Studies on the Expression Regulation Mechanisms of Epimastigote Stage Specific Genes in African Trypanosome

（アフリカトリパノソーマエピマスティゴート型虫体特異的発現遺伝子の発現調節機構に関する研究）

2013

The United Graduate School of Veterinary Sciences, Gifu University
(Obihiro University of Agriculture & Veterinary Medicine)

SUGANUMA, Keisuke

CONTENTS

ABBREVIATIONS

A	AAT:	Animal African trypanosomosis
	actin:	Actin gene
	ActD:	Actinomycin D
	ALBA:	Acetylation lowers binding affinity
	α-TcUBP1:	Anti-TcUBP1
	α-Pf1-Cys-Prx	Anti-*Plasmodium falciparum* 1cys-peroxiredoxine
	ARE:	Adenosine and uridine rich element
	ARE-M:	ARE-middle element
	α-tub:	Alpha tubulin gene
B	BARP:	*Brucei* alanine-rich protein
	BBB:	Blood-brain barrier
	BSA:	Bovine serum albumin
	BSF:	Blood stream form
C	cDNA:	complementary DNA
	CESP:	*Congolense* epimastigote specific protein
	cesp:	CESP gene
	CNS:	Central nervous system
	CSF:	Cerebrospinal fluid

E	EDTA:	Ethylendiaminetetraacetic acid
	eGFP:	Enhanced green fluorescence protein
	egfp:	eGFP gene
	EGTA:	ethylene glycol tetraacetic acid
	ELISA:	Enzyme-linked immunosorbent assay
	EMEM:	Eagle's minimal essential medium
	EMF:	Epimastigote form
	EST:	Expression sequence tag
F	FBS:	Fetal bovine serum
	Front:	Front element
G	GARP:	Glutamic acid/alanine-rich protein
	GRE:	Guanine rich element
	GST:	Glutathione S-transferase
H	HAT:	Human African trypanosomosis
	HEPES:	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
	HI-FBS:	Heart inactivated-FBS
	HIM-9:	Hirumi's modified Iscove's medium-9
	HRP:	Horseradish peroxidase
	hig^R:	Hygromycin B phosphotransferase resistance gene
I	ICR:	Inter coding region

	IFAT:	Indirect fluorescence antibody test
	Ig:	Immunoglobulin
	IMDM:	Iscove's modified Dulbecco's medium
	i. v.:	Intravenous injection
L	LAMP:	Loop-mediated isothermal amplification
M	mAb:	Monoclonal antibody
	MCF:	Metacyclic form
	MCS:	Multiple cloning site
N	NECT:	Nifurtimox/ eflronithine combination treatment
O	ORF:	Open reading frame
P	PBS:	Phosphate-buffered saline
	PBS-T:	PBS containing 0.02 % Tween 20
	PCF:	Procyclic form
	PCR:	Polymerase chain reaction
	PSG:	PBS containing 1% glucose
	PVDF:	Polyvinylidene difluoride
R	RBPs:	RNA binding proteins
	REMSA:	RNA electro mobility shift assay
	rTcUBP1:	Recombinant TcUBP1
	RT-PCR:	Reverse transcriptase-PCR

	RRM:	RNA recognition motif
	rRNA:	Ribosomal RNA
S	SAS:	Saturated ammonium sulfate
	SDS:	Sodium dodecyl sulfate
	SDS-PAGE:	SDS-polyacrylamide gel electrophoresis
	SIT:	Sterile insect technique
	SMUG:	Small mucin family
	smug:	SMUG gene
	SS:	Splicing site
T	TBV:	Transmission blocking vaccine
	Tc*ppdk*:	*Trypanosoma congolense* phosphate dikinase gene
	TcUBP1:	*Trypanosoma congolense* uridine binding protein 1
	Tc*ubp1:*	TcUBP1 gene
	TNE:	Tris-NaCl-EDTA
	TVM-1:	*Trypanosoma vivax* medium-1
U	UBP:	Uridine binding protein
	UTR:	Untranslated region
V	VSG	Variant surface glycoprotein

UNIT ABBREVIATIONS

B	bp:	Base pair
D	ºC:	Degree Celsius
H	h:	Hour
K	kb:	Kilobase
	kbp:	Kilobase pair
	kDa:	Kilodalton
L	l:	Liter
M	µg:	Microgram
	mg:	Milligram
	min:	Minute
	µl:	Microliter
	µm:	Micrometer
	µM:	Micromolar
	µsec:	Microsecond
N	ng:	Nanogram
	nt:	Nucleotide
V	V:	Volt

GENERAL INTRODUCTION

I. African trypanosomosis

Importance of the disease

African trypanosomosis is one of the most important and widespread infectious diseases in sub-Saharan Africa. Human African trypanosomosis (HAT) known as sleeping sickness is lethal if left untreated. Africa experienced three severe sleeping sickness epidemics in 20th century (82), and around 10,000 new cases were reported in 2009 but many cases left undiagnosed and untreated (79). Meanwhile, animal African trypanosomosis (AAT) in domesticated and wild animals causes the estimated economic loss of total agricultural gross domestic product US$ 4.75 billion per year (Food and Agriculture Organization of the United Nations: http://www.fao.org/home/en/). HAT and AAT are classified as tsetse-transmitted African trypanosomosis. Therefore, the distribution of the trypanosomoses in Africa corresponds to the area of tsetse fly (*Glossina* spp.) distribution ranging over 8 million km^2 at latitude 14 degrees North to 20 degrees South including 40 sub-Saharan countries (82).

While, surra is a member of non-tsetse transmitted animal trypanosomosis in domestic and wild animals. The distribution of the disease is not only Africa but also almost of all tropical and subtropical regions including the Asia and South America

(24).

Pathogens and vectors

The causative agents of the disease are spindle-shaped, hemoflagellated protozoan parasites referred as African trypanosome, which belong to the genus *Trypanosoma* that is within phylum Euglenozoa, class Kinetoplastidae and order Trypanosomatida. HAT is caused by *Trypanosoma brucei rhodesiense* in East Africa and *T. b. gambiense* in West and Central Africa (79). On the other hand, AAT is caused by *T. b. brucei*, *T. congolense*, *T. vivax, T. simiae* and *T. godfreyi* (81). All of these trypanosomes are biologically transmitted by tsetse flies. *Glossina* consists of 31 species and three groups, namely *fusca, palpalis* and *morsitans* (91) and distributes from the south of Sahara to north of Kalahari Deserts. Since tsetse distribution in Africa looks like belt-shape, it is so-called "Tsetse belt".

Phylogenetic reconstruction based on the genes coding for the small subunit ribosomal RNA suggests that all Salivarian trypanosome separated from other trypanosome species approximately 300 million years ago (33). Because of the appearance of tsetse flies approximately 35 million years ago, trypanosome has an opportunity to transmit to mammals. This long coexistence of both African trypanosomes and wild animals in Africa may cause the trypanotolerance of wild animals (51). In contrast, domestic animals have yet been unable to develop tolerance

or resistance to trypanosome infections within the 13,000 years of their breeding (82).

Clinical symptoms

HAT is divided into two stages, the hemolymphatic stage (stage 1) and meningoencephalitic stage (stage 2). After tsetse bite, trypanosome migrates to the draining lymph node and reaches blood circulation. Following proliferation in blood circulation, trypanosomes cross the blood-brain barrier (BBB) and invade central nervous system (CNS) (5, 55, 61, 72). *T. b. gambiense* infection follows chronic course over several months or years without specific symptoms in stage 1. After invasion to CNS (stage 2), disease progresses with immunosuppression, severe headache, sleep disorder, weight loss and endocrine abnormalities. Progressive CNS involvement culminates in coma and then death in untreated cases (5). The infection of *T. b. rhodesiense* follows more acute course and manifests higher fever, rapid coma and death within a few weeks (5).

In case of AAT, the most consistent clinical features in livestock are intermittent fever and anemia. There is a general leukopenia, enlarged spleen and liver, and weight loss. Chronically infected animals lose appetite, become lethargic and emaciated, and die usually of congestive heart failure (62). *T. congolense* and *T. vivax* only proliferate in blood circulation, while, *T. b. brucei* and *T. b. rhodesiense* were found also CNS (62, 94).

Surra in camels and horses is usually acute with high fever, anemia, weakness and death. On the other hand, Surra in cattle and buffalos has been considered as mild, chronic, or asymptomatic disease. However, Surra in India was reported high mortality rate (94) and experimental infection of *T. evansi* Indonesian strain also leads death, with occasional nervous signs (85).

Diagnosis

The rapid diagnosis of African trypanosomosis is important for reducing the risk of disease progression, choice of medicine as well as prevention of the disease. Microscopic examination technique for direct detection of trypanosome in blood and/or cerebrospinal fluid (CSF) is a classical but reliable diagnosis method. This method is inexpensive and field-applicable, however, has low-sensitivity and not suitable for mass screening (93). Serological techniques such as indirect fluorescence antibody test (IFAT), enzyme-linked immunosorbent assay (ELISA) and card agglutination test (CATT), has been utilized for diagnosis and epidemiological surveillance of trypanosomosis (14, 56). Although these anti-trypanosome antibody detection methods are suitable for mass screening, they cannot distinguish past from current infections. Recently, polymerase chain reaction (PCR)-based method has been applied to trypanosome detection (68, 93). The advantages of this method are high-sensitivity, specificity, and rapid identification of trypanosome species (25). In addition to PCR,

loop-mediated isothermal amplification of DNA (LAMP) has developed as more sensitive, specific, cost-effective and field-applicable diagnosis method than PCR (63, 64, 84).

Treatment

In case of HAT treatment, pentamidine and suramine are commonly used for hemolymphatic stage. Pentamidine is the first choice for *T. b. gambiense*, while suramine is especially used for *T. b. rhodesiense.* In case of CNS stage, melarsoprol is traditionally used for both *T. b. gambiense* and *T. b. rhodesiense.* Since melarsoprol, a derivative of organic arsenical, causes a reactive encephalopathy in about a fifth of all patients receiving treatment, a substantial proportion of patients (2-12%) die due to adverse side effects (13). The ornithine decarboxylase inhibitor, eflonithine, was registered to treat CNS-stage of *T. b. gambiense* infection since 1990 following first reports of treatment in 1984 (12). Currently, the first line treatment for *T. b. gambiense* switch gradually from melarsoprol to eflonithine (79). In addition to melarsoprol and eflonithine treatment, clinical trial for nifurtimox/eflronithine combination treatment (NECT) was on-going since 2003 (71). The results showed that NECT presented no adverse effects and safer than eflronithine monotherapy (79). Regarding treatment of AAT, three drugs were commonly used. Isometamidium and homidium were used for prophylactic and therapeutic, while, diminazene is only

suitable for treatment (32).

The drug resistance trypanosome has been reported and is increasing both HAT and AAT in different field studies (3, 22, 23, 70).

Disease control

Until now, all attempts to establish effective vaccine against blood stream form (BSF) trypanosome have been unsuccessful, because BSF evades from adaptive host immune responses by changing the antigenicity of a dense variant surface glycoprotein (VSG) (40). Therefore, development of effective vaccines for a realistic field setting has not been achieved (48). Recently, Transmission Blocking Vaccine (TBV) is expected as a new prevention measure for vector-born diseases. TBV aims at interfering and/or blocking pathogen development within vector, halting transmission to non-infected mammalian host (17). Proteins derived from insect stage parasite and tick vector were reported as TBV candidates (17, 60). The TBV for trypanosomosis is not yet developed, however, insect stage trypanosome and/or tsetse proteins might be potential TBV candidates (80).

In addition to developing vaccines, vector control for interrupting the life cycle of trypanosome in tsetse has long been carried out. However, tsetse traps have shown only limited effectiveness in controlling tsetse population (1, 66). As an alternative approach, sterile insect technique (SIT) was developed, and thousands of male tsetse

flies are released to target areas after sub-lethal irradiation. The recent success of SIT was eradication of *G. austeni* from island of Zanzibar, Tanzania, which has demonstrated the feasibility and applicability of SIT in integrated tsetse control programs (90). The Pan African Tsetse and Trypanosomiasis Eradication Campaign was planned to eliminate the tsetse from vast area of sub-Sahara (~10 million km^2) by SIT and other methods (9). However, the huge costs associated with the SIT-based eradication project are also a concern as the most of sub-Saharan countries, which belong to the most heavily indebted poor countries in the world (9).

II. Life cycle of African trypanosome

African trypanosomes are transmitted by tsetse flies and undergo cell differentiation events during their life cycle in tsetse and vertebrate host. The developmental stages of African trypanosome are traditionally divided into four stages, namely BSF, procyclic form (PCF), epimastigote form (EMF) and metacyclic form (MCF). In order to survive in host and vector environments, each developmental stage of trypanosome expresses developmental stage-specific proteins, especially cell surface proteins (88, 89). The VSG is expressed in BSF and MCF, which infect and proliferate in the vertebrate host blood circulation, for immune evasion. In case of *T. brucei*, BSF stages are morphologically and biologically distinguished as long slender form and short stumpy form. Long slender form proliferates in blood circulation and

differentiates into non-proliferative short stumpy form by unidentified induction factor known as stumpy induction factor (29, 87). Only stumpy form can establish parasitism in tsetse midgut and different into PCF (29). Such morphological differences have not yet been reported in *T. congolense* BSF. When BSFs are ingested by tsetse flies, they differentiate into the PCFs and the glutamic acid/alanine-rich protein (GARP) (*T. congolense*) or EP/GPEET procyclins (*T. brucei*) are expressed (6, 75). Then PCFs are differentiated into long-shape PCFs in adjacent area of proventriculus, known as mesocyclic forms (86). It was reported that the mesocyclic trypanosome undergoes asymmetric cell division to long EMF and short EMF (67, 74). In the case of *T. brucei*, only a few PCFs successfully migrate from the midgut to the salivary glands, where they differentiate into the EMFs (73). Since tsetse flies are more susceptible to infections by *T. congolense* than by *T. brucei* (67), many more *T. congolense* PCFs successfully migrate from the midgut to the proboscis, where they differentiate into EMFs. EMF expresses stage-specific surface proteins on their cell surface, namely *brucei* alanine-rich protein (BARP) (*T. brucei*) or *congolense* epimastigote specific protein (CESP) (*T. congolense*), and eventually differentiates into MCF (37, 75, 85). Interestingly, EMFs are the only adherent form throughout the life cycle of the parasite. *T. congolense* EMF parasites adhere tightly via their flagella to the lining of the labrum located in the mouthparts of tsetse fly by hemidesmosome-like structures (27).

The life cycle of *T. vivax* in tsetse is simpler. Briefly, *T. vivax* does not invade tsetse midgut, and BSF differentiates directly into EMF followed by MCF at the proboscis (39).

III. Gene expression regulation mechanism of trypanosome

Gene transcription in Kinetoplastida parasites, including trypanosomes, occurs constitutively and initiates bi-directionally in regions between two divergent gene clusters (59). Different sets of genes are transcribed as large polycistronic RNA precursors that are processed into mature mRNA by coupled mechanisms involving *trans*-splicing and polyadenylation (21, 46, 47). As a consequence of polycistronic transcription by RNA polymerase II, transcription of all protein coding genes occurs at similar rates (47). In post-transcriptional gene expression regulation, each mature mRNA interacts with different sets of RNA binding proteins (RBPs) that interact with the *cis*-element within its 3' untranslated region (UTR) and regulate gene expression (21, 46, 47). The majority of *cis*-elements affecting mRNA half-life and translation have been detected within the 3'UTR. Recently, studies have focused on RBPs interacting with *cis*-elements. For instance, *T. brucei* procyclin is regulated by the *cis*-element in 3'UTR and RBPs (31, 92). Genome-wide studies also reported that the majority of candidate regulation motifs are located within the 3'UTR (58). EMF stage-specific gene expression of the *T. cruzi* small mucin family (SMUG) is regulated

by the adenosine and uridine rich element (ARE) and the guanosine rich element (GRE) of their 3'UTR and by *T. cruzi* uridine binding protein (UBP) 1 and *T. cruzi* UBP2 (18, 19). AREs are well known regulatory elements for mRNA stability in eukaryotic cells (4). AREs are sequence elements of 50 - 150 nt that are rich in adenosine and uridine bases. The AUUUA pentamer motif and certain uridine enrichment are two important characteristics of an ARE in many eukaryotes (4). *T. cruzi* UBP has a single RNA recognition motif (RRM), which is the most abundant RNA binding motif in eukaryotes (15, 19). It has been reported that stage-specific RNA-RBP complexes containing *T. cruzi* UBP regulate *T. cruzi smug* mRNA stabilization or destabilization (18). In addition, the cell differentiation in tsetse digestive tract was controlled by a series of RBPs (45, 83). Therefore, *cis*-element RNA and RBPs plays important roles for adaptation to the external environment and cell differentiation. Since, the lack of *in vitro* culture system for *T. brucei* EMF and MCF stages, the gene expression regulation mechanisms of these two stages have not been well understood yet.

IV. Objectives of this study

The objective of this study is to clarify the molecular mechanisms of stage-specific gene expression regulation of EMF stage of African trypanosome. Specific objectives are: 1) To clarify the role of *cesp* 3'UTR in EMF stage-specific gene expression regulation (Chapter I); 2) To identify and characterize a novel *cis*-element for EMF

stage-specific gene expression regulation located within *cesp* 3'UTR and to reveal the interaction between the *cis*-element RNA and stage-specific RBPs (Chapter II); and 3) To identify and characterize the candidates of the RBPs which interact with the *cis*-element and regulate EMF stage-specific gene expression (Chapter III).

CHAPTER I

The epimastigote stage-specific expression of CESP is tightly regulated by its 3'UTR.

1-1: Introduction

African trypanosomes are transmitted by tsetse flies (*Glossina* spp.) and undergo cell differentiation events during their life cycle. It is known that these parasites express developmental stage-specific proteins in order to survive in different host environments as described in general introduction (6, 37, 75, 86). Gene expression in trypanosomes is regulated post-transcriptionally, wherein genes are initially transcribed as polycistronic RNA precursors. The maturation of the polycistronic RNAs to monocistronic mature mRNAs requires *trans*-splicing and polyadenylation. Previous studies have revealed that developmental stage-specific gene expression is regulated by *cis*-elements located in the 3'UTR and by RBPs (21, 46).

For analysis of gene expression regulation mechanisms of African trypanosome, *T. brucei in vitro* culture and the gene modification systems of *T. brucei* were utilized (49, 95, 96). However, lack of stable *in vitro* culture of EMF and MCF in *T. brucei*, the gene expression regulation mechanisms of these stages were still unknown. On the other hand, *T. congolense* IL3000 strain was established all life cycle *in vitro* cultivation system (38, 77). In addition, the gene modification methods of *T. congolense* were

reported previously (16, 41, 77).

In this chapter, role of *cesp* 3'UTR in EMF-stage specific gene expression was investigated.

1-2: Materials and Methods

Reagents and plastic wares

Unless otherwise stated, all reagents were molecular biology grade and were obtained from Wako Pure Chemicals Industries Ltd. (Osaka, Japan). Cultures flasks (Cat No. 156340 and 156472, Nunc EasYFlask) were obtained from Thermo Fisher Scientific Inc., Waltham, MA, USA. Restriction endonucleases are purchased from Roche Diagnostics (Upper Bavaria, Germany) and New England Biolabs (Ipswich, MA, USA). PCR primers were purchased from Sigma-Aldrich.

Parasites

Trypanosoma congolense IL3000, a Savannah type strain isolated near the Kenya/Tanzania border in 1966 (according to the records of the Biological Service Unit at the International Livestock Research Institute, Nairobi, Kenya), was used in this study. PCF and EMF stages were propagated at 27 ºC in air using *Trypanosoma vivax* Medium (TVM)-1 medium composed of Eagle's minimum essential medium (EMEM, M4655, Sigma-Aldrich Co., St. Louis, Mo, USA) supplemented with 20% heat-inactivated fetal bovine serum (HI-FBS), 60 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES, Sigma-Aldrich) and 10 mM L-proline (Sigma-Aldrich) as previously described (38). The PCF were maintained by diluting 3 ml of log phase parasite suspension with 7 ml of fresh medium

every other day. The adherent EMF culture was maintained by replacing the entire culture supernatant with fresh medium every other day. Living PCF was collected by centrifugation at 3,000 x*g* for 10 min at room temperature. In order to prepare single cell suspensions of the EMF, cells were washed by phosphate-buffered saline (PBS) containing 0.02 % ethylenediaminetetraacetic acid (EDTA) and followed by treatment with PBS containing 0.02 % EDTA and 0.25 % papain. The papain-treated EMF cell culture was kept at room temperature for 10 min. Detached EMF cells were collected by centrifugation at 3,000 x*g* for 10 min at room temperature. MCF stage, which continuously differentiated from EMF and accumulates in the culture supernatants, was purified from EMF culture supernatants by DE52 anion-exchange column chromatography (Whatman Plc., Maidstone, Kent, UK) (52).

Nucleic acid extraction

The trypanosome pellets were resuspended in DNA extraction buffer [150 mM NaCl, 10 mM Tris-HCl (pH 8.0), 10 mM EDTA and 0.1 % sodium dodecyl sulfate (SDS)] containing Protease K (100 μg/ml) and incubated over night at 55 ºC with gentle agitations. The suspension was extracted twice with phenol-chloroform-isoamylalcohol (Sigma-Aldrich), twice with chloroform and genomic DNA was precipitated by adding 0.1 volume 3 M sodium acetate (pH 5.5) and 1 volume of isopropyl alcohol. After centrifugation at 22,500 x*g* for 20 min at 4 ºC, the pellet was rinsed with 70 % ethanol,

air-dried and dissolved in distilled water. Meanwhile, total RNAs of transgenic and wild type trypanosome were extracted using TRIzol reagent (Invitrogen, Carlsbad, CA, USA) following manufacture instructions and dissolved in DNase/RNase-free water.

In silico analysis

The *actin* 5' and 3' splicing sites (SS) were predicted from *T. congolense* Expressed Sequence Tag (EST) database (TriTrypDB: http://tritrypdb.org/tritrypdb/) and TrypsRU (http://tryps.rockefeller.edu). Followed by 3' and 5' SS prediction, the *actin* 5' and 3' UTR were determined. On the other hand, *cesp* 5' and 3' UTR were predicted using previous report of *cesp* sequence data (The National Center for Biotechnology Information DB: http://www.ncbi.nlm.nih.gov) (76). Schematic diagrams of *cesp* and *actin* open reading frame (ORF) and UTR were shown in Fig. 1.

Vector construction

The transgene expression vector optimized for *T. congolense*, namely pSAK (77), was modified for this study. In order to introduce a new multiple cloning site (MCS) into pSAK, pSAK was digested by *Bcl* II and *Bbs* I, and ligated with a double strand artificial synthetic oligonucleotides (Fig. 2). The pSAK with the new MCS was named as pSAK2/MCS and utilized for transgene construction as shown in Fig. 2. The full-length enhanced green fluorescence protein gene (*egfp*) fragment was amplified

from pSAK plasmid by PCR with the *egfp* primer pair (Table 1), and was cloned into the corresponding restriction sites in pSAK2/MCS (pSAK2/*egfp*). For construction of *egfp* expression vector which fused with variable 5' and 3'UTR, each fragment of the *actin* and *cesp* 5' and 3'UTR were amplified form DNA extracted from wild type trypanosome by PCR using appropriate primer pairs (Table 1) and cloned into pCR®2.1-TOPO® vector (Invitrogen). The DNA sequences of all constructs were confirmed using an ABI Prism 3100 Genetic Analyzer (Applied Biosystems Inc., Foster City, CA, USA). These fragments were cloned into the corresponding restriction sites in pSAK2/*egfp* and new *egfp* expression vectors were named as pSAK2/AA, pSAK2/AC, pSAK2/CA and pSAK2/CC (Fig. 2).

