



**ÉTUDE DES ÉLÉMENTS RÉGULATEURS *CIS* ET *TRANS*  
IMPLIQUÉS DANS LA STABILITÉ DU TRANSCRIT DE  
L'AMASTINE AU STADE INTRACELLULAIRE CHEZ  
*LEISHMANIA*.**

Thèse

Aurélien Dupé

**Doctorat en Biologie cellulaire et moléculaire**  
Philosophiae Doctor (Ph.D.)

Québec, Canada

© Aurélien Dupé, 2013



## Résumé

Le genre *Leishmania* regroupe des parasites protozoaires transmis par piqûre d'un insecte vecteur et qui sont responsables des leishmanioses. Le cycle de *Leishmania* alterne entre promastigotes dans l'appareil digestif de l'insecte et amastigotes dans les phagolysosomes des macrophages d'un hôte mammifère. Les delta-amastines sont une famille de protéines membranaires qui jouent potentiellement un rôle dans la virulence. L'expression exclusivement au stade intracellulaire de l'un de ces gènes est permise par une accumulation préférentielle de l'ARNm et la stimulation de la traduction, toutes deux chez les amastigotes. L'objectif de cette thèse est de caractériser les mécanismes permettant l'expression différentielle de l'ARNm de la delta-amastine. Ces organismes ont divergé rapidement des autres eucaryotes, ce qui engendre plusieurs différences fonctionnelles, dont notamment l'absence de régulation transcriptionnelle. Notre hypothèse est que la présence d'une région riche en uridines (URE) dans l'extrémité 3' non traduite (3'UTR) du transcrit peut être impliquée dans la dégradation de l'ARNm. Nous démontrons que le URE est responsable d'une dégradation du transcrit au stade promastigote, par un phénomène indépendant de la déadénylation. Nous avons identifié une protéine à domaine Alba, *LiAlba20*, liant l'ARNm de la delta-amastine dans une région proche de l'URE. La suppression de cette protéine réduit l'accumulation du transcrit au stade amastigote. Ainsi, deux mécanismes complémentaires sont responsables de l'expression différentielle de ce transcrit. Le génome de *Leishmania* code pour une seconde protéine à domaine Alba, *LiAlba13*. Ces protéines interagissent ensemble, mais *LiAlba13* n'affecte pas l'abondance de l'ARNm de la delta-amastine. Les protéines Alba ont une évolution exceptionnelle puisqu'elles stabilisent l'ADN chez les *Archaea*, et sont retrouvées dans les complexes RNase P/MRP chez les eucaryotes supérieurs. Nos résultats montrent qu'elles régulent l'expression de protéines spécifiques du stade amastigote, ce qui concorde avec les récents travaux chez d'autres parasites protozoaires. Ces protéines sont cytoplasmiques dans les deux stades de développement. Cependant, pendant la différenciation, elles s'accumulent dans le flagelle et le nucléole, respectivement décrits comme senseur et coordinateur de la réponse au stress. Nos travaux suggèrent donc l'implication du flagelle et du nucléole dans la coordination de la régulation de facteurs de virulence pendant la différenciation du parasite.



## Abstract

The *Leishmania* genus encompasses protozoan parasites which are transmitted through the bite of an insect vector and are responsible for leishmaniasis. The *Leishmania* life cycle alternates between promastigote forms within the gut of the insect vector and amastigotes which multiply in the phagolysosomal vacuoles of the mammalian host's macrophages. Delta-amastins are part of a multigenic family of membrane proteins that potentially act in parasite virulence. One of the delta-amastin's exclusive expression in the intracellular stage is mediated by mRNA accumulation and translation stimulation, both taking place in the amastigote stage. The aim of this thesis is to characterize the mechanisms implicated in the differential expression of delta-amastin mRNA. *Leishmania* splits early in evolution from other eukaryotes and this split correlates with many functional differences, including the absence of transcriptional control of gene expression. Our hypothesis is that the presence of a uridine-rich element (URE) within the 3' untranslated region (3'UTR) of the transcript might be implicated in an mRNA decay mechanism. We reveal that the URE is responsible for a fast mRNA decay only in the promastigote stage, performed by an unusual deadenylation-independent pathway. We next identified an Alba domain protein, *LiAlba20*, which binds to the delta-amastin mRNA in a region flanking the URE. Depletion of this protein leads to a reduced mRNA accumulation in the amastigote stage specifically. Therefore, we identified two complementary mechanisms taking part in the transcript's differential expression. The *Leishmania* genome encodes a second Alba domain protein, *LiAlba13*. These proteins interact together, but *LiAlba13* does not affect the delta-amastin mRNA level during the parasite life cycle. Alba domain proteins have a remarkable evolution, being involved in DNA stabilization in *Archaea* and subunits of the RNAses P/MRP complexes in higher eukaryotes. In addition, our data show that these proteins regulate stage-specific protein expression, which is in agreement with recent works in other protozoan parasites. Alba domain proteins are constitutively expressed in the cytoplasm of both parasite life cycle stages. Nevertheless, during the differentiation, those proteins accumulate in flagellar and nucleolus compartments, respectively described as sensor and stress response coordinators in higher eukaryotes. Our work suggests that the flagellum is implicated in the coordination of stage-specific transcript expression in response to stress in *Leishmania*.



# Table des matières

Résumé.....	iii
Abstract.....	v
Table des matières .....	vii
Liste des tableaux.....	xi
Liste des figures.....	xiii
Liste des abréviations .....	xv
Remerciements .....	xvii
Avant-propos.....	xix
<b>Chapitre 1 : <i>Leishmania</i> et les leishmanioses. ....</b>	<b>1</b>
1.1. Le genre <i>Leishmania</i> .....	1
1.2. Insecte vecteur et répartition géographique. ....	2
1.3. Manifestations cliniques des leishmanioses. ....	4
1.3.1. Les leishmanioses cutanées. ....	5
1.3.2. La leishmaniose viscérale. ....	6
1.3.3. Cas particulier : La co-infection <i>Leishmania</i> / VIH. ....	8
1.4. Diagnostics.....	8
1.5. Préventions et traitements.....	9
<b>Chapitre 2 : Cycle de vie de <i>Leishmania</i> et stratégies de survie parasitaire. ....</b>	<b>13</b>
2.1. Vue globale du cycle de vie.....	14
2.2. Au sein de l'insecte vecteur.....	15
2.3. Dissémination et survie des parasites dans l'hôte mammifère. ....	18
2.3.1. Inoculation des parasites.....	18
2.3.2. La résistance au complément.....	18
2.3.3. Dissémination des parasites dans l'organisme infecté.....	19
2.3.4. Interaction de <i>Leishmania</i> avec les macrophages.....	22
2.3.4.1. Les différents types de macrophage.....	22
2.3.4.2. Phagocytose par les macrophages et fusion phagosomes – lysosomes. ....	24
2.3.4.3. Les stress environnementaux et la différenciation du parasite en sa forme amastigote. ....	25
2.3.5. Implication de virus de <i>Leishmania</i> dans la sévérité de l'infection.....	27
<b>Chapitre 3 : Origine et organisation cellulaire des <i>Kinetoplastidae</i>. ....</b>	<b>29</b>
3.1. Origine des <i>Kinetoplastidae</i> .....	32
3.2. Organisation cellulaire des <i>Trypanosomatidae</i> .....	34
3.2.1. Organisation générale et organites.....	34
3.2.2. Le noyau et le nucléole.....	35
3.2.3. La mitochondrie.....	37
3.2.4. Flagelle et poche flagellaire.....	39
3.2.4.1. La poche flagellaire .....	39
3.2.4.2. La membrane flagellaire .....	39
3.2.4.3. Matrice et cytosquelette flagellaire.....	40

<b>Chapitre 4 : Organisation du génome des <i>Trypanosomatidae</i> et régulation de l'expression génique.</b> .....	<b>45</b>
4.1. Le génome des <i>Trypanosomatidae</i> .....	45
4.1.1. Organisation du génome et transcription polycistronique .....	45
4.1.2. Comparaison du contenu du génome des <i>Trypanosomatidae</i> .....	48
4.1.3. Plasticité génomique. ....	49
4.1.3.1. Amplification génique et aneuploïdie. ....	49
4.1.3.2. Les éléments transposables. ....	51
4.1.4. Production d'ARNm matures par <i>trans</i> -splicing et polyadénylation. ....	54
4.2. Régulation de l'expression génique au niveau de l'ARNm. ....	57
4.2.1. Régulation de l'épissage des transcrits. ....	57
4.2.2. Exportation des ARNm vers le cytoplasme. ....	59
4.2.3. La dégradation des ARNm. ....	60
4.2.3.1. Mécanismes de dégradation des ARNm connus chez les eucaryotes. ....	61
4.2.3.2. Les complexes enzymatiques impliqués dans la dégradation des ARNm. .	62
4.2.3.3. La surveillance des ARNm. ....	66
4.2.4. Les granules à ARNm. ....	67
4.2.5. Régulation de l'expression génique par les petits ARN. ....	70
4.2.6. La régulation des ARNm par les éléments riches en AU. ....	72
4.2.7. Participation des rétroposons inactifs dans la régulation des ARNm endogènes. ....	75
4.2.8. Les interactions ARNm–protéines dans la régulation de l'expression génique chez les <i>Trypanosomatidae</i> . ....	75
4.3. Régulation traductionnelle et post-traductionnelle. ....	76
4.3.1. La traduction. ....	76
4.3.1.1. Les facteurs de liaison de la coiffe. ....	80
4.3.1.2. Les protéines de liaison aux queues poly(A). ....	82
4.3.1.3. De multiples possibilités de formation du complexe d'initiation de la traduction. ....	83
4.3.1.4. Réponse au stress et différenciation en amastigote par phosphorylation d'eIF2 $\alpha$ . ....	83
4.3.2. Les ARNr antisens sont régulateurs de la traduction. ....	85
4.3.3. Modifications post-traductionnelles. ....	86
4.3.4. Dégradation des protéines. ....	88
<b>Chapitre 5 : Les amastines.</b> .....	<b>91</b>
5.1. Structure et fonctions potentielles des amastines. ....	92
5.2. Distribution et classification des amastines. ....	93
5.3. Les amastines de <i>Trypanosoma cruzi</i> . ....	98
5.4. Les amastines chez <i>Leishmania</i> . ....	100
<b>Chapitre 6 : Les protéines à domaine Alba</b> .....	<b>103</b>
6.1. Les protéines Alba chez les <i>Archaea</i> . ....	104
6.2. Les protéines à domaine Alba chez les parasites protozoaires. ....	106
6.2.1. <i>Toxoplasma gondii</i> (agent de la toxoplasmose) .....	106
6.2.2. <i>Plasmodium</i> spp (agent de la malaria). ....	109
6.2.3. <i>Trypanosoma</i> spp. ....	111
6.2.4. <i>Leishmania</i> spp. ....	113

6.3. Les protéines à domaine Alba chez les eucaryotes supérieurs comme sous unités des RNase P / MRP.....	113
<b>Chapitre 7 : Hypothèses et objectifs.....</b>	<b>117</b>
<b>Chapitre 8 : Dégradation stade spécifique des ARNm par un mécanisme indépendant de la déadénylation chez <i>Leishmania</i>. .....</b>	<b>121</b>
Résumé.....	121
Article .....	122
Abstract.....	122
Introduction.....	123
Materials and Methods.....	125
Results.....	127
Discussion.....	132
Acknowledgments .....	136
References.....	137
Figure Legends .....	141
Tables.....	145
Figures .....	146
Supplementary Figures .....	153
<b>Chapitre 9 : Une protéine à domaine Alba est impliquée dans la stabilité stade spécifique des transcrits des Amastines chez <i>Leishmania infantum</i>.....</b>	<b>157</b>
Résumé.....	157
Article .....	158
Summary .....	158
Introduction.....	159
Results.....	161
Discussion.....	166
Experimental Procedures .....	169
Acknowledgements.....	173
References.....	174
Figure Legends .....	179
Tables.....	183
Figures .....	184
Supporting Information.....	191
<b>Chapitre 10 : Analyses structurelle et fonctionnelle des protéines à domaine Alba chez <i>Leishmania</i>.....</b>	<b>203</b>
Résumé.....	203
Article .....	204
Abstract.....	204
Introduction.....	205
Results.....	207
Discussion.....	216
Materials and Methods.....	220
Acknowledgements.....	224
References.....	225
Figure Legends .....	229

Tables .....	232
Figures .....	233
Supplementary data .....	239
<b>Chapitre 11 : Discussion générale.....</b>	<b>245</b>
11.1. Les mécanismes impliqués dans la régulation de l'expression de l'amastine spécifiquement au stade amastigote. ....	247
11.1.1. Implication du <i>LiSIDER1</i> et d'une seconde séquence de 100 nt dans la traduction préférentielle au stade amastigote. ....	247
11.1.2. L'élément riche en uridines dégrade l'ARNm de l'amastine spécifiquement chez les promastigotes. ....	249
11.1.3. Recherche de protéines liant l'URE du transcrit de l'amastine. ....	250
11.1.4. L'implication des protéines à domaine Alba dans la régulation du transcrit de l'amastine. ....	252
11.1.4.1. La protéine <i>Rpp25-like</i> : <i>LiAlba20</i> .....	252
11.1.4.2. La protéine <i>Rpp20-like</i> : <i>LiAlba13</i> .....	255
11.1.5. Autres éléments <i>cis</i> communs aux 3'UTR des $\delta$ -amastines.....	255
11.2. La famille des protéines Alba chez les parasites protozoaires .....	257
11.3. La localisation des protéines Alba lors de la différenciation du parasite et en réponse au stress.....	260
11.3.1. Localisation flagellaire.....	261
11.3.2. Localisation nucléolaire .....	262
<b>Conclusions.....</b>	<b>265</b>
<b>Bibliographie.....</b>	<b>267</b>
<b>Annexe 1.....</b>	<b>295</b>
<b>Annexe 2.....</b>	<b>297</b>
<b>Annexe 3.....</b>	<b>299</b>
<b>Annexe 4.....</b>	<b>301</b>
<b>Annexe 5.....</b>	<b>303</b>
<b>Annexe 6.....</b>	<b>305</b>
<b>Annexe 7.....</b>	<b>307</b>

## Liste des tableaux

Tableau 1 : Type de protéines retrouvées dans la surface et la matrice flagellaire chez la forme sanguine de <i>T. brucei</i> . .....	41
Tableau 2 : Principaux composants des complexes impliqués dans la dégradation des ARNm chez les eucaryotes et comparaison avec ceux identifiés chez les <i>Trypanosomatidae</i> . .....	64
Tableau 3 : Contenu des <i>P-bodies</i> et granules de stress identifiés chez les trypanosomes. 69	
Tableau 4 : Résumé des principales interactions ARNm-protéine étudiées chez les <i>Kinetoplastidae</i> . .....	77
Tableau 5 : Nombre de gènes codant pour des protéines Amastines retrouvés par génome de <i>Trypanosomatidae</i> . .....	94
Tableau 6 : Les protéines à domaine Alba chez les <i>Apicomplexa</i> et les <i>Kinetoplastidae</i> . 107	



# Liste des figures

Figure 1-1 : Classification du genre <i>Leishmania</i> .....	2
Figure 1-2 : Phlébotome et répartition géographique des leishmanioses. ....	3
Figure 1-3 : Les trois principales formes de leishmanioses. ....	5
Figure 2-1 : Cycle de vie des parasites <i>Leishmania</i> .....	14
Figure 2-2 : Cycle de vie de <i>Leishmania</i> de type suprapylarien au sein de l'insecte vecteur. .....	15
Figure 2-3 : Type des cellules infectées et dissémination de <i>Leishmania</i> dans l'organisme. .....	20
Figure 2-4 : Susceptibilité au parasitisme par <i>Leishmania</i> des macrophages en fonction de la réponse immunitaire. ....	23
Figure 3-1 : Comparaison des cycles parasitaires des <i>Kinetoplastidae</i> infectieux <i>T. brucei</i> , <i>T. cruzi</i> et <i>Leishmania</i> spp. ....	30
Figure 3-2 : Relations évolutives des <i>Kinetoplastidae</i> . ....	32
Figure 3-3 : Schéma général de la structure cellulaire des <i>Trypanosomatidae</i> . ....	35
Figure 3-4 : Structure du flagelle de <i>Leishmania</i> chez les promastigotes et les amastigotes. .....	42
Figure 4-1 : Organisation des gènes sur le chromosome 1 de <i>Leishmania major</i> . ....	46
Figure 4-2 : Diagramme présentant la spécificité des gènes en fonction des espèces de <i>Leishmania</i> .....	48
Figure 4-3 : Différence de ploïdie observée pour chaque chromosome en fonction des souches de <i>Leishmania</i> spp. ....	50
Figure 4-4 : Types de rétroposons retrouvés chez les <i>Trypanosomatidae</i> . ....	52
Figure 4-5 : Maturation des ARNm chez les <i>Trypanosomatidae</i> par épissage en <i>trans</i> et polyadénylation. ....	55
Figure 4-6 : Mécanismes de dégradation des ARNm dans les cellules eucaryotes. ....	61
Figure 4-7 : Signalisation impliquée dans la régulation de la dégradation des ARNm contenant des ARE. ....	73
Figure 5-1 : Topologie prédite de la structure des protéines amastines. ....	92
Figure 5-2 : Arbre phylogénique "maximum likelihood" des amastines de 7 espèces de <i>Trypanosomatidae</i> . ....	96
Figure 6-1 : Arbre phylogénique de type <i>maximum-likelihood</i> de la super famille Alba des <i>Archaea</i> jusqu'aux eucaryotes supérieurs. ....	103
Figure 6-2 : Composition des principaux complexes des RNAses P et MRP humaines. ...	115
Figure 11-1 : Représentation du transcrit de la $\delta$ -amastine LinJ.34.1010 et des facteurs <i>cis</i> et <i>trans</i> impliqués dans la régulation de l'expression génique. ....	247
Figure 11-2 : Modèle mécanistique suggéré de l'expression de l'amastine LinJ.34.1010 en fonction du stade de vie chez <i>Leishmania</i> . ....	254



## Liste des abréviations

ADNc:	ADN complémentaire
ADNk:	ADN kinétoplastique
ARE:	Élément riche en bases adénosines et uridines
ARNm:	ARN messenger
ARNnc:	ARN non codant
ARNP:	ARN polymérase
ARNr:	ARN ribosomal
ARNras :	ARN ribosomal antisens
ARNrs :	ARN ribosomal sens
ARNsn:	Petit ARN nucléaire
ARNt:	ARN de transfert
AUBP:	Protéine de liaison aux ARE
AUF:	Facteur liant les éléments riches en AU
BRF:	<i>butyrate-regulated factor</i>
CBP:	Protéine de liaison à la coiffe
CPB:	Cystéine peptidase B
CR :	Récepteur du complément
CRE:	Élément rétrotransposable de <i>Crithidia</i>
DIRE:	<i>Degenerate ingi/L1Tc-Related Elements</i>
EDC:	<i>Enhancer of decapping</i>
eIF:	Facteur d'initiation de la traduction des eucaryotes
EJC:	Complexe de jonction des exons
ELAVL1:	<i>embryonic lethal abnormal vision-like 1</i>
GP63 :	glycoprotéine de 63 kDa
GPI:	glycosylphosphatidylinositol
GT:	Transporteur de glucose
HSP:	Protéine de choc thermique
IFN $\gamma$ :	Interféron $\gamma$
IFT:	Transport intraflagellaire
IL :	Interleukine
IRES:	Site interne d'entrée des ribosomes
L1Tc:	Élément LINE 1 de <i>Trypanosoma cruzi</i>
LINE:	<i>long interspersed repetitive elements</i>
LPG :	lipophosphoglycan
LRV:	Virus à ARN de <i>Leishmania</i>
LSU	Grande sous unité
LTR:	<i>Long terminal repeat</i>
MAPK:	<i>Mitogen-activated protein kinase</i>
miARN:	microARN
miRISC:	<i>miRNA-induced silencing complex</i>
mRNP:	complexes ribonucléoprotéiques
NGD:	<i>no-go mRNA decay</i>
NMD:	<i>nonsense-mediated mRNA decay</i>
NSD:	<i>non-stop mRNA decay</i>

nt :	nucléotide
ORF:	Cadre de lecture ( <i>Open reading frame</i> )
PABP:	Protéine de liaison aux queues polyadénylées
PAN:	<i>PABP-dependent poly(A) nuclease</i>
PARN:	<i>Poly(A)-specific ribonuclease</i>
pb :	paire de base
PFR:	<i>Paraflagellar rod</i>
piRNA:	ARN interagissant avec PIWI
PPG :	protéophosphoglycan
PRMT:	<i>protein arginine methyltransferase</i>
PSG :	<i>promastigote secretory gel</i>
PTB:	<i>polypyrimidine tract-binding protein</i>
RIME:	<i>Ribosomal Mobile Element</i>
RNAi:	Interférence à l'ARN
RPP :	Protéine de la RNase P
RPR :	ARN structurel de la RNase P
SIDER:	<i>Short Interspersed degenerated retroposon</i>
siRNA:	Petit ARN interférant
SL-RNA:	<i>Splice leader RNA</i>
SMP:	<i>Small myristoylated protein</i>
SSU:	Petite sous unité
TIA-1:	<i>T-cell intracellular antigen 1</i>
TREX:	Complexe de Transcription / Export
TTP:	tristetraproline
UBP:	Protéine de liaison aux uridines
URE:	Élément riche en uridine
UV :	Ultraviolet
VIH:	virus de l'immunodéficience humaine
VSG:	glycoprotéine de surface variable

## Remerciements

Après ces quelques années de thèse au Centre de recherche en Infectiologie, mes travaux touchent à leur fin. Ceci n'aurait pas été possible sans le soutien et l'expertise de bien des personnes et organismes que je tiens à citer ici.

Tout d'abord, je remercie ma directrice de thèse, la Dre Barbara Papadopoulou, de m'avoir encadré pendant cette période. Sa disponibilité et le partage de son expertise scientifique ont été très appréciés pour faire progresser puis aboutir mes travaux. Je lui suis particulièrement reconnaissant de m'avoir soutenu par son écoute, sa patience et sa compréhension dans les diverses difficultés professionnelles ou personnelles que j'ai rencontrées.

Ces travaux n'auraient pas été possibles sans le soutien financier des Instituts de Recherche en Santé du Canada (IRSC-CIHR) et du Centre de recherche sur les interactions hôte-parasite (CHPI), que je remercie donc vivement.

Mes remerciements s'adressent ensuite aux membres du jury, la Dre Angela Kaysel Cruz (Université de São Paulo, Brésil), le Dr Marc Ouellette, le Dr Dave Richard et la Dre Barbara Papadopoulou, pour avoir effectué la lecture, la critique et la correction de ma thèse. J'ai été honoré de pouvoir défendre et discuter mes résultats de recherche avec des spécialistes de la parasitologie qui ont tous d'excellentes publications scientifiques, notamment dans le domaine de *Leishmania*.

Je remercie chaleureusement tous les membres de l'équipe de recherche du Dre Papadopoulou pour leur aide dans mes travaux, leurs conseils, et leur bonne humeur. Tout d'abord Carole Dumas, ma voisine de paillasse pendant toutes ces années, qui m'a beaucoup aidé à finaliser les expériences ces derniers mois, et le Dr Simon Haile, avec qui j'ai travaillé sur le premier article et dont les connaissances en parasitologie et en techniques sur l'ARN ont donné une nouvelle dimension à nos recherches. Je remercie aussi le Dr Prasad Padmanabhan, le Dr Serge Cloutier, le Dr Martin Smith, la Dre Michaela Muller, le Dr Mukesh Samant, Hiva Azizi, la Dre Annie Rochette, la Dre Ouafa Zghidi-Abouزيد, le Dr Francois McNicoll, le Dr Conan Chow et Nathalie Boilard, pour leur aide dans mes

recherches, ma formation et les bons moments que nous avons passés ensemble. Je n'oublie pas le Dr Osvaldo de Melo Neto que je remercie pour le partage de son expertise dans le domaine des protéines chez les parasites, ainsi que les nombreux autres brésiliens qui sont passés par ici et ont stimulé mes réflexions dans mes recherches : la Dre Tamara da Costa Lima, le Dr Nilmar Moretti, Tiago Ferreira, la Dre Tatiany Romão et la Dre Karen Charret.

Le centre de recherche a accueilli bien des personnes durant ma thèse, beaucoup d'entre elles me sont aujourd'hui chères et je ne les oublierai pas. Je n'aurai pas la place de mentionner tout le monde, mais je tiens encore à remercier toute l'équipe de recherche du Dr Marc Ouellette, dont les Drs Rubens do Monte et Nuno Santarém pour leur sympathie et nos diverses discussions scientifiques enrichissantes. Enfin, je n'oublie pas mes amis qui formaient souvent avec moi l'équipe de nuit au laboratoire : le Dr Masayuki Fujino et Anissa Cheikh. Je remercie chaleureusement la Dre Manel Ben Aissa de m'avoir accompagné durant la rédaction de mes travaux et pour ses critiques et suggestions sur le contenu de ma thèse.

Je remercie ma famille pour son soutien, ses encouragements et sa compréhension malgré cette longue période d'absence. Je suis désolé d'avoir été si éloigné d'elle pendant toutes ces années, mais je la sais très heureuse maintenant de voir mes travaux de thèse arriver à leur terme. Il me tarde de la revoir.

Enfin, j'ai noté que le romancier Bernard Werber liste à la fin de ses livres les œuvres qui l'ont inspiré durant l'écriture. La rédaction d'une thèse est très différente de celle d'un roman, mais j'aimerais ici me prêter à l'exercice pour mentionner quelques œuvres qui m'ont inspirées durant ma période de rédaction : la trilogie de polar scientifiques *Le syndrome [E]*, *[GATACA]* et *ATOM[KA]* de Franck Thilliez, *Comment ne plus être en retard* de Jane Burka et Leonora Yuen, l'album *The wall* de Pink Floyd, toute la discographie de Megadeth, et quelques titres de différents groupes tels que Led Zeppelin, Walk Off The Earth ou Lucky Uke.

## Avant-propos

Les travaux présentés dans *Deadenylation-independent stage-specific mRNA degradation in Leishmania* (Chapitre 8) ont été effectués par le Dr Simon Haile et moi-même, sous la supervision de la Dre Barbara Papadopoulou. J'ai réalisé les mutations dans le 3'UTR du transcrite de l'amastine, et toutes les étapes d'intégration génomique et de clonage de *Leishmania*. J'ai ensuite effectué la majorité des Northern blots pour définir le niveau d'expression basal et la mesure de la demi-vie des transcrits de l'ARNm, ainsi que les essais luciférase. Le Dr Simon Haile a effectué des Northern blots supplémentaires pour confirmer mes résultats, les mesures de demi-vie avec ajout de la sinefungin, et toutes les expériences sur la déadénylation et sur la dégradation de 5' vers 3'. Nous avons tous les trois participé à l'analyse des résultats. Le Dr Haile a rédigé l'article, puis a été à l'écoute des critiques du Dre Papadopoulou et moi-même pour les corrections. Cet article a été publié en février 2008 dans la revue *Nucleic Acids Research*.

Dans l'étude *An Alba domain protein is involved in stage-regulated stability of the amastin transcripts in Leishmania infantum* (Chapitre 9), j'ai réalisé la majorité du travail sous la supervision du Dre Barbara Papadopoulou. J'ai adapté le protocole de chromatographie à l'ARN en m'inspirant des recommandations publiées, puis effectué les expériences. J'ai généré les souches mutantes *LiAlba13<sup>-/-</sup>* et *LiAlba20<sup>-/-</sup>* ainsi que les expressions épisomales. Carole Dumas a effectué ici les confirmations par Southern blot. J'ai réalisé tous les Northern blots, Western blot et immunoprécipitations présentés dans l'article ainsi que leurs répétitions. La spectrométrie de masse a été effectuée par le service de protéomique du CHUL. Les expériences de puce à ADN ont été effectuées par la plateforme de transcriptomique du CHUL (Éric Madore). D'autres expériences liées à cette étude et non publiées ont été effectuées par Carole Dumas, telles que l'intégration de constructions luciférase dans les souches *LiAlba13<sup>-/-</sup>* et *LiAlba20<sup>-/-</sup>*, puis des Northern blots et des RT-PCR en temps réel. La Dre Papadopoulou et moi-même avons analysé les résultats, puis j'ai rédigé cet article sous sa supervision. Cet article sera bientôt soumis pour publication dans le journal *Molecular Microbiology*.

Les travaux du chapitre 10 (*Structural and functional analysis of Alba domain proteins in Leishmania*) ont été réalisés par Carole Dumas et moi-même, sous la direction

de la Dre Barbara Papadopoulou. J'ai réalisé les alignements de séquences, les courbes de croissance et les immunoprécipitations. J'ai conçu le protocole de clonage présenté, et Carole Dumas a effectué les constructions, les immunolocalisations, et nous avons partagé les expériences de localisation de protéines fluorescentes. J'ai réalisé les montages des figures, présenté une première version du manuscrit au Dre Barbara Papadopoulou qui a proposé ses corrections. Cet article sera prochainement soumis dans le journal *Plos One*.

Toutes les annexes présentent des travaux supplémentaires que j'ai réalisés, à l'exception des extensions d'amorces de l'Annexe 1, généreusement effectuées par le Dr Prasad Padmanabhan dans le cadre de mes études.

# Chapitre 1 : *Leishmania* et les leishmanioses.

Les travaux présentés dans cette thèse portent sur l'étude de la régulation de l'expression de la protéine amastine chez le parasite protozoaire *Leishmania infantum*, agent de la leishmaniose viscérale. Ce premier chapitre documente les différents types de leishmanioses recensés et introduit les espèces responsables de ces maladies ainsi que leurs insectes vecteurs qui conditionnent la répartition géographique de ces infections. Les quelques diagnostics et traitements disponibles seront énumérés. Le développement d'un vaccin et la recherche de nouveaux traitements pour contourner à la fois les effets secondaires des médicaments actuels et l'émergence de résistance sont une priorité pour le traitement de ces maladies négligées.

## 1.1. Le genre *Leishmania*.

Les parasites *Leishmania* font partie de la famille des *Trypanosomatidae* (voir Chapitre 3) et sont responsables de maladies appelées leishmanioses. Ces infections présentent plusieurs variations épidémiologiques et cliniques principalement en fonction de l'espèce parasitaire, de l'inoculum infectieux et des mécanismes de défense des hôtes infectés. Ces parasites sont transmis par piqûre d'un insecte hématophage et se multiplient dans le phagolysosome des macrophages de l'hôte mammifère infecté (voir Chapitre 2).

Environ 30 espèces de *Leishmania* ont été recensées dont plus des deux tiers sont des pathogènes pour l'homme (Figure 1-1). Le sous-genre *Vianna*, notamment le complexe *L. braziliensis*, est généralement responsable de la forme mucocutanée de la leishmaniose. Le sous-genre *Leishmania* regroupe à la fois les agents de la forme viscérale (complexe *L. donovani*) et de la forme cutanée (complexes *L. tropica*, *L. major*, *L. aethiopica* et *L. mexicana*). Parmi la dizaine d'espèces inoffensives pour l'homme, la plus étudiée est *L. tarentolae*, du sous-genre *Sauroleishmania*, qui infecte le Gekko *Tarentolae annularis*. La répartition géographique de ces espèces différencie celles d'Amérique centrale et du Sud (Nouveau Monde) et celles d'Afrique, d'Europe et d'Asie (Ancien Monde). La dislocation de la Pangée (-200 Millions d'années) semble être à l'origine de la séparation géographique d'espèces primitives de type *Sauroleishmania*, mais les espèces infectant les mammifères

ont utilisé différentes voies de migration jusqu'à l'holocène (-11 500 ans) [1]. Cependant les fortes homologues entre les espèces *L. infantum* (Ancien Monde) et *L. chagasi* (Nouveau Monde) suggèrent également une dissémination effectuée par l'homme lors de la conquête du Nouveau Monde [2].

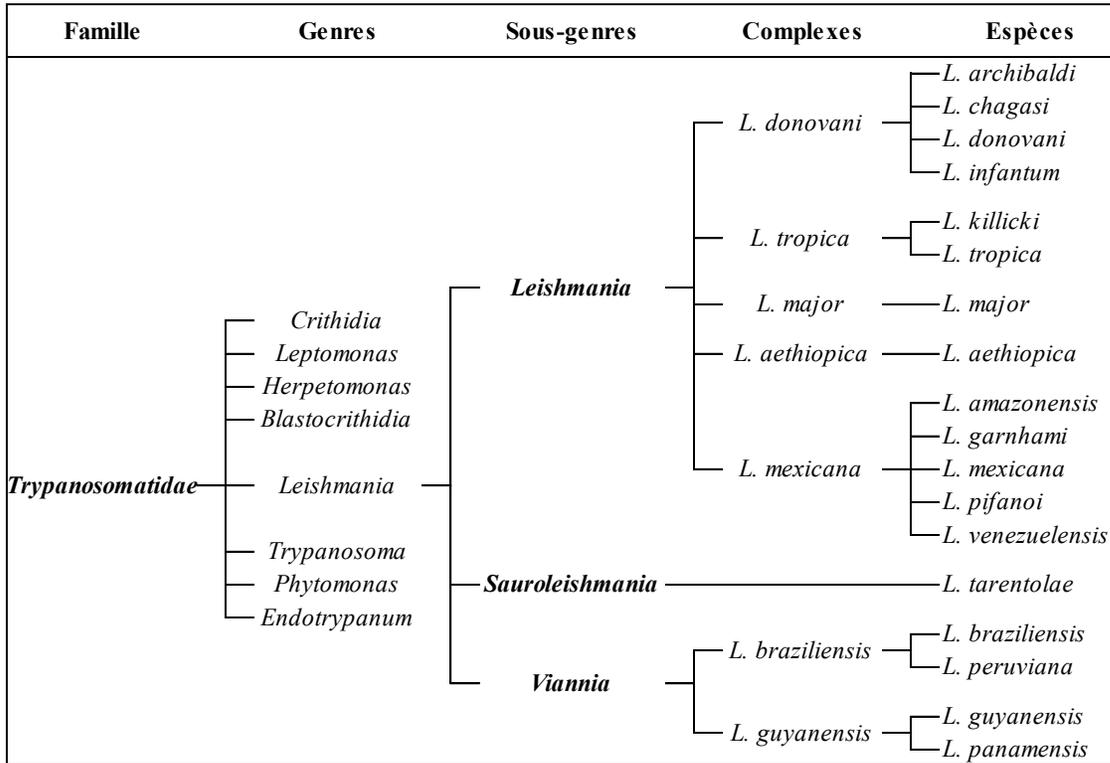


Figure 1-1 : Classification du genre *Leishmania*.

Mise à jour de la classification de l'OMS de 1990 effectuée sur des données immunologiques, génétiques et biochimiques. Adapté de [3].

## 1.2. Insecte vecteur et répartition géographique.

Les seuls vecteurs connus des leishmanioses sont des diptères, de la sous-famille des *Phlebotominae*, communément appelés « Mouche des sables » ou phlébotome (Figure 1-2A). Deux genres issus de cette sous-famille, regroupant 30 espèces, peuvent transmettre les parasites. Habituellement, chaque espèce de phlébotome transmet une seule espèce de *Leishmania* et provoque donc un type particulier de leishmaniose. Le genre *Phlebotomus* sévit dans l'Ancien Monde alors que le genre *Lutzomyia* est plus caractéristique du

Nouveau Monde [4]. La prévalence des infections par *Leishmania* dans les populations de phlébotomes est généralement faible, autour de 0,1% [5].

À l'état adulte, les phlébotomes ne mesurent que 2 à 3 mm de long et sont couverts d'un dense duvet de poils. Au repos, ils gardent leurs ailes sous une forme caractéristique en « V » au dessus de leur dos. Les pièces buccales sont courtes, pointues et orientées vers le bas, et par conséquent bien adaptées à la succion. Les mâles et les femelles adultes obtiennent leur source de carbohydrates de la sève des plantes, mais les femelles sont hémaphages avant la ponte des œufs, période pendant laquelle se fait la transmission des agents pathogènes [6]. Les sites de ponte du vecteur sont des lieux humides et riches en matière organique. Ces insectes sont très sensibles à la déshydratation, la plupart d'entre eux est donc nocturne et les habitations humaines font ainsi partie de leurs refuges connus pendant le jour [7]. Les parasites *Leishmania* circulent dans des zones géographiques où coexistent phlébotomes et réservoirs parasitaires (hôtes mammifères porteurs asymptomatiques des parasites).

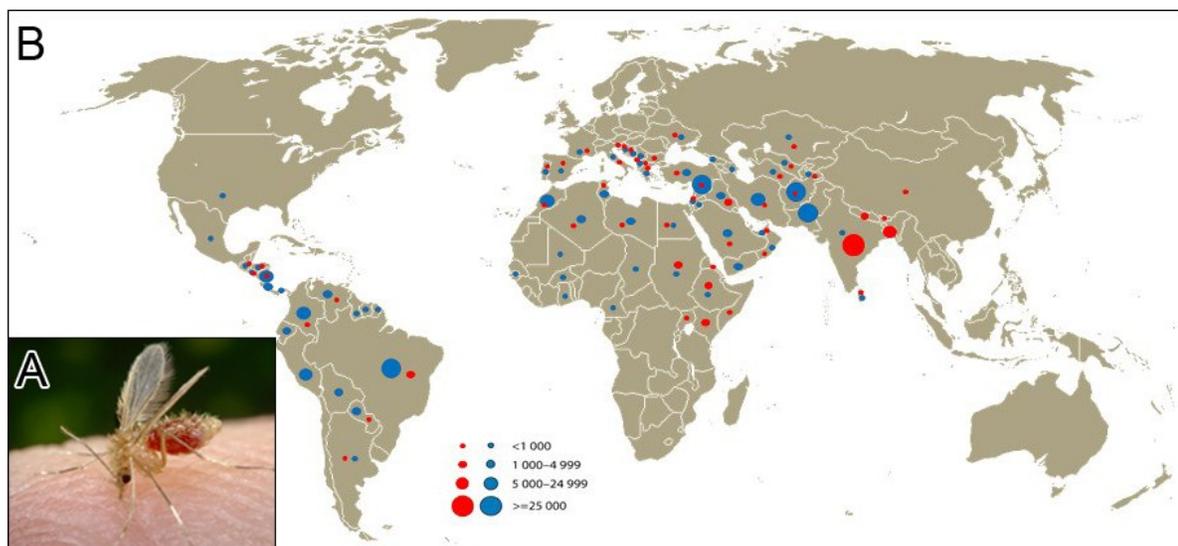


Figure 1-2 : Phlébotome et répartition géographique des leishmanioses.

(A) Mouche des sables *Phlebotomus papatasi* femelle prenant un repas sanguin. (Image adaptée de Public Health Image Library : <http://phil.cdc.gov/phil/home.asp>). (B) Carte du monde indiquant la répartition des cas de leishmaniose viscérale (rouge) et cutanée ou muco-cutanée (bleu). Moyenne des cas notifiés entre 2005 et 2009. Montage effectué par Rubens Do Monte (Ph.D.) à partir des cartes de l'organisation mondiale de la santé.

Les leishmanioses sont endémiques dans 88 pays, sur 4 continents. Elles sont principalement distribuées dans les régions tropicales et subtropicales d'Amérique Centrale, d'Amérique du Sud, d'Afrique et d'Asie et se retrouvent sur le pourtour méditerranéen en Europe (Figure 1-2B). La majorité des leishmanioses sont des zoonoses et les hôtes réservoirs peuvent être différentes espèces de mammifères sauvages ou domestiques. Ces réservoirs sont bien adaptés à *Leishmania*, développant des infections bénignes et persistantes sur plusieurs années. Le chien, hôte réservoir de l'espèce *L. infantum*, fait ici exception en développant une leishmaniose viscérale généralisée et fatale. L'homme est infecté par *L. infantum* et est aussi un hôte réservoir de leishmanioses viscérales ou cutanées causées par *Leishmania tropica*.

Les zones à risque s'étendent à cause de plusieurs facteurs, tels que le mouvement des populations et les changements climatiques, potentiellement responsables d'une modification de la portée et de la densité des insectes vecteurs et des réservoirs parasitaires. Par exemple, en France, 6 départements étaient touchés par la leishmaniose canine en 1988, contre 19 en 2000 [8]. Cette extension a été également décrite en Tunisie et en Israël. Les humains exposés à la piqûre d'un phlébotome sont de plus en plus nombreux. Il est estimé que 350 millions de personnes sont dans des zones à risque de leishmaniose.

### **1.3. Manifestations cliniques des leishmanioses.**

Les manifestations cliniques, la pathogénicité et la virulence des leishmanioses dépendent des espèces infectieuses. Globalement, trois formes prédominent : la forme cutanée, la forme mucocutanée et la leishmaniose viscérale (Figure 1-3). Chez l'homme et les autres hôtes mammifères, *Leishmania* est un parasite intracellulaire qui se multiplie dans les cellules phagocytaires du système immunitaire (voir Chapitre 2). Les symptômes cliniques seront influencés par le type de réponse immunitaire développée par l'hôte, l'âge et la nutrition des personnes infectées [9]. Les mécanismes moléculaires impliqués dans le tropisme des parasites au sein de son hôte mammifère sont encore peu compris. Chaque espèce engendre préférentiellement un type de leishmaniose, mais il existe souvent des contre-exemples qui rendent difficile la compréhension de la dissémination du parasite.

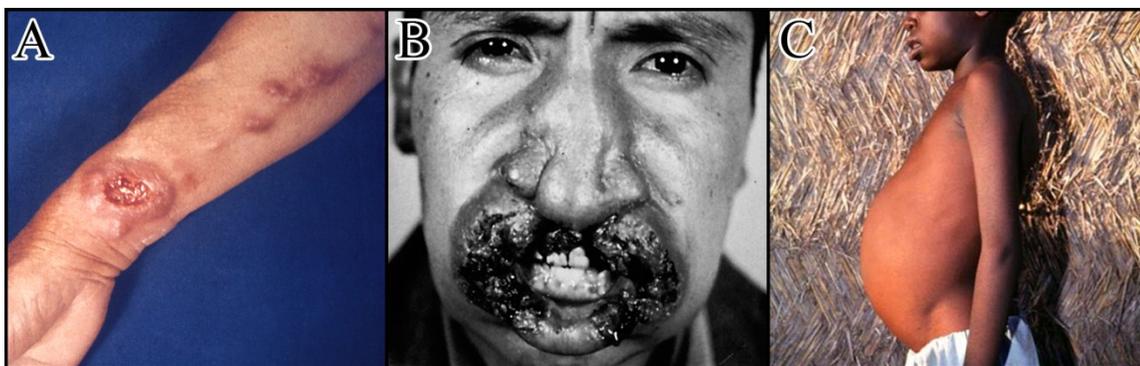


Figure 1-3 : Les trois principales formes de leishmanioses.

(A) Ulcération de la peau par une leishmaniose cutanée (*Images tirées de Public Health Image Library : <http://phil.cdc.gov/phil/home.asp>*). (B) Patient atteint de leishmaniose mucocutanée. La maladie se développe au niveau des muqueuses de la bouche et du nez et provoque une défiguration. (C) Enfant souffrant de leishmaniose viscérale, avec un gonflement caractéristique de l'estomac dû notamment à une splénomégalie.

### 1.3.1. Les leishmanioses cutanées.

Quatre vingt dix pour cent des cas de leishmaniose cutanée sont contractés en Afghanistan, au Pakistan, en Syrie, en Arabie Saoudite, en Algérie, en Iran, au Brésil et au Pérou [10]. Les espèces dites dermatotropes, communément retrouvées dans cette forme de la maladie, sont *L. mexicana*, *L. braziliensis* et *L. panamensis* dans les Amériques, *L. major* et *L. tropica* en Afrique et Eurasie. Les leishmanioses cutanées ne sont pas létales et dans de rares cas, un traitement inadéquat peut engendrer un risque de leishmaniose mucocutanée. Les lésions peuvent causer de graves défigurations et stigmatisation sociale difficile à vivre pour les patients.

Les lésions engendrées par cette maladie peuvent apparaître plusieurs semaines, voir plusieurs mois, après la piqûre d'un phlébotome et se retrouvent fréquemment sur les parties du corps exposées comme le visage, les bras et les jambes. Dans la forme classique des leishmanioses cutanées, chaque lésion correspond à une piqûre indépendante de phlébotome infecté [10]. Les lésions apparaissent initialement comme des petites papules rouges de 5 à 10 mm. Selon l'espèce infectieuse, ces papules peuvent se résorber d'elles-mêmes ou se développer soit en nodules érythémateux, soit en ulcères non douloureux à périphérie bombée. Ces altérations cutanées peuvent être sèches ou accompagnées d'exsudats [11]. Les espèces de l'Ancien Monde ont des manifestations cliniques limitées

comparativement à celles du Nouveau Monde. Par ailleurs, il existe une grande diversité de développement et d'évolution des formes cutanées de leishmaniose, qui sont bien décrites et illustrées dans plusieurs revues de littérature récentes [11, 12] :

- *La leishmaniose cutanée localisée* : c'est la forme la plus répandue qui se caractérise par une à dix lésions ulcéreuses retrouvées sur les parties exposées du corps. Les espèces responsables sont tous les *Leishmania* dermatropiques.
- *Recidiva cutis* (*L. tropica*, *L. amazonensis*, *L. panamensis* et sous-genre *Viannia*) : plusieurs mois ou années après guérison d'une forme cutanée, des lésions papillaires ou vésiculaires apparaissent sur ou autour de la cicatrice.
- *La leishmaniose disséminée* (*L. braziliensis*) : elle est caractérisée par la présence de 10 à 300 lésions de formes variables, principalement sous forme de papule ou d'acnée, dans plusieurs régions du corps.
- *La leishmaniose cutanée diffuse* (*L. mexicana*, *L. amazonensis* et *L. aethiopica*) : c'est une forme anergique sévère de la maladie. Elle est caractérisée par la présence de plusieurs lésions nodulaires, non ulcératives et riches en parasites.
- *La leishmaniose mucocutanée* (*L. braziliensis* majoritairement, *L. panamensis*, *L. guyanensis* et *L. amazonensis*) : également connue sous le nom d'« *espundia* », cette forme apparaît parfois plusieurs années après une leishmaniose cutanée, en fonction de la virulence des souches et de l'état immunitaire du patient. Elle est le résultat d'une réponse inflammatoire incontrôlée, conduisant à l'infiltration massive de cellules inflammatoires détruisant les muqueuses mucopharyngales. Elle est caractérisée par la destruction des cavités orales, nasales et pharyngales et peut évoluer en lésions défigurantes (Figure 1-3B). Ces destructions sont irréparables, et sont parfois associées à des infections bactériennes secondaires et létales.

### **1.3.2. La leishmaniose viscérale.**

Plus de 90 % des cas de leishmaniose viscérale sont regroupés dans cinq pays : Bangladesh, Brésil, Inde, Népal et Soudan. Cette forme est causée principalement par

*L. donovani* (Inde et Afrique de l'Est) et *L. infantum* (pourtour méditerranéen, Asie centrale) [13]. La leishmaniose viscérale est également retrouvée en Amérique du Sud par *L. chagasi*, espèce très proche de *L. infantum* [2]. Plus rarement, ces formes de la maladie ont été décrites lors d'infections avec *L. mexicana* et *L. braziliensis* [13]. Les leishmanioses viscérales provoquent parfois de véritables épidémies meurtrières pouvant toucher des centaines de milliers de personnes, comme observé dans l'état du Bihar en Inde dans les années 1970 où 200 000 personnes ont été infectées. Ce pic d'infection a sans doute été causé par l'arrêt de pulvérisation d'insecticide contre la malaria. Dans la fin des années 1980, au Sud Soudan, la forme viscérale a causé la mort d'environ 100 000 personnes dans une région de 300 000 habitants, où les leishmanioses viscérales ne sont pourtant pas courantes [14].

Selon l'espèce parasitaire et le système immunitaire de l'hôte, la maladie peut se présenter sous la forme d'affection tégumentaire ou systémique. Les manifestations cliniques sont très variables, allant du portage asymptomatique prolongé à la forme viscérale proprement dite, mortelle en absence de traitement. À elle seule, la leishmaniose viscérale représente plus de 500 000 cas annuels et 50 000 décès.

Une triade clinique est généralement associée à la leishmaniose viscérale. Elle comprend une fièvre irrégulière (« fièvre folle »), une splénomégalie et une altération de l'état général. Une anémie et une hépatomégalie sont parfois observables. En outre, une hyperpigmentation cutanée concerne presque exclusivement l'infection par *L. donovani* en Inde. Elle est alors appelée « *kala-azar* » (fièvre noire en Hindi) ou « *Dumdum fever* ».

La maladie est due à la multiplication de parasites viscérotropiques dans le foie, la rate, les nodules lymphatiques et la moelle osseuse, voir même dans la totalité des organes en cas d'immunosuppression. Le gonflement de l'abdomen est dû à la splénomégalie et l'hépatomégalie (Figure 1-3C).

La leishmaniose viscérale peut évoluer en forme cutanée appelée leishmaniose cutanée post kala-azar. Cette migration, surtout observée avec *L. donovani*, joue un rôle important dans la réinfestation du vecteur. Cette évolution concerne 50 % des cas au

Soudan et 5 à 10 % des cas en Inde [13]. Elle est spontanément résolutive et présente des lésions principalement papulo-nodulaires et quelquefois maculopapulaires.

### **1.3.3. Cas particulier : La co-infection *Leishmania* / VIH.**

Une co-infection de *Leishmania* et du virus de l'immunodéficience humaine (VIH) représente un cas particulier, mais fréquemment rencontrée puisque les distributions géographiques de ces agents infectieux se concentrent souvent dans les mêmes régions. La co-infection peut être responsable d'une viscéralisation des leishmanioses cutanées, ou à la dissémination des parasites vers les muqueuses. Ces observations ne sont cependant pas catégorisées comme une forme mucocutanée. En effet, dans le cas des leishmanioses viscérales, il y a une dissémination inhabituelle des parasites qui infectent les bronches, les voies digestives ou le système tégumentaire et la triade clinique semble plus rare. La co-infection permet également une évolution plus rapide de l'infection par le VIH, puisque ces deux pathogènes ciblent les mêmes cellules et exercent un effet délétère synergique sur la réponse immunitaire. Aussi, les traitements anti-leishmaniens de type antimonies peuvent aggraver l'évolution de l'infection rétrovirale en stimulant la réplication du VIH-1 [15].

## **1.4. Diagnostics.**

Lors de leishmaniose cutanée, le diagnostic correspond à la mise en évidence de la présence de parasite sous forme amastigote (voir Chapitre 2) dans des cellules infectées. Pour ce faire, des prélèvements tissulaires sont colorés au Giemsa et analysés au microscope. Il est quasiment impossible de différencier les différentes espèces de *Leishmania* sur des bases morphologiques. Puisque l'identification des espèces infectieuses est importante pour permettre un traitement adapté de la maladie, différents tests complémentaires basés sur la détection d'ADN ou d'antigènes sont disponibles. Les tests sérologiques ne sont pas assez sensibles pour l'identification des espèces lors de leishmaniose cutanée.

Le Montenegro, ou *Leishmanin skin test* (LST), est parfois utilisé dans les pays endémiques. Il consiste en l'injection intradermique au niveau de l'avant bras d'une préparation d'antigènes de *Leishmania* traités au phénol. Le test s'avère positif si, 48 heures

après injection, un érythème de plus de 5 mm est présent au niveau du site d'injection. Ce test ne différencie pas une infection active d'une hypersensibilité tardive [10].

Dans le cas de leishmaniose viscérale, les techniques sérologiques présentent une bonne spécificité. La stratégie utilisée est basée sur des méthodes de détection rapide et simple par test ELISA (pour *Enzyme Linked Immuno Sorbent Assay*) dont les caractéristiques varient en fonction des antigènes utilisés. Des immuno-empreintes ou *Western Blot* sont effectuées comme techniques de confirmation dans quelques laboratoires spécialisés. Des techniques sérologiques par bandelette ont également été testées pour les contextes sanitaires difficiles. Le diagnostic formel de leishmaniose viscérale requiert la détection du parasite dans un prélèvement pathologique, les prélèvements profonds offrant une meilleure sensibilité (ponction ganglionnaire, prélèvement de moelle osseuse ou ponction splénique). Ici encore, la coloration au Giemsa et l'analyse microscopique sont privilégiées, mais la mise en culture sur des milieux adaptés permet l'obtention de formes promastigotes prolifératives, augmentant ainsi la sensibilité. Ces cultures permettent aussi de typer les souches par analyses isoenzymatique ou par PCR [13].

## **1.5. Préventions et traitements.**

Les leishmanioses font partie des maladies négligées, et sont même considérées par l'OMS comme émergentes et incontrôlées. Elles ne font pas l'objet de recherche à la hauteur de l'épidémie par les industriels en raison du faible retour économique que les ventes de traitements pourraient engendrer, surtout que les populations à risque sont souvent très pauvres. Ceci explique en partie pourquoi peu d'avancées thérapeutiques ont été réalisées depuis 1920 comparativement à d'autres infections. Étant donné qu'aucun vaccin n'est encore disponible, la prévention de la leishmaniose reste avant tout une lutte anti-vectorielle.

Le traitement historique et de première ligne des leishmanioses est l'antimoine pentavalent, administré par injection intraveineuse ou intramusculaire sur une durée de 20 à 28 jours [16]. Ce traitement engendre toutefois des intolérances et est également toxique. L'intolérance, bénigne, est caractérisée par des fièvres, des myalgies ou asthénies observées dans les premiers jours du traitement. Une intoxication, plus rare, peut être fatale. Elle

correspond à des atteintes graves cardiaques, rénales ou pancréatiques. Ce traitement a longtemps permis d'obtenir des taux de guérison supérieurs à 95 % pour un coût inférieur aux autres thérapies [9]. Cependant de plus en plus de résistances sont observées, notamment en Inde où elles concernent 35% des infections [17]. Le mécanisme d'action de ces molécules implique à la fois des facteurs de l'hôte et du parasites afin de produire un antimoine trivalent qui inhibe la trypanothione reductase parasitaire, inhibant ainsi la glycolyse et l'oxydation des acides gras, mais les détails de ces mécanismes restent flous [18, 19].

L'Amphotéricine B a des taux de succès pouvant dépasser les 95% chez les immunocompétents [20] et peut éliminer les souches résistantes aux antimoinés pentavalents. L'injection d'une forme liposomale permet de réduire sa toxicité rénale et d'augmenter son efficacité, mais cette molécule reste toxique et présente des effets secondaires. Les coûts de production élevés de l'Amphotéricine B ne permettent pas son utilisation à large échelle dans les régions pauvres. Les rares échecs de ce médicament ne semblent pas associés à l'émergence de résistance du parasite ce qui rend éventuellement possible le retraitement avec le même médicament [21]. L'Amphotéricine B cible les ergostérols de la membrane du pathogène et crée des pores pour lyser les parasites.

Depuis une dizaine d'années, la leishmaniose peut être traitée par voie orale grâce aux comprimés de miltéfosine. Son efficacité est supérieure à 90% en Inde chez les immunocompétents [13]. Malgré une bonne tolérance de la molécule, il existe des effets secondaires digestifs et tératogènes, et la longue demi-vie de la molécule est propice à l'émergence de résistance [22].

La pentamidine est moins efficace que l'Amphotéricine B [23]. Si elle fut un traitement de seconde ligne très utilisé par le passé, elle est aujourd'hui surtout prescrite en Amérique. Les parasites succombent à cette molécule par l'interférence à la synthèse de leur ADN et la dissolution de la membrane mitochondriale.

Enfin, la paramomycine, un aminoglycoside à large spectre utilisé pour contrer les infections bactériennes, est efficace pour traiter les leishmanioses viscérales en Inde [24]. Elle est injectée par voie intramusculaire et présente des effets secondaires.

L'obtention d'un vaccin à faible coût de production serait idéale pour pouvoir protéger les populations dans les zones à risque. Plusieurs études ont été effectuées, utilisant principalement la souris comme modèle animal. Les immunisations effectuées en utilisant des antigènes, de l'ADN ou des lysats parasitaires présentent des résultats prometteurs dans ce modèle, mais aucune immunisation préventive n'a été efficace chez l'homme [25]. Si les leishmanioses cutanées sont souvent résolutive et confèrent une immunité contre de nouvelles infections, cette dernière reste une immunisation concomitante puisque les parasites de la première infection persistent silencieusement et à un faible niveau dans l'organisme [26].



## Chapitre 2 : Cycle de vie de *Leishmania* et stratégies de survie parasitaire.

*Leishmania* présente un cycle parasitaire principalement dimorphique, la forme promastigote étant retrouvée dans l'insecte vecteur et la forme amastigote dans les macrophages des hôtes infectés. Cependant, au sein de l'insecte vecteur, différentes morphologies et activités multiplicatives sont observables pour permettre le déplacement du parasite dans l'intestin de l'insecte. Trois protéines de surface du pathogène vont jouer un rôle important dans sa survie, aussi bien chez l'insecte vecteur que dans l'hôte mammifère : la glycoprotéine GP63, le protéophosphoglycan (PPG) et le lipophosphoglycan (LPG).

Ce chapitre se concentre sur le cycle parasitaire, et ne traitera pas du cycle de division cellulaire. Notons que des travaux sont effectués sur les cycles cellulaires des *Trypanosomatidae* et sur les variations morphologiques qui en découlent. Les séquences de ce cycle ont été notamment détaillées pour *Trypanosoma cruzi* [27], *Trypanosoma brucei* [28], *L. mexicana* [29] ou *L. major* [30]. Il en ressort que la division cellulaire commence par la naissance d'un second flagelle, puis se poursuit par la mitose et la duplication de l'ADN mitochondrial. Chez *Leishmania* principalement, le cycle diffère notamment par la réduction de la phase G2, indiquant un passage rapide de la mitose à la cytokinèse. Cette dernière est incomplète, ou finalisée tardivement, dans 10% des cas pour des *Leishmania* cultivés en laboratoire [29].

Récemment, une étude a montré que *Leishmania* pouvait avoir un cycle de reproduction sexué dans son insecte vecteur, mais le mécanisme reste encore à décrire [31]. Différents *Leishmania* peuvent ainsi se rencontrer lors du passage dans l'insecte vecteur et des parasites hybrides peuvent être générés et transmis à l'hôte mammifère. Cet échange peut avoir un impact important au niveau épidémiologique, puisqu'un brassage génétique pourrait permettre au parasite de faciliter son adaptation à de nouvelles niches écologiques, ou à de nouveaux vecteurs ou hôtes [32].

## 2.1. Vue globale du cycle de vie

Lors d'un repas sanguin sur un hôte mammifère, le phlébotome infecté va injecter des parasites sous forme promastigote métacyclique, bien adaptée à l'infection. Cette forme est internalisée dans le phagolysosome des macrophages où elle se transforme en amastigote multiplicatif. La multiplication parasitaire engendre la lyse de la cellule immunitaire. Les amastigotes dans la circulation sanguine sont alors internalisés dans une nouvelle cellule phagocytaire. Lors d'une nouvelle piqûre par un autre phlébotome, l'insecte ingère des macrophages infectés, dont la digestion va libérer les amastigotes dans son intestin (Figure 2-1). Les amastigotes subissent plusieurs événements de différenciation, permettant leur adaptation, multiplication et migration vers le *stomodeum* (intestin antérieur) de l'insecte, et se préparent enfin à l'infection d'un nouveau mammifère.

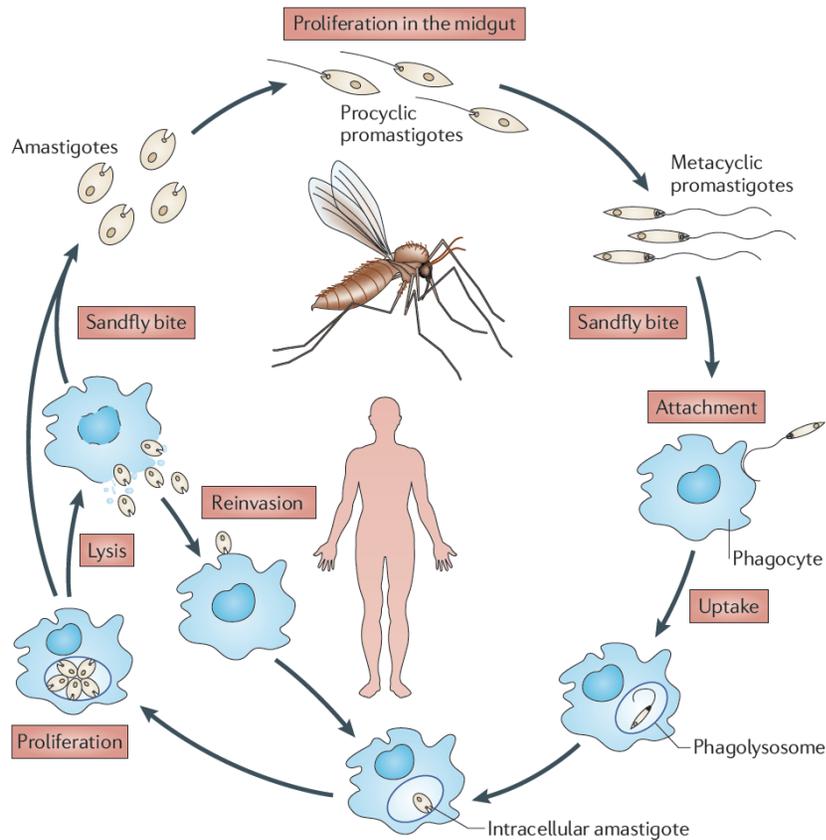


Figure 2-1 : Cycle de vie des parasites *Leishmania*.

Adapté de [26].

## 2.2. Au sein de l'insecte vecteur.

Hors de son hôte mammifère, *Leishmania* est confiné dans le système digestif des phlébotomes. En fonction des espèces, les parasites vont envahir l'insecte de différentes manières. Les parasites du sous-genre *Leishmania* présentent une infection suprapylarique, c'est-à-dire que leur développement s'effectue dans le mésentéron (intestin moyen) de l'insecte. Les *Leishmania* du sous-genre *Viannia* telle que *L. braziliensis* sont dits peripylarique car ils atteignent le *proctodeum* (intestin postérieur) avant de retourner vers le mésentéron [33]. La description ci-dessous représente le type suprapylarique, qui concerne donc la majorité des espèces de *Leishmania*, dont *L. infantum*.

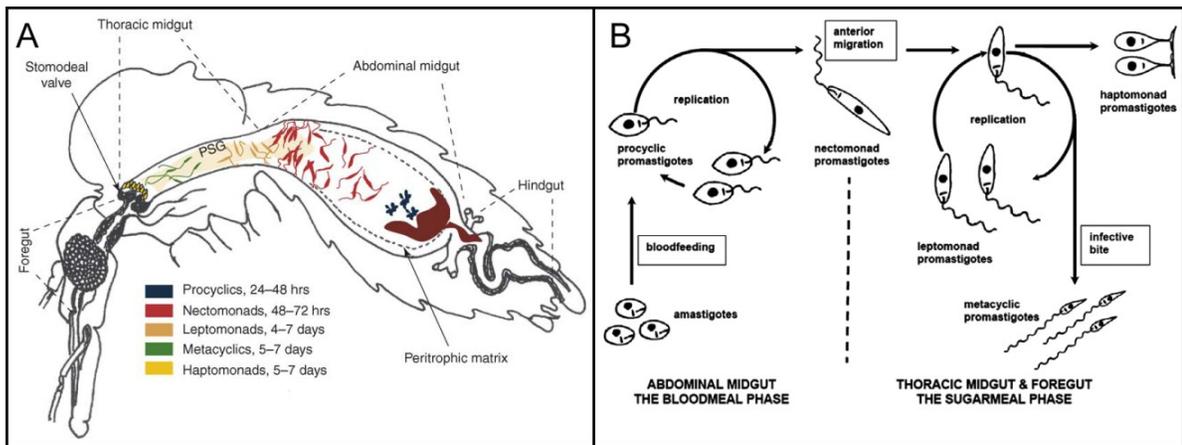


Figure 2-2 : Cycle de vie de *Leishmania* de type suprapylarique au sein de l'insecte vecteur.

(A) Localisation des principales formes promastigotes dans l'intestin d'un phlébotome femelle. Les temps indiqués correspondent aux moments d'observation de ses formes suite au repas sanguin. Tiré de [33]. (B) Séquences de développement et cycles réplicatifs des formes promastigotes. Tiré de [4].

Le cycle débute lorsque la mouche ingère du sang contenant des macrophages infectés avec des parasites *Leishmania* sous forme amastigote. Le repas atteint le mésentéron abdominal pour y être digéré. En six à neuf jours, les parasites vont remonter jusqu'au *stomodeum*, ce qui correspond à la partie insecte de leur cycle parasitaire (Figure 2-2).

Les amastigotes vont d'abord se différencier en promastigotes dits procycliques. Cette forme encore arrondie développe un court flagelle et effectue ses premiers cycles réplicatifs. Repas et parasites procycliques sont séparés de l'intestin par la synthèse d'une

poche de protéines et glycoprotéines protectrices, appelée matrice péritrophique, émise physiologiquement par l'insecte. La résistance aux attaques des enzymes digestives, l'évasion aux lectines sécrétées par l'insecte et l'inhibition du péristaltisme intestinal sont ici les déterminants de la survie parasitaire [33, 34].

Les procycliques deviennent ensuite promastigotes nectomonades. Cette forme effilée et non multiplicative permet aux parasites de se faufiler à travers les mailles de la matrice péritrophique, évitant ainsi la défécation et marquant l'infection de l'insecte [4]. Ils migrent ainsi jusqu'à la partie thoracique du mésentéron et se fixent aux cellules épithéliales du vecteur. La sécrétion de chitinase par les parasites est nécessaire pour permettre la dégradation de la chitine composant la matrice péritrophique [35].

Les nectomonades se transforment en promastigotes leptomonades, forme un peu plus courte qui effectue de nouveaux cycles réplicatifs [36]. Le LPG exprimé par le parasite est la glycoprotéine majoritaire de surface des promastigotes. Les galectines présentes sur la surface intestinale de l'insecte lui confèrent une immunité contre les pathogènes, mais *Leishmania* utilise l'une de ces protéines, PpGalec, comme récepteur des LPG afin de se lier à la paroi intestinale [37]. Cette interaction va définir la compétence des vecteurs à transmettre la leishmaniose, mais présente également une certaine spécificité d'affinité entre les espèces de parasites et d'insectes. Les leptomonades sécrètent un gel (PSG pour « *promastigote secretory gel* », Figure 2-2A) composé majoritairement de filaments de PPG et jouant un rôle important pour la transmission à l'hôte mammifère tel que décrite ci-dessous [38].

Les haptomonades sont une forme élargie et perdent leur motilité pour obstruer la valve stomodéale. La perte de la motilité est liée à la spécialisation de l'apex du flagelle en une forme de type hemi-desmosome [39]. Leur descendance des nectomonades est incertaine et leur importance dans la transmission des parasites métacycliques n'est sans doute pas essentielle. Ils semblent faciliter la génération et le maintien du PSG. La dégradation de la valve stomodéale par les chitinases sécrétés par les haptomonades dérègle les habitudes alimentaires de l'insecte « parasitiquement modifié » (voir ci-après), conduisant à l'augmentation du nombre de piqûres et à des périodes de succions prolongées [5].

Enfin, les métacycliques dérivent des nectomonades et sont la forme promastigote préparée à l'infection d'un hôte mammifère. Ils sont non multiplicatifs, de forme effilée et pourvus d'un long flagelle. Ces parasites sont regroupés devant la valve stomodéale et le long du *stomodeum*. La métacyclogénèse [40] est l'étape essentielle pour la transmission du parasite. La modification de l'expression génique et l'adaptation morphologique lors de cette étape permettent au parasite d'acquérir des mécanismes de défense, notamment contre la lyse par le complément et les activités microbicides dépendantes ou indépendantes de l'oxygène (voir la partie suivante). Plusieurs études comparatives au niveau protéomique ont été effectuées avec des parasites métacycliques pour mieux comprendre les facteurs de virulence exprimés spécifiquement durant la métacyclogénèse [41-43].

Les modifications de comportement de l'hôte intermédiaire par un parasite afin d'optimiser la pérennité de son cycle de vie sont connues sous le terme de manipulation parasitaire [44]. Parmi les manipulations les plus étonnantes, on retrouve le nématomorphes *Paragordius tricuspidatus* dont la larve va inciter son hôte, le grillon, à se « suicider » par noyade afin de continuer son cycle de vie dans un milieu aquatique, mais également la petite douve du foie du mouton, *Dicrocoelium dendriticum*, qui va obliger son hôte intermédiaire, la fourmi, à rester sur les herbes mangées par les moutons. Chez *Toxoplasma gondii*, les rats infectés ne sont plus apeurés par l'odeur de l'urine de chat, l'hôte définitif, mais sont attirés et sexuellement stimulés par cette odeur afin de favoriser le passage du rat vers le chat [45]. Chez *Leishmania*, certains auteurs considèrent la forme haptomonade comme altruiste, puisqu'elle ne sera pas adaptée à l'infection de l'hôte mammifère et semble uniquement optimiser la transmission des métacycliques [4]. Ceci n'est pas sans rappeler le sacrifice de certains *D. dendriticum* non infectieux, migrant dans le système nerveux de la fourmi pour prendre le contrôle de ses mouvements et permettre à ses semblables infectieux de parasiter le mouton. Ces observations suggèrent que, comme d'autres organismes parasites, les promastigotes ont une fine régulation de leur différenciation et sans doute une communication intercellulaire au sein des phlébotomes, qui confèrent un comportement d'apparence social des parasites dans le but d'optimiser la survie de l'espèce au dépend de celle de l'individu.

## **2.3. Dissémination et survie des parasites dans l'hôte mammifère.**

La transmission des parasites par piqûre entraîne une dégradation de la peau de l'hôte mammifère. La réparation tissulaire, réponse habituelle de l'hôte aux blessures, va rapidement recruter au site d'inoculation des neutrophiles et des macrophages, indépendamment de la présence d'agents infectieux. Les parasites vont co-évoluer lors de la réparation tissulaire pour pouvoir atteindre leur principal objectif : s'installer et se développer dans les phagolysosomes des macrophages [46]. Les macrophages sont une niche très intéressante pour le parasite car ils possèdent une longue durée de vie, leur laissant amplement le temps de se transformer sous forme amastigote et de se multiplier. Quelques études mentionnent une localisation qui semble cytoplasmique des parasites amastigotes, mais des confirmations restent nécessaires pour pouvoir l'affirmer, et l'avantage ou le détriment d'un tel phénomène pour le parasite n'est pas élucidé [47].

### **2.3.1. Inoculation des parasites.**

Des études effectuées avec le phlébotome *Lutzomyia longipalpis*, infecté par *L. mexicana*, ont estimé qu'environ 1 000 parasites, fortement enrichis en métacycliques, étaient libérés par l'insecte lors d'un repas sur membrane en laboratoire [48]. Des mesures de l'*inoculum* ont été effectuées par PCR en temps réel lors de piqûres d'oreilles de souris avec le phlébotome *Phlebotomus duboscqi* infecté par *L. major*. Dans la majorité des cas, 600 parasites étaient transmis par une piqûre, mais certaines permettent le passage de 1 000 jusqu'à 100 000 parasites. Les souris infectées avec une faible dose développent une pathologie limitée, mais constituent un bon réservoir à long terme [49]. En plus des parasites infectieux, l'*inoculum* contient des parasites morts ou apoptotiques, des filaments de PPG, et des composés de la salive de l'insecte qui vont avoir un rôle important dans la modulation de la réponse immunitaire de l'hôte mammifère. Les phlébotomes peuvent également co-transmettre des virus ou des bactéries qui influenceront la réponse immunitaire [26].

### **2.3.2. La résistance au complément.**

Une fois déversés dans le milieu extracellulaire de l'hôte, la première épreuve rencontrée par les parasites est de contourner la lyse par le complément. Le complément est

un complexe multi protéique, activable par différentes voies, qui conduit à la fixation du complexe d'attaque membranaire (C5b-C9), permettant de créer des pores sur les cellules pathogènes. Il y a donc lyse de la cellule par perte de son homéostasie (revue dans [50]). Le système du complément peut également conduire à l'opsonisation des pathogènes, qui dans ce cas, ne sont pas lysés mais phagocytés par les macrophages via les récepteurs du complément (CR). Les quatre CR1-4 reconnaissent notamment des protéines du complément C3b et C3bi (version inactive du C3b) avec différentes affinités [51].

L'évasion au système du complément semble être effectuée différemment en fonction des espèces de *Leishmania*. Dans le cas de *L. major*, la métacyclogénèse permet entre autre d'augmenter de deux à trois fois la longueur des protéines LPG [52], ce qui interdit l'ancrage du complexe C5b-C9 du complément, et donc empêche la lyse cellulaire. Les parasites procycliques sont eux plus sensibles à cette lyse [53]. Une alternative a été proposée pour *L. donovani*, la résistance à la lyse par le complément s'explique par la fixation préférentielle du C3bi, inhibant la lyse par le complément [54] et favorisant l'opsonisation. La GP63, dix fois plus abondante chez les métacycliques que chez les procycliques, participe à la conversion du C3b en C3bi. De plus, plusieurs protéines kinases du parasite inhibent la cascade de réaction du complément par phosphorylation de ses membres [55]. *Leishmania* sous forme métacyclique ne fait donc pas que déjouer la lyse par le système du complément, il profite de ce dernier pour entrer dans les cellules phagocytaires. Ceci permet au parasite d'infecter « silencieusement » le macrophage, c'est-à-dire sans activer le système immunitaire. Cependant, tous les parasites ne résistent pas à l'attaque du complément, et il en résulte une forte pression de sélection pour les parasites les mieux adaptés. En effet, la lyse par le complexe d'attaque membranaire est visible dès une minute après contact avec le sérum, et 90% des parasites succombent en 2 min 30 [56].

### **2.3.3. Dissémination des parasites dans l'organisme infecté.**

Trois principaux types de cellules phagocytaires vont ensuite entrer en jeu et internaliser les parasites : les neutrophiles, les cellules dendritiques et les macrophages (Figure 2-3). C'est uniquement au sein des cellules phagocytaires mononucléées que *Leishmania* se différencie, se multiplie et survit sur du long terme. Une certaine

chronologie se dessine entre ces différents acteurs au cours de l'infection, qui ne semble pas pour autant répondre à la loi du « premier arrivé, premier servi » [57].

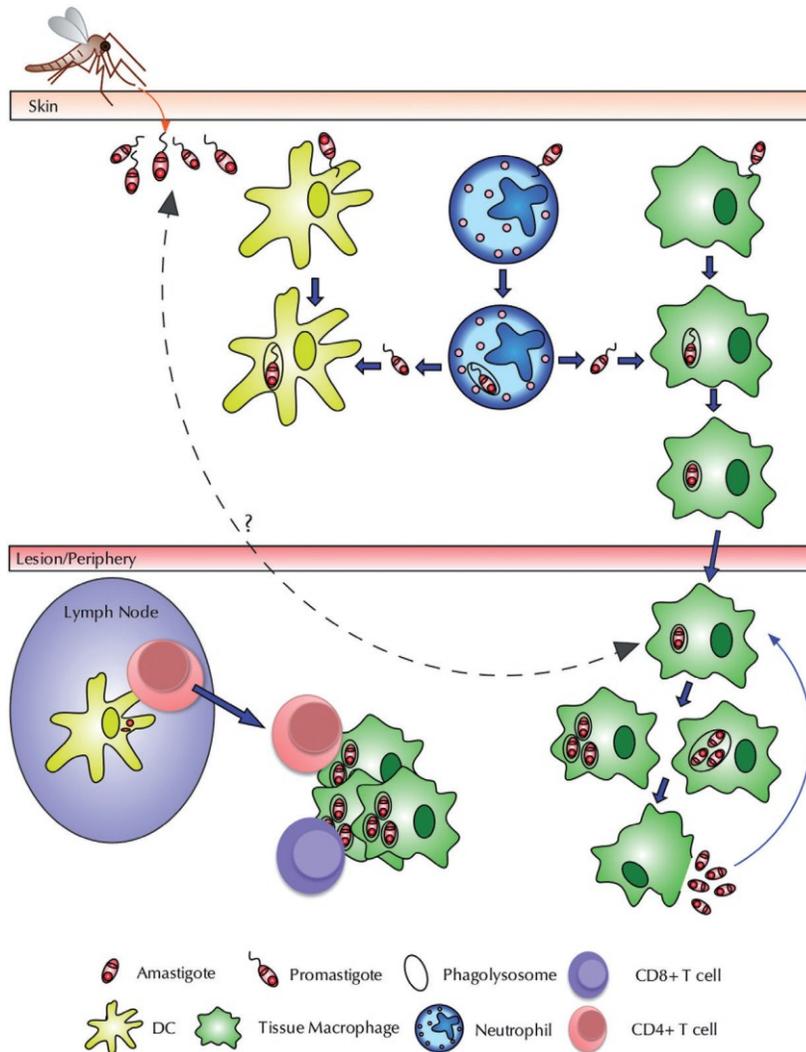


Figure 2-3 : Type des cellules infectées et dissémination de *Leishmania* dans l'organisme.

Adapté de [57].

La majorité des parasites est d'abord phagocytée par les neutrophiles qui infiltrent rapidement le site de lésion. Les parasites ne se différencient pas en forme amastigote et ne se multiplient pas au sein de cette lignée cellulaire [58], mais les neutrophiles infectés deviennent apoptotiques et constituent alors un repas phagocytaire pour les macrophages. Les neutrophiles apoptotiques sont phagocytés par l'intermédiaire de récepteurs qui n'induisent pas l'établissement de défense antimicrobienne au sein des macrophages,

permettant ici encore une entrée silencieuse des parasites au sein du phagosome des macrophages [26]. La survie de *L. major* dans les phagosomes de neutrophiles semble liée à l'inhibition de la fusion avec les granules impliqués dans l'acidification du phagosome et la production de radicaux libres [59]. Les neutrophiles participent cependant à l'élimination de *Leishmania* en sécrétant le piège extracellulaire des neutrophiles (NET, de l'anglais *neutrophil extracellular traps*), correspondant à un piège fibreux d'ADN, d'histones et de granules protéiques relargués par le neutrophile lors de leur lyse [60].

Si les neutrophiles sont les principales cellules phagocytaires infectées suite à l'inoculation, la fréquence de leur infection décroît simultanément à l'accroissement de celle des macrophages. En observant ce phénomène, il a été proposé que *Leishmania* se serve des neutrophiles comme cheval de Troie pour infecter discrètement les macrophages, afin d'éviter la mise en place d'une réponse immunitaire adaptative. Ce modèle est de plus en plus délaissé pour favoriser celui du « lapin de Troie » [61], dans lequel le parasite n'entre pas dans le macrophage en étant au sein même du neutrophile, mais plutôt en s'échappant du neutrophile juste avant l'arrivée du macrophage. Ceci lui permet d'être co-phagocyté sans pour autant activer la cellule immunitaire [26, 57, 61]. L'intérêt du passage par le neutrophile dans l'adaptation parasitaire n'est pas encore élucidé, il n'est pas non plus obligatoire *in vitro* pour l'infection des macrophages puisque celle-ci est également observable chez des souris déficientes en neutrophiles.

Les cellules dendritiques phagocytent également des parasites durant les premières heures de l'infection, mais les macrophages sont la population cellulaire infectée dominante après 24h [62]. Puisque les macrophages et cellules dendritiques résidants dans la peau sont en nombre trop limité pour permettre la multiplication parasitaire, la progression de l'infection semble impliquer les précurseurs immédiats de ces cellules, les monocytes, et leur recrutement est justement régulé par les neutrophiles [26]. Il semble que la majorité des monocytes qui arrivent sur le site d'infection se différencie en cellules dendritiques (MDDC, de l'anglais *monocyte-derived dendritic cell*). Lorsque ces cellules sont infectées par *L. braziliensis*, il y a une inhibition des fonctions permettant le développement d'une réponse immunitaire acquise, notamment en inhibant la surproduction du complexe majeur d'histocompatibilité de classe II et en inhibant la sécrétion d'Interleukine-12 (IL-12),

cytokine qui dirige le système immunitaire vers la voie Th1, réponse protectrice pour l'organisme (Figure 2-4).

Il est encore difficile d'affirmer que ce sont les cellules dendritiques et les macrophages qui transportent les parasites du derme vers les nœuds lymphatiques, passage obligatoire pour développer une réponse immunitaire acquise [26]. Comme précisé dans le chapitre 1, les mécanismes permettant le tropisme des parasites au sein de leur hôte mammifère sont encore très peu compris.

Une fois dans le phagolysosome des macrophages, *Leishmania* se transforme en sa forme amastigote afin de s'adapter à son nouvel environnement et de proliférer (détaillé dans les parties suivantes). La multiplication diffère en fonction des parasites. En effet, les espèces de l'Ancien Monde (*L. major*, *L. donovani*) vivent dans des petites vésicules contenant un seul amastigote alors que celles du Nouveau Monde prolifèrent dans une seule vésicule remplie d'amastigotes [63]. Les parasites doivent maintenir une balance rigoureuse entre prolifération pour augmenter les chances de transmission à un nouvel insecte et discrétion afin d'éviter la détection par le système immunitaire. Pour les espèces de viscéralisation (*L. donovani*, *L. infantum*, *L. chagasi*), la prolifération est synchrone à la dissémination du lieu de piqûre aux organes internes, tels que le foie, la rate et la moelle osseuse. Faute de bon modèle animal et des limites de l'imagerie médicale [57], ces mécanismes sont encore très peu étudiés actuellement.

La multiplication des parasites provoque la lyse de la cellule immunitaire. Les amastigotes ainsi libérés vont alors être phagocytés par d'autres cellules phagocytaires et continuer leur prolifération. Le cycle au sein de l'hôte mammifère se termine pour certains parasites lorsqu'un nouveau phlébotome vient prendre un repas sanguin et ingère des pathogènes et des macrophages infectés.

### **2.3.4. Interaction de *Leishmania* avec les macrophages.**

#### **2.3.4.1. Les différents types de macrophage.**

Contrairement aux infections bactériennes, les infections parasitaires au niveau du derme sont principalement contrôlées par une réponse immunitaire adaptative. Les réponses immunitaires protectrices sont généralement dirigées par les cellules T plutôt que par la

production d'anticorps. Les cellules T CD4<sup>+</sup> constituent la population cellulaire effectrice la plus importante dans les infections par *Leishmania* et dirigent la réponse de l'hôte à l'infection. Chez l'homme et la souris, les cellules T CD4<sup>+</sup> peuvent être subdivisées en au moins deux sous groupes de cellules T helper (Th) polarisées, Th1 et Th2, en fonction de leur sécrétion de cytokines [64]. Ce sont ces cytokines qui vont dicter la maturation des macrophages.

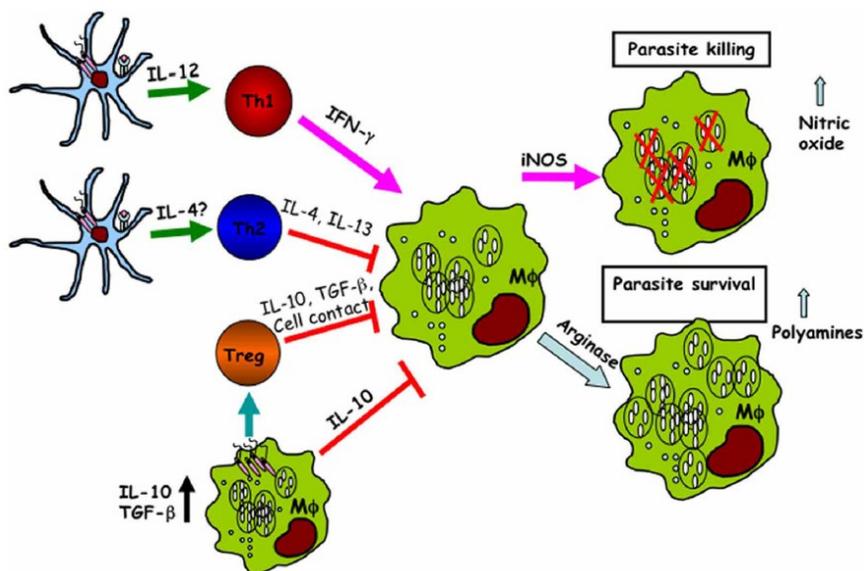


Figure 2-4 : Susceptibilité au parasitisme par *Leishmania* des macrophages en fonction de la réponse immunitaire.

Suite à une infection, les macrophages et les cellules dendritiques phagocytent les parasites *Leishmania*, provoquant différents types de réponse immunitaire. Les cellules dendritiques infectées produisent de l'IL-12, qui est un élément essentiel pour le développement des cellules Th1. L'IFN-γ produit par ces dernières permet une activation classique des macrophages, ce qui entraîne la surexpression d'iNOS puis la production d'oxyde nitrique et autres radicaux libres importants pour permettre la lyse des parasites dans le macrophage. Par contre, la production d'IL-4 développe des cellules Th2 productrices d'IL-4 et d'IL-13. Ces interleukines stimulent l'activité arginase, l'activation alternative des macrophages et la production de polyamines, favorisant la prolifération et la survie intracellulaire du parasite. Enfin, les cellules T régulatrices et les macrophages infectés produisent également des cytokines, dont l'IL-10 et le TGF-β, qui augmentent également la survie parasitaire. Adapté de [65].

En fonction de leur mécanisme d'activation, divers types de macrophages sont produits, ayant chacun un rôle immunologique différent [66]. L'activation classique est

effectuée par l'interféron  $\gamma$  (IFN $\gamma$ ), sécrété par des cellules Th1 ou Natural Killer (NK). Ces macrophages, tissulaires, possèdent une activité microbicide et sont nécessaires pour contrer les infections à *Leishmania*. L'activation par la voie alterne est effectuée via l'IL-4, sécrétée par des cellules Th2. Ce second type d'activation produit des macrophages impliqués dans la réparation tissulaire, sans activité microbicide. La sécrétion d'IL-10 par des cellules T régulatrices participe à la maturation des macrophages « régulateurs », spécialisés dans la réponse anti-inflammatoire de l'organisme [66]. Seuls ces deux derniers types d'activation produisent des macrophages non microbicides, permissifs aux infections par *Leishmania*. La résistance ou la susceptibilité aux infections à *L. major* est donc respectivement associée au développement d'une réponse immunitaire à dominance Th1 ou Th2 [32, 46] (Figure 2-4).

#### **2.3.4.2. Phagocytose par les macrophages et fusion phagosomes – lysosomes.**

La phagocytose de *Leishmania* par le macrophage nécessite l'interaction de molécules de surface du parasite avec les récepteurs membranaires du macrophage et plusieurs liaisons ont été décrites. Les protéines GP63 et LPG interagissent directement avec les récepteurs de la cellule immunitaire CR1 et CR3, ou indirectement suite à l'opsonisation du parasite. CR3 semble favorisé puisqu'il lie le C3bi, et de plus, il ne provoque pas de stress oxydatif pendant la phagocytose [67]. Les récepteurs aux fibronectines, au mannose-fructose, ou encore à la protéine C réactive interviennent également (revu dans [55, 68]). Les amastigotes ont des molécules de surface différentes de celles des promastigotes métacycliques. En plus des interactions mentionnées, le système immunitaire a eu le temps de mettre en place une réponse immunitaire humorale et les amastigotes peuvent être phagocytés en utilisant les récepteurs Fc du macrophage, suite à l'opsonisation du parasite par des anticorps spécifiques [69]. La phagocytose de *Leishmania* par le macrophage est un mécanisme qui reste silencieux favorisant la survie du parasite [70]. Suite à l'attachement, les parasites sont internalisés dans la vacuole parasitophore (ou phagosome) via des pseudopodes générés par le macrophage (revu dans [47]).

Parallèlement à la phagocytose, il a été montré que les exosomes sécrétés par les parasites pouvaient libérer des facteurs de virulence au sein du cytoplasme des

macrophages et contribuent à la pathogénicité. Les exosomes sont des vésicules qui bourgeonnent de la membrane plasmique retrouvés dans une grande variété de cellules de mammifères et de pathogènes (revu dans [71]). Les analyses protéomiques effectuées sur les exosomes de *Leishmania* [72] ont montré que ces vésicules contiennent la grande majorité des facteurs de virulence. L'administration des exosomes de *Leishmania* avant l'infection montre un rôle immunosuppresseur du contenu de ces vésicules, inhibant la production de TNF- $\alpha$  et stimulant celle de IL-10 chez la souris [73].

Les microdomaines lipidiques de la membrane influencent le contact entre cellules et jouent un rôle dans la signalisation cellulaire. Les radeaux lipidiques des macrophages semblent avoir un rôle dans la pénétration des facteurs de virulence de *Leishmania* dans le cytosol. Des analyses du transcriptome ont montré que la voie de synthèse des lipides chez le macrophage était perturbée lors de l'infection par *Leishmania* [74]. *L. donovani* est capable d'interférer sur la présentation antigénique des macrophages en modulant ces microdomaines [75]. Les LPG de *L. donovani* s'associent également à ces microdomaines sur la membrane du phagosome, inhibant ainsi la fusion de la vacuole parasitophore avec les lysosomes [76]. Ce ralentissement de la fusion va permettre aux métacycliques d'initier leur différenciation en amastigotes, une forme plus résistante aux hydrolases concentrées dans le lysosome. La transformation de promastigote en amastigote, associée à l'arrêt de l'expression de LPG par le parasite, est concomitante à l'arrêt de l'inhibition de la fusion lysosome-phagosome [77].

#### **2.3.4.3. Les stress environnementaux et la différenciation du parasite en sa forme amastigote.**

La transformation morphologique et physiologique des parasites en amastigotes est la cytodifférenciation la plus marquée durant tout le cycle parasitaire de *Leishmania*. Elle est associée à une forte atrophie du flagelle (voir Chapitre 3), une réduction de la taille de la cellule, son arrondissement et un remaniement des protéines membranaires. Les nutriments disponibles varient entre l'intestin des insectes et le phagolysosome des macrophages, demandant également des adaptations métaboliques de *Leishmania* entre stades promastigotes et amastigotes (revu dans [78, 79]). La différenciation de *L. donovani* en milieu axénique a montré une surexpression des enzymes de la néoglucogénèse, une

diminution des capacités glycolytiques, une augmentation de la voie de la chaîne respiratoire mitochondriale et des phosphorylations oxydatives [80].

En plus des différences observées dans les nutriments disponibles, d'autres stress environnementaux sont rencontrés par le parasite au cours de son cycle parasitaire, mais le pH et la température semblent être les principaux facteurs de différenciation de promastigote vers amastigote (revus dans [81]). L'insecte vecteur étant poïkilotherme, la température est dépendante des conditions météorologiques (généralement définie à 25°C dans le cadre expérimental) et le pH est relativement basique, autour de 8.0 dans l'intestin de l'insecte. Au sein du phagolysosome d'organismes thermostables, la température est plus constante et plus élevée (34°C à 37°C chez l'homme en fonction de la pénétration dermique ou viscérale), le pH acide se situe entre 4.5 – 6.0 et les parasites sont en plus exposés aux attaques oxydatives.

En laboratoire, il est possible d'induire la différenciation de certaines souches de *Leishmania* en amastigote, dont *L. infantum* sans utiliser de cellule phagocytaire. Ces conditions axéniques sont obtenues en cultivant les parasites à un pH de 5.5, à 37°C avec un taux de CO<sub>2</sub> de 5 à 7% (revu dans [82]). Tout comