Mitochondrial gene expression system involving extensive frameshift in a shellfish pathogen *Perkinsus*

貝類寄生虫 ***Perkinsus*** における高頻度なフレームシフトを伴う

ミトコンドリア遺伝子発現系

増田　功

CONTENTS

ABBREVIATIONS

ADP	adenosine diphosphate
AOX	alternative oxidase
ATP	adenosine triphosphate
BLAST	basic local alignment search tool
BSA	bovine serum albumin
COX	cytochrome *c* oxidase
COX1	cytochrome *c* oxidase subunit 1
cDNA	complementary DNA
EDTA	ethylenediaminetetraacetic acid
ETC	electron-transport chain
EST	expressed sequence tag
gDNA	genomic DNA
GFP	green fluorescent protein
indel	insertion and deletion
LSU rDNA	large subunit ribosomal DNA
mRNA	messenger RNA
mt gene	mitochondrial gene
mt genome	mitochondrial genome
NAD(H)	nicotiamide dinucleotide
NCBI	National Center for Biotechnology Information
NDH	NADH dehydrogenase
NDH2	type II NADH dehydrogenase
PCR	polymerase chain reaction

QCR	ubiquinol-cytochrome *c* reductase
RACE	rapid amplificaion of cDNA end
rDNA	ribosomal DNA
rRNA	ribosomal RNA
SDH	succinate dehydrogenase
SDS	sodium dodecyl sulfate
SQR	succinate-ubiquinone reductase
TBE	tris-borate EDTA
TBS-T	tris-buffered saline with 0.05% (v/v) Tween-20
TCA	tricarboxylic acid
TIGR	The Institute for Genomic Research
Tris	tris(hydroxymethyl)amino methane
tRNA	transfer RNA
WHO	World Health Organization

ABSTRACT

Apicomplexan parasites are the causative agents of human infectious diseases including malaria. Modified metabolic pathways and unique gene expression systems of the parasites' mitochondria are promising drug targets. From an evolutionary viewpoint, in this study, I investigated mitochondria of *Perkinsus*, a relative of apicomplexans. The enrichment method for *Perkinsus* mitochondria was preliminarily established using Percoll density gradient ultracentrifugation. *Perkinsus* cells obviously showed enzyme activity ensuring the presence of the active mitochondrial electron-transport chain, and homologs of major subunits of mitochondrial electron-transport chain proteins of *P. marinus* were found from a public database. Although the sequence searches did not detect possible *Perkinsus* mitochondrial genes, I could determine the full-length mRNA sequence of *cox1*. Similar to apicomplexans, this gene lacked canonical start and stop codons in terminal regions. Curiously, this mRNA was not translated in a single reading frame with standard codon usage. Careful examination of the nucleotide sequence and its three-frame translation suggested that the reading frame must be shifted 10 times, at every AGG and CCC codon to yield a consensus COX1 protein. Two possible mechanisms were proposed for the frameshifts: ribosomal frameshifts in which stalled ribosomes skip the first bases of these codons or specialized tRNAs recognizing non-triplet codons, AGGY and CCCCU. Notably, this is the most extensive frameshift case described to date, which would utilize active and efficient machinery. It is valuable that this is a fundamental study on *Perkinsus* mitochondria and would provide evolutionary insights into mitochondrial functions of *Perkinsus* and their relatives including apicomplexans.

Chapter 1. GENERAL INTRODUCTION

1.1. The apicomplexan parasites

The spread of infectious diseases has long been one of the big worldwide problems. The most devastating infectious diseases are AIDS (caused by virus), tuberculosis (by bacterium), and malaria (by parasitic protist). Despite the generations of efforts in preventing the outspread of the diseases, they led to the deaths of as many as 4,000,000 people in total, annually (WHO, 2009) and also cause serious economic losses. Among the three prevalent and potentially fatal infectious diseases, malaria is majorly widespread in tropical and subtropical areas, and most prevalent in sub-Saharan Africa. According to the annual report on malaria by World Health Organization (WHO), there were 243 million cases of malaria and 863,000 people died in the world per year (WHO, 2009).

The etiological agents of malaria are the parasitic protists belonging to the genus *Plasmodium*. This parasite is a member of the phylum Apicomplexa, one of the largest assemblages comprised of parasitic protists (Levine *et al*., 1980; Levine, 1988; Adl *et al*., 2005). This phylum encompasses diverse parasitic species which cause human infectious disease. For example, *Toxoplasma gondii* induces influenza-like symptoms and *Cryptosporidium parvum* causes watery diarrhea; remarkably, patients with weakened immune system become critically ill. In order to treat and control these parasitic diseases, it is an urgent task to characterize the apicomplexan parasites' biological features (e.g., ultrastructure, life cycle, metabolic network, and gene expression) and to discover antiparasitic compounds and their targets. For this purpose, a number of molecular biological and biochemical studies have extensively been undertaken along with ecological, epidemiological, and evolutionary studies.

1.2. Mitochondria, the center of energy metabolism in a eukaryotic cell

Eukaryotic cells contain various organelles forming a complicated network with a number of metabolic pathways. Among them, mitochondria take part in various important functions like iron-sulfur cluster assembly and beta-oxidation of fatty acids. Most importantly, mitochondria take a pivotal role in the generation of adenosine triphosphate (ATP) in the cell. The tricarboxylic acid (TCA) cycle is operated in mitochondria, which is the final steps in oxidation of carbohydrates (reviewed in Sweetlove *et al*., 2010). The reduction potential from TCA cycle provided in nicotiamide dinucleotide (NADH) and succinate facilitates the generation of proton gradient across the mitochondrial inner membranes through the electron-transport chain (ETC, or respiratory chain). The proton gradient drives the ATP synthase, finally producing ATP from adenosine diphosphate (ADP) by oxidative phosphorylation (for a detailed review, Saraste, 1999; Schultz and Chan, 2001 and references therein). ATP thus produced in mitochondria is utilized as the main energy source around the cell.

Mitochondria contain important processes for the energy conversion of organisms and the dysfunction of these reactions can give rise to negative effect on the organisms' proliferation. Therefore, apicomplexan mitochondria have been studied as a promising target for chemotherapy for parasitic diseases (Kita *et al*., 2001; Kita *et al*., 2007; Mather *et al*., 2007; Sen and Majumder, 2008). Notably, it has been discovered that apicomplexan parasites have modified metabolic pathways in their mitochondria to survive in the host cells. *C. parvum* seems to lack conventional ETC complexes I-IV and instead has the cyanide-resistant, alternative oxidase (AOX) (reviewed in Mogi and Kita, 2010). One most recent study characterizes the unique, branched TCA cycle in *Plasmodium* mitochondria (Oslzewski *et al*., 2010). In fact, several drugs are reported to

show inhibitory effects on the apicomplexan mitochondrial ETCs. For example, ascofuranone, a potent inhibitor of mitochondrial AOX of the causative agent of African sleeping sickness *Trypanosoma brucei*, was shown to inhibit the *Cryptosporidium* AOX as well (Suzuki *et al*., 2004). Atovaquone and its analogs specifically bind to mitochondrial complex III and effectively stop the electron flow in *Plasmodium* (Barton *et al*., 2010). However, there are several problems like still high cost of drugs for poor people and the emergence of drug-resistant strain (Perry *et al*., 2009). Hence, it is still necessary to characterize the parasite's physiology and thereby to find additional drug targets.

1.3. Unique mt genome of apicomplexan parasites

Most mitochondria possess their own DNA called mitochondrial genomes (mt genomes). It is generally accepted that mitochondria are the reduced form of an endosymbiotic α-proteobacterial progenitor (endosymboitic theory, Sagan, 1967) and that the mt genomes are derived from the ancestor's genomic DNA. It is known that another notable feature of apicomplexan mitochondria is observed in their mt genomes. The mt genome is most commonly organized into a circular genome encoding 40 to 50 genes on average across the eukaryotic taxa (Burger *et al*., 2003). Mt genomes encode a series of genes (mt genes) mainly for respiratory ETC proteins, ATP synthase proteins, ribosomal RNA (rRNA), and transfer RNA (tRNA). These components are essential for mitochondrial ETC and mt gene expression, and the mt genes are actually transcribed followed by translation. The importance of the accurate expression of the mt genes has been emphasized, as several human diseases are associated with the mutations both in nuclear genes encoding the proteins for mitochondrial expression machinery and mt

genes (reviewed in Jacobs and Turnbull, 2005; Taylor and Turnbull, 2005).

In contast to the conventional mt genomes, the mt genomes of apicomplexans are generally comprised of 6- to 8-kb long linear molecules, the smallest of the known mt genomes (Hikosaka *et al*., 2010). The first completely sequenced mt genome of apicomplexa is the 6-kb element of *Plasmodium falciparum* (Feagin, 1992), which was later found to exist as a concatemer (Preiser *et al*., 1996). The apicomplexan mt genome element contains a strictly limited set of genes: only three protein-coding genes, *cox1*, *cox3*, and *cob*, and several fragmented rRNA genes. These mt genes are polycistronically transcribed with 3' oligoadenylation (Ji *et al*., 1996; Gillespie *et al*., 1999; Rehkopf *et al*., 2000). Curiously, the three protein-coding genes lack canonical start (AUG) and stop (UAA, UAG, UGA) codons in the 5' and 3' terminal regions, respectively (Conway *et al*., 2000; Rehkopf *et al*., 2000). Transfer RNA genes have not been detected in any of the apicomplexan mt genomes. The gene expression system for the apicomplexan mt genes is thus divergent and remains to be characterized well. Given the importance of the gene products and the probable difference in the expression mechanism between parasites and mammalian hosts, they can also be potential drug targets (Mather *et al*., 2007).

1.4. Uniformity of biological features of apicomplexans and related lineages

Some of the characteristic features of the apicomplexan mitochondria are shared in the related lineages. Phylogenetically, apicomplexans form a sister clade with dinoflagellates, and further with their outer group ciliates, and comprise the superphylum Alveolata (Cavalier-Smith, 1993; Figure 1-1). In particular, the sister lineage dinoflagellates share some features seen in mt genomes with apicomplexans. The dinoflagellate mt genomes encode the same set of genes as apicomplexans, and the

canonical start and stop codons are absent in protein-coding genes (this topic will be revisited in Chapter 4). Although investigations on each species of these lineages are being performed, it only can characterize the physiology only of the species among the diverse assemblage. Meanwhile, researchers have recently paid attentions to the ancestral species of apicomplexans and dinoflagellates, as it can be anticipated that a comprehensive understanding on physiology of many apicomplexan and dinoflaegllate species is obtained. The closest species to the common ancestor of the two groups described to date are a photosynthetic coral symbiont *Chromera* (Moore *et al*., 2008; Janouškovec *et al*., 2010; Figure 1-1) and a shellfish pathogen *Perkinsus* (Mackin *et al*., 1950; Levine, 1978; Figures 1-1 and 1-2). According to this close relationship, *Perkinsus* is expected as a model species that would characterize the parasitic apicomplexans' physiology, elucidating the evolution and function of mitochondria.

1.5. General description of a shellfish pathogen *Perkinsus*

All species belonging to the genus *Perkinsus* are parasites infecting shellfish such as clams, oysters and abalones and they have been found in American, European, Australian, and Asian coastal areas (Choi and Park, 2010; Table 1-1). The most studied species *P. marinus* was initially described as a parasitic fungus *Dermocystidium marinum*, causing the eastern oyster *Crassostrea virginica* "Dermo" disease in the Gulf of Mexico (Mackin *et al*., 1950). After reclassification as *Labyrinthomixa marina* (Mackin and Ray, 1966), ultrastructural study revealed that this species possessed an apical complex in its cell (Perkins, 1976), based on which this species was included in the phylum Apicomplexa. This parasite was renamed *Perkinsus marinus* under the class Perkinsea, by Levine (1978). It was later demonstrated that *Perkinsus* was more related to dinoflagellates than to apicomplexans, according to molecular phylogenetic

approaches (Reece *et al*., 1997; Siddall *et al*., 1997; Saldarriaga *et al*., 2003; Figure 1-1). More recently, Bråte *et al*. (2010) detected novel ribosomal DNA (rDNA) sequences possibly belonging to Perkinsea (including *Perkinsus* and another parasitic protist *Parvilucifera*) from marine and freshwater environments, indicating the underlying high biodiversity of the class. As this organism is isolated from shellfish worldwide threatening the fisheries with great economic lossess (Villalba *et al*., 2004), countermeasures are needed to prevent proliferating of the parasite. In Japan, also, *P. olseni* has been isolated from a Manila clam *Ruditapes phillipinarum* (Hamaguchi *et al*., 1998). A recent work by Shimokawa *et al*. (2010) demonstrated that juvenile *Ruditapes* clams showed mortality by infection with *P. olseni* prezoosporangia.

As suggested from the close relationship with apicomplexans, *Perkinsus* actually possesses several biological characters similar to those of apicomplexans. Regarding the physical structure, the presence of the apical complex structure observed in apicomplexans is evident in flagellated zoospores of *Perkinsus* (Perkins, 1976; Siddall *et al*., 1997; Leander and Keeling, 2003). As *Plasmodium* merozoites intrude into the host erythrocyte using this structure, *Perkinsus* invades hemocyte of the host shellfish (Sunila *et al*., 2001; Figure 1-3). These resemblances imply the possibility that they share metabolic and genomic features as well. The expressed sequence tag (EST) survey (mentioned later) and *in silico* analyses actually detect the important enzyme genes homologous to those of apicomplexans (Joseph *et al*., 2010; Mogi and Kita, 2010).

1.6. Advantages in using *Perkinsus* as a related organism to apicomplexans

In studying *Perkinsus* as an apicomplexan-related organism, there are several advantages; *Perkinsus* is much easier to deal with, than apicomplexans and dinoflagellates. The axenic *Perkinsus* culture without any host cell is available in which

Perkinsus trophozoites can proliferate by binary fission and schizogony (Gauthier and Vasta, 1995). On the other hand, *Plasmodium* cells collected from red blood cell cultures contain heme-derived hemozoin, which should be often removed before biochemical experiments (Kobayashi *et al*., 2007). Using this *in vitro* culture, pure *Perkinsus* cells are easily obtained with low cost, and this property makes it easy to collect enough volume of samples for biochemical studies. Furthermore, the transfection method for *Perkinsus* is already available (Fernández-Robledo *et al*., 2008a and 2008b) which will enable us to establish transformed cell lines with fluorescent signals. Molecular studies on *Perkinsus* are not so abundant currently, but due to its industrial and phylogenetical significance, the genome project for *P. marinus* is being undertaken by scientists at the J. Craig Venter Institute (formerly The Institute for Genomic Research, TIGR) and scientists at the Department of Microbiology and Immunology, University of Maryland School of Medicine/Institute of Marine and Environmental Technology (formerly at the Center of Marine Biotechnology). The study on the EST analysis has also just been published (Joseph *et al*., 2010), and available molecular data on *Perkinsus* have been increasing.

Moreover, it is thus anticipated that *Perkinsus* will provide insights into dinoflagellate physiology more easily than using dinoflagellates, in addition to the physiology of *Perkinsus* itself. Considering that related dinoflagellates include notorious harmful algal bloom species causing red tides and shellfish poisoning (Hallegraeff, 1993), investigations of these organisms should be urgently needed in the aspects of physiology, ecology, and toxicology. However, researchers have had difficulties in preparing fine protein samples from dinoflagellates, for cells contain large amount of interfering compounds such as DNA, polysaccharide (contained in thecal plates), and pigments. Although improved methods have been proposed recently (Lee

and Lo, 2008; Wang *et al*., 2009), these properties have been obstacles to smooth progress in dinoflagellate proteomics. In this respect, protein analysis of *Perkinsus* is not so difficult because the cell is not armored with thecal plates and seems to lack photosynthetic pigment.

As such, the apicomplexan relative *Perkinsus* i) is easier to use by virtue of the availability of *in vitro* culture where they axenically grow resulting in pure cell materials sufficient for biochemical studies assisted also by genomic data, and ii) can contribute to understanding of biological characters of a diverse species both from apicomplexans and dinoflagellates because of its ancestral position in the phylogenetic relationship. These properties are highly advantageous and are not found in apicomplexans and also dinoflagellates. Totally, *Perkinsus* is an ideal, advantageous organism in pursuing a comprehensive understanding of apicomplexans and dinoflagellates.

1.7. The purpose of this study

As mentioned above, it is of great significance to characterize the mitochondrial energy conversion system and mt gene expression system of the shellfish pathogen *Perkinsus*. However, most of current studies on *Perkinsus* do not focus on mitochondria and therefore neither annotated mt gene sequence nor a basic preparation method is currently available. In this background, I investigated mitochondria of *Perkinsus* in this study in order to build the foundation for studying mitochondrial metabolic and gene expression systems of this organism.

I first tested several methods to establish the basic protocol for preparation of enriched mitochondria from *P. marinus*. Then I performed BLAST-based searches in *P. marinus* draft genome databse to identify *Perkinsus* homologs for the mitochondrial

ETC protein and to outline the organization of its ETC. It was revealed that *Perkinsus* possessed many core subunits of mitochondrial ETC proteins. I also demonstrated the enzyme activities of ETC complexes by biochemical assays. Although the above-mentioned BLAST-based sequence searches did not detect *P. marinus* homologs for COX1, COX3 and cytochrome *b*, which are exclusively encoded by mt genomes in related organisms, I succeeded in determining the full-length messenger RNA (mRNA) sequence of one of the possible *P. marinus* mt genes. Analyzing the sequence, I realized that *Perkinsus* mt gene lacked canonical start and stop codons. Additionally, I discovered an unusual phenomenon in mitochondrial gene expression system; it was suggested that *Perkinsus* mt genes require unusually frequent frameshifts in translation, ten times per a gene.

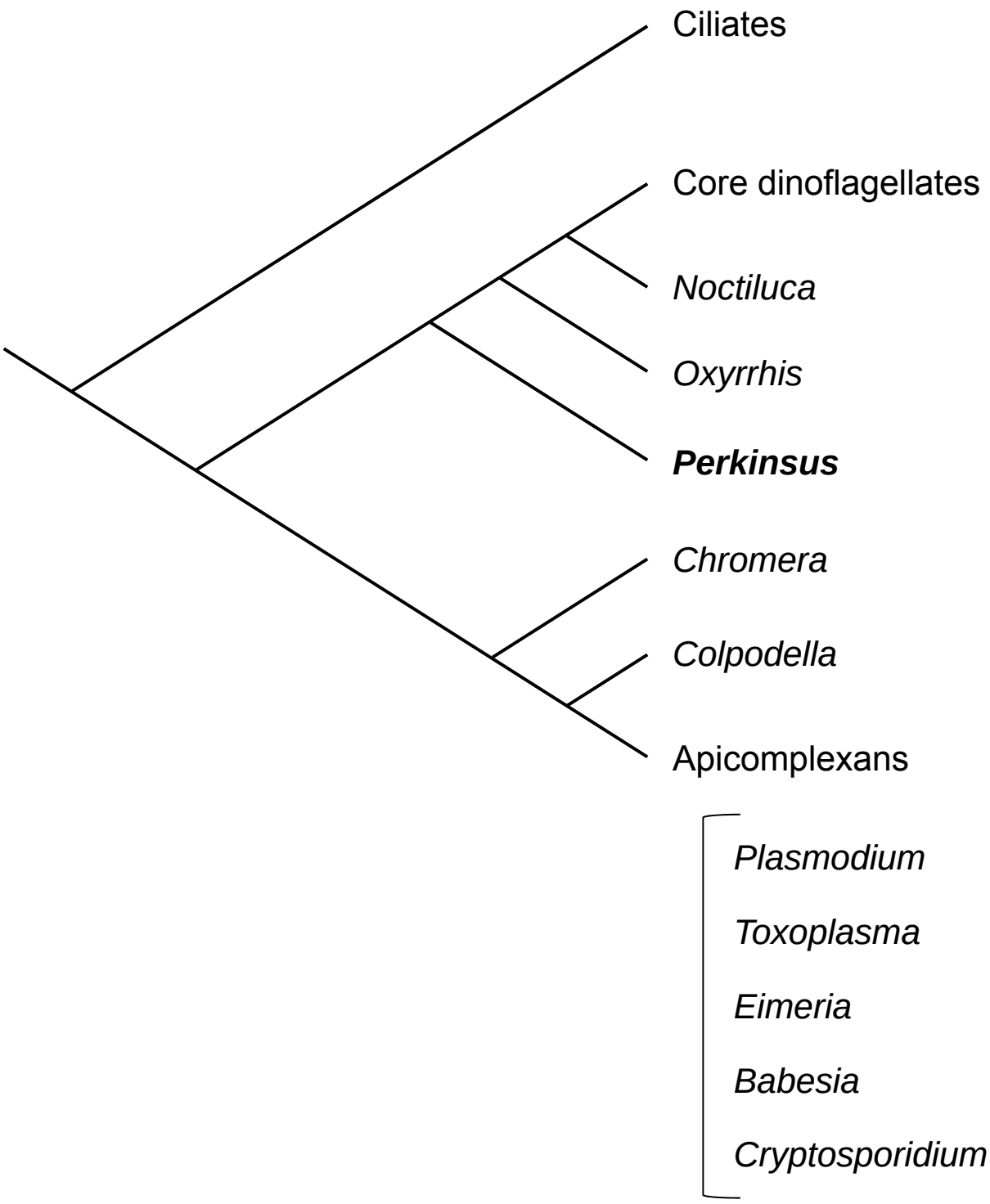

Figure 1-1.
Phylogenetic relationship of Alveolata.

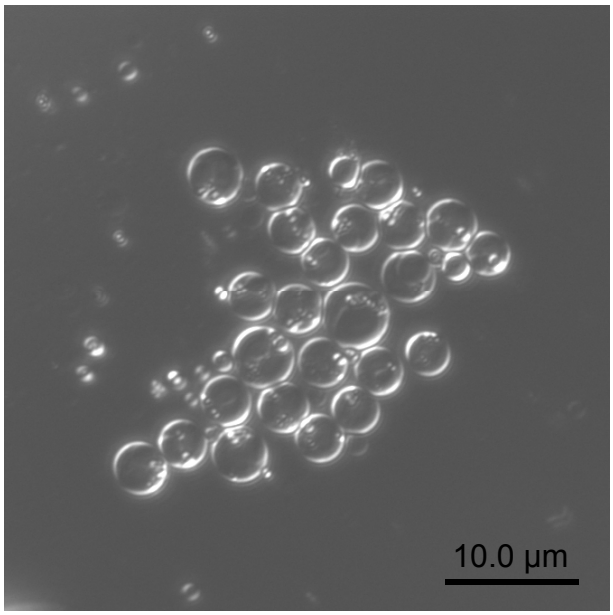

Figure 1-2.
Perkinsus marinus culture *in vitro*.

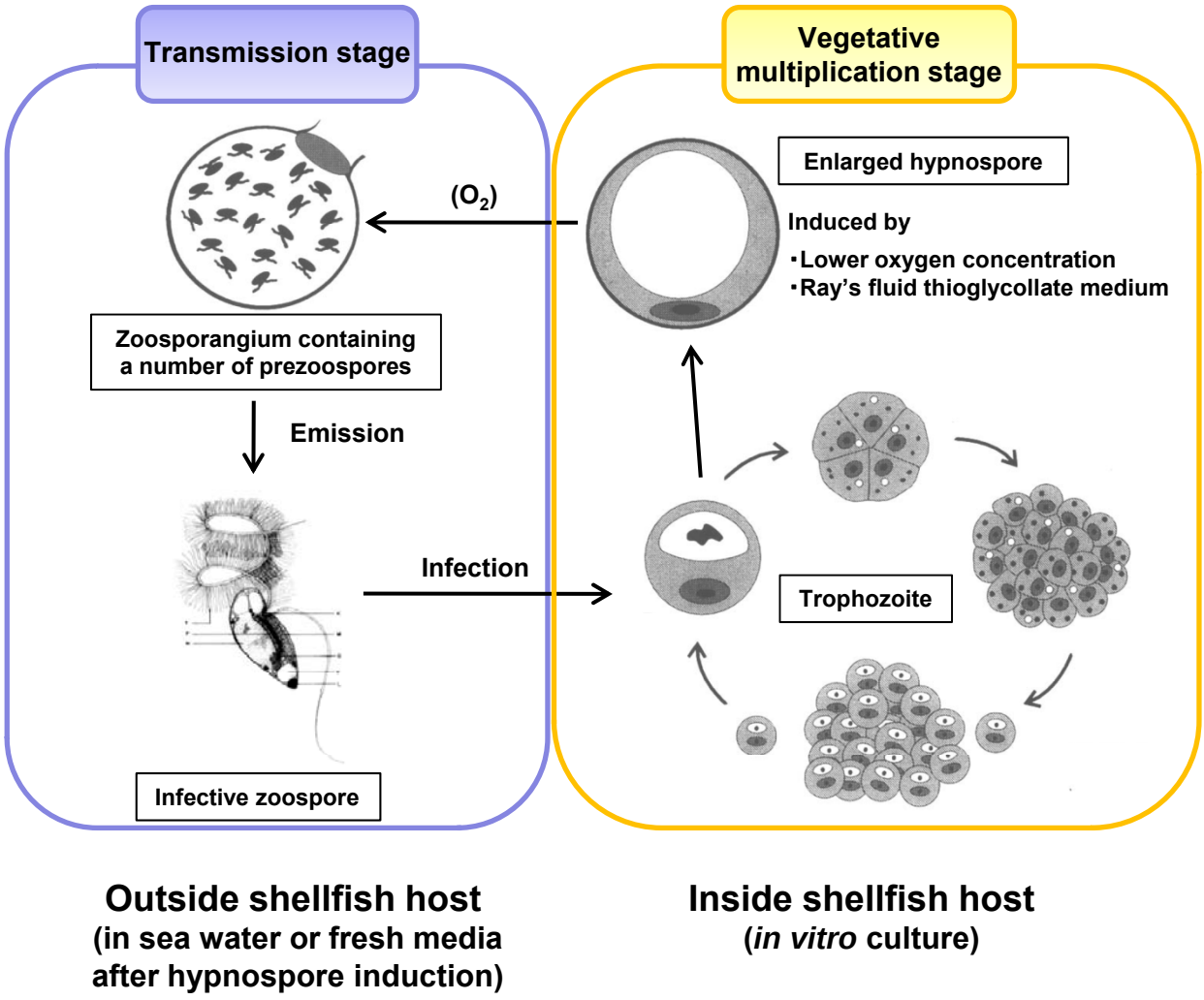

Figure 1-3.
Proposed life cycle of *Perkinsus*. Figures are adapted from Perkins and Menzel, 1967 (the drawing of zoospore) and Sunila *et al*., 2001 (others).

Chapter 2. PRELIMINARY ENRICHMENT OF *PERKINSUS* MITOCHONDRIA

2.1. Introduction

To specifically investigate properties of an organelle, the enrichment and/or purification of the organelle is preferably required. Whereas mitochondria enrichment/purification methods are available for several model higher eukaryotes, mitochondria preparation method for *Perkinsus* is presently not established; there are only a few studies on possible *Perkinsus* mitochondria-targeted genes (Wright *et al*., 2002; Schott and Vasta, 2003) and none on mitochondria themselves. Also, the presence of *Perkinsus* mitochondria is observed only microscopically (Sunila *et al*., 2001; Matsuzaki *et al*., 2008), but there is no report about biochemical evidence for the presence of *Perkinsus* mitochondria. Due to this situation, the biochemical evidence should be addressed and the enrichment method is preferably needed as a starting point to study *Perkinsus* mitochondria.

Several studies have been reported to purify protist mitochondria by subcellular fractionation have been performed to specifically analyze the property of mitochondria: *Chlamydomonas* (Eriksson *et al*., 1995), *Leishmania* (Horváth *et al*., 2002), *Tetrahymena* (Kobayashi and Endoh, 2005; Balabaskaran *et al*., 2010), *Toxoplasma* (Esseiva *et al*., 2004; Pino *et al*., 2010), and *Plasmodium* (Takashima *et al*., 2001; Kobayashi *et al*., 2007). Referring to these previous studies on other protists, I tried to enrich/purify *Perkinsus* mitochondria for molecular biological and biochemical studies in this chapter. *Perkinsus* cells were disrupted by the N_2 cavitation method, and mitochondria were partially enriched by Percoll density gradient ultracentrifugation, using the enzyme activity of a mitochondrial ETC protein as a marker.

2.2. Materials and methods

2.2.1. Strain and culture conditions

The *P. marinus* strain CRTW-3HE was purchased from the American Type Culture Collection (ATCC, no. 50439) and regularly maintained at 26°C in 10 mL of ATCC medium 1886 in a 25 cm^2 flask (BD Biosciences). Discontinued products were substituted as follows: Lipid Mixture (1000×; L5146; Sigma) replaced Lipid Concentrate (100×; 21900-014; Gibco), and Instant Ocean Sea Salt (Aquarium Systems) replaced artificial seawater (S1649; Sigma). For mass culture, cells were cultivated in 200 ml media in a 500 ml Erlenmeyer flask, shaken at 150 rpm.

2.2.2. Preparation for enriched *Perkinsus* mitochondria

The *Perkinsus* cells were grown up to 1×10^8 cells/ml in 200 ml culture and were collected by centrifugation at 800 × *g* for 5 min. This collection method was used unless otherwise stated. To mildly disrupt the cells and obtain intact mitochondria, a chemical disruption method using 0.05% or 0.1% (w/v) digitonin was first tried, which is employed in obtaining *Toxoplasma* mitochondria (Esseiva *et al*., 2004). Digitonin-treatment was performed also in combination with either 0.1% or 1.0% (w/v) cellulase, as *Perkinsus* cells are surrounded by the cell wall (La Peyre *et al*., 1993). Also, several physical disruption means were tried: glass-teflon homogenization, the homogenization by vortex in a solution containing glass beads, and the N_2 cavitation method. Among the three methods, the most preferable *Perkinsus* organellar fraction was obtained by the N_2 cavitation method (see 2.3. Results for details), and the method was optimized for *Perkinsus* because it was assumed that there should be differences in cellular properties between *Perkinsus* and *Plasmodium*. In the previous studies on *Plasmodium* (Takashima *et al*., 2001 and Kobayashi *et al*., 2007), the N_2-dissolving

pressure was 1200 psi. However, with this pressure, almost no crude organellar fraction was obtained, which indicated most of organelles including mitochondria are also burst. Eventually the N_2 pressure of 300 psi was selected to disintegrate the cells as mildly as possible. Also, by the microscopic observation, mannitol-sucrose buffer with ethylenediaminetetraacetic acid (EDTA) was found hypotonic to the *Perkinsus* cells, therefore 600 mM sucrose buffer was used. After these optimizations, the enrichment of mitochondria was performed as follows (summarized in Figure 2-1).

The *Perkinsus* cells from the 200 ml mass culture were collected, washed with 2 × phosphate buffered saline, and disrupted using 4639 Cell Disruption Bomb (Parr Instrument Company) at 300 psi for 20 min at 4°C. After homogenization, the cell lysate was incubated on ice with 5 µl (125 U) of Benzonase (Novagen) in the presence of 5 mM $MgCl_2$ for 20 min to remove DNA and RNA which are the main cause of aggregate formation during the ultracentrifugation. After incubation, cell debris and nuclear fraction were removed by centrifugation at 1,000 × *g* for 10 min at 4°C, followed by collecting crude organellar fraction by centrifugation at 5,000 × *g* for 20 min at 4°C. A 200 µl of organellar fraction was ultracentrifuged in 8.5 ml of 28% (v/v) Percoll solution (Percoll PLUS, GE Healthcare) adjusted to 600 mM with 2.5 M sucrose solution, at 23,000 rpm (50,000 × *g*) for 1 h at 4°C using Himac CP80WX with RP55T rotor (HITACHI). The ultracentrifuged sample was separated into 24 fractions (almost 350 µl each) using a peristaltic pump (PST-100, Iwaki), from bottom to top. The protein concentration of each fraction was determined using Protein Assay kit (BIO-RAD) according to the manufacturer's instruction with bovine serum albumin (BSA) as a standard.

2.2.3. Enzyme activity assay

As a marker of mitochondria, the succinate-dehydrogenase (SDH) activity of the mitochondrial ETC complex II was measured by monitoring the absorbance change of 2-(4,5-dimethyl-2-thiazolyl)-3,5-diphenyl-2*H*-tetrazolium bromide (extinction coefficient = 17 mM^{-1} cm^{-1}) at 570 nm in the presence of 120 μg/ml phenazine methosulfate and 2 mM KCN (the specific inhibitor of the complex IV), in 100 mM Tris-HCl buffer pH 8.0 (Kita *et al.*, 1989; Miyadera *et al*., 2003). Assays were conducted in 1 ml solution at 25°C using the UV-3000 spectrophotometer (Shimadzu).

2.3. Results

The *Perkinsus* cell membranes and cell walls were hardly disrupted by the digitonin-treatment, irrespective of the presence or absence of cellulase. Glass-teflon homogenization did not disrupt *Perkinsus* cells at all, probably because of the small size of *Perkinsus* trophozoites. Meanwhile, the homogenization by vortex in a solution containing glass beads disrupted the cells intensely, resulting in the small amount of crude organellar fraction. The only method among those tested so far that might yield the appropriate organellar fraction was N_2 cavitation method used also for preparation of *Plasmodium* mitochondria (Takashima *et al*., 2001 and Kobayashi *et al*., 2007). The *Perkinsus* organellar fraction obtained by this method was ultracentrifuged and fractionated, and the SDH activity was measured on each fraction.

Figure 2-2 is one example of the result of fractionation, showing the distribution of the SDH activity and protein concentration of each fraction. The fraction No. 7 contained the largest amount of proteins (135 μg/ml), and the fractions No. 6 and Nos. 8-14 showed high protein concentration ranging from 94.7 to 135 μg/ml. Other fractions contained 45 μg/ml on average. The highest total SDH activity was observed in fractions Nos. 12 and 14 (0.294 nmol/min) followed by the fraction No. 13 (0.235

nmol/min). The fractions Nos. 10, 11 and 15 also showed high activity (176, 176, 118 nmol/min, respectively). Higher activities were also observed from fractions Nos. 18, 19, 21-24. Higher specific activities were detected from the fractions Nos. 12-14 (0.0993, 0.0787, 0.103 μmol/min/mg of protein). The fractions Nos. 18-21, 23, and 24 showed high specific activities ranging from 0.0762-1.47 μmol/min/mg of protein.

2.4. Discussion

2.4.1 Partial enrichment of *Perkinsus* mitochondria

From *Perkinsus* organellar ultracentrifugal fractions, the obvious enzyme activity of the mitochondrial ETC protein was detected. This is the first finding that *Perkinsus* possesses mitochondria that contain an active ETC. The fractions Nos. 6-14 contained the higher amount of proteins, and the higher total SDH activity was observed in fractions Nos. 10-15. Also, higher specific activities were detected from the fractions Nos. 12-14. The higher enzyme activities are also observed in the fractions Nos. 19-24, but they are probably attributed to some unknown contaminants because these fractions contained low amount of proteins. Based on these observations, it was indicated that mitochondria in the crude organellar fraction are mainly concentrated in the fractions Nos. 12-14. In other words, the partial enrichment of *Perkinsus* mitochondria has been accomplished. The SDH activity showed relatively a broader distribution, and it was suggested that this protocol needs more refinements and the verification by additional methods described below, to establish a purification protocol.

2.4.2. Toward further refinements of the purification protocol

Firstly, the intactness of mitochondria must be ensured especially for analysis on molecules functioning in mitochondria. The subcellular fractionation using the density

gradient centrifugation is performed on the assumption that the organelle of interest has a uniform density. As the density of organelles will be changed if it is broken, the intact organelles should be prepared for fractionation. One good marker is the cytochrome *c* residing in the mitochondrial intermembrane space; its immunoblot detection will support the intactness of mitochondria. I already confirmed the cross-reactivity of the anti-plant cytochrome *c* antibody on *Perkinsus* organellar fraction.

Secondly, the specific distribution of mitochondria will be further endorsed by some other markers for mitochondria and other organelles. The SDH activity can be detected independently of the presence of the mitochondrial membranes because this activity only depends on the soluble parts of complex II (Fp and Ip subunits). On the other hand, the succinate-ubiquinone reductase (SQR) activity requires all the complex II subunits including the membrane-anchor subunits (CybL and CybS). Therefore, the SQR activity is a better marker for mitochondrial membranes. The distribution of mitochondria is also confirmed by mitochondria-specific proteins which can be detected by immunoblot assays. So far, I have confirmed that the antibody raised against a plant mitochondrial alternative oxidase (AOX) can crossreact on *Perkinsus* organellar fractions by immunoblot. For purification, it should be ensured that other organelles like plastids and endoplasmic reticulum are not co-distributed with mitochondria. Additionally, parameters during fractionation like Percoll concentration, the osmolarity of the solution, and the centrifugation time and speed, should be rearranged. The reproducibility of the distribution must be ensured as well.

Lastly, I will use the transformed *Perkinsus* cell lines with green fluorescent protein (GFP) signal in organelles to evaluate the homogenization and distribution results visually. Presently I cannot confirm the presence of mitochondria outside the cells after homogenization by MitoTracker-staining. This implies that mitochondria might be also

disrupted by homogenization, or that my experimental condition is unsuitable for staining. To obtain stable fluorescent signals in organelles, *P. marinus* strain CRTW-3HE transformed with the *Perkinsus*-specific vectors (constructed by Dr. M. Matsuzaki, modified from Fernández-Robledo *et al*., 2008a) would be useful. We have a transformed *P. marinus* strain with pMOE-mfGFP vector containing the plastid-targeted gene *ispC* (Matsuzaki *et al*., 2008), together with multi-functional GFP gene (Kobayashi *et al*., 2008) as a marker. The GFP signals make it easy to see if the cell is broken (Figure 2-3). Similarly, I am now preparing additional strains which have GFP signals in mitochondria. I prepared transformation vectors ligated with each of two mitochondria-targeted genes, AOX, elongation factor Tu and transformation has been carried out using them. GFP signals were observed as a network extending all around the cell in transformed *Perkinsus* cells, which possibly indicated the mitochondria. These cell lines are now being established, and they would greatly help us confirm the efficiency in homogenization and organellar distribution in subcellular fractionation.

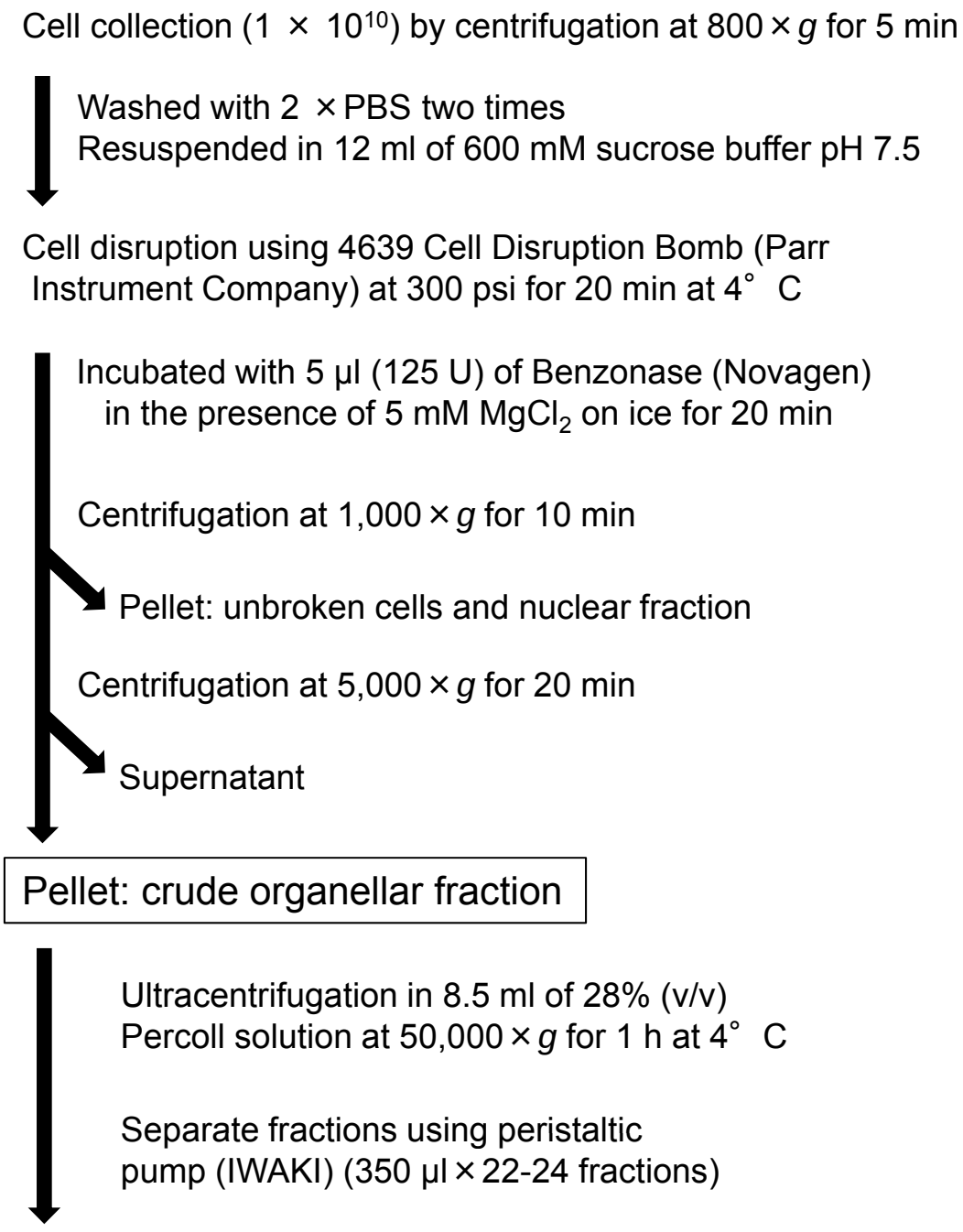

Figure 2-1.
Summary of the procedure of mitochondrial fractionation arranged for *Perkinsus.*

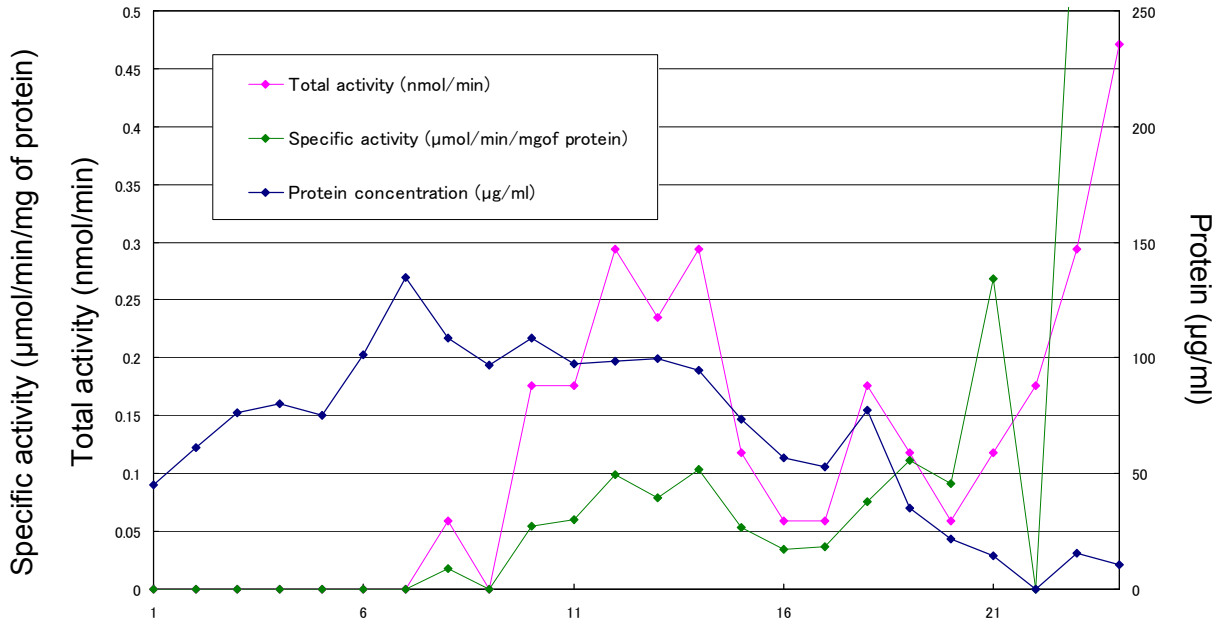

Figure 2-2.
Distribution of the SDH activity and the protein concentration in the ultracentrifugal fractions.

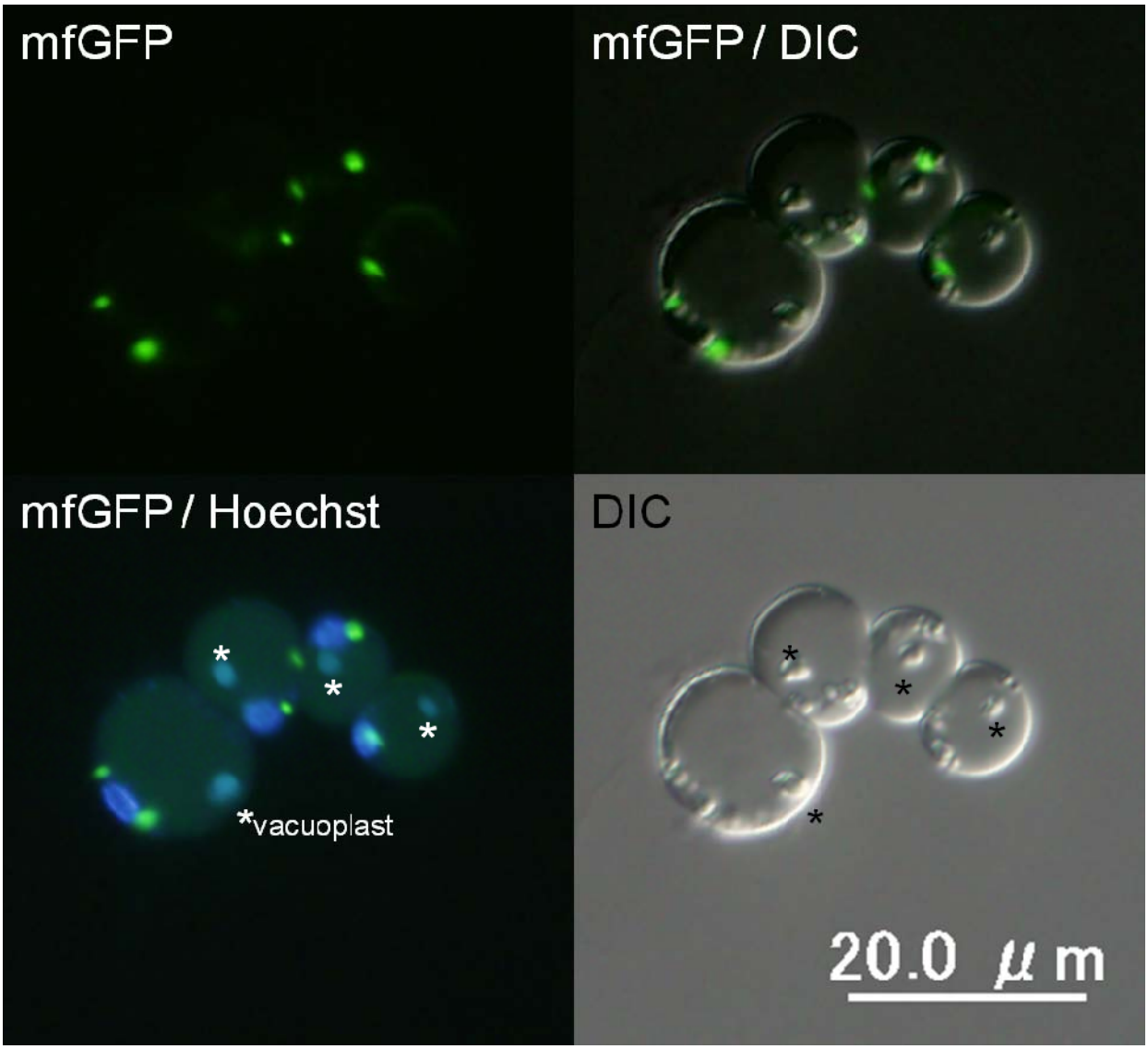

Figure 2-3.
Differential interference contrast image (bottom right) and fluorescent images of *P. marinus* strain *ispC*/mfGFP (pictures by the courtesy of Dr. M. Matsuzaki). Note that the dotted GFP signals are observed around the cells, probably localizing to the plastids. Blue signals in the bottom left panel is nuclear DNA stained with Hoechst dye.

Chapter 3. GENOMIC AND BIOCHEMICAL CHARACTERIZATION OF *P. MARINUS* MITOCHONDRIAL ETC PROTEINS

3.1. Introduction

3.1.1. General description of the mitochondrial ETC

Mitochondria take an important role in energy production in cells. Especially, the function of the ETC directly contributes to the energy production by generating proton gradient, which in turn, drives the ATP synthase (Figure 3-1; summarized in Schultz and Chan, 2001). Generally, the mitochondrial ETC is composed of four membrane-bound complexes. The ETC complex I is the NADH dehydrogenase (NDH), which catalizes the oxidation of NADH generated in the TCA cycle and glycolysis, and reduces quinone to quinol. The complex II is called SDH or SQR. This complex mediates the dehydrogenation of succinate to fumarate as a part of the TCA cycle (SDH activity), and transfer the electron from the reaction to quinone (SQR activity). This is the only enzyme which is involved both in the ETC and the TCA cycle. The electron in quinol is then transferred to the ubiquinol-cytochrome *c* reductase (QCR), the complex III, and further transferred to cytochrome *c* in the mitochondrial intermembrane space. Finally, the cytochrome *c* oxidase (COX), the complex IV, receives the electron from cytochrome *c*, and reduces the oxygen to water.

During this electron-flow, protons are translocated from mitochondrial matrix to the mitochondrial intermembrane space by the proton-motive force of the complexes I, III and IV. The proton gradient across the mitochondrial innermembrane described above, enables the oxidative phosphorylation; the ATP synthase phosphorylates ADP to ATP which is utilized as the energy source around the cell. The genes for subunits of these ETC complexes are coded in both nuclear and mt genomes (Burger *et al.*, 2003), and the

nuclear-genome-encoded components are transported into mitochondria from cytosol after translation, by virtue of the mitochondrial targeting sequence (Neupert and Herrmann, 2007).

3.1.2. Modified ETCs found in apicomplexan parasites

The conventional organization of the ETC mentioned above is observed in many eukaryotic species, and a various types of modifications are also found in diverse organisms. One divergent example in mitochondrial ETCs is the participation of the type II NADH dehydrogenase (NDH2) instead of the complex I, found sporadically from several prokaryotes, and mitochondria of fungi, plants, and protists (Melo *et al*., 2004). NDH2 oxidizes NADH and transfers electrons to quinones, but is not involved in proton pumping (Figure 3-2). Distinct from the conventional multi-subunit, rotenone-sensitive NDH, the NDH2 consists of a single subunit, and is resistant to rotenone. Another example is the cyanide-resistant AOX found in some plants, fungi, and protists such as *Trypanosoma* (Moore and Albury, 2008). This enzyme acts as a ubiquinol-oxygen oxidoreductase; it receives electrons from quinols and generates water by reducing oxygen as a terminal oxidase of the ETC (Figure 3-2). It was demonstrated that the inhibition of *Trypanosoma* AOX activity led to the parasite death, which indicated that this enzyme may take an essential role for survival (Minagawa *et al.*, 1997).

In apicomplexan parasites, also, these unconventional enzymes are reported to reside in the mitochondrial ETCs. *Plasmodium* possesses the NDH2 at the external surface of the mitochondrial innermembrane (Biagini *et al*., 2006) and the *Toxoplasma* NDH2 is suggested to be involved in maintaining the mitochondrial membrane potential (Lin *et al.*, 2009). *C. parvum* homologs of the NDH2 and AOX are found in its

complete genome sequence (Abrahamsen *et al.*, 2004) and the involvement of AOX in the parasite's survival is suggested (Roberts *et al.*, 2004). These observations imply that *Perkinsus* might have homologs of these enzymes in addition to the common enzymes.

So far, the report on the EST analysis of *P. marinus* does not contain any discussion about the mitochondrial ETC at all (Joseph *et al.*, 2010) and only a few *Perkinsus* ETC proteins are mentioned in published documents (Mogi and Kita, 2010). There is no report on sequences of *Perkinsus* mt genome which is expected to encode some of the mitochondrial ETC proteins. Because of this situation, I first conducted sequence searches from *P. marinus* draft genomic data using the basic local alignment search tool (BLAST) to pursue ETC protein sequences supposedly encoded both in nuclear and mt genomes. Then I performed biochemical assays on some of the mitochondrial ETC enzymes found, and confirmed the activities of several ETC proteins in *Perkinsus*.

3.2. Materials and methods

3.2.1. Sequence search of *Perkinsus* mitochondrial ETC proteins

To obtain *Perkinsus* mitochondrial ETC enzyme sequences, BLASTp-based searches in the public database at the National Center for Biotechnology Information (NCBI; http://blast.ncbi.nlm.nih.gov/Blast.cgi) were performed using the amino acid sequences of ETC complex proteins from related species as queries (*P. falciparum*, a dinoflagellate *Karlodinium micrum*, a diatom *Thalassiosira pseudonana*, and an apicomplexa *C. parvum*). The searches were conducted regarding homologs of major subunits of mitochondrial complexes I-IV and F_OF_1-ATP synthase, and of cytochrome *c*, NDH2, and AOX. Accession numbers of the query sequences are described in Table 3-1. After obtaining mitochondrial ETC protein candidates, the probability of export to mitochondria was estimated on each protein using MitoProt II v1.101 (Claros and

Vincens, 1996; http://ihg2.helmholtz-muenchen.de/ihg/mitoprot.html).

3.2.2. Enzyme activity assay of *Perkinsus* mitochondrial ETC complexes

The samples for enzyme activity assays were prepared as follows. The *P. marinus* strain CRTW-3HE cells were grown up to 1×10^8 cells/ml in 200 ml culture and were collected and washed. 1×10^{10} cells were resuspended in 12 ml of 600 mM sucrose buffer containing 0.1 mM EDTA, 2 mM tris(hydroxymethyl)amino methane (Tris)-HCl pH 7.5 and 1 mM phenylmethylsulfonyl fluoride, and subsequently homogenized using 4639 Cell Disruption Bomb (Parr Instrument Company) at 1000 psi for 1 h at 4°C. Crude organellar fraction was obtained as described in Chapter 2. Among the mitochondrial ETC complexes for which the subunit genes were found, biochemical assays were carried out for enzyme activities of complexes II-IV on the organellar fraction. The protein concentration was determined using Protein Assay kit (BIO-RAD) according to the manufacturer's instruction with BSA as a standard and specific activities were calculated using the concentration.

The SDH activity of the complex II was measured according to the method described in Chapter 2. All assays were conducted in 1 ml solution at 25°C using the UV-3000 spectrophotometer (Shimadzu) unless otherwise stated.

The SQR activity of the complex II was measured in 50 mM potassium phosphate buffer pH 7.5 by monitoring the absorbance change of 50 μM 2,6-dichlorophenolindophenol (extinction coefficient = 21 mM^{-1} cm^{-1}) at 600 nm in the presence of 60 μM ubiquinone-2 and 2 mM KCN (Kita *et al.*, 1989; Miyadera *et al.*, 2003).

The activity of succinate-cytochrome *c* oxidoreductase, i.e., the electron-flow from succinate to cytochrome *c* mediated by the complexes II and III was measured. This

assay was performed by tracing the absorbance change of cytochrome *c* (extinction coefficient = 19 mM^{-1} cm^{-1}) at 550 minus 540 nm, in 30 mM potassium phosphate buffer pH 7.5, in the presence of 2 mM KCN. After tracing for a few minutes, 3 mM malonate, the competitive inhibitor of the complex II, was added. The reactions of these three assays were started by the addition of 10 mM succinate to the reaction mixture (modified from Trounce *et al*., 1996).

The quinol oxidase activity, which is the combined activity of the complexes III-IV and the AOX, was also measured. This assay was performed in 100 mM Tris-HCl buffer pH 7.4, by recording the absorbance change of ubiquinol-1 (extinction coefficient = 15 mM^{-1} cm^{-1}) at 278 nm (Matsumoto *et al.*, 2006). After monitoring for five minutes, 200 nM ascofuranone was added to the mixture to inhibit the AOX, and monitoring the absorbance change was resumed.

The COX activity was measured as the O_2 uptake ratio in a 300 µl buffer containing 125 mM sucrose, 65 mM KCl, 10 mM Hepes-KOH (pH 7.2), 5 mM $MgCl_2$, 2 mM potassium phosphate, 0.5 mM ethylene glycol tetraacetic acid (Vercesi *et al*., 1998; Uyemura *et al*., 2004), using MicroxTX3 Trace micro oxygen sensor (Presense GmbH). 0.1 µg antimycin A and 200 nM ascofuranone were added to the reaction mixture, which are the specific inhibitors for complex III and AOX, respectively. The reaction was initiated by the addition of 300 µM *N*,*N*,*N*',*N*'-tetramethyl-*p*-phenylenediamine dihydrochloride and 500 µM ascorbate. After monitoring, the inhibition of the activity by the addition of 2 mM KCN was examined.

3.3. Results

3.3.1. Repertoire of *P. marinus* mitochondrial ETC proteins

First, sequence searches were performed using the *P. falciparum* sequences as

queries. These BLASTp-based searches yielded many homologous sequences of conventional ETC complex proteins with quite low E-value (Table 3-1): Fp, Ip (complex II), Rieske FeS, cytochrome c_1, hinge protein, MPP (complex III), COX2N, COX2C, COX5, COX6, COX7 (complex IV), and cytochrome *c*. Also, the homologs for core subunits of F_OF_1-ATP synthase, α, β, γ, δ, ε and OSCP, were found. The nomenclature of these subunits of the ATP synthase is according to mitochondrial homologs (Devenish *et al.*, 2008), not to bacterial homologs. Homologs of subunits of conventional muti-subunit NDH were not found, but the homolog of NDH2 was found instead. As *Plasmodium* has no AOX homolog, the *C. parvum* AOX sequence was used as a query and a homolog in *Perkinsus* was found. Meanwhile, CybL and CybS homologs were not identified by searches utilizing *T. pseudonana* sequences as queries instead, because of the absence of obvious homologs in *Plasmodium*. Curiously, none of the homologs generally encoded on mt genomes in apicomplexans and dinoflagellates (COX1, COX3, and cytochrome *b*) were found using either of *Plasmodium* and *Karlodinium* sequence. In total, twenty mitochondrial ETC protein homologs were discovered.

MitoProt can estimate the probability of mitochondrial targeting by analyzing the amino acid sequences in the N terminal regions of proteins. Twelve out of twenty proteins showed high probability of mitochondrial targeting (0.7940-0.9981). Meanwhile, other three exhibited moderate scores (0.3207-0.4836), and the remaining five were with rather low scores (0.000-0.1760).

3.3.2. Enzyme activities of *P. marinus* mitochondrial ETC complexes

Using the protein concentration and the total enzyme activity, the specific enzyme

activity on the crude organellar fraction was determined (summarized in Table 3-2). The SDH and SQR activities which were attributed to the complex II, was 215 and 169 nmol/min/mg of protein, respectively. The succinate-cytochrome *c* reductase activity was 292 nmol/min/mg of protein. The addition of malonate reduced the activity to 49.4 nmol/min/mg of protein with 83.1% inhibition rate. The quinol oxidase activity was 12.6 nmol/min/mg of protein. After the addition of ascofuranone, the activity went down to 7.22 nmol/min/mg of protein with 58.0% inhibition rate. The cytochrome *c* oxidase activity was 9.47 nmol/min/mg of protein, and this activity was completely inhibited by the addition of KCN.

3.4. Discussion

3.4.1. *Perkinsus* has an active mitochondrial ETC

Twenty *Perkinsus* protein homologs were found which constitute the core of the mitochondrial ETC complexes by BLAST-based sequence searches (Table 3-1). They were the NDH2, 2 subunits of the complex II, 4 subunits of the complex III, cytochrome *c*, 5 subunits of the complex IV, the AOX, and 6 subunits of the F_OF_1-ATP synthase. Whereas the twelve mitochondrial ETC protein candidates showed high probability of mitochondrial localization by prediction using MitoProt, the values of the other eight candidates were not so high (Table 3-1). One possible reason for this is that the annotation regarding the N termini of the eight proteins is not correct and the targeting signals reside in true N termini. Alternatively, these eight genes possess non-N-terminal targeting signals. C terminal targeting signal has been reported (Lee *et al*, 1999) and the involvement of the internal sequence in targeting to mitochondria and hydrogenosome (mitochondria-related organelle) has also been observed (Fölsch *et al*., 1996; Mentel *et al*., 2008). Although there is no program available for predicting such non-N-terminal

signals, the eight proteins might be translocated into mitochondria by such signals.

The homologs for Fp and Ip subunits of the complex II were found, and the SDH activity that was attributed to the two subunits was confirmed. The SQR activity, another activity of the complex II was also clearly observed. Also, the electron flow from succinate to cytochrome *c* via the complexes II and III was detected as succinate-cytochrome *c* oxidoreductase activity, and the activity was greatly inhibited by the addition of malonate, the competitive inhibitor of the complex II. These three kinds of activities involving the complex II were rather higher than those of *Plasmodium* (Kawahara *et al.*, 2009) and even than cultured human cells (Tomitsuka *et al*., 2003), although the values in these references were obtained using purified mitochondria. Though the physiolosical significance of these high activities remains unclear, it is noteworthy that the high activities can be utilized as good markers for *Perkinsus* mitochondria.

Unexpectedly, candidates for CybL and CybS the membrane anchor subunits of the complex II, were not detected with E-value < 0.01. In contrast to the SDH activity, the SQR activity which transfers the electrons from succinate to the quinol in the intermembrane space requires all the four subunits of the complex II (Nakamura *et al.*, 1996). The transmembrane proteins are said to be more divergent among species, and the highly divergent candidates for CybL and CybS are suggested for *Plasmodium* (Kawahara *et al*., 2009; Mogi and Kita, 2009). To prove the suggested presence of these two subunits, it should be needed to purify the complex II from *Perkinsus*, *Plasmodium* and/or related species and to determine the amino acid sequences.

The oxidase activity was also experimentally confirmed. Although candidates for two subunits were not discovered, the COX activity was detected. It was verified that the activity was completely inhibited by the addition of KCN, which is the shared

feature of the conventional COX. The quinol oxidase activity indicated that the AOX and the complexes III-IV were active in *Perkinsus* mitochondrial ETC. By the inhibition assay using the AOX-specific inhibitor ascofuranone, it was confirmed that the contribution ratio of the AOX to quinol oxidation was apparently around 40% in my experimental condition. The AOX does not have the proton-pumping activity and is not involved in the ATP generation (Moore and Albury, 2008). In other organisms, it is known that AOX is expressed selectively for functions other than energy generation. Each gene of the AOX families are expressed in a stress-induced manner (Clifton *et al*., 2006), and the AOX is actively expressed and seems to participate exclusively in the oxidation of reducing equivalents generated in glycolysis in bloodstream forms *Trypanosoma brucei* and possibly in microsporidia (Chaudhuri *et al.*, 2006; Williams *et al*., 2010). For the elucidation of the physiological role of *Perkinsus* AOX, the gene expression level and the enzyme activity of AOX should be investigated using parasites at other stages in its life cycle such as zoospore.

As far as I investigated here, the complexes II-IV of the *Perkinsus* mitochondrial ETC are active, and their functions would probably produce the driving force of the F_OF_1-ATP synthase. The presence of alternative enzymes NDH2 and AOX supports the similarity in the mitochondrial ETC organization of apicomplexans and *Perkinsus*.

3.4.2. Missing *Perkinsus* mt genes should be present and expressed

As with CybL and CybS subunits I discussed above, I did not find any homologous sequences for COX1, COX3, and cytochrome *b* in the sequence searches (Table 3-1). Now that the enzyme activities of the complexes III and IV are experimentally demonstrated, *Perkinsus* is likely to have the the three ETC elements. Especially, cytochrome *b* and COX1 are essential components of complex III and IV, respectively.

Cytochrome *b* directly participates in the oxidation of the quinol: it binds the quinol and transfers the electrons using hemes (Crofts, 2004). COX1 is also the highly important element, as this subunit contains the heme *a*3-Cu_B binuclear center reducing molecular oxygen to water using electrons from cytochrome *c*, and transports protons from the mitochondrial matrix to the intermembrane space (Iwata *et al*., 1995). These functional importance strongly indicated the presence of the missing subunits of *Perkinsus* mitochondrial ETC complexes. Intriguingly, the three missing subunits are encoded exclusively in mt genomes in apicomplexans and dinoflagellates (coded as *cox1*, *cox3*, and *cob*, respectively). Taken all together, it is appropriate to anticipate that *Perkinsus* mt genome sequences were not retrieved from the public database by BLAST-based searches.

I have two possible reasons for this situation. The first possibility is that the mt genome data of *Perkinsus* have not been added to the public database. Especially, it is sometimes observed that the nucleotide sequences with high AT content are a little hard to be cloned in large-scale genome sequencing (Glöckner, 2000; Gardner, 2001). Considering that mt genomes are generally AT-rich, it can be expected that the *Perkinsus* mt genome is highly AT-rich and its most parts may not be sequenced. The second is that *Perkinsus* mt gene sequences are quite divergent and therefore were not found by BLAST-based sequence searches.

In either case, the three genes must be present and expressed. To solve this issue, I will have to examine carefully the candidate sequences with lower sequence similarity to the orthologs regarding the conservation of essential residues and/or perform molecular biological experiments like polymerase chain reaction (PCR). In the next chapter, I investigated the *P. marinus* mt genome sequence and characterized one of the *P. marinus* mt genes, *cox1*.

Enzyme	Subunit	Query	RefSeq ID of *P. marinus* homolog	E-value	Mitochondrial targeting score by MitoProt
Type I NADH dehydrogenase (Complex I)	NAD1-5	*	none	-	-
Type II NADH dehydrogenase	-	*Pf*	XP_002784460	6.00E-50	0.3207
Succinate dehydrogenase (Complex II)	Fp	*Pf*	XP_002784717	0.00	0.8775
	Ip	*Pf*	XP_002765391	1.00E-103	0.9635
	CybL	*	none	-	-
	CybS	*	none	-	-
Ubiquinol-cytochrome *c* oxidoreductase (Complex III)	cyt b^{Ψ}	*	none	-	-
	Rieske FeS	*Pf*	XP_002778583	1.00E-90	0.7940
	cyt c_1	*Pf*	XP_002787441	6.00E-95	0.0245
	hinge	*Pf*	XP_002775032	1.00E-17	0.1341
	MPP	*Pf*	XP_002780898	4.00E-122	0.9333
Cytochrome *c*	-	*Pf*	XP_002769625	2.00E-39	0.1760
Cytochrome *c* oxidase (Complex IV)	$COX1^{\Psi}$	*	none	-	-
	COX2N	*Pf*	XP_002767155	4.00E-37	0.8331
	COX2C	*Pf*	XP_002788770	3.00E-60	0.4836
	$COX3^{\Psi}$	*	none	-	-
	COX5	*Pf*	XP_002769485	3.00E-57	0.9981
	COX6	*Pf*	XP_002771317	5.00E-28	0.0577
	COX7	*Pf*	XP_002780296	6.00E-07	0.0000
Alternative oxidase	-	*Cp*	XP_002788619	8.00E-72	0.9040
ATP synthase	α	*Pf*	XP_002780357	0.00	0.9954
	β	*Pf*	XP_002782393	0.00	0.8239
	γ	*Pf*	XP_002778693	3.00E-60	0.9490
	δ	*Pf*	XP_002785641	4.00E-23	0.9170
	ε	*Pf*	XP_002776346	5.00E-09	0.4348
	OSCP	*Pf*	XP_002778045	5.00E-22	0.9975

Table 3-1.
Candidates for major ETC complex proteins and ATP synthase of *Perkinsus* retrieved from the public database. Only the sequences with the highest alignment score among the BLAST search results are shown for each protein. Query sequences from *P. falciparum* (*Pf*) are listed with accession numbers in Mogi and Kita, 2009. "none" means that there were no hits with E-value less than 0.01 by sequence searches using homologous sequences from protists to higher eukaryotes as queries (indicated as an asterisk for query). For the proteins which were not retrieved by BLAST search using the *P. falciparum* sequence, further searches were performed using genes of either of a dinoflagellate *Karlodinium micrum* (*Km*) and a diatom *Thalassiosira pseudonana* (*Tp*) as queries. For AOX, I used the sequence from an apicomplexa *C. parvum* (*Cp*) as a query because *Plasmodium* does not have any AOX homolog. Accession numbers for these sequences are DQ186202 (*Tp nad1-5*), CM000649 (*Tp* SDH3), CM000642 (*Tp* SDH4), EF443012 (*Km* CYT *B*), AM773802 (*Km* COX1), EF443023 (*Km* COX3), and AB118216 (*Cp* AOX). The proteins with Ψ are coded by mt genomes in apicomplexans and dinoflagellates.

Enzyme	Specific activity (nmol/min/mg of protein)
Succinate dehydrogenase (Complex II)	215
Succinate-ubiquinone reductase (Complex II)	169
Succinate-cytochrome *c* oxidoreductase (Complex II-III)	292
+ Malonate	49.4
Quinol oxidase (Complex III-IV and AOX)	12.6
+ Ascofuranone	7.22
Cytochrome c oxidase (Complex IV)	9.47
+ KCN	0

Table 3-2.
ETC enzyme activities of *Perkinsus* crude organellar fraction. The values were obtained from five independently prepared samples.

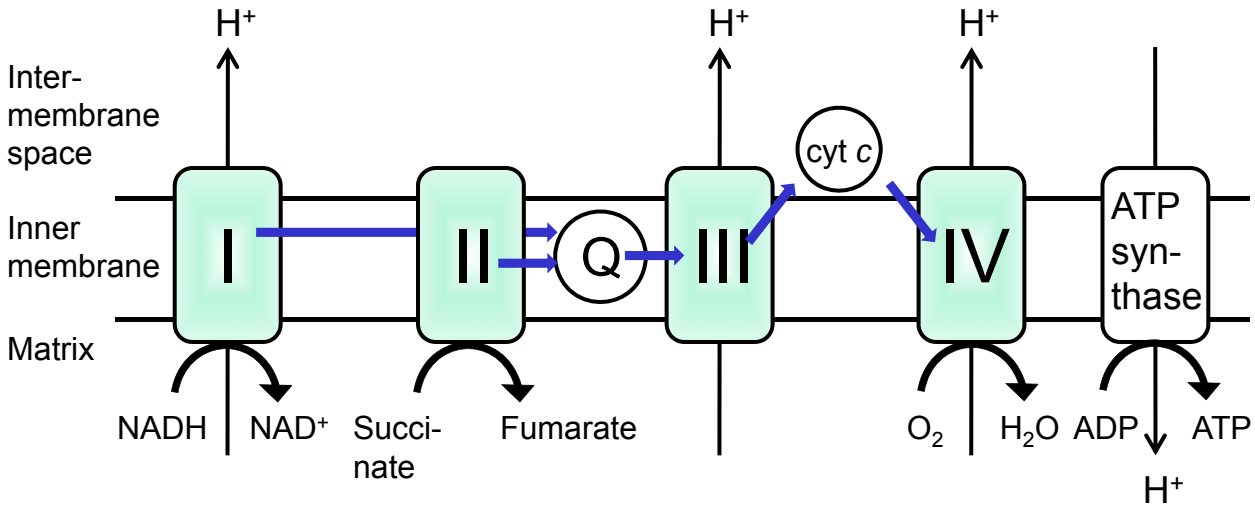

Figure 3-1.
Schematic description of the general mitochondrial ETC. ETC complexes are colored in light blue. I, Complex I: NADH dehydrogenase; II, Complex II: succinate-quinone dehydrogenase; III, Complex III: quinone-cytochrome *c* reductase; IV, Complex IV: cytochrome *c* oxidase; Q, ubiquinone; cyt *c*, cytochrome *c*. Blue arrows indicate the electron flow.

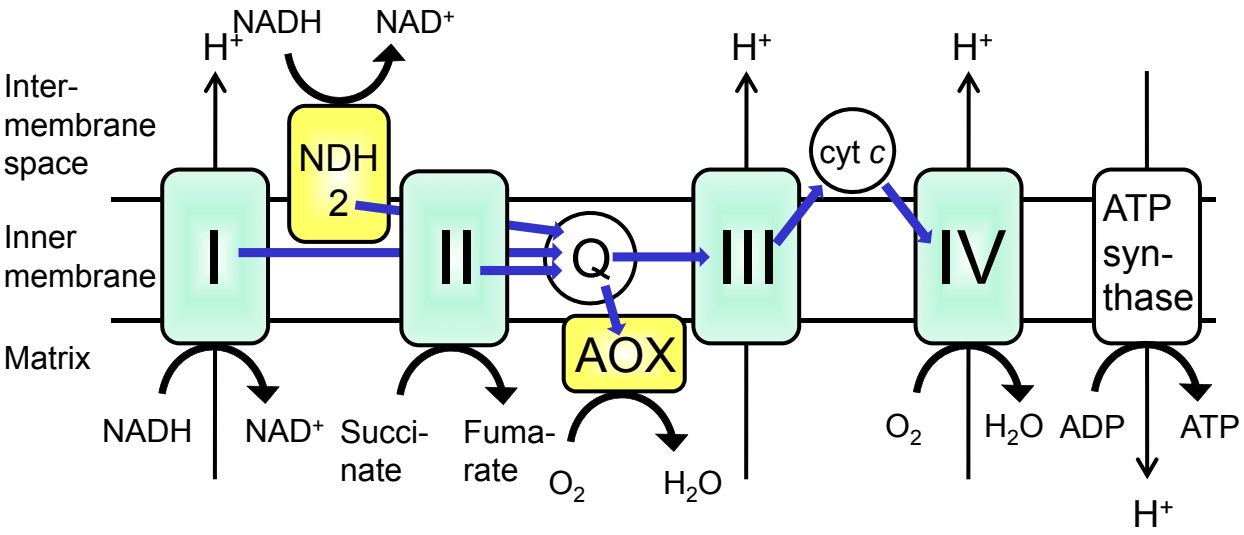

Figure 3-2.
Schematic description of the mitochondrial ETC with unconventional enzymes. All are described as in Figure 3-1, except for the unconventional enzymes: the type II NADH dehydrogenase (NDH2) and the cyanide-resistant terminal oxidase AOX colored in yellow. Note that the NDH2 is described as being located at the external surface of the mitochondrial inner membrane as observed in *Plasmodium*, although the NDH2 homolog is found also at the internal surface of the mitochondrial inner membrane in *Saccharomyces* (Marres *et al.*, 1991).

Chapter 4. ANALYSIS OF *P. MARINUS* MITOCHONDRIAL GENE AND ITS SPECIFIC EXPRESSION SYSTEM

4.1. Introduction

4.1.1. Mt genome diversity and nonstandard decoding events

Mt genomes are very diverse with regard to the physical structure, genome size, and gene content (reviewed in Lang *et al.*, 1999; Burger *et al*., 2003), due to independent evolutionary events across eukaryotic taxa (e.g., gene loss, gene transfer to the nucleus, and genome reorganization). For example, mt genomes of land plants are highly expanded (up to 2.4 Mbp in muskmelon) (Ward *et al*., 1981), and the smallest mt genome reported is a 6-kb long linear molecule in apicomplexan parasites (Feagin, 1992). An mt genome with unusual organization—several hundred linear DNA molecules coding one or a few genes—is found in the icthyosporean *Amoebidium* (Burger *et al*., 2003). In Euglenozoan flagellate *Diplonema*, one mt gene is separated into multiple fragments, each encoded on a different mini circlular molecule (Marande *et al*., 2005; Marande and Burger, 2007).

