inner membrane space and associated transmembrane potential. The pmf and associated transmembrane potential is used by F-type ATP synthase complex to produced ATP from ADP and inorganic phosphate (Pi) [33, 34].

Among apicomplexan parasites, mitochondrial metabolism, especially electron transport chain and associated functions are well studied in *Plasmodium* sp. and *T. gondii* [35-37]. Comparative genomics studies have revealed that apicomplexan parasites possess most components of a canonical mtETC, except complex I (Figure-3) [26-27, 38-39]. The canonical complex I is a multimeric transmembrane protein complex, which transfers electrons from reduced $NADH_2$ to ubiquinone while pumping protons across the inner membrane, into the inner membrane space. Complex I is sensitive to rotenone, an inhibitor of mtETC. Instead of the canonical multimeric complex I, apicomplexan parasites possess a single protein, non-

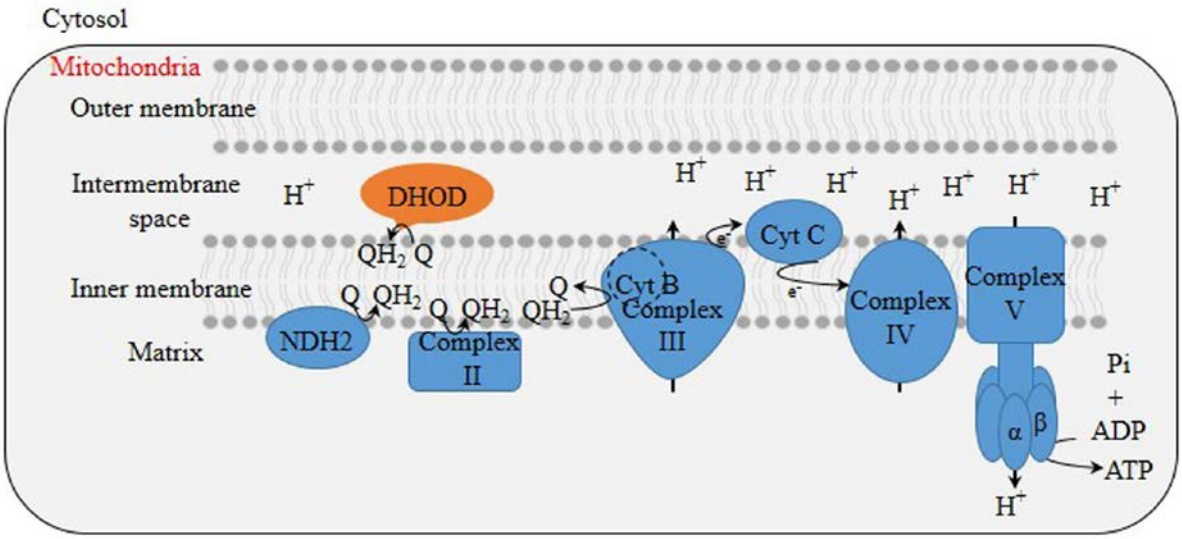


Figure-3: Schematic representation of the electron transport chain in the apicomplexan parasites.

proton pumping complex I called as the type-II NADH dehydrogenase, which is also present in plants and fungi.

The type II NADH dehydrogenase lacks the transmembrane domain, is present as a peripheral membrane protein on the matrix side of the mitochondrial inner membrane and is insensitive to rotenone [40-43]. The presence of a type II NADH dehydrogenase is not unique to apicomplexan parasites; in fact the entire alveolate clade, except for ciliates, which are basal to all other alveolates, possess the type II NADH dehydrogenase enzyme [26]. In *T. gondii* and many other coccidian parasites such as *Neospora* and *Eimeria*, two or more genes encoding this enzyme are present, while other apicomplexan parasites possess only a single copy of this enzyme [39, 44]. It is reported that 1-hydroxy-2-dodecyl-4-(1H) quinolone (HDQ), a quinolone like compound identified as an inhibitor of the yeast type-II NADH dehydrogenase enzyme, also inhibits the enzyme activity in *P. falciparum* and *T. gondii* [44]. Also the type II NADH dehydrogenase enzyme is dispensable in asexual stage of *P. berghei* while the sexual development is affected [45]. The two isoforms of this enzyme are each individually dispensable in *T. gondii* and the KO parasites exhibit only a modest fitness defect. Interestingly, complementing the individual KOs with the other isoform could not reverse the fitness defect and restore normal growth, indicating that they may be functionally non-redundant [46].

Complex II of the mtETC is the succinate dehydrogenase enzyme, which is also associated with the TCA cycle. Complex II transfer electron from $FADH_2$ to ubiquinone and the apicomplexan enzyme is orthologous to other eukaryotic enzymes. In addition to Complex I and II enzymes, other dehydrogenases such as the dihydroorotate dehydrogenase (DHODH), Malate:quinone oxidoreductase (MQO) and glycerol-3-phosphate dehydrogenase are also capable of directly reducing oxidized ubiquinone.

Complex III, also called the cytochrome bc_1 complex, is an essential component of

mtETC, which is responsible for oxidizing the reduced ubiquinone by transferring the electron from reduced ubiquinone to cytochrome c protein *via* the Q cycle [47-48]. In apicomplexan parasites, complex III activity is the first major mechanism by which the proton gradient and associated membrane potential across the mitochondrial inner membrane is established. Complex III contains three main catalytic domains called the Rieske iron sulfur protein, cytochrome c1, and cytochrome b. The mitochondrial genome encoded cytochrome b subunit is a target for the antimalarial drug atovaquone [49-50]. Other compounds such as pyridones and endochin like quinolones [51-52] also inhibit complex III in apicomplexan parasites.

Structural studies on *P. falciparum* cytochrome b indicate that atovaquone binds to the Qo site of the protein, which can block the regeneration of oxidized ubiquinone necessary for accepting the electron from reduced substrates [53]. However cytochrome b mutations are known to confer resistance to atovaquone in blood stage *P. falciparum*. Interestingly, these mutations appear to compromise cytochrome b function resulting in arrested sexual development of the parasite in the mosquito [54], while blood stage parasites appear to have normal growth. This agrees with the fact that glycolysis is the main source of ATP for blood stage parasites, while mitochondrial oxidative phosphorylation is the main source of ATP in the mosquito where glucose can be limiting.

Next the component of the mtETC is the Complex IV or cytochrome c oxidase, which is again a multi-subunit protein complex involved in the oxidation of cytochrome c and generation of proton gradient and associated membrane potential across the inner membrane space. It is also responsible for reduction of molecular oxygen into water molecules. The final component of the mtETC is the complex V or the F-type ATP synthase complex, which utilizes the transmembrane potential and proton gradient generated by the mtETC to produce ATP from ADP and Pi, which will be discussed in detail in the following section.

Complex V

The complex V or F_1F_O-ATP synthase complex which is also called as F-type ATP synthase (F-ATPase) enzyme is a ubiquitous nanomotor, found in the inner membrane of mitochondrion, thylakoid membrane of chloroplast and the bacterial plasma membrane [55-60]. The F-ATPase fulfills the important role of catering to bulk synthesis of adenosine triphosphate (ATP) [55-60], which is the primary source of cellular energy for all living organisms. This enzyme is capable of hydrolysis of ATP into ATD and Pi, and so its activity needs to be tightly regulated. The F-ATPase is similar in structure and function to the A-type ATPase (archaeal ATPase) responsible for ATP synthesis in archaea and the eukaryotic V-type ATPase (vacuolar ATPase) present on the membranes of vacuoles which degrade ATP and utilize the energy to generate ion gradient across the vacuolar membrane [57, 61].

The energy required to drive the F-ATPase nanomotor is harnessed from the proton motive force and associated membrane potential [62], which are generated by the component of respiratory complexes present in inner membrane of mitochondria, chloroplast and the bacterial plasma membrane [33-34]. The structure and function of the multisubunit F-ATPase is well characterized from bacteria, yeast, bovine and many more organisms using genetic, biochemical, crystallographic and cryo-electron microscopic techniques. As a result of these studies, it appears that the general organization, architecture and catalytic mechanisms of the F-ATPase is conserved across bacteria, unicellular eukaryotes and complex multicellular organism.

Architecture of F-type ATP synthase

Irrespective of the source of the F-ATPase enzyme complex, its structure and function are highly conserved. The simplest form of F-ATPase is present in bacteria and chloroplast membranes, which contain 8-9 subunits, while the mitochondrial F-ATPase has 16-18 subunits

[56-60]. In addition to the orthologs for the subunits present in the bacterial and chloroplast F-ATPase, the mitochondrial F-ATPase also contain other subunits required for assembly and function of the enzyme complex (Figure-4).

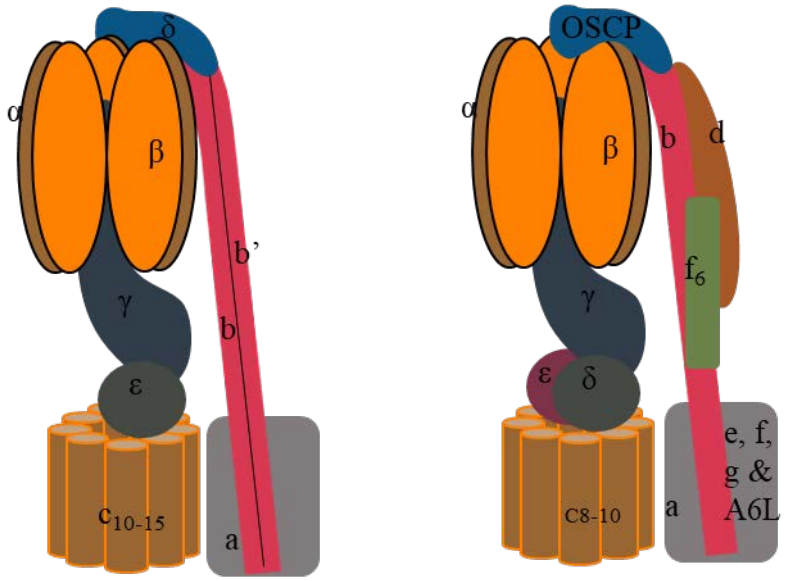


Figure 4: **Schematic representation of F-type ATP synthase enzyme** subunit composition from bacteria, chloroplast (left) and mitochondria (right). Adapted and modified from (57).

The F-ATPase comprises of two distinct domains called the F_1 and F_O sectors. The F_1 sector is a globular, membrane extrinsic, hydrophilic domain contain catalytic core involved in ATP formation or hydrolysis, and the central stalk which connects the catalytic core to the F_O sector. The F_O sector contains the membrane spanning hydrophobic domain through which proton translocation happens and the peripheral stalk which connects the F_O sector to F_1 sector.

The catalytic core of the F_1 sector comprises of three copies each of α and β subunits, which are arranged alternatively to form a globular heterodimer. In the catalytic core the formation or hydrolysis of ATP takes place at the interface of non-catalytic α and catalytic β subunit [55-59]. The central stalk is composed of single copies of γ, δ and ε subunits in the mitochondrial F-ATPase, but only γ and ε subunits in case of bacterial and chloroplast F-ATPase. The elongated α helical structure of γ subunit protrude from the catalytic core toward

the F_O sector, at the base of which are present the δ and ε subunits. In case of the mitochondrial F-ATPase the δ and ε subunits are firmly attached to the membrane bound oligomeric subunit c. The F_O sector contains hydrophobic membrane spanning oligomeric subunit c having the stoichiometry of c_{10-12} depending on the species, and a single copy each of subunits a, b, d, f, 8, h, and OSCP as is the case for yeast and bovine mitochondrial F-ATPase. Similarly the F_O sector subunits in bacterial and chloroplast F-ATPase are composed of oligomeric subunit c_{10-15} along with single copies of subunit a, b, b' and δ (Figure-4). The mitochondrial OSCP subunit is a counterpart of the δ subunit present in bacterial and chloroplast F-ATPase. The F_O sector subunit c and a form the proton channel through which proton translocation happens [63-65], while rest of the F_O sector subunits belong to the peripheral stalk [58-59,66-67]. The oligomeric subunit c along with the central stalk act as the rotor component of the enzyme. The stator structure of the enzyme is made up of the α and β catalytic core, the F_O subunit a, and the peripheral stalk subunit components [56-59, 66].

Mechanism of ATP synthase and hydrolysis

Detailed structural and biochemical studies indicate that the F-ATPase is mechanistically divided into two parts - the rotor and stator. The rotor made up of the oligomeric subunit c from F_O sector and subunit γ, δ and ε from the central stalk, while the stator part includes the $\alpha_3\beta_3$ catalytic core of the F_1 sector and subunit a, b, d, F6 and OSCP of F_O sector in mitochondrial F-ATPase at least in yeast and bovine [59]. During ATP synthesis, the rotary motion of the nanomotor, which is powered by the electrochemical gradient, is used to generate a torque force with respect to the stationary catalytic core, which by way of conformational changes facilitates the synthesis of ATP in the catalytic site. The energy required to generate the rotary motion and the associated torque force is produced by translocation of proton in the F_O sector at the interface of subunit c and a [63-65]. The rotary motion induces conformational change in the asymmetric γ subunit of the central stalk. The

conformational change in γ subunit facilitates the confirmation change in the β subunit of catalytic core of the enzyme, which ultimately results in ATP synthesis [55-59]. The torque force is counterbalanced by the peripheral stalk of the stator, which has the crucial role of holding the catalytic core in position relative to the rotating central stalk [66-67]. The mechanism of ATP synthesis or hydrolysis is well understood and is referred to as the binding change mechanism [55, 61 , 68].

As illustrated by the binding change mechanism, the catalytic β subunit sequentially undergoes three different conformational changes resulting in the binding of ADP and Pi and the formation of ATP molecules. At any given instance, one of the β subunits is present in the β-ATP confirmation (with tightly bound ATP; also called the "T" state), another β subunit is present in the β-ADP + Pi confirmation (loose binding of substrates; also called "L" state), and the third β subunit is present in the β-empty or open confirmation (very loose binding; also called the "O" state). Proton translocation induces rotation of γ subunit as indicated by the green arrowhead (Figure-5), which interacts with each of the three β subunit sequentially, resulting in subsequent conformational change in each of the β subunit. For example, in the "O" state, the ADP and Pi substrates are able to enter the catalytic site and bind very loosely. As this β subunit changes conformation, due to its contact with the rotating γ subunit, the catalytic site goes from the "O" state to "L" state. In the "L" state, there is very close association between ADP and Pi, which effectively reduces the activation energy required for formation of ATP in the "T" state. Further conformational change of the β subunit from "L" to "T" state facilitates very tight association between ADP+ Pi resulting in the formation of ATP within the catalytic site. Finally, the conformational change of the β subunit once again to the "O" state releases the bound ATP from the catalytic site and the empty site is again available for binding of ATD+ Pi. In summary the complete 360° rotation of γ subunit induces the formation of 3 molecules of ATP [55-56, 59, 61 and 68].

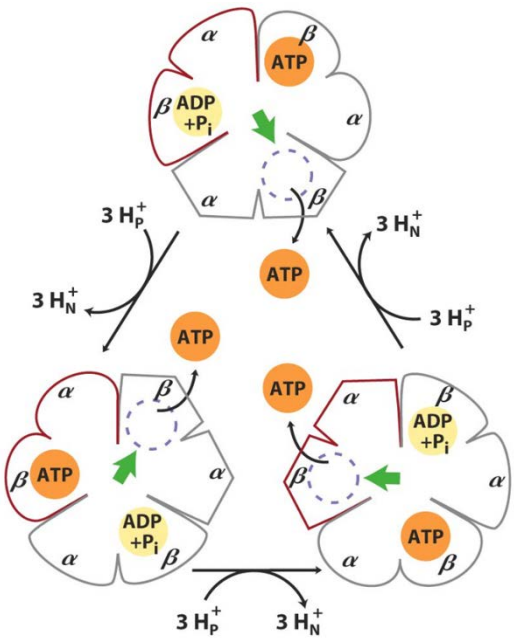


Figure-5: The binding change mechanism of the ATP synthesis or hydrolysis. The green depict the direction of rotation leading to synthesis of ATP molecule. When a similar process happens in reverse direction, it results in ATP hydrolysis. Adopted from [69].

The reverse process results in ATP hydrolysis activity of F-ATPase. For this, the F_1 sector uses the stored chemical energy available due to tightly bound ATP in the catalytic site to generate a torque force which subsequently induces the F_O sector to rotate and act as a proton pump in the reverse direction. This reverse activity of the F-ATPase represents a very powerful mechanism to generate or maintain the indispensable proton gradient across the inner mitochondrial membrane. Certain anaerobic bacteria which lack a functional respiratory chain use the reverse activity of F-ATPase to hydrolyze ATP for maintaining ion gradient across their cell membrane. This is necessary for driving vital bioenergetic processes such as chemotaxis and membrane transport [57, 61]. In bacteria hydrolysis of ATP is also coupled to maintenance of intracellular ion concentration and pH homeostasis [61, 70].

Regulation of F-ATPase activity

The bacterial F-ATPase carries out both ATP synthesis and hydrolysis depending on whether the organism is growing in aerobic or anaerobic condition. In contrast to bacterial F-

ATPase, the mitochondrial and chloroplast F-ATPase enzymes mainly carry out synthesis of ATP molecules. The wasteful hydrolysis of ATP by F-ATPase in prevented by binding of Mg^{2+}-ADP to one of the catalytic sites in mitochondrial and chloroplast F-ATPAse, when the transmembrane potential and matrix pH is low [57]. It is reported that in chloroplast the ATP hydrolysis activity is light dependent and in the absence of light the transmembrane proton gradient generated by photorespiration is absent. In this condition the γ subunit of chloroplast F-ATPase forms disulfide linkage with the α subunit of the catalytic core resulting in inhibition of ATP hydrolysis. The thioredoxin regulated reduction of disulfide bond happens again when day light is restored [57, 71].

Similarly, the futile cycle resulting in the hydrolysis of ATP by mitochondrial F-ATPase is well studied, at least in yeast and mammalian systems. The respiring mitochondria have high membrane potential (~150-180 mV) and high matrix pH, which together favor ATP synthesis. However, in instances when the mitochondrial homeostasis is disrupted, such as during very low oxygen supply, the membrane potential decreases below a critical level and this triggers ATP hydrolysis and proton pumping in the reverse direction by F-ATPase, which helps to restore the membrane potential. The hydrolysis of ATP and the coupled reverse proton pumping activity of F-ATPase is regulated by the inhibitory factor 1 (IF1) protein, which is also a mitochondrial matrix pH sensor protein. The IF1 protein is present in two forms - the active dimeric form and inactive tetrameric form. At higher mitochondrial matrix pH (i.e., pH = ~8.0), this protein is present in its tetrameric form and is unable to interact, *via* its N-terminal inhibitory domain, with the catalytic core in the F_1 sector of F-ATPase. When the mitochondrial matrix pH becomes acidic, the tetrameric IF1 dissociates into dimers, and the dimeric IF1 (which is the active form of the protein) is capable of interacting with the F_1 sector catalytic core of F-ATPase.

Detailed structural studies on bovine mitochondrial F-ATPase and IF1 interaction indicate that in its active form, IF1 interacts mainly with the F_1 γ subunit and the catalytic center at the interface between ADP bound α and β subunits. This interaction prevents the wasteful hydrolysis of ATP by inhibiting the conformational change required for ATP hydrolysis [57,59,72]. Some studies also suggest that the IF1 protein is involved in dimerization of the mitochondrial F-ATPase complex, but the exact role of IF1 in this process is not clear. For example, despite IF1 being absent in *Saccharomyces cerevisiae*, the yeast F-ATPase is capable of assembling into dimers on the inner mitochondrial membrane.

Supramolecular organization of mitochondrial F-ATPase

The F-ATPase in bacteria and chloroplast are exclusively present in the monomeric form, and the subunits required for dimerization are absent in this case [57, 59-60, 73]. In contrast to bacterial and chloroplast F-ATPase, the mitochondrial F-ATPase assembles into dimers and other oligomeric forms, as demonstrated by detailed biochemical and structural studies using cryo electron microscopy [74]. The first evidence of dimeric and oligomeric organization of mitochondrial F-ATPase was reported in the EM study of *Paramecium multimicronucleatum* mitochondria [75]. Furthermore the dimeric and oligomeric form of mitochondrial F-ATPase are also reported in phylogenetically distant organisms that include yeast, protozoans, algae, plants, and mammals [76-82]. The detailed structural studies on purified dimeric F-ATPase from different species revels that the two F_1 headpieces are separated by an angle between 35°- 90° [80]. Now it is widely accepted that the dimeric F-ATPase are involved in mitochondrial inner membrane cristae formation (Figure-6) [82-84]. Detailed biochemical analysis of purified dimeric F-ATPase complex from yeast has revealed the supernumeratory proteins such as subunit e and g are involved in dimerization of monomeric F-ATPase complex [85-87]. Subunit e and g are hydrophobic single transmembrane proteins and their transmembrane domains have the GXXXG motif.

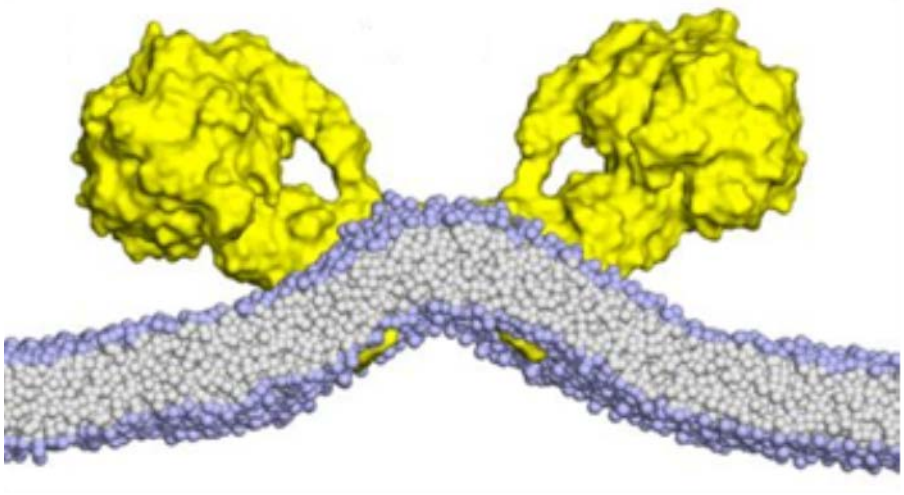


Figure-6: Membrane curvature induced by ATP synthase dimers. Cross sections through the membrane patch showing the curvature profile of the of the lipid bilayer. Adapted from [82].

This motif is important for the dimerization and stability of dimeric F-ATPase [87]. The yeast mutants, which are devoid of subunit e and g, show defective dimerization of F-ATPase and loss of cristae in the inner mitochondrial membrane, indicating a crucial role for these proteins in dimerization and mitochondrial inner membrane morphology formation [85-86]. The corresponding orthologs of these two proteins are also present in bovine F-ATPase [88-89]. It is also reported that in *in vivo* condition, yeast mitochondrial F-ATPase can dimerize even in the absence of subunit e, and this is probably mediated by peripheral stalk components of the enzyme [90]. It is proposed that the inner membrane cristae formation due to oligomeric form of F-ATPase increase the surface area of the inner membrane, leading to increased presence of OXPHOS components and increased proton pumping into the inter-membrane space, which in turn results in efficient production of ATP.

Divergence of mitochondrial F-ATPase

Despite the conserved structural features and functional mechanism of F-ATPase from bacteria to higher eukaryotes, significant divergence exists in the subunit composition of mitochondrial F-ATPase from diverse species. The mitochondrial F-ATPase enzyme is well

characterized from yeast and bovine species. In brief, the yeast and bovine mitochondrial F_1 sector include subunit (α_3,β_3,γ,δ and ε) and the F_O sector comprised of oligomeric subunit c $_{10-12}$ along with subunits a, b, d, f, 8, h, and OSCP (oligomycin sensitivity–conferring protein). In addition, other accessory subunits such as e, g, i, j, and k are also associated with the enzyme. Orthologs of all yeast F_O sector subunits, except subunits j and k, are present in the mammalian counterpart (bovine enzyme), which additionally has subunits AGP and MLQ [90, 91]. Recent studies on organisms from different phyla have indicated that the F_1 sector catalytic core is highly conserved across species. However, it has also become apparent that the orthologs for many of the F_O sector subunits of yeast or bovine mitochondria F-ATPase are absent in most other eukaryotic species. For example, in case of the mitochondrial F-ATPase from photosynthetic free living algae *Chlamydomonas reinhardtii* and Polytomella species, except for subunits c, a and OSCP, yeast and bovine orthologs of other F_O sector subunits are absent. It is now known that at least 9 novel proteins are present as F_O sector subunits in these species, likely in place of the missing yeast and bovine counterparts, and these are conserved across the chlorophyceae algae [93-94]. The divergence in F_O sector subunit composition is not unique to chlorophyceae algae. The protozoan parasite *Trypanosoma brucei* (phylum kinetoplastida) also shows the presence of intact canonical F_1 sector subunits and only subunit c and OSCP from F_O sector. Mass spectrometry based proteomics analysis on purified F-ATPase from *T. brucei* identified at least 14 novel proteins as subunit components of the enzyme [95]. Similarly the Euglenoidea algae *Euglena gracilis* shows presence of 20 novel subunits as part of its F-ATPase while many of the canonical F_O sector subunits are missing [96]. *In silico* analysis of these 20 novel F-ATPase subunits from *E. gracilis* reveals that 7 of these subunits are common in both *T. brucei* and *E. gracilis* [96]. Divergence in F-ATPase subunit composition is also reported in case of *Tetrahymena thermophila* (phylum ciliate; super-phylum alveolata). Mass spectrometry analysis of purified *T. thermophila* mitochondrial F-ATPase identified 13 novel

subunits, including the novel mitochondria genome encoded hydrophobic protein Ymf66 as the counterpart of yeast/bovine subunit a [97].