Transfection

The transgenes were transfected into the *T. congolense* genome by electroporation using an ECM-830 device (BTX: Harvard Apparatus, Holliston, MA, USA). Fifty μg of purified pSAK vectors was linearized by *Bgl* II and *Nar* I restriction enzyme double digestion and dissolved in 50 μl of distilled water. PCFs (2.4×10^7 cells) at log phase were collected by centrifugation at 3,000 x*g* for 10 min at 4 ºC, washed once with Cytomix buffer [120 mM KCl, 0.15 mM $CaCl_2$, 10 mM K_2HPO_4, 10 mM KH_2PO_4, 25 mM HEPES, pH 7.6, 2 mM ethylene glycol tetraacetic acid (EGTA, Sigma-Aldrich), 5 mM $MgCl_2$, 100 μg/ml bovine serum albumin (BSA, Sigma-Aldrich) and 1 mM

hypoxanthine (Sigma-Aldrich)] and resuspended in 400 µl Cytomix buffer. The PCFs and linearized plasmids were mixed and subjected to an electric pulse of 1.8 kV/100 µsec. Thereafter, the PCFs were immediately transferred into 10 ml of TVM-1 medium.

Drug selection and cultivation of transgenic trypanosome

Hygromycin was added to the transgenic PCF cultures at 5 µg/ml from 2 days after the transfection. The transgenic PCF cultures were maintained every 4 days until the wild type PCFs elimination, and then every 2 days as described in materials and methods. The transgenic EMFs were differentiated from the transgenic PCFs. The transgenic MCFs were cell differentiated from transgenic EMF and purified from transgenic EMF culture supernatant as described above.

Southern blot analysis

For analysis of copy number of *T. congolense* pyruvate phosphate dikinase gene (Tc*ppdk)* (accession number: HE575324, locus tag: TCIL3000_11_6820), 10 µg of genomic DNA which was extracted from wild type PCF were digested by restriction enzymes that did not cut at Tc*ppdk* ORF (*Pst* I), single cut (*Hin* dIII, *Bgl* I and *Bcl* I) and double cut (*Ppu* MI and *Afe* I). Meanwhile, for analysis of transgenes locus in transgenic *T. congolense* genome, 10 µg of genomic DNA that were extracted from

transgenic and wild type PCF were digested with *Msc* I. Digested DNA samples were separated on a 1% agarose gel, and blotted onto Hybond-N^+ nylon membrane (GE Healthcare Co., Little Chalfont, UK). Each probe that was PCR amplified using Tc*ppdk*-F and Tc*ppdk*-R, *egfp*-F and *egfp*-R primer pairs, respectively (Table 1) was labeled using the AlkPhos direct labeling kit (GE Healthcare Co.) and hybridized following the manufacturer's instruction. Chemiluminescence signal was generated using the CDP-*Star* detection reagent (GE Healthcare Co.)

Flow cytometry

The transgenic and wild type PCFs were collected from culture supernatant. MCFs were purified from EMF culture supernatant as described above. Single cell EMFs were prepared by papain treatment as described above. Each trypanosome was washed three times with PBS by centrifugation at 3,000 x*g* for 10 min at room temperature and resuspended with PBS. 5 x 10^6 cells/ml of the parasites were analyzed by flow cytometry (EPICS XL, Beckman Coulter, Inc., Brea, CA, USA).

Confocal laser scanning microscopy

The transgenic and wild type PCFs and EMFs were collected from culture supernatant. MCF was purified from EMF culture supernatant as described above. After collection trypanosome by centrifugation at 3,000 x*g* for 5 min at 4 ºC,

trypanosomes were resuspended and kept in 1 h with PBS with 1 % glucose (PSG) containing 2 μg/ml of Hoechst 33342 (Dojindo Co. Ltd., Kumamoto, Japan) by staining Nucleus and kinetoplast DNA. Following nucleus staining steps, trypanosome was collected by centrifugation at 3,000 x*g* for 5 min at 4 ºC and resuspended in PSG. Living trypanosomes were subjected to confocal scanning microscopic observations without being dried or fixed. Briefly, 1 μl of culture supernatant containing trypanosomes was applied on slide glass. After putting cover glass on, the specimens were observed by confocal laser scanning microscopy (Leica TCS SP5, Leica Microsystems, Wetzlar, Germany).

Reverse transcription

Total RNA samples extracted from each stage of the wild type and four transgenic trypanosomes were subjected to reverse transcription using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems Inc.).

Real-time PCR

For analysis of copy number of transgene, Tc*ppdk* was utilized as an internal control single copy gene. Two-fold serial dilutions from 10,000 to 312.5 copies of plasmids cloned with full-length *egfp* or Tc*ppdk* were used for writing standard curves. The *egfp* and *Tc*ppdk fragments were amplified from 1 ng of genomic DNA that were

extracted from wild type and transgenic PCFs using *egfp* and Tc*ppdk* real-time primer pairs, respectively (Table 1).

For analysis about the relative amount of mRNA expression level, the mRNA of α-*tubulin* was utilized as an internal control (10). *A-tubulin, egfp, actin* and *cesp* fragments were amplified from complementary DNA (cDNA), which was synthesized from 1 μg of total RNA as described above, using α-*tubulin, egfp, actin* and *cesp* real-time primer pairs, respectively (Table 1). The relative quantity of *actin, cesp* and *egfp* mRNA were calculated using the amount of α-*tubulin* mRNA and were statistically evaluated in each developmental stage by Thuky's multiple comparison test followed by one-way factorial ANOVA.

1-3: Results

Confirmation of the copy number of Tc*ppdk*

For confirmation of the copy number of transgene in trypanosome genome, the copy number of Tc*ppdk* was confirmed by Southern blot analysis for internal reference gene in preliminary experiment. Restriction sites were predicted Tc*ppdk* ORF by TriTrypDB. Only single signal was observed in *Pst* I treated samples (Fig. 3-A, lane 1). While, double signals were observed in *Hin* dIII, *Bgl* I and *Bcl* I treated samples (Fig. 3-A, lane 2 - 4). In addition, triple signals which including one expected size signals were clearly observed in *Ppu* MI and *Afe* I treated samples (Fig. 3-A, lane 5 and 6, and Fig. 3-B).

Transgenic trypanosome

The schematic diagrams of *egfp* expression cassettes were shown in Fig. 2. These 4 *egfp* expression cassettes (pSAK2/AA, pSAK2/AC, pSAK2/CA and pSAK2/CC) were transfected into wild type PCF, and transgenic PCFs were cultivated in TVM-1 medium supplemented with 5 μg/ml hygromycin B for drug selection. Wild type PCFs were eliminated within two weeks, while drug-resistant PCFs proliferated. Four different of transgenic PCF cell lines were established namely AA, AC, CA and CC, respectively. Transgenic EMFs were differentiated from transgenic PCF culture.

Transgenic MCFs were purified from transgenic EMF culture supernatant by DE52 anion-exchange column.

Firstly, the integration site of transgene was confirmed by Southern blot. *Msc* I digested genomic DNA that was extracted from wild type and transgenic PCFs were separated on agarose gel and blotted to Hybond-N^+ membrane. The blots were probed by *egfp* coding sequence. As a result, any signal was not observed wild type (Fig. 4-A, WT). While, the strong signals were detected at 2.8 kbp (Fig. 4-A, AA and CA) and 4.8 kbp (Fig. 4-A, AC and CC) as expected sizes (Fig. 4-B).

Secondly, the copy number of transgene in transgenic trypanosome genome was determined by real-time PCR. The *egfp* and Tc*ppdk,* which is an internal control single copy gene confirmed by Southern blot, were amplified from genomic DNA extracted from each trypanosome using each specific real-time primer pairs. As a result, estimated copy numbers of each *egfp* expression cassette were in the range of 0.96 – 1.27 (Table 2).

The comparison of eGFP fluorescence

The eGFP fluorescence intensity in transgenic PCF, EMF and MCF cell lines was analyzed by flow cytometry and confocal laser scanning microscopy. Because the EMF and MCF stages of the parasite were derived from parental PCF cell lines, genome organization of the expression cassettes in each transgenic cell line has to be identical in

each lineage regardless of their life cycle stage. Transgenic trypanosomes which were transfected *egfp* fused with the *actin* 3′ UTR were expressed eGFP throughout the insect stage at a constant level (Fig. 5, AA and CA). On the other hand, the eGFP expression was exclusively observed in the EMF and MCF stages when *egfp* were fused with the *cesp* 3′ UTR (Fig. 5, AC and CC). In AC and CC, the eGFP fluorescence intensity in EMF was relatively higher than that of the MCF (Fig. 5-A, AC and CC).

The comparison of the amount of *egfp* mRNA

The amount of *egfp* mRNA in the transgenic trypanosome cell lines was quantified by real-time PCR analysis. *Actin, cesp, egfp* and *α-tubulin* (Internal control gene) were amplified from cDNA using each specific real-time primer pairs (Table 1). There are no significant differentiations of *actin* mRNA expression among developmental stages (Fig. 6-A). The *cesp* mRNA was statistically higher expressed in EMF than other stages (Fig. 6-B) ($p < 0.05$). In AC and CC, unlike the eGFP protein expression pattern shown in Fig. 5, a significantly large quantity of the *egfp* mRNA fused with the *cesp* 3′UTR was exclusively detected in the EMF stage (Fig. 6-C) ($p < 0.05$). The relative amount of *egfp* mRNA in the EMF was 50-fold greater than that of the PCF (Fig. 6-C).

1-4: Discussion

African trypanosome undergoes cell differentiation in mammalian blood circulation and in tsetse digestive tract. During cell differentiation, stage-specific genes were expressed for adaptation in external environment. In Kinetoplastida including trypanosome, UTR plays an important role for the regulation of stage-specific gene expression (21, 46, 47).

Initially, for decision of *cesp* and *actin* 5' and 3'UTRs, SS database and EST database were utilized. These UTR sequences were cloned into pSAK *egfp* expression vector, which is optimized for *T. congolense* transgene expression. The 4 different *egfp* expression cassettes were transfected into wild type PCF, and transgenic trypanosome cell lines were established, namely AA, AC, AC and CC. For analysis of real-time PCR to confirm of the transgene copy number in transgenic trypanosomes, Tc*ppdk* was chosen as internal control gene. In previous report suggested that *T. brucei ppdk* was confirmed as single copy gene (11). The Tc*ppdk* was indicated also single copy gene by Southern blot analysis (Fig. 3). The real-time PCR analysis of *egfp* copy number in transgenic trypanosomes showed only one copy of *egfp* expression cassette was integrated into transgenic trypanosome. In addition, Southern blot analysis showed that *egfp* was integrated into *α*- and *β-tubulin* tandem repeat locus as expected. These results indicated *egfp* expression cassettes were equally transcribed among each transgenic trypanosome.

In flow cytometry and confocal laser scanning microscopic observation, eGFP protein was strongly expressed in both EMF and MCF which was transfected with *egfp* fused with *cesp* 3'UTR (Fig. 5, AC and CC). While, the differentiation of 5'UTR did not effect of eGFP protein expression (Fig. 5 and Fig. 6). From these results suggested that 5'UTR was not significant in EMF and MCF stage-specific gene expression regulation. While, real-time reverse transcription-PCR (RT-PCR) analysis showed that *egfp* mRNA fused with the *cesp* 3'UTR was highly up regulated in EMF (Fig. 4, AC and CC). These results indicated that eGFP protein detected from MCF was simply the residual of eGFP expressed in the parental EMF cytosol.

1-5: Summary

The 5' and 3'UTR of *actin* and *cesp* were predicted and cloned. The pSAK2 serious, which fused with various UTRs were constructed. These *egfp* expression cassettes were transfected into wild type *T. congolense* genomic DNA and established 4 different transgenic trypanosome cell lines. The correct single integration of transgene into *T. congolense* genome was confirmed by Southern blot and real-time PCR using genomic DNA extracted from transgenic trypanosomes. Transgenic trypanosome, which was transfected with *egfp* fused with *cesp* 3'UTR specifically expressed eGFP in EMF and MCF. On the other hand, *egfp* mRNA was exclusively expressed in EMF as well as *cesp* mRNA.

In conclusion, these results suggested that the *cis*-elements which tightly regulate EMF specific gene expression by selective stabilization of transcripts expressed exclusively in the EMF are located in *cesp* 3' UTR. EMF stage-specific gene expression might be also regulated by RBPs which stabilize/destabilize mRNAs throughout the life stages of the parasite.

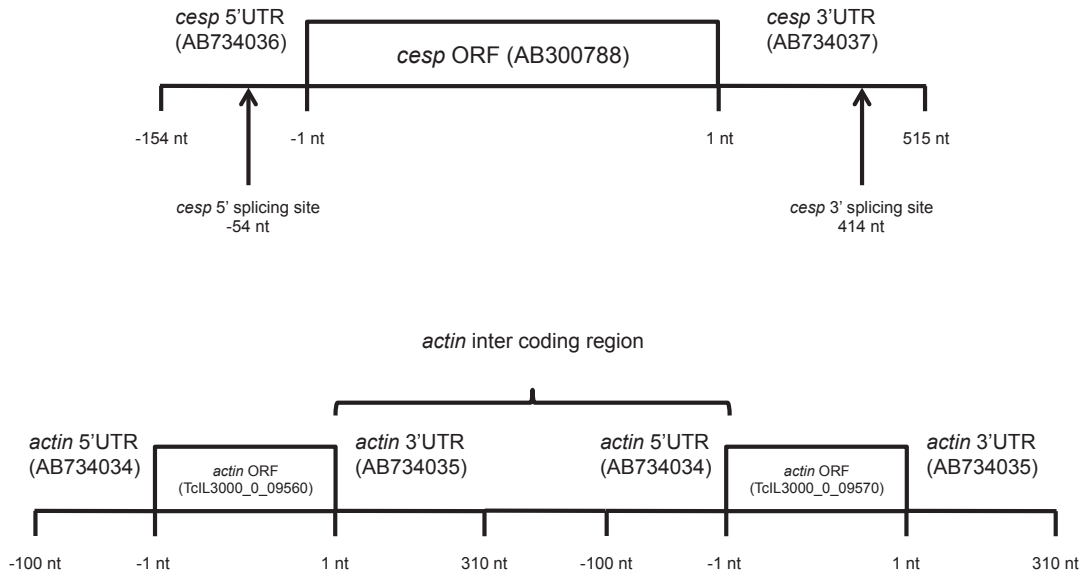

Fig. 1. Schematic diagrams of *cesp* and *actin* ORFs and UTRs. 5' and 3'UTR were predicted from database. Accession numbers of each UTR and gene are shown in parentheses.

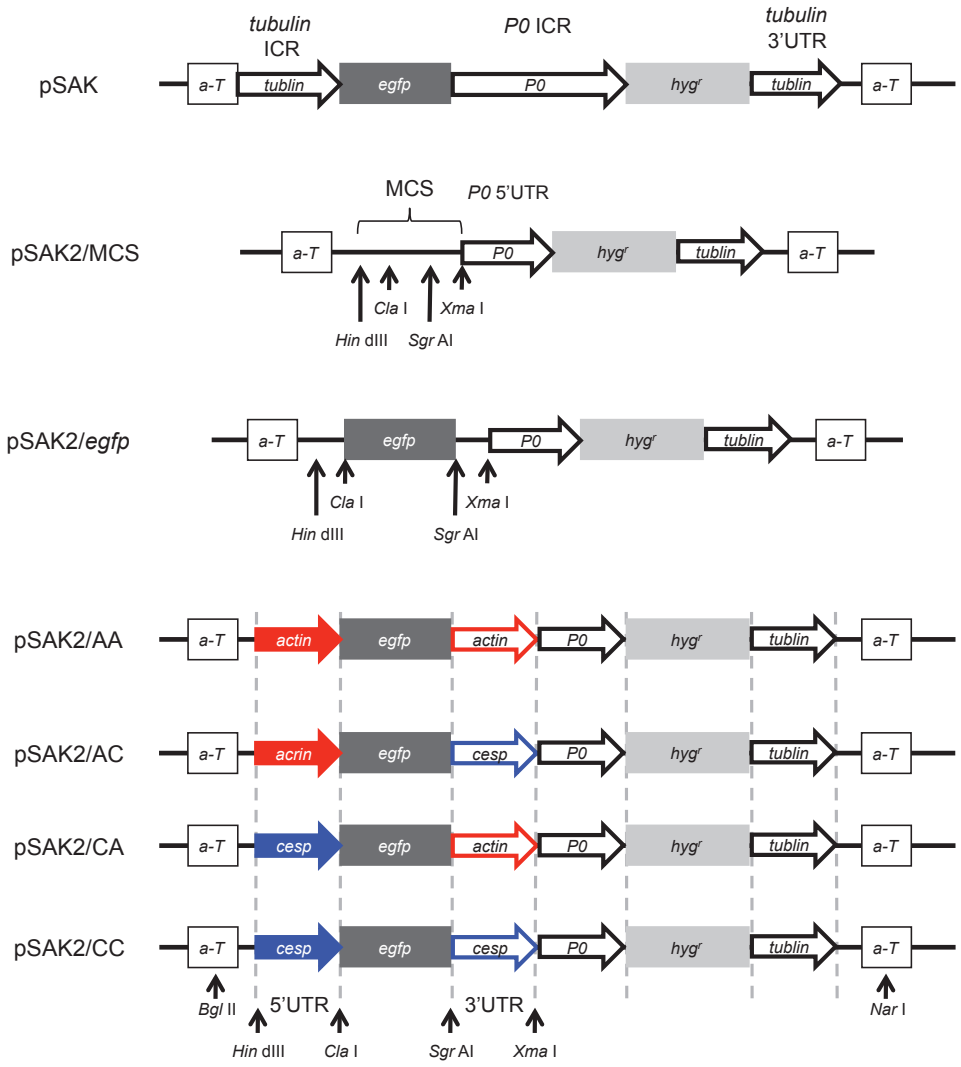

Fig. 2. Schematic diagrams of modification of pSAK vector. A MCS composed of four different restriction sites was constructed with in pSAK (pSAK2/MCS). *egfp* was cloned and inserted into pSAK2/MCS (pSAK2/*egfp*). Each 5' and 3'UTR element were cloned and inserted into pSAK2/*egfp*. The positions of each restriction enzyme site (*Bgl* II, *Hin* dIII, *Cla* I, *Sgr* AI, *Xma* I and *Nar* I), UTRs and/or ICRs of *actin, cesp, P0* and *tubulin* and hygromycin B phosphotransferase gene (*hygr*) are shown.

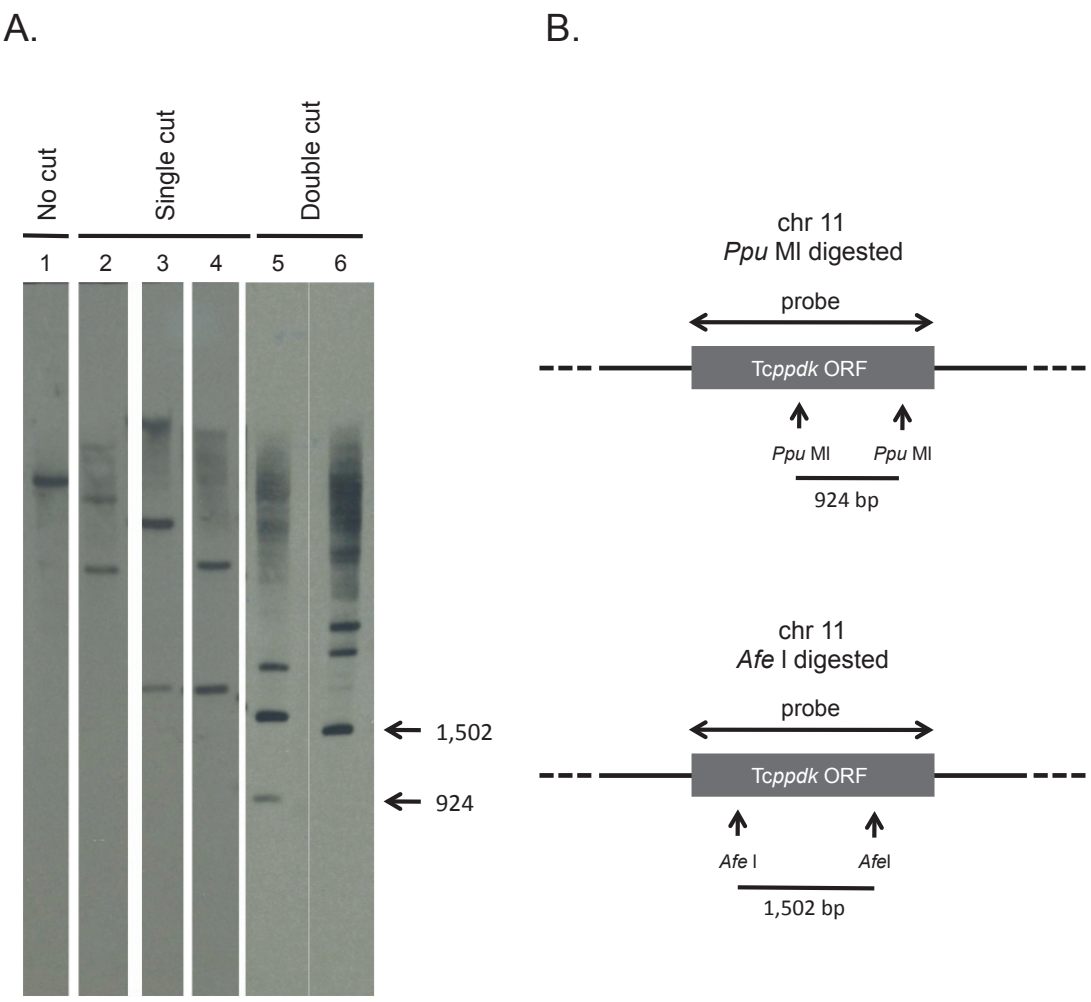

Fig. 3. Southern blot analysis of Tc*ppdk*. (A) Genomic DNA extracted from wild-type PCF was digested by restriction enzymes with *Pst* I (lane 1), *Hin* dIII (lane 2), *Bgl* II (lane 3), *Bcl* I (lane 4), *Ppu* MI (lane 5) and *Afe* I (lane 6). (B) Schematic diagrams of *Ppu* MI and *Afe* I treated Tc*ppdk* ORF are shown.

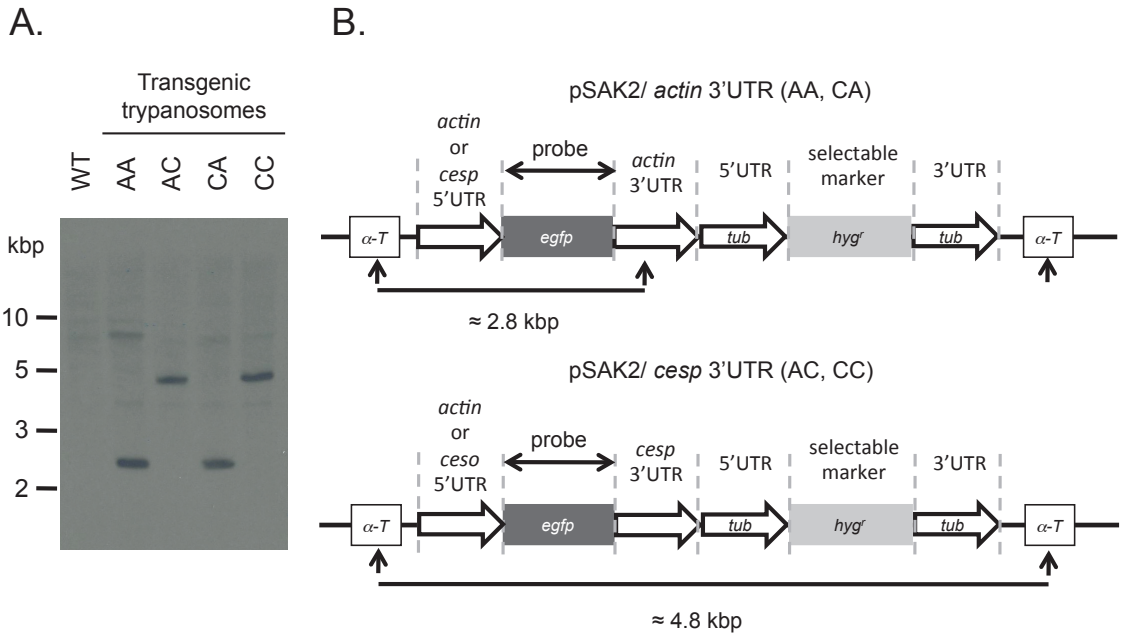

Fig. 4. Southern blot analysis of transgene integration locus. (A) Genomic DNA from wild type and each transgenic PCF cell lines were treated with *Msc* I (indicated by upward arrows). The molecular sizes (2, 3, 5 and 10 kbp) are shown at the left side of the Southern blot image. (B) Schematic diagrams of *Msc* I treated *egfp* expression cassettes are shown.