Mt gene expression is distinct from that in the nucleus, and mitochondria are notable for having alternative genetic codes. One well-known code alteration is codon reassignment in which codons are not decoded as designated in the standard codon table. For example, the UGA codon in mitochondria of many eukaryotes (other than land plants) codes for tryptophan rather than a stop (Sengupta *et al*., 2007); AGR codons (R = A or G) in inchordata mitochondria code for glycine rather than arginine (Kondow *et al*., 1999); and CUN codons (N = A, U, G or C) in yeast mitochondria code for threonine instead of leucine (Osawa *et al*., 1990). Some codon reassignments, even those that result in the same coding change, are suggested to have evolved

independently in separate taxa; one example is the reassignment of UAG codon to leucine in chlorophycean and in fungal mitochondria (Laforest *et al*., 1997; Kück *et al*., 2000; Sengupta *et al*., 2007; Ohama *et al*., 2008).

There is another category of nonstandard events in decoding: recoding (Gesteland *et al*., 1992). Whereas codon reassignments are usually due to the translational apparatus and it is irrespective the mRNA context, recoding is dependent on some signals like mRNA sequence contexts and the presence of nascent peptides (Atkins and Baranov, 2010). For example, a selenocysteine residue is coded by UGA codon with the aid of unique mRNA sequence signal (Howard *et al*., 2005). Another recoding event is known as ribosomal frameshift or programmed frameshift, a phenomenon observed in a wide range of organisms which results in a shift forward or backward in the reading frame during translation (Farabaugh, 2000; a list for reported frameshifts is found in Baranov *et al*., 2002). Ribosomal frameshifts have also been found in mt genes from various animals, and a +1 frameshift is suggested at specific codons (Härlid *et al*., 1997; Mindell *et al*., 1998; Beckenbach *et al*., 2005; Russell and Beckenbach, 2008; Milbury and Gaffney, 2008; Rosengarten *et al*., 2008). As just described, several types of nonstandard decoding events are occurring in mitochondrial gene expression systems.

4.1.2. Enigmatic mt genomes of apicomplexans and dinoflagellates

As mentioned in Chapter 1, the apicomplexan mt genome is composed of unusually reduced 6-kb element and its gene expression system is not characterized well. Interestingly, some of the unusual characteristics of apicomplexan mt genomes are shared with those of dinoflagellates, albeit with significant differences in mt genome organization (Gray *et al*., 2004; Waller and Jackson, 2009; Figure 4-1). Although the overall structure of dinoflagellate mt genome is not yet determined, they are suggested

to be composed of a number of heterogeneous DNA molecules that resulted from rampant homologous recombination (Nash *et al*., 2008; Waller and Jackson, 2009). The entire mt genome size is estimated to be at least 30 kb but is probably much larger (Nash *et al*., 2007). The genome encodes the same three protein-coding genes as apicomplexans, as well as fragmented rRNA genes; the protein-coding genes also lack canonical start and stop codons (Nash *et al*., 2007; Jackson *et al*., 2007; Nash *et al*., 2008; Waller and Jackson, 2009; Kamikawa *et al*., 2009). Transfer RNA genes have not been detected in any of the dinoflagellate mt genomes, and most of these dinoflagellate mt genomes comprise non-coding and pseudogene sequences (Norman and Gray, 2001; Nash *et al*., 2008; Kamikawa *et al*., 2009).

Recent studies on two basally-branching dinoflagellates have further highlighted the complexity of these mt genomes; the mt genomes of both *Oxyrrhis marina* and *Amphidinium carterae* are comprised of a number of DNA molecules bearing multiple copies of the three protein-coding genes with different intergenic contexts to one another (Nash *et al*., 2007; Slamovits *et al*., 2007). Particularly in the latter species, long intergenic sequences containing extensive inverted repeats are predicted to form many stem-loop structures (Nash *et al*., 2007). Although the mt genomes of these two sister lineages, which share unusual features, have not been fully characterized for the mechanisms of gene expression, the shared gene content suggests that the drastic gene reduction in genome content occurred before the divergence of these lineages. In contrast, the significant difference in apicomplexan and dinoflagellate mt genome structures indicates that drastic mt genome reorganization events occurred after the two lineages split and independently diverged from their common ancestor (Waller and Jackson, 2009).

4.1.3. Analysis of *Perkinsus* mt genome

The presence of DNA in *Perkinsus* mitochondria was previously ensured by 4',6-diamidino-2-phenylindole staining (Matsuzaki *et al.*, 2008). But contrary to my expectation, the BLAST-based sequence searches performed in Chapter 3 did not detect any of the possible *Perkinsus* mt genes. The obvious enzyme activities strongly indicated the presence of these genes and their translated products. Also, considering the phylogenetic importance, analyzing *Perkinsus* mt genome and its gene expression system may provide valuable insights into mt genome evolution of apicomplexans and dinoflagellates. In this chapter, I describe the detection and general features of *P. marinus* mitochondrial *cox1* gene (*Pmcox1*), the first annotated *Perkinsus* mt gene. I obtained the full-length mRNA sequence for *Pmcox1* by PCR and rapid amplification of cDNA end (RACE). The primary sequence of this mRNA shared several features with orthologs from related species, and together with Southern hybridization data, the codon usage suggested that this gene resides in the *P. marinus* mt genome. Unexpectedly, multiple sequence alignments and a three-frame translation indicated that the translation of this mRNA employs a modified decoding system. I discussed the primary sequence features of this mRNA and further described the possibility of a unique, modified translational decoding system in *Perkinsus* mitochondria.

4.2. Materials and methods

4.2.1. Strains and culture conditions

The *P. marinus* strain CRTW-3HE was maintained in the same manner described in Chapter 2. Strains of *P. honshuensis* and *P. olseni* were provided by Dr. T. Yoshinaga (Graduate School of Agricultural and Life Sciences, The University of Tokyo) and maintained in the same manner as for *P. marinus*.

4.2.2. Nucleic acid preparation

Collected *Perkinsus* cells were resuspended in extraction buffer containing 100 mM Tris, 100 mM boric acid and 50 mM EDTA, pH 8.0. Cell suspensions were treated with sodium dodecyl sulfate (SDS) at 60°C for 30 min. Proteins were removed by centrifugation at 2000×g for 5min after the mixing with twice volume of Tris / EDTA (TE) -saturated phenol. The aqueous layer containing the total genomic DNA (gDNA) was extracted once using an equal volume of 1:1 (v/v) phenol: chloroform mixture, and once with chloroform only, and then DNA was precipitated with 2 volumes of 100% (v/v) ethanol and washed with 70% (v/v) ethanol. Total RNA was prepared using TRIzol Reagent (Invitrogen) according to the manufacturer's protocols, followed by the poly(A)$^+$-RNA enrichment with PolyATract mRNA Isolation System III (Promega). Complementary DNA (cDNA) was synthesized with SMART RACE cDNA amplification kit (Clontech) following manufacturers' instructions.

4.2.3. PCR, RACE, cloning and sequencing

PCR was performed using Takara Ex *Taq* (Takara Bio) or *PfuUltra* II HS DNA polymerase (Stratagene). Reaction mixtures were prepared according to the manufacturers' instructions. Amplification was performed as follows: denaturation at 94°C for 4 min followed by 35 cycles of 94°C for 30 s, a primer annealing gradient from 40 to 50°C for 30 s, and extension at 72°C for 1 min, followed by a final extension at 72°C for 7 min. Primer set Pmcox1F1 and Pmcox1R3, based on the *cox1*-like sequences of nuclear DNA of mitochondrial origin (Numt) found in database (see 4.3. Results; Figure 4-2), was used to amplify the partial sequence of the *Perkinsus* mitochondrial *cox1* (Figure 4-3). Pmcox1R3 was then used in combination with a

degenerate primer cox1-3f, which was designed based on *cox1* orthologs from related species (Figure 4-3), to additionally sequence the upstream region of *Perkinsus* mitochondrial *cox1*. After determining the full *P. marinus cox1* mRNA sequence, The primer set Pmcox1fullF and Pmcox1fullR was designed for use in PCR of the nearly full-length *P. marinus cox1* both from gDNA and cDNA. Primer sequences are listed in Table 4-1.

RACE experiments were performed using Takara Ex *Taq* with *P. marinus* cDNA as the template. Reaction mixtures were prepared according to the instructions of the cDNA synthesis kit manufacturer. Reaction conditions were 35 cycles of 94°C for 30 s, 48°C for 30 s, and 72°C for 3 min, and a final extension at 72°C for 7 min. Primers for 5' and 3' RACE were Pmcox1-5RACE and Pmcox1-3RACE, respectively (Table 4-1).

PCR and RACE products were separated by electrophoresis on 1.2% agarose gel containing 1 × tris-borate EDTA (TBE) buffer and target products were extracted with the MagExtractor PCR & Gel Clean up kit (Toyobo). The gel-purified products were then cloned using the TOPO TA cloning kit for Sequencing (Invitrogen). The recombinant plasmids containing PCR or RACE products were extracted from transformed *E. coli* (strain DH5α) using MagExtractor Plasmid (Toyobo). Both strands of cloned products were sequenced with the DYEnamic ET Terminator Cycle Sequencing kit (GE Healthcare) on an ABI310 automatic sequencer (Applied Bioscience). Sequences were determined from more than three clones, unless otherwise stated. For nearly full-length gene fragments, direct sequencing was performed on four independently obtained PCR products. Consensus sequences were determined from the alignments of multiple sequences. The assembled full-length mRNA sequence was deposited to the DNA Data Bank of Japan (accession no. AB513789).

4.2.4. Sequence analysis

Sequences were aligned with Clustal X 1.83 (Thompson *et al*., 1997) and amino acid sequences were predicted using the ExPASy translate tool (http://www.expasy.org/tools/dna.html). Codon usage in several *P. marinus* genes was calculated using the Countcodon program (Kazusa DNA Res. Inst., http://www.kazusa.or.jp/codon/countcodon.html). Accession numbers for *P. marinus* nuclear genes are as follows: *ispC* (AB284362), *sod1* (AY095212), *sod2* (AY095213), and *act1* (AY436364).

After cloning and determining the putative *P. marinus* mt gene sequence, sequence searches were carried out in NCBI database (http://blast.ncbi.nlm.nih.gov/Blast.cgi) using the *P. marinus* mt gene sequence to further obtain possible *P. marinus* mt genome sequences. Note that the TIGR database was closed and unavailable after Oct 2009.

4.2.5. Southern hybridization

DNA fragments for use as probes were amplified by PCR using the following primer sets (for primer sequences, see Table 4-1): Pmcox1pF and Pmcox1pR for *P. marinus cox1*, nucLSU-7f and nucLSU-7r for large subunit ribosomal DNA (LSU rDNA), PmNumt1F and PmNumt1R for *cox1*-like Numt1 and its flanking regions, and PmNumt2F and PmNumt2R for Numt2 and its flanking regions. The amplified fragments were cloned as described above. The extracted plasmids were digested with EcoRI overnight except for two Numt-plasmids, which were digested with both NotI and PstI, and the fragments were purified, labeled, and hybridized to *P. marinus* gDNA with or without restriction enzyme digestion. The probes were used for detection with the AlkPhos Direct Labelling and Detection System with CDP-Star (GE Healthcare) as follows. First, 1 μg *P. marinus* gDNA was digested with each restriction enzyme

overnight at 37°C. Digested and uncut DNA was subjected to electrophoresis on a 0.3% agarose gel and transferred onto Hybond N^+ nylon membrane (GE Healthcare) overnight. Purified probe (100 ng) was labeled with alkaline phosphatase and hybridized to the membrane-linked gDNA overnight at 42°C. The membrane was washed and incubated with the substrate CDP-Star, and the chemiluminescence signal was detected using LAS-4000 (Fujifilm).

4.3. Results

4.3.1. Primary sequence of *P. marinus cox1*

Preliminary searches for possible mt genome fragments of *P. marinus* in the NCBI databases and the *P. marinus* draft genome database at TIGR, using mt gene sequences of dinoflagellates and apicomplexans as queries, did not produce any sequences that were supported with statistical significance (E-value <0.01). These less-E-value sequences were checked carefully by eye while referring to the amino acid alignment of COX1 from related species to identify highly conserved amino acid residues in the partial sequences, and two contigs were found to harbor *cox1*-like fragments, albeit these were only partial and tiny fragments (Figure 4-2). Contig no. 22713 (available as part of AAXJ01000589 in Genbank/EMBL/DDBJ) contained a fragment with 75.0% AT, that showed 68% predicted-amino-acid identity (17/25 residues) with *O. marina* COX1 (ABK57983) and was found to include functionally essential amino acid residues His276 and Glu278 (amino acid numbers according to Iwata *et al.*, 1995). Another fragment in contig no. 22822 (available as part of AAXJ01000147) had 70.7% AT and showed 52% predicted-amino-acid identity (20/38 residues) with *O. marina* COX1 and conserved His325 and His326 (Figure 4-2). It was realized that the base composition of these *cox1*-like fragments differed from those of the flanking regions (<55% AT). The

flanking regions did not show sequence similarity to *cox1* and were discovered to harbor nuclear genes like RNA helicase gene and clathrin-associated protein gene, the former of which contained the *cox1*-like fragment in one of its intronic regions (Figure 4-3A). These observations imply that these *cox1*-like, AT-rich fragments are nuclear DNA of mitochondrial origin (Numts), which are DNA fragments that had been transferred from mt genomes into the nucleus and, in many cases, have become transcriptionally inactive (Lopez *et al*., 1994). They are usually found in eukaryotic nuclear genomes (Richly and Leister, 2004).

Because the *cox1*-like Numts and the true mitochondrial *cox1* are likely to have similar sequences, two primer sets were used for PCR: i) Pmcox1F1 and Pmcox1R3, both of which were derived from the Numt sequences and ii) Pmcox1R3 and a degenerate primer cox1-3f, which was designed based on *cox1* sequences of closely related species. In each case, there was a distinct single DNA amplification from total *P. marinus* DNA template. Sequencing of these PCR products confirmed the lengths at 167 bp and 434 bp, respectively, with the former being completely included in the latter. The 434-bp fragment was also amplified from *P. honshuensis* and *P. olseni* gDNA. The nucleotide sequence identity amongst the 434-bp fragments from the three species was more than 96%. To obtain the full-length sequence of this gene, 5' and 3' RACE were performed using internal primers Pmcox1-5RACE and Pmcox1-3RACE, respectively, with *P. marinus* cDNA as the template. After cloning and sequencing 5 clones for each of the RACE products (approximately 700 bp each), the nearly full-length sequences (approximately 1400 bp) of both gDNA and cDNA were amplified using specific primer sets (Pmcox1fullF and Pmcox1fullR) followed by direct sequencing of multiple independent PCR products. The sequences of the RACE products and the nearly full-length sequence were manually assembled to determine the full-length mRNA

sequence (1434 bp) of this gene, which was confirmed to contain sequences identical to the PCR and RACE fragments obtained above. Conversely, this mRNA contained regions which are similar to but not identical to Numt sequences, and their flanking regions were completely different from each other (Figure 4-3B). There were no substitutions, insertions and deletions between sequences from gDNA and cDNA, suggesting that RNA editing does not occur in this gene. The overall AT content of this gene was 80.9%. As a whole, this gene was similar to *cox1* of dinoflagellates and apicomplexans with an E-value less than 10^{-70}; hereafter, this sequence is referred to as *Pmcox1* mRNA.

4.3.2. Genomic localization of *Pmcox1*

To determine the localization of *Pmcox1* in *P. marinus* genomes (nucleus or organelles) and to infer the *Perkinsus* mt genome structure, Southern hybridization was conducted using total DNA because it was difficult to isolate pure mt DNA from *P. marinus*. *Pmcox1* signals constituted a smear in the low molecular-weight region (<10 kb) of uncut genomic DNA, which is far lower than the expected position for chromosomal DNA (Figure 4-4). Similarly, *Pmcox1* signals formed a smear for the digestion of total DNA with BamHI, EcoRI, or HindIII. A distinct signal was only observed (1-2 kb region) for the digestion of total DNA with AccI. Given the high AT content of *Pmcox1*, it is natural that AccI was the only restriction enzyme tried here which cut the DNA sequences around *Pmcox1*.

In sharp contrast to the *Pmcox1* probe, the probe for the nuclear LSU rDNA hybridized to the stacked, high molecular-weight, chromosomal DNA in the uncut DNA sample (Figure 4-4). Moreover, one or two distinct LSU rDNA band(s) were detected in gDNA digested with AccI, BamHI, EcoRI, or HindIII. The LSU rDNA signals indicate

the high quality of gDNA and that the restriction digests were complete. Similarly, probes for Numts and its flanking regions hybridized to the undigested chromosomal DNA without a smear signal, indicating that they reside on chromosomal DNA (Figure 4-5). The smear signals from the *Pmcox1* probe suggest that *Pmcox1* resides on small (<10 kb) heterogeneous non-chromosomal DNA.

4.3.3. Prediction of amino acid sequence

The amino acid sequence predicted to be encoded by the primary *Pmcox1* mRNA sequence unexpectedly could not be translated in its entirety using the standard codon table in a single reading frame; several stop codons appeared in all three frames (Figure 4-6). The BLASTx-based search using the entire *Pmcox1* mRNA sequence as a query identified several partial COX1-like amino acid sequences that appeared separately in all three reading frames (gray boxes in Figure 4-6). In total, I found eleven COX1-like "coding-blocks" (gray boxes numbered I – XI in Figure 4-7) that cover almost the entire sequence of *Pmcox1*, though discontinuously.

To understand the discontinuity in the COX1-like amino acid sequences, I aligned the *Pmcox1* mRNA sequence with *cox1* sequences of related species (Figure 4-8). Among the four *cox1* sequences, *Pmcox1* was the most divergent and contained the largest number of insertions and deletions (indels). Curiously, there were ten 1- or 2-base indels specifically in the *Pmcox1* mRNA that occurred in the context of AGGY (8 of 10) or CCCCU (2 of 10) motifs (shown on a black background in Figure 4-8). Most intriguingly, these 10 regions appeared to coincide with the transitions between coding-blocks, and AGG and CCC appeared in-frame preceded by the predicted COX1-coding blocks (Figure 4-6).

Based on these observations, I postulated that modified decoding, which could shift

the reading frame, occurs in the translation of *Pmcox1* mRNA. In my hypothesis, specifically, when an in-frame AGG or CCC appears, the reading frame should be shifted forward by 1 or 2 bases, respectively. Accordingly, I prepared a putative PmCOX1 amino acid sequence in the following manner. I eliminated the A residues of the AGGY motifs and made a +1 frameshift, making GGY instead of AGG in-frame. I also deleted the first two C residues of CCCCU motifs and made a +2 frameshift, making CCU instead of CCC in-frame. This model accounts for all the *Perkinsus*-specific 1- and 2-base indels and connects the 11 "blocks" into one consecutive coding sequence. The alignment of our putative PmCOX1 sequence with counterparts from related organisms shows the conservation of functionally important amino acid residues (Figure 4-9, black boxes). This sequence also conserves the glycine and proline residues, which are most common in the proximity of the AGGY and CCCCU motifs. The potential mechanisms for these frameshifts will be further discussed later.

4.3.4. Codon usage and base composition

Based on the amino acid sequence deduced in 3.3.3., I analyzed the codon usage and base composition. Around the 5' terminal regions, no AUG codon that is likely to act as start codon was identified (Figure 4-10). Canonical stop (UAA, UAG and UGA) codons were not observed in 3' terminal regions, as is often the case with mt genes of dinoflagellates and apicomplexans (Figure 4-10). Comparison of the COX1 amino acid alignment and nucleotide sequence also showed well-conserved tryptophan residues among related species that appeared to be coded by UGA codons in *Pmcox1* (open boxes in Figure 4-9, Figure 4-6, and Table 4-2). The predicted PmCOX1 sequence did not contain any cysteines; whereas the 19 other amino acids occurred at least once.

Codon usage calculation on the entire sequence revealed that *Pmcox1* is highly divergent among *cox1* orthologs from related protists. *Pmcox1* utilizes only 35 kinds of codons, a much smaller number than those in related species (48-58) and in *P. marinus* nuclear genes (53-60) listed in Table 4-2. For example, leucine is coded exclusively by UUA codon in *Pmcox1*, whereas each of the other species examined uses 5 or 6 kinds of codons to code for leucine. Also, arginine was coded solely by AGA and alanine by GCU in *Pmcox1*.

Since the biased codon usage is most likely related to the base composition, I examined base composition of *Pmcox1* in detail and compared it to those of related species (Table 4-3). The overall AT content of *Pmcox1* was 80.9%, higher than those of *cox1* orthologs from related organisms (64% to 72%) and much higher than those of *Perkinsus* nuclear genes (43% to 50%). The AT skew of *Pmcox1* was -0.009, meaning that *Pmcox1* mRNA contained almost the same number of A and T residues on the sense strand. Contrastingly, the AT skew of other *cox1* sequences ranged from -0.157 to -0.297, reflecting that they used 1.3 to 1.8 times more T than A.

To determine whether the base composition bias is associated with codon usage, base composition was investigated for each codon position. A and T were used in 89.8% of the 3rd codon positions of *Pmcox1*, while they were used in 77.4% and 75.9% of the 1st and 2nd codon positions, respectively. This AT bias is reflected in codon usage; only nine *Pmcox1* codons had a G or C at their 3rd position. In contrast, the 3rd codon positions of *P. marinus* nuclear genes have A or T much less frequently (39.1%). The AT bias at the 3rd codon position is higher in *Pmcox1* than in *cox1* orthologs from other species (73.9% to 87.7% AT) (Table 4-3).

4.3.5. BLAST-based searches of the database for other possible *Perkinsus* mt

genome sequences

Sequence searches with BLASTn and BLASTx program were performed using the *Pmcox1* mRNA sequence in the NCBI database as of November 2009. From *P. marinus* whole-genome shotgun assemblies, I obtained only one contig containing a fragment which perfectly matches our *Pmcox1* (accession number: AAXJ01004741). This contig possesses 5' and 3' flanking regions of *Pmcox1*, in which no protein-coding or structural RNA sequence was found.

Subsequent sequence searches using the 5' flanking region of *Pmcox1* as a query recovered a few other sequence with high similarity. I reiterated the sequence search process using newly recovered sequences as queries. Rounds of searches recovered several contigs, and the AT content of these sequences was as high as 80%. Almost all the sequences recovered were non-coding (data not shown), but I found one entry which contained a fragment similar to *cob* of dinoflagellates (AAXJ01022806). As opposed to *Pmcox1* which is similar in length to orthologous *cox1* genes, this gene fragment was shorter than the full-length orthologous genes found in other organisms. The *cob*-like fragment from *P. marinus* sequence was less than 600 bp, while *cob* is generally around 1100 bp in related species (Jackson *et al*., 2007; Zhang *et al.*, 2008).

4.4. Discussion

4.4.1. *Pmcox1* shares the features with apicomplexan and dinoflagellate mt genes

Using the newly determined sequence of *Pmcox1* mRNA and nearly the full-length sequence of its genomic counterpart, I found several data to suggest that *Pmcox1* is located in the mt genome and recognized that this gene conserved several features shared with apicomplexan and dinoflagellate mt genes.

First, Southern hybridization of total gDNA from *P. marinus* shows the localization

of *Pmcox1* that is distinct from that of the nuclear LSU rDNA. Signal from *Pmcox1* probes formed a smear in the relatively low molecular-weight regions of uncut total gDNA, while LSU rDNA probe hybridized to stacked, uncut DNA with high molecular weight, i.e., chromosomal DNA (Figure 4-4). These results indicate that *Pmcox1* resides on the relatively small DNA molecules distinct from chromosomal, nuclear DNA. The present hybridization data (Figure 4-4) is congruent with previously reported results on other dinoflagellates (Norman and Gray, 1997; Chaput *et al*., 2002; Jackson *et al*., 2007), suggesting that *Pmcox1* is encoded on multiple heterogeneous DNA molecules, which is similar to the structure found for other dinoflagellate mt genomes.

Second, canonical start and stop codons are not found in the terminal regions of *Pmcox1* (Figure 4-10). As the mt genes of dinoflagellates and apicomplexans do not possess AUG start codon and stop codons, these are assumed to utilize alternative start and stop mechanisms (Norman and Gray, 1997; Rehkopf *et al*., 2000; Chaput *et al*., 2002; Jackson *et al*., 2007; Kamikawa *et al*., 2009). All of the *Perkinsus* nuclear genes examined here had AUG start and stop codons in the expected positions based on comparisons to orthologs from related species. These observations support that *Pmcox1* resides in the mt genome and is expressed by probably similar mechanism to those in apicomplexan and dinoflagellate mitochondria.

Lastly, overall codon usage and base composition showed significant differences between *Pmcox1* and *Perkinsus* nuclear genes (Tables 4-2 and 4-3). Moreover, several UGA codons, which typically function as stop codons in nuclear genes but often code for tryptophan in mt genes, were present in the *Pmcox1* mRNA and appeared to code for tryptophan (Figures 4-6 and 4-9).

These multiple lines of evidence strongly suggest that this gene is not located in the nuclear but in the extra-chromosomal DNA which has similar structure to the

dinoflagellate mt genomes, although no direct evidence has been obtained presently. To fully prove the mt genome localization of *Pmcox1*, direct evidence is required using approaches such as fluorescence *in situ* hybridization. It was also strongly indicated that the gene expression system for *Pmcox1* shares features with those in apicomplexans and dinoflagellates.

4.4.2. Mt gene translation of *Perkinsus* involves multiple frameshifts

Surprisingly, the *Pmcox1* mRNA is apparently not translated in a single reading frame. Because I detected the cyanide-sensitive enzyme activity of COX as described in Chapter 3, the functional COX1 protein most likely exists in *Perkinsus* mitochondria. Furthermore, I obtained the fragments of *Pmcox1* orthologs from a related species *P. honshuensis* and *P. olseni* (Figure 4-11). They showed high sequence similarity to one another (> 96%), and conserved one AGGY motif. In addition, the longer sequence of the *P. olseni* ortholog, which possibly covered almost the entire sequence of *Pmcox1*, was found in the draft whole genome assembly of *P. olseni* we have. There were only four gaps in the 5' terminal region in the alignment, and the nucleotide sequence identity of the unambiguously aligned region was 98%. All the AGGY and CCCCU motifs were conserved, and all substitutions were synonymous, indicating the selective pressure to conserve the amino acid sequence in *Pmcox1* and this ortholog. Taken together with there being no *cox1*-like sequence other than *Pmcox1*, these results further emphasize that *Pmcox1* is functional and is translated with the aid of an unusual mechanism that requires multiple frameshifts.

At present, I am unable to show direct evidence that translation of *Pmcox1* mRNA requires frameshifts because we have not directly sequenced the PmCOX1 protein. However, the predicted PmCOX1 amino acid sequence reinforces the validity of my

frameshift model. As mentioned in 3.4.2, COX1 takes the major role in reducing oxygen to water. The amino acid sequence of PmCOX1 predicted by the frameshift model retains the conserved residues that are essential for these reactions (see Figure 4-9 and its legend) (Iwata *et al*., 1995). The reading frame is possibly shifted back by 1 base (-1 frameshift) at the CCCCU motif; this -1 frameshift also adjusts the frame so as to read NCC CCU and the COX1-like amino acid sequence would be followed in the same reading frame. But this is less likely because it would result in the insertion of one extra amino acid residue, coded by NCC, into the amino acid alignment.

Moreover, this frameshift motif may be conserved in another mt gene. I identified a *cob*-like fragment from *P. marinus* whole-genome shotgun assemblies (AAXJ01022806) in a BLAST-based search using dinoflagellate mt gene sequences. This fragment included five conserved AGGY motifs where the reading frame appeared to be shifted forward by 1 base to connect discontinuous COB-like amino acid sequences to form a plausible COB protein (Figure 4-12). The orthologous gene fragment, although shorter, was found in the draft whole genome assembly of *P. olseni*. The fragment also conserved three AGGY motifs, and the deduced amino acid sequence was identical to that of *P. marinus* counterpart (Figure 4-12). In contrast, the deduced amino acid sequences for *Perkinsus* nuclear genes shown in Table 4-2 did not include such translational frameshifts. These observations strongly indicate that an unconventional event occurred during translation, specifically in mitochondria of *P. marinus*, and also of other *Perkinsus* species. Notably, this is the first report of a frameshift-involving translation system in protist mitochondria.

4.4.3. Possible mechanisms suggested for frameshift in *Perkinsus* mt genes

If *Pmcox1* mRNA is read in all three frames to generate PmCOX1, an

unconventional mechanism must exist in the *Perkinsus* mitochondrial translation system to shift the reading frame systematically. One possible mechanism is a ribosomal frameshift. In the case of +1 ribosomal frameshift, a rarely used codon or a stop codon in the ribosome A site is suggested to induce the ribosome to stall and allow the reading frame to be subsequently shifted forward by skipping 1 base (Belcourt and Farabaugh, 1990). Based on previous studies, I hypothesized that ribosomes in *Perkinsus* mitochondria skip the A residue in the first position of the in-frame AGG in the shared AGGY motif and the first two C residues in the CCCCU motif by shifting forward by one base at in-frame CCC (Figure 4-13A). These two types of frameshifts at the rarely used AGG and CCC codons change the reading frame and allow the discontinuous COX1-like amino acid sequences to be joined, which produces the preferred amino acid residues at the frameshift sites (Figure 4-9).

Alternatively, specialized tRNAs that recognize non-triplet codons may be utilized at frameshift sites during translation. Naturally occurring deviant tRNAs recognize four-base codons and act as suppressors of nonsense mutations, and artificial tRNAs bearing modified loops can recognize quadruplet and even quintuplet codons (Hohsaka *et al*., 2001; Magliery *et al*., 2001; Anderson *et al*., 2004; Wang *et al*., 2006; Atkins and Björk, 2009). In the case of *Pmcox1*, specialized tRNAs may recognize AGGY (for glycine) and CCCCU (for proline) to enable the proposed frameshifts (Figure 4-13B). One possible process for the generation of $tRNA^{Gly}$ recognizing AGGY is that the nucleotides of the anticodon and its adjacent regions of $tRNA^{Gly}$ (anticodon: GCC) are modified post-transcriptionally to decode four nucleotide codon. Two most common modified nucleotides in the tRNA anticodon region are inosine and pseudouridine, which are synthesized by enzymatic modification from adenine and uridine, respectively (various types of modifications are described in Grojean *et al*., 2010). They

have extended pairing capacities; inosine can pair with U, C and A. It can be anticipated that some modifications like them are present in the anticodon loop of the *Perkinsus* mitochondrial $tRNA^{Gly}$. Also, it can be possible that the $tRNA^{Arg}$ (CCU) has a modified anticodon loop, and is charged with glycine instead of arginine by a specific aminoacyl-tRNA synthase. Regarding the frameshift at CCCCU motifs, $tRNA^{Pro}$ (GGG or GGA) with modified anticodon loop is expected to be present. With these tRNAs, the reading frame would be shifted at AGGY and CCCCU motifs by 1 and 2 base(s) respectively, and one contiguous COX1 protein would be translated. Such specialized tRNAs with altered decoding capacity may be used in *Perkinsus* mitochondria, although mitochondrial tRNAs involved in frameshifting have not yet been identified from any organism.

Regardless of the mechanism, it should be noted that the 100% frameshift efficiency has never been observed, although the efficiency of translational frameshift depends on the nucleotide sequence and the abundance of tRNAs (Sundararajan *et al*., 1999). Lower frameshift efficiencies are not lethal to organisms known to have frameshift-dependent genes because there is only one (most cases) or at most two (for nuclear genes of some ciliates like *Euplotes* described in Klobutcher, 2005) ribosomal frameshifts per one gene. In contrast, frameshift must occur at as many as ten sites to produce a complete COX1 protein in *Perkinsus*, which is a surprisingly high number. If one frameshift failure occurs at any of the 10 sites due to low efficiency, only a truncated COX1 protein, and not the full-length protein, will be synthesized to deleterious effect on respiratory function of *Perkinsus*. It is known that "stimulatory" elements such as upstream Shine-Dalgarno-like sequences or downstream pseudoknot structures promote efficient frameshifts (Gesteland and Atkins, 1996; Giedroc and Cornish, 2009). There are, however, no such sequences found around the frameshift sites in *Pmcox1*.

Based on these observations, I suggest that the complete translation of *Pmcox1*, a *Perkinsus* mt gene, requires a quite accurate mechanism for highly frequent and efficient frameshifts. “Ten times per gene” is by far the highest frequency among the reported translational frameshifts. Considering this unusually high frequency, I suggest that the function of the frameshift mechanism in *Perkinsus* mitochondria is far more efficient and active than that of the frameshifts in other organisms. In the aspect of the efficiency, the tRNA-mediated frameshift mechanism is more likely rather than the ribosomal frameshift due to the insufficient amount of tRNA.

To confirm the frameshift model and also to identify the start and stop codons within *Pmcox1* mRNA, the actual amino acid sequence of *Pmcox1* should be determined. I will also investigate the translational machinery in *Perkinsus* mitochondria to understand the mechanisms that promote these “extensive” frameshifts.

Name	Sequence (5'-3')
PCR and RACE for sequencing	
Pmcox1F1	TTCAACGGTGATGTGGTATTATAAC
Pmcox1R3	CTAAACCTACTGTATATATATGATGAGC
cox1-3f	ACNGGITGRACIYTITAYCCNCCNYT
Pmcox1-5RACE	AAGTGTTGATATAATACTACATCACC
Pmcox1-3RACE	ATGGTGATGTAGGTATTATATCAACAC
Pmcox1fullF	CTAATCAGTAATCGTGATACGCTAACC
Pmcox1fullR	TGAACCAATAGATGATATTAAATTCCATAC
Probe amplification for southern hybridization	
Pmcox1pF	GCCTAGGTTTATATGGTGGTATACC
Pmcox1pR	CTTCTAGGGGCATTACATTGAAACC
nuLSU-7f	TCCTGAGGGAAACTTCGGAGG
nuLSU-7r	GATAGCAACAAGTACCGTGAGG
PmNumt1F	GTCGAGGCAGCACAGTACTG
PmNumt1R	CGATTCTCGACATCACAGTCACC
PmNumt2F	CCAAGAGACATGCATGAGTAGC
PmNumt2R	GGTGACTGTGATACGACTGCC

Table 4-1.
Primer sequences used in Chapter 4.

Codon	Amino acid	Dinoflagellates and Apicomplexa				Diatom	*P. marinus* nuclear genes			
		Pmar	*Acat*	*Omar*	*Pfal*	*Tpse*	*ispC*	*sod1*	*sod2*	*actin*
UUU	F	14.9	80.2	60.9	79.5	84.2	5.6	4.4	7.1	0
UUC		25.6	22.6	39.9	10.5	20	18.6	53.1	25	39.9
UUA	L*	119.4	67.9	60.9	115.1	70.1	0	4.4	0	0
UUG		0	16.5	21	6.3	12	9.3	31	32.1	23.9
CUU	L*	0	41.2	54.6	6.3	16	11.1	4.4	7.1	8
CUC		0	14.4	12.6	0	0	27.8	13.3	46.4	18.6
CUA		0	10.3	8.4	18.8	12	9.3	0	7.1	2.7
CUG		0	6.2	4.2	2.1	2	39	13.3	28.6	13.3
AUU	I	74.6	28.8	44.1	58.6	56.1	16.7	26.5	10.7	26.6
AUC		6.4	41.2	33.6	2.1	16	37.1	22.1	17.9	42.6
AUA		185.5	24.7	16.8	43.9	20	14.8	4.4	10.7	0
AUG	M	23.5	26.7	18.9	37.7	36.1	29.7	17.7	25	47.9
GUU	V	8.5	20.6	33.6	18.8	34.1	18.6	13.3	17.9	18.6
GUC		0	10.3	8.4	2.1	0	24.1	31	17.9	29.3
GUA		49	24.7	21	43.9	30.1	13	13.3	7.1	5.3
GUG		0	4.1	2.1	0	8	16.7	4.4	21.4	16
UCU	S	25.6	37	54.6	31.4	32.1	13	13.3	25	26.6
UCC		4.3	16.5	4.2	0	2	9.3	0	3.6	16
UCA		23.5	37	25.2	31.4	16	7.4	4.4	3.6	2.7
UCG		2.1	2.1	4.2	2.1	4	20.4	4.4	21.4	2.7
CCU	P	10.7	14.4	16.8	18.8	28.1	20.4	17.7	17.9	18.6
CCC		0	2.1	2.1	0	0	9.3	0	10.7	23.9
CCA		14.9	26.7	16.8	16.7	14	11.1	13.3	14.3	5.3
CCG		0	6.2	0	2.1	6	11.1	17.7	10.7	2.7
ACU	T	2.1	20.6	21	37.7	24	14.8	31	17.9	42.6
ACC		0	4.1	6.3	4.2	18	9.3	22.1	14.3	18.6
ACA		27.7	37	21	35.6	18	9.3	8.8	3.6	2.7
ACG		0	0	8.4	0	4	11.1	4.4	10.7	10.6
GCU	A*	14.9	20.6	54.6	20.9	38.1	22.3	66.4	25	39.9
GCC		0	2.1	4.2	2.1	0	39	17.7	17.9	23.9
GCA		0	12.3	23.1	18.8	36.1	27.8	8.8	28.6	2.7
GCG		0	0	2.1	0	0	31.5	8.8	10.7	0
UAU	Y	72.5	32.9	37.8	29.3	28.1	1.9	22.1	7.1	21.3
UAC		10.7	4.1	4.2	8.4	14	14.8	17.7	32.1	13.3
UAA	Stop	0	0	0	0	0	0	0	0	0
UAG		0	0	0	0	0	0	0	0	0
CAU	H	12.8	26.7	21	25.1	24	7.4	35.4	21.4	13.3
CAC		4.3	4.1	6.3	4.2	10	9.3	8.8	10.7	10.6
CAA	Q*	8.5	8.2	12.6	10.5	4	9.3	13.3	0	2.7
CAG		0	2.1	0	0	2	24.1	17.7	14.3	26.6
AAU	N	70.4	32.9	21	41.8	24	7.4	17.7	17.9	10.6
AAC		6.4	8.2	21	6.3	10	14.8	53.1	17.9	13.3
AAA	K*	25.6	10.3	4.2	10.5	24	3.7	4.4	7.1	13.3
AAG		0	6.2	6.3	0	0	44.5	53.1	28.8	47.9
GAU	D*	23.5	16.5	21	12.6	24	29.7	17.7	28.6	31.9
GAC		0	2.1	2.1	8.4	2	24.1	31	25	26.6
GAA	E*	10.7	12.3	14.7	14.6	12	16.7	0	14.3	16
GAG		0	2.1	0	0	2	27.8	48.7	28.6	58.5
UGU	C$^{\Psi}$	0	8.2	2.1	12.6	2	3.7	0	3.6	2.7
UGC		0	0	4.2	2.1	0	14.8	8.8	3.6	8
UGA	Stop/W	10.7	0	0	0	10	0	0	0	0
UGG	W	4.3	14.4	12.6	16.7	14	7.4	22.1	32.1	10.6
CGU	R*	0	0	2.1	6.3	2	9.3	8.8	10.7	13.3
CGC		0	2.1	4.2	0	0	5.6	8.8	3.6	18.6
CGA		0	0	2.1	0	6	9.3	0	25	2.7
CGG		0	0	2.1	0	2	11.1	0	21.4	0
AGU	S	21.3	12.3	14.7	18.8	20	14.8	17.7	25	10.6
AGC		0	8.2	10.5	4.2	2	26	13.3	3.6	5.3
AGA	R*	12.8	14.4	6.3	14.6	10	13	4.4	3.6	5.3
AGG		0	2.1	4.2	0	0	5.6	8.8	21.4	8
GGU	G	64	14.4	27.3	33.5	58.1	16.7	26.5	10.7	42.6
GGC		4.3	6.2	8.4	0	6	31.5	26.5	21.4	16
GGA		4.3	57.6	18.9	50.2	26.1	13	17.7	10.7	13.3
GGG		0	12.3	2.1	2.1	4	24.1	0	32.1	5.3
gene length (bp)		1410	1458	1428	1434	1497	1617	678	840	1128
total codon number		470	486	476	478	499	539	226	280	376
codon kind used		35	55	58	48	54	60	53	59	56
A+T content		81%	68%	68%	72%	67%	43%	50%	45%	49%

Table 4-2.

Codon usage of *Pmcox1*, *cox1* from related organisms, and *Perkinsus* nuclear genes. The numbers represent codon frequencies per thousand. Asterisk indicates the amino acid coded only by one codon species in PmCOX1. Ψ indicates the amino acid which is not contained in PmCOX1. The codon usage of PmCOX1 was calculated after eliminating the eight A residues in AGGY motifs and the two CCs in CCCCU motifs.

Base		*Pmar*	*Omar*	*Acat*	*Pfal*	*Tpse*	*Pmar* nuclear genes
A		40.1	22.8	26.3	30.5	25.7	22.5
T		40.8	42.0	39.6	41.4	40.7	23.8
C		8.2	20.1	18.4	12.9	15.6	25.7
G		10.9	15.1	15.7	15.2	18.0	28.0
A + T	1st position	77.4	59.0	61.7	66.1	59.1	44.1
	2nd position	75.9	61.3	60.9	61.7	59.7	55.7
	3rd position	89.8	73.9	75.1	87.7	80.4	39.1
	total	80.9	64.8	65.9	71.9	66.4	46.3
AT skew		-0.009	-0.297	-0.200	-0.157	-0.225	-0.024
GC skew		-0.139	0.141	0.078	-0.084	-0.070	-0.044

Table 4-3.
Base composition of several *cox1* sequences and *P. marinus* nuclear genes. All values are depicted as percentages except for the AT and GC skews. *P. marinus* nuclear genes are the same as listed in Table 4-2.

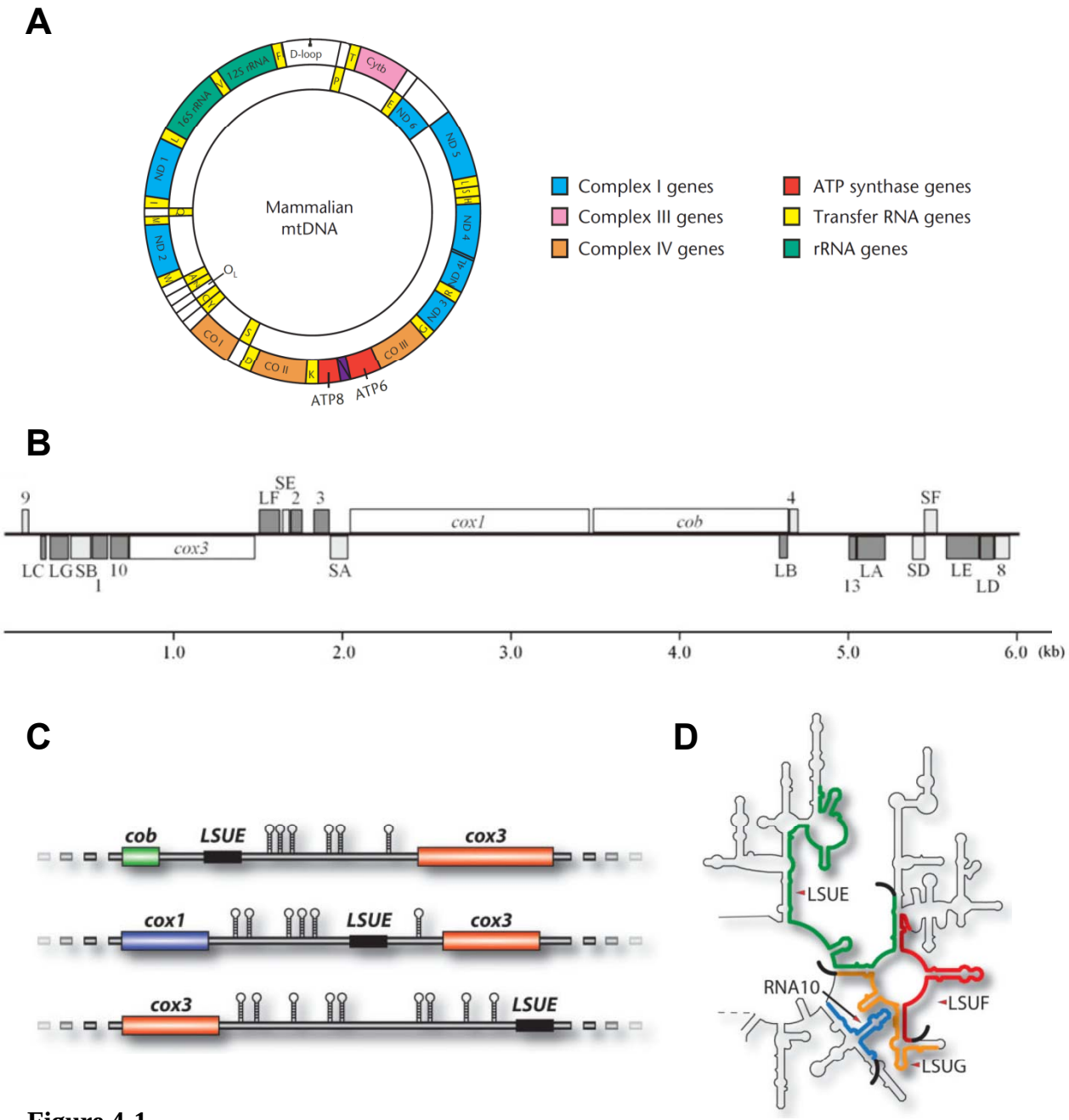

Figure 4-1.
Schematic descriptions of mt genomes. (A) Circular human mt genome. The details are described in Saccone, 2005. (B) Mt genome of *Plasmodium* (adapted from Hikosaka *et al.*, 2010). The three protein-coding genes *cox1*, *cox3* and *cob* are described as white boxes. Dark and light gray boxes are genes for rRNA. (C) Partial mt genome organization of dinoflagellates (Waller and Jackson, 2009). The entire mt genome is composed of heterogeneous DNA molecules as shown here, and duplicated genes are distributed among the DNA molecules with stem-loop-rich intergenic regions. (D) The mitochondrial rRNA genes are fragmented in apicomplexans and dinoflagellates, and each of the transcripts is suggested to constitute the rRNA (Waller and Jackson, 2009).

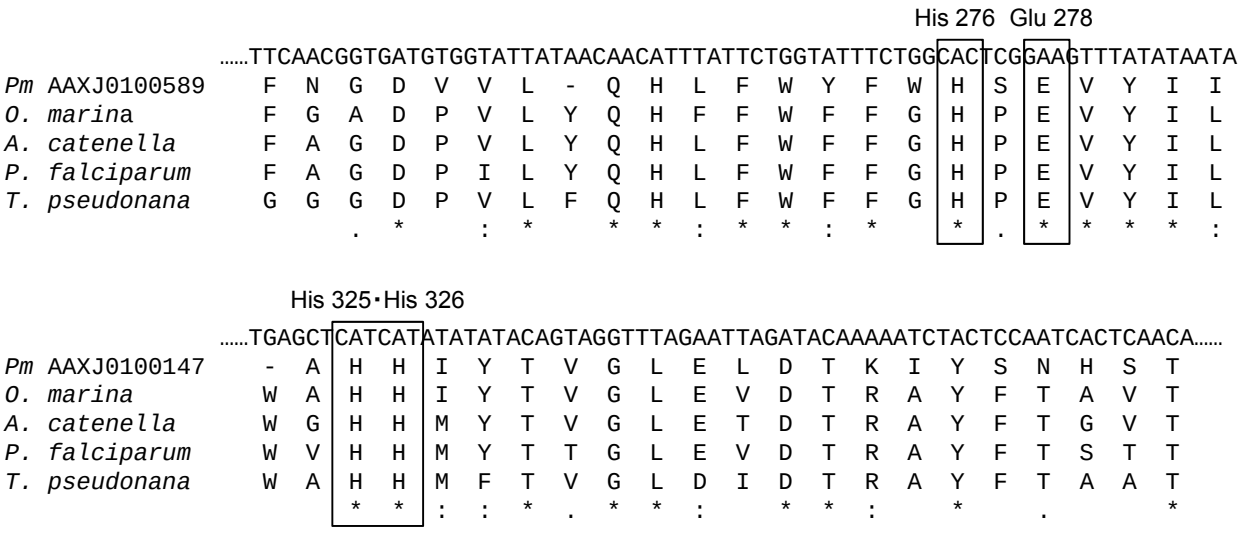

Figure 4-2.
cox1-like Numt sequences (represented as NCBI entry number due to the unavailability of TIGR database) and the deduced COX1-like amino acid sequence aligned with COX1 sequences from related species. Asterisks, colons and dots indicate identical residues, conserved and semi-conserved substitutions, respectively. *cox1* sequences were obtained from the NCBI database for *Oxyrrhis marina* (EF680822), *Alexandrium catenella* (AB374235), *P. falciparum* (AY282930) and *T. pseudonana* (DQ186202).

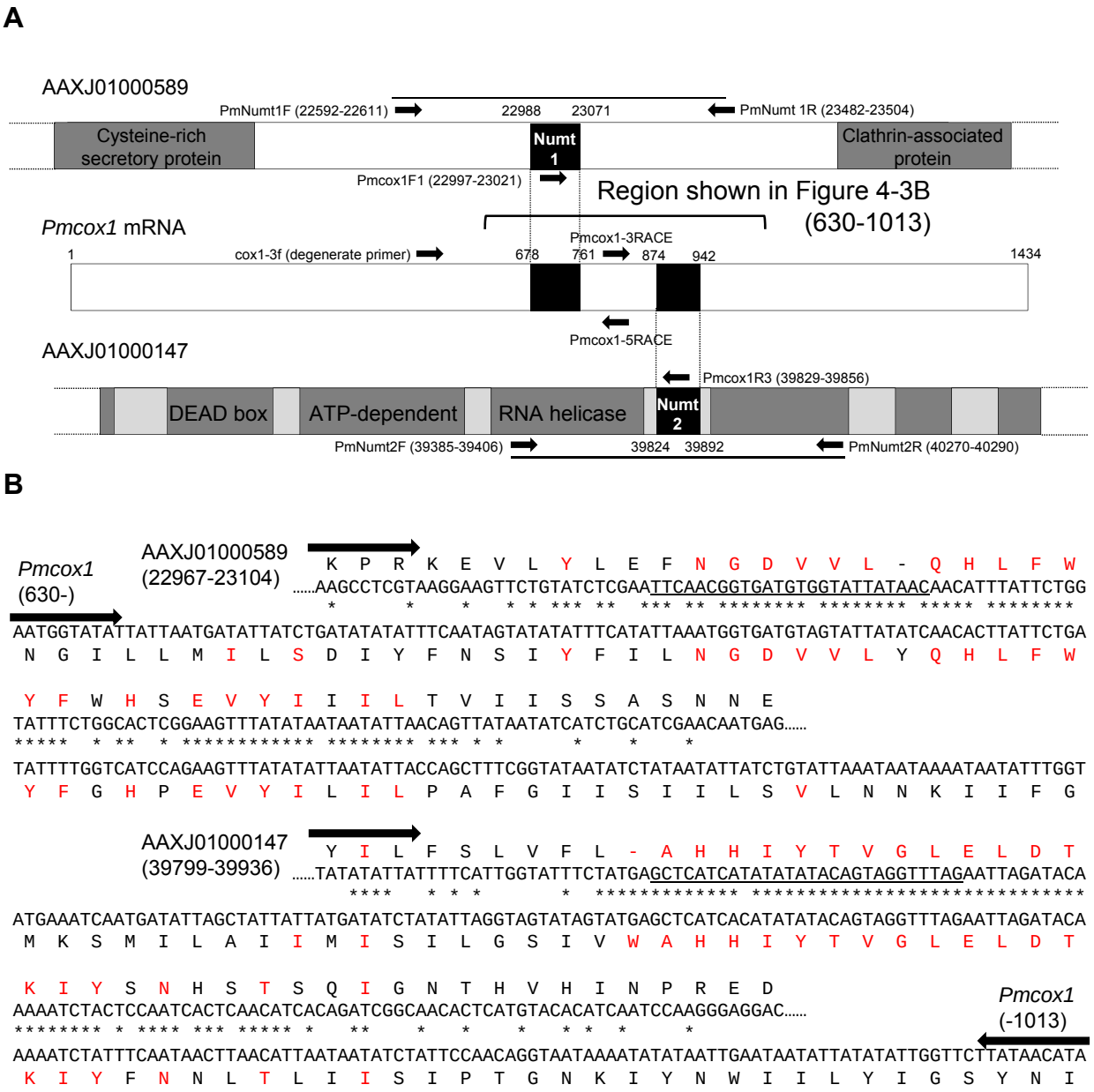

Figure 4-3.
(A) Schematic model of Numt-containing contigs and *Pmcox1* (not scaled). Dark and light gray boxes indicate exons and introns, respectively, and black boxes indicate *cox1*-like Numts. Arrows stand for PCR primers, and regions used as probe for southern hybridization are indicated by bars. (B) Pairwise alignment of each of Numt sequences and *Pmcox1*, and their corresponding amino acid sequences. Identical nucleotides and amino acid residues are represented by asterisks and red characters, respectively.

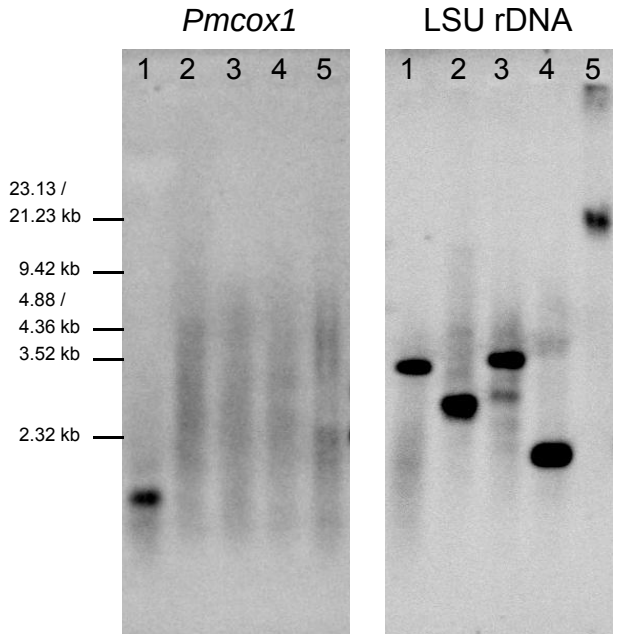

Figure 4-4.
Southern hybridization with *Pmcox1* (left) and a control nuclear LSU rDNA (right) probe. Lanes 1–4, *P. marinus* gDNA digested with AccI (1), BamHI (2), EcoRI (3) and HindIII (4); lane 5, uncut gDNA.

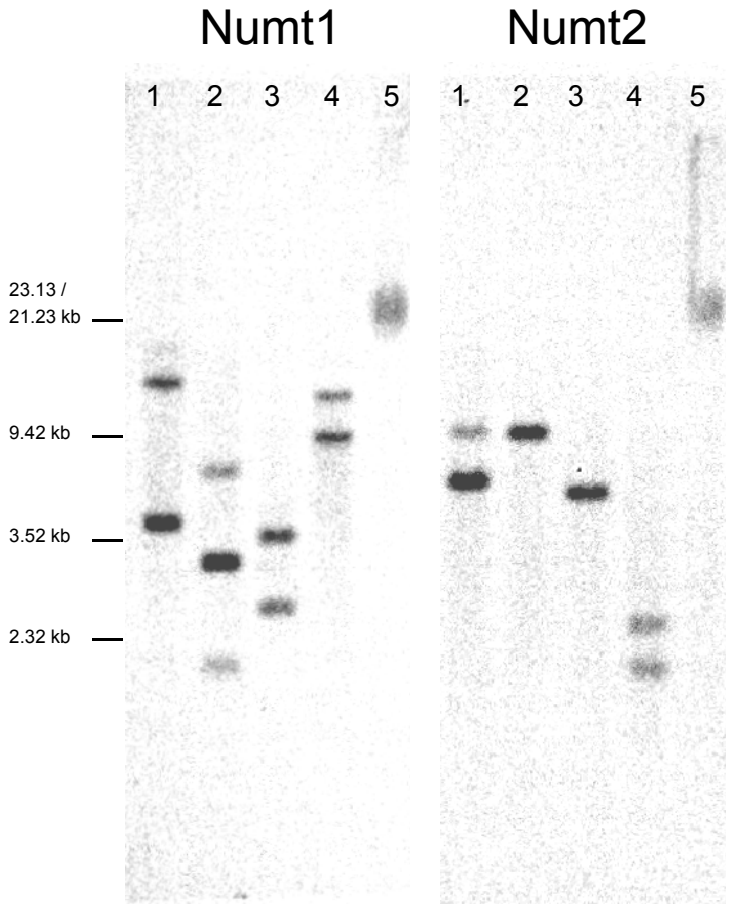

Figure 4-5.
Southern hybridization using Numt1 (left) and Numt2 (right) probes. Lanes 1-4, *P. marinus* gDNA digested with BamHI (1), HindIII (2), PstI (3), and SalI (4); Lane 5, uncut gDNA.

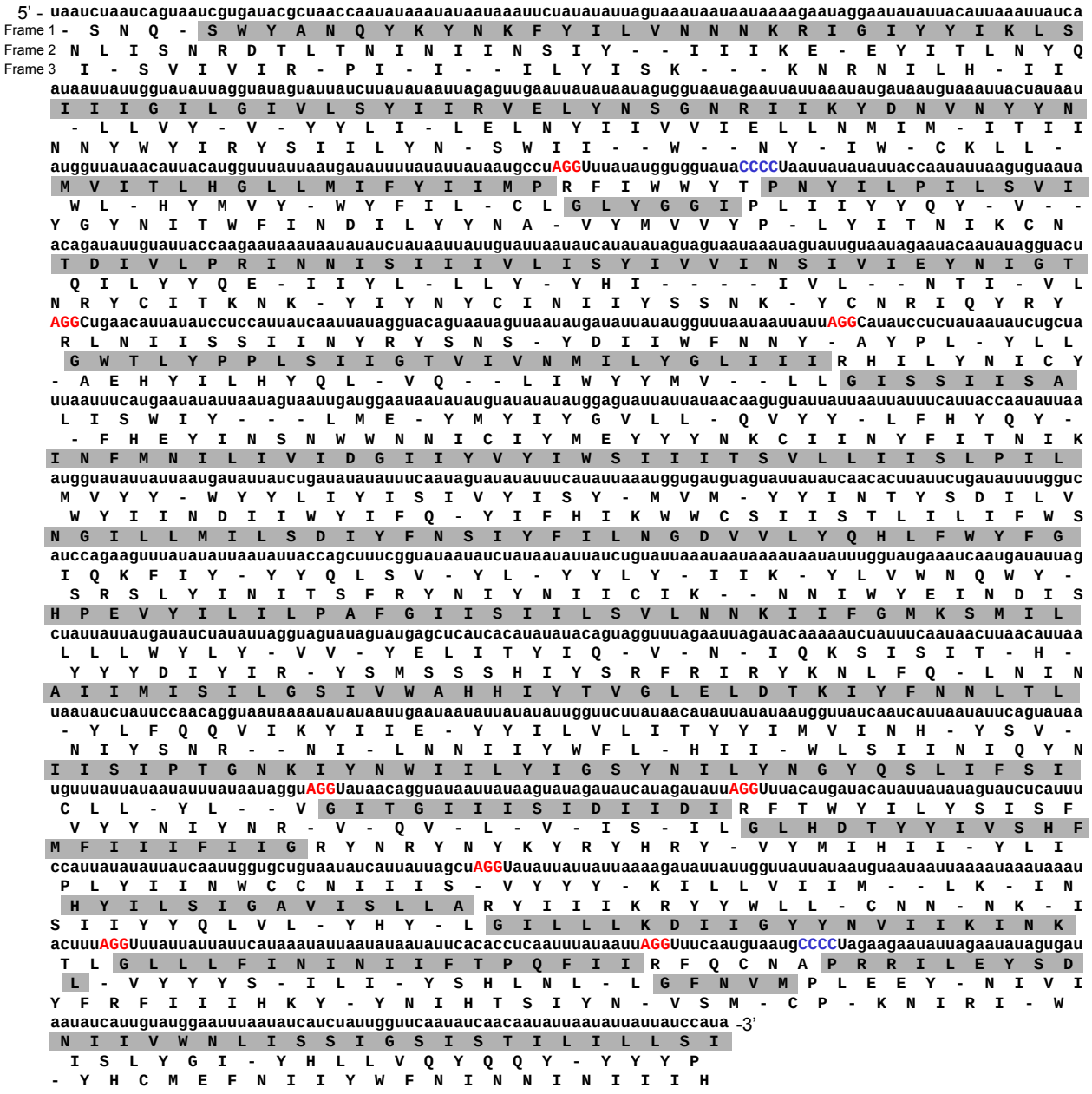

Figure 4-6.
Nucleotide sequence of *Pmcox1* and its three frame translation. Gray boxes denote COX1-like amino acid sequences. Nucleotides in uppercase represent the sequence motifs found in transition areas of the reading frame where COX1-like amino acid sequences are deduced. AGG and CCC codons in the motifs are indicated in red and blue, respectively.

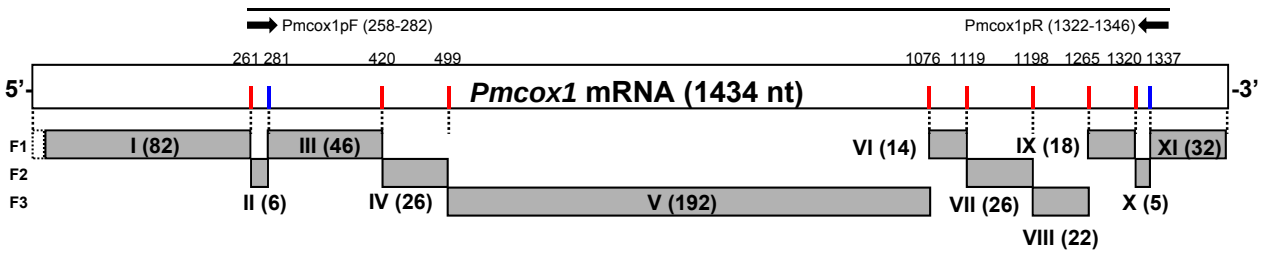

Figure 4-7.
Schematic model of *Pmcox1* mRNA and the COX1-like blocks. COX1-like amino acid sequences distributed across three reading frames are represented by gray boxes labeled with Roman numerals. Numbers in parenthesis adjacent to the Roman numerals indicate the number of the amino acid residues. Red bars at the end of COX1-like blocks indicate AGGY motifs containing in-frame AGG, and blue bars indicate CCCCU motifs including in-frame CCC. Numbers above the open box indicate the positions of the first nucleotide in each motif. The uppermost bar indicates the probe region used for Southern hybridization with arrows indicating the primer locations.

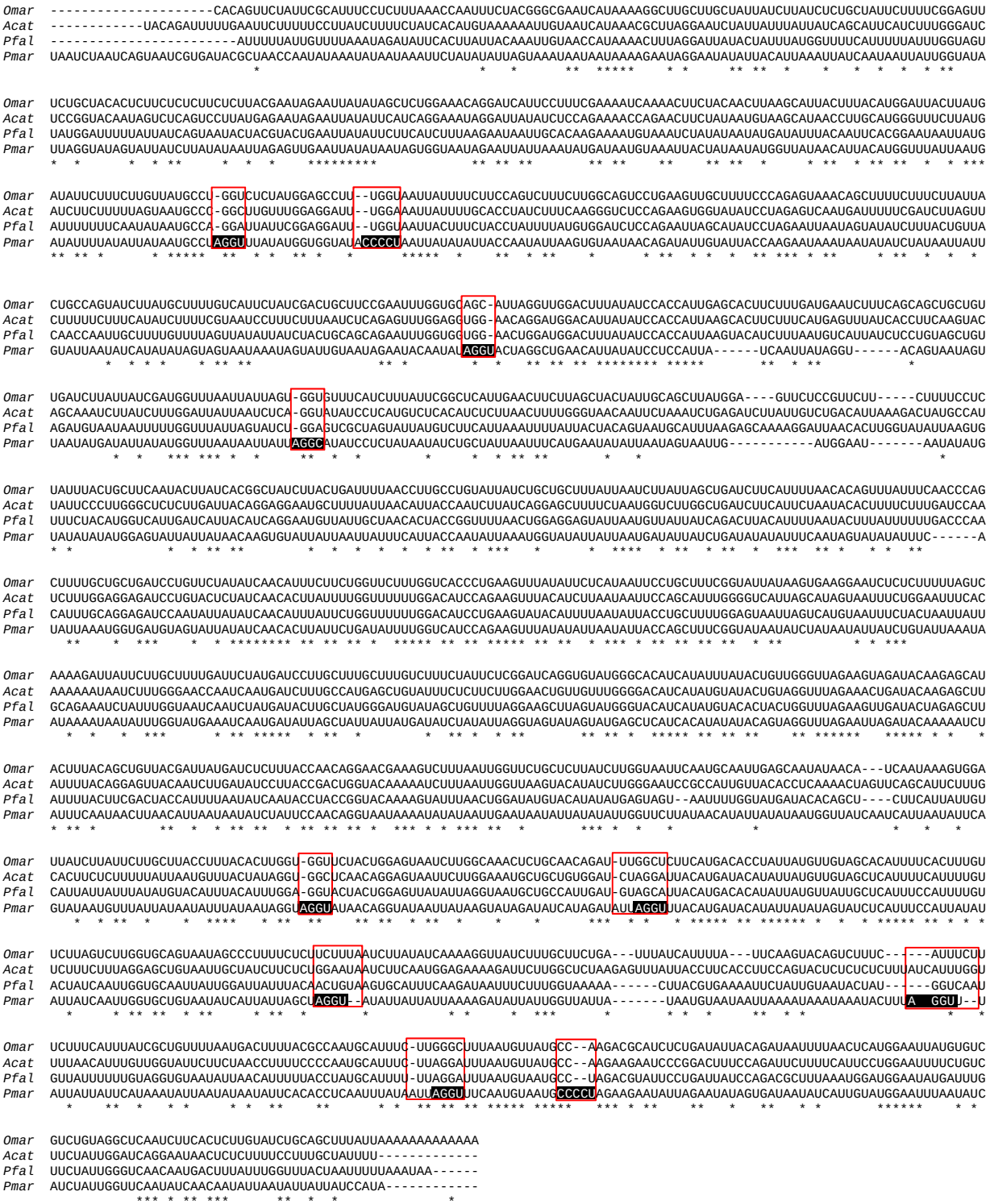

Figure 4-8.
Alignment of several *cox1* sequences and *Pmcox1* mRNA sequence. Asterisks denote conserved nucleotides among four sequences. Regions containing one- or two-base indels specific in *Pmcox1* are indicated with red squares. AGGY and CCCCU motifs in the vicinities of the indels are highlighted in black. The accession numbers of the orthologs from related organisms are described in the legend of Figure 4-2.

Figure 4-9.
Alignment of multiple COX1 amino acid sequences of *P. marinus* and related protists. The sequence for *Perkinsus* was predicted based on the frameshift model. Symbols are described in the same way as in Figure 3-2. Residues highlighted in black are conserved amino acid essential for cytochrome *c* oxidase function: His94 (ligand for heme a), His276 (ligand for CuB), Glu278 (D-channel), His 325, His326 (ligand for CuB), Lys354 (for K-channel), His411 (ligand for heme a3) and His413 (ligand for heme a) The amino acid numbers are according to the *Paracoccus denitrificans* homolog (Iwata *et al.*, 1995). The red and blue arrowheads indicate the motifs (AGGY and CCCCU, respectively) where the reading frame is hypothetically shifted. Note that the glycine and proline residues at the frameshift sites are often highly conserved (highlighted in red and blue, respectively). The open boxes indicate tryptophans coded by UGA codons in *Pmcox1* and those conserved at the homologous positions in related species. The accession numbers of the orthologs from related organisms are described in the legend of Figure 4-2. Note that the ciliate genes are not included because they are highly divergent and the gene length also differs greatly from those of related alveolates.

A

```
Pm mRNA 5′-(15 nt)-UCGUGAUACGCUAACCAAUAUAAAUAUAAUAAAUUCUAUAUAUUAGUAAAUAAUAAUAAAAGAAUAGGAAUAUAUUACAUUAAAUUAUCAAUAAUUAUU-(96 nt)-AUG
Pm          -----  S  W  Y  A  N  Q  Y  K  Y  N  K  F  Y  I  L  V  N  N  N  K  R  I  G  I  Y  Y  I  K  L  S  I  I  I -(32 aa)- M
Om          -----  -  -  H  S  S  I  R  I  S  S  L  N  Q  F  L  R  A  N  H  K  R  L  A  C  Y  Y  L  I  S  A  I  L  F -(41 aa)- M
Ac          ----Y  R  F  L  N  S  F  S  L  S  F  L  S  H  V  K  N  C  N  H  K  R  L  G  I  Y  Y  L  L  S  A  F  I  F -(10 aa)- M
Pf          -----  -  -  -  M  F  I  V  L  N  R  Y  S  L  I  T  N  C  N  H  K  T  L  G  L  Y  Y  L  W  F  S  F  L  F
Tp          MATYS  P  Q  L  N  P  I  Y  K  F  V  T  R  W  L  F  S  T  N  H  K  D  I  G  T  L  Y  L  I  F  G  A  I  S
                                                                 .           *  :  *     :  .        *  :        .     :
```

B

```
Pm mRNA AGAAGAAUAUUAGAAUAUAGUGAUAAUAUCAUUGUAUGGAAUUUAAUAUCAUCUAUUGGUUCAAUAUCAACAAUAUUAAUAUUAUUAUCCAUA- 3′
Pm          R  R  I  L  E  Y  S  D  N  I  I  V  W  N  L  I  S  S  I  G  S  I  S  T  I  L  I  L  L  S  I
Om          R  R  I  S  D  I  T  D  N  F  N  S  W  N  Y  V  S  S  V  G  S  I  F  T  L  V  S  A  A  L  L Kⁿ
Ac          R  R  I  P  D  F  P  D  S  F  H  S  W  N  F  L  S  S  I  G  S  G  I  T  L  F  S  F  A  I  L Kⁿ
Pf          R  R  I  P  D  Y  P  D  A  L  N  G  W  N  M  I  C  S  I  G  S  T  M  T  L  F  G  L  L  I  F -(13 nt)-Kⁿ
Tp          R  R  I  P  D  Y  P  D  A  Y  F  T  F  N  K  I  A  S  W  G  S  Y  V  S  A  I  S  S  L  F  F -(54 nt)-Stop
            *  *  *     :     .  *              :  *     :  .  *     *  *        :     .
```

Figure 4-10.
Alignment of partial COX1 amino acid sequences from *Perkinsus*, dinoflagellates, apicomplexa, and diatom. (A) The 5’ region of *Pmcox1* mRNA and the N terminal region of COX1 from *P. marinus*, dinoflagellates, apicomplexa, and diatom. (B) The 3’ region of *Pmcox1* mRNA and the C terminal region of COX1 from *P. marinus*, dinoflagellates, an apicomplexa, and a diatom. Symbols are described in the same way as in Figure 4-2. AUA and AUU codons in the *Pmcox1* mRNA coding for isoleucine residues are highlighted in gray boxes. The accession numbers of the orthologs from related organisms are described in the legend of Figure 4-2.

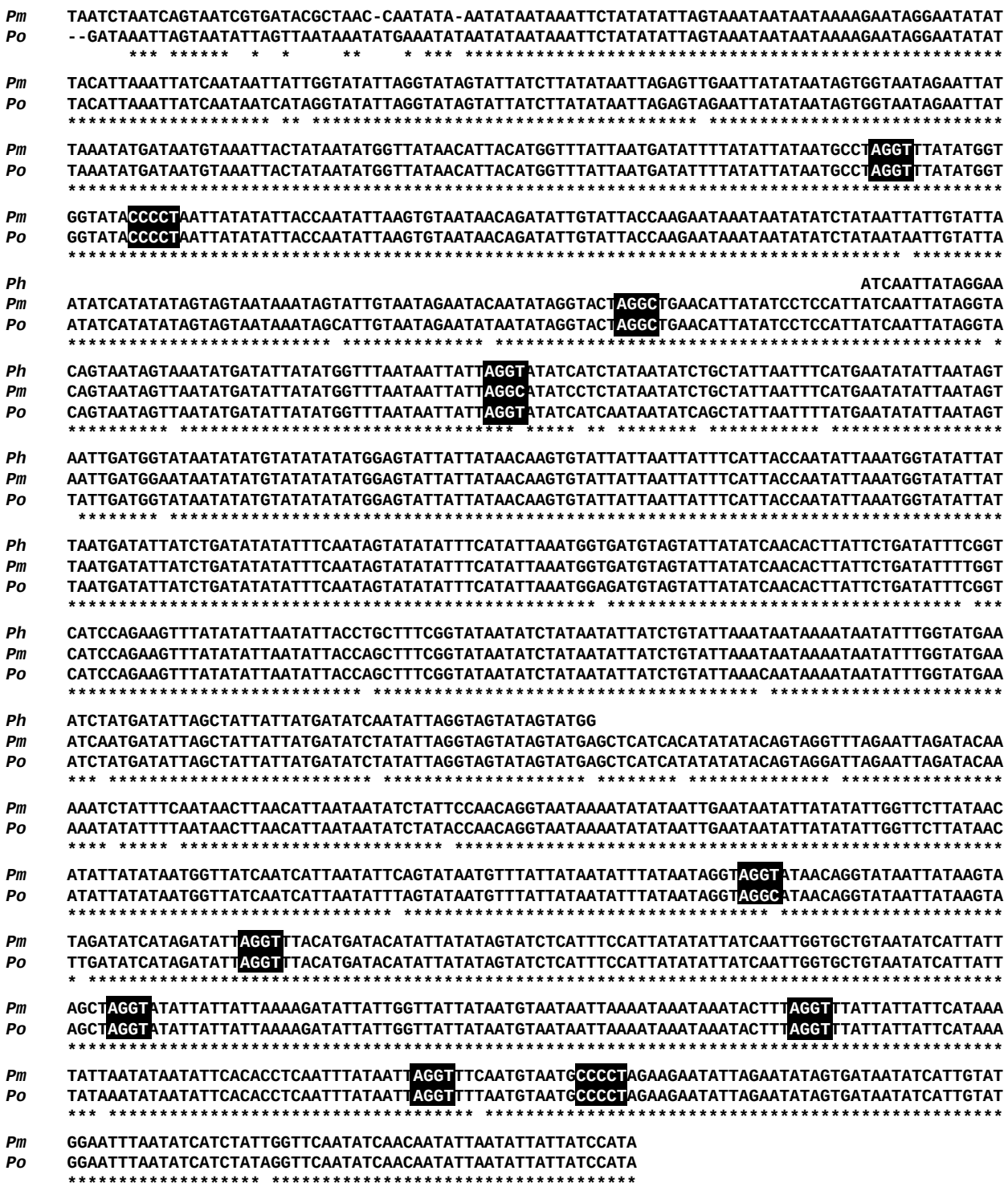

Figure 4-11.

Multiple alignment of *Pmcox1* (*Pm*) and its orthologous sequences obtained from *P. olseni* (*Po*) and *P. honshuensis* (*Ph*). Asterisks show the conserved nucleotides. The motifs which seem to be involved in frameshift are shown in black background. Note that the sequences are described as the DNA sequences.

A

```
5' -Pm ataataataaacttatatatgaatattaacaatggttataaaggaataatatatataatcaaagaaatatattatAGGTtatatattaagatatatacataataa
   Po ----------------------------------------------------------------------------------------------------
Frame 1  I  I  I  N  L  Y  M  N  I  N  N  G  Y  K  G  I  I  Y  I  I  K  E  I  Y  Y  R  L  Y  I  K  I  Y  T  -  -
Frame 2   -  -  -  T  Y  I  -  I  L  T  M  V  I  K  E  -  Y  I  -  S  K  K  Y  I  I  G  Y  I  L  R  Y  I  H  N  N
Frame 3    H  N  N  K  L  I  Y  E  Y  -  Q  W  L  -  R  N  N  I  Y  N  Q  R  N  I  L  -  V  I  Y  -  D  I  Y  I

   Pm taatagtacattaatatatgtagtagtatatttacatattattagaaatttatattataaaacatattattataatatattaatatgatattctggaatgataat
   Po ----------------------------------------------------------------------------------------------------
       -  -  Y  I  N  I  C  S  S  I  F  T  Y  Y  -  K  F  I  L  -  N  I  L  L  -  Y  I  N  M  I  F  W  N  D  N
          N  S  T  L  I  Y  V  V  V  Y  L  H  I  I  R  N  L  Y  Y  K  T  Y  Y  Y  N  I  L  I  W  Y  S  G  M  I  M
       I  I  V  H  -  Y  M  -  -  Y  I  Y  I  L  L  E  I  Y  I  I  K  H  I  I  I  I  Y  -  Y  D  I  L  E  -  -

                                  *************** ************************************************** **
   Pm gttatatcaattaataataattggtttcataggttatatattaggttggggtcaattatcatattggAGGTataacagtaattataaacttaatatctAGGCata
   Po ----------------------------taggttatatattaggatggggtcaattatcatattggAGGTataacagtaattataaacttaatatctAGGTatt
       V  I  S  I  N  N  N  W  F  H  R  L  Y  I  R  L  G  S  I  I  I  L  E  V  -  Q  -  L  -  T  -  Y  L  G  I
          L  Y  Q  L  I  I  I  G  F  I  G  Y  I  L  G  W  G  Q  L  S  Y  W  R  Y  N  S  N  Y  K  L  N  I  -  A  Y
       C  Y  I  N  -  -  -  L  V  S  -  V  I  Y  -  V  G  V  N  Y  H  I  G  G  I  T  V  I  I  N  L  I  S  R  H

       ********************************************************************************** **************************
   Pm ccatatttaatattattaatatct AGGTaattattatataacaattgtaactattaaaagattatatatagtacatttcatattacctataatttttaatatatgt
   Po ccatatttaatattattaatatct AGGTaattattatataacaattgtaactattaaaagattatatatagtacattttatattacctataatttttaatatatgt
       P  Y  L  I  L  L  I  S  R  -  L  L  Y  N  N  C  N  Y  -  K  I  I  Y  S  T  F  H  I  T  Y  N  F  N  I  C
          H  I  -  Y  Y  -  Y  L  G  N  Y  Y  I  T  I  V  T  I  K  R  L  Y  I  V  H  F  I  L  P  I  I  L  I  Y  V
       T  I  F  N  I  I  N  I  -  V  I  I  I  -  Q  L  -  L  L  K  D  Y  I  -  Y  I  S  Y  Y  L  -  F  -  Y  M

       ********************************************** ***************** ***************
   Pm agaaataatacatgtatattatattcattatttaataaataataatatagtagaatataatgttaataataaaataatattcaataattatatattagtaaaaga
   Po agaaataatacatgtatattatattcattatttaataaataataacatagtagaatataatgtaaataataaaataata----------------------------
       R  N  N  T  C  I  L  Y  S  L  F  N  K  -  -  Y  S  R  I  -  C  -  -  -  N  N  I  Q  -  L  Y  I  S  K  R
          E  I  I  H  V  Y  Y  I  H  Y  L  I  N  N  N  I  V  E  Y  N  V  N  N  K  I  I  F  N  N  Y  I  L  V  K  D
       -  K  -  Y  M  Y  I  I  F  I  I  -  -  I  I  I  -  -  N  I  M  L  I  I  K  -  Y  S  I  I  I  Y  -  -  K

   Pm taataatAGGTataa-3'
   Po --------------
       -  -  -  V  -
          N  N  R  Y
       I  I  I  G  I
```

B

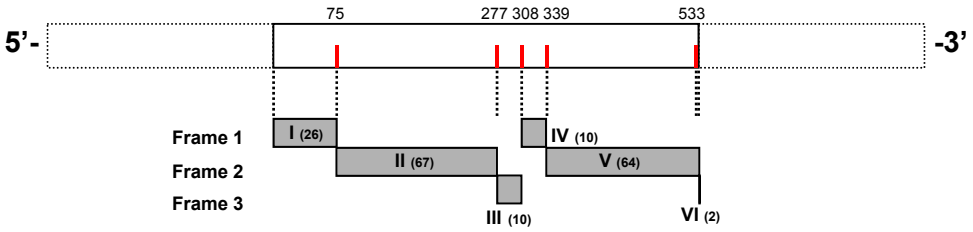

Figure 4-12.
(A) Nucleotide sequences and the three frame translation of *cob*-like fragments from *P. marinus* and *P. olseni*. Amino acid sequences were deduced using the *P. marinus* nucleotide sequence. Note that the sequences are described as the DNA sequences. Asterisks show the conserved nucleotides. (B) Schematic model of the *P. marinus cob*-like fragment and the COB-like blocks dispersed in three reading frames. Detail descriptions are found in the legends of Figures 4-6 and 4-7, respectively.

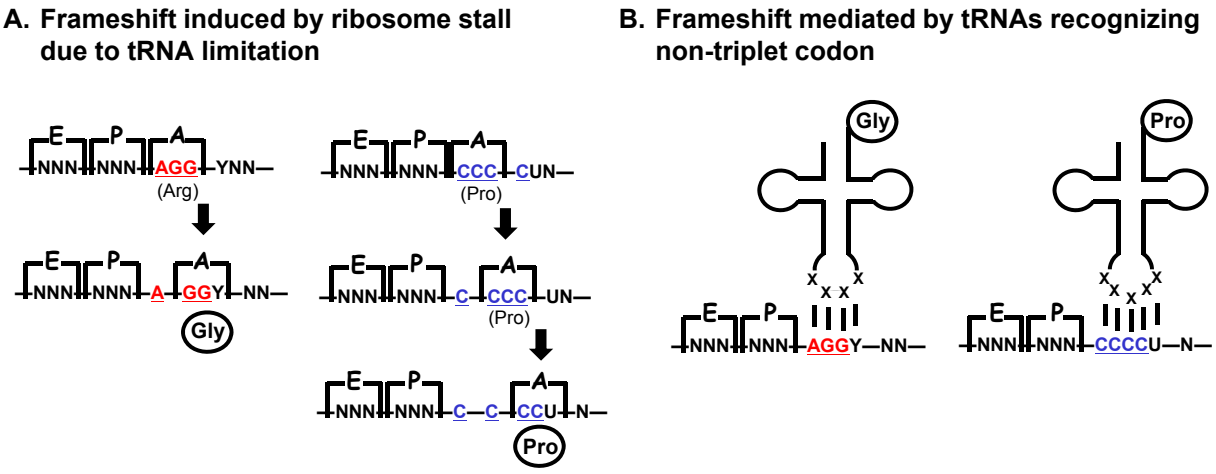

Figure 4-13.
Possible mechanisms of frameshift-dependent translation at AGG and CCC codons of *Perkinsus* mt genes. (A) A ribosome stalled by tRNA limitation induces a ribosomal frameshift. When an in-frame AGG or CCC moves into the ribosome A site during translation, the ribosome stalls due to the limitation of the corresponding tRNA molecules, and the reading frame is subsequently shifted to the +1 frame. Translation then restarts in the new frame. In the case of CCCCU, two consecutive frameshifts occur at CCC codons. (B) Specialized tRNAs recognize non-triplet codons, AGGY and CCCCU, and the reading frame shifts forward by one (at AGGY) or two (at CCCCU) bases.

SUMMARY AND CONCLUSIONS

In this thesis, I analyzed mitochondria of a shellfish pathogen *Perkinsus*, a relative of apicomplexans, as a pioneering work. Initially, I established the preliminary protocol for the enrichment of *Perkinsus* mitochondria using ultracentrifugal fractionation and I obtained the first biochemical evidence for the mitochondrial ETC from *Perkinsus*. Although this protocol needs refinements, the enrichment/purification of mitochondria would surely contribute to the further specific study. Next I investigated the entire mitochondrial ETC organization using both genomic and biochemical approaches. By BLAST-based sequence searches, it was revealed that *P. marinus* is likely to possess the mitochondrial ETC complexes II-IV, ATP synthase, NDH2, and AOX. Although a few of essential subunits were missing, the enzyme activities of the complexes II-IV were experimentally demonstrated. Also, the unconventional enzymes NDH2 and the AOX would probably contribute to the *Perkinsus* ETC. This is the first report on the active mitochondrial ETC of *Perkinsus*, and the existence of these unconventional enzymes implies the similarity in the mitochondrial ETC between *Perkinsus* and apicomplexan parasites. These observations, along with the ease in handling, prompt us to consider that *Perkinsus* may be a highly suitable material to investigate the properties of these enzymes, which are the candidates of drug targets in apicomplexan parasites.

To characterize the genomic aspects of *Perkinsus* mitochondria, I determined the full-length mRNA sequence of *Pmcox1*, the first mt gene sequence of *Perkinsus*. This gene was extremely AT-rich, and did not bear canonical start and stop codons in the terminal regions apparently. The absence of these codons is observed in mt genes of apicomplexans and dinoflagellates, and it is assumed that *Perkinsus* and these evolutionarily related organisms may share alternative start and stop mechanisms in

translation, which might be another candidate of the antiparasitic drug target. Surprisingly, the close examination of its nucleotide sequence suggested that an unusually frequent translational frameshift is occurring in *Perkinsus* mitochondria. The frameshifts at all AGG and CCC codons during translation, ten times in total, can produce the plausible COX1 amino acid sequence. The conservation of the frameshift motifs in the *P. olseni cox1* and the partial *cob* sequences both from *P. marinus* and *P. olseni* strongly endorsed this frameshift model. This unusual feature is only observed in *Perkinsus* and not in related species, therefore this organism acquired this unique character independently of other related species. Most remarkably, it should be noted that this frameshift is infinitely more frequent than any other case reported to date. It was therefore proposed that active and efficient machinery allowing the frequent frameshifts is utilized during translation in *Perkinsus* mitochondria (published as Masuda *et al*., 2010). In-depth characterization of the unprecedented frameshift mechanism in *Perkinsus* may elucidate the mechanism of translational reading frame maintenance and of accurate codon recognition, and they will be invaluable insights into the essence of translation. I would be able to unveil the mechanism by analyzing mitochondrial translational components with the aid of the improved method.

As just described, this study encompasses basic and essential information in studying *Perkinsus* mitochondria, ensures the biological similarity in mitochondria amongst the related organisms, and presents a discovery of a unique and tantalizing phenomenon. Future studies on *Perkinsus* mitochondria would probably provide valuable information on the derived mitochondrial ETC and mt gene expression system of apicomplexans and dinoflagellates, and elucidate the uncharacterized mechanism in translation that we have never imagined.

PERSPECTIVES

As this is the first study that focuses on *Perkinsus* mitochondria, future studies on *Perkinsus* mitochondria would be operated referring to this study as a starting point. I demonstrated several similarities in mitochondria amongst *Perkinsus*, apicomplexans and dinoflagellates. From the viewpoints of parasitology and health science, the insights from *Perkinsus* will surely serve as a good reference in analyzing the divergent mitochondrial metabolic pathways with particular emphasis on the ETC, and the highly derived gene expression system for antiparasitic strategies.

The extremely frequent translational frameshift that I discovered represents an unusual biological phenomenon. I will characterize the mt gene expression system of *Perkinsus* specifically, and identify the components involved in translation in *Perkinsus* mitochondria like tRNAs, ribosomes, and aminoacyl-tRNA synthases. For this, I need to develop the more sophisticated purification method for mitochondria. Elucidation of the mechanism of unprecedented translational events would delineate modified molecular machinery, which extends our general view on decoding with a great scientific impact on basic molecular biology. Furthermore, considering the phylogenetic position of *Perkinsus* and its threat to fisheries, this exciting theme and future related studies must be highly informative also for evolutionary protistology and fisheries sciences. Thus, this study is expected to lead new ways in and contribute to a wide range of science fields in the near future.

REFERENCES

Abrahamsen,M.S., Templeton,T.J., Enomoto,S., Abrahante,J.E., Zhu,G., Lancto,C.A., Deng,M., Liu,C., Widmer,G., Tzipori,S., Buck,G.A., Xu,P., Bankier,A.T., Dear,P.H., Konfortov,B.A., Spriggs,H.F., Iyer,L., Anantharaman,V., Aravind,L. and Kapur,V. Complete genome sequence of the apicomplexan, *Cryptosporidium parvum*. *Science*, **304**, 441-445. (2004)

Adl,S.M., Simpson,A.G., Farmer,M.A., Andersen,R.A., Anderson,O.R., Barta,J.R., Bowser,S.S., Brugerolle,G., Fensome,R.A., Fredericq,S., James,T.Y, Karpov,S., Kugrens,P., Krug,J., Lane,C.E., Lewis,L.A., Lodge,J., Lynn,D.H., Mann,D.G., McCourt,R.M., Mendoza,L., Moestrup,O., Mozley-Standridge,S.E., Nerad,T.A., Shearer,C.A., Smirnov,A.V., Spiegel,F.W. and Taylor,M.F. The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. *J. Eukaryot. Microbiol*., **52**, 399-451. (2005)

Anderson,J.C., Wu,N., Santoro,S.W., Lakshman,V., King,D.S. and Schultz,P.G. An expanded genetic code with a functional quadruplet codon. *Proc. Natl. Acad. Sci. USA*, **101**, 7566-7571. (2004)

Atkins,J.F. and Baranov,P.V. The distinction between recoding and codon reassignment. *Genetics*, **185**, 1535-1536. (2010)

Atkins,J.F. and Björk,G.R. A gripping tale of ribosomal frameshifting: extragenic suppressors of frameshift mutations spotlight P-site realignment. *Microbiol. Mol. Biol. Rev*., **73**, 178-210. (2009)

Balabaskaran Nina,P., Dudkina,N.V., Kane,L.A., van Eyk,J.E., Boekema,E.J., Mather,M.W. and Vaidya,A.B. Highly divergent mitochondrial ATP synthase complexes in *Tetrahymena thermophila*. *PLoS Biol*., **8**, e1000418. (2010)

Baranov,P.V., Gesteland,R.F. and Atkins,J.F. Recoding: translational bifurcations in gene expression. *Gene*, **286**, 187-201. (2002)

Barton,V., Fisher,N., Biagini,G.A., Ward,S.A. and O'Neill,P.M. Inhibiting *Plasmodium* cytochrome bc_1: a complex issue. *Curr. Opin. Chem. Biol.*, **14**, 440-446. (2010)

Beckenbach,A.T., Robson,S.K. and Crozier,R.H. Single nucleotide +1 frameshifts in an apparently functional mitochondrial cytochrome *b* gene in ants of the genus *Polyrhachis*. *J. Mol. Evol.*, **60**, 141-152. (2005)

Belcourt,M.F. and Farabaugh,P.J. Ribosomal frameshifting in the yeast retrotransposon Ty: tRNAs induce slippage on a 7 nucleotide minimal site. *Cell*, **62**, 339-352. (1990)

Biagini,G.A., Viriyavejakul,P., O'neill,P.M., Bray,P.G. and Ward,S.A. Functional characterization and target validation of alternative complex I of *Plasmodium falciparum* mitochondria. *Antimicrob. Agents Chemother.*, **50**, 1841-1851. (2006)

Bråte,J., Logares,R., Berney,C., Ree,D.K., Klaveness,D., Jakobsen,K.S. and Shalchian-Tabrizi,K. Freshwater Perkinsea and marine-freshwater colonizations revealed by pyrosequencing and phylogeny of environmental rDNA. *ISME J.*, **4**, 1144-1153. (2010)

Burger,G., Forget,L., Zhu,Y., Gray,M.W. and Lang,B.F. Unique mitochondrial genome architecture in unicellular relatives of animals. *Proc. Natl. Acad. Sci. USA*, **100**, 892-897. (2003)

Burger,G., Gray,M.W. and Lang,B.F. Mitochondrial genomes: anything goes. *Trends Genet.*, **19**, 709-716. (2003)

Cavalier-Smith,T. Kingdom protozoa and its 18 phyla. *Microbiol. Rev.*, **57**, 953-994. (1993)

Chaput,H., Wang,Y. and Morse,D. Polyadenylated transcripts containing random gene fragments are expressed in dinoflagellate mitochondria. *Protist*, **153**, 111-122. (2002)

Chaudhuri,M., Ott,R.D. and Hill,G.C. Trypanosome alternative oxidase: from molecule to function. *Trends Parasitol.*, **22**, 484-491. (2006)

Choi,K.S. and Park,K.I. Review on the Protozoan Parasite *Perkinsus olseni* (Lester and Davis 1981) Infection in Asian Waters. In: Ishimatsu,E.A. and Lie,H.-J. (eds), *Coastal*

Environmental and Ecosystem Issues of the East China Sea, 269–281. (2010)

Claros,M.G. and Vincens,P. Computational method to predict mitochondrially imported proteins and their targeting sequences. *Eur. J. Biochem.*, **241**, 779-786. (1996)

Clifton,R., Millar,A.H. and Whelan,J. Alternative oxidases in *Arabidopsis*: a comparative analysis of differential expression in the gene family provides new insights into function of non-phosphorylating bypasses. *Biochim. Biophys. Acta*, **1757**, 730-741. (2006)

Conway,D.J., Fanello,C., Lloyd,J.M., Al-Joubori,B.M., Baloch,A.H., Somanath,S.D., Roper,C., Oduola,A.M., Mulder,B., Povoa,M.M., Singh,B. and Thomas,A.W. Origin of *Plasmodium falciparum* malaria is traced by mitochondrial DNA. *Mol. Biochem. Parasitol.*, **111**, 163-171. (2000)

Crofts,A.R. The cytochrome *bc*1 complex: function in the context of structure. *Annu. Rev. Physiol.*, **66**, 689-733. (2004)

Devenish,R.J., Prescott,M. and Rodgers,A.J. The structure and function of mitochondrial F_1F_O-ATP synthases. *Int. Rev. Cell Mol. Biol.*, **267**, 1-58. (2008)

Eriksson,M., Gardestrom,P. and Samuelsson,G. Isolation, purification, and characterization of mitochondria from *Chlamydomonas reinhardtii*. *Plant Physiol.*, **107**, 479-483. (1995)

Esseiva,A.C., Naguleswaran,A., Hemphill,A. and Schneider,A. Mitochondrial tRNA import in *Toxoplasma gondii*. *J. Biol. Chem.*, **279**, 42363-42368. (2004)

Farabaugh,P.J. Translational frameshifting: implications for the mechanism of translational frame maintenance. *Prog. Nucleic Acid Res. Mol. Biol.*, **64**, 131-170. (2000)

Feagin,J.E. The 6-kb element of *Plasmodium falciparum* encodes mitochondrial cytochrome genes. *Mol. Biochem. Parasitol.*, **52**, 145-148. (1992)

Fernández-Robledo,J.A., Lin,Z. and Vasta,G.R. Transfection of the protozoan parasite *Perkinsus marinus*. *Mol. Biochem. Parasitol.*, **157**, 44-53. (2008)

Fernández-Robledo,J.A., Schott,E.J. and Vasta,G.R. *Perkinsus marinus* superoxide dismutase 2 (PmSOD2) localizes to single-membrane subcellular compartments. *Biochem. Biophys. Res. Commun.*, **375**, 215-219. (2008)

Fölsch,H., Guiard,B., Neupert,W. and Stuart,R.A. Internal targeting signal of the BCS1 protein: a novel mechanism of import into mitochondria. *EMBO J.*, **15**, 479-487. (1996)

Gardner,M.J. A status report on the sequencing and annotation of the *P. falciparum* genome. *Mol. Biochem. Parasitol.*, **118**, 133-138. (2001)

Gauthier,J.D. and Vasta,G.R. *In vitro* culture of the eastern oyster parasite *Perkinsus marinus*: optimization of the methodology. *J. Invertebr. Pathol.*, **66**, 156-168. (1995)

Gesteland,R.F., Weiss,R.B. and Atkins,J.F. Recoding: reprogrammed genetic decoding. *Science*, **257**, 1640-1641. (1992)

Gesteland,R.F. and Atkins,J.F. Recoding: dynamic reprogramming of translation. *Annu. Rev. Biochem.*, **65**, 741-768. (1996)

Giedroc,D.P. and Cornish,P.V. Frameshifting RNA pseudoknots: structure and mechanism. *Virus Res.*, **139**, 193-208. (2009)

Gillespie,D.E., Salazar,N.A., Rehkopf,D.H. and Feagin,J.E. The fragmented mitochondrial ribosomal RNAs of *Plasmodium falciparum* have short A tails. *Nucleic Acids Res.*, **27**, 2416-2422. (1999)

Glöckner,G. Large scale sequencing and analysis of AT rich eukaryote genomes. *Curr. Genomics*, **1**, 289-299. (2000)

Gray,M.W., Lang,B.F. and Burger,G. Mitochondria of protists. *Annu. Rev. Genet.*, **38**, 477-524. (2004)

Grosjean,H., de Crécy-Lagard,V. and Marck,C. Deciphering synonymous codons in the three domains of life: co-evolution with specific tRNA modification enzymes. *FEBS Lett.*, **584**, 252-64. (2010)

Hallegraeff,G.M. A review of harmful algal blooms and their apparent global increase. *Phycologia*, **32**, 79–99. (1993)

Hamaguchi,M., Suzuki,N., Usuki,H. and Ishioka,H. *Perkinsus* protozoan infection in short-necked clam *Tapes* (=*Ruditapes*) *philipinarum* in Japan. *Fish Pathol.*, **33**, 473-480. (1998)

Härlid,A., Janke,A. and Arnason,U. The mtDNA sequence of the ostrich and the divergence between paleognathous and neognathous birds. *Mol. Biol. Evol*., **14**, 754-761. (1997)

Hikosaka,K., Watanabe,Y., Tsuji,N., Kita,K., Kishine,H., Arisue,N., Palacpac,N.M., Kawazu,S., Sawai,H., Horii,T., Igarashi,I. and Tanabe,K. Divergence of the mitochondrial genome structure in the apicomplexan parasites, *Babesia* and *Theileria*. *Mol. Biol. Evol.*, **27**, 1107-1116. (2010)

Hohsaka,T., Ashizuka,Y., Murakami,H. and Sisido,M. Five-base codons for incorporation of nonnatural amino acids into proteins. *Nucleic Acids Res*., **29**, 3646-3651. (2001)

Horváth,A., Nebohácova,M., Lukeš,J. and Maslov,D.A. Unusual polypeptide synthesis in the kinetoplast-mitochondria from *Leishmania tarentolae*. Identification of individual *de novo* translation products. *J. Biol. Chem.*, **277**, 7222-7230. (2002)

Howard,M.T., Aggarwal,G., Anderson,C.B., Khatri,S., Flanigan,K.M. and Atkins,J.F. Recoding elements located adjacent to a subset of eukaryal selenocysteine-specifying UGA codons. *EMBO J.*, **24**, 1596-1607. (2005)

Iwata,S., Ostermeier,C., Ludwig,B. and Michel,H. Structure at 2.8 A resolution of cytochrome *c* oxidase from *Paracoccus denitrificans*. *Nature,* **376**, 660-669. (1995)

Jackson,C.J., Norman,J.E., Schnare,M.N., Gray,M.W., Keeling,P.J. and Waller,R.F. Broad genomic and transcriptional analysis reveals a highly derived genome in dinoflagellate mitochondria. *BMC Biol.,* **5**, 41. (2007)

Jacobs,H.T. and Turnbull,D.M. Nuclear genes and mitochondrial translation: a new class of genetic disease. *Trends Genet.*, **21**, 312-314. (2005)

Janouškovec,J., Horák,A., Oborník,M., Lukeš,J. and Keeling,P.J. A common red algal origin of the apicomplexan, dinoflagellate, and heterokont plastids. *Proc. Natl. Acad. Sci. USA*, **107**, 10949-10954. (2010)

Ji,Y.E., Mericle,B.L., Rehkopf,D.H., Anderson,J.D. and Feagin,J.E. The *Plasmodium falciparum* 6 kb element is polycistronically transcribed. *Mol. Biochem. Parasitol.*, **30**, 211-223. (1996)

Joseph,S.J., Fernández-Robledo,J.A., Gardner,M.J., El-Sayed,N.M., Kuo,C.-H., Schott,E.J., Wang,H., Kissinger,J.G. and Vasta,G.R. The alveolate *Perkinsus marinus*: Biological insights from EST gene discovery. *BMC Genomics*, **11**, 228. (2010)

Kamikawa,R., Nishimura,H. and Sako,Y. Analysis of the mitochondrial genome, transcripts, and electron transport activity in the dinoflagellate *Alexandrium catenella* (Gonyaulacales, Dinophyceae). *Phycol. Res.*, **57**, 1-11. (2009)

Kawahara,K., Mogi,T., Tanaka,T.Q., Hata,M., Miyoshi,H. and Kita,K. Mitochondrial dehydrogenases in the aerobic respiratory chain of the rodent malaria parasite *Plasmodium yoelii yoelii*. *J. Biochem.*, **145**, 229-237. (2009)

Kita,K., Miyadera,H., Saruta,F. and Miyoshi,H. Parasite mitochondria as a target for chemotherapy. *J. Health Sci.*, **47**, 219–239. (2001)

Kita,K., Shiomi,K. and Omura,S. Advances in drug discovery and biochemical studies. *Trends Parasitol*, **23**, 223-229. (2007)

Kita,K., Vibat,C.R., Meinhardt,S., Guest,J.R. and Gennis,R.B. One-step purification from *Escherichia coli* of complex II (succinate: ubiquinone oxidoreductase) associated with succinate-reducible cytochrome b_{556}. *J. Biol. Chem.*, **264**, 2672-2677 (1989)

Klobutcher,L.A. Sequencing of random *Euplotes crassus* macronuclear genes supports a high frequency of +1 translational frameshifting. *Eukaryot. Cell*. **4**, 2098-2105. (2005)

Kobayashi,T. and Endoh,H. A possible role of mitochondria in the apoptoticlike programmed nuclear death of *Tetrahymena thermophila*. *FEBS J*., **272**, 5378–5387. (2005)

Kobayashi,T., Morone,N., Kashiyama,T., Oyamada,H., Kurebayashi,N. and Murayama,T. Engineering a novel multifunctional green fluorescent protein tag for a wide variety of protein research. *PLoS One*, **3**, e3822. (2008)

Kobayashi,T., Sato,S., Takamiya,S., Komaki-Yasuda,K., Yano,K., Hirata,A., Onitsuka,I., Hata,M., Mi-ichi,F., Tanaka,T., Hase,T., Miyajima,A., Kawazu,S., Watanabe,Y. and Kita,K. Mitochondria and apicoplast of *Plasmodium falciparum*: behaviour on subcellular fractionation and the implication. *Mitochondrion*, **7**, 125-132. (2007)

Kondow,A., Suzuki,T., Yokobori,S., Ueda,T. and Watanabe,K. An extra $tRNA^{Gly}$(U*CU) found in ascidian mitochondria responsible for decoding non-universal codons AGA/AGG as glycine. *Nucleic Acids Res.*, **27**, 2554-2559. (1999)

Kück,U., Jekosch,K. and Holzamer,P. DNA sequence analysis of the complete mitochondrial genome of the green alga *Scenedesmus obliquus*: evidence for UAG being a leucine and UCA being a non-sense codon. *Gene*, **253**, 13-18. (2000)

La Peyre,J.F., Faisal,M. and Burreson,E.M. *In vitro* propagation of the protozoan *Perkinsus marinus*, a pathogen of the eastern oyster, *Crassostrea virginica*. *J. Euk. Microbiol.*, **40**, 304-310, (1993)

Laforest,M.J., Roewer,I., and Lang,B.F. Mitochondrial tRNAs in the lower fungus *Spizellomyces punctatus*: tRNA editing and UAG 'stop' codons recognized as leucine. *Nucleic Acids Res.*, **25**, 626-632. (1997)

Lang,B.F., Gray,M.W. and Burger,G. Mitochondrial genome evolution and the origin of eukaryotes. *Annu. Rev. Genet.*, **33**, 351-397. (1999)

Leander,B.S. and Keeling,P.J. Morphostasis in alveolate evolution. *Trends Ecol. Evol.*, **18**, 395-402. (2003)

Lee,C.M., Sedman,J., Neupert,W. and Stuart,R.A. The DNA helicase, Hmi1p, is transported into mitochondria by a C-terminal cleavable targeting signal. *J. Biol. Chem.*, **274**, 20937-20942. (1999)

Lee,F.W. and Lo,S.C. The use of Trizol reagent (phenol/guanidine isothiocyanate) for producing high quality two-dimensional gel electrophoretograms (2-DE) of dinoflagellates. *J. Microbiol. Methods.*, **73**, 26-32. (2008)

Levine,N.D. *Perkinsus* gen. n. and other new taxa in the protozoan phylum Apicomplexa. *J. Parasitol.*, **64**, 549. (1978)

Levine,N.D. Progress in taxonomy of the Apicomplexan protozoa. *J. Protozool.*, **35**, 518-520. (1988)

Levine,N.D., Corliss,J.O., Cox,F.E., Deroux,G., Grain,J., Honigberg,B.M., Leedale,G.F., Loeblich,A.R. 3rd, Lom,J., Lynn,D., Merinfeld,E.G., Page,F.C., Poljansky,G., Sprague,V., Vavra,J. and Wallace,F.G. A newly revised classification of the protozoa. *J. Protozool.*, **27**, 37-58. (1980)

Lin,S.S., Gross,U. and Bohne,W. Type II NADH dehydrogenase inhibitor 1-hydroxy-2-dodecyl-4(1*H*)quinolone leads to collapse of mitochondrial inner-membrane potential and ATP depletion in *Toxoplasma gondii*. *Eukaryot. Cell*, **8**, 877-887. (2009)

Lopez,J.V., Yuhki,N., Masuda,R., Modi,W. and O'Brien,S.J. Numt, a recent transfer and tandem amplification of mitochondrial DNA to the nuclear genome of the domestic cat. *J. Mol. Evol.*, **39**, 174–190. (1994)

Mackin,J.G., Owen,H.M. and Collier,A. Preliminary note on the occurrence of a new protistan parasite, *Dermocystidium marinum* n. sp. in *Crassostrea virginica* (Gmelin). *Science*, **111**, 328-329. (1950)

Mackin,J.G. and Ray,S.M. The taxonomic relationships of *Dermocystidium marinum* Mackin, Owen, and Collier. *J. Invertebr. Pathol.*, **8**, 544-545. (1966)

Magliery,T.J., Anderson,J.C. and Schultz,P.G. Expanding the genetic code: selection of efficient suppressors of four-base codons and identification of "shifty" four-base codons with a library approach in *Escherichia coli*. *J. Mol. Biol*., **307**, 755-769. (2001)

Marande,W. and Burger,G. Mitochondrial DNA as a genomic jigsaw puzzle. *Science*, **318**, 415. (2007)

Marande,W., Lukeš,J. and Burger,G. Unique mitochondrial genome structure in diplonemids, the sister group of kinetoplastids. *Eukaryot. Cell*, **4**, 1137-1146. (2005)

Marres,C.A., de Vries,S. and Grivell,L.A. Isolation and inactivation of the nuclear gene encoding the rotenone-insensitive internal NADH: ubiquinone oxidoreductase of mitochondria from *Saccharomyces cerevisiae*. *Eur. J. Biochem.*, **195**, 857-862. (1991); Erratum in: *Eur. J. Biochem*., **200**, 812. (1991)

Masuda,I., Matsuzaki,M. and Kita,K. Extensive frameshift at all AGG and CCC codons in the mitochondrial cytochrome *c* oxidase subunit 1 gene of *Perkinsus marinus* (Alveolata; Dinoflagellata). *Nucleic Acids Res.*, **38**, 6186-6194. (2010)

Mather,M.W., Henry,K.W. and Vaidya,A.B. Mitochondrial drug targets in apicomplexan parasites. *Curr. Drug Targets*, **8**, 49-60. (2007)

Matsumoto,Y., Muneyuki,E., Fujita,D., Sakamoto,K., Miyoshi,H., Yoshida,M. and Mogi,T. Kinetic mechanism of quinol oxidation by cytochrome *bd* studied with ubiquinone-2 analogs. *J. Biochem.*, **13**9, 779-788. (2006)

Matsuzaki,M., Kuroiwa,H., Kuroiwa,T., Kita,K. and Nozaki,H. A cryptic algal group unveiled: a plastid biosynthesis pathway in the oyster parasite *Perkinsus marinus*. *Mol. Biol. Evol.*, **25**, 1167-1179. (2008)

Melo,A.M., Bandeiras,T.M. and Teixeira,M. New insights into type II NAD(P)H:quinone oxidoreductases. *Microbiol. Mol. Biol. Rev*., **68**, 603-616. (2004)

Mentel,M., Zimorski,V., Haferkamp,P., Martin,W. and Henze,K. Protein import into hydrogenosomes of *Trichomonas vaginalis* involves both N-terminal and internal

targeting signals: a case study of thioredoxin reductases. *Eukaryot. Cell*, **7**, 1750-1757. (2008)

Milbury,C.A. and Gaffney,P.M. Complete mitochondrial sequence of the eastern oyster *Crassostrea virginica*. *Mar Biotechnol (NY).*, **7**, 697-712. (2008)

Minagawa,N., Yabu,Y., Kita,K., Nagai,K., Ohta,N., Meguro,K., Sakajo,S. and Yoshimoto,A. An antibiotic, ascofuranone, specifically inhibits respiration and *in vitro* growth of long slender bloodstream forms of *Trypanosoma brucei brucei*. *Mol. Biochem. Parasitol.*, **84**, 271-280. (1997)

Mindell,D.P., Sorenson,M.D. and Dimcheff,D.E. An extra nucleotide is not translated in mitochondrial ND3 of some birds and turtles. *Mol. Biol. Evol.*, **15**, 1568-1571. (1998)

Miyadera,H., Shiomi,K., Ui,H., Yamaguchi,Y., Masuma,R., Tomoda,H., Miyoshi,H., Osanai,A., Kita,K. and Omura,S. Atpenins, potent and specific inhibitors of mitochondrial complex II (succinate-ubiquinone oxidoreductase). *Proc. Natl. Acad. Sci. U S A*, **100**, 473-477. (2003)

Mogi,T. and Kita,K. Identification of mitochondrial Complex II subunits SDH3 and SDH4 and ATP synthase subunits *a* and *b* in *Plasmodium* spp. *Mitochondrion*, **9**, 443-453. (2009)

Mogi,T and Kita,K. Diversity in mitochondrial metabolic pathways in parasitic protists *Plasmodium* and *Cryptosporidium*. *Parasitol. Int.*, **59**, 305-12. (2010)

Moore,A,L, and Albury,M.S. Further insights into the structure of the alternative oxidase: from plants to parasites. *Biochem. Soc. Trans.*, **36**, 1022-1026. (2008)

Moore,R.B., Oborník,M., Janouškovec,J., Chrudimský,T., Vancová,M., Green,D.H., Wright,S.W., Davies,N.W., Bolch,C.J., Heimann,K., Slapeta,J., Hoegh-Guldberg,O., Logsdon,J.M. and Carter,D.A. A photosynthetic alveolate closely related to apicomplexan parasites. *Nature*, **451**, 959-963 (2008); Erratum in: *Nature*, **452**, 900. (2008)

Nakamura,K., Yamaki,M., Sarada,M., Nakayama,S., Vibat,C.R., Gennis,R.B.,

Nakayashiki,T., Inokuchi,H., Kojima,S. and Kita,K. Two hydrophobic subunits are essential for the heme *b* ligation and functional assembly of complex II (succinate-ubiquinone oxidoreductase) from *Escherichia coli*. *J. Biol. Chem.*, **271**, 521-527. (1996)

Nash,E.A., Barbrook,A.C., Edwards-Stuart,R.K., Bernhardt,K., Howe,C.J. and Nisbet,R.E. Organization of the mitochondrial genome in the dinoflagellate *Amphidinium carterae*. *Mol. Biol. Evol.*, **24**, 1528-1536. (2007)

Nash,E.A., Nisbet,R.E., Barbrook,A.C. and Howe,C.J. Dinoflagellates: a mitochondrial genome all at sea. *Trends Genet.*, **24**, 328-335. (2008)

Neupert,W. and Herrmann,J.M. Translocation of proteins into mitochondria. *Annu. Rev. Biochem.*, **76**, 723-749. (2007)

Norman,J.E. and Gray,M.W. The cytochrome oxidase subunit 1 gene (*cox1*) from the dinoflagellate, *Crypthecodinium cohnii*. *FEBS Lett.*, **413**, 333-338. (1997)

Norman,J.E. and Gray,M.W. A complex organization of the gene encoding cytochrome oxidase subunit 1 in the mitochondrial genome of the dinoflagellate, *Crypthecodinium cohnii*: homologous recombination generates two different *cox1* open reading frames. *J. Mol. Evol.*, **53**, 351-363. (2001)

Ohama,T., Inagaki,Y., Bessho,Y. and Osawa,S. Evolving genetic code. *Proc Jpn Acad Ser B Phys Biol Sci.*, **84**, 58-74. (2008)

Olszewski,K.L., Mather,M.W., Morrisey,J.M., Garcia,B.A., Vaidya,A.B., Rabinowitz,J.D. and Llinás,M. Branched tricarboxylic acid metabolism in *Plasmodium falciparum*. *Nature*, **466**, 774-778. (2010) Erratum in: *Nature*, **469**, 432. (2011)

Osawa,S., Collins,D., Ohama,T., Jukes,T.H. and Watanabe K. Evolution of the mitochondrial genetic code. III. Reassignment of CUN codons from leucine to threonine during evolution of yeast mitochondria. *J. Mol. Evol.*, **30**, 322-328. (1990)

Perkins,F.O. Zoospores of the oyster pathogen, *Dermocystidium marinum*. I. Fine structure of the conoid and other sporozoan-like organelles. *J. Parasitol.*, **62**, 959-974. (1976)

Perkins,F.O. and Menzel,R.W. Ultrastructure of sporulation in the oyster pathogen *Dermocystidium marinum*. *J. Invertebr. Pathol.*, **9**, 205–229. (1967)

Perry,T.L., Pandey,P., Grant,J.M. and Kain,K.C. Severe atovaquone-resistant *Plasmodium falciparum* malaria in a Canadian traveller returned from the Indian subcontinent. *Open Med*., **3**, e10-e16. (2009)

Pino,P., Aeby,E., Foth,B.J., Sheiner,L., Soldati,T., Schneider,A. and Soldati-Favre,D. Mitochondrial translation in absence of local tRNA aminoacylation and methionyl tRNA Met formylation in Apicomplexa. *Mol. Microbiol.,* **76**, 706-718. (2010)

Preiser,P.R., Wilson,R.J., Moore,P.W., McCready,S., Hajibagheri,M.A., Blight,K.J., Strath,M. and Williamson,D.H. Recombination associated with replication of malarial mitochondrial DNA. *EMBO J.*, **15**, 684-693. (1996)

Reece,K.S., Siddall,M.E., Burreson,E.M. and Graves,J.E. Phylogenetic analysis of *Perkinsus* based on actin gene sequences. *J. Parasitol*.. **83**, 417-423. (1997)

Rehkopf,D.H., Gillespie,D.E., Harrell,M.I. and Feagin,J.E. Transcriptional mapping and RNA processing of the *Plasmodium falciparum* mitochondrial mRNAs. *Mol. Biochem. Parasitol.,* **105**, 91-103. (2000)

Richly,E. and Leister,D. NUMTs in sequenced eukaryotic genomes. *Mol. Biol. Evol.*, **21**, 1081–1084. (2004)

Roberts,C.W., Roberts,F., Henriquez,F.L., Akiyoshi,D., Samuel,B.U., Richards,T.A., Milhous,W., Kyle,D., McIntosh,L., Hill,G.C., Chaudhuri,M., Tzipori,S. and McLeod,R. Evidence for mitochondrial-derived alternative oxidase in the apicomplexan parasite *Cryptosporidium parvum*: a potential anti-microbial agent target. *Int. J. Parasitol*., **34**, 297-308. (2004)

Rosengarten,R.D., Sperling,E.A., Moreno,M.A., Leys,S.P. and Dellaporta,S.L. The mitochondrial genome of the hexactinellid sponge *Aphrocallistes vastus*: evidence for programmed translational frameshifting. *BMC Genomics*, **9**, 33. (2008)

Russell,R.D. and Beckenbach,A.T. Recoding of translation in turtle mitochondrial genomes: programmed frameshift mutations and evidence of a modified genetic code. *J. Mol. Evol.*, **67**, 682-695. (2008)

Saccone,C. Mitochondrial genome. In: John Wiley & Sons, Ltd., *Encyclopedia of Life Science 2005*. (2005)

Sagan,L. On the origin of mitosing cells. *J. Theor. Biol.*, **14**, 255-274. (1967)

Saldarriaga,J.F., McEwan,M.L., Fast,N.M., Taylor,F.J. and Keeling,P.J. Multiple protein phylogenies show that *Oxyrrhis marina* and *Perkinsus marinus* are early branches of the dinoflagellate lineage. *Int. J. Syst. Evol. Microbiol.,* **53**, 355-365. (2003)

Saraste,M. Oxidative phosphorylation at *fin de siècle*. *Science*, **283**, 1488–1493. (1999)

Schott,E.J. and Vasta,G.R. The *PmSOD1* gene of the protistan parasite *Perkinsus marinus* complements the *sod2Δ* mutant of *Saccharomyces cerevisiae,* and directs an iron superoxide dismutase to mitochondria. *Mol. Biochem. Parasitol.*, **126**, 81-92. (2003)

Schultz,B.E. and Chan,S.I. Structures and proton-pumping strategies of mitochondrial respiratory enzymes. *Annu. Rev. Biophys. Biomol. Struct.*, **30**, 23-65. (2001)

Sen,N. and Majumder,H.K. Mitochondrion of protozoan parasite emerges as potent therapeutic target: exciting drugs are on the horizon. *Curr. Pharm. Des.*, **14**, 839-846. (2008)

Sengupta,S., Yang,X. and Higgs,P.G. The mechanisms of codon reassignments in mt genetic codes. *J. Mol. Evol.,* **64**, 662-688. (2007)

Siddall,M.E., Reece,K.S., Graves,J.E. and Burreson,E.M. 'Total evidence' refutes the inclusion of *Perkinsus* species in the phylum Apicomplexa. *Parasitology*, **115**, 165-176. (1997)

Slamovits,C.H., Saldarriaga,J.F., Larocque,A. and Keeling,P.J. The highly reduced and fragmented mitochondrial genome of the early-branching dinoflagellate *Oxyrrhis marina* shares characteristics with both apicomplexan and dinoflagellate mt genomes. *J. Mol. Biol.*, **372**, 356-368. (2007)

Sundararajan,A., Michaud,W.A., Qian,Q., Stahl,G. and Farabaugh,P.J. Near-cognate peptidyl-tRNAs promote +1 programmed translational frameshift in yeast. *Mol. Cell*., **4**, 1005-1015. (1999)

Sunila,I., Hamilton,R.M. and Dungan,C.F. Ultrastructural characteristics of the *in vitro* cell cycle of the protozoan pathogen of oysters, *Perkinsus marinus*. *J. Eukaryot. Microbiol.*, **48**, 348–361. (2001)

Suzuki,T., Hashimoto,T., Yabu,Y., Kido,Y., Sakamoto,K., Nihei,C., Hato,M., Suzuki,S., Amano,Y., Nagai,K., Hosokawa,T., Minagawa,N., Ohta,N. and Kita,K. Direct evidence for cyanide-insensitive quinol oxidase (alternative oxidase) in apicomplexan parasite *Cryptosporidium parvum*: phylogenetic and therapeutic implications. *Biochem. Biophys. Res. Commun.*, **313**, 1044-1052. (2004)

Sweetlove,L.J., Beard,K.F., Nunes-Nesi,A., Fernie,A.R. and Ratcliffe,R,G. Not just a circle: flux modes in the plant TCA cycle. *Trends Plant. Sci.*, **15**, 462-470. (2010)

Takashima,E., Takamiya,S., Takeo,S., Mi-ichi,F., Amino,H. and Kita,K. Isolation of mitochondria from *Plasmodium falciparum* showing dihydroorotate dependent respiration. *Parasitol. Int.*, **50**, 273-278. (2001)

Taylor,R.W. and Turnbull,D.M. Mitochondrial DNA mutations in human disease. *Nat. Rev. Genet.*, **6**, 389-402. (2005)

Thompson,J.D., Gibson,T.J., Plewniak,F., Jeanmougin,F. and Higgins,D.G. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.*, **25**, 4876-4882. (1997)

Tomitsuka,E., Goto,Y., Taniwaki,M. and Kita,K. Direct evidence for expression of type II flavoprotein subunit in human complex II (succinate-ubiquinone reductase). *Biochem. Biophys. Res. Commun.*, **311**, 774-779. (2003)

Trounce,I.A., Kim,Y.L., Jun,A.S. and Wallace,D.C. Assessment of mitochondrial oxidative phosphorylation in patient muscle biopsies, lymphoblasts, and transmitochondrial cell lines. *Methods Enzymol*., **264**, 484-509. (1996)

Uyemura,S.A., Luo,S., Vieira,M., Moreno,S.N. and Docampo,R. Oxidative phosphorylation and rotenone-insensitive malate- and NADH-quinone oxidoreductases in *Plasmodium yoelii yoelii* mitochondria *in situ*. *J. Biol. Chem.*, **279**, 385-393. (2004)

Vercesi,A.E., Rodrigues,C.O., Uyemura,S.A., Zhong,L. and Moreno,S.N. Respiration and oxidative phosphorylation in the apicomplexan parasite *Toxoplasma gondii*. *J. Biol. Chem*., **273**, 31040-31047. (1998)

Villalba,A., Reece,K.S., Ordás,M.C., Casas,S.M. and Figueras,A. Perkinsosis in molluscs: a review. *Aquat. Living. Resour*., **17**, 411-432. (2004)

Vimaladithan,A. and Farabaugh,P.J. Special peptidyl-tRNA molecules can promote translational frameshifting without slippage. *Mol. Cell. Biol*., **14**, 8107-8116. (1994)

Waller,R.F. and Jackson,C.J. Dinoflagellate mt genomes: stretching the rules of molecular biology. *Bioessays*, **31**, 237-245. (2009)

Wang,D.Z., Lin,L., Chan,L.L. and Hong,H.S. Comparative studies of four protein preparation methods for proteomic study of the dinoflagellate *Alexandrium* sp. using two-dimensional electrophoresis. *Harmful Algae*, **8**, 685–691. (2009)

Wang,L., Xie,J. and Schultz,P.G. Expanding the genetic code. *Annu. Rev. Biophys. Biomol. Struct*., **35**, 225-249. (2006)

Wang X, Huang Y, Lavrov DV, Gu X. Comparative study of human mitochondrial proteome reveals extensive protein subcellular relocalization after gene duplications. *BMC Evol. Biol.*, **9**, 275. (2009)

Ward,B.L., Anderson,R.S. and Bendich,A.J. The mitochondrial genome is large and variable in a family of plants (cucurbitaceae). *Cell*, **25**, 793-803. (1981)

Williams,B.A., Elliot,C., Burri,L., Kido,Y., Kita,K., Moore,A.L. and Keeling,P.J. A broad distribution of the alternative oxidase in microsporidian parasites. *PLoS Pathog.*, **6**, e1000761. (2010)

World Health Organization. Global summary of the HIV/AIDS epidemic, December 2008. http://www.who.int/hiv/data/2009_global_summary.gif (2009)

World Health Organization. Tuberculosis facts 2009. http://www.who.int/tb/publications/2009/factsheet_tb_2009update_dec09.pdf (2009)

World Health Organization. World Malaria Report 2009. http://www.who.int/malaria/publications/atoz/9789241563901/en/index.html (2009)

Wright,A.C., Ahmed,H., Gauthier,J.D., Silva,A.M. and Vasta,G.R. cDNA cloning and characterization of two iron superoxide dismutases from the oyster parasite *Perkinsus marinus*. *Mol. Biochem. Parasitol.*, **123**, 73-77. (2002)

Zhang,H. and Lin,S. mRNA editing and spliced-leader RNA trans-splicing groups *Oxyrrhis*, *Noctiluca*, *Heterocapsa*, and *Amphidinium* as basal lineages of dinoflagellates. *J. Phycol*., **44**, 703-711. (2008)

ACKNOWLEDGEMENTS

I express my sincere gratitude to my supervisor Professor Kiyoshi Kita, for generously giving me helpful and critical comments and suggestions with inestimable value, throughout my research. I also express my deepest appreciation to my great mentor, Dr. Motomichi Matsuzaki. He provided me with a series of experimental methods and skills for *Perkinsus*, both basic and latest information about *Perkinsus*, and knowledge on general science. He also perseveringly continued giving me critical and suggestive comments on the experiment schemes and results, composing papers, this thesis, and presentations, and thus totally supported my research.

I really thank Dr. Yoh-ichi Watanabe (The University of Tokyo) for helpful comments and discussions about codon recognition, Dr. Tatsushi Mogi (The University of Tokyo) for valuable information in analyzing COX1 sequence. I also thank Dr. Shigeo Yoshinari and Professor Tsutomu Suzuki (The University of Tokyo) for informative suggestion for experimental strategies, and Dr. Ryoma Kamikawa (University of Tsukuba) for helpful comments on mt genomes of dinoflagellates. I am grateful also to Dr. Tomoyoshi Yoshinaga (The University of Tokyo) for providing the isolates of *P. honshuensis* and *P. olseni*. I deeply thank Professor John F. Atkins (The University of Utah) for interesting discussions, kind encouragement, and the wonderful gift, the latest book summarizing recoding events which he edited. My great appreciation goes also to my peers Mr. Emmanuel O. Balogun, Mr. Masayuki Hata, and Ms. Chika Sakai. They entered the Ph.D. course together with me and always helped me during my three years at the Department of Biomedical Chemistry. I thank all other members in the Department of Biomedical Chemistry, Graduate School of Medicine, The University of Tokyo for providing me their constructive discussions and

encouragements. Financially, this work was partly supported by Grants-in-Aid for Creative Scientific Research (18GS0314) and for JSPS Fellows (2105920) from the Japan Society for the Promotion of Science (JSPS).

Lastly, I truly thank my family for supporting all of my life. They generously allowed me to spend another three years as a student at The University of Tokyo, far away from hometown, and kindly have kept in touch with me by telephone, letter and e-mail encouraging me. Without their financial and mental support, I couldn't have accomplished my Ph.D. study. I again thank my family with my whole heart.