The *T. gondii* mitochondrial F-ATPase

Comparative genomics studies on *T. gondii* and many other alveolate species, including all apicomplexan parasites, revealed that many of the canonical F_O sector subunits are missing in this entire super-phylum (Figure-7) [26-27, 98-99]. This is intriguing since this super-phylum includes many diverse species, parasitic and free-living, belonging to phylum apicomplexa, dinoflagellata, chromerida and ciliata. Surprisingly, orthologs for the novel subunits identified from the ciliate F-ATPase were not present in other alveolate species, indicating key evolutionary difference within the alveolate clade [97]. In contrast to the F_O sector subunits, the F_1 sector subunits are well conserved across all alveolate phyla, except in the cryptosporidium species, which lack a true mitochondrion organelle [26-27,98-99].

The missing subunits of the *T. gondii* F-ATPase are F_O sector subunit a, which is involved in proton translocation along with the conserved subunit c [63-64], and b, d, F6 subunits, which are core components of the peripheral stalk required for counterbalancing the torque generated during the rotary motion of the central stalk by holding the catalytic core in relative position to subunit a [66-67], and other additional subunits corresponding to the supernumeratory proteins present in yeast and bovine mitochondrial F-ATPase which are involved in oligomerization of the F-ATPase complex [85-87].

Given that glycolysis is a major source of ATP, at least in *P. falciparum* and *T. gondii*, and the fact that many essential F-ATPase subunits are missing in apicomplexan species, it was suggested that this F-ATPase may not be functional as a genuine ATP synthase. However, subsequent studies in *T. gondii* established that glycolysis is not an essential source of ATP [18-19]. In the absence of glycolysis, oxidative phosphorylation appears to be the main source

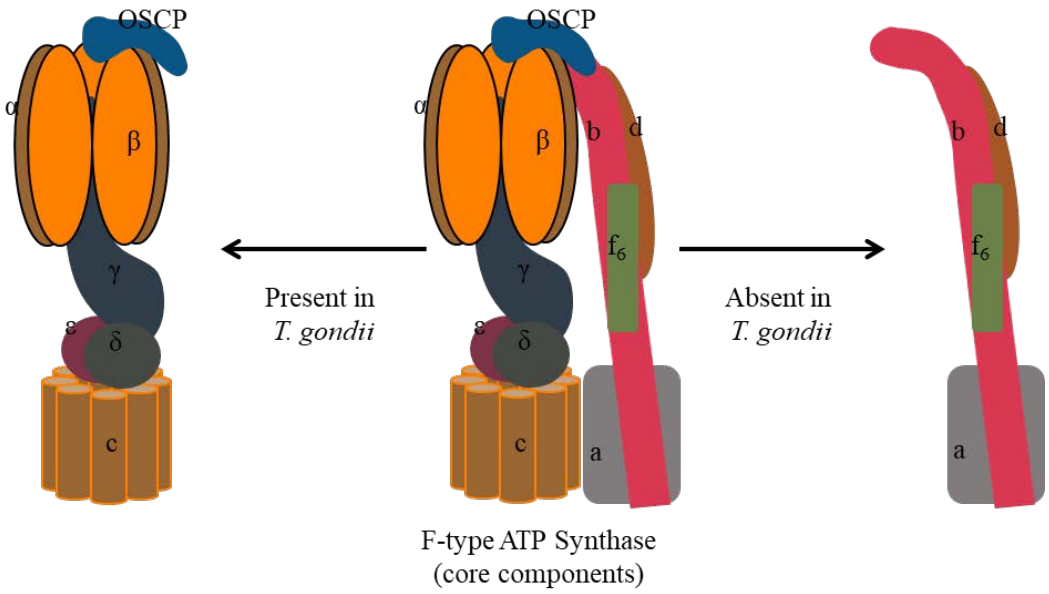


Figure-7: Schematic representation of core subunit composition of *T. gondii* F-type ATP synthase. The different subunits are color coded and annotated. Middle, the complete set of canonical subunits of the core F-ATPase enzyme; left and right, subunits identified *in silico* to be present and absent from *T. gondii*.

of cellular ATP in *T. gondii*, despite the absence of crucial subunits required for assembly and function of mitochondrial F-ATPase. An active oxidative phosphorylation was demonstrated in *Plasmodium* sp. [35-37], and in *P. falciparum* the presence of monomeric and dimeric F-ATPase from detergent solubilized mitochondrial lysate was demonstrate [99]. The same study also reported that the F_1 sector β and γ subunit are essential for asexual development of *P. falciparum*.

In summary, It appears that a full complement of subunits, typical for the eukaryotic enzyme, is present in *P. falciparum* and likely in other apicomplexan parasites as well [99]. Surprisingly another study in the rodent malaria *P. berghei*, the ATP synthase β subunit is dispensable for blood stage asexual development but essential for completion of sexual development in the mosquito [100]. Thus it appears probable that in all apicomplexan parasites the mitochondrial F-ATPase is fully functional and the missing subunits might be complemented by novel and highly diverged proteins as reported in other eukaryotes [93-97].

Scope of thesis

The mitochondrial F-ATPase is highly conserved and required for synthesis of ATP by oxidative phosphorylation. Comparative genomics reveals the missing or incomplete nature of F-ATPase in apicomplexan parasites. Despite this, the oxidative phosphorylation is functionally intact in apicomplexan parasites. Previous studies in the human malaria parasite *P. falciparum* have revealed the presence of monomeric and dimeric form of F-ATPase, but identity of the novel subunits is not known. This study also showed that the enzyme is essential for parasite survival. Thus it is important to carry out a detailed study on the subunit composition, structure and function of the F-ATPase from apicomplexan parasites.

For this study, we have decided characterize the mitochondrial F-ATPase from the model apicomplexan parasite *T. gondii*. It is very important to understand the complete subunit composition and the role of ATP synthase in apicomplexan parasites, since this enzyme is already identified as a valid drug target. Thus, a detailed understanding of the *T. gondii* mitochondrial F-ATPase will facilitate the discovery of new antiparasitic drugs. The specific aim and objective of this thesis work are listed below.

1. Carryout complete mapping of *T. gondii* F-type ATP synthase subunits and its phyletic profile across the apicomplexan phylum.
2. Demonstrate that the *T. gondii* F-type ATP synthase is actually involved in ATP synthesis *via* oxidative phosphorylation.
3. Generate transgenic parasite expressing tagged (fluorescent/epitope) version of *T. gondii* F-ATPase and using these parasites for characterizing the oligomeric status and identifying novel subunit components of the enzyme.
4. Validate the newly identified subunits of *T. gondii* F-ATPase and assign the functional equivalents among these novel proteins to canonical eukaryotic counterparts.

Research Methodology and Results

Research Methodology:

Molecular reagents and methods: Genomic DNA (gDNA) and total RNA from tachyzoite stage of *T. gondii*, and plasmid DNA from recombinant *E. coli* (DH5α) were isolated using Qiagen kits (Germany). From total RNA complementary DNA (cDNA) was prepared using the reverse transcription kit from Thermo Fisher Scientific, United States of America (USA). The manufacturer's protocol was followed while using the various kits. gDNA and cDNA PCRs were done using the proof reading DNA polymerases AccuPrime™ *Pfx* from Thermo Fisher Scientific (USA) or LA Polymerase from Takara Bio (Japan) respectively. All primers used in this study (Table-1) were obtained from Integrated DNA Technologies (USA). Ligase and various restriction endonucleases used in this study were purchased from either New England BioLabs (NEB), (USA) or Promega (USA).

Plasmid constructs: Topo 2.1 plasmid backbone from Thermo Fisher Scientific (USA), were used to clone various PCR products. The PCR product were first cloned either into pCR™ Blunt II-TOPO® vector Thermo Fisher Scientific (USA), or pGEM®-T Easy Vector Systems, Promega (USA), and sequenced before sub cloning, digested with specific enzymes shown as underline sequence (Table-1) and cloned into Topo 2.1 plasmid. Topo 2.1 plasmid were used for either modification of target genomic loci or ectopic expression of full-length cDNA for gene of interest (GOI) in *T. gondii*. The plasmid construct used for tagging the 3' end of F-ATPase subunit β (*Tg*atpβ) and oscp (*Tg*atposcp) genes with YFP-HA tag had the following features: 2,606 bp (BglII and AvrII) and 1,831 bp (BamHI and AvrII) PCR fragments amplified from the 3' end of *Tg*atpβ and *Tg*atposcp loci, respectively, without the stop codon, were cloned into the Topo 2.1 plasmid, upstream of a YFP-HA tag coding sequence, followed by the 3' UTR of the *T. gondii* dihydrofolate reductase-thymidilate synthase (*Tg*dhfr-ts) gene [101]. For generating stable transgenic parasite lines, the plasmid which was used for transfecting the parasite contained the DHFR cassette (a mutant version of the *Tg*dhfr cDNA that confers

pyrimethamine resistance [102]) as a NotI fragment. A modified Topo 2.1 plasmid backbone were used for ectopic expression of selected full-length cDNA of *T. gondii* gene. The plasmid contains the *T. gondii* β-tubulin gene promoter for constitutive expression of GOI, HA- 3' UTR from *Tg*dhfr-ts gene and CAT cassette. The CAT cassette expresses the chloramphenicol acetyl transferase gene, which confers resistance to chloramphenicol [103]. The full-length cDNA of selected *T. gondii* genes encoding the novel F-ATPase subunits identified from this work were cloned as BamHI and XbaI/NheI fragments in frame with HA-3' UTR for constitutive expression. Chloramphenicol was used to generate and select stable transgenic parasite lines. For gene knockout studies using CSISPR/ Cas9 technique, the previously published [104] plasmid and protocol were used.

***T. gondii* culture and genetic manipulation:** *T. gondii* parasites Type-I RH [RH-wild type (RH-wt)] and Type-I RH ΔKu80 (RHΔKu80) [105] strains were used in this study. In brief human foreskin fibroblasts (HFF) cells were used as host cell for parasites infection. The HFF cells were grown in Dulbecco's Modified Eagle Medium (DMEM) with High glucose, supplemented with 2mM GlutaMAX, 25mM HEPES, 50 μg/ml Gentamicin and 10% heat-inactivated fetal bovine serum. The HFF cells were maintained into T25 flasks at 37 °C in a humidified atmosphere containing 5% CO_2. The tachyzoite stage of parasites were propagated by inoculating approximately 10^5 freshly harvested tachyzoites into T25 flask containing 2 week old confluent HFF cells. The medium used for parasites propagation is similar to that used for host cell propagation except the medium lack serum components. All cell culture grade reagents were procured from Thermo Fisher Scientific (USA).

To obtain tachyzoite stage of parasites devoid of host cell debris, the HFF monolayer were scrapped after 48 hours post infection of tachyzoites, the cell suspension were physically disrupted by passing it through a 22-gauge needle and filtered it through a 3 μM nucleopore membrane obtained from Whatman, GE Healthcare, (USA). Following isolation of tachyzoites

the number of parasites present per 1 ml of suspension was estimated from cell counts obtained using a hemocytometer. For generating transgenic parasites expressing F-ATPase subunit *Tg*ATPβ and *Tg*ATPOSCP tagged with YFP-HA proteins from the endogenous loci, the RHΔKu80 parasites were transfected with the respective tagging plasmids. Before transfecting, the plasmids DNA containing β (*Tg*atpβ) and oscp (*Tg*atposcp) in frame with YFP-HA epitope tag were linearized with BstZ17I and BstBI restriction enzymes respectively and ethanol precipitated. Transfection was achieved by electroporating approximately 10^7 freshly harvested tachyzoite-stage parasites resuspended in 400 µl parasite culture medium containing 50 µg of linearized sterile plasmid DNA, using a BioRad Gene Pulser system (USA), having 10 µF capacitance, ∞ ohms resistance, and 1.5 kV voltage settings. Transfected parasites were immediately inoculated into a T25 flask containing HFF monolayer and allowed to invade and replicate for 12-15 hours before beginning drug selection with 1 µM pyrimethamine. For ectopic expression of cDNA with HA tags for selected genes, RH-*wt* parasites were transfected with the respective plasmid constructs, and stable transfected line were selected using 20 µM chloramphenicol. Clonal lines of stable transgenic parasites were isolated using the limiting dilution technique [106].

Microscopy and western blotting: For visualizing the expression and subcellular localization of selected newly identified ATP synthase associated proteins (ASAPs), the respective genes were fused to either YFP-HA or HA tags, and expressed from the endogenous locus or ectopically from a plasmid backbone. The expressed F-ATPase subunit proteins were imaged by microscopy. The transgenic parasites expressing reporter tagged genes of interest were allowed to infect confluent HFF monolayers cultured on glass coverslips. After 24 hours post infection the HFF monolayer were treated with 250 nM Mitotracker Red Thermo Fisher Scientific (USA), for 20 minutes. Following that the monolayer was washed with 1× phosphate-buffered saline (PBS) and fixed for 1 hour at ambient temperature using 4%

paraformaldehyde. The coverslip were washed with water and mounting on glass slides were done using the fluoroshield reagent obtained from Sigma (USA). The slides were imaged using the 63× oil immersion objective fitted to the Axio Observer inverted fluorescent microscope, from Carl Zeiss (Germany), and images were processed using the Zen software Carl Zeiss, (Germany). Intrinsic fluorescence from YFP and Mitotracker Red were visualized using the excitation/emission filter combination of 493/520 and 578/599, respectively. HA-tagged proteins were visualized by immunofluorescence staining, using rabbit α-HA primary antibodies (1:1,000) followed by Alexa 488 conjugated goat anti rabbit secondary antibody (1:1,000), both purchased from Thermo Fisher Scientific (USA). After fixing, the cells were permeabilized for 5-10 minutes using 0.25% Triton X-100 in 1× PBS. They were then treated with 2% fetal bovine serum in 1× PBS for 30 minutes, followed by primary antibody for 1 hour, 3 times washing with 1× PBS containing 0.25% Triton X-100, and secondary antibody for 1 hour. Finally, the coverslips were washed 3 times with 1× PBS containing 0.25% Triton X-100, followed by a water wash, before mounting them on glass slides for imagining.

For western blotting, mitochondria isolated from transgenic parasites expressing endogenous tagged F-ATPase subunits with YFP-HA were used. The purified mitochondria were solubilized in 1× Laemmli buffer (120 mM Tris-HCL pH 6.8, 2% SDS, 10% glycerol, and 0.01% w/v bromophenol blue), denatured at 94 °C for 10 minutes and resolved on 10% SDS-PAGE. Following this, electro-transfer to PVDF membrane was done using Tris-glycine buffer pH 8 containing 20% methanol at 65 mA for approximately 1 hour and 20 minutes. The blot was kept in blocking buffer (5% skimmed milk in 1× PBS) overnight at 4 °C and then treated with primary antibody (1:5,000; rabbit α-HA monoclonal antibody, Thermo Fisher Scientific, USA) for 1 hour at ambient temperature. Following, this the blot was washed 3 times with 1× PBS containing 0.1% Tween 20. Subsequently the blot was incubated with horseradish peroxidase-coupled secondary antibody (1:5,000; donkey anti-rabbit antibody, Nif 824 from

GE Healthcare, USA), for 1 hour at ambient temperature, and washed 3 times with 1× PBS containing 0.1% Tween 20. The blot was then developed using either the 3,3'-diaminobenzidine substrate form Sigma (USA) or the chemiluminescent ECL western blotting kit from GE Healthcare (USA).

Measurement of intracellular ATP: Tachyzoites stage of RHΔKu80 (parental strain) and RHΔKu80 parasites expressing F-ATPase subunit *Tg*ATPβ and *Tg*ATPOSCP tagged with YFP-HA (transgenic line) were propagated in Minimum Essential Medium (MEM) supplemented with 2mM GlutaMAX, 25mM HEPES, 50 µg/ml Gentamicin. After 48 hours of post infection, parasites were harvested, and the freshly isolated extracellular tachyzoite were further washed with DMEM without glucose. The tachyzoite were incubated in culture media that either contained 5.5 mM glucose and 4 mM glutamine or 0 mM glucose and 4mM glutamine, in order to evaluate the ability of the parasites to maintain ATP homeostasis. The mtETC of the parasite was inhibited using the antiparasitic drug atovaquone [50], which specifically inhibits the Cytb protein leading to inhibition of mitochondrial ATP synthesis produced by oxidative phosphorylation. After 2 hours of incubation at 37 °C in a humidified atmosphere containing 5% CO_2, total cellular ATP was measured using the ViaLight Plus Cell Proliferation and Cytotoxicity Bioassay Kit obtained from Lonza (Switzerland), as per the manufacturer's protocol. The luminescence readout was quantified using the Varioskan Flash plate reader from Thermo Fisher Scientific (USA).

Preparation of mitochondria from *T. gondii*: Mitochondria were isolated from freshly harvested tachyzoite-stage transgenic parasites expressing either *Tg*ATPβ-YFP-HA or *Tg*ATPOSCP-YFP-HA tagged epitope. Approximately 10^9 parasites were used for mitochondria preparation. The harvested tachyzoite stage parasites were first washed with PBS and resuspended in 1-2 ml of hypotonic lysis buffer (15 mM phosphate buffer pH 7.5 and 2 mM Glucose). The cells were lysed by sonication in an iced water bath for 30 minutes, and

samples were centrifuged at 2,000 g for 2 minutes at 4 °C to remove unbroken parasites. The supernatant was recovered into a new tube and centrifuged at 21,000 g for 15 minutes at 4 °C to pellet the crude mitochondria. The unbroken parasites again resuspended in hypotonic lysis buffer and repeat the lysis procedure till complete lysis of parasites was observed. The mitochondrial pellet obtained from each steps was then resuspended in 500 - 1000 µl of mitochondria storage buffer (320 mM sucrose, 1 mM EDTA, 10 mM Tris pH 7.4) [107]. Total protein in the mitochondria suspension was estimated by Bradford method [108], and the mitochondrial lysate was stored in −80 °C until further use. When required, the mitochondria were recovered from storage buffer by centrifuging at 21,000 g for 15 minutes at 4 °C and the mitochondrial pellet further solubilized using mitochondria solubilization buffer A (MSB-A; 50 mM NaCl, 50 mM Imidazole, 2 mM 6-Aminohexoanic acid, 1 mM EDTA, pH 7.0) containing β-dodecyl maltoside (DDM) detergent in total protein to detergent ratio of 1:5. Solubilization was allowed to continue overnight with rocking at 4 °C, and the solubilized fraction was separated by centrifuging at 100,000 g for 20 minutes at 4 °C.

Blue Native PAGE (BNP) separation, in-gel ATPase activity assay, and Native western blotting: In order to identify the monomeric and dimeric forms of the F-ATPase in *T. gondii*, the supernatant from detergent solubilized mitochondrial lysate was subject to one-dimensional BNP analysis [109_ 110]. Samples were prepared by adding 50% glycerol to a final concentration of 5%, and Coomassie Blue G 250 5% (w/v) dye was added such that DDM detergent to dye ratio was 8:1. Solubilized samples were separated on a 3% - 12% native PAGE from Thermo Fisher Scientific (USA), at 150 V setting for 30 minutes in cathode buffer A (50 mM Tricine, 7.5 mM imidazole pH 7.0, 0.02% Coomassie Blue G 250 dye) which was then switched to cathode buffer B (similar to cathode buffer A except the Coomassie Blue G 250 dye was 0.002%), while the imidazole (25 mM, pH 7.0) was used as the anode buffer, and separation was continued till the blue dye exited the gel. The gel was visualized under white

light to identify the various dye stained bands.

For in-gel ATPase assay, after sample separation by BNP, the gel was incubated overnight in ATPase assay buffer (35 mM Tris-HCL pH 7.8, 270 mM glycine, 14 mM $MgSO_4$, 0.2% $Pb(NO_3)_2$, and 8 $_{mM}$ ATP) The gel was then washed with water, and the ATPase activity of F-ATPase was observed by the formation of a milky white precipitate, visible against a black background, at the region corresponding to the expected monomeric and dimeric enzymes complex protein bands[110].

For native western blotting after BNP separation, proteins were electro-transferred onto a PVDF membrane using Tris-glycine buffer pH 8.0 without methanol at 65 mA for approximately 1 hour and 20 minutes. The membrane was then fixed in 8% acetic acid for 15 minutes, air dried at room temperature for 30 minutes, and washed with methanol several times to remove the Coomassie dye stain [99]. Further steps were similar to that described above for SDS PAGE western blotting.

Immunoprecipitation (IP) of F-ATPase complex: The transgenic *T. gondii* parasites expressing YFP-HA tagged *Tg*ATPOSCP subunit was used for immunoprecipitating the F-ATPase complex. The supernatant from detergent solubilized mitochondrial preparation were subjected for immunoprecipitation of the F-ATPase complex using α-HA antibodies obtained from Thermo Fisher Scientific (USA). The antibodies were first cross-linked to Protein A/G Plus agarose beads from Santa Cruz Biotechnology (USA), using the following protocol. Approximately 50 µl of Protein A/G Plus agarose beads were washed twice in 500 µl 1× PBS (pH 7.4) at 4 °C and mixed with 1 ml of 1× PBS containing 5-7 µg of rabbit α-HA antibody and incubated overnight at 4 °C with gentle mixing on rocker. The beads were then separated by centrifugation at 1,000 g for 2 minute at 4 °C, equilibrated with 0.2M Triethanolamine (pH 8.2) for 2 minutes, and then washed twice with the same buffer. The bound antibodies were

then crosslinked to the beads using 20 mM DMP from Sigma (USA), in 1 ml of 0.2M Triethanolamine at room temperature for 45 minutes, followed by washing with 1 ml of 50 mM Tris pH 7.5 twice for 15 minutes to quench the cross-linking reaction. Unbound antibodies are removed using 3 quick washes with 0.2M glycine buffer pH 2.3. The antibody conjugated beads were then washed twice with 1× PBS and equilibrate with MSB containing the detergent DDM at critical micelle concentration. The supernatant from solubilized mitochondria was added to the antibody coupled beads and incubated overnight at 4 °C with gentle mixing on rocker. The beads were then washed 3 times with 1× PBS containing DDM at critical micelle concentration. Proteins captured by the antibodies were eluted in 100 µl of 0.2M glycine buffer (pH 2.3) in 3 rounds, and the pooled eluate was neutralized with 1M Tris (pH 8.0) and concentrated using a 3 KDa cut-off concentrator purchases from Amicon Millipore (Germany). The samples were equilibrated with 0.1% RapiGest from Waters (USA), in 50 mM ammonium bicarbonate buffer in preparation for LC-MS analysis.

Purification of native F-ATPase by chromatography: The transgenic *T. gondii* parasites expressing YFP-HA tagged *Tg*ATPβ subunit were used for purification of F-ATPase complex. The protocol that was previously reported for the purification of the F-ATPase from the *Polytomella sp*., [111] was used. About 13 mg of crude mitochondrial sample were solubilized overnight at 4 °C in 3 ml of MSB-B (MSB-B; 50 mM Tris HCl [pH 8.0], 1 mM $MgCl_2$, and DDM at protein to detergent ration 1:5). The supernatant was loaded onto a HiTrap DEAE sepharose column GE Healthcare Life sciences (USA), having a bed volume of 1 ml and equilibrated with MSB-C (MSB-C; 50 mM Tris HCl [pH 8.0] $MgCl_2$ 1 mM, 0.05% [w/v] DDM). After completion of loading, the column was first washed with 10 column volumes of MSB-C_{20} (50 mM Tris HCl [pH 8.0], 1 mM $MgCl_2$, 20 mM NaCl, and 0.05% [w/v] DDM) and eluted with a 10 column volume linear gradient using MSB-C_{20} and MSB-C_{500} (50 mM Tris HCl [pH 8.0], 1 m $MgCl_2$, 500 mM NaCl, and 0.05% [w/v] DDM) at a flow rate of 0.5 ml per

minute. Fractions (500 µl) were collected and checked by western blotting using α-HA antibodies. Peak positive fractions were identified, pooled, and concentrated to a final volume of 200 µl using the Vivaspin concentrator 100 kDa cut-off; GE Healthcare Life sciences (USA), and separated by size exclusion chromatography using the Superose 6 Increase 3.2/300 column GE Healthcare Life sciences (USA), having a bed volume of 2.4 ml. A 50 µl sample was loaded on the column previously equilibrated with MSB-C_{20} and then eluted with 1.5 column volumes of the same buffer at a 50 µl flow rate. Fractions (100 µl) were collected, and positive fractions, identified by western blotting using α-HA antibodies, were pooled and concentrated using a 3 KDa cutoff concentrator Amicon Millipore (Germany). Samples were then equilibrated with 0.1% RapiGest in 50 mM ammonium bicarbonate buffer in preparation for LC-MS analysis. All chromatography steps were carried out using the AKTApure system GE Healthcare Life sciences (USA).

Liquid Chromatography tandem Mass Spectrometry (LC-MS/MS) proteomics: To identify the subunit components of the F-ATPase from *T. gondii*, LC-MS/MS proteomic analysis was performed on the partially purified enzyme sample obtained from BNP, IP, and chromatographic purification. Sample preparation for LC-MS analysis was based on a previous report [112]. The bands corresponding to the dimeric and monomeric forms of the F-ATPase enzyme were excised out from BNP gel, cut to small pieces, destained with 50% acetonitrile in 50 mM ammonium bicarbonate, and dehydrated with 100% acetonitrile. The gel pieces were then treated with 10 mM DTT in 50 mM ammonium bicarbonate for 45-60 minutes at 56 °C to reduce the proteins, followed by alkylation in the dark with 55 mM Iodoacetamide in 50 mM ammonium bicarbonate at ambient temperature for 45 minutes. The alkylated gel pieces were wash with 50 mM ammonium bicarbonate, and the gel pieces were dehydrated as stated above. Then, trypsinization of protein was carried out, starting with rehydration of the dehydrate gel pieces at 4 °C for 30 min in approximately 100 µl of 50 mM ammonium bicarbonate containing

12.5 ng/ µl trypsin procured from Sigma (USA). The rehydrated gel pieces were further covered by adding 50 mM ammonium bicarbonate and incubate overnight at 37 °C. Trypsinization were stop by acidifying the solution and then peptides were extracted using 50% acetonitrile in 2% formic acid. The extracts were dried using a speedvac, and the peptides were reconstituted in 50 µl of 50 mM ammonium bicarbonate, acidified with HCl, and desalted using the C_{18} Zip tip columns Millipore (Germany).

Peptide desalting was done using C_{18} Zip tip column, by first dehydrating the column with 100% acetonitrile solution by pipetting and dispensing the solution 3-4 times. Following this the column was equilibrated with hydrating solution (50% acetonitrile in 0.1%Trifluoroacetic acid [TFA]) as mentioned above. The hydrated column was washed 3-4 times with 0.1% TFA, and the tryptic digested protein samples were loaded by pipetting in the peptide sample. The column was then washed with 0.1% TFA and the bound peptides were eluted in hydrating solution. The samples was then dried in a speedvac and stored at −80 °C until further use.

Samples obtained from IP and size exclusion chromatography were immediately equilibrated with 0.1% RapiGest as stated above. The mixtures were then heated to 80 °C and maintained at this temperature for 15 minutes. The denatured proteins were reduced by 100 mM DTT at 60 °C for 15 minutes, followed by alkylation with 200 mM iodoacetamide at ambient temperature in the dark for 30 minutes. Samples were then treated overnight with trypsin at 20:1 substrate to enzyme ratio at 37 °C. Trypsinization was stopped by addition of 2 µl of 1 N HCL and incubated at 37 °C for 20 minutes, after which tryptic digested samples were vortexed and centrifuged at 21,000 g for 10 minutes at room temperature. The peptide samples were desalted using the C_{18} Zip tip Millipore (Germany), dried, and stored at −80 °C until further use.

The peptide samples were analyzed using LC-MS^E workflow on the nano ACQUITY (UPLC) Synapt system Waters Corporation, (USA). The digested, desalted, speedvac dried peptide samples were reconstituted in reconstitution solution (3% acetonitrile in 0.1% formic acid) having approximately 100 ng/µl as peptide concentration. The reconstituted sample (4 µl) was injected into a 5 µm Symmetry C_{18} trapping column (180 µm × 20 mm) at a flow rate of 5 µl/ minute, and peptides were eluted by using the following protocol: 3% - 40% B (B composition: 100% acetonitrile with 0.1% formic acid) for 90 minutes, 40% - 85% B from 90 - 105 minutes and 97% A (A composition: 0.1% formic acid in water), and 3% B from 105 - 120 minutes, on a bridge ethyl hybrid C_{18} column (75 µm × 250 mm, 1.7 µm) at a flow rate of 250 nl/minute. The system was coupled to the Synapt High Definition Mass Spectrometer Waters Corporation (USA), with a nanoLockspray as the ion source. Standard glu-fibrinopeptide B (Sigma, USA) as the lock mass calibrate peptide was infused into ion source having 500 nl/minute flow rate and sampled every 30 seconds. Acquisition of LC-MS^E data was performed in positive "v" mode with mass range m/z 50 - 1,900, having a scan time of 0.75 second, a constant low energy of 4 V for MS mode, and 20 - 40 V of collision energy during high-energy MS^E mode scan. A capillary voltage of 3.5 kV and cone voltage 38 V were maintained during analysis. LC-MS^E data were acquired and processed using Mass lynx and Protein Lynx Global Server (PLGS Version 2.5.3, Waters Corporation, USA) software respectively, for identifying the proteins with reference to the *T. gondii* proteome taken from ToxoDB.org release 36.

Gene coexpression analysis and phylogenetic studies: Normalized signal intensity values from microarray hybridizations performed at 13 different time points in the tachyzoite stage [113] were obtained directly from EupathDB.org. Normalized microarray intensities from time series data for intraerythrocytic stage *P. falciparum* were retrieved from published work [114]. Pearson correlation coefficients were calculated for all gene pairs using R (R-project.org). *P*-

values for gene coexpression was calculated using the Mann-Whitney U test [115]. Coexpression between genes encoding only the ATP synthase subunits, or coexpression between all other genes encoded by the parasite (i.e., all non-ATP synthase genes), or coexpression between genes encoding ATP synthase genes and non-ATP synthase genes was calculated. Since minute differences in very large samples may lead to very small *p*-values, the effect of size was therefore tested using Cohen's d test [116], for which values above 0.8 denote large effects due to size.

For phylogenetic analysis, orthologs for the novel subunits of *T. gondii* ATP synthase were identified from other alveolates, as reported previously [27]. Sequences were aligned with mafft (v7.222) [117] and Neighbor-Joining trees were constructed using Clustalw (2.1) [118] using 1,000 bootstrap replicates. The expected true phylogeny was adopted from a previous study [27]. The phylogenetic tree data has been deposited in TreeBASE (TB2:S22877).

Primer Name	Primer sequence
TGME49_261950_ F	TACAGATCTTCTGAACGGACTGCCTCGCG
TGME49_261950_ R	ACTCCTAGGCTTTCCGCTCGCCGCTTCCTGCG
TGME49_261950_ConF	CTGCGGCGCCGAATCCCGGAAAGAAACCAGC
TGME49_284540_ F	TACGGATCCGAAGCTCGCGATTTCGTGTTTCGG
TGME49_284540_R	ACTCCTAGGAAGAGGAGCCAGCAGCTGCGACTGC
TGME49_284540_ConF	CTGTGAACCACCGGGAGAAAACAACGCGGC
YFP_R	CCATGATATAGACGTTGTGGCTGTTGTAG
DHFR_ F	TCTGCGGCCGCCGCCAGGCTGTAAATCCCG
DHFR_R	ATCGCGGCCGCTCCTGCAAGTGCATAGAAGG
TGME49_245450_F	AATTCGGATCCATGCTGAACTTCATCCCGAAAAGATGCC
TGME49_245450_R	GAATTTTCTAGACTTGATGTTTTTCCCCTGGATGGGGTACG
TGME49_282180_F	AATTCGGATCCATGTCGCCGGTCGGACGCCTCTTTTTG
TGME49_282180_R	GAATTTGCTAGCTTTCGTCGTCGGGATGAAGACATCCGTG
TGME49_290030_F	AATTCGGATCCATGGGGCTCTCCCCGGCCTTC
TGME49_290030_R	GAATTTTCTAGAATGGTGTCCAGCGGCTTCCTC
TGME49_223040_F	AATTCGGATCCATGGCAGAGACTCGCGAAGGG
TGME49_223040_R	GAATTTTCTAGAAGAGTACTGCAGATCGGGCGCTCC
TGME49_261950 (β) gRNA	ACTCTTTGCCGACCAGTCGAGTTTTAGAGCTAGAAATAGC
DHFR_F with 30 bp overlap to TGME49_261950	*GAAGACCTACTCGATTCACCGCGCAGCTC*CGCCAGGCTGTAAATCCCG
DHFR_R with 30 bp overlap to TGME49_261950	*TATGCAACATAGGAAGACGCACCGGGTT*TCCTGCAAGTGCATAGAAGG
TGME49_249720 (C) gRNA	TCCGTCGGCGGTGTTGCGCAGTTTTAGAGCTAGAAATAGC
DHFR_F with 30 bp overlap to TGME49_249720	*GCAAGTCTTAGCGCCGCCATTGCTCTGATG*CGCCAGGCTGTAAATCCCG
DHFR_R with 30 bp overlap to TGME49_249720	*GAAACAAGAGCAGCAAAGAGAGACCCGATA*TCCTGCAAGTGCATAGAAGG

Table -1: List of primer used in this study: The restriction enzymes sites are underline in the 5' end of primer sequences; gRNA sequence is shown in red color; gene specific sequences are shown in blue.

Results

Comprehensive *in silico* analysis for identification of F-ATPase subunits from *T. gondii*:

The mitochondrial metabolic pathways are mapped very well across the eukaryotes, and this information is available in repositories such as the KEGG database [https://www.genome.jp/kegg/]. Using available genome sequence datasets for various apicomplexan parasites and few other related organisms belonging to the alveolate infrakingdom, a comprehensive *in silico* analysis on presence and absence of key enzymes involved in mitochondrial metabolisms in apicomplexan parasites and the within alveolate clade has been previously carried out [26-27, 98-99]. These studies have highlighted the missing subunits of the F-ATPase enzyme from these organisms. These findings were further confirmed in this study by efforts to identify the orthologs of corresponding yeast and bovine F-ATPase subunits from *T. gondii* and other alveolate species. These ortholog mapping studies revealed that all F_1 sector subunits of F-ATPase were present in these organisms. These are the α (TGME49_204400), β (TGME49_261950), δ (TGME49_226000), ε (TGME49_314820), and γ (TGME49_231910) subunits. However, in case of the F_O sector, only subunit c (TGME49_249720) and OSCP (TGME49_284540) were readily identified in *T. gondii* and other alveolate species. This suggested that all the subunits necessary for assembling the catalytic core ($\alpha_3\beta_3$), central stalk (δ, ε, and γ), and the oligomeric subunit c portions of the F-ATPase are encoded by *T. gondii*. However, as reported previously [26-27,98-99], it was not possible to identify the orthologs for critical F_O sector subunits involved in proton translocation and formation of the peripheral stalk structure of the enzyme in *T. gondii* and other alveolate species based on protein sequence similarity. (Figure-8). It is likely that the missing F_O subunits are either highly divergent in sequence or a completely novel set of proteins are present in place of the missing ones.

					Alveolate infrakingdom													
F-type ATP synthase module	Sub-unit	*Scer*	*Btau*	*Atha*	*Tgon*	*Ncan*	*Pfal*	*Pviv*	*Tpar*	*Bbov*	*Cmur*	*Cpar*	*Chom*	*Cvel*	*Pmar*	*Tthe*	*Crei*	*Tbru*
F_1 Sector	α	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	β	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	γ	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	+
	ε	+	+	+	+	+	+	+	+	+	+	-	-	+	+	-	+	+
	δ	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	+
F_O Sector	c	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	+
	OSCP	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	+
	a	+	+	+	-	-	-	-	-	-	-	-	-	-	-	+	+	+
	b	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	+
	d	+	+	+	-	-	-	-	-	-	-	-	-	-	-	+	-	-
	f	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	8/A6L	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	h/f6	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Other associated proteins	g	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	IF1	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	e	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	j	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	k	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	AGP	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	MLQ	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Novel subunits																13	9	14

Apicomplexa | Chromerida | Dinoflagellata | Ciliophora | Chlorophyceae

Kinetoplastida | Newly discovered & unique to phylum

Figure-8: Table showing the presence/absence of orthologs of yeast and bovine F-ATPase subunits in selected species from various alveolate clades: The alveolate species are highlighted in orange color. The different phylums are shown in color and identified below the ortholog table. Gray background with (+) symbol indicate the corresponding ortholog for yeast and bovine F-ATPase subunit is present, while white box with (-) indicates absence of the ortholog. Blue color indicates diverged (no clear orthologs) functionally equivalent species-specific novel protein identified experimentally; numeric values within the blue box in last row denotes the total number of novel subunits. Species names: *Scer*, *Saccharomyces cerevisiae*; *Btau*, *Bos taurus*; *Atha*, *Arabidopsis thaliana; Tgon*, *T. gondii*; *Ncan*, *Neospora caninum*; *Pfal*, *P. falciparum*; *Pviv*, *Plasmodium vivax*; *Tpar*, *Theileria parva*; *Bbov*, *Babesia bovis*; *Cmur*, *Cryptosporidium muris*; *Cpar*, *Cryptosporidium parvum*; *Chom*, *Cryptosporidium hominis*; *Cvel*, *Chromera velia*; *Pmar*, *Perkinsus marinus*; *Tthe*, *T. thermophila*; *Crei*, *Chlamydomonas reinhardtii*; *Tbru*, *T. brucei*. IF1, inhibitory factor 1; OSCP, oligomycin sensitivity-conferring protein.

The missing of key F_O subunits is not exclusive to alveolates, since many other unicellular eukaryotes also appear not to possess the corresponding orthologs for many of the yeast or bovine F_O subunits. In fact, novel F-ATPase associated proteins have been previously identified from *Tetrahymena* [97] belong to phylum ciliophora within alveolate infrakingdom, *Chlamydomonas* [93-94] a free living chlorophyceae algae and *Trypanosoma* [95] another protozoan belong to kinetoplastid group. Therefore, it is probable that a similar situation exists in the apicomplexan parasite *T. gondii*. To verify this, as detailed below, the *T. gondii* F-ATPase enzyme was purified in native form from isolated *T. gondii* mitochondria and its subunit composition was identified using mass spectrometry based proteomics approach.

Generating transgenic *T. gondii* parasites expressing F-ATPase β (*Tg*ATPβ) and OSCP (*Tg*ATPOSCP) subunits with yellow fluorescent protein plus hemagglutinin (YFP-HA) tag:

To facilitate the purification of the *T. gondii* F-ATPase enzyme in native form, tachyzoite stage *T. gondii* parasites were engineered to express in-frame genomic YFP-HA tags at 3' end of the respective genes encoding the ATP synthase F_1 β and F_O OSCP subunits (Figure-9). The modified genes continued to be expressed under the control of their endogenous promoters. The correct insertion of the tags was confirmed by genomic PCRs (Figure-10) from clonal isolates of the respective transgenic parasites.

Genomic locus tagging strategy

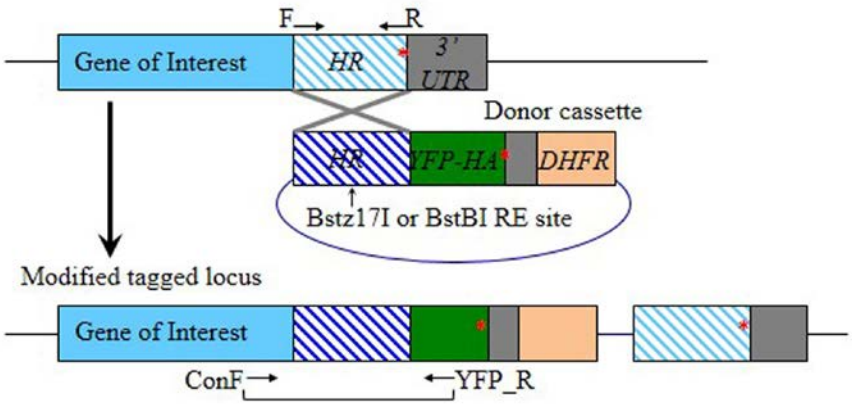


Figure-9: Schematic representation for genomic locus tagging for TGME49_261950 (*Tgatpβ*) and TGME49_284540 (*Tgatposcp*) with YFP-HA epitope tag. HR- Homologous region; YFP-HA-yellow fluorescent protein plus hemagglutinin; DHFR- *T. gondii* dihydrofolate reductase-thymidilate synthase; Red (*) - Stop codon.

Confirmation of endogenous gene tagging by diagnostic PCR

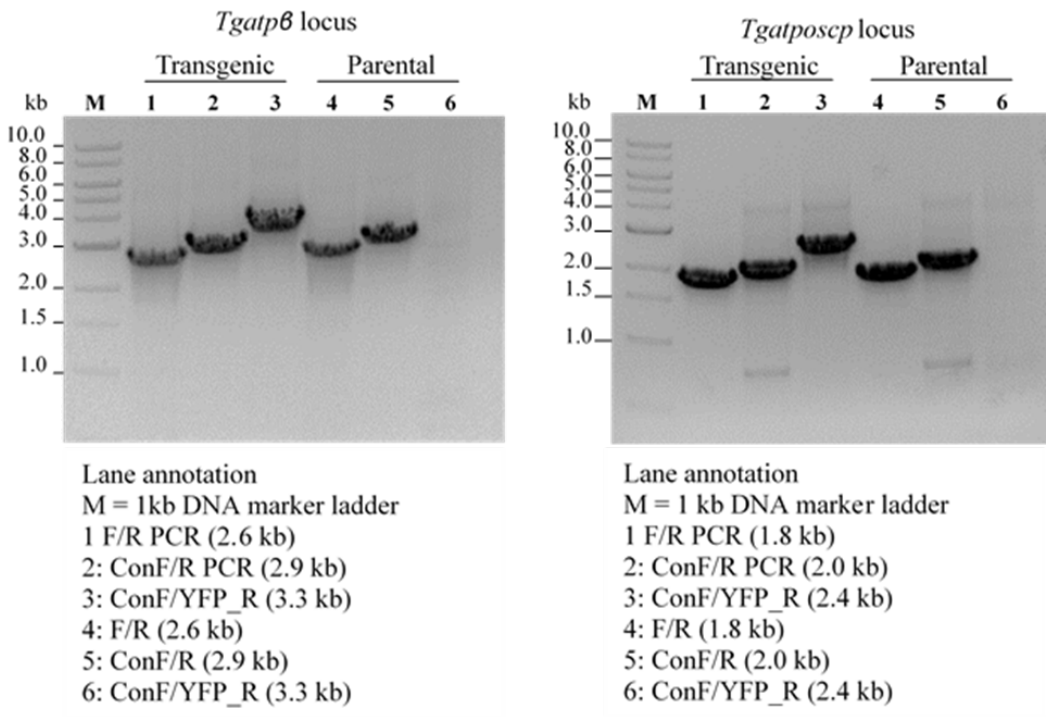


Figure -10: Confirmation of endogenous gene modification by diagnostic PCR. Using the respective forward (F) and reverse (R) primer pairs, the 3' end regions for the two genes was PCR amplified without stop codon, cloned into donor plasmid and linearized with respective enzymes. After transfection, the presence of the desired genomic locus modification was confirmed by genomic PCRs performed using the ConF and YFP_R primer pairs for the respective genes (see Figure 9). The two gel pictures show the results from genomic PCR amplifications confirming the endogenous tagging, which is evident from the presence and absence of the PCR products in lanes 3 (transgenic) and 6 (parental), respectively.

Expression of the desired YFP-HA tagged proteins and their mitochondrial localization was confirmed by Western blotting (Figure -11A) and microscopy (Figure-11B), respectively.

Confirmation of endogenous gene tagging by western blotting and its mitochondrial localization

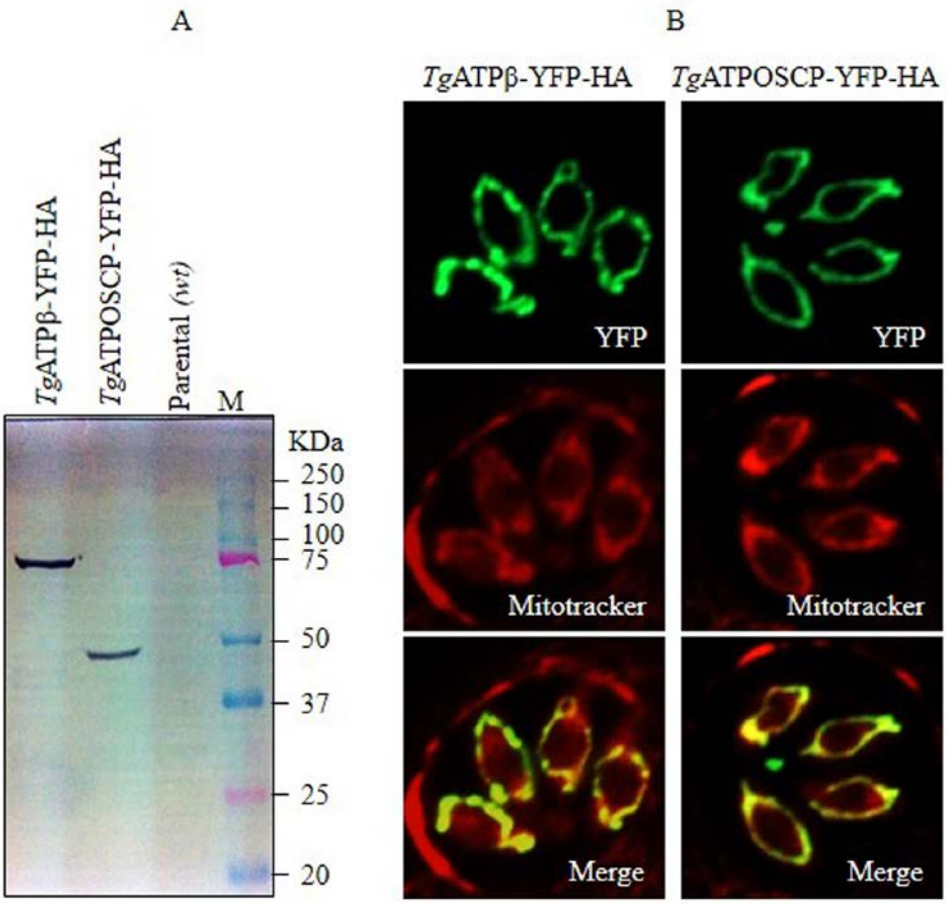


Figure-11: Confirming the expression of the *Tg*ATPβ-YFP-HA and *Tg*ATPOSCP-YFP-HA proteins by SDS-PAGE western blotting (A) and its mitochondrial localization (B). Partially purified mitochondrial sample from transgenic parasites were used to confirmed the expression of endogenous tagged proteins, from left to right cell lysate from *Tg*ATPβ-YFP-HA; *Tg*ATPOSCP-YFP-HA; parental parasites and M, Molecular weight size markers respectively (A). Mitotracker red was used to visualize the mitochondrion (B).

Functional validation of mitochondrial ATP synthesis: By comparing the ability of the parental and transgenic strains of the parasites to maintain total cellular ATP levels in the presence/absence of glucose, it is possible to find out whether the ATP synthesis function modified F-ATPase is intact in the transgenic parasites. Although *T. gondii* tachyzoites are known to prefer glucose as the primary nutrient source, in the absence of glucose, they can be

switch to glutaminolysis for carbon and energy supply. In the latter case, ATP is obtained by the parasites *via* oxidative phosphorylation, and this can be inhibited using atovaquone (Figure-12) [450, 98,119]. Based on these earlier observations, we designed an assay in which host cell free tachyzoite stage parasites were incubated in the presence or absence of glucose and atovaquone for a short period, before estimating the total cellular ATP levels. It was observed that similar to parental parasites, transgenic parasites expressing YFP-HA-tagged ATP synthase F_1 β and F_O OSCP subunits were capable of maintaining cellular ATP levels *via* oxidative phosphorylation in the absence of glucose, and this was susceptible to inhibition by atovaquone. These results confirm that the modification of either F1 β or F_O OSCP proteins with YFP-HA tag had no detrimental effect on the function of the enzyme, implicating that the structure of the tagged enzyme remained intact (Figure-13).

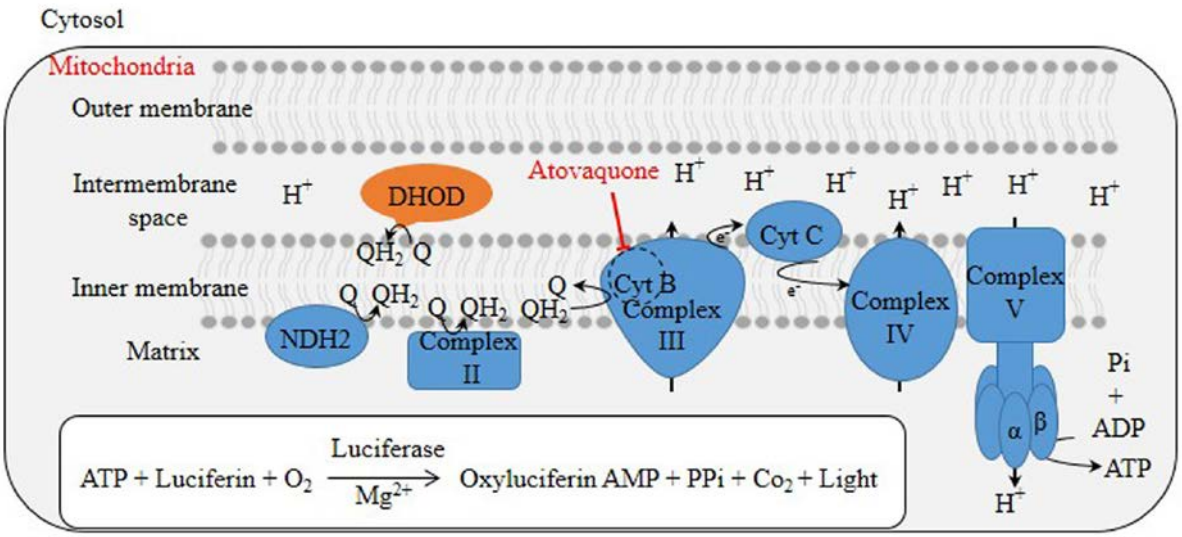


Figure-12: Schematic representation of mETC in *T. gondii.* Atovaquone inhibits Cyt b in complex III and blocks electron transfer downstream from ubiquinone, resulting in inhibition of ATP production by oxidative phosphorylation. White box with black border in mitochondria indicate the luciferase assay used to determine the intracellular ATP production in *T. gondii.*

Identification of monomer and dimer forms of *T. gondii* F-ATPase by Blue Native PAGE (BNP) analysis: The fully assembled F-ATPase is known to exist in dimeric and monomeric

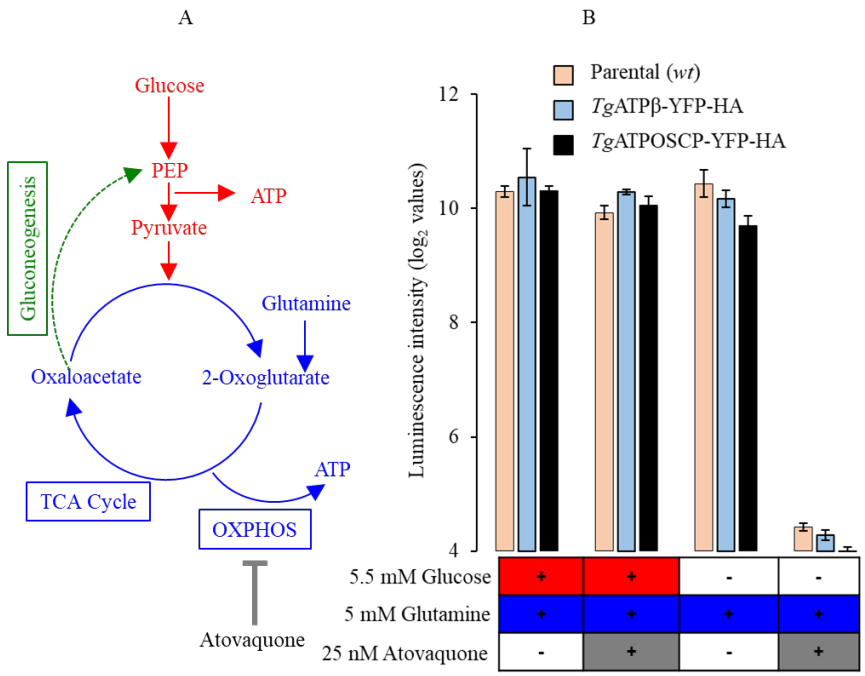


Figure-13: Nutrient metabolism in *T. gondii* (A) and functional validation of F-ATPase in parental (*wt*), *Tg*ATPβ-YFP-HA and *Tg*ATPOSCP-YFP-HA expressing transgenic parasites (B). *T. gondii* can utilize both glucose and glutamine as carbon source to produce ATP by glycolysis and oxidative phosphorylation. The later is selectively inhibited by the antiparasitic drug atovaquone (A). Comparison of intracellular ATP production in all three strains exhibited similar response to atovaquone treatment in the presence and absence of glucose, indicating that the YFP-HA tag had no effect on mitochondrial ATP synthesis. (+/-) indicate presence and absence of glucose and atovaquone in respective growth conditions. Table color shading: Red - ATP produced by glycolysis as well as by oxidative phosphorylation; blue- ATP produced by oxidative phosphorylation only; gray- oxidative phosphorylation inhibited by atovaquone (B).

forms on the inner mitochondrial membrane. The dimeric form of F-ATPase is known to influence the characteristic cristae formation of the inner mitochondrial membrane [80,82-84]. Dimeric and monomeric form of F-ATPase has been observed in the free living ciliate *T. thermophila* and the apicomplexan parasite *P. falciparum* [97, 99]. To find out whether the *T. gondii* enzyme can assemble into dimeric and monomeric form, BNP analysis was carried out on detergent-solubilized mitochondrial preparations obtained from transgenic parasite expressing YFP-HA tagged F-ATPase subunits. After resolving the samples on a 3% - 12%

gradient BNP gel capable of resolving a molecular weight range between 20 kDa to approximately 1,200 kDa (Figure-14 A lane M), Coomassie staining revealed prominent bands at the sizes corresponding to dimeric and monomeric forms of the enzyme (Figure-14 A lane A). This was further confirmed by native western blotting after BNP separation using α-HA antibodies (Figure-14 A lane B), which indicated the presence of both dimeric and monomeric forms of the enzyme, with the dimeric form being more abundant than the monomeric form.

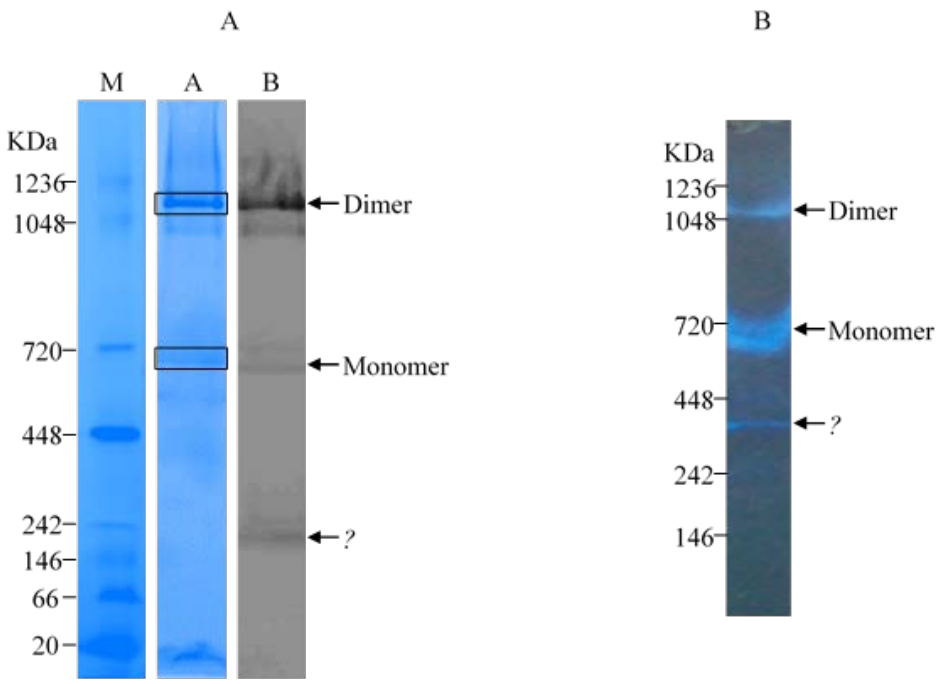


Figure-14: Identifying the dimeric and monomeric forms of *T. gondii* F-ATPase by BNP analysis. (A) Mitochondria lysates prepared from tachyzoites stage transgenic parasites expressing *Tg*ATPβ-YFP-HA protein was separated by BNP. Lane M, native molecular weight markers; lane A, Coomassie blue staining of BNP gel; lane B, Western blotting of BNP gel using α-HA antibodies. The dimer and monomer forms are indicated by arrows. The boxed regions in lane A correspond to the excised gel pieces, which were processed for LC-MS/MS analysis. (B) In-gel ATPase activity assays following BNP separation confirms that the dimer and monomer forms of F-ATPase are functionally intact.

Next, in-gel activity assays were carried out to detect the ATPase activity associated with the isolated native enzyme. Results from these assays confirmed that the ATPase activity was intact, for both dimeric and monomeric forms of the enzyme, after BNP separation (Figure-

14 B). Although the dimeric form of the enzyme was more abundant, its ATPase activity was less than that of the monomeric form. This is supported by the fact that the F_1 ATPase activity of dimeric form of enzyme is inhibited *in vivo* to minimize the risk of ATP hydrolysis and favor ATP synthesis [71, 120]. Since the in-gel activity assay is an ATP hydrolysis assay, the dimeric form has less of this activity. Based on BNP mobility, the size of the dimeric form of the enzyme was deduced to be 1-1.2 MDa and that of the monomeric form to be approximately 600 kDa. This is in agreement with what has been reported previously for *P. falciparum* [99], *S. cerevisiae* [85], and bovine [77] enzymes and suggests that a full complement of F_O subunits is present in *T. gondii* F-ATPase. Further, the regions in the BNP corresponding to the dimeric and monomeric forms of the enzyme were excised and processed for identifying the proteins in the gel band by LC-MS analysis

Identification of *T. gondii* F-ATPase associated proteins (ASAPs) by liquid chromatography tandem mass spectrometry (LC-MS/MS) proteomics: A standard protocol was followed for in-gel LC-MS/MS analysis of the gel bands corresponding to the dimeric and monomeric forms of the enzyme identified by BNP. This was done on samples prepared from both *Tg*ATPβ-YFP-HA and *Tg*ATPOSCP-YFP-HA expressing transgenic parasite lines. Since the dimeric form of the enzyme was more abundant than the monomeric form after BNP separation, LC-MS/MS analysis of the gel band corresponding to the dimeric form of the enzyme yielded better success. A total of 96 proteins were identified with high confidence from consensus data obtained from multiple in-gel LC-MS/ MS analysis. Proteins corresponding to all F_1 subunits and F_O OSCP were detected, while F_O subunit c was not detected in any of the experiments. In addition, many proteins of unknown function were also detected. However, due to co-migration of several nonspecific proteins, such as myosin, in the area corresponding to the gel band processed for LC-MS/MS analysis, it is likely that some of these proteins of unknown function are not *bona fide* subunits of the F-ATPase. A full list of

the proteins identified from these samples, along with the details of the peptides detected, is available online in the following link- (https://doi.org/10.1371/ journal.pbio.2006128.s006).

In order to specifically identify the protein subunits of F-ATPase, the native F- type ATP synthase complex was purified from transgenic parasites expressing YFP-HA tagged ATP synthase subunits by chromatographic separation (Figure-15 A and B) and Immunoprecipitation (IP) using the α-HA antibody. In contrast to BNP analysis, after size exclusion chromatography, the dimeric form of the enzyme was less abundant than the monomeric form (Figure-15B). This is likely due to stability issues with the dimeric from, which might progressively fall apart to monomeric forms during the process of chromatographic separation. The fractions corresponding to the dimeric and monomeric forms of the enzyme were pooled and concentrated separately before processing for LC-MS/MS analysis. A total of 64 proteins were detected (Figure-16) from the fractions corresponding to the monomeric form of the enzyme, while the data from dimeric from of enzymes is not reliable, due to very low protein concentration. LC-MS/MS analysis on the enzyme enriched by IP resulted in the identification of a total of 29 proteins, of which 28 were also detected from either BNP or chromatography samples (Figure-16). Details for the peptides detected from chromatography and IP samples are available online on following link- (https://doi.org/10.1371/journal.pbio.2006128.s006).

It is notable that 19 proteins, including all F_1 subunits and F_O OSCP of F-ATPase, were detected in samples prepared by all three methods. Out of the 29 proteins detected in IP, only 3 were found to be nonspecific proteins and not related to F-ATPase, based on the protein domain annotation. Out of the remaining 26 proteins, 5 were known F_1 subunits (α, β, δ, ε, γ), 1 was F_O OSCP, and the remaining 20 were proteins of unknown function. It should be noted that we were unable to identify F_O subunit c in any of our LC-MS/MS analyses, probably owing to its highly hydrophobic nature. Based on the consensus of proteins identified by the three

different approaches, the F-ATPase from *T. gondii* appears to be comprised of at least 27 protein subunits. We have coined the term "ATP synthase-associated proteins" (or "ASAPs") to refer to the 20 novel subunit components of the enzyme identified in this study. A complete list of all *T. gondii* F-ATPase subunits identified by mass spectrometry, along with their annotation from ToxoDB are given in (Table -2).

Identification of ATP synthase F_O subunits a, b, and d in *T. gondii* As shown in (Figure-7) in addition to subunits c and OSCP, the F_O sector consists of at least 6 other subunits in yeast, mammalian, and plant enzymes. Subunit a is essential for proton conductance, which it facilitates along with subunit c [63-65]. Subunit b and d forms the core of the stator structure, which is essential for holding the catalytic α3β3 structure in place during the rotary motion of the central stalk [66-67]. Our interest was to find out which of the ASAPs correspond to these three subunits in *T. gondii* F-ATPase. Since none of the ASAPs had any sequence identity to the yeast F_O a, b, and d subunits, conserved structure-based identification using pairwise comparison of profile hidden Markov models as implemented in HHPred was used for identification [121,122]. For this analysis, the corresponding tool available from web-based MPI bioinformatics toolkit (toolkit.tuebingen.mpg. de) was used. Each of the ASAPs was analyzed using this tool and this resulted in the successful identification of F_O subunits a, b, and d based on previously known structures for these proteins from other species. The ASAP TGME49_310360 was identified as F_O subunit a (Figure-17 A), and it was possible to generate sequence alignments for the C-terminal domain of this protein, which showed the conservation of the key amino acid residues arginine (important for proton translocation) and glutamine (yellow highlight in Figure-17 B). Similarly, searches with TGME49_231410 and TGME49_268830 came up with hits for F_O subunits b and d, respectively (Figure-18 A and B). Sequence alignments as well as the secondary structure prediction suggested that these proteins are likely to be ATP synthase F_O subunit b and d, based on the best hits from

HHPred (Figure-19 A-D).

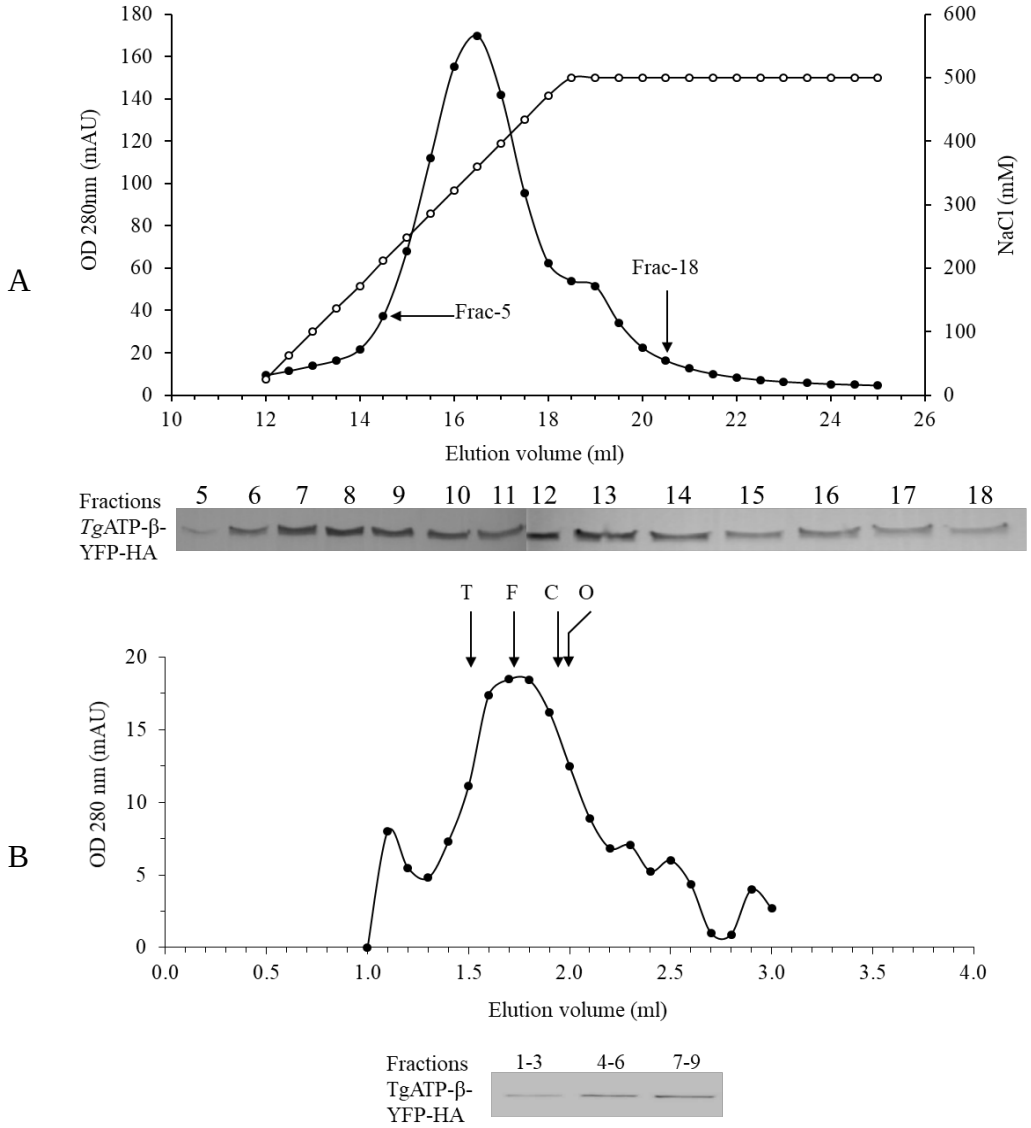


Figure-15: Partial purification of dimeric and monomeric forms of *T. gondii* F-ATPase by chromatography: (A) Ion exchange (DEAE sepharose) separation of mitochondrial lysates prepared from *Tg*ATPβ-YFP-HA expressing transgenic parasites. Absorbance at 280 nm (filled circles) and NaCl concentration (open circles) are plotted for each fraction. Fractions 5 and 18 are marked with arrows. Bottom panel shows SDS-PAGE western blotting for fractions 5-18 to find out the elution profile of *Tg*ATPβ-YFP-HA. (B) Size exclusion profile of the pooled fractions from ion exchange chromatography. The absorbance at 280 nm is plotted for each fraction. The size exclusion column was calibrated with the following native markers: thyroglobulin ("T"-660 kDa), ferritin ("F"- 440 kDa), conalbumin ("C"-75 kDa), ovalbumin ("O"-45 kDa). Peak elution volume for each marker is

indicated by arrow. Fractions 1-3, 4-6, and 7-9 were pooled, concentrated, and subject to SDS-PAGE western blotting to detect *Tg*ATPβ-YFP-HA, as shown in bottom panel.

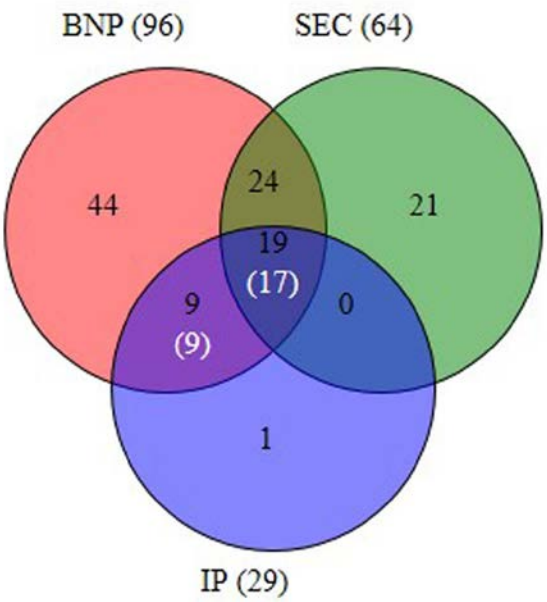


Figure-16: Identification of novel *T. gondii* F-ATPase subunits from LC-MS/MS analysis. The Venn diagram shows shared identification of proteins following BNP, SEC, and IP sample preparation. Total number of proteins identified in each technique is given within brackets outside the Venn diagram. The numbers shown in white font and within brackets are the final set of proteins assigned as subunit components of *T. gondii* F-ATPase. BNP data are a combination of experiments done with both *Tg*ATP-β-YFP-HA and *Tg*ATPOSCP-YFP-HA expressing transgenic parasites. SEC and IP data are from *Tg*ATPβ-YFP-HA and *Tg*ATPOSCP-YFP-HA expressing transgenic parasites, respectively.

Mitochondrial localization of ASAPs and gene essentiality in *T. gondii*: The expected localization of the F-ATPase enzyme complex is the mitochondrion, and accordingly localization studies confirmed that the β and OSCP subunits of the *T. gondii* F-ATPase are localized to the parasite mitochondrion (Figure-11 B). In order to confirm that the novel subunits identified from the LC-MS/MS analysis of *T. gondii* F-ATPase are *bona fide* subunits of the enzyme complex, mitochondrial localization studies were carried out. First, the presence of mitochondria targeting signal sequence was predicted for all twenty ASAPs using the Mitoprot tool [123], which revealed the presence of mitochondrial signal sequence in 12

ASAPs. Five of the subunits including TGME49_231410 (F_O subunit b like protein), were previously shown to localize in the mitochondrion [126], and four other subunits were confirmed experimentally to localize to the parasite mitochondrion in this study (Figure-20).

S. No.	Gene ID	Product description
1	TGME49_204400	F_1alpha subunit
2	TGME49_261950	F_1beta subunit
3	TGME49_226000	F_1delta subunit
4	TGME49_314820	F_1epsilon subunit
5	**TGME49_231910**	**F_1 gamma subunit**
6	TGME49_284540	F_0 OSCP subunit
7	**TGME49_310360**	**hypothetical protein**
8	TGME49_231410	hypothetical protein
9	TGME49_268830	hypothetical protein
10	TGME49_223040	hypothetical protein
11	TGME49_247410	hypothetical protein
12	**TGME49_260180**	**hypothetical protein***
13	**TGME49_218940**	**hypothetical protein**
14	TGME49_282180	hypothetical protein
15	TGME49_285510	hypothetical protein
16	TGME49_215610	hypothetical protein
17	**TGME49_290030**	**hypothetical protein**
18	**TGME49_310180**	**hypothetical protein**
19	**TGME49_214930**	**hypothetical protein**
20	TGME49_245450	hypothetical protein
21	**TGME49_208440**	**hypothetical protein**
22	TGME49_201800	hypothetical protein
23	TGME49_225730	membrane protein
24	TGME49_263080	hypothetical protein
25	TGME49_263990	hypothetical protein
26	**TGME49_270360**	**hypothetical protein**

Table-2: List of genes encoding the ASAPs identified in this work by mass spectrometry analysis: Gene ID and product description details were taken from Toxodb.org (release 36). The entries shown in bold indicate that the protein was detected with high confidence from only BNP and IP samples. Asterisk indicates that the corresponding peptides were detected in only one of the replicate runs for SEC sample.

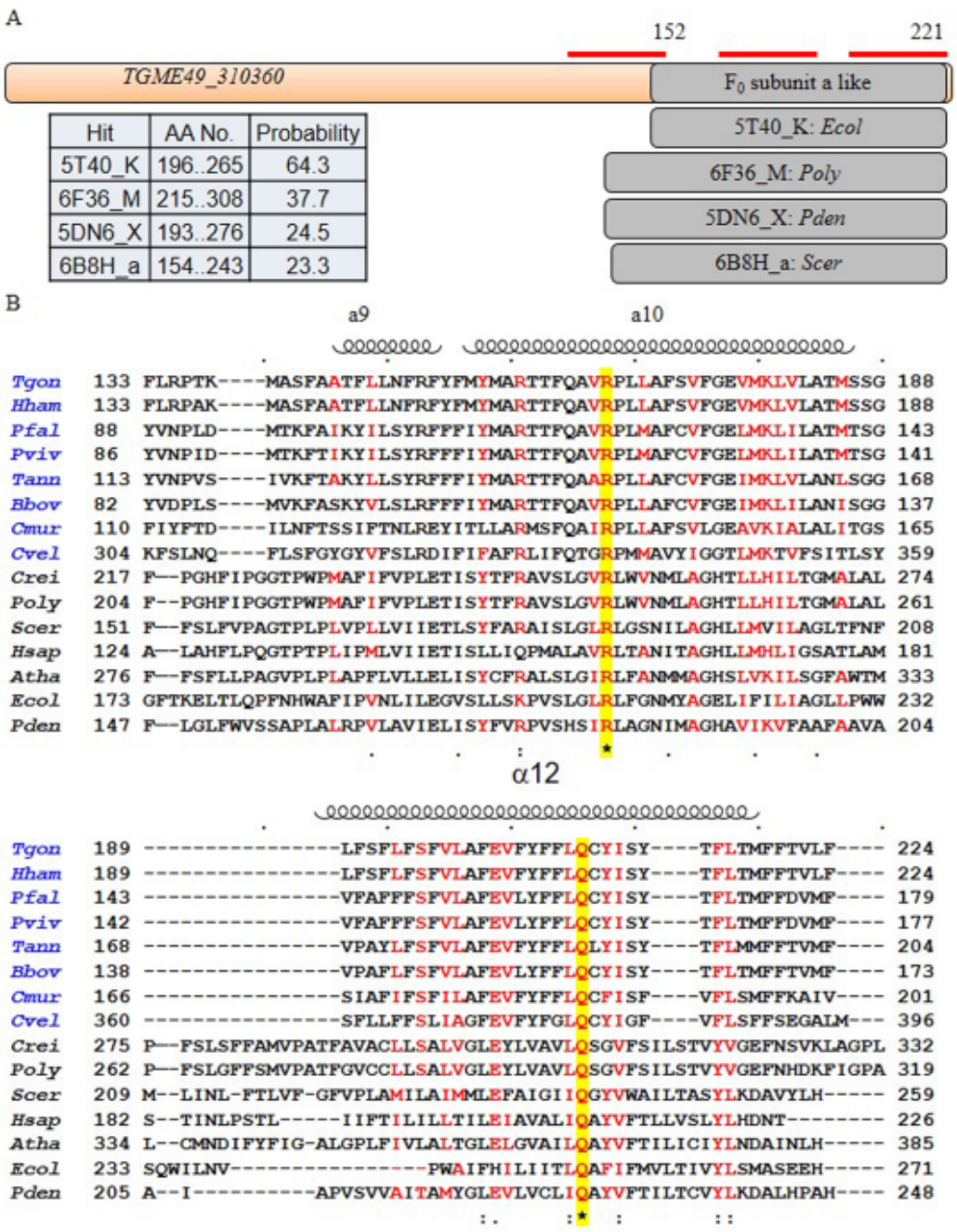
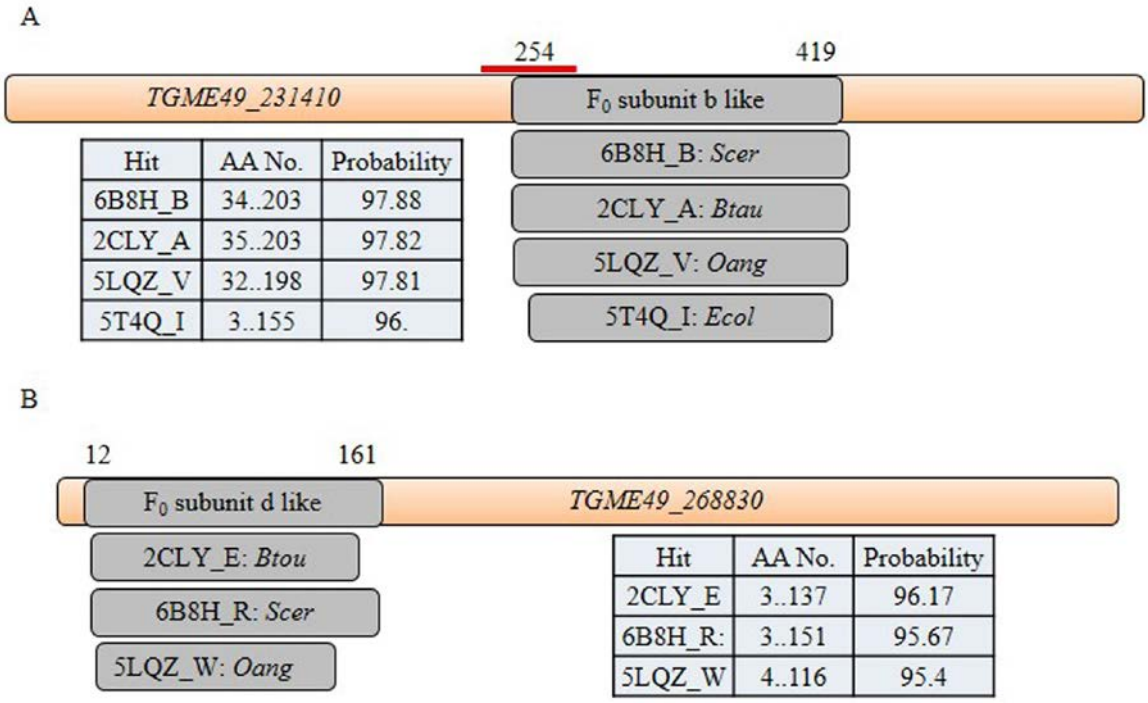


Figure-17: HHPred identification of novel F_O subunit a based on conserved structural features. (A) Representation of the pairwise sequence alignments generated by HHPred for the putative F_O subunit a from *T. gondii* and 7 other F_O subunit a proteins for which the structure is known. The table provides details of the amino acid length and a probability score for the prediction from the hit alignments. The red lines indicate the 3 transmembrane domains present in the *T. gondii* protein. (B) Protein sequence alignments and secondary structure information were made using the Clustal Omega [124] and ESPript [125] software, and only the C-terminal portion of respective proteins is shown. The names of alveolate species are highlighted in blue. For the non alveolate species included in the alignment, the F_O subunit a is either readily identified from sequence (*Scer*, *Hsap*,

Atha, *Ecol*, and *Pten*) or has been experimentally determined (*Crei* and *Poly*). Positions with similar amino acids are highlighted in red in the alignment. The arginine and glutamine residues, highlighted in yellow, are conserved in all species and important for function. The helices (α9, α10, and α12) shown are from the structure of *Poly* F-ATPase. Species names: *Tgon*, *T. gondii*; *Hham*, *Hammondia hammondi*; *Pfal*, *P. falciparum*; *Pviv*, *P. vivax*; *Tann*, *Theileria annulata*; *Bbov*, *B. bovis*; *Cmur*, *C. muris*; *Cvel*, *C. velia*; *Crei*, *C. reinhardtii*; *Poly*, *Polytomella* sp.; *Scer*, *S. cerevisiae*; *Hsap*, *Homo sapiens*; *Atha*, *A. thaliana*; *Ecol*, *Escherichia coli*; *Pden*, *Paracoccus denitrificans*.



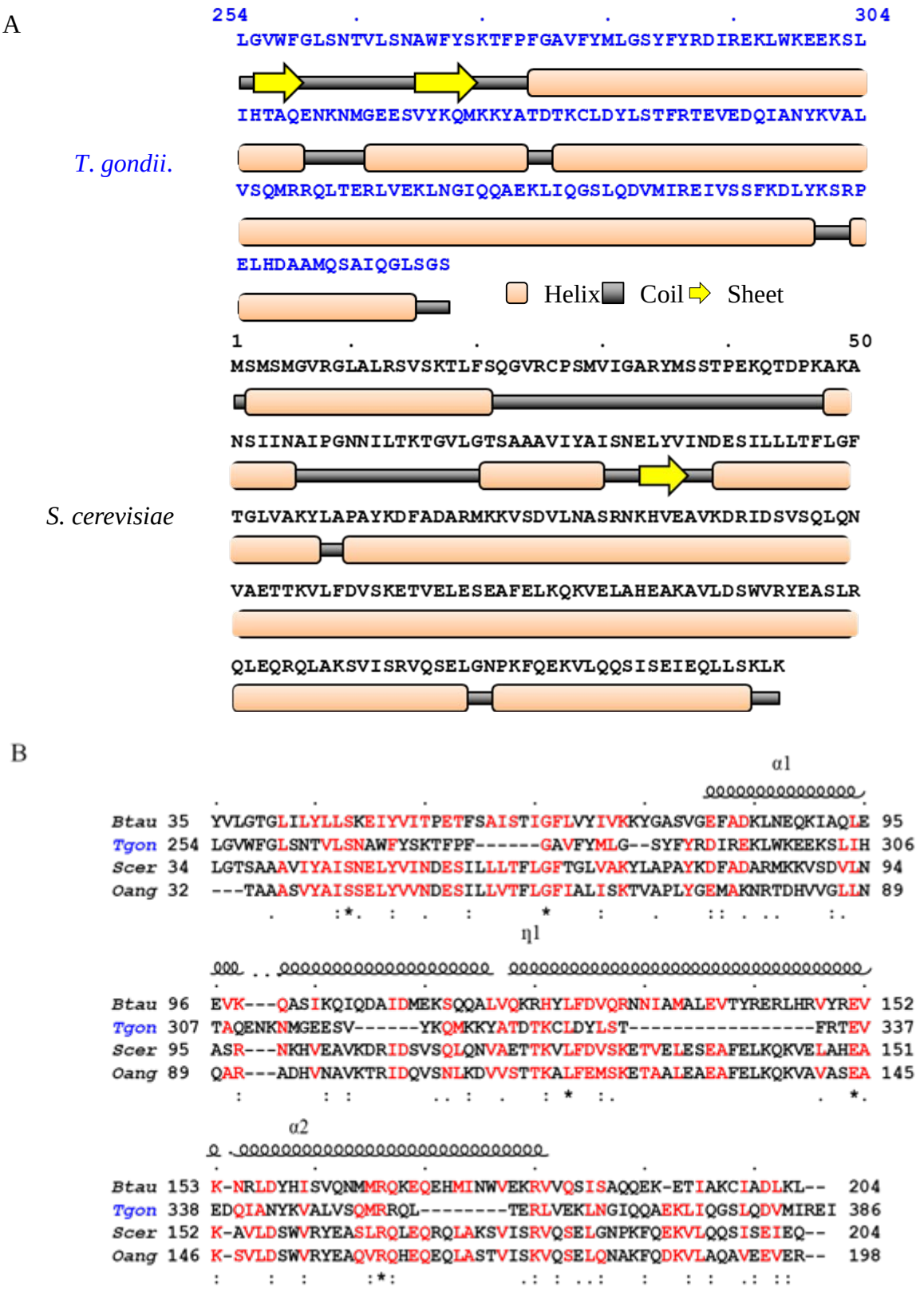
Hit	AA No.	Probability
6B8H_B	34..203	97.88
2CLY_A	35..203	97.82
5LQZ_V	32..198	97.81
5T4Q_I	3..155	96.

Hit	AA No.	Probability
2CLY_E	3..137	96.17
6B8H_R:	3..151	95.67
5LQZ_W	4..116	95.4

Figure-18: Homology detection structure prediction by HHPred and its comparison for Fo subunit b and d like protein from *T. gondii*. (A) TGME49_231410 (B) TGME49_268830. The table provides the details of the amino acid length and a probability score for the prediction from the hit alignments.

Due to the importance of the F-ATPase enzyme, it is reasonable to expect that the enzyme would be essential for survival of *T. gondii*. Interestingly, it was reported that the β subunit of the enzyme is not essential for blood stage parasite growth, but essential for sexual development of parasite, in the rodent malaria parasite *P. berghei* [100]. Our multiple attempts to disrupt the *T. gondii* genes encoding the β and c subunits of the F-ATPase using the CRISPR/Cas9 technique (Figure-22) proved futile, indicating the essentiality of these proteins

for *T. gondii* survival. Interestingly, all ASAPs identified in this study, except ASAP-19 and 20, were shown to be essential proteins in a genome-wide CRISPR/Cas9 screen in another study [126].



C

T. gondii.

12 62

SRLLPRRDSWMSLAPFVAPNNAAAWRKLRDGAQEVQTVIERQSTPGKPQQ

IDWAKWESQIAHKDILNCLKTFYTNQVQILDRALGALETAKTPAPCEGAE

KGWALFDAALSACAKSVEKSEELLSNGARALWVSCSNPPVWKVNTNEWLD

SDQYWQAFVEK

Helix Coil Sheet

S. cerevisiae

1 50

MSLAKSAANKLDWAKVISSLRITGSTATQLSSFKKRNDEARRQLLELQSQ

PTEVDFSHYRSVLKNTSVIDKIESYVKQYKPVKIDASKQLQVIESFEKHA

MTNAKETESLVSKELKDLQSTLDNIQSARPFDELTVDDLTKIKPEIDAKV

EEMVKKGKWDVPGYKDRFGNLNVM

D

```
                    α1          η1            α2                α3
                 ooooo      oooooooooooooooooo  ooooo        oooooo
           .         .         .         .         .         .         .
Btau 3     RKLALKTIDWVAFGEIIP--RNQKAVANSLKSWNETLT--SRLATLPEKPPAIDWAYYKA 54
Tgon 12    SRLLPRRDSWMSLAPFVA--PNNAAAWRKLRDGAQEVQTVIERQSTPGKPQQIDWAKWES 69
Scer 3     AKSAANKLDWAKVISSLRITGSTATQLSSFKKRNDEAR--RQLLELQSQPTEVDFSHYRS 56
Oang 4     AKSTANKLDWTKIVSKLGLSGQTAAALTSFKKRNDEAK--RILFELKQQPSNVDFAFYKS 55
            :   .  .*  .   :    .  :   .::.  :             :*  :*:: :.:

                    α4                                    α5
           ρ     oooooooooo                        oooooooooooooooo
           .         .         .         .         .         .         .
Btau 55    NVAKAGLVDDFEKKFNAL--------------KVPIPEDKYTAQVDAEEKEDVKSCAEFL 200
Tgon 70    QIAHKDILNCLKTFYTNQVQILDRALGALETAKTPAPCEGAEKGWALFD-AALSACAKSV 228
Scer 57    VLKNTSVIDKIESYVKQY-----------KPVKIDASKQ--LQVIESFEKHAMTNAKETE 203
Oang 56    TLKNTAIVDKIQSDVSKF-----------TPSKANLSKQ--LNLIESFEAKALENAKETE 102
            : :  ::: ::.  .               *     :         :   :  . :

           oooooooooooooooooo
           .         .         .         .         .
Btau 201   TQSKTRIQEYEKELEKMRNIIPFDQMTIEDLNE----------- 137
Tgon 229   EKSEELLSNGARALWVSCSNPPVWKVNTNEWLDSDQYWQAFVEK 172
Scer 204   SLVSKELKDLQSTLDNIQSARPFDELTVDDLTKIKPEIDAKVEE 151
Oang 103   SVVLAELT------------------------------------ 116
                :
```

Figure-19: The predicted secondary structure features and multiple sequence alignment. Secondary structure prediction of putative F-ATPase subunit b (TGME49_231410) and d (TGME49_268830); and its comparison to *Scer* counterpart (A and C) respectively. Multiple sequence alignment of putative F_O subunit b and d protein from *T. gondii* with *Btau*, *Scer*, and *Oang* counterparts using Culstal Omega [48] and ESPript [49]. The

helices shown are derived from the crystal structure of *Btau* protein. Species names: *Btau*, *B. taurus*; *Tgon*, *T. gondii*; *Scer*, *S. cerevisiae*; *Oang*, *Ogataea angusta*.

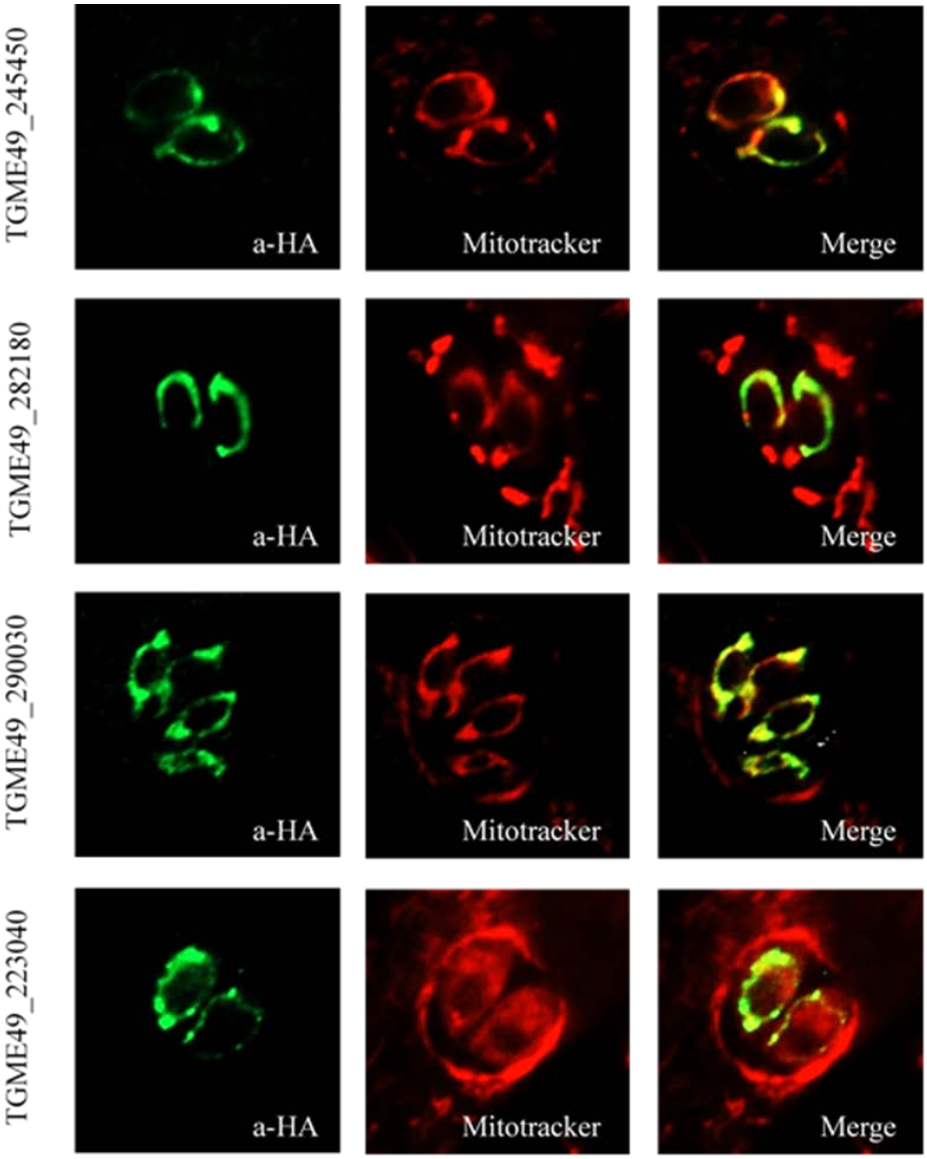


Figure-20: Confirming mitochondrial localization for selected ASAPs. The cDNA for TGME49_245450, TGME49_282180, TGME49_290030, and TGME49_223040 were constitutively expressed from a plasmid under the β-tubulin promoter as C-terminal HA-tagged proteins in tachyzoites stage parasites. Mitochondrial localization was confirmed by colocalization with Mitotracker. Immunostaining was carried out as described in the Methods section.

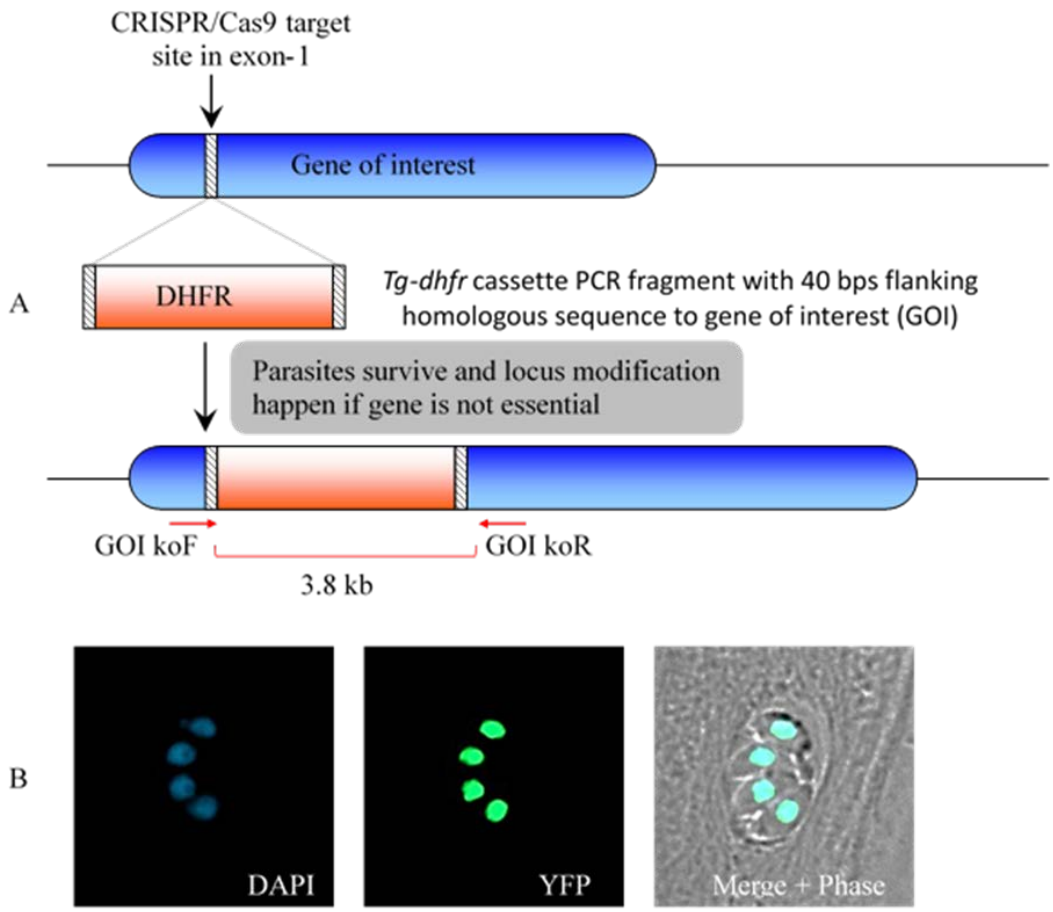
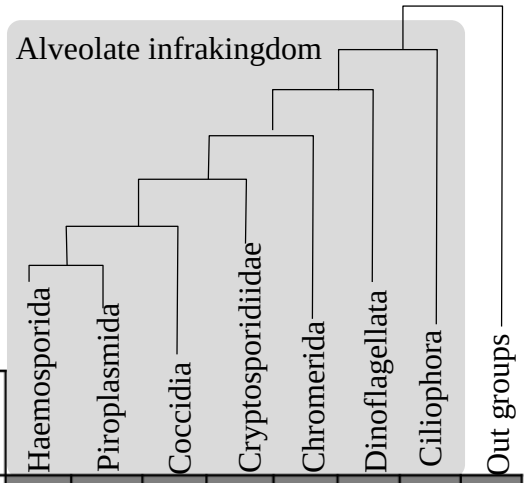


Figure-21: Schematic representation for knockout strategy in *T. gondii* and expression of CRISPR/Cas9. (A) The schematic representation of gene knockout strategy using CRISPR/Cas9 technique. The guide RNA sequence (Table-1) was designed to target exon 1 of gene of interest (TGME49_261950 - *Tgatp-β* and TGME49_249720 - *Tgatpc*). (B) Expression of Cas9 protein and its localization to parasite nucleus.

Phylogenetic analysis reveals conserved F-ATPase subunit composition in 3 major alveolate taxons: From ortholog identification attempts carried out in this work, and from previous studies [26-27], it was apparent that the F-ATPase subunits missing in *T. gondii* were also missing from all other alveolate species (Figure-8). In fact, the novel subunit components identified from *Tetrahymena* (Alveolata; Ciliophora) enzyme were found to be unique to ciliates and not conserved in other alveolate organisms [97]. Therefore, it was of interest to study the phyletic profile of the novel ASAPs identified in this study in order to find out whether these proteins are unique to *T. gondii* F-ATPase are present in other alveolates also. Ortholog identification revealed that many of the ASAPs were conserved across 3 major



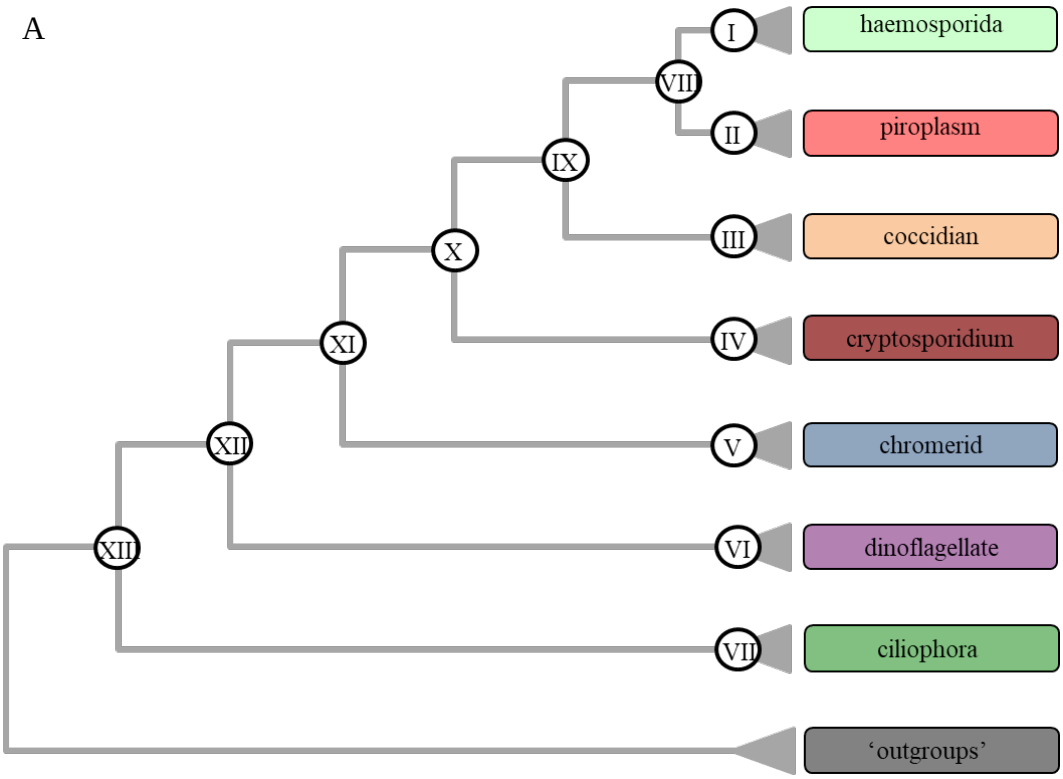
Gene ID	Subunit Annotation	CRISPR Pheno Score*	Mito. location
TGME49_204400	F_1 alpha	-3.84	Mito#
TGME49_261950	F_1 beta	-4.84	**Mito#**
TGME49_226000	F_1 delta	-4.57	Mito#
TGME49_314820	F_1 epsilon	-3.21	Mito#
TGME49_231910	F_1 gamma	-3.94	Mito#
TGME49_249720	F_0 c	-2.98	Mito
TGME49_284540	OSCP	-3.94	**Mito#**
TGME49_310360	ASAP-1 (F_0 a?)	-4.49	Mito
TGME49_231410	ASAP-2 (F_0 b?)	-5.37	Mito*#
TGME49_268830	ASAP-3 (F_0 d?)	-3.10	Mito#
TGME49_223040	ASAP-4	-4.49	**N#**
TGME49_247410	ASAP-5	-3.90	N*#
TGME49_260180	ASAP-6	-4.07	Mito*#
TGME49_218940	ASAP-7	-3.92	N*#
TGME49_282180	ASAP-8	-2.86	**Mito#**
TGME49_285510	ASAP-9	-1.87	N#
TGME49_215610	ASAP-10	-3.55	Mito*#
TGME49_290030	ASAP-11	-3.88	**Mito#**
TGME49_310180	ASAP-12	-3.40	N#
TGME49_214930	ASAP-13	-1.37	N#
TGME49_245450	ASAP-14	-2.95	**Mito#**
TGME49_208440	ASAP-15	-3.54	Mito#
TGME49_201800	ASAP-16	-4.01	Mito#
TGME49_225730	ASAP-17	-3.65	Mito#
TGME49_263080	ASAP-18	-3.10	Mito#
TGME49_263990	ASAP-19	0.22	N#
TGME49_270360	ASAP-20	0.32	N#

Figure-22: Phylogenetic profile of the alveolate infrakingdom for all *T. gondii* F-ATPase subunits. Alveolate clades are highlighted with a gray background, and their expected phylogenetic relationship is indicated by a tree structure above. Gray and white boxes indicate the presence and absence of the corresponding ortholog,

respectively. The hatched boxes represent the presence of the ortholog in *C. muris* only and absence in *C. parvum* and *C. hominis*. The table on the left lists the gene ID for all ASAPs, along with their annotation, essentiality phenotypes (phenotype score from CRISPR/Cas9 knockout study [126]), and protein localization. The key for localization annotation is as follows; Mito- mitochondria targeting signal sequence present, N- no mitochondrial signal sequence present, Bold- localized in this study, * localized in previous studies, # mitochondrial localization confirmed in Hyper LOPIT study (unpublished evidence Dr. Ross Waller Cambridge University, UK).

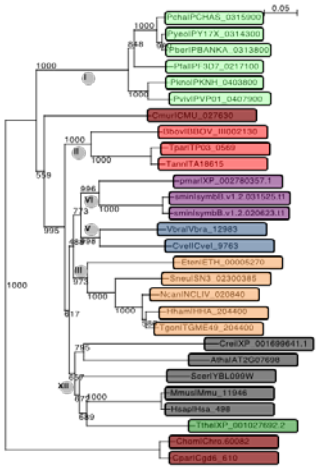
alveolate taxons - Apicomplexa, Chromerida, and Dinoflagellata (Figure-22). A list of all orthologs identified from selected species belonging to these taxons is available on this link- (https://doi.org/10.1371/jozurnal.pbio.2006128.s007). Out of the 20 ASAPs, 15 were conserved in all apicomplexan clades, except in the case of *Cryptosporidium*, in which only *C. muris* contained orthologs for 10 of these proteins. A few ASAPs were unique to the Coccidian clade. More importantly, all ASAPs, except one, were conserved in Chromerida, and at least 9 and 12 ASAPs were also conserved in *Symbiodinium* and *Perkinsus*, respectively. To obtain further insights on the evolutionary origin of these proteins, neighbor- joining phylogenetic trees were generated for ortholog sequences of all F_1 and F_O subunits from representative species of Haemosporida, Piroplasmida, Coccidia, Cryptosporidiidae, Chromerida, Dinoflagellata, and Ciliophora (Figure-23 A and B). *S. cerevisiae, H. sapiens, M. musculus, C. reinhardtii* and *A. thaliana* species were included as outgroups while constructing the phylogenetic trees for the highly conserved F_1 subunits. Even though the topology of most trees did not reflect the expected evolutionary relationship between the included species, monophyletic grouping was observed in general at the taxon level, and importantly, this was evident for the conserved F_1/F_O subunits, as well as the novel ASAPs (Figure-23). Thus, the evolutionary origin of the newly identified highly divergent ASAPs in Apicomplexa, Chromerida, and Dinoflagellata clades appears to be ancient.

A



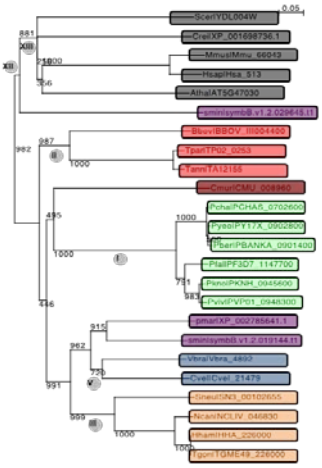
B

TGME49_204400
F_1 alpha



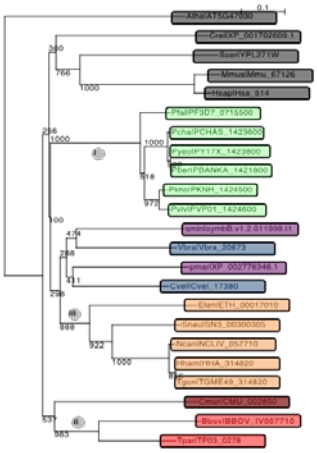
TGME49_261950
F_1 beta

TGME49_226000
F_1 delta

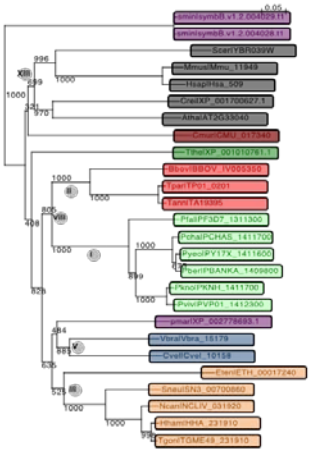


B *(continued)*

TGME49_314820
F_1 epsilon

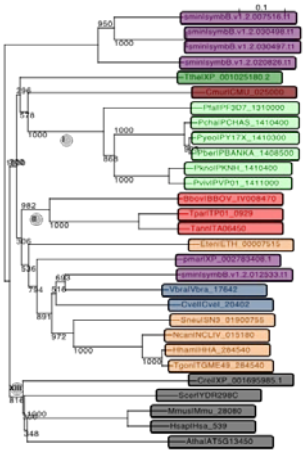


TGME49_231910
F_1 gamma

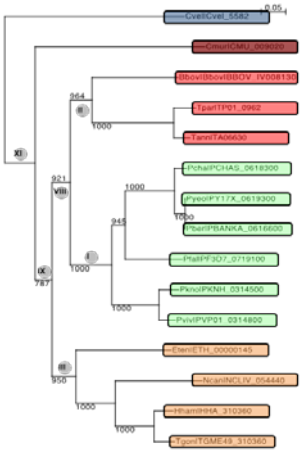


TGME49_249720
F_0 c

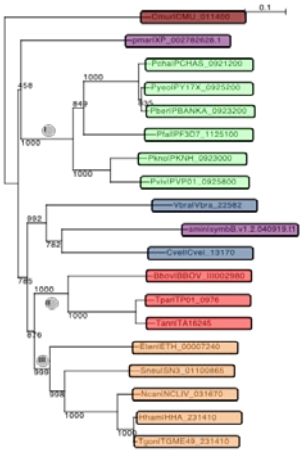
TGME49_284540
OSCP



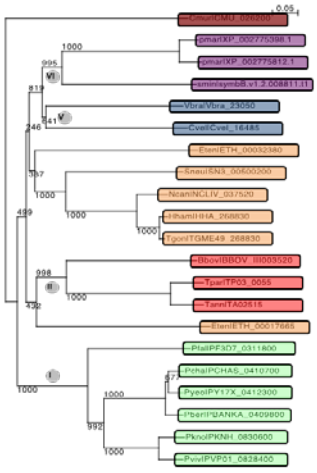
TGME49_310360
ASAP-1 (F_0 a?)



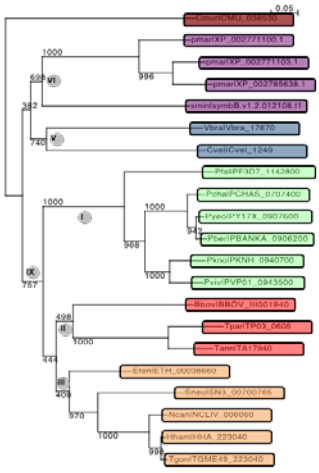
TGME49_231410
ASAP-2 (F_0 b?)



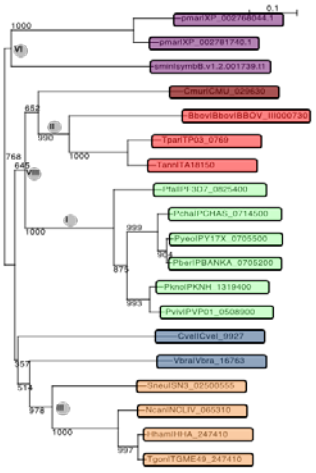
TGME49_268830
ASAP-3 (F_0 d?)



TGME49_223040
ASAP-4

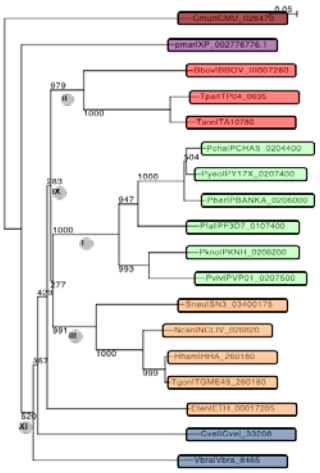


TGME49_247410
ASAP-5

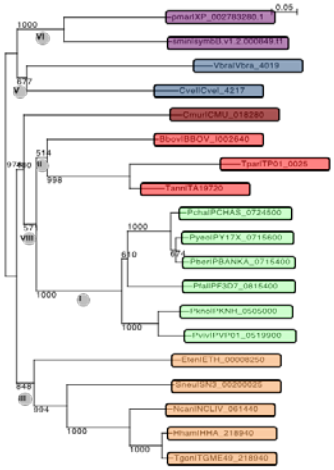


B *(continued)*

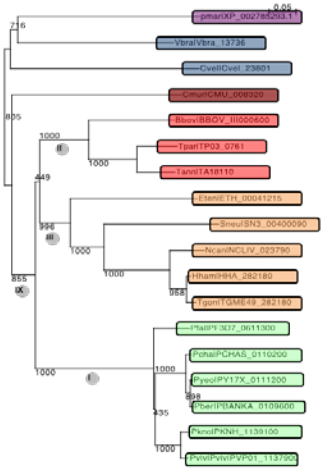
TGME49_260180
ASAP-6



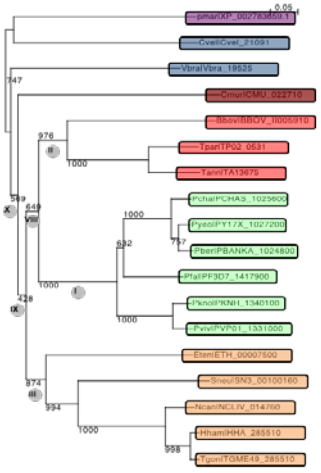
TGME49_218940
ASAP-7



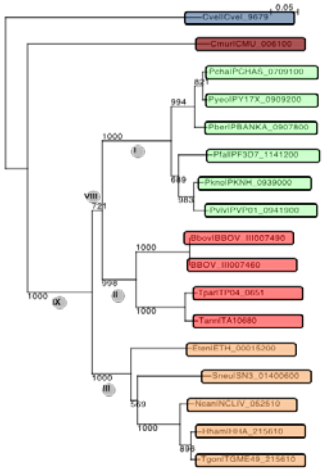
TGME49_282180
ASAP-8



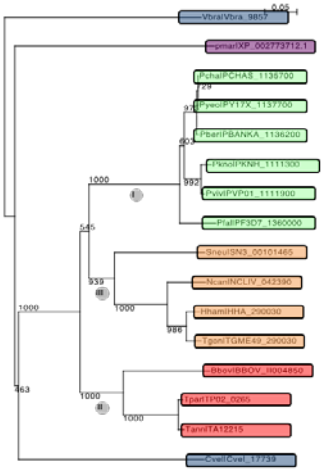
TGME49_285510
ASAP-9



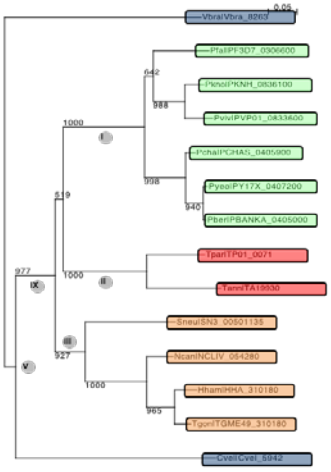
TGME49_215610
ASAP-10



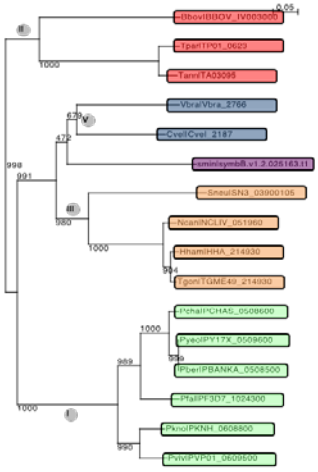
TGME49_290030
ASAP-11



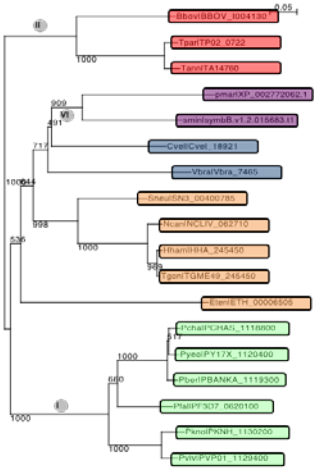
TGME49_310180
ASAP-12



TGME49_214930
ASAP-13



TGME49_245450
ASAP-14



B *(continued)*

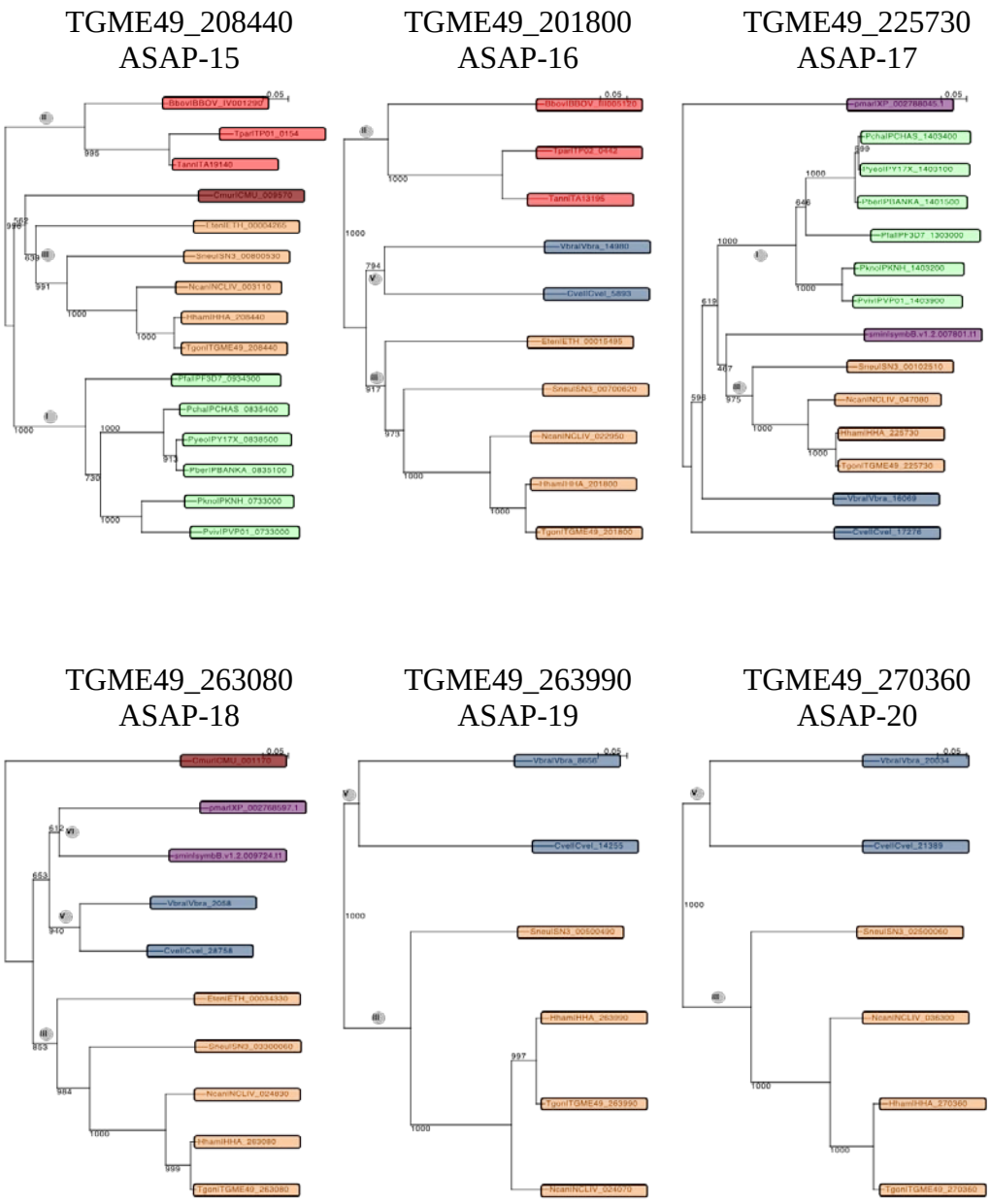


Figure-23: Neighbor-Joining phylogenetic trees showing the evolutionary relationship for all orthologs of ASAPs identified from representative species of Haemosporidia, Piroplasma, Coccidia, Cryptosporidiidae, Chromerida, Dinoflagellates, Ciliophora, and outgroups. (A) Cladogram representation of the expected phylogenetic relation for the selected species as previously published [27]. I-XIII represent the nodes on the cladogram and are used to denote the monophyletic of taxon-specific sequences in the individual trees for each ASAP ortholog set shown in (B). The taxon color coding is same in (A) and (B). The numbers in (B) indicate bootstrapping support. ASAP- ATP synthase associated protein.

Gene coexpression analysis of ASAPs from *T. gondii* and *P. falciparum:* In order to obtain further independent evidence for ASAPs as *bona fide* subunits of F-ATPase, transcriptome coexpression correlation analysis was carried out using publicly available gene expression datasets for *T. gondii* and *P. falciparum* [113-114]. This was carried out based on the assumption that the ASAPs interacting with one other would show significant pairwise correlation profiles in their expression levels. Strikingly, very good correlation in gene coexpression profiles can be detected among the F-ATPase subunits, in comparison to coexpression with other unrelated gene pairs, in both *T. gondii* and *P. falciparum* transcriptome datasets (Figure-24). This finding further supports the fact that the novel ASAPs are indeed coexpressed and are likely *bona fide* subunits of the unusual F-ATPase enzyme from *T. gondii* and in *P. falciparum*.

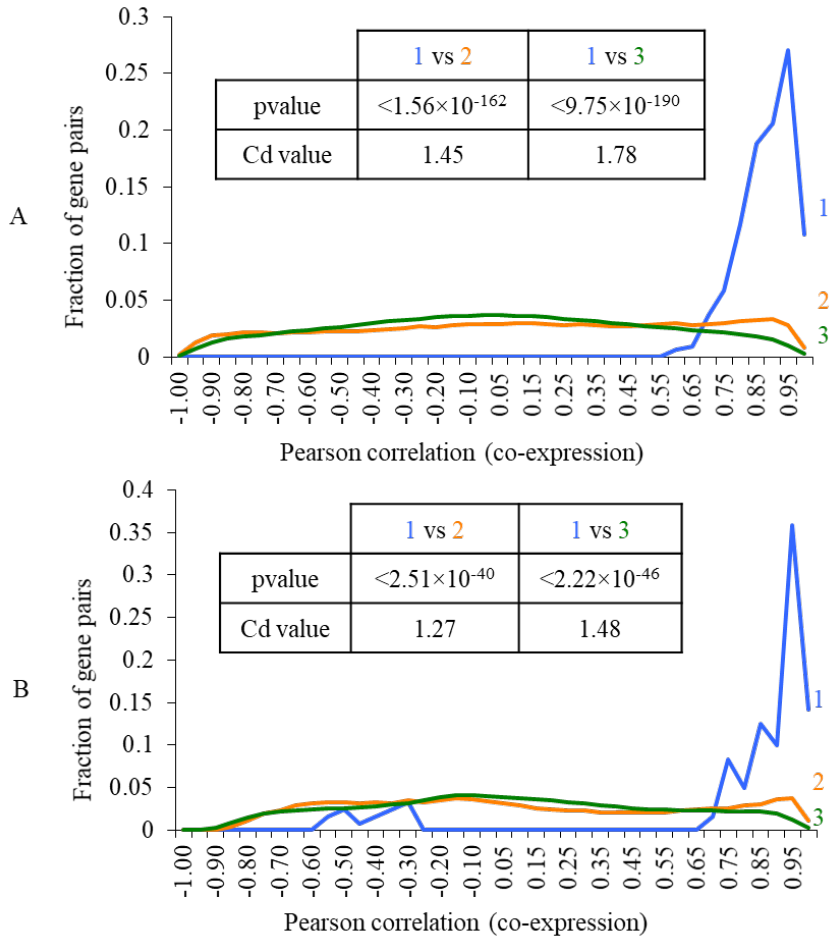


Figure-24: Gene coexpression analysis for *T. gondii* and *P. falciparum* F-ATPase subunits. The distribution of coexpression values as measured by the Pearson correlation coefficients are plotted for the three gene pairs categories, as shown for *T. gondii* (A) and *P. falciparum* (B). The statistical support from MWU and Cd values are shown in the table within each plot. Blue (1), coexpression correlation between ATP synthase genes; Orange (2), coexpression correlation between ATP synthase genes and non-ATP synthase genes; Green (3), coexpression correlation between non-ATP synthase genes. MWU, Mann-Whitney U *p*-values; Cd, Cohen's d.

Discussion

Discussion:

All life forms need energy for survival, which is mostly available in the form of ATP. In most eukaryotes, bulk cellular ATP is mainly produced by mitochondrial oxidative phosphorylation, where the multimeric F-ATPase complex facilitates the production of ATP from ADP and Pi. The energy required to drive ATP production by F-ATPase complex is harnessed from the movement of protons down the electrochemical energy gradient from the mitochondrial inter-membrane space to the mitochondrial matrix side [55-59, 61-63]. The enzyme consists of two functionally distinct regions called the F_1 and F_O sectors, which act in concert to convert the electrochemical energy into mechanical energy to facilitate ATP synthesis [55-59, 61]. The F_1 sector comprises of three copies each of α and β subunits which make the catalytic core of the enzyme complex, and single copy of each of γ, δ and ε subunits, which makeup the central stalk of the enzyme complex. The core components of the F_O sector includes the oligomeric assembly of subunit c along with single copy of subunits a, b, d, h, F6 and OSCP, as seen in yeast F-ATPase and its mammalian counterpart [59]. The central stalk bridges the catalytic core with the rotary motor formed by oligomeric subunit c [55-59, 61]. During proton translocation the rotation of F_O sector subunit c leads to conformational change in the asymmetric γ subunit of the rotating central stalk, which induces the conformational changes in the $\alpha_3\beta_3$ catalytic core and facilitates ATP synthesis. The stator structure is comprised of the catalytic core of F_1 sector, the F_O sector subunits a, along with the peripheral stalk components b, d, h F6 and OSCP and other accessory subunits [55-59, 61, 67]. The peripheral stalk of the enzyme helps in holding the $\alpha_3\beta_3$ catalytic core in place while the rotor part of the enzyme rotates [67-68]. Site directed mutagenesis and gene knockout/knockdown studies in model organisms reveals that, all subunit components of this unique enzyme are critical for proper assembly and efficient ATP synthesis.

Apart from the core subunit composition described above, other accessory proteins

associated with mitochondrial F-ATPase are required for assembly of oligomeric forms of enzymes complex, specifically the V shaped dimeric form [85]. The first evidence of presence of dimeric and oligomeric forms of F-ATPase came from freeze-fracture and deep-etching electron microscopy studies on *P. multimicronucleatum* mitochondrial membrane. This studied indicated the paired arrangement of the F_1 complex (spaced by 12 nm) on the inner mitochondrial membrane [75]. Subsequently the dimeric and oligomeric forms of mitochondrial F-ATPase was confirmed by blue native page (BNP) analysis, atomic force microscopy and cryo electron microscopy studies on detergent solubilized mitochondrial preparations from various organisms such as yeast, bovine and *Polytomella Sp.*, [83, 85,111,127]. In fact, the dimeric form of F-ATPase facilitates cristae formation in the inner mitochondrial membrane [80, 82-84] and is implicated in maintenance of the mitochondrial membrane potential [128]. However, the identity of the proteins responsible for dimer formation has not been ascertained in most species. For example, in the yeast *S. cerevisiae*, dimerization and oligomerization of F-ATPase enzyme complex is mediated by the accessory subunits e, g, and k. Genetic studies on mutant yeast lacking these subunits reveals that subunit e and g are essential for dimerization of F-ATPase [85-86]. The ortholog for these proteins are also present in mammalian species. Interestingly the yeast or mammalian F-ATPase dimer specific subunits are absent in other organisms such as free living algae, ciliates and parasitic protozoans, despite the presence of dimeric form of F-ATPase in these organisms [93-95, 97]. This indicates that the proteins responsible for dimerization may have different evolutionary origins.

The mitochondrial F-ATPase complex purified by BNP and size exclusion chromatography studies from yeast revealed that the average molecular size of the dimeric and monomeric form of the enzyme is around 1 MDa and between 500–600 kDa, respectively [85]. Considering the complex structure and assembly of F-ATPase enzymes, subunit composition

studies in yeast and mammals revealed that the complex is made up of around 20 different subunits [91-92]. Even though the structure and function of F-ATPase is conserved across all eukaryotes, the subunit composition can vary among them. In many taxa, including in unicellular eukaryotes such as free living algae and parasitic protista, only some of the F-ATPase subunits are detectable by sequence, i.e., all F_1 sector subunits and only subunit c and OSCP from the F_O sector, as per the data available from the (KEGG) database (kegg.jp). It appears that the remaining subunits of the F_O sector are poorly conserved and are not readily identified from sequence similarity.

In case of apicomplexan parasites also a similar situation prevails and accordingly, other than subunit c and OSCP other F_O sector are missing [26-27, 98-99]. By extending the analysis to other alveolate phyla, such as ciliata, dinoflagellata and chromerida, it became apparent that the F_O sector subunits missing in apicomplexa were also missing in these organisms. A detailed study on the subunit composition of F-ATPase purified from detergent solubilized mitochondrial sample obtained from the ciliate *T. thermophila*, an alveolate species, resulted in the identification of 13 novel proteins [97]. However, unexpectedly the novel subunits identified for *T. thermophila* F-ATPase were not conserved within alveolate infrakingdom, despite the evolutionary relatedness. Based on this finding, it was expected that maybe a unique set of subunit components maybe present in *T. gondii* and other apicomplexan F-ATPase.

The molecular size of well-characterized F-ATPase from yeast and bovine species is around 500-600 kDa for the monomeric form and ~1 MDa for the dimeric form. The fact that the *P. falciparum* F-ATPase assembles into both monomer and dimer forms with similar size range [99] to the model enzymes suggested two things: one is that the parasite enzyme is assembling into intact complexes needed to be functional and the other is that, novel proteins have likely replaced the missing F_O sector subunits of the enzyme. The *P. falciparum* F-ATPase

was also found to be essential for survival of asexual blood stage parasites, even though in the rodent malaria parasite *P. berghei*, it was found to be essential only for sexual development in the mosquito [100]. Similarly studies on metabolic mutants in *T. gondii* have revealed that mitochondrial oxidative phosphorylation is an important source of ATP, especially when glucose is absent and glycolytic flux is not a source of ATP [18, 19]. Mitochondrial ATP synthesis is also likely essential for formation and maintenance of tissue cyst forms of *T. gondii* in infected hosts. Despite the importance of F-ATPase enzyme in apicomplexan parasites, not much is known about its subunit composition, structure, and function in these parasites.

This study has revealed that the F-ATPase enzyme is functional as an ATP synthase in *T. gondii* and is functionally integrated mtETC in *T. gondii*. The native enzyme purified from detergent-solubilized parasite mitochondria preparations was found to exist in both dimeric and monomeric forms, as is the case in other eukaryotic F-ATPase [99], as identified by BNP analysis. The F-ATPase from *T. gondii* was partially purified using BNP, IP and chromatography techniques, and subsequently LC/MS/MS analysis on tryptic digested samples resulted in the identification of 20 novel proteins. These proteins are referred to as ATP synthase associated proteins (ASAPs) and are listed in Table-2. Some of these proteins are probable functional equivalents of the missing subunits of F_O sector such as subunit a, b and d (which form the core components of the stator structure), while others are likely to be accessory proteins required for oligomerization and regulation of the enzyme activity. As these novel proteins had almost no sequence similarity to any of the yeast or bovine F_O sector subunits, it was not possible to identify the functional equivalents of key subunits such as a, b and d by sequence conservation. Therefore, the presence of conserved structural features for canonical F_O sector subunits a, b and d was used to identify the corresponding functional equivalents for *T. gondii* F-ATPase. This approach led to the successful identification of putative F_O subunits a (TGME49_310360), b (TGME49_231410), and d (TGME49_268830). These three subunits

have a central role in the functioning of the F-ATPase as subunit a along with subunit c involved in proton translocation process, while subunit b and d is important component of peripheral stalk.

Predicted secondary structure features of the putative subunit a revealed that this protein has 3 transmembrane domains (Figure-17A), while its counterpart from yeast and many other eukaryotes has 5-7 transmembrane domains. Nevertheless, the critical arginine residue, which is essential for proton translocation [64], appears to be conserved (Figure-17 B). Similarly, subunit b from yeast and many other eukaryotes has the hallmark structural feature of an extended α-helix, which extends from the membrane to the catalytic core of F_1 sector. In addition, this protein is also known to possess at least 1-2 transmembrane domains. From the predicted secondary structure features of putative subunit b of *T. gondii*, despite very poor sequence conservation with respect to yeast and mammalian counterpart, very high structural similarity (>97% probability in HHPred) was detectable (Figure-18 A). The F_O sector subunit d is known to interact with subunit b via parallel/ antiparallel, coiled coil helical domains [59], and along with subunit b, h, F6 and OSCP it forms the peripheral stalk structure [66-67]. The putative *T. gondii* F_O subunit d was found to have highly similar secondary structure features over the conserved regions with that of the bovine enzyme (Figure-18 B). Interestingly, the putative subunit b and d from *T. gondii* are >500 amino acids in length, while the corresponding proteins from yeast and mammals is <300 amino acid in length. How these longer proteins are structurally accommodated into the fully assembled F-ATPase complex is an intriguing question. Thus, in order to fully validate the subunit composition of the enzyme, understand the functional role of the novel ASAPs and study their structural features, obtaining a high resolution 3D structure of the *T. gondii* F-ATPase becomes imperative.

Further, the functional association of the novel ASAPs with the readily identified F_1 sector subunits and F_O subunit c and OSCP was ascertained in three different ways. First, in a

genome wide essentiality screen in *T. gondii*, all ASAPs and the known subunits of the F-ATPase were shown to be essential for parasite survival (Figure-22) [126]. Next, the mitochondrial localization of all the ASAPs was confirmed by considering evidence from *in silico* prediction of mitochondrial localization signals as well as from experimental localization of selected ASAPs (Figure-20 and 22). Lastly, gene co-expression analysis showed that, both in *P. falciparum* and in *T. gondii*, the expression of ASAPs was highly correlated with the known subunits of the F-ATPase (Figure-24). In summary, evidence from gene essentiality, subcellular localization, and transcript co-expression analysis provide independent levels of support for validating ASAPs as *bona fide* subunits of *T. gondii* F-ATPase enzyme. Moreover, the authenticity of at least 10 of the ASAPs as subunits of *T. gondii* F-ATPase was also confirmed by an independent and parallel study [129]. This study also revealed that the knockdown of putative F_O sector subunit b lead to defective assembly of F-ATPase and decreased in mitochondrial cristae formation.

Phylogenetic analysis revealed that the ASAPs identified from *T. gondii* are conserved across phylum apicomplexa except in *Cryptosporidium* species which lack a true mitochondria (Figure-22) [129]. Furthermore, the phylogenetic studies also revealed that orthologs of the *T. gondii* ASAPs are also conserved in two other alveolate phyla, Chromerida and Dinoflagellata. This is interesting from an evolutionary perspective, since given the critical role of the F-ATPase in energy metabolism, it appears that the origin of these novel ASAPs is ancient and likely predates the evolutionary branching of the apicomplexan clade. However, the intriguing observation is that the ciliates, which are basal alveolates, have a completely different subunit composition of F-ATPase [97]. Why would the ciliate F-ATPase have a different subunit composition from that of other alveolate species. It turns out that this is not an exclusive case for the F-ATPase enzyme only. Many features of the ciliate mitochondria appears to be different from other alveolates, such as the size of the mitochondrial genome, the number of

protein coding genes encoded by the mitochondrial genome, and the subunit composition of the respiratory chain complexes [26] (Figure-25).

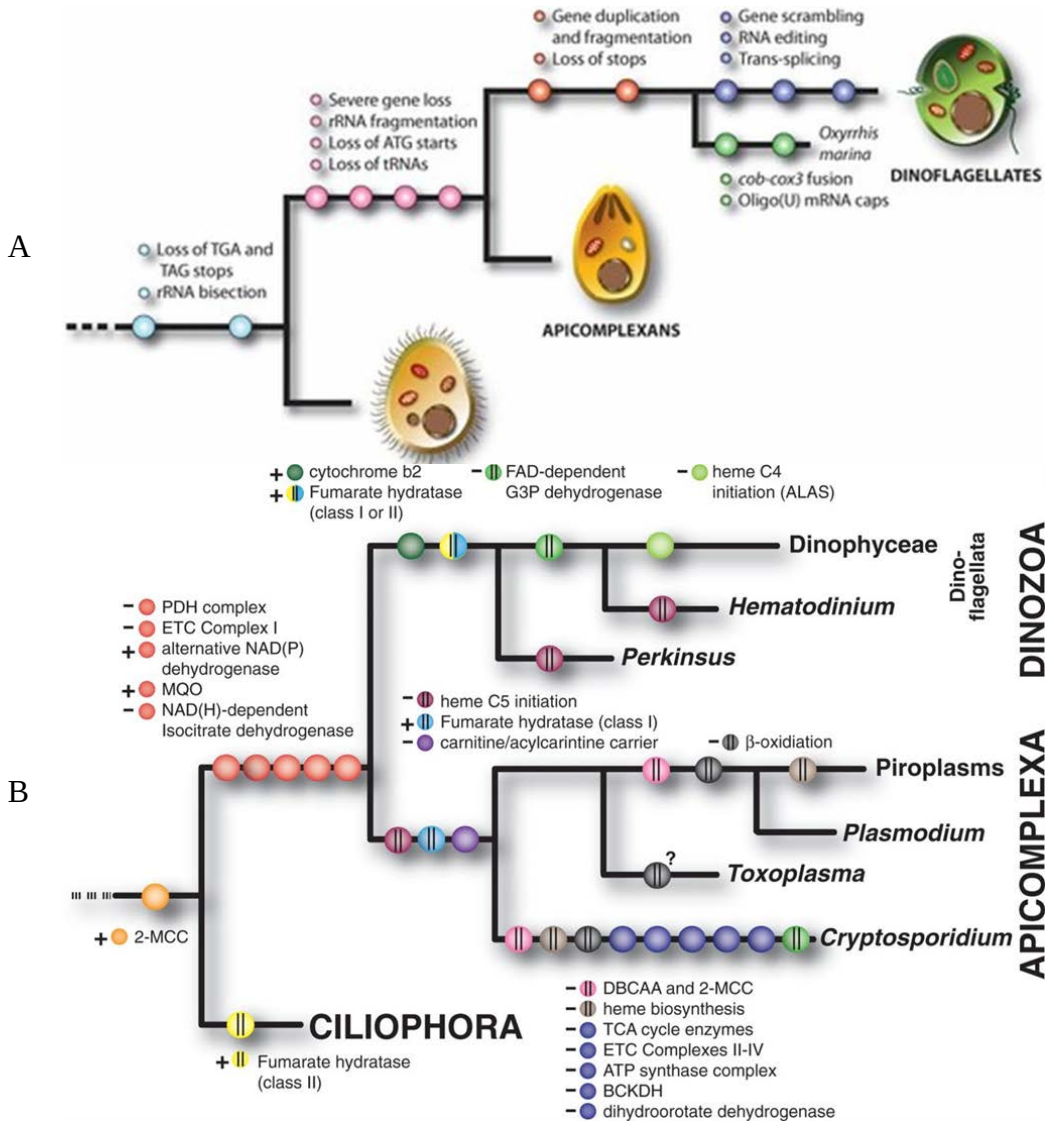


Figure-25: Mitochondrial genome evolution and divergence of metabolic capability within the alveolate clade. (A) Variations in molecular processes associated with the mitochondrion in different alveolate clades. Adapted from [133]. (B) The cladogram shows the presence (+) and absence (−) of mitochondrial metabolic characters based on genomic data, mapped onto tree branches. The parallel gain or loss of genes in different metabolic pathways is indicated by double lines through circles. "?," uncertain presence/absence of pathway due to the presence of incomplete set of genes for the pathways; BCKDH, branched-chain keto-acid dehydrogenase; DBCAA, degradation of branched-chain amino acids; ETC, electron transport chain; 2-MCC, 2-methyl citrate cycle; MQO, malate quinone oxidoreductase; PDH, pyruvate dehydrogenase. Adapted from [26].

For example in the ciliate *T. thermophile*, the mitochondrial genome is approximately 47 kb in size and contains 45 protein coding genes, including genes which code for the F-ATPase F_O sector subunits c and a [97,130]. In contrast, the mitochondrial genomes of other alveolates from phylum Apicomplexa, Chromerida, and Dinoflagellata are highly reduced and encode only 3 or 2 protein coding genes, none of which are F-type ATP synthase subunits [131-133]. This dramatic difference between the ciliates and other alveolates might signify an important bottleneck during evolutionary radiation within the alveolate infrakingdom. The mitochondrion of ancestral alveolate species might have retained many of the complex features and functions associated with the mitochondrion of the closely related stramenopiles. Subsequently, after the divergence of ciliates, it appears that there was massive reduction in the mitochondrial complexity, including the reduction in genome size and metabolic streamlining, in the common ancestor of apicomplexa, dinoflagellata and chromerida [1].

Regardless of the independent evolutionary histories of the apicomplexa, dinoflagellata and chromerida clades, and the different life styles of these organisms – either parasitic or free living – it is striking that the mitochondrion associated features and functions have remained very similar. So, what might be the evolutionary origin of these novel ASAPs in these organisms? Understanding the evolutionary history and driving force behind the highly similar status of the mitochondrion in these organisms may be key to answering this question. One possibility is that the novel ASAPs originated *de novo* in the common ancestor for apicomplexa, dinoflagellata and chromerida species. But why was it necessary to innovate on the F-ATPase and other aspects of mitochondrial biology beyond what was already present in the common ancestor with the ciliates? It is also striking that the conservation of mitochondrial features parallels the presence of the plastid organelle, the apicoplast in these organisms.

It is interesting that the ASAPs identified from *T. gondii* are conserved across all plastid-bearing alveolate clades, suggesting a probable link between the two events. Studies on

the origin of the plastid have revealed it to be acquired from an ancestral red algal endosymbiont (Figure-26). The evolutionary history of the alveolate plastid, however, appears to be complex, with recent studies supporting serial higher order endosymbiotic events (Figure-27) [134-136] rather than a single ancestral endosymbiotic event and subsequent vertical descent [137]. Regardless of these complexities, once the common ancestor of apicomplexa, chromerida and dinoflagellata acquired the plastid, it resulted in large-scale acquisition of genes from the endosymbiont, which are now detectable in the nuclear genome of these organisms. These lateral gene transfer events resulted in the acquisition of genes from the nuclear as well as organelle genomes (as evidenced from genes encoding plastid associated proteins) of the endosymbiont.

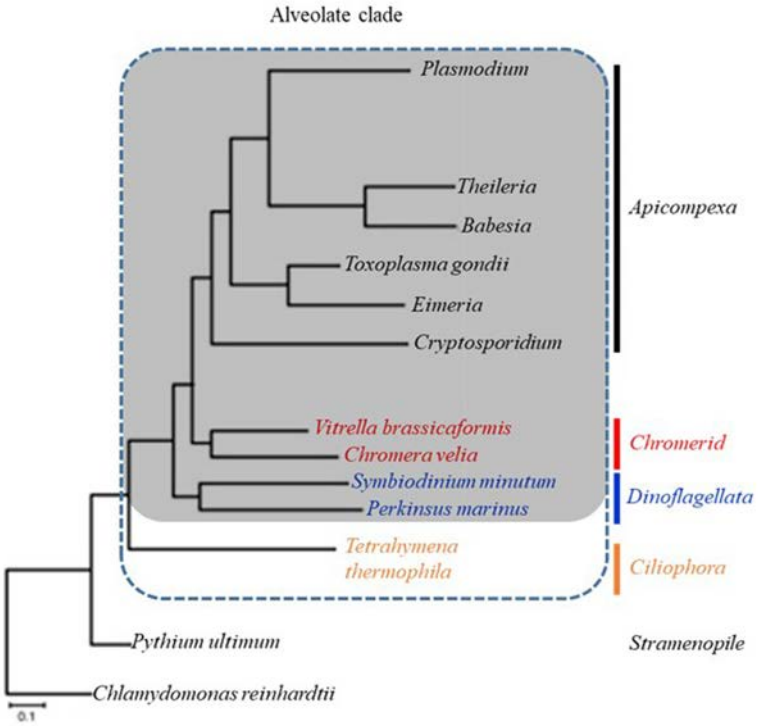


Figure-26: Phylogeny of alveolate infrakingdom and its relationship with free-living *Stramenopiles*. Alveolate species are highlighted by the dotted outline box and the phyla are marked by different color codes on the right. The phylum with red algae derived plastids and the conservation of *T. gondii* F-ATPase ASAPs are shown with gray background. Adapted with modifications from [27].

Thus it is likely that the novel ASAPs present in alveolate F-ATPase may have been acquired from the secondary endosymbiotic event, which gave rise to the plastid organelle in these organisms.

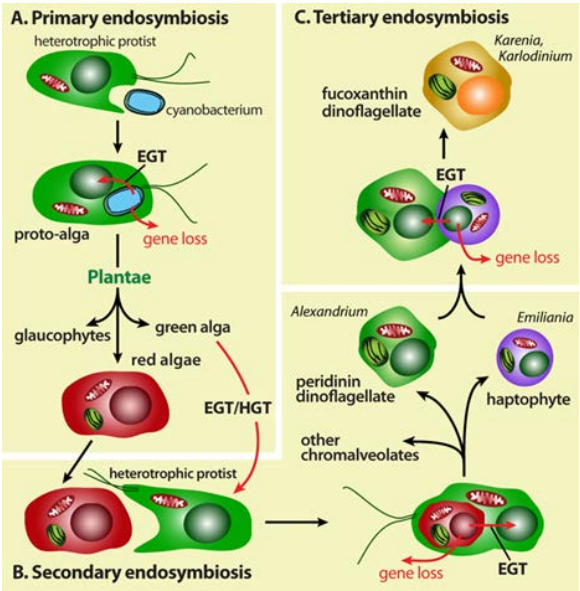


Figure-27: Schematic representation of the evolutionary events resulting in the acquisition of the plastid organelle by alveolates. These events where originally described by the chromalveolate hypothesis which postulated a single secondary endosymbiotic event. But subsequent studies have shown that tertiary and possibly other higher order endosymbiotic events might have played a role in the origin of the plastid in dinoflagellates, and likely in other alveolate species. Adapted from [136].

To summarize, the presence of novel ASAPs in all apicomplexan parasites provides an unprecedented opportunity to study mitochondrial energy metabolism in these parasites. *T. gondii* will be an excellent model organism for these studies, since mutants that are dependent on mitochondrial oxidative phosphorylation as the only bulk source of ATP are already available (for example, the hexokinase enzyme knockout mutants). This will also facilitate inhibitor identification studies from whole cell screens. Given that many subunits of the parasite F-ATPase are completely different from the human enzyme, they will be excellent targets for discovering the next generation of antiparasitic agents, particularly as antimalarial therapy.

Summary and future perspective

Summary:

In this study, we have successfully identified the novel subunit constituents of F-type ATP synthase from *T. gondii*, a model apicomplexan parasite. First, intact monomeric and dimeric forms of the enzyme were identified by BNP analysis of detergent solubilized parasite mitochondrial preparations. Twenty novel hypothetical proteins were identified as subunit constituents of the *T. gondii* F-type ATP synthase, by LC-MS/MS analysis on enzyme that was partially purified using BNP, IP and SEC techniques. These were named as ATP synthase associated proteins (ASAPs). Some of these ASAPs are likely counterparts of the missing F_O subunits, while others are most likely accessory proteins required for dimerization or oligomerization of the enzyme. Importantly, putative F_O sector subunits a, b and d can be identified based on conserved structural features despite extreme sequence divergence in comparison to yeast or mammalian counterparts. Results from CRISPR/Cas9 knock out study indicated that the F1 subunits β and F_O subunit c were essential in *T. gondii*. Interestingly, a genome wide CRISPR/Cas9 mediated gene essentiality screen found that all known subunit of F-type ATP synthase, along with 18 ASAPs were essential in *T. gondii*. In the same study, five ASAPs were localized to the mitochondrion. We have interrogated the mitochondrial localization for all subunits of *T. gondii* F-type ATP synthase by first considering evidence from *in silico* prediction of mitochondrial localization signals, followed by experimental localization of four selected ASAPs. Furthermore, evidence from transcript co-expression analysis indicated that the expression levels of all the novel subunits of F-type ATP synthase were highly correlated with the known F_1 subunits, independently in both *T. gondii* and *P. falciparum* transcriptome datasets. This strongly supports the possibility of functional interaction between these proteins.

The phylogenetic analysis of the highly diverged ASAPs identified in *T.gondii* revealed that most ASAPs identified from *T. gondii* are conserved across the phylum *Apicomplexa* and

also in two other plastid bearing alveolate phylums, Chromerida and Dinoflagellata. This scenario is interesting from an evolutionary perspective, since the ciliates, which are basal alveolates, have a completely different, unique set of ASAPs. Interestingly, the ciliates also dramatically differ from other alveolates by way of missing the secondary endosymbiotic plastid and having a more complex mitochondrion. In summary, evidence from gene essentiality, subcellular localization, transcript co-expression and phylogenetic analysis together provide independent levels of support for the identification of ASAPs as *bona fide* subunits of F-type ATP synthase in *T. gondii* and other alveolate organisms, except the ciliates. This novel and highly diverged parasite F-ATPase is important as a drug target and hence further structure function studies on this protein are warranted.

Future perspectives:

The mitochondrion of apicomplexan parasite is already a valid drug target. In this study, we have shown that the novel F-ATPase ASAPs identified in *T. gondii* are conserved across all apicomplexan parasites including the deadly human malaria parasite *P. falciparum*, which easily develops resistance to currently available anti-parasitic drug. It is very interesting that orthologs of the novel ASAPs identified in this study are restricted only in alveolate species and most importantly are absent in host organism. The novel parasite F-ATPase is essential for survival of the parasite and so provides a selective target for inhibition with new drugs. Therefore, gaining a more complete understanding of the function of F-ATPase enzyme is a critical area of investigation that may facilitate the development of newer antiparasitic therapeutics.

For obtaining a deeper understanding of the structure and function of this enzyme, we are currently pursuing the cryo electron microscopy technique. We hope that the availability of the dimeric structure of the *T. gondii* F-ATPase will facilitate the discovery of novel anti-

parasitic compounds with pan-apicomplexan effect. Since, the novel ASAPs are also conserved in other free-living alveolates, such as *Chromera velia* and *Perkinsus marinus*, the enzyme can be purified in bulk from these species for structural and functional studies. These studies are also likely to shed light on the shared evolution of this important and unusual enzyme in parasitic as well as free-living organisms.

References

1. Dorrell RG, Butterfield ER, Nisbet RE, Howe CJ. Evolution: unveiling early alveolates. Curr Biol. 2013 Dec 16;23(24):R1093-6.
2. Gould SB, Tham WH, Cowman AF, McFadden GI, Waller RF. Alveolins, a new family of cortical proteins that define the protist infrakingdom Alveolata. Mol Biol Evol. 2008 Jun;25(6):1219-30.
3. Alastair G. B. Simpson, David J. Patterson. Ultrastructure and identification of the predatory flagellate Colpodella pugnax Cienkowski (Apicomplexa) with a description of Colpodella turpis n. sp. and a review of the genus. Systematic Parasitology, 1996, Volume 33, Number 3, Page 187.
4. Moore RB, Oborník M, Janouskovec J, Chrudimský T, Vancová M, Green DH, *et al*. A photosynthetic alveolate closely related to apicomplexan parasites. Nature. 2008 Feb 21;451(7181):959-63.
5. Meissner M. The asexual cycle of apicomplexan parasites: new findings that raise new questions. Curr Opin Microbiol. 2013 Aug;16(4):421-3.
6. Weiss LM, Dubey JP. Toxoplasmosis: A history of clinical observations. Int J Parasitol. 2009 Jul 1;39(8):895-901.
7. Dubey JP. Advances in the life cycle of Toxoplasma gondii. Int J Parasitol. 1998 Jul;28(7):1019-24.
8. Hunter CA, Sibley LD. Modulation of innate immunity by Toxoplasma gondii virulence effectors. Nat Rev Microbiol. 2012 Nov;10(11):766-78.
9. Lunghi M, Galizi R, Magini A, Carruthers VB, Di Cristina M. Expression of the glycolytic enzymes enolase and lactate dehydrogenase during the early phase of Toxoplasma differentiation is regulated by an intron retention mechanism. Mol Microbiol. 2015 Jun;96(6):1159-75.
10. Clough B, Frickel EM. The Toxoplasma Parasitophorous Vacuole: An Evolving Host-Parasite Frontier. Trends Parasitol. 2017 Jun;33(6):473-488.
11. Polonais V, Soldati-Favre D. Versatility in the acquisition of energy and carbon sources by the Apicomplexa. Biol Cell. 2010 Apr 23;102(8):435-45.
12. Oppenheim RD, Creek DJ, Macrae JI, Modrzynska KK, Pino P, Limenitakis J, *et al*. BCKDH: the missing link in apicomplexan mitochondrial metabolism is required for full virulence of Toxoplasma gondii and Plasmodium berghei. PLoS Pathog. 2014 Jul 17;10(7):e1004263.

13. Crawford MJ, Thomsen-Zieger N, Ray M, Schachtner J, Roos DS, Seeber F. Toxoplasma gondii scavenges host-derived lipoic acid despite its de novo synthesis in the apicoplast. EMBO J. 2006 Jul 12;25(13):3214-22.
14. Foth BJ, Stimmler LM, Handman E, Crabb BS, Hodder AN, McFadden GI. The malaria parasite Plasmodium falciparum has only one pyruvate dehydrogenase complex, which is located in the apicoplast. Mol Microbiol. 2005 Jan;55(1):39-53.
15. Fleige T, Fischer K, Ferguson DJ, Gross U, Bohne W. Carbohydrate metabolism in the Toxoplasma gondii apicoplast: localization of three glycolytic isoenzymes, the single pyruvate dehydrogenase complex, and a plastid phosphate translocator. Eukaryot Cell. 2007 Jun;6(6):984-96.
16. Al-Anouti F, Tomavo S, Parmley S, Ananvoranich S. The expression of lactate dehydrogenase is important for the cell cycle of Toxoplasma gondii. J Biol Chem. 2004 Dec 10;279(50):52300-11.
17. Pomel S, Luk FC, Beckers CJ. Host cell egress and invasion induce marked relocations of glycolytic enzymes in Toxoplasma gondii tachyzoites. PLoS Pathog. 2008 Oct;4(10):e1000188.
18. Blume M, Rodriguez-Contreras D, Landfear S, Fleige T, Soldati-Favre D, Lucius R, *et al*. Host-derived glucose and its transporter in the obligate intracellular pathogen Toxoplasma gondii are dispensable by glutaminolysis. Proc Natl Acad Sci U S A. 2009 Aug 4;106(31):12998-3003.
19. Shukla A, Olszewski KL, Llinás M, Rommereim LM, Fox BA, Bzik DJ, *et al*. Glycolysis is important for optimal asexual growth and formation of mature tissue cysts by Toxoplasma gondii. Int J Parasitol. 2018 Oct;48(12):955-968.
20. Shen B, Sibley LD. Toxoplasma aldolase is required for metabolism but dispensable for host-cell invasion. Proc Natl Acad Sci U S A. 2014 Mar 4;111(9):3567-72.
21. Slavic K, Straschil U, Reininger L, Doerig C, Morin C, Tewari R, *et al*. Life cycle studies of the hexose transporter of Plasmodium species and genetic validation of their essentiality. Mol Microbiol. 2010 Mar;75(6):1402-13.
22. Sagan L. On the origin of mitosing cells. 1967. J NIH Res. 1993 Mar;5(3):65-72.
23. McBride HM, Neuspiel M, Wasiak S. Mitochondria: more than just a powerhouse. Curr Biol. 2006 Jul 25;16(14):R551-60.
24. Owen OE, Kalhan SC, Hanson RW. The key role of anaplerosis and cataplerosis for citric acid cycle function. J Biol Chem. 2002 Aug 23;277(34):30409-12.

25. Akram M. Citric acid cycle and role of its intermediates in metabolism. Cell Biochem Biophys. 2014 Apr;68(3):475-8.
26. Danne JC, Gornik SG, Macrae JI, McConville MJ, Waller RF. Alveolate mitochondrial metabolic evolution: dinoflagellates force reassessment of the role of parasitism as a driver of change in apicomplexans. Mol Biol Evol. 2013 Jan;30(1):123-39.
27. Woo YH, Ansari H, Otto TD, Klinger CM, Kolisko M, Michálek J, *et al.* Chromerid genomes reveal the evolutionary path from photosynthetic algae to obligate intracellular parasites. Elife. 2015 Jul 15;4:e06974.
28. MacRae JI, Sheiner L, Nahid A, Tonkin C, Striepen B, McConville MJ. Mitochondrial metabolism of glucose and glutamine is required for intracellular growth of Toxoplasma gondii. Cell Host Microbe. 2012 Nov 15;12(5):682-92.
29. Ke H, Lewis IA, Morrisey JM, McLean KJ, Ganesan SM, Painter HJ, *et al*. Genetic investigation of tricarboxylic acid metabolism during the Plasmodium falciparum life cycle. Cell Rep. 2015 Apr 7;11(1):164-74.
30. Yang J, Kalhan SC, Hanson RW. What is the metabolic role of phosphoenolpyruvate carboxykinase? J Biol Chem. 2009 Oct 2;284(40):27025-9. PMID: 19636077.
31. Nitzsche R, Günay-Esiyok Ö, Tischer M, Zagoriy V, Gupta N. A plant/fungal-type phosphoenolpyruvate carboxykinase located in the parasite mitochondrion ensures glucose-independent survival of Toxoplasma gondii. J Biol Chem. 2017 Sep 15;292(37):15225-15239.
32. Macrae JI, Dixon MW, Dearnley MK, Chua HH, Chambers JM, Kenny S *et al*. (2013) Mitochondrial metabolism of sexual and asexual blood stages of the malaria parasite Plasmodium falciparum. BMC Biol. 2013 Jun 13;11:67.
33. Hatefi Y. The mitochondrial electron transport and oxidative phosphorylation system. Annu Rev Biochem. 1985;54:1015-69.
34. Saraste M. Oxidative phosphorylation at the fin de siècle. Science. 1999 Mar 5;283(5407):1488-93.
35. Vercesi AE, Rodrigues CO, Uyemura SA, Zhong L, Moreno SN. Respiration and oxidative phosphorylation in the apicomplexan parasite Toxoplasma gondii. J Biol Chem. 1998 Nov 20;273(47):31040-7.
36. Krungkrai J, Burat D, Kudan S, Krungkrai S, Prapunwattana P. Mitochondrial oxygen consumption in asexual and sexual blood stages of the human malarial parasite, Plasmodium falciparum. Southeast Asian J Trop Med Public Health. 1999 Dec;30(4):636-42.

37. Uyemura SA, Luo S, Moreno SN, Docampo R. Oxidative phosphorylation, Ca(2+) transport, and fatty acid-induced uncoupling in malaria parasites mitochondria. J Biol Chem. 2000 Mar 31;275(13):9709-15.
38. Seeber F, Limenitakis J, Soldati-Favre D. Apicomplexan mitochondrial metabolism: a story of gains, losses and retentions. Trends Parasitol. 2008 Oct;24(10):468-78.
39. Flegontov P, Michálek J, Janouškovec J, Lai DH, Jirků M, Hajdušková E, *et al*. Divergent mitochondrial respiratory chains in phototrophic relatives of apicomplexan parasites. Mol Biol Evol. 2015 May;32(5):1115-31.
40. Kerscher SJ. Diversity and origin of alternative NADH:ubiquinone oxidoreductases. Biochim Biophys Acta. 2000 Aug 15;1459(2-3):274-83.
41. Rasmusson AG, Soole KL, Elthon TE. Alternative NAD(P)H dehydrogenases of plant mitochondria. Annu Rev Plant Biol. 2004;55:23-39.
42. Melo AM, Bandeiras TM, Teixeira M. New insights into type II NAD(P)H:quinone oxidoreductases. Microbiol Mol Biol Rev. 2004 Dec;68(4):603-16.
43. Kerscher S, Dröse S, Zickermann V, Brandt U. The three families of respiratory NADH dehydrogenases. Results Probl Cell Differ. 2008; 45:185-222.
44. Saleh A, Friesen J, Baumeister S, Gross U, Bohne W. Growth inhibition of Toxoplasma gondii and Plasmodium falciparum by nanomolar concentrations of 1-hydroxy-2-dodecyl-4(1H)quinolone, a high-affinity inhibitor of alternative (type II) NADH dehydrogenases. Antimicrob Agents Chemother. 2007 Apr; 51 (4):1217-22.
45. Boysen KE, Matuschewski K. Arrested oocyst maturation in Plasmodium parasites lacking type II NADH:ubiquinone dehydrogenase. J Biol Chem. 2011 Sep 16; 286 (37):32661-71.
46. Lin SS, Gross U, Bohne W. Type II NADH dehydrogenase inhibitor 1-hydroxy-2-dodecyl-4(1H) quinolone leads to collapse of mitochondrial inner-membrane potential and ATP depletion in Toxoplasma gondii. Eukaryot Cell. 2009 Jun; 8(6):877-87.
47. Trumpower BL, Gennis RB. Energy transduction by cytochrome complexes in mitochondrial and bacterial respiration: the enzymology of coupling electron transfer reactions to transmembrane proton translocation. Annu Rev Biochem. 1994; 63:675-716.
48. Hunte C, Solmaz S, Palsdóttir H, Wenz T. A structural perspective on mechanism and function of the cytochrome bc (1) complex. Results Probl Cell Differ. 2008; 45:253-78.

49. Srivastava IK, Rottenberg H, Vaidya AB. Atovaquone, a broad spectrum antiparasitic drug, collapses mitochondrial membrane potential in a malarial parasite. J Biol Chem. 1997 Feb 14; 272(7):3961-6.
50. Baggish AL, Hill DR. Antiparasitic agent atovaquone. Antimicrob Agents Chemother. 2002 May; 46(5):1163-73.
51. Doggett JS, Nilsen A, Forquer I, Wegmann KW, Jones-Brando L, Yolken RH, *et al*. Endochin-like quinolones are highly efficacious against acute and latent experimental toxoplasmosis. Proc Natl Acad Sci U S A. 2012 Sep 25; 109(39):15936-41.
52. Nilsen A, LaCrue AN, White KL, Forquer IP, Cross RM, Marfurt J, *et al*. Quinolone-3-diarylethers: a new class of antimalarial drug. Sci Transl Med. 2013 Mar 20; 5(177):177ra37.
53. Birth D, Kao WC, Hunte C. Structural analysis of atovaquone-inhibited cytochrome bc1 complex reveals the molecular basis of antimalarial drug action. Nat Commun. 2014 Jun 4; 5:4029.
54. Goodman CD, Siregar JE, Mollard V, Vega-Rodríguez J, Syafruddin D, Matsuoka H, *et al*. Parasites resistant to the antimalarial atovaquone fail to transmit by mosquitoes. Science. 2016 Apr 15; 352(6283):349-53.
55. Boyer PD. The ATP synthase a splendid molecular machine. Annu Rev Biochem. 1997; 66:717-49.
56. Capaldi RA, Aggeler R. Mechanism of the F(1)F(0)-type ATP synthase, a biological rotary motor. Trends Biochem Sci. 2002 Mar; 27(3):154-60.
57. Walker JE. The ATP synthase: the understood, the uncertain and the unknown. Biochem Soc Trans. 2013 Feb 1; 41(1):1-16.
58. Stock D, Leslie AG, Walker JE. Molecular architecture of the rotary motor in ATP synthase. Science. 1999 Nov 26;286(5445):1700-5.
59. Devenish RJ, Prescott M, Rodgers AJ. The structure and function of mitochondrial F1F0-ATP synthases. Int Rev Cell Mol Biol. 2008; 267:1-58.
60. Hahn A, Vonck J, Mills DJ, Meier T, Kühlbrandt W. Structure, mechanism, and regulation of the chloroplast ATP synthase. Science. 2018 May 11; 360(6389).
61. von Ballmoos C, Cook GM, Dimroth P. Unique rotary ATP synthase and its biological diversity. Annu Rev Biophys. 2008; 37:43-64.
62. Mitchell P. Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism. Nature. 1961 Jul 8;191:144-8.

63. Cain BD, Simoni RD. Proton translocation by the F1F0ATPase of Escherichia coli. Mutagenic analysis of the a subunit. J Biol Chem. 1989 Feb 25;264(6):3292-300.
64. Valiyaveetil FI, Fillingame RH. On the role of Arg-210 and Glu-219 of subunit a in proton translocation by the Escherichia coli F0F1-ATP synthase. J Biol Chem. 1997 Dec 19;272(51):32635-41.
65. Rastogi VK, Girvin ME. Structural changes linked to proton translocation by subunit c of the ATP synthase. Nature. 1999 Nov 18;402(6759):263-8.
66. Weber J. ATP synthase: subunit-subunit interactions in the stator stalk. Biochim Biophys Acta. 2006 Sep-Oct;1757(9-10):1162-70.
67. Walker JE, Dickson VK. The peripheral stalk of the mitochondrial ATP synthase. Biochim Biophys Acta. 2006 May-Jun;1757(5-6):286-96.
68. Sielaff H, Börsch M. Twisting and subunit rotation in single F(O)(F1)-ATP synthase. Philos Trans R Soc Lond B Biol Sci. 2012 Dec 24;368(1611):20120024.
69. Nelson, D. L., Lehninger, A. L., & Cox, M. M. (2013). Lehninger's principles of biochemistry (6th ed.). New York: W.H. Freeman.
70. Mesbah NM, Wiegel J. The Na(+)-translocating F_1F_o-ATPase from the halophilic, alkalithermophile Natranaerobius thermophilus.Biochim Biophys Acta. 2011 Sep;1807(9):1133-42.
71. Nalin CM, McCarty RE. Role of a disulfide bond in the gamma subunit in activation of the ATPase of chloroplast coupling factor 1. J Biol Chem. 1984 Jun 10;259(11):7275-80.
72. Campanella M, Parker N, Tan CH, Hall AM, Duchen MR. IF(1): setting the pace of the F(1)F(o)-ATP synthase. Trends Biochem Sci. 2009 Jul;34(7):343-50.
73. Daum B, Kühlbrandt W. Electron tomography of plant thylakoid membranes. J Exp Bot. 2011 Apr;62(7):2393-402.
74. Wittig I, Schägger H. Supramolecular organization of ATP synthase and respiratory chain in mitochondrial membranes. Biochim Biophys Acta. 2009 Jun;1787(6):672-80.
75. Allen RD, Schroeder CC, Fok AK. An investigation of mitochondrial inner membranes by rapid-freeze deep-etch techniques. J Cell Biol. 1989 Jun;108(6):2233-40.
76. Schägger H, Pfeiffer K. Supercomplexes in the respiratory chains of yeast and mammalian mitochondria. EMBO J. 2000 Apr 17;19(8):1777-83.
77. Krause F, Reifschneider NH, Goto S, Dencher NA. Active oligomeric ATP synthases in mammalian mitochondria. Biochem Biophys Res Commun. 2005 Apr 8;329(2):583-90.

78. Pfeiffer K, Gohil V, Stuart RA, Hunte C, Brandt U, Greenberg ML, *et al*. Cardiolipin stabilizes respiratory chain supercomplexes. J Biol Chem. 2003 Dec 26;278(52):52873-80.
79. Eubel H, Heinemeyer J, Sunderhaus S, Braun HP. Respiratory chain supercomplexes in plant mitochondria. Plant Physiol Biochem. 2004 Dec;42(12):937-42.
80. Dudkina NV, Heinemeyer J, Keegstra W, Boekema EJ, Braun HP. Structure of dimeric ATP synthase from mitochondria: an angular association of monomers induces the strong curvature of the inner membrane. FEBS Lett. 2005 Oct 24;579(25):5769-72.
81. Dudkina NV, Heinemeyer J, Sunderhaus S, Boekema EJ, Braun HP. Respiratory chain supercomplexes in the plant mitochondrial membrane. Trends Plant Sci. 2006 May;11(5):232-40.
82. Davies KM, Anselmi C, Wittig I, Faraldo-Gómez JD, Kühlbrandt W. Structure of the yeast F1Fo-ATP synthase dimer and its role in shaping the mitochondrial cristae. Proc Natl Acad Sci U S A. 2012 Aug 21;109(34):13602-7.
83. Dudkina NV, Sunderhaus S, Braun HP, Boekema EJ. Characterization of dimeric ATP synthase and cristae membrane ultrastructure from Saccharomyces and Polytomella mitochondria. FEBS Lett. 2006 Jun 12;580(14):3427-32.
84. Paumard P, Vaillier J, Coulary B, Schaeffer J, Soubannier V, Mueller DM, *et al*. The ATP synthase is involved in generating mitochondrial cristae morphology. EMBO J. 2002 Feb 1;21(3):221-30.
85. Arnold I, Pfeiffer K, Neupert W, Stuart RA, Schägger H. Yeast mitochondrial F1F0-ATP synthase exists as a dimer: identification of three dimer-specific subunits. EMBO J. 1998 Dec 15;17(24):7170-8.
86. Brunner S, Everard-Gigot V, Stuart RA. Su e of the yeast F1Fo-ATP synthase forms homodimers. J Biol Chem. 2002 Dec 13;277(50):48484-9.
87. Arselin G, Giraud MF, Dautant A, Vaillier J, Brèthes D, Coulary-Salin B, *et al*. The GxxxG motif of the transmembrane domain of subunit e is involved in the dimerization/oligomerization of the yeast ATP synthase complex in the mitochondrial membrane. Eur J Biochem. 2003 Apr;270(8):1875-84.
88. Collinson IR, Runswick MJ, Buchanan SK, Fearnley IM, Skehel JM, van Raaij MJ, *et al*. Fo membrane domain of ATP synthase from bovine heart mitochondria: purification, subunit composition, and reconstitution with F1-ATPase. Biochemistry. 1994 Jun 28;33(25):7971-8.

89. Walker JE, Lutter R, Dupuis A, Runswick MJ. Identification of the subunits of F1F0-ATPase from bovine heart mitochondria.Biochemistry. 1991 Jun 4;30(22):5369-78.
90. Gavin PD, Prescott M, Devenish RJ. F1F0-ATP synthase complex interactions in vivo can occur in the absence of the dimer specific subunit e. J Bioenerg Biomembr. 2005 Apr;37(2):55-66.
91. Velours J, Arselin G. The Saccharomyces cerevisiae ATP synthase. J Bioenerg Biomembr. 2000 Aug;32(4):383-90.
92. Meyer B, Wittig I, Trifilieff E, Karas M, Schägger H. Identification of two proteins associated with mammalian ATP synthase. Mol Cell Proteomics. 2007 Oct;6(10):1690-9.
93. van Lis R, Atteia A, Mendoza-Hernández G, González-Halphen D. Identification of novel mitochondrial protein components of Chlamydomonas reinhardtii. A proteomic approach. Plant Physiol. 2003 May;132(1):318-30.
94. Vázquez-Acevedo M, Cardol P, Cano-Estrada A, Lapaille M, Remacle C, González-Halphen D. The mitochondrial ATP synthase of chlorophycean algae contains eight subunits of unknown origin involved in the formation of an atypical stator-stalk and in the dimerization of the complex. J Bioenerg Biomembr. 2006 Dec;38(5-6):271-82.
95. Zíková A, Schnaufer A, Dalley RA, Panigrahi AK, Stuart KD. The F(0)F(1)-ATP synthase complex contains novel subunits and is essential for procyclic Trypanosoma brucei. PLoS Pathog. 2009 May;5(5):e1000436.
96. Yadav KNS, Miranda-Astudillo HV, Colina-Tenorio L, Bouillenne F, Degand H, Morsomme P, *et al*. Atypical composition and structure of the mitochondrial dimeric ATP synthase from Euglena gracilis. Biochim Biophys Acta. 2017 Apr;1858(4):267-275.
97. Balabaskaran Nina P, Dudkina NV, Kane LA, van Eyk JE, Boekema EJ, Mather MW, *et al*. Highly divergent mitochondrial ATP synthase complexes in Tetrahymena thermophila. PLoS Biol. 2010 Jul 13;8(7):e1000418.
98. Mather MW, Henry KW, Vaidya AB. Mitochondrial drug targets in apicomplexan parasites. Curr Drug Targets. 2007 Jan;8(1):49-60.
99. Balabaskaran Nina P, Morrisey JM, Ganesan SM, Ke H, Pershing AM, Mather MW, *et al*. ATP synthase complex of Plasmodium falciparum: dimeric assembly in mitochondrial membranes and resistance to genetic disruption. J Biol Chem. 2011 Dec 2;286(48):41312-22.

100. Sturm A, Mollard V, Cozijnsen A, Goodman CD, McFadden GI. Mitochondrial ATP synthase is dispensable in blood-stage Plasmodium berghei rodent malaria but essential in the mosquito phase. Proc Natl Acad Sci U S A. 2015 Aug 18;112(33):10216-23.
101. Roos DS. Primary structure of the dihydrofolate reductase-thymidylate synthase gene from Toxoplasmagondii. J Biol Chem. 1993 Mar 25; 268(9):6269-80.
102. Donald RG, Roos DS. Stable molecular transformation of Toxoplasma gondii: a selectable dihydrofolate reductase-thymidylate synthase marker based on drug-resistance mutations in malaria. Proc Natl Acad Sci U S A. 1993 Dec 15; 90(24):11703-7.
103. Kim K, Soldati D, Boothroyd JC. Gene replacement in Toxoplasma gondii with chloramphenicol acetyltransferase as selectable marker. Science. 1993 Nov 5; 262(5135):911-4.
104. Shen B, Brown K, Long S, Sibley LD. Development of CRISPR/Cas9 for Efficient Genome Editing in Toxoplasma gondii. Methods Mol Biol. 2017; 1498:79-103.
105. Huynh MH, Carruthers VB. Tagging of endogenous genes in a Toxoplasma gondii strain lacking Ku80. Eukaryot Cell. 2009 Apr; 8(4):530-9.
106. Roos DS, Donald RG, Morrissette NS, Moulton AL. Molecular tools for genetic dissection of the protozoan parasite Toxoplasma gondii. Methods Cell Biol. 1994; 45:27-63.
107. Sims NR, Anderson MF. Isolation of mitochondria from rat brain using Percoll density gradient centrifugation. Nat Protoc. 2008;3(7):1228-39.
108. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem. 1976 May 7; 72:248-54.
109. Wittig I, Braun HP, SchaÈgger H. Blue native PAGE. Nat Protoc. 2006; 1(1):418-28.
110. Nijtmans LG, Henderson NS, Holt IJ. Blue Native electrophoresis to study mitochondrial and other protein complexes. Methods. 2002 Apr; 26(4):327-34.
111. Allegretti M, Klusch N, Mills DJ, Vonck J, KuÈhlbrandt W, Davies KM. Horizontal membrane-intrinsic α-helices in the stator a-subunit of an F-type ATP synthase. Nature. 2015 May 14; 521(7551):237-40.
112. Korwar AM, Vannuruswamy G, Jagadeeshaprasad MG, Jayaramaiah RH, Bhat S, Regin BS, *et al*. Development of Diagnostic Fragment Ion Library for Glycated Peptides of Human Serum Albumin: Targeted Quantification in Prediabetic, Diabetic,

and Microalbuminuria Plasma by Parallel Reaction Monitoring, SWATH, and MSE. Mol Cell Proteomics. 2015 Aug; 14(8):2150-9.

113. Behnke MS, Wootton JC, Lehmann MM, Radke JB, Lucas O, Nawas J, *et al*. Coordinated progression through two subtranscriptomes underlies the tachyzoite cycle of Toxoplasma gondii. PLoS ONE. 2010 Aug 26;5(8):e12354.
114. Bozdech Z, Llinás M, Pulliam BL, Wong ED, Zhu J, DeRisi JL. The transcriptome of the intraerythrocytic developmental cycle of Plasmodium falciparum. PLoS Biol. 2003 Oct;1(1):E5.
115. Hollander M, Wolfe DA, Chicken E, editors. Nonparametric Statistical Methods. 3rd ed. Hoboken, N.J.:Wiley; 2013.
116. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, N.J.: L. Erlbaum Associates; 1988.
117. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013 Apr; 30(4):772-80.
118. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, *et al*. ClustalWand Clustal X version 2.0. Bioinformatics. 2007 Nov 1; 23(21):2947-8.
119. McFadden DC, Tomavo S, Berry EA, Boothroyd JC. Characterization of cytochrome b from Toxoplasma gondii and Q (o) domain mutations as a mechanism of atovaquone-resistance. Mol Biochem Parasitol. 2000 Apr 30;108(1):1-12.
120. Pullman ME, Monroy GC. A Naturally Occurring Inhibitor Of Mitochondrial Adenosine Triphosphatase. J Biol Chem. 1963 Nov;238:3762-9.
121. Zimmermann L, Stephens A, Nam SZ, Rau D, Kübler J, Lozajic M, *et al*. A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. J Mol Biol. 2017 Dec 16.
122. Söding J. Protein homology detection by HMM-HMM comparison. Bioinformatics. 2005 Apr 1;21(7):951-60. Erratum in: Bioinformatics. 2005 May 1;21(9):2144.
123. Claros MG, Vincens P. Computational method to predict mitochondrially imported proteins and their targeting sequences. Eur J Biochem. 1996 Nov 1;241(3):779-86.
124. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, *et al*. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2011 Oct 11;7:539.
125. Robert X, Gouet P. Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res. 2014 Jul;42(Web Server issue):W320-4.

126. Sidik SM, Huet D, Ganesan SM, Huynh MH, Wang T, Nasamu AS, *et al*. A Genome-wide CRISPR Screen in Toxoplasma Identifies Essential Apicomplexan Genes. Cell. 2016 Sep 8;166(6):1423-1435.e12.
127. Allen RD, Schroeder CC, Fok AK. An investigation of mitochondrial inner membranes by rapid-freeze deep-etch techniques. J Cell Biol. 1989 Jun;108(6):2233–40.
128. Bornhövd C, Vogel F, Neupert W, Reichert AS. Mitochondrial membrane potential is dependent on the oligomeric state of F1F0-ATP synthase supracomplexes. J Biol Chem. 2006 May 19;281(20):13990–8.
129. Huet D, Rajendran E, van Dooren GG, Lourido S. Identification of cryptic subunits from an apicomplexan ATP synthase. Elife. 2018 Sep 11;7. pii: e38097.
130. Brunk CF, Lee LC, Tran AB, Li J. Complete sequence of the mitochondrial genome of Tetrahymena thermophila and comparative methods for identifying highly divergent genes. Nucleic Acids Res. 2003 Mar 15;31(6):1673–82.
131. Vaidya AB, Mather MW. Mitochondrial evolution and functions in malaria parasites. Annu Rev Microbiol. 2009;63:249–67.
132. Oborník M, Lukeš J. The Organellar Genomes of Chromera and Vitrella, the Phototrophic Relatives of Apicomplexan Parasites. Annu Rev Microbiol. 2015;69:129–44.
133. Waller RF, Jackson CJ. Dinoflagellate mitochondrial genomes: stretching the rules of molecular biology. Bioessays. 2009 Feb;31(2):237–45.
134. Petersen J, Ludewig AK, Michael V, Bunk B, Jarek M, Baurain D, *et al*. Chromera velia, endosymbioses and the rhodoplex hypothesis—plastid evolution in cryptophytes, alveolates, stramenopiles, and haptophytes (CASH lineages). Genome Biol Evol. 2014 Mar;6(3):666–84.
135. Waller RF, Gornik SG, Koreny L, Pain A. Metabolic pathway redundancy within the apicomplexan-dinoflagellate radiation argues against an ancient chromalveolate plastid. Commun Integr Biol. 2015 Dec 8;9(1):e1116653.
136. Chan, C. X. & Bhattacharya, D. (2010) The Origin of Plastids. Nature Education 3(9):84.
137. Cavalier-Smith T. Principles of protein and lipid targeting in secondary symbiogenesis: euglenoid, dinoflagellate, and sporozoan plastid origins and the eukaryote family tree. J Eukaryot Microbiol. 1999 Jul-Aug;46(4):347–66.