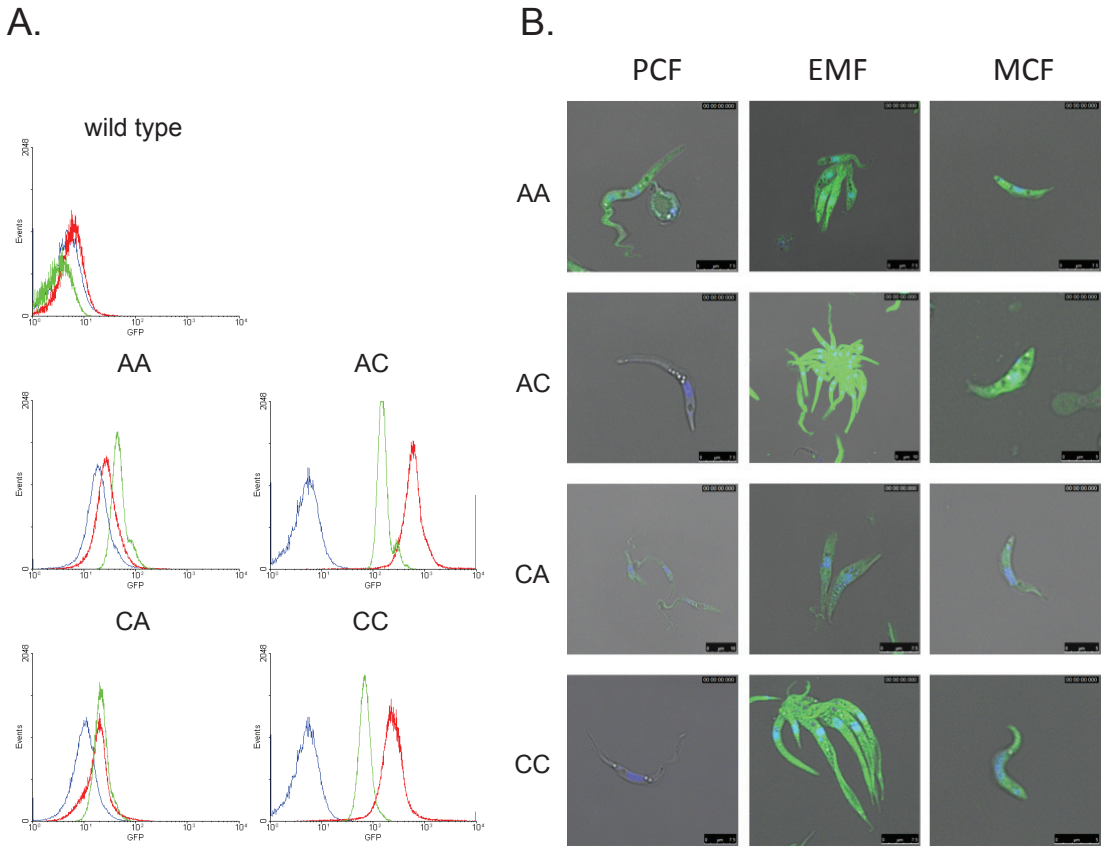

Fig. 5. Comparison of eGFP fluorescence intensity by flow cytometry and confocal laser scanning microscopy. (A) eGFP fluorescence intensity in the live PCF, EMF and MCF stages were analyzed by flow cytometry. Colored lines show eGFP expression intensity in the PCF (Blue), the EMF (Red) and the MCF (Green). (B) Confocal laser scanning microscopic observation of each stage of transgenic cell lines. Observations were made in live trypanosome cells. Nuclear and kinetoplast DNA were stained by Hoechst 33342. Each image in each transgenic cell line was captured by the same gains and voltage of the photomultiplier tubes of the microscope.

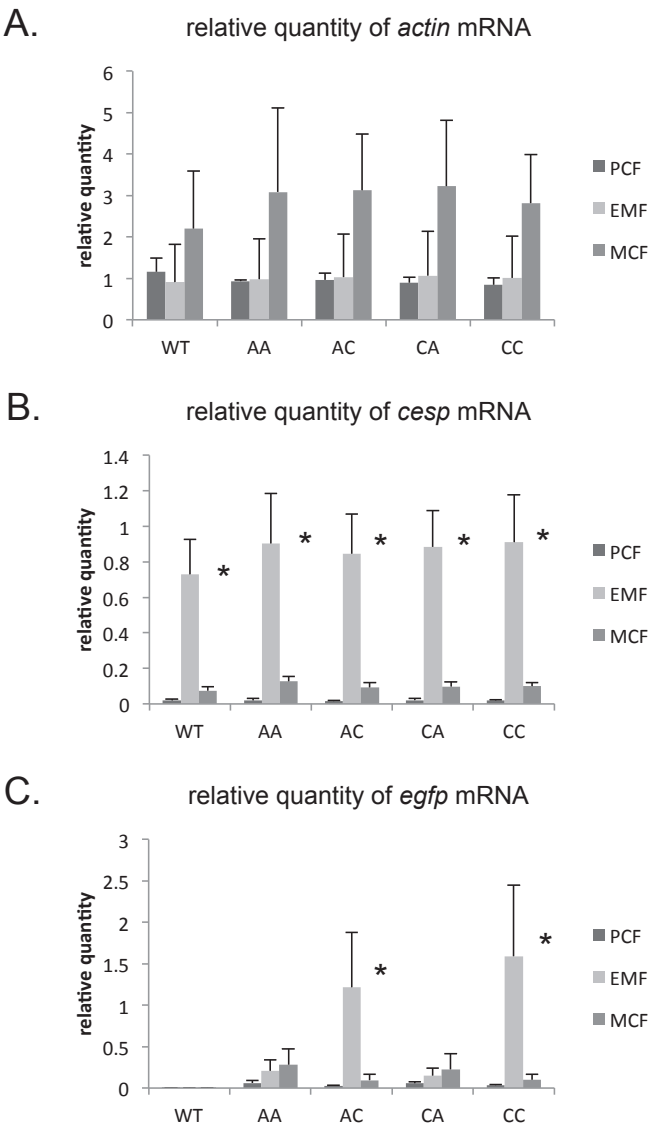

Fig. 6. Comparison of *egfp* mRNA by real-time PCR. Relative quantity of *actin*, *cesp* and *egfp* mRNA against *α-tubulin* mRNA (internal control) were measured by real-time PCR. Asterisks indicate statistical significance between other stages by Tukey's multiple comparison test (*: $p<0.05$). (A) Relative quantity of *actin* mRNA. (B) Relative quantity of *cesp* mRNA. (C) Relative quantity of *egfp* mRNA.

Table 1. DNA oligomers used in this experiment

Name	Sequence	Restriction Enzyme
MCS-sense	5'-ATCCAAAGCTTTTAATTAAATCGATTTCGAACACCG GTGGCTAGCCCCGGGCAAATGT-3'	
MCS-antisense	5'-GATCATCAATTGCCCGGGGCTAGCCACCGGTGTTC GAAATCGATTTAATTAAAAGCTTT-3'	
egfp-F	5'-ATCGATATGGTGAGCAAGGGC-3'	*Cla* I
egfp-R	5'-CACCGGTGTTACTTGTACAGCTGGT-3'	*Sgr* AI
actin 5'UTR-F	5'-AAGCTTTAAGATTCTCCTAACTTTCATCCT-3'	*Hin* dIII
actin 5'UTR-R	5'-ATCGATTGTAGCGAAAGTGGTATTTCT-3'	*Cla* I
actin 3'UTR-F	5'-CACCGGTGTTACAACGTGTTGTATGGTTCTCATT-3'	*Sgr* AI
actin 3'UTR-R	5'-CCCGGGCAAAGTGTAAAACATTTTCCAAATTCA-3'	*Xma* I
cesp 5'UTR-F	5'-AAGCTTTTTGGCGGTGGTGGCTGCGTGT-3'	*Hin* dIII
cesp 5'UTR-R	5'-ATCGATCACCACGACGCGAAAGGCACTTGT-3'	*Cla* I
cesp 3'UTR-F	5'-CACCGGTGGACCGTGAAGGGCAGTTAGTAG-3'	*Sgr* AI
cesp 3'UTR-R	5'-CCCGGGGGTAAGGAAAGGAAAGGCACAC-3'	*Xma* I
Tc*ppdk*-F	5'-ATGGAGACTAAAAAGTGGGTTTAC-3'	
Tc*ppdk*-R	5'-TTAGAGCTTGGCTGCAATC-3'	
egfp real-time F	5'-TCAAGATCCGCCACAACATC-3'	
egfp real-time R	5'-TCTCGTTGGGGTCTTTGCTC-3'	
Tc*ppdk* real-time F	5'-TCTGCGCATCGATCCCTATC-3'	
Tc*ppdk* real-time R	5'-CTTTGCCGACTCAGCATCAA-3'	
α-*tubulin* real-time F	5'-CGCCACGGCAAGTACATG-3'	
α-*tubulin* real-time R	5'-ATCGTGCGCTTCGTCTTGAT-3'	

Underlines indicate additional restriction enzyme sites.

Table 2. The copy number of egfp expression cassette in transgenic trypanosome

Transgenic trypanosome cell strain	The copy number of transgene
Wild type	0.01 ± 0.00
AA	1.27 ± 0.05
AC	1.04 ± 0.07
CA	0.98 ± 0.04
CC	0.96 ± 0.03

The copy numbers of *egfp* in transgenic trypanosomes were calculated using Tc*ppdk* as internal single copy gene. The data were expressed as mean *egfp* copy number ± standard deviation.

CHAPTER II

Adenosine-uridine-rich element is one of the required *cis*-element for epimastigote stage-specific gene expression of the CESP

2-1: Introduction

T. congolense causes animal African trypanosomosis in cattle and wild animals known as Nagana. For adaptation of their external environments in mammal blood circulation and tsetse digestive tract, trypanosome undergoes cell differentiation and expressed for stage-specific genes as described in general introduction. The stage-specific gene expression was regulated by *cis*-element RNA located within its 3'UTR and RBPs. After *trans*-splicing, *cis*-element RNA-RBPs complex determined that mature mRNA stabilize and translate to protein, or destabilized and degradation (21, 46, 47). Previous studies revealed that the common *cis*-element RNA sequence is located in same stage expression genes and regulated gene expression at one time (58). These common *cis*-elements constructed the characteristic RNA secondary structure. These RNA secondary structures have been reported as binding motif of RBPs in *T. cruzi* (65). Moreover, the ARE sequence was reported as one of the major *cis*-elements that control mRNA stability in eukaryotes, including trypanosome (4, 19). Previous studies showed that 44-nt ARE *cis*-element and 27-nt GRE sequences were bound to RBPs and these RNA-RBPs complexes were responsible for mRNA

stabilization and/or destabilization in EMF stage *T. cruzi* (18, 20). The EST and proteome analysis of *T. congolense* reveled that many EMF stage-specific gene were expressed besides *cesp* (28, 36). However, the required *cis*-element and RBPs for African trypanosome EMF stage-specific gene expression regulation are still unknown.

In Chapter I, it was revealed that EMF stage-specific *cesp* expression was regulated by its 3'UTR. This result suggested that the *cis*-element for *cesp* expression regulation was located within the *cesp* 3'UTR. In this chapter, therefore, one of the required *cis*-elements, which regulated EMF stage-specific expression of *cesp*, was identified and characterized. In addition, some stage-specific RBPs, which interacted with the required *cis*-element, were found.

2-2: Materials and Methods

Reagents and plastic wares

Reagents, PCR primers and plastic wears were purchased or prepared as described in Chapter I.

Parasites

The BSF stage of *T. congolense* IL3000 was propagated at 33 ºC in air using Hirumi's modified Iscove's medium (HMI)-9 medium composed of Iscove's modified Dulbecco's medium (IMDM, I1339, Sigma-Aldrich Co., St. Louis, Mo, USA) supplemented with 20% HI-FBS, 60 mM HEPES (Sigma-Aldrich), 1 mM pyruvic acid sodium salt (Sigma-Aldrich), 0.1 mM bathocuproine (Sigma-Aldrich), 1 mM hypoxanthine and 16 μM thymidine (HT supplement, Invitrogen, Carlsbad, CA, USA), 10 μg/L insulin, 5.5 μg/L transferrin and 6.7 ng/L sodium selenite (ITS-X, Invitrogen), 0.0001 % 1-β-mercaptoethanol (Sigma-Aldrich), 0.4 g/L BSA (Sigma-Aldrich) and 2 mM L-cysteine (Sigma-Aldrich) as previously reported (38). The BSF cultures were maintained by replacing the entire culture supernatant with fresh medium every other day. The PCF and EMF of *T. congolense* IL3000 were maintained *in vitro* as described in Chapter I. The MCF was isolated from EMF supernatant as described in Chapter I.

***In silico* analysis**

The *cesp* 3'UTR sequence was analyzed using GENETIX software (GENETYX CORPORATION, Tokyo, Japan). The secondary structure of complete and modified *cesp* and *actin* 3'UTR were predicted by CentroidHomfold software (The National Institute of Advanced Industrial Science and Technology: http://www.ncrna.org/).

Vector construction

The pSAK2/AC (containing the 3'UTR of *cesp* located downstream of the *egfp* reporter gene) and pSAK2/AA (containing the 3'UTR of *actin* located downstream of the *egfp* reporter gene) plasmids were constructed in the Chapter I. The pSAK2/ACΔARE and pSAK2/ACΔFront, which are deletion mutants of the ARE and Front regions from the *cesp* 3'UTR, respectively, were constructed by artificial gene synthesis (Medical and Biological Laboratories Co., LTD, Toyohashi, Japan) (Fig. 8). For construction of deletion mutants from the *cesp* 3'UTR, each fragment of the *cesp* 3'UTR was amplified by PCR using appropriate primer pairs (Table 3 and Fig. 8) and cloned into pCR®2.1-TOPO® vector. The DNA sequences of all constructs were confirmed using an ABI Prism 3100 Genetic Analyzer (Applied Biosystems Inc., Foster City, CA, USA). These fragments for deletion mutants were cloned into the corresponding restriction sites in pSAK2/AC (pSAK2/ACΔLoop 1, pSAK2/ACΔLoop 2 and pSAK2/ACΔLoop 1/2) or in pSAK2/AA (pSAK2/AA+Front, pSAK2/AA+ARE,

pSAK2/AA+AREΔLoop 1, pSAK2/AA+AREΔLoop 2 and pSAK2/AA+AREΔLoop 1/2) (Fig. 8).

Transfection and drug selection of transgenic trypanosome

The transgenes were transfected into the *T. congolense* genome and transgenic trypanosome cell lines were established as described in Chapter I.

Southern blot analysis

The transgene integration locus of transgenic trypanosomes was confirmed by Southern blot as described in Chapter I.

Confocal laser scanning microscopy

The PCF and EMF cells were collected from each culture supernatant, and washed 3 times with PBS. Parasites were fixed on ice for 5 min with PBS containing 4% paraformaldehyde. Following the fixation step, Nucleus and kinetoplast DNA were stained by 2 μg/ml of Hoechst 33342 (Dojindo Co. Ltd., Kumamoto, Japan). The eGFP protein expression in transgenic parasites was observed by confocal laser scanning microscope (Leica TCS SP5, Leica Microsystems, Wetzlar, Germany). The eGFP fluorescence intensity was measured by LAS-AF software (Leica Microsystems).

Northern blot analysis

For analysis of relative amount of *egfp* mRNA, 10 μg of total RNA which were extracted from wild type and transgenic trypanosome. Total RNA samples were separated on a 1.5% denatured agarose gel, and blotted onto Hybond-N^+ nylon membrane (GE Healthcare Co., Little Chalfont, UK). The probes that were amplified using primer pairs of *egfp*-F and *egfp*-R and 18S rRNA-F and 18S rRNA-R (Table 3), were labeled using the AlkPhos direct labeling kit (GE Healthcare Co.) and hybridized according to the manufacturer's instructions. Chemiluminescence signal was generated using the CDP-*Star* detection reagent (GE Healthcare Co.). Densitometry was performed using Image J software (National Institutes of Health). The relative amount of *egfp* mRNA was calculated using the density of 18S rRNA.

RNA stability analysis by Northern blot

For analysis *egfp* mRNA stability, the 18S rRNA was utilized as an internal control (18, 26). Transgenic and wild type EMFs were cultivated in a 6-well cell culture plate until confluent. Confluent EMFs were cultivated TVM-1 supplement with actinomycin D (ActD, Sigma-Aldrich) at a final concentration of 10 μg/ml, which is known to inhibit transcription in trypanosome (7). Attached EMFs were collected at different times after addition of the ActD and total RNA were extracted using TRIzol reagent (Invitrogen) according to the manufacturer's instructions. The Northern blot

analysis and the calculation of the relative amount of *egfp* were performed as described above.

Reverse transcription

For *in vitro* transcription, the target fragments were amplified by PCR using primer pairs containing the T7 RNA promoter sequence (Table 3) and cloned into pCR2.1®TOPO® vector, and the DNA sequences were confirmed. The plasmids were digested by *Eco* Rl and purified by Qiaquick® gel extraction kit (Qiagen, Hilden, Germany). Transcription of biotin-labeled sense sequences was performed using 100 ng of *Eco* Rl-digested fragment and T7 RNA polymerase (New England BioLabs, Ipswich, MA, USA) in the presence of 2.5 mM Biotin-16-UTP (Roche Diagnostics, Upper Bavaria, Germany), 7.5 mM UTP (New England BioLabs), and 10 mM ATP (New England BioLabs), GTP (New England BioLabs) and CTP (New England BioLabs). Transcripts were ethanol-precipitated and resuspended in RNase free water.

Crude protein preparation

To prepare the parasite cell extracts, 5 x 10^8 trypanosomes were resuspended in 1 ml of lysis buffer [0.75% CHAPS detergent (Sigma-Aldrich), 1 mM $MgCl_2$, 1 mM EGTA (Sigma-Aldrich), 5 mM β-mercaptoethanol (Sigma-Aldrich), 10 mM Tris-HCl (pH 7.6) and 10% glycerol] supplemented with a protease inhibitor cocktail (Roche

Diagnostics). After 30 minutes incubation on ice, the extract was centrifuged at 21,500 x*g* at 4 ºC for 10 min and the supernatant was stored at -80 ºC until use (18). The protein concentration of each cell lysate was measured by BCA assay (Thermo Fisher Scientific Inc., Waltham, MA, USA).

RNA electro mobility shift assay (REMSA)

Binding reactions were performed with 4 μg of clued protein extracted from each stage of wild type trypanosome, 0.2 ng of biotin-labeled *in vitro* transcribed RNA probes, 1 % glycerol, 2 μg tRNA and 1x RNA REMSA binding buffer (Thermo Fisher Scientific Inc.) in a 20 μl final volume. For competition assays, the clued proteins were incubated with unlabeled and biotin-labeled RNAs. The incubation time was 30 minutes in room temperature. Each reactant was separated by 4% non-denaturing polyacrylamide gel electrophoresis in 4 ºC according to the manufacturer's instruction for REMSA. The separated reactant was electro-blotted onto the Hybond-N^+ membrane (GE Healthcare Co.) and detected using Chemiluminescent Nucleic Acid Detection Module (Thermo Fisher Scientific Inc.) according to the manufacturer's instructions.

Statistical analysis

The eGFP fluorescence intensities of PCF and EMF were statistically evaluated by Student's *t*-test (Fig. 12). The ratio of *egfp* mRNA amounts after ActD treatment was statistically evaluated by Student's *t*-test (Fig. 14). The eGFP fluorescence intensity and the amount of *egfp* mRNA between different transgenic cell lines was statistically evaluated by Tukey's multiple comparison test followed by one-way factorial ANOVA using GraphPad Prism soft wear (GraphPad Software, Inc., La Jolla, CA, USA) (Fig. 12 and 13).

2-3: Results

Transfection of a series of *egfp* expression cassettes to *T. congolense* PCF

The nucleotide sequence of the *cesp* 3'UTR, predicted secondary structures and schematic diagrams of eGFP expression cassettes are shown in Fig. 7, Fig.9-B and Fig. 8, respectively. Using GENETYX software, an adenosine and uridine rich region was identified in the middle part of the *cesp* 3'UTR, and an AUUUA pentamer (the eukaryotic ARE motif) was also identified within this region. Therefore, this region was named as the *T. congolense* putative ARE (ARE-like). A region from the stop codon of the *cesp* ORF to the first nucleotide of the ARE-like element was designated as the front element (Front) (Fig. 7). The secondary structures of each modified 3'UTR in all transgenic trypanosomes were predicted using CentroidHomfold software (Fig. 9 and Fig. 10). As a result of the secondary structure prediction of complete *cesp* 3'UTR and ARE-like elements, two stem loop structures were found in ARE-like and named Loop 1 and Loop 2 (Fig. 9-B). Moreover, these stem loop structures in Front and ARE-like element (Loop 1 and Loop 2) were maintained in each modified 3'UTR series (Fig. 9 and Fig. 10). To identify the required *cis*-elements in the *cesp* 3'UTR, I modified the pSAK2 *egfp* expression cassette which was used in Chapter I. Ten modified 3'UTR from the complete *cesp* 3'UTR sequence were constructed (Fig. 8), namely the pSAK2/modified *cesp* 3'UTR plasmids (pSAK2/ACΔARE, pSAK2/ACΔFront, pSAK2/ACΔLoop 1, pSAK2/ACΔLoop 2 and pSAK2/ACΔLoop

1/2) and the pSAK2/modified *actin* 3'UTR plasmids (pSAK2/AA+ARE, pSAK2/AA+Front, pSAK2/AA+AREΔLoop 1, pSAK2/AA+AREΔLoop 2 and pSAK2/AA+AREΔLoop 1/2). These 10 plasmids were linearized by *Bgl* ll and *Nar* l restriction enzyme double digestion and transfected into wild type PCF by electroporation. Thereafter, transgenic PCFs were cultivated and differentiated into transgenic EMFs as described in Chapter I. Correct integration of the *egfp* expression cassettes in the *T. congolense* genome was confirmed by Southern blot analysis using 5 μg each of genomic DNA samples extracted from the 10 transgenic PCF cell lines. Clear signals were detected at the expected sizes from each transgenic parasite (4.8 kbp in pSAK2/ modified *cesp* 3'UTR series and 2.8 kbp in pSAK2/ modified *actin* 3'UTR series) (Fig. 11). Therefore, this result confirmed that the *egfp* expression cassette was integrated into the *α*- and *β-tubulin* tandem repeat locus of each transgenic trypanosome genome as expected.

ARE-like contains the required *cis*-element for EMF stage-specific gene expression regulation

The eGFP protein and *egfp* mRNA expression levels of different transgenic trypanosomes were quantified and compared. Transgenic trypanosomes transfected with plasmids containing *egfp* fused with the *cesp* 3'UTR (AC) express eGFP not in PCF but in the EMF stage. While transgenic trypanosomes transfected with *egfp* fused

with *actin* 3'UTR (AA) stably expressed eGFP throughout the life cycle stages. In the PCF stage, transgenic trypanosomes transfected with modified *cesp* 3'UTR series (ACΔARE, ACΔFront, ACΔLoop 1, ACΔLoop 2 and ACΔLoop 1/2) were unable to express eGFP (Fig. 12). Transgenic trypanosomes transfected with modified *actin* 3'UTR series (AA+ARE, AA+AREΔLoop 1, AA+AREΔLoop 2 and AA+AREΔLoop 1/2), were also unable to express eGFP in the PCF stage, while AA and AA+Front were observed as weak eGFP fluorescence in this stage (Fig. 12). However, during the EMF stage of transgenic trypanosomes transfected with modified *cesp* 3'UTR, eGFP expression was only abrogated by the ACΔARE (Fig. 12). All six transgenic trypanosomes transfected with complete and modified *actin* 3'UTR expressed eGFP in the EMF stage. The eGFP expression was almost stable in AA, whereas this expression was significantly elevated in the other transgenic trypanosomes from PCF to EMF (Student's *t*-test: $p<0.01$) (Fig. 12-B). A significant difference in eGFP expression levels was observed between modified *cesp* 3'UTR series of PCFs and EMFs, except the ACΔARE (Student's *t*-test: $p<0.01$) (Fig. 12-B). The transgenic EMF transfected with pSAK2/ACΔFront showed reduced eGFP fluorescence levels, approximately 4.5-fold lower than the control transgenic trypanosomes (pSAK2/AC) (Tukey's multiple comparison test: $p<0.001$) (Fig. 12-B). In addition, the eGFP fluorescence level of pSAK2/ACΔLoop 1 was 2-fold lower than control (Tukey's multiple comparison test: $p<0.001$). No significant differences in eGFP fluorescence

levels were observed for pSAK2/ACΔLoop 2 or pSAK2/ACΔLoop 1/2 compared to the control (Fig. 12-B).

To compare *egfp* mRNA expression levels among transgenic EMF samples, Northern blot analysis was performed with *egfp* and 18S rRNA probes (Fig. 13). The relative amount of *egfp* mRNAs was calculated using the amount of 18S rRNA. There was a tendency towards a positive correlation between the eGFP fluorescent intensity and the amount of *egfp* mRNA. Significant differences were observed between control transgenic trypanosomes AC and ACΔARE, ACΔFront and ACΔLoop 1 (Tukey's multiple comparison test: $p<0.05$), whereas no significant differences were observed between AC and ACΔLoop 2 and ACΔLoop 1/2 (Fig. 13-B). Additionally, in transgenic trypanosomes, which were transfected with modified *actin* 3'UTR series, significant differences were observed between control transgenic trypanosomes AC and the others (Tukey's multiple comparison test: $p<0.05$) (Fig. 13-B).

Front and ARE-like elements regulate mRNA stability in the EMF stage

Analysis of mRNA stability during EMF from the complete (AC) and modified *cesp* 3'UTR constructs (ACΔARE, ACΔFront, ACΔLoop 1, ACΔLoop 2 and ACΔLoop 1/2) were performed. The amount of *egfp* mRNA extracted from ActD-treated EMF cultures was determined by Northern blot analysis (Fig. 14). The relative amounts of *egfp* mRNA in each transgenic trypanosome were plotted every 1 hour for 5 hours (Fig.

14-B). The amount of each time point of *egfp* mRNA was normalized by that of 18S rRNA. The *egfp* mRNA levels extracted from AC, ACΔLoop 2 and ACΔLoop 1/2 were kept at a steady state level during the experimental period. No significant differences were observed between the ACΔLoop 2 and ACΔLoop 1/2 and control transgenic EMF (AC) at each time point (Fig. 14-B). However, the *egfp* mRNA levels extracted from ACΔARE, ACΔFront and ACΔLoop 1 were decreased over time to differing extents. Stability of the *egfp* mRNA was significantly lowest in the ACΔARE, followed by ACΔFront and ACΔLoop 1 (Fig 14-B). There was a tendency towards a positive correlation in the *egfp* mRNA stability, the amount of *egfp* mRNA and the eGFP fluorescence intensity.

RBPs had sequence-specific interactions with the required *cis*-element

The ARE-like sequence located in the *cesp* 3'UTR was characterized as the responsible sequence for EMF stage-specific gene expression. To investigate RNA-RBPs interaction in ARE-like sequence, the ARE-like element was divided into three parts based on their predicted RNA secondary structures (Fig. 9-B). Each part of the ARE-like element was *in vitro* transcribed, namely Loop 1-RNA, Loop 2-RNA and ARE-middle-RNA (ARE-M-RNA). These RNAs were used to perform REMSA. When these three RNAs were reacted with the EMF cell lysate, the different gel-shift signal patterns were observed by REMSA (Fig. 15-A, lanes 2, 5 and 8). Then,

ARE-M-RNA was reacted with the trypanosome cell lysates from BSF, PCF and EMF. The gel-shift patterns were different among each cell lysate (Fig. 15-B, lanes 2 – 4). These gel-shift signals disappeared by adding unlabeled ARE-M-RNA as a competitor RNA (Fig. 15-A, lanes 3, 6 and 9, and Fig. 15-B, lanes 5 – 7). In order to clarify the sequence specificity of the RNA-RBPs interaction, I performed a cross-competition assay. A 100-fold molar excess of unlabeled ARE-M-RNA mostly prevented the formation of the ARE-M-RNA-RBPs complex (Fig 16, lane 3). A 500-fold or higher molar excess of unlabeled ARE-M-RNA completely prevented the ARE-M-RNA-RBPs complex formation (Fig 16, lanes 4 and 5). However, a 100-fold molar excess of unlabeled Loop 1-RNA and Loop 2-RNA did not prevent the formation of the ARE-M-RNA-RBPs complex (Fig 16, lanes 6 and 9). Adding 500-fold or higher molar excess of unlabeled Loop 1-RNA and Loop 2-RNA, the gel shift pattern, which shows the complex of ARE-M-RNA and RBPs, still remained (Fig 16, lanes 7, 8, 10 and 11).

2-4: Discussion

Differentiation of trypanosomes from PCFs to EMFs requires drastic changes in cell biology including expression of many EMF stage-specific proteins in order to parasitize in the proboscis (*T. congolense*) or in the salivary gland (*T. brucei*) of the tsetse fly. *T. congolense* EMF parasites adhere tightly via their flagella to the lining of the labrum located in the mouthparts by hemidesmosome-like structures (27). Previous studies suggested that CESP was exclusively translated in the EMF stage and that CESP might be one of these adhesion molecules (76). In addition, I demonstrated the involvement of *cesp* 3'UTR in the stage-specific expression of the *egfp* reporter gene in Chapter I. Similar mechanisms controlling mRNA stability and translation by 3'UTR sequences have also been described for other trypanosome genes (21, 46, 47).

The adenosine and uridine rich region was found in the middle part of the *cesp* 3'UTR. Moreover, an AUUUA pentamer, which is an essential motif of the eukaryotic ARE, was found in this region. Therefore, this adenosine and uridine rich element was designated, as *T. congolense* ARE-like sequence. By RNA secondary structure analysis, two stem loop elements, namely Loop 1 and Loop 2, were found in the ARE-like.

Transgenic trypanosomes in which *egfp* was fused with the 3'UTR containing the ARE-like element, specifically expressed eGFP fluorescence during the EMF stage (Fig. 12). However, *cesp* 3'UTR lacking the ARE-like element did not confer eGFP

expression in both PCF and EMF stages (Fig. 12, ACΔARE). EMF stage-specific eGFP fluorescence was observed even after deletion of Front from the *cesp* 3'UTR, although the eGFP fluorescence intensity of this transgenic EMF was significantly lower than that of the complete *cesp* 3'UTR (Fig. 12, AC and ACΔFront). The eGFP fluorescence of AA and AA+Front was expressed not only in PCF but also in EMF. The eGFP fluorescence level in the PCF stage was the same for AA and AA+Front, however, it was only enhanced in AA+Front during the EMF stage (Fig. 12). These results suggested that the *T. congolense* ARE-like sequence was one of the responsible elements for EMF stage-specific gene expression of *cesp*. The *T. congolense* ARE-like sequence, similar to that of the higher eukaryotic cell ARE, acts as a *cis*-element for gene expression regulation, therefore, I designated this element as ARE. However, the Front element was not required for EMF stage-specific gene expression, but did enhance gene expression in EMF. Previous studies showed that the *cis*-element contributed to characteristic RNA secondary structures and that RBPs bound to the *cis*-element through such RNA structures (65). Two stem loop structures, Loop 1 and Loop 2, were predicted in ARE (Fig. 9-B). However, the deletion mutants of two Loop structures (ΔLoop 1, ΔLoop 2 and ΔLoop 1/2) still showed EMF stage-specific eGFP expression (Fig. 12, ACΔLoop 1, ACΔLoop 2, ACΔLoop 1/2, AA+AREΔLoop 1, AA+AREΔLoop 2 and AA+AREΔLoop 1/2). Moreover, sequence deletion in *cesp* 3'UTR and insertion in *actin* 3'UTR did not change the characteristic

secondary structures (Fig. 8 and 9), suggesting that these predicted secondary structures (Loop 1 and Loop 2 in ARE) might not be responsible for EMF stage-specific gene expression. Therefore, it was concluded that this required *cis*-elements is located in between Loop 1 and Loop 2 of ARE, and this required *cis*-element was designated as ARE-M.

The eGFP fluorescence intensity and the amount of *egfp* mRNA varied in each transgenic EMF (Fig. 12 and Fig. 13). Moreover, the stability of *egfp* mRNA also varied in each transgenic EMF (Fig. 14). The stability of *egfp* mRNA showed a tendency towards positive correlation with the levels of both eGFP fluorescence intensity and *egfp* mRNA. In general, transcription of all trypanosome genes by RNA polymerase II occurs at similar rates (47). In this study, since the *egfp* cassettes of each transgenic trypanosome cell line were integrated into the *α*- and *β-tubulin* tandem repeat locus transcribed by RNA polymerase II from the internal promoter site (Fig. 11), transcription levels of each *egfp* expression cassette were virtually equal in all transgenic trypanosome cell lines. Therefore, it was concluded that the varied eGFP expression levels (both protein and mRNA) in each transgenic EMF were caused by varied *egfp* mRNA stability.

Cellular factors regulating mRNA stability have been reported in many eukaryotes, including trypanosomes (47, 69). In this chapter, it was clarified that ARE-M in the *cesp* 3'UTR is one of the required *cis*-elements for gene expression in an EMF

stage-specific manner. This result implied that RBPs interact with the ARE-M-RNA. The ARE-M sequence containing the characteristic AUUUA pentamer motif was cloned and *in vitro* transcribed with 16-biotin UTP. Biotin-labeled ARE-M-RNA was reacted with crude protein extracted from each trypanosome stage. RNA gel-shift patterns varied among BSF, PCF and EMF (Fig. 15-B, lanes 2 - 4). Moreover, the RNA gel-shift reacted with EMF protein was inhibited by addition of lower molar excess of unlabeled ARE-M-RNA (Fig. 16, lanes 3 – 5). By comparison, other competitor RNAs could not completely inhibit this RNA gel-shift even after adding higher molar excess of unlabeled RNA (Fig. 16, lanes 6 - 11). Taken together, these results suggested that the interaction of ARE-M-RNA and RBPs that interact with ARE-M-RNA was stronger than that of other RNA sequence and these RBPs. Tough, it was concluded that this ARE-M-RNA and RBPs interaction is sequence-specific. In addition, the ARE-M-RNA and RBP complexes may differ between each trypanosome stage. An AU rich sequence, that is similar to the ARE motif and known to be a regulatory element of eukaryotic RNA stability, was found in the 3'UTR of Tc*smug*, which is specifically expressed during the EMF stage (26). Stage-specific RBPs express and interact with ARE-RNA (18). In this chapter, it has shown shown that one of the required *cis*-elements in ARE and RBPs interact with the *cis*-element. This result implied that the common sequence element regulated gene expression in EMF stage-specific manner.

2-5: Summary

Characteristic secondary structures and AU rich element in *cesp* RNA were predicted using *in silico* analysis. Ten modified *egfp* expression cassettes were constructed. These modified pSAK2 series were transfected into wild type PCF and 10 different transgenic trypanosomes were established. The correct integration of pSAK2 into *T. congolense* genome was confirmed by Southern blot analysis. The eGFP florescence intensity and *egfp* mRNA of these transgenic trypanosomes and transgenic trypanosomes AC and AA which were established in Chapter I were analyzed by confocal laser scanning microscopy and Northern blot, respectively. Transgenic trypanosome, which was transfected with *egfp* containing with ARE-M element in *cesp* 3'UTR were exclusively expressed e*gfp* mRNA in EMF. In addition, *egfp* mRNA which lacking of ARE element from 3'UTR was more quickly degraded than *egfp* mRNA which containing ARE element. Furthermore, *cis*-element RNA and RBPs interaction was analyzed by REMSA. Unknown RBPs that has stage specificity were interacted with ARE-M-RNA.

In conclusion, these results suggested that EMF stage-specific expression of *cesp* was regulated by ARE-M element in *cesp* 3'UTR and RBPs complex. It was hypothesized that other EMF stage-specific genes expression were also regulated by unknown RBPs and ARE-M like element complex.

Fig. 7. *In silico* analysis of *cesp* 3'UTR sequence. The *cesp* 3'UTR sequence used in this study was derived from the *cesp* gene. Major sequence features are indicated in black (Front element: Front), orange (Loop 1 and Loop 2), and surrounded by a blue square (AU rich element: ARE). The ARE-M sequence is indicated in green characters. An AUUUA pentamer located in ARE-M is indicated by capital letters.

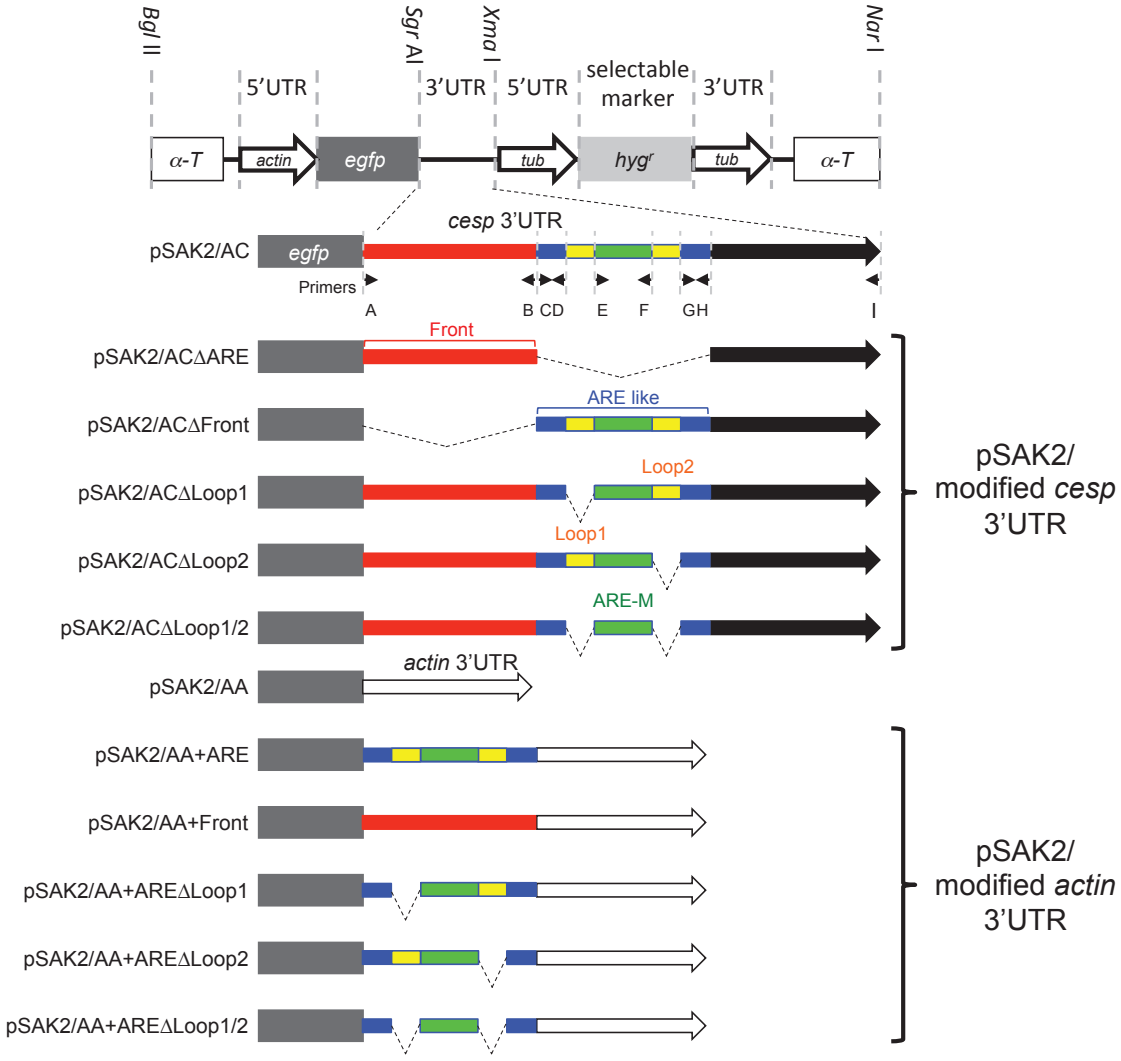

Fig. 8. Schematic diagrams of primer positions and modified pSAK2 plasmids. The letters (A – I) and arrows under *cesp* 3'UTR indicate the locations of each primer shown in Table 3. Major sequence features in *cesp* 3'UTR are indicated by the red box (Front) and in yellow (Loop 1 and Loop 2), and surrounded by a blue square (ARE like). The ARE-M sequence is indicated by the green box. The positions of each restriction enzyme site (*Bgl* II, *Sgr* AI, *Xma* I and *Nar* I), 5'UTRs of *actin* and *tubulin* (*actin* and *tub*) and hygromycin B phosphotransferase gene (*hygr*) are shown. These expression cassettes were targeted into the *α*- and *β-tubulin* tandem repeat locus of the *T. congolense* genome and transcribed by internal promoter.

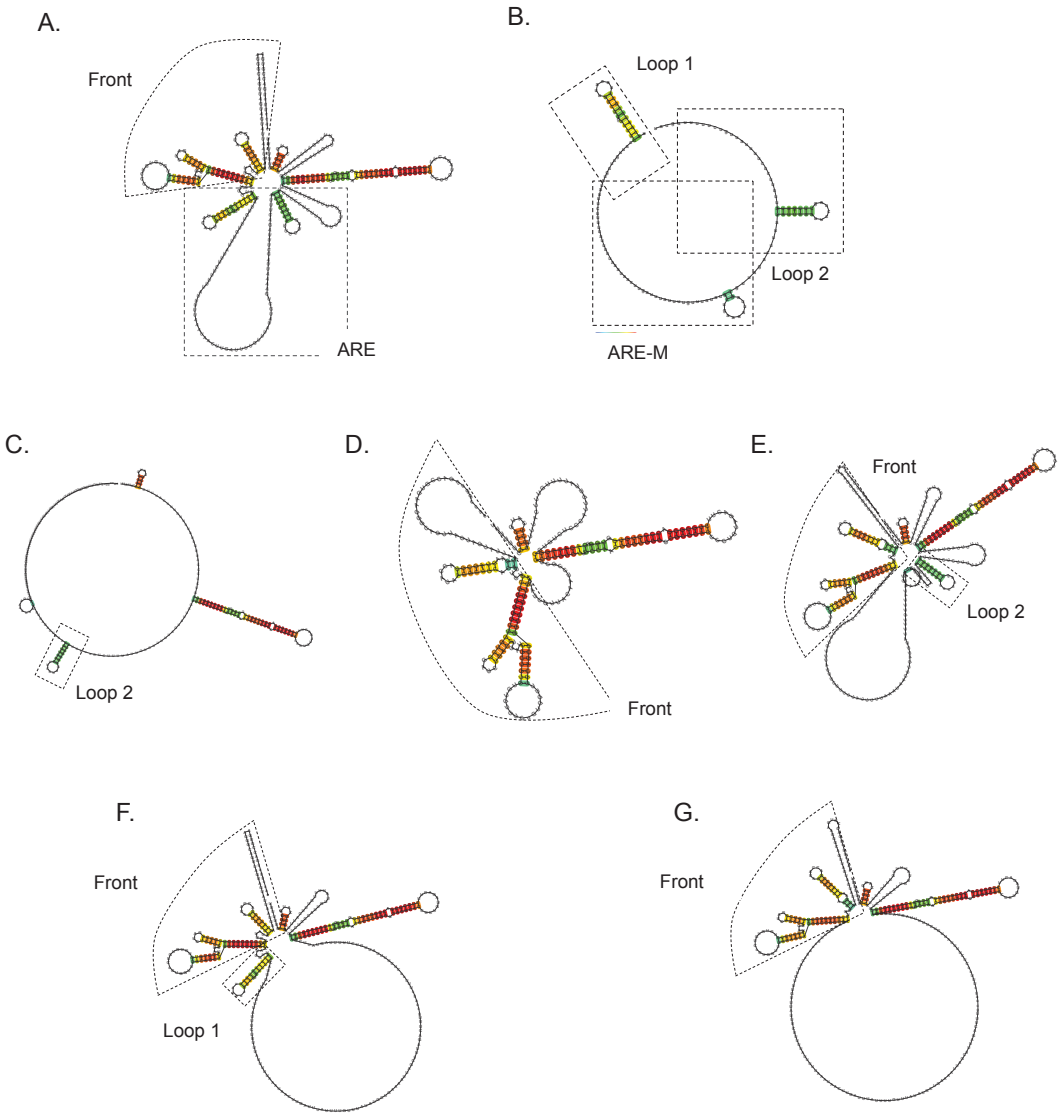

Fig. 9. The predicted secondary structures of complete and modified *cesp* 3'UTR. RNA secondary structures were predicted by CentroidHomfold software. A: complete *cesp* 3'UTR, B: ARE sequence in *cesp* 3'UTR, C: *cesp* 3'UTR lacking Front (ΔFront), D: *cesp* 3'UTR lacked with ARE (ΔARE), E: *cesp* 3'UTR lacked with Loop 1 (ΔLoop 1), F: *cesp* 3'UTR lacked with Loop 2 (ΔLoop 2), G: *cesp* 3'UTR lacked with Loop 1 and 2 (ΔLoop 1/2). Major sequence features are surrounded by dashed line (Front, Loop 1, Loop 2, ARE and ARE-M).

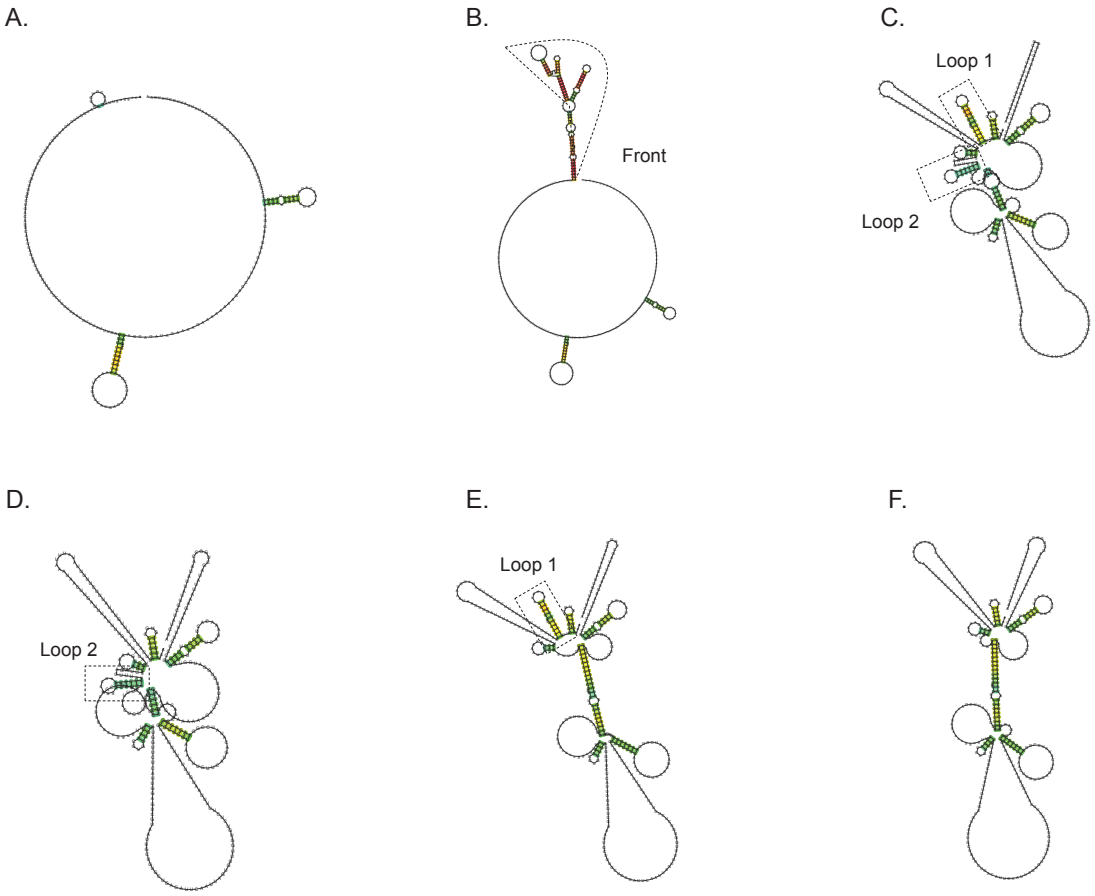

Fig. 10. The predicted secondary structures of complete and modified *actin* 3'UTR. RNA secondary structures were predicted by CentroidHomfold software. A: complete *actin* 3'UTR, B: *actin* 3'UTR fused with Front (+Front), C: *actin* 3'UTR fused with ARE (+ARE), D: *actin* 3'UTR fused with ARE lacking Loop 1 (+AREΔLoop 1), E: *actin* 3'UTR fused with ARE lacking Loop 2 (+AREΔLoop 2) and F: *actin* 3'UTR fused with ARE lacking Loop 1 and 2 (+AREΔLoop 1/2). Major sequence features are surrounded by dashed line (Front, Loop 1, Loop 2, ARE and ARE-M).

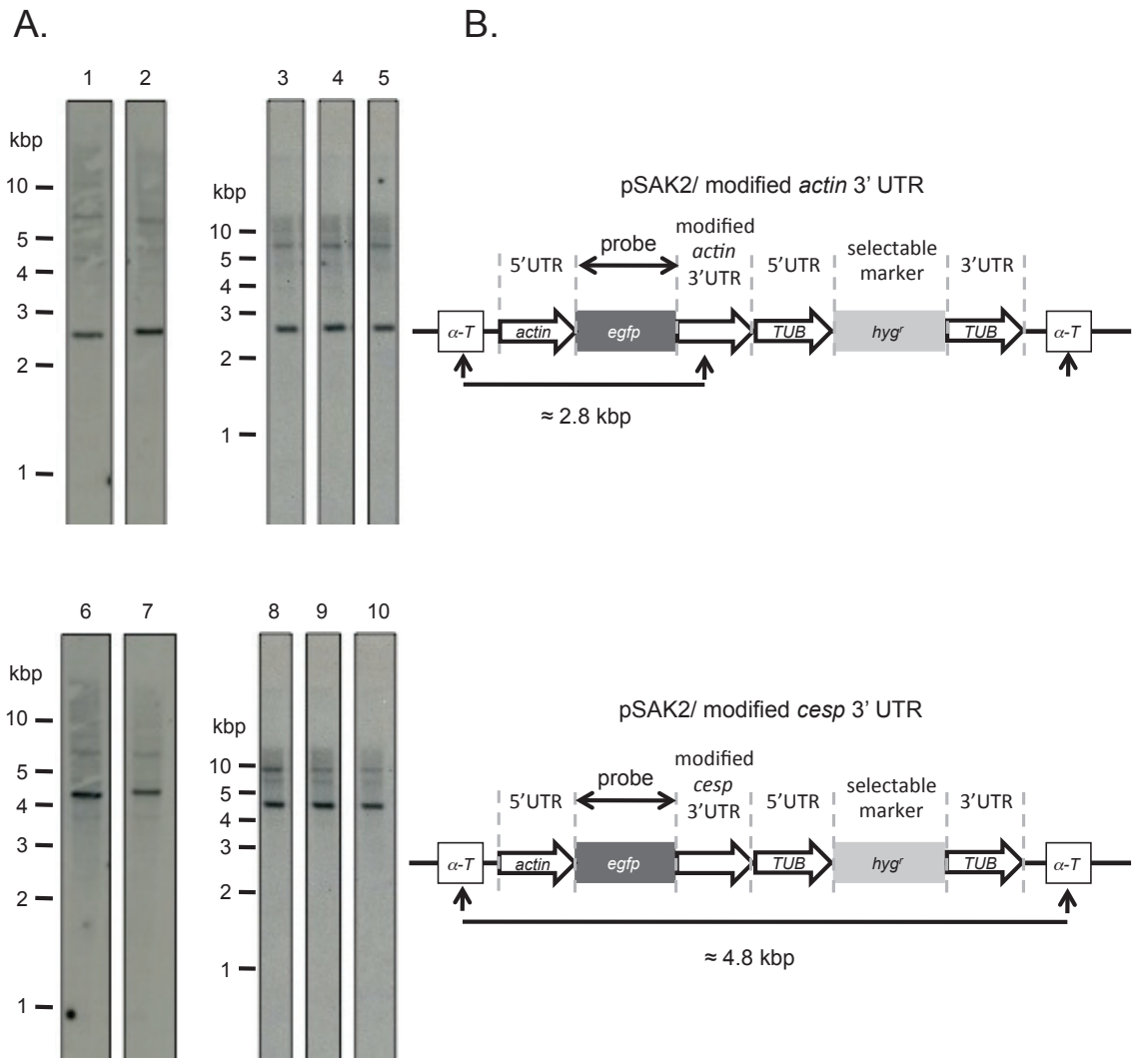

Fig. 11. Southern blot analysis of transgene integration locus. (A) Genomic DNA from wild type and each transgenic PCF cell lines were treated with *Msc* I (indicated by upward arrows). The molecular sizes (1, 2, 3, 4, 5 and 10 kbp) are shown at the left side of the Southern blot image. 1: pSAK2/AA+Front, 2: pSAK2/AA+ARE, 3: pSAK2/AA+AREΔLoop 1, 4: pSAK2/AA+AREΔLoop 2, 5: pSAK2/AA+AREΔLoop 1/2, 6: pSAK2/ACΔFront, 7: pSAK2/ACΔARE, 8: pSAK2/ACΔLoop 1, 9: pSAK2/ ACΔLoop 2 and 10: pSAK2/ACΔLoop 1/2. (B) Schematic diagrams of *Msc* I treated *egfp* expression cassettes are shown.

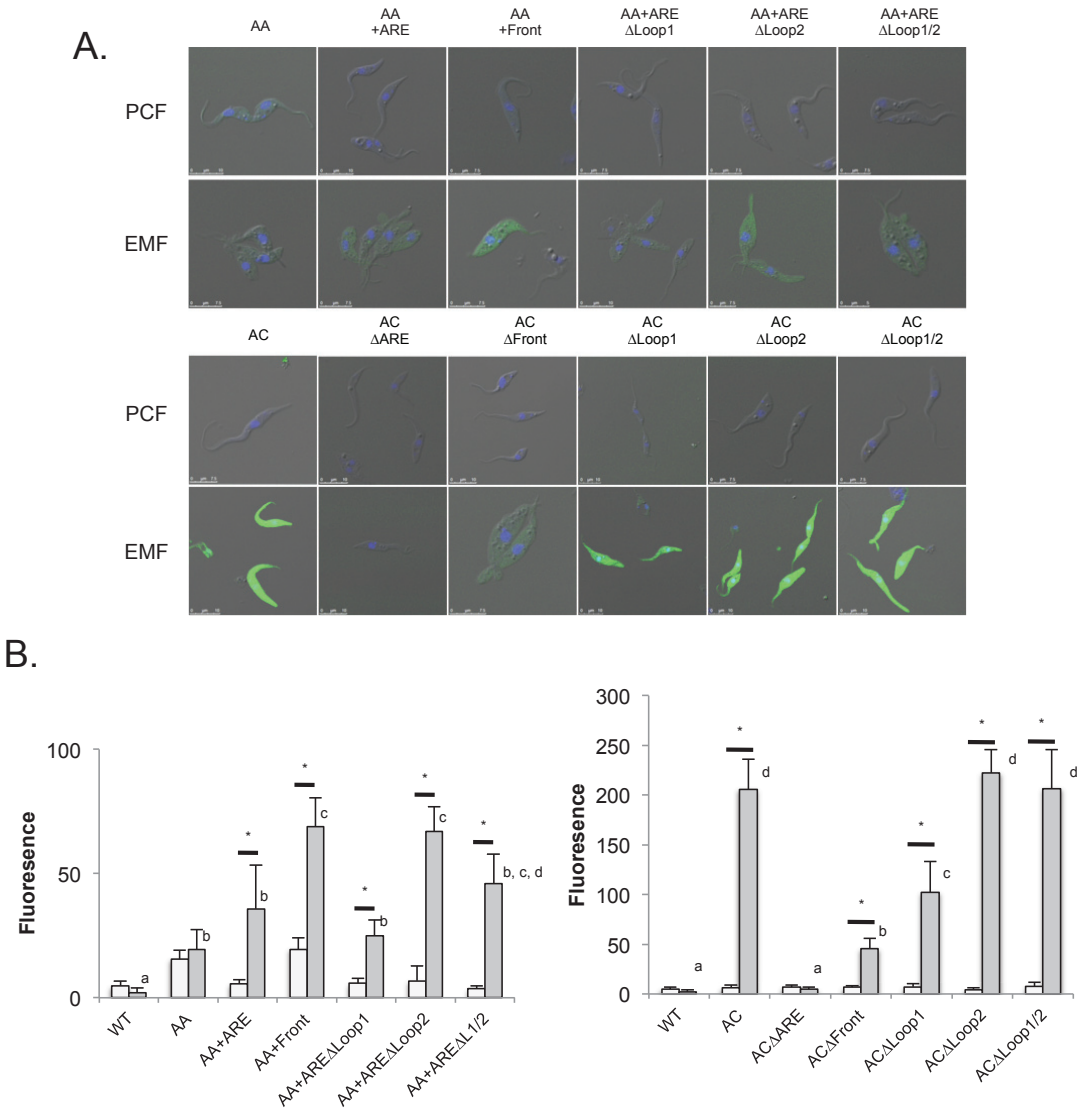

Fig. 12. Comparison of eGFP fluorescence intensity by confocal laser scanning microscopy. (A) Confocal laser scanning microscopy of each transgenic cell line. Fixed trypanosome cells were observed. Nucleus and kinetoplast DNA were stained by Hoechst 33342. Each image in each transgenic cell was captured by the same gains and voltage of the photomultiplier tubes of the microscope. (B) A comparison of eGFP fluorescence levels for each transgenic cell lines was performed using LAS-AF software. The data were expressed as mean eGFP fluorescence levels ± standard deviation. Asterisks indicate statistical significance between PCF and EMF by Student's *t*-test (*: $p<0.05$). Different letters indicate statistical significance between each EMF by Tukey's multiple comparison test ($p<0.05$).

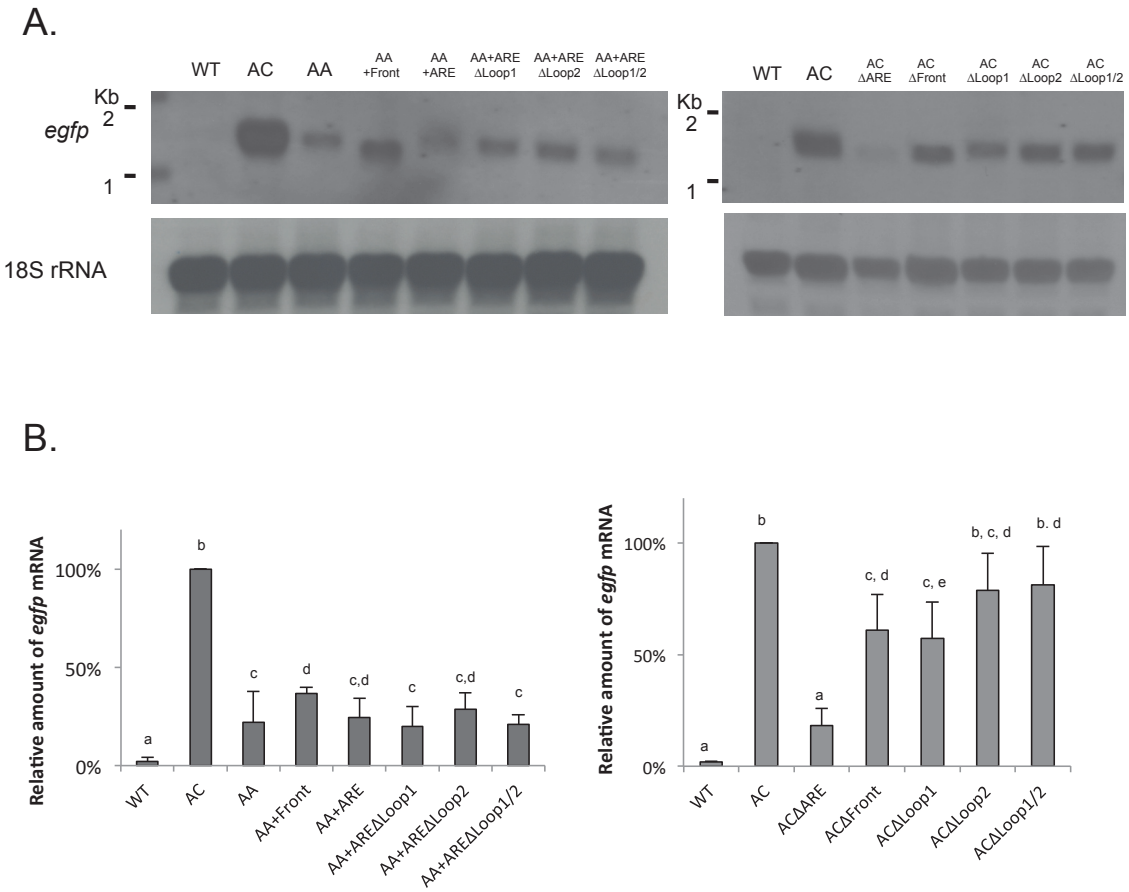

Fig. 13. The relative amount of *egfp* mRNA of transgenic EMFs. (A) Northern blot performed with total RNA extracted from EMF of wild type and transgenic trypanosome cell lines using *egfp* and 18S rRNA sequences as probes. The amount of *egfp* mRNA was normalized using 18S rRNA. (B) The data were expressed as the mean relative amount of *egfp* mRNA ± standard deviation after correction with 18S rRNA. Different letters indicate statistical significance between each EMF by Tukey's multiple comparison test ($p<0.05$).

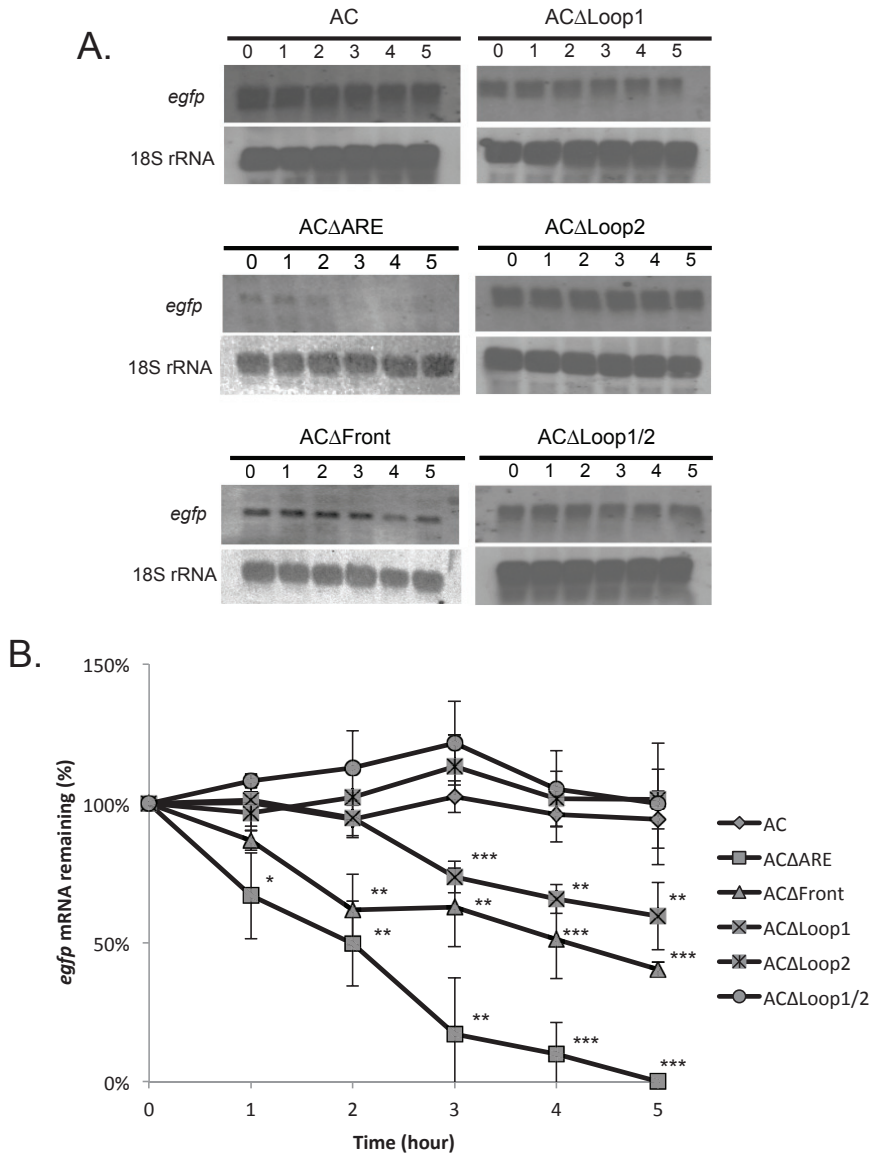

Fig. 14. The relative amount of *egfp* mRNA of transgenic EMFs. (A) Northern blot of total RNA extracted from 6 different types of transgenic trypanosome cell lines using *egfp* and 18S rRNA sequences as probes. EMFs were treated with 10 μg/ml ActD, and total RNA was prepared at indicated times (0, 1, 2, 3, 4 and 5 hours) for Northern blotting. (B) Quantification of *egfp* mRNA levels from the Northern blot. The data are shown as the mean relative amount of *egfp* mRNA ± standard deviation at each time point after correction with 18S rRNA. Asterisks indicate statistical significance between control transgenic trypanosome (AC) and other transgenic trypanosomes by Student's *t*-test (*: $p<0.05$, **: $p<0.01$, ***: $p<0.001$).

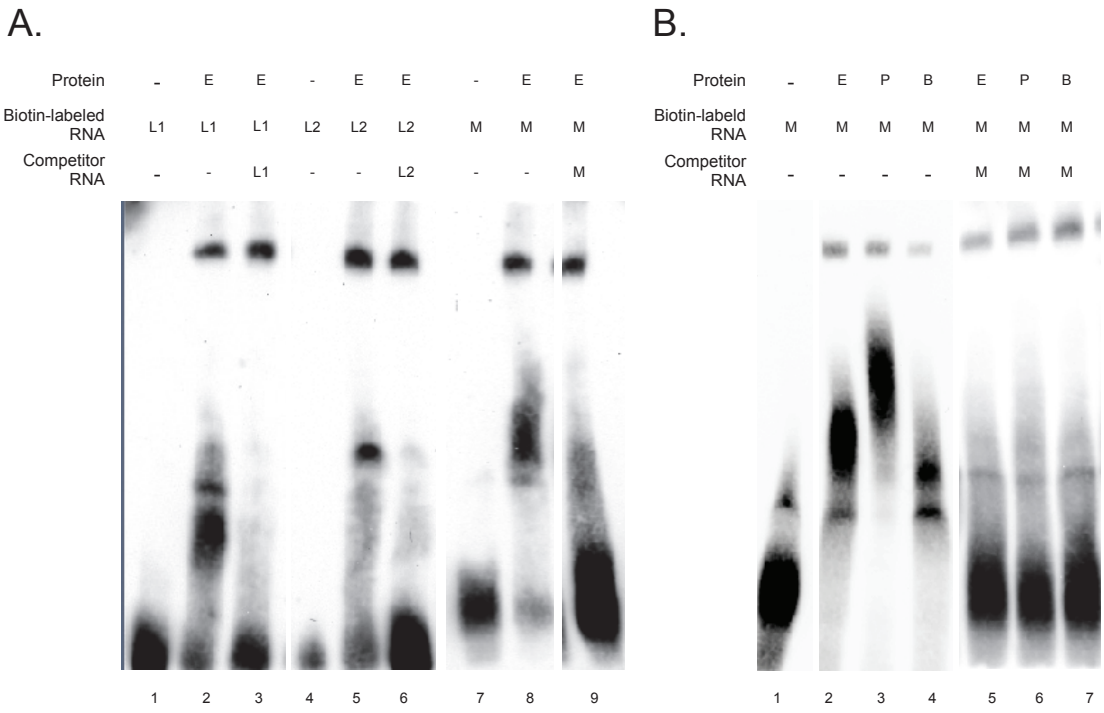

Fig. 15. The interaction of RNA and RNA binding proteins. (A) REMSA with EMF cell lysate. Three different *in vitro* transcribed elements in ARE were biotin labeled. Lanes 1, 4, 7 are biotin-labeled RNA only. REMSA was performed with 4 µg of EMF cell lysate in lanes 2, 5 and 8. REMSA was performed with 4 µg of EMF cell lysate and competitors in lanes 3, 6 and 9. L1: Loop 1-RNA, L2: Loop 2-RNA, M: ARE-M-RNA, E: EMF protein. (B) REMSA with cell lysates from different stages. The biotin-labeled *in vitro* transcribed ARE-M RNA was incubated with total protein extracted from the three life cycle parasite stages. Lane 1 is biotin-labeled ARE-M-RNA only. Lanes 2 – 4 represent REMSA with 4 µg of cell lysate (EMF cell lysate in lane 2, PCF cell lysate in lane 3, BSF cell lysate in lane 4). Lanes 5 – 7 represent REMSA with 4 µg of cell lysate and competitor ARE-M (EMF cell lysate in lane 5, PCF cell lysate in lane 6, BSF cell lysate in lane 7). M: ARE-M-RNA, E: EMF protein, P: PCF protein, B: BSF protein.

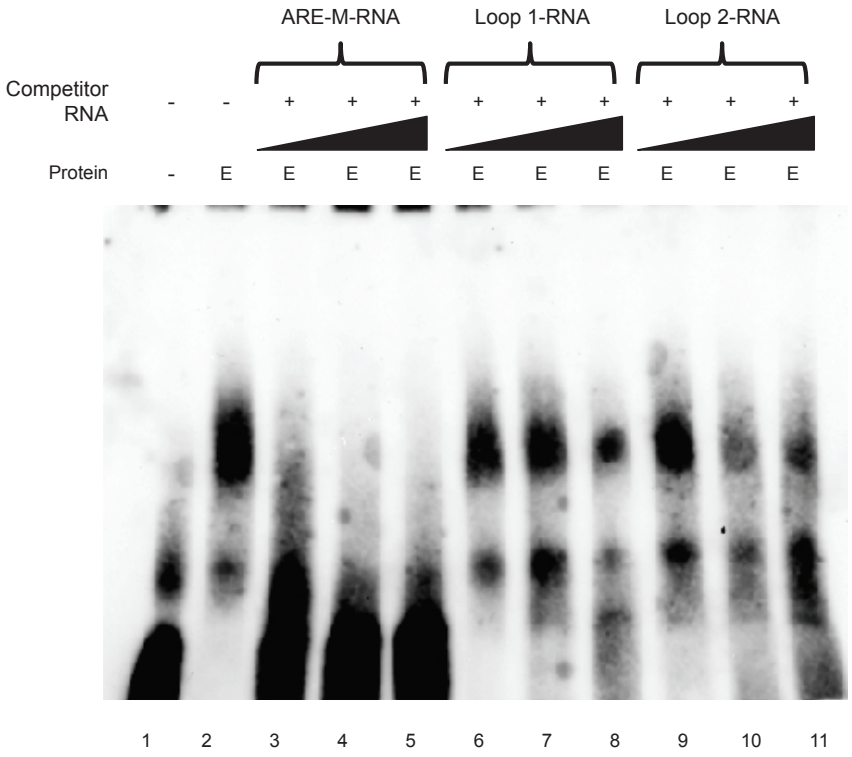

Fig. 16. The ARE-M-RNA interacts stage specific RNA binding proteins. Cross-competition assay. REMSA was performed with biotin-labeled ARE-M-RNA. Lane 1 is biotin-labeled RNA only. Lanes 2 – 11 represent REMSA with 4 µg of EMF cell lysate in the absence or presence of cross-competitors. Lanes 3 – 5, lanes 6 – 8, lanes 9 - 11 represent cross-competition with increasing amount of competitors (100-fold molar excess in lanes 3, 6 and 9; 500-fold molar excess in lanes 4, 7 and 10; 1,000-fold molar excess included in lanes 5, 8 and 11). M: ARE-M-RNA, L1: Loop 1-RNA, L2: Loop 2-RNA, E: EMF protein.

Table 3. DNA oligomers used in this experiment

Name	Primer	Sequence	Underline
Front-F	A	5'-CCCGGGCACCGGTGGACCGTGAA-3'	*Xma* I, *Sgr* AI
Front-R	B	5'-CACCGGTGCCCGGGGCGGATTGTAAC-3'	*Sgr* AI, *Xma* I
ARE-F	C	5'-CCCGGGCACCGGTGGTGCCATCGTT-3'	*Xma* I, *Sgr* AI
Loop1 front-R	D	5'-CCCGGGGCTAGCAACGATGGCACG-3'	*Xma* I, *Nhe* I
Loop middle-F	E	5'-GCTAGCCATTGTTTGTTTATTTCTTTTTATTTTTA-3'	*Nhe* I
Loop middle-R	F	5'-CCCGGGGCTAGCAAAGTCTAACGCTAAT-3'	*Xma* I, *Nhe* I
Loop2 back-F	G	5'-GCTAGCTTTTTCTCTTTTATTGTCTTTT-3'	*Nhe* I
ARE-R	H	5'-CACCGGTGCCCGGGTCGAATACTAAAAA-3'	*Sgr* AI, *Xma* I
cesp 3'UTR-R	I	5'-CCCGGGGGTAAGGAAAGGAAAGGCACAC-3'	*Xma* I
egfp-F		5'-ATCGATATGGTGAGCAAGGGC-3'	*Cla* I
egfp-R		5'-CACCGGTGTTACTTGTACAGCTGGT-3'	*Sgr* AI
18S rRNA-F		5'-GATCTGGTTGATTCTGCCAG-3'	
18S rRNA-R		5'-AAATGAGCCAGCTGCAGGTTC-3'	
T7+ARE		5'-CTAATACGACTCACTATAGGGGTTTCTTTTTATTTT	T7 RNA promoter
middle-F		TAAA-3'	
ARE middle-R		5'-TTTTCAGCATATAAACAAAATA-3'	
T7+Loop1-F		5'-CTAATACGACTCACTATAGGGTGCCATCGTT-3'	T7 RNA promoter
Loop1-R		5'-AAACAAACAATGACAATCC-3'	
T7+Loop2-F		5'-CTAATACGACTCACTATAGGGCGTTAGACTT-3'	T7 RNA promoter
Loop2-R		5'-CAATAAAAGAGAAAAATATCAAAC-3'	

Underlines indicate additional restriction enzyme sites and T7 RNA promoter sequence. Primers correspond to the alphabets of Supplementary Fig. 1.

CHAPTER III

Identification and characterization of an RNA binding protein, *T. congolense* uridine binding protein-1

3-1: Introduction

In general, the gene expression of trypanosome is regulated not in transcriptional level but in post-transcriptional level. The RBPs regulate mRNA stability via *cis*-element RNA in 3'UTR. Some *cis*-elements located in 3'UTR were reported, while the RBPs were not well identified yet. RBPs recognize the RNA sequence and/or RNA secondary structure, and interact with *cis*-element RNA. One RBP regulates many genes that have common *cis*-element sequence and/or RNA secondary structure (47). For example, *T. brucei* RBP6 that expressed at proventriculus and salivary grand might regulate cell differentiation and metacyclogenesis (45). The RBPs containing an accetylation lowers binding affinity (ALBA) domain, which has known as a nucleic acid binding affinity domain (2), were also stage-specifically expressed and regulated gene expression (57, 83). UBP1 was known as a RBP and regulating mRNA stabilization in *T. brucei* and *T. cruzi* (19, 20, 35). In addition, *T. cruzi* UBP was identified as one of a gene expression regulatory proteins for *T. cruzi* SMUG (20, 53). Moreover, *T. cruzi* UBP regulated over 30 common-metabolic-pathway associated genes, which contained characteristic RNA

motif. (65). These RNA-RBPs complex is usually constructed from multiple proteins and regulates mRNA stability (20, 83). However lack of *in vitro* culture system for *T. brucei* EMF stage, the molecular mechanisms of EMF stage-specific gene expression regulation in African trypanosome were still unknown.

In Chapter II, it was revealed that the *cis*-element RNA involving specific stabilization of mRNA in EMF was located in *cesp* 3'UTR and designated as ARE-M. In addition, specific interaction between unknown stage-specific RBPs and ARE-M-RNA was clarified. These results suggested that these unknown stage-specific RBPs regulated gene expression in EMF stage. In this chapter, in order to characterize these RBPs, expression profiles of RBPs, interaction between RBPs and ARE-M-RNA, and interaction between RBP and *T. congolense* UBP (TcUBP) 1 was investigated.

3-2: Materials and Methods

Reagents and plastic wears

Reagents, PCR primers and plastic wear were purchased or prepared as described in Chapter I.

Parasites

The PCF and EMF of *T. congolense* IL3000 were maintained *in vitro* as described in Chapter I. The MCF was isolated from EMF supernatant as described in Chapter I. The BSF was maintained *in vitro* as described in Chapter II.

Experimental animals

Eight weeks old female BALB/c mice (CLEA Japan Inc., Tokyo, Japan) were used for immunization.

Myeloma cells

The mouse myeloma cell line SP/0-Ag14 (44, 78) was cultivated at 37 ºC using GIT medium supplemented with 20 % HI-FBS.

Nucleic acid extraction

The genomic DNA and total RNA were extracted as described in Chapter I and II.

Expression and purification of recombinant proteins

Recombinant protein was expressed as a glutathione S-transferase (GST)-fusion protein using pGEX6P-1 expression vector (GE Healthcare Co., Little Chalfont, UK) in *Escherichia coli* (*E. coli*). The expression of GST-fusion recombinant protein was induced in *E. coli* by adding 100 μM isopropyl-β-thiogalactoside at 18 ºC for over night with gentle agitation. Recombinant protein was purified using Sepharose 4B beads (GE Healthcare Co.) according to manufacture's instructions. Briefly, *E. coli* was collected by a centrifugation at 13,000 x*g* for 10 min at 4 ºC and resuspended in 20 ml of Tris-NaCl-EDTA (TNE) buffer [50 mM Tris-HCl (pH 7.5), 100 mM NaCl, 0.5 % Triton-X and 2 mM EDTA]. One hundred μl of lysozyme (10 mg/ml) was added to the suspension and incubated for over night at 4 ºC with gentle agitation. Consequently, the suspension was sonicated for 10 min on ice. The pellet containing insoluble recombinant protein was resuspended with 8 M urea buffer [8M Urea, 1mM dithiothreitol, 50 mM Tris-HCl (pH 8.0) and 1 mM EDTA] and sonicated for 10 min on ice. The supernatant containing denaturated-soluble recombinant protein was refolded by dialysis. The soluble recombinant protein was immobilized on ice-chilled PBS washed Sepharose 4B beads at 4 ºC for over night with gentle agitation. The recombinant protein immobilized beads were washed 3 times with ice-chilled PBS. For removing GST from the GST-fusion recombinant protein, the GST-fusion

recombinant protein immobilized beads were incubated with precision protease buffer [50 mM Tris-HCl (pH 7.0), 150 mM NaCl, 1 mM EDTA, 1 mM DTT] containing precision protease (GE Healthcare Co.) at 4 ºC for over night with gentle agitation. The supernatant containing recombinant protein was collected by centrifugation at 1,000 x*g* 5 min at 4 ºC. After dialysis, recombinant protein was collected and stored at -80 ºC.

Recombinant proteins

Full-length TcUBP1 was expressed as a GST-fusion protein using a bacterial protein expression system. Entire ORF of TcUBP1 was amplified by PCR using appropriate PCR primer pairs (Table 4). The PCR product was cloned into a pGEX6P-1 plasmid vector for recombinant protein expression. The recombinant TcUBP1 (rTcUBP1) was expressed and cleaved GST as described above, and used to immunize mice for production of polyclonal antisera. In addition the recombinant proteins were mixed with equal volume of 2x SDS-polyacrylamide gel electrophoresis (SDS-PAGE) sample buffer [120 mM Tris-HCl (pH 6.8), 20 % glycerol, 4 % SDS, 10 % 2-β-mercaptoethanol (Sigma-Aldrich), 0.1 % bromophenol blue] and incubated at 100 ºC for 10 min.

Immunization

The rTcUBP1 was emulsified with an equal volume of TiterMax Gold adjuvant (TiterMax, USA Inc., Norcross, GA, USA) and used to immunize mice. Female mice were given primary subcutaneous injections of 100 µl followed by two booster injections by the same route at 2 weeks intervals. One week after the second booster injection, the blood was collected by cardiac puncture and incubated at room temperature for 1 h to allow clotting. Anti TcUBP1 (α-TcUBP1) sera were isolated by centrifugation at 17,000 x*g* for 10 min at 4 ºC and the sera were stored at -30 ºC until use.

Monoclonal antibody (mAb) production and purification

The rTcUBP1 was immunized to BALB/c mice as described above. After two boosters, 50 µl of rTcUBP1 solution was immunized to mice by intravenous injection (i. v.) through tail vain. Three days after the i. v., the spleen was extracted in order to isolate splenocytes. Hybridoma cells were produced by fusion of harvested splenocytes to Sp2/0-Ag14 myeloma cells (44, 78) and screened mAb production using ELISA and Western blot. Isotyping of mAbs was performed using IsoStrip mouse monoclonal antibody is typing (Roche Diagnostics, Upper Bavaria, Germany).

Saturated ammonium sulfate (SAS) solution were added to equal volume of the culture supernatant of hybridoma drop by drop with gentle string to bring the final

ammonium sulfate concentration to 50 %. The mixture was stirred gently for overnight at 4 ºC. The 50 % SAS precipitates were collected by centrifugation at 9,000 x*g* for 15 min at 4 ºC. The pellet was resuspended with Binding/Wash buffer (20 mM Na_2HPO_4, 0.15M NaCl and adjusted pH 8.0). The mAb was purified by Protein G column (GE Healthcare Co.) according to the manufactures instructions.

The binding affinities of α-TcUBP1 mAbs were determined by measuring the dissociation constant (K_d) as described before (30, 34). Briefly, constant amounts of mAbs were incubated with various concentrations of rTcUBP1 until the equilibrium was reached. The remaining unsaturated mAbs were measured by indirect ELISA. K_d was determined using Klotz plot (30, 34).

Crude protein preparation

The crude protein from BSF, PCF and EMF were extracted as described in Chapter II.

Gel electrophoresis

Protein samples were separated on SDS-PAGE (50). After electrophoresis, the gel was stained with Coomassie brilliant blue R-250.

Immunoblotting

Prior to immunoblotting, protein samples were prepared and separated by SDS-PAGE as described above. The proteins separates on the gel were electrophoretically transferred to a polyvinylidene difluoride (PVDF) membrane (Hybond-P, GE Healthcare Co.) using a semi-dry blotting system AE-6677 (ATTO Co, Tokyo, Japan). Briefly, the membrane was blocked by 5 % skim milk in PBS containing 0.02 % Tween 20 (PBS-T) at 4 ºC overnight and was then incubated with primary antibody diluted 1,000 times with PBS-T for 1 h at room temperature. After washing with PBS-T, the membrane was incubated with secondary antibody (ECL^{TM} anti-mouse IgG, Horseradish Peroxidase linked whole antibody from sheep, GE Healthcare Co.) diluted 5,000 times with PBS-T for 1 h at room temperature. The result was visualized by 3, 3-diaminobenzidine and H_2O_2.

Northern blot analysis

Total RNA was extracted from BSF, PCF and EMF of wild type trypanosome and blotted onto Hybond-N$^+$ membrane (GE Healthcare Co.) as described in Chapter II. The full-length Tc*ubp1* was labeled using the AlkPhos direct labeling kit and hybridized according to the manufacturer's instructions as described in Chapter II. Chemiluminescence signal was generated using the CDP-*Star* detection reagent (GE Healthcare Co.).

Indirect fluorescent antibody test (IFAT)

Each stage of trypanosome was obtained from *in vitro* culture as described in Chapter I, II and above. The trypanosomes were suspended in PSG. The cell suspensions were spread over glass slides (Glass slide printed with highly water-repellent mark, Matsunami Glass Ind., Ltd., Tokyo, Japan), air-dried and fixed with 100 % methanol for 10 min at room temperature. For blocking, the specimens were incubated with PBS containing 5 % skim milk at 4 ºC over night. The specimens were incubated with primary antibodies diluted with 40 times with PBS containing 0.5 % skim milk for 1 h at 37 ºC. After washing with PBS, they were incubated with secondary antibody [Alexa Fluor 488 goat anti-mouse IgG (H+L), Invitrogen, Carlsbad, CA, USA] diluted 500 times with PBS containing 0.5 % skim milk and 1 μg/ml Hoechest 33342 (Dojundo, Co. Ltd., Kumamoto, Japan) for 1 h at room temperature. The specimens were observed by confocal laser scanning microscopy (Leica TCS SP5, Leica Microsystems, Wetzlar, Germany).

RNA electro mobility shift assay (REMSA)

REMSA were preformed with 0.5 μM of rTcUBP1 and other component as described in Chapter II. For analysis of protein-specific Ab interaction, REMSA was

preformed with 2.7 μM α-TcUBP1 mAb #2, 0.5 μM of rTcUBP1 and other component as described in Chapter II.

Co-immunoprecipitation assay

Forty μg of purified α-TcUBP1 mAb #2 or anti-*Plasmodium falciparum* 1-Cys peroxiredoxin (α-Pf1-Cys-Prx) (34) were directory cross-linked by Protein A/G agarose beads (Thermo Fisher Scientific Inc., Waltham, MA, USA) according to the manufactures instruction. For intact protein complex that is targeted by the antibodies was captured onto the agarose via the antibodies. Crude trypanosome protein (1,000 μg) was reacted with antibody cross-linked beads at 4 ºC for overnight. After capturing reaction, the beads were washed by crude protein extraction buffer (as described in Chapter II) without detergent three times. The captured protein complexes were eluted by elution buffer as following manufacture instruction and separated by SDS-PAGE. The gel was stained by silver staining kit according to the manufactures instruction.

3-3: Results

Expression of recombinant TcUBP1 and production of anti-TcUBP1 monoclonal antibodies

For identification of these unknown RBPs, the orthologues of RBPs in trypanosome were analyzed *in silico*. The UBP1 orthologous gene was found in *T. congolense* genome through *in silico* analysis and named as *T. congolense ubp1* (Tc*ubp1*). The RRM and two RNP domains were conserved in TcUBP1 (Fig. 17).

Full-length of Tc*ubp1* was cloned into protein expression vector and expressed as GST-fusion protein (rTcUBP1-GST). The rTcUBP1-GST was cleaved by precision protease in order to remove GST. Before and after protease treatments, the recombinant protein was analyzed by SDS-PAGE. The recombinant proteins that had apparent molecular masses of 50 kDa (rTcUBP1-GST: Fig. 18-A, lane 1) and 24 kDa (rTcUBP1: Fig 18-A, lane 2) were obtained in full agreement with predicted masses. The α-TcUBP1 mAb hybridoma clones were produced from rTcUBP1-immunized mice splenocytes fused with SP2/14-Ag myeloma cells. Four hybridoma clones were selected by ELISA (data not shown) and Western blot (Fig. 18-B). Among 4 positive clones, 3 clones were selected and α-TcUBP1 mAbs were purified from the culture supernatant. The mAb #2 showed the highest affinity with K_d 45 nM (Table 5) calculated by Klotz plot analysis. While mAbs #1 and #3 showed lower affinity than that of mAb #2 (Table 5).

Expression profile of TcUBP1

Firstly, the mRNA expression profile was analyzed by Northern blot using full-length of *Tcubp1* probe. The result showed that Tc*ubp1* mRNA was equally transcribed in EMF, BSF and PCF (Fig. 19-A). Next, the TcUBP1 protein expression pattern was analyzed by Western blot. The result clearly showed TcUBP1 was transcribed in EMF, BSF and PCF (Fig. 19-B). Finally, the cellular localization and expression profile of TcUBP1 protein were analyzed by confocal laser scanning microscopy. The results showed that TcUBP1 was expressed in all stage of *T. congolense* as scattered localization in cytosol.

TcUBP1 interacted with *cis*-element RNA, ARE-M-RNA

To reveal of RNA binding ability of TcUBP1, the ARE-M-RNA was reacted with equal molar amount of rTcUBP1 or BSA. As a result, the gel-shift signal was only observed by REMSA when ARE-M-RNA was reacted with rTcUBP1 (Fig. 20, lane 2). Moreover, the gel-shift signal was super-shifted to higher molecular mass position by adding purified α-TcUBP1 mAb #2 (Fig. 20, lane 4). When the RNase inhibitor was added to the reactant to protect the degradation of ARE-M-RNA, the gel-shift signal patterns were not changed (Fig. 20, lanes 3 and 5).

TcUBP1 interacts with some stage-specific proteins

TcUBP1 interacted with ARE-M-RNA, however TcUBP1 expressed in all developmental stages of the parasite as described above. Therefore, I hypothesized that the EMF stage-specific protein complex containing TcUBP1 recognized ARE-M-like RNA sequence and regulated EMF stage-specific gene expression. The intact TcUBP1-protein complexes were purified from crude protein extracted from BSF, PCF and EMF by co-immunoprecipitation assay. The common proteins were observed at the molecular masses of 25 kDa, 37 kDa and 50 kDa among all developmental stages. Because mAb #2 directory cross-linked to Protein A/G agarose beads, the antibody was not eluted. Therefore, these 25 kDa and 50 kDa proteins were not heavy and light chains of mAb #2 but TcUBP1-protein complexes. In addition to the common proteins, some stage-specific proteins were also detected. The 20 kDa and 18 kDa proteins were observed in BSF (Fig. 21, lane 1). The 15 kDa and 13 kDa proteins were observed in PCF (Fig. 21, lane 3). Furthermore, 13 kDa, 18 kDa and 20 kDa proteins were observed in EMF (Fig. 21, lane 5). On the other hand, these common and stage-specific proteins were not observed using α-Pf1-Cys-Prx (34) as negative control mAb (Fig. 21, lanes 2, 4 and 6).

3-4: Discussion

The stage-specific gene expression was regulated by *cis*-element RNA and RBPs complex (21, 46, 47). An RNA binding protein, namely UBP also regulates stage-specific mRNA stability with other proteins in trypanosome (19, 20, 35). In Chapter II, it was revealed that the stage-specific RBPs interacted with *cis*-element RNA, ARE-M-RNA. In this chapter, gene expression profile, RNA binding activity and protein-protein interaction of TcUBP1 were analyzed.

The Tc*ubp1* was characterized through *in silico* analysis. From the sequence analysis, the RRM and two RNP domains were well conserved in TcUBP1 as other UBP1 orthologs (Fig. 17). Because these RNA binding motifs and domains were important for RNA-protein interaction (15), TcUBP1 was also predicted as an RBP.

From gene expression profile analysis, TcUBP1 was expressed in each stage trypanosome in cytosol. These results suggested that TcUBP1 was constitutively expressed in all developmental stages of trypanosomes. It was clearly shown by REMSA that rTcUBP1 interacted with ARE-M-RNA. This result suggested that TcUBP1 is a component of the RBPs, which interacts with ARE-M-RNA. However, expression profile of TcUBP1 did not show stage specificity. Therefore, the other proteins that interact with TcUBP1 may be required in order to express genes in stage-specific manner. To analyze the protein-protein interaction of TcUBP1, co-immunopresipitation assay was performed. The results suggested that some

common and stage-specific proteins interacted with TcUBP1. Although the key molecules for regulation of EMF stage-specific gene expression are still unknown, stage-specific proteins that interact with TcUBP1 were found in this study. Therefor the key protein molecules will be identified by mass spectrometry in future.

3-5: Summary

The UBP1 orthologue gene, TcUBP1 was characterized *in silico* analysis. From the sequence analysis, the RNA recognition motifs and two RNP domains were well conserved in TcUBP1 as other UBP1 orthologs (Fig. 17). Three α-TcUBP1 mAb clones were established and were purified (Fig. 18). One of them (mAb #2) showed high affinity (Table 5). Northern blot analysis, Western blot analysis and immunofluorescent microscopy clearly showed that TcUBP1 was expressed as scattered cytosolic localization in each stage of trypanosome. REMSA using rTcUBP1 and ARE-M-RNA clearly showed that rTcUBP1 interacted with ARE-M-RNA. This result implied that TcUBP1 was a component of the RBPs that interacts with ARE-M-RNA. To analyze the protein-protein interaction of TcUBP1, co-immunopresipitation assay was performed. The results suggested that some common and stage-specific proteins interacted with TcUBP1.

In conclusion, the stage-specific RBPs, which were component from TcUBP1 and other unknown RBPs, might regulate EMF stage-specific mRNA stabilization and destabilization via interacting with ARE-M-RNA.

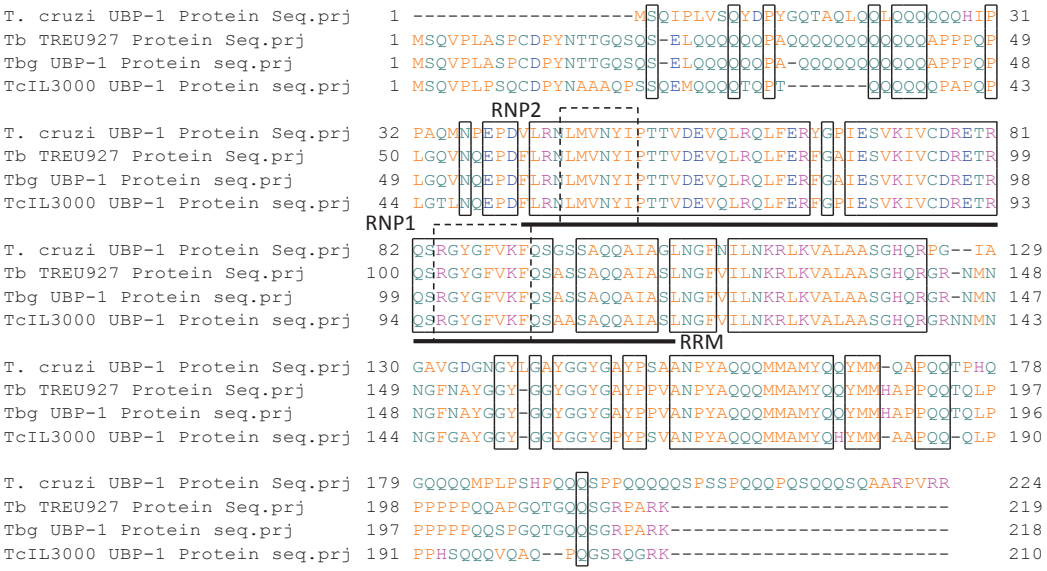

Fig. 17. Sequence analysis of RNA binding proteins. Amino acid sequence alignment of *T. cruzi* UBP1, *T. brucei* TREU927 strain UBP1, *T. brucei gambience* UBP1 and *T. congolense* UBP1. Dashed line boxes are indicate RNP domain 1 octamer and RNP domain 2 hexamer, respectively. The RNA recognition motif (RRM) is underlined. The identicial residues between the four UBP1 orthologue sequences are boxed.

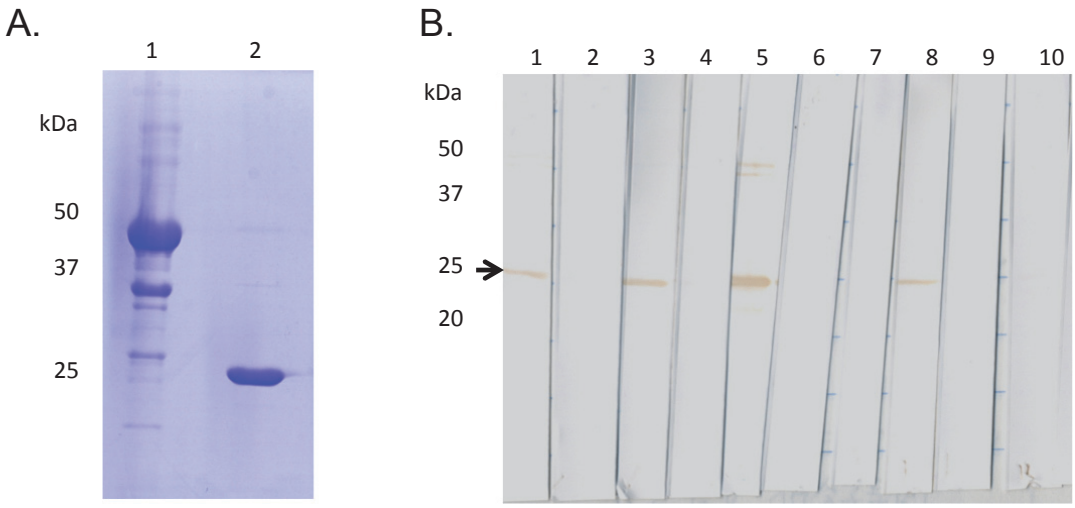

Fig. 18. Monoclonal antibody purification. (A) Recombinant TcUBP1 was expressed and purified using *E. coli* recombinant expression system. The molecular sizes (25, 37 and 50 kDa) are shown in the left side of CBB stained SDS-PAGE image. Purified rTcUBP1-GST by Sepharose 4B beads (lane 1). Purified rTcUBP1 after presision protease treatment (lane 2). (B) Monoclonal antibody (mAb) expression hybridoma cells were selected by western blot. The molecular sizes (25, 37 and 50 kDa) are shown in the left side of western blot image. rTcUBP1 was immnoblotted using the hybridoma supernatant (lane 1 – 10). Four hybridoma clones express anti-TcUBP1 mAb (lane 1, 3, 5 and 8). Arrow indicates the signals of rTcUBP.

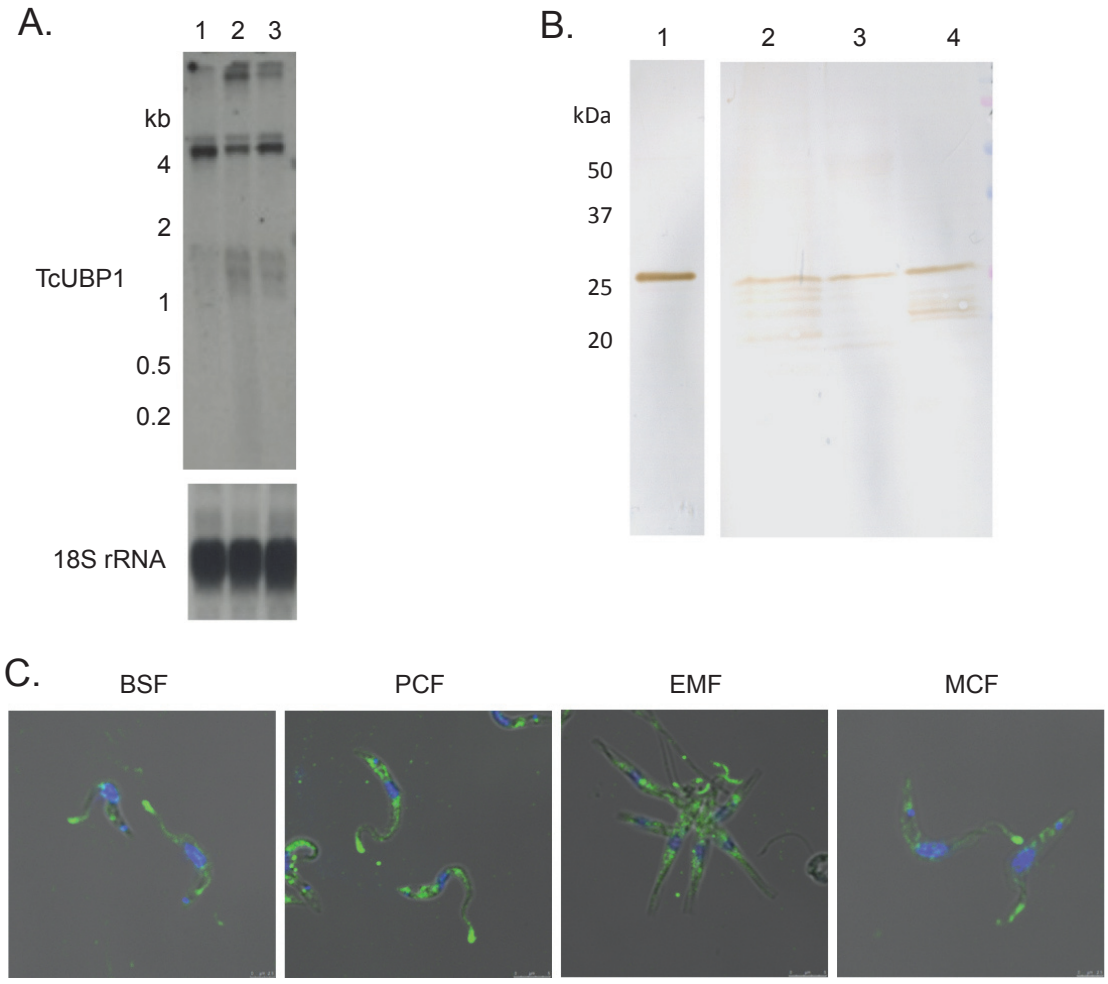

Fig. 19. Expression and localization analysis of TcUBP1 in *T. congolense*. (A) The mRNA expression profile of Tc*ubp1* was analyzed by northern blot. The molecular sizes (0.2, 0.5, 1, 2 and 4 kb) are shown in the left side of northern blot image. Total RNA was extracted from each stage (EMF RNA in lane 1, BSF RNA in lane 2 and PCF RNA in lane 3). (B) The expression profile of TcUBP1 was analyzed by western blot. The molecular sizes (20, 25, 37 and 50kDa) are shown in the left side of western blot image. rTcUBP1 was in lane 1. Total cell lysate was extracted from each stage trypanosome (EMF cell lysate in lane 2, BSF cell lysate in lane 2 and PCF cell lysate in lane 4). (C) The localization of TcUBP1 was analyzed by IFAT. Methanol fixed each stage of trypanosome was observed. Nucleus and kinetoplast DNA were stained by Hoechst 33342.

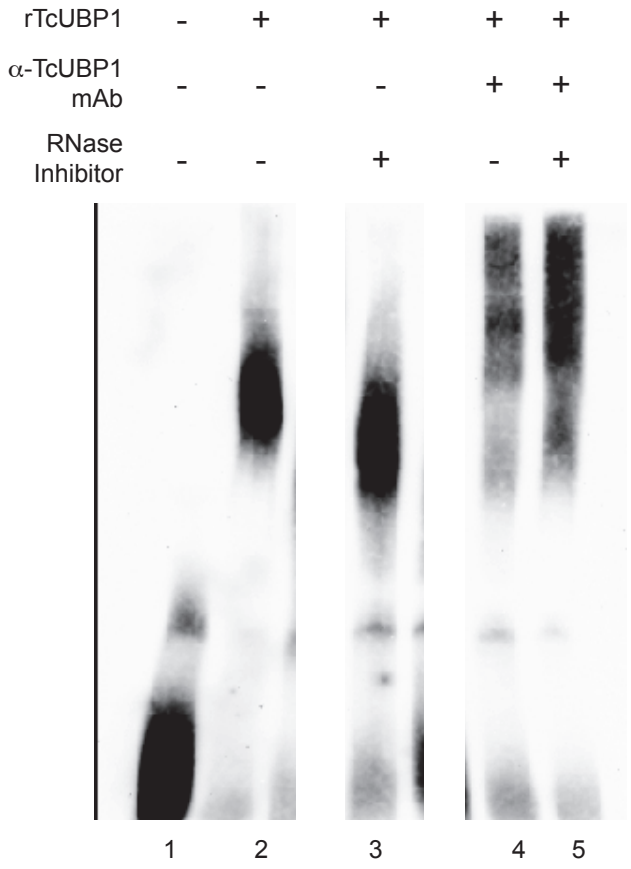

Fig. 20. The interaction of ARE-M-RNA and rTcUBP1. REMSA with rTcUBP1. *In vitro* transcribed element in ARE-M-RNA was biotin labeled. REMSA was performed with rTcUBP1 (lanes 2 – 5) and RNase inhibitor (lanes 3 and 5). When adding the α-TcUBP1 mAb, these gel shift signals are super-shifted (lanes 4 and 5).

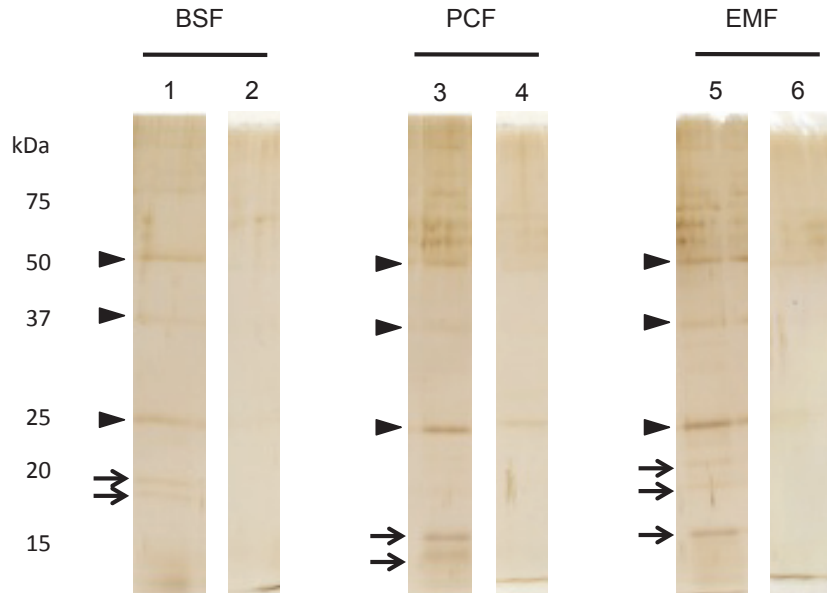

Fig. 21. Protein-Protein interaction analysis by co-immunoprecipitation assay. The TcUBP1 and other protein interaction was analyzed by co-immunoprecipitation assay. The molecular sizes (15, 20, 25, 37, 50 and 75 kDa) are shown in the left side of silver stained gel image. Protein G sepharose was immobinized by α-TcUBP1 mAb (lane 1, 3 and 5) or α-*Plasmodium falciparum*1-Cys-Prx mAb (lane 2, 4 and 6). The allow heads indicate common protein bands among stages. While, the allows indicate stage specific bands. Co-immunoprecipitation assay was performed using BSF protein, PCF protein, or EMF protein.

Table 4. Dissociation constants of monoclonal antibodies determined by ELISA

mAb*	Isotype	Concentration**	Kd (nM)***
1	IgG l/κ	2.5 μg/μl	6485 nM
2	IgG l/κ	0.8 μg/μl	45 nM
3	IgG l/κ	1.0 μg/μl	825 nM

*: Monoclonal antibodies tested, **: The concentration of mAb solution was measured by BCA assay, ***: The dissociation constants (Kd) were calculated by Klotz plot.

GENERAL DISCUSSION

The genome sequence of *T. b. brucei* was published in 2005 (8). Until now, the genome sequencing projects of *T. b. gambiense, T. congolense* IL3000 and *T. vivax* were completed (42, 43). The genomic information combined with cutting-edge molecular biological techniques has lead greatly to understanding cell biology of African trypanosomes. Nevertheless, many people are still suffering from trypanosomosis in sub-Saharan African countries. Because trypanosome undergoes antigenic variation of VSG in blood circulation to evade from adaptive immunity, all attempts for vaccine development have been failed (40). Insect stages of trypanosome express stage-specific genes for adaptation to external environment, however they do not undergo their surface antigenic variation. Therefore, insect stages of trypanosome have been focused as newly trypanosome control targets (80).

This study was aimed at clarification of the molecular mechanisms of expression regulation of EMF stage-specific genes. In particular, *cis*-element for regulation of EMF stage-specific gene expression and RBPs, which interact with *cis*-element, were identified. The overall goal of this study was establishment of new control measures against African trypanosomosis by disrupting cell cycle progression and/or transmission of insect stage trypanosome.

Stage-specific genes play important roles for adaptation to their external

environments, proliferation and cell differentiation into next stage (46, 54). Prior to this study, the gene expression mechanisms of African trypanosome were revealed using *T. brucei* PCF and BSF because of ease of their *in vitro* cultivation. On the other hand, lack of stable *in vitro* culture systems for *T. brucei* EMF and MCF, the molecular mechanisms of cell differentiation from PCF to EMF and EMF stage-specific gene expression of African trypanosome were still unknown. Recently, the *cesp* as a novel EMF stage-specific expression gene in *T. congolense* was reported (76) and useful genetic manipulation tools were also reported in *T. congolense* (16, 41, 77). Consequently, it was analyzed that the molecular mechanisms of gene expression regulation of *cesp* as the model of EMF stage-specific gene. Four different *egfp* expression cassettes were transfected into wild type *T. congolense* and *egfp* expressed transgenic trypanosomes were established. Correct single copy integration was confirmed by Southern blot analysis and real-time PCR. This result suggested that the amount of transcript was stable among all the developmental stages of each transgenic trypanosome. Transgenic trypanosome that was transfected with *egfp* fused with *actin* 3'UTR constantly expressed eGFP throughout the developmental stages. While transgenic trypanosome that was transfected with *egfp* fused with *cesp* 3'UTR only expressed eGFP in EMF and MCF. However, the *egfp* mRNA expression of this transgenic trypanosome was exclusively higher in EMF. This *egfp* mRNA expression pattern was completely in accord with *cesp* expression pattern. Therefore it was

suggested that eGFP in MCF was remnant from parental EMF cell and EMF stage-specific expression of *cesp* was regulated by its 3'UTR by EMF stage-specific mRNA stabilization. These results implied that *cis*-element for EMF stage-specific gene expression regulation was located with in *cesp* 3'UTR.

Next, the interaction between the *cis*-element located within *cesp* 3'UTR and RBPs were analyzed. Characteristic sequences and secondary structures were found in the middle parts of *cesp* 3'UTR by *in silico* analysis. Twelve different of *egfp* expression cassettes were constructed and transgenic trypanosome cell lines were established. Comparative analyses of eGFP fluorescence intensity, eGFP expression pattern and the amount of *egfp* mRNA between transgenic trypanosomes, the ARE-M element was identified as a novel *cis*-element for EMF stage-specific *cesp* expression regulation. Moreover, it was revealed that gene expression in EMF was regulated by stage-specific mRNA stabilization via ARE-M-RNA and RBPs complex.

Finally, the expression profile and RNA binding activity of TcUBP1 were analyzed. TcUBP1 was identified as an orthologous gene of *T. cruzi* and *T. brucei* UBP1. UBP1 in *T. cruzi* and *T. brucei* was reported as RBP, and regulates stage-specific gene expression (19, 20, 35). RRM and RNPs that recognize and interact with RNA (15) were conserved in TcUBP1 and this result suggested TcUBP1 was RBPs. In addition, the ARE-M-RNA binding ability of rTcUBP1 was revealed by REMSA. Moreover, the result from co-immunoprecipitation assay using α-TcUBP1

mAb #2 suggested that TcUBP1 interacted with some unknown stage-specific proteins. These results suggested that stage-specific RBPs, which consist from TcUBP1 and other proteins, interacted with ARE-M. However, the expression analysis clearly showed that TcUBP1 was constantly expressed during life cycle. Therefore, it was revealed that TcUBP1 is not essential protein factor for *cesp* expression regulation and unknown proteins, which interact with TcUBP1, may regulate *cesp* expression.

The study in this thesis clearly showed that ARE-M element in *cesp* 3'UTR was a novel *cis*-element for EMF stage-specific gene expression regulation. In addition, it was shown that some RBPs including of TcUBP1 interacted with this element. Since many stage-specific gene expressions switch during cell differentiation from PCF to EMF, this molecular mechanism that regulates *cesp* expression might also regulate other EMF stage-specific expression genes. Furthermore *T. congolense* ARE like sequence and UBP1 were already reported as gene expression regulation factor of *T. cruzi*. Therefore, it was hypothesized that the same mechanisms of gene expression regulation were extrapolate other trypanosome EMF stage-specific genes. Although the ARE-M-like sequence in other *T. congolense* EMF stage-specific expression genes were not investigated. Because TcUBP1 was constitutively expressed during life cycle, the key protein that regulated EMF stage-specific gene expression and cell differentiation from PCF to EMF will be clarified.

CONCLUSION

This thesis mainly reported the identification and characterization of a novel *cis*-element RNA located within *cesp* 3'UTR, namely ARE-M, and TcUBP1 that interacted with ARE-M-RNA.

Firstly, for clarification of gene expression regulation ability of *cesp* 3'UTR, pSAK *egfp* expression plasmid was modified and transfected into *T. congolense* genome. The correct single integration of these *egfp* expression cassettes was confirmed by Southern blot analysis and real-time PCR. Comparison of *egfp* mRNA expression between transgenic trypanosomes by real-time RT-PCR, significantly higher *egfp* mRNA expression was observed when *egfp* fused with *cesp* 3'UTR. These results showed that *cis*-element for EMF stage-specific *cesp* expression might be located in its 3'UTR. (Chapter I)

For identification of *cis*-element in *cesp* 3'UTR, *egfp* expression cassettes which was fused with modified 3'UTR were constructed. The analysis of eGFP fluorescence intensity and the amount of *egfp* mRNA combined with *egfp* mRNA stability test showed that ARE-M was a required *cis*-element for EMF stage-specific *cesp* expression. Additional analysis using REMSA revealed that stage-specific RBPs recognize ARE-M sequence and interact with this element. (Chapter II)

For clarification of unknown RBPs, which interacted with ARE-M, Tc*ubp1* was

cloned and expressed as rTcUBP1. The rTcUBP1 interacted with ARE-M and some unknown stage-specific proteins. This result suggested that protein-protein complex which composed TcUBP1 and other proteins regulated *cesp* expression. However, immunofluorescent microscopy and Western blot analysis using α-TcUBP1 combined with Northern blot analysis clearly showed that TcUBP1 was constantly expressed in cytosol during life cycle. Therefore it was suggested that unknown protein(s) is the most important regulatory factor(s) for EMF stage-specific *cesp* expression. (Chapter III)

Here it was revealed that ARE-M-RNA in *cesp* 3'UTR is a novel *cis*-element and the EMF stage-specific *cesp* expression is regulated by RNA-EMF stage-specific RBPs complex which consist of the combination of ARE-M-RNA, TcUBP and some unknown RBPs. This molecular mechanism of EMF stage-specific gene expression might regulate other EMF stage-specific genes, which is necessary for the trypanosome undergo cell differentiation from PCF into EMF.

ACKNOWLEGEMENTS

This research work was carried out at the National Research Center for Protozoan Diseases (NRCPD), Obihiro University of Agriculture and Veterinary Medicine. This study was financially supported by founds from a Grant-in-Aid for scientific Research from JSPS.

I owe my gratitude to my supervisor Prof. Shin-ichiro Kawazu (NRCPD) for intellectual guidance, constructive criticism and encouragement during the period of study. I express my deepest gratitude to Prof. Noboru Inoue (NRCPD) for his perennial supervision, spirit of forbearance and teaching me profundity of biology. I am grateful to Assoc. Prof. Shinya Fukumoto for valuable comments and suggestion during group seminar. I am thankful to Dr. Tatsuya Sakurai for establishment powerful and useful transgenic tools about trypanosomes and fundamentally teaching me cellular biology. I also thank Dr. Junya Yamagishi, Dr. Tadashi Okada, Dr. Masahito Asada, Dr. Tatsunori Masatani and Dr. Madoka Seki-Ichikawa to support my Ph. D. researches.

I am thankful to Prof. Ikuo Igarashi, Prof. Naoaki Yokoyama, Prof. Xuenan Xuan, Assoc. Prof. Makoto Igarashi, Assoc. Prof. Yoshifumi Nishikawa and Assistant Prof. Rika Shirafuji-Umemiya for their critical reviews of my progress report during the NRCPD seminar.

My special thanks to Prof. Hiroshi Suzuki (NRCPD), Prof. Tadashi Itagaki (Iwate University), Prof. Kazuaki Takehara (Tokyo University of Agriculture and Technology) and Assoc. Prof. Yasuhiro Takashima (Gifu University) for their valuable suggestion and their patience in revising this dissertation.

I thank former and current members of the Research Unit for Advanced Preventive Medicine and Research Unit for Vector Biology.

Lastly, I thank my parents Hideaki Suganuma and Kayoko Suganuma, my junior brother Shogo Suganuma and Masahiro Suganuma for everlasting love of family. I grateful thank my wife Hirono Suganuma for helping and supporting me every time.

REFERENCES

1. Aksoy, S. (2003). Control of tsetse flies and trypanosomes using molecular genetics. Vet Parasitol. 115, 125-145.
2. Aravind, L., Iyer, L. M., Anantharaman, V. (2003). The two faces of Alba: the evolutionary connection between proteins participating in chromatin structure and RNA metabolism. Genome Biol. 4, R64.
3. Baker, N., De Koning, H. P., Maser, P., Horn, D. (2013). Drug resistance in African trypanosomiasis: the melarsoprol and pentamidine story. Trends Parasitol. 29, 110-118.
4. Barreau, C., Paillard, L., Osborne, H. B. (2005). AU-rich elements and associated factors: are there unifying principles? Nucleic Acids Res. 33, 7138-7150.
5. Barrett, M. P., Burchmore, R. J., Stich, A., Lazzari, J. O., Frasch, A. C., Cazzulo, J. J., Krishna, S. (2003). The trypanosomiases. Lancet. 362, 1469-1480.
6. Beecroft, R. P., Roditi, I., Pearson, T. W. (1993). Identification and characterization of an acidic major surface glycoprotein from procyclic stage *Trypanosoma congolense*. Mol Biochem Parasitol. 61, 285-294.
7. Berberof, M., Vanhamme, L., Tebabi, P., Pays, A., Jefferies, D., Welburn, S., Pays, E. (1995). The 3'-terminal region of the mRNAs for VSG and procyclin can confer stage specificity to gene expression in *Trypanosoma brucei*. EMBO J. 14, 2925-2934.
8. Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D. C., Lennard, N. J., Caler, E., Hamlin, N. E., Haas, B., Bohme, U., Hannick, L., Aslett, M. A., Shallom, J., Marcello, L., Hou, L., Wickstead, B., Alsmark, U. C., Arrowsmith, C., Atkin, R. J., Barron, A. J., Bringaud, F., Brooks, K., Carrington, M., Cherevach, I., Chillingworth, T. J., Churcher, C., Clark, L. N., Corton, C. H., Cronin, A., Davies, R. M., Doggett, J., Djikeng, A., Feldblyum, T., Field, M. C., Fraser, A., Goodhead, I., Hance, Z., Harper, D., Harris, B. R., Hauser, H., Hostetler, J., Ivens, A., Jagels, K., Johnson, D., Johnson, J., Jones, K., Kerhornou, A. X., Koo, H., Larke, N., Landfear, S., Larkin, C., Leech, V., Line, A., Lord, A., Macleod, A., Mooney, P. J., Moule, S., Martin, D. M., Morgan, G. W., Mungall, K., Norbertczak, H., Ormond, D., Pai, G., Peacock, C. S., Peterson, J., Quail, M. A., Rabbinowitsch, E., Rajandream, M. A., Reitter, C., Salzberg, S. L., Sanders, M., Schobel, S., Sharp, S., Simmonds, M., Simpson, A. J.,

Tallon, L., Turner, C. M., Tait, A., Tivey, A. R., Van Aken, S., Walker, D., Wanless, D., Wang, S., White, B., White, O., Whitehead, S., Woodward, J., Wortman, J., Adams, M. D., Embley, T. M., Gull, K., Ullu, E., Barry, J. D., Fairlamb, A. H., Opperdoes, F., Barrell, B. G., Donelson, J. E., Hall, N., Fraser, C. M., Melville, S. E., El-Sayed, N. M. (2005). The genome of the African trypanosome *Trypanosoma brucei*. Science. 309, 416-422.

9. Bhalla, N. (2002). Pan African group takes lead against the tsetse fly. Lancet. 359, 686.
10. Brenndorfer, M., Boshart, M. (2010). Selection of reference genes for mRNA quantification in *Trypanosoma brucei*. Mol Biochem Parasitol. 172, 52-55.
11. Bringaud, F., Baltz, D., Baltz, T. (1998). Functional and molecular characterization of a glycosomal PPi-dependent enzyme in trypanosomatids: pyruvate, phosphate dikinase. Proc Natl Acad Sci U S A. 95, 7963-7968.
12. Burri, C., Brun, R. (2003). Eflornithine for the treatment of human African trypanosomiasis. Parasitol Res. 90 Supp 1, S49-52.
13. Burri, C., Nkunku, S., Merolle, A., Smith, T., Blum, J., Brun, R. (2000). Efficacy of new, concise schedule for melarsoprol in treatment of sleeping sickness caused by *Trypanosoma brucei gambiense*: a randomised trial. Lancet. 355, 1419-1425.
14. Chappuis, F., Pittet, A., Bovier, P. A., Adams, K., Godineau, V., Hwang, S. Y., Magnus, E., Buscher, P. (2002). Field evaluation of the CATT/Trypanosoma brucei gambiense on blood-impregnated filter papers for diagnosis of human African trypanosomiasis in southern Sudan. Trop Med Int Health. 7, 942-948.
15. Clery, A., Blatter, M., Allain, F. H. (2008). RNA recognition motifs: boring? Not quite. Curr Opin Struct Biol. 18, 290-298.
16. Coustou, V., Guegan, F., Plazolles, N., Baltz, T. (2010). Complete *in vitro* life cycle of *Trypanosoma congolense*: development of genetic tools. PLoS Negl Trop Dis. 4, e618.
17. Coutinho-Abreu, I. V., Ramalho-Ortigao, M. (2010). Transmission blocking vaccines to control insect-borne diseases: a review. Mem Inst Oswaldo Cruz. 105, 1-12.
18. D'orso, I., Frasch, A. C. (2001). Functionally different AU- and G-rich *cis*-elements confer developmentally regulated mRNA stability in *Trypanosoma cruzi* by interaction with specific RNA-binding proteins. J Biol Chem. 276, 15783-15793.
19. D'orso, I., Frasch, A. C. (2001). TcUBP-1, a developmentally regulated

U-rich RNA-binding protein involved in selective mRNA destabilization in trypanosomes. J Biol Chem. 276, 34801-34809.
20. D'orso, I., Frasch, A. C. (2002). TcUBP-1, an mRNA destabilizing factor from trypanosomes, homodimerizes and interacts with novel AU-rich element- and Poly(A)-binding proteins forming a ribonucleoprotein complex. J Biol Chem. 277, 50520-50528.
21. De Gaudenzi, J. G., Noe, G., Campo, V. A., Frasch, A. C., Cassola, A. (2011). Gene expression regulation in trypanosomatids. Essays Biochem. 51, 31-46.
22. Delespaux, V., Dinka, H., Masumu, J., Van Den Bossche, P., Geerts, S. (2008). Five-fold increase in *Trypanosoma congolense* isolates resistant to diminazene aceturate over a seven-year period in Eastern Zambia. Drug Resist Updat. 11, 205-209.
23. Delespaux, V., Geysen, D., Van Den Bossche, P., Geerts, S. (2008). Molecular tools for the rapid detection of drug resistance in animal trypanosomes. Trends Parasitol. 24, 236-242.
24. Desquesnes, M., Holzmuller, P., Lai, D. H., Dargantes, A., Lun, Z. R., Jittaplapong, S. (2013). and Surra: A Review and Perspectives on Origin, History, Distribution, Taxonomy, Morphology, Hosts, and Pathogenic Effects. Biomed Res Int. 2013, 194176.
25. Desquesnes, M., Mclaughlin, G., Zoungrana, A., Davila, A. M. (2001). Detection and identification of *Trypanosoma* of African livestock through a single PCR based on internal transcribed spacer 1 of rDNA. Int J Parasitol. 31, 610-614.
26. Di Noia, J. M., D'orso, I., Sanchez, D. O., Frasch, A. C. (2000). AU-rich elements in the 3'-untranslated region of a new mucin-type gene family of *Trypanosoma cruzi* confers mRNA instability and modulates translation efficiency. J Biol Chem. 275, 10218-10227.
27. Evans, D. A., Ellis, D. S., Stamford, S. (1979). Ultrastructural studies of certain aspects of the development of *Trypanosoma congolense* in Glossina morsitans morsitans. J Protozool. 26, 557-563.
28. Eyford, B. A., Sakurai, T., Smith, D., Loveless, B., Hertz-Fowler, C., Donelson, J. E., Inoue, N., Pearson, T. W. (2011). Differential protein expression throughout the life cycle of *Trypanosoma congolense*, a major parasite of cattle in Africa. Mol Biochem Parasitol. 177, 116-125.
29. Fenn, K., Matthews, K. R. (2007). The cell biology of *Trypanosoma brucei* differentiation. Curr Opin Microbiol. 10, 539-546.
30. Friguet, B., Chaffotte, A. F., Djavadi-Ohaniance, L., Goldberg, M. E. (1985). Measurements of the true affinity constant in solution of

antigen-antibody complexes by enzyme-linked immunosorbent assay. J Immunol Methods. 77, 305-319.
31. Furger, A., Schurch, N., Kurath, U., Roditi, I. (1997). Elements in the 3' untranslated region of procyclin mRNA regulate expression in insect forms of *Trypanosoma brucei* by modulating RNA stability and translation. Mol. Cell. Biol. 17, 4372-4380.
32. Geerts, S., Holmes, P. H., Eisler, M. C., Diall, O. (2001). African bovine trypanosomiasis: the problem of drug resistance. Trends Parasitol. 17, 25-28.
33. Haag, J., O'huigin, C., Overath, P. (1998). The molecular phylogeny of trypanosomes: evidence for an early divergence of the Salivaria. Mol Biochem Parasitol. 91, 37-49.
34. Hakimi, H., Nguyen, T. T., Suganuma, K., Masuda-Suganuma, H., Angeles, J. M., Inoue, N., Kawazu, S. (2013). Development of Monoclonal Antibodies That Target 1-Cys Peroxiredoxin and Differentiate *Plasmodium falciparum* from *P. vivax* and *P. knowlesi*. Trop Med Health. 41, 55-59.
35. Hartmann, C., Benz, C., Brems, S., Ellis, L., Luu, V. D., Stewart, M., D'orso, I., Busold, C., Fellenberg, K., Frasch, A. C., Carrington, M., Hoheisel, J., Clayton, C. E. (2007). Small trypanosome RNA-binding proteins TbUBP1 and TbUBP2 influence expression of F-box protein mRNAs in bloodstream trypanosomes. Eukaryot Cell. 6, 1964-1978.
36. Helm, J. R., Hertz-Fowler, C., Aslett, M., Berriman, M., Sanders, M., Quail, M. A., Soares, M. B., Bonaldo, M. F., Sakurai, T., Inoue, N., Donelson, J. E. (2009). Analysis of expressed sequence tags from the four main developmental stages of *Trypanosoma congolense*. Mol Biochem Parasitol. 168, 34-42.
37. Hendry, K. A., Vickerman, K. (1988). The requirement for epimastigote attachment during division and metacyclogenesis in *Trypanosoma congolense.* Parasitol Res. 74, 403-408.
38. Hirumi, H., Hirumi, K. (1991). *In vitro* cultivation of *Trypanosoma congolense* bloodstream forms in the absence of feeder cell layers. Parasitology. 102 Pt 2, 225-236.
39. Hoare, C. A. (1972). The trypanosomes of mammals. A zoological monograph.: Blackwell Scientific Publications.
40. Horn, D., Mcculloch, R. (2010). Molecular mechanisms underlying the control of antigenic variation in African trypanosomes. Curr Opin Microbiol. 13, 700-705.
41. Inoue, N., Otsu, K., Ferraro, D. M., Donelson, J. E. (2002).

Tetracycline-regulated RNA interference in *Trypanosoma congolense*. Mol Biochem Parasitol. 120, 309-313.

42. Jackson, A. P., Berry, A., Aslett, M., Allison, H. C., Burton, P., Vavrova-Anderson, J., Brown, R., Browne, H., Corton, N., Hauser, H., Gamble, J., Gilderthorp, R., Marcello, L., Mcquillan, J., Otto, T. D., Quail, M. A., Sanders, M. J., Van Tonder, A., Ginger, M. L., Field, M. C., Barry, J. D., Hertz-Fowler, C., Berriman, M. (2012). Antigenic diversity is generated by distinct evolutionary mechanisms in African trypanosome species. Proc Natl Acad Sci U S A. 109, 3416-3421.
43. Jackson, A. P., Sanders, M., Berry, A., Mcquillan, J., Aslett, M. A., Quail, M. A., Chukualim, B., Capewell, P., Macleod, A., Melville, S. E., Gibson, W., Barry, J. D., Berriman, M., Hertz-Fowler, C. (2010). The genome sequence of *Trypanosoma brucei gambiense*, causative agent of chronic human african trypanosomiasis. PLoS Negl Trop Dis. 4, e658.
44. Kohler, G., Milstein, C. (1976). Derivation of specific antibody-producing tissue culture and tumor lines by cell fusion. Eur J Immunol. 6, 511-519.
45. Kolev, N. G., Ramey-Butler, K., Cross, G. A., Ullu, E., Tschudi, C. (2012). Developmental progression to infectivity in *Trypanosoma brucei* triggered by an RNA-binding protein. Science. 338, 1352-1353.
46. Kramer, S. (2012). Developmental regulation of gene expression in the absence of transcriptional control: the case of kinetoplastids. Mol Biochem Parasitol. 181, 61-72.
47. Kramer, S., Carrington, M. (2011). *Trans*-acting proteins regulating mRNA maturation, stability and translation in trypanosomatids. Trends Parasitol. 27, 23-30.
48. La Greca, F., Magez, S. (2011). Vaccination against trypanosomiasis: can it be done or is the trypanosome truly the ultimate immune destroyer and escape artist? Hum Vaccin. 7, 1225-1233.
49. Lacount, D. J., Donelson, J. E. (2001). RNA interference in African trypanosomes. Protist. 152, 103-111.
50. Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature. 227, 680-685.
51. Lambrecht, F. L. (1985). Trypanosomes and Hominid Evolution. BioScience. 35, 640-646.
52. Lanham, S. M., Godfrey, D. G. (1970). Isolation of salivarian trypanosomes from man and other mammals using

DEAE-cellulose. Exp Parasitol. 28, 521-534.
53. Li, Z. H., De Gaudenzi, J. G., Alvarez, V. E., Mendiondo, N., Wang, H., Kissinger, J. C., Frasch, A. C., Docampo, R. (2012). A 43-nucleotide U-rich element in 3'-untranslated region of large number of *Trypanosoma cruzi* transcripts is important for mRNA abundance in intracellular amastigotes. J Biol Chem. 287, 19058-19069.
54. Macgregor, P., Matthews, K. R. (2012). Identification of the regulatory elements controlling the transmission stage-specific gene expression of PAD1 in *Trypanosoma brucei*. Nucleic Acids Res. 40, 7705-7717.
55. Maclean, L., Myburgh, E., Rodgers, J., Price, H. P. (2013). Imaging African trypanosomes. Parasite Immunol. 35, 283-294.
56. Magnus, E., Vervoort, T., Van Meirvenne, N. (1978). A card-agglutination test with stained trypanosomes (C.A.T.T.) for the serological diagnosis of *T. b. gambiense* trypanosomiasis. Ann Soc Belg Med Trop. 58, 169-176.
57. Mani, J., Guttinger, A., Schimanski, B., Heller, M., Acosta-Serrano, A., Pescher, P., Spath, G., Roditi, I. (2011). Alba-domain proteins of *Trypanosoma brucei* are cytoplasmic RNA-binding proteins that interact with the translation machinery. PLoS One. 6, e22463.
58. Mao, Y., Najafabadi, H. S., Salavati, R. (2009). Genome-wide computational identification of functional RNA elements in *Trypanosoma brucei*. BMC Genomics. 10, 355.
59. Martinez-Calvillo, S., Yan, S., Nguyen, D., Fox, M., Stuart, K., Myler, P. J. (2003). Transcription of *Leishmania major* Friedlin chromosome 1 initiates in both directions within a single region. Mol Cell. 11, 1291-1299.
60. Merino, O., Alberdi, P., Perez De La Lastra, J. M., De La Fuente, J. (2013). Tick vaccines and the control of tick-borne pathogens. Front Cell Infect Microbiol. 3, 30.
61. Mulenga, C., Mhlanga, J. D., Kristensson, K., Robertson, B. (2001). *Trypanosoma brucei brucei* crosses the blood-brain barrier while tight junction proteins are preserved in a rat chronic disease model. Neuropathol Appl Neurobiol. 27, 77-85.
62. Naessens, J. (2006). Bovine trypanotolerance: A natural ability to prevent severe anaemia and haemophagocytic syndrome? Int J Parasitol. 36, 521-528.
63. Njiru, Z. K., Mikosza, A. S., Armstrong, T., Enyaru, J. C., Ndung'u, J. M., Thompson, A. R. (2008). Loop-mediated isothermal

amplification (LAMP) method for rapid detection of *Trypanosoma brucei rhodesiense*. PLoS Negl Trop Dis. 2, e147.

64. Njiru, Z. K., Ouma, J. O., Bateta, R., Njeru, S. E., Ndungu, K., Gitonga, P. K., Guya, S., Traub, R. (2011). Loop-mediated isothermal amplification test for *Trypanosoma vivax* based on satellite repeat DNA. Vet Parasitol. 180, 358-362.
65. Noe, G., De Gaudenzi, J. G., Frasch, A. C. (2008). Functionally related transcripts have common RNA motifs for specific RNA-binding proteins in trypanosomes. BMC Mol Biol. 9, 107.
66. Okiria, R., Okuna, N. M., Magona, J. W., Mayende, J. S. (2002). Sustainability of tsetse control by subsequent treatment of 10% of a previously treated Ugandan cattle population with 1% w/v deltamethrin. Trop Anim Health Prod. 34, 105-114.
67. Peacock, L., Cook, S., Ferris, V., Bailey, M., Gibson, W. (2012). The life cycle of *Trypanosoma* (Nannomonas) *congolense* in the tsetse fly. Parasit Vectors. 5, 109.
68. Penchenier, L., Simo, G., Grebaut, P., Nkinin, S., Laveissiere, C., Herder, S. (2000). Diagnosis of human trypanosomiasis, due to *Trypanosoma brucei gambiense* in central Africa, by the polymerase chain reaction. Trans R Soc Trop Med Hyg. 94, 392-394.
69. Pichon, X., Wilson, L. A., Stoneley, M., Bastide, A., King, H. A., Somers, J., Willis, A. E. (2012). RNA binding protein/RNA element interactions and the control of translation. Curr Protein Pept Sci. 13, 294-304.
70. Pinder, M., Authie, E. (1984). The appearance of isometamidium resistant *Trypanosoma congolense* in West Africa. Acta Trop. 41, 247-252.
71. Priotto, G., Kasparian, S., Mutombo, W., Ngouama, D., Ghorashian, S., Arnold, U., Ghabri, S., Baudin, E., Buard, V., Kazadi-Kyanza, S., Ilunga, M., Mutangala, W., Pohlig, G., Schmid, C., Karunakara, U., Torreele, E., Kande, V. (2009). Nifurtimox-eflornithine combination therapy for second-stage African *Trypanosoma brucei gambiense* trypanosomiasis: a multicentre, randomised, phase III, non-inferiority trial. Lancet. 374, 56-64.
72. Rodgers, J., Mccabe, C., Gettinby, G., Bradley, B., Condon, B., Kennedy, P. G. (2011). Magnetic resonance imaging to assess blood-brain barrier damage in murine trypanosomiasis. Am J Trop Med Hyg. 84, 344-350.
73. Roditi, I., Lehane, M. J. (2008). Interactions between trypanosomes and

tsetse flies. Curr Opin Microbiol. 11, 345-351.
74. Rotureau, B., Subota, I., Buisson, J., Bastin, P. (2012). A new asymmetric division contributes to the continuous production of infective trypanosomes in the tsetse fly. Development. 139, 1842-1850.
75. Ruepp, S., Furger, A., Kurath, U., Renggli, C. K., Hemphill, A., Brun, R., Roditi, I. (1997). Survival of *Trypanosoma brucei* in the tsetse fly is enhanced by the expression of specific forms of procyclin. The Journal of Cell Biology. 137, 1369-1379.
76. Sakurai, T., Sugimoto, C., Inoue, N. (2008). Identification and molecular characterization of a novel stage-specific surface protein of *Trypanosoma congolense* epimastigotes. Mol Biochem Parasitol. 161, 1-11.
77. Sakurai, T., Tanaka, M., Kawazu, S., Inoue, N. (2009). Establishment of an *in vitro* transgene expression system in epimastigotes of *Trypanosoma congolense*. Parasitol Int. 58, 110-113.
78. Shulman, M., Wilde, C. D., Kohler, G. (1978). A better cell line for making hybridomas secreting specific antibodies. Nature. 276, 269-270.
79. Simarro, P. P., Diarra, A., Ruiz Postigo, J. A., Franco, J. R., Jannin, J. G. (2011). The human African trypanosomiasis control and surveillance programme of the World Health Organization 2000-2009: the way forward. PLoS Negl Trop Dis. 5, e1007.
80. Simo, G., Herder, S., Cuny, G., Hoheisel, J. (2010). Identification of subspecies specific genes differentially expressed in procyclic forms of *Trypanosoma brucei* subspecies. Infect Genet Evol. 10, 229-237.
81. Stevenes, J. R., Brisse, S. (2004). The Trypanosomiases. Wallingford, Oxfordshire OX10 8DE, UK: CABI Publishing.
82. Steverding, D. (2008). The history of African trypanosomiasis. Parasit Vectors. 1, 3.
83. Subota, I., Rotureau, B., Blisnick, T., Ngwabyt, S., Durand-Dubief, M., Engstler, M., Bastin, P. (2011). ALBA proteins are stage regulated during trypanosome development in the tsetse fly and participate in differentiation. Mol Biol Cell. 22, 4205-4219.
84. Thekisoe, O. M., Omolo, J. D., Swai, E. S., Hayashida, K., Zhang, J., Sugimoto, C., Inoue, N. (2007). Preliminary application and evaluation of loop-mediated isothermal amplification (LAMP) for detection of bovine theileriosis and trypanosomosis in Tanzania. Onderstepoort J Vet Res. 74, 339-342.

85. Tuntasuvan, D., Sarataphan, N., Nishikawa, H. (1997). Cerebral trypanosomiasis in native cattle. Vet Parasitol. 73, 357-363.
86. Van Den Abbeele, J., Claes, Y., Van Bockstaele, D., Le Ray, D., Coosemans, M. (1999). *Trypanosoma brucei* spp. development in the tsetse fly: characterization of the post-mesocyclic stages in the foregut and proboscis. Parasitology. 118 (Pt 5), 469-478.
87. Vassella, E., Reuner, B., Yutzy, B., Boshart, M. (1997). Differentiation of African trypanosomes is controlled by a density sensing mechanism which signals cell cycle arrest via the cAMP pathway. J Cell Sci. 110 (Pt 21), 2661-2671.
88. Vickerman, K. (1985). Developmental cycles and biology of pathogenic trypanosomes. Br Med Bull. 41, 105-114.
89. Vickerman, K., Tetley, L., Hendry, K. A., Turner, C. M. (1988). Biology of African trypanosomes in the tsetse fly. Biol Cell. 64, 109-119.
90. Vreysen, M. J., Saleh, K. M., Ali, M. Y., Abdulla, A. M., Zhu, Z. R., Juma, K. G., Dyck, V. A., Msangi, A. R., Mkonyi, P. A., Feldmann, H. U. (2000). *Glossina austeni* (Diptera: *Glossinidae*) eradicated on the island of Unguja, Zanzibar, using the sterile insect technique. J Econ Entomol. 93, 123-135.
91. Vreysen, M. J., Seck, M. T., Sall, B., Bouyer, J. (2013). Tsetse flies: their biology and control using area-wide integrated pest management approaches. J Invertebr Pathol. 112 Suppl, S15-25.
92. Walrad, P., Paterou, A., Acosta-Serrano, A., Matthews, K. R. (2009). Differential trypanosome surface coat regulation by a CCCH protein that co-associates with procyclin mRNA *cis*-elements. PLoS Pathog. 5, e1000317.
93. Wastling, S. L., Welburn, S. C. (2011). Diagnosis of human sleeping sickness: sense and sensitivity. Trends Parasitol. 27, 394-402.
94. Wellde, B. T., Reardon, M. J., Kovatch, R. M., Chumo, D. A., Williams, J. S., Boyce, W. L., Hockmeyer, W. T., Wykoff, D. E. (1989). Experimental infection of cattle with *Trypanosoma brucei rhodesiense*. Ann Trop Med Parasitol. 83 Suppl 1, 133-150.
95. Wirtz, E., Clayton, C. (1995). Inducible gene expression in trypanosomes mediated by a prokaryotic repressor. Science. 268, 1179-1183.
96. Wirtz, E., Leal, S., Ochatt, C., Cross, G. A. (1999). A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in *Trypanosoma brucei.* Mol Biochem Parasitol. 99, 89-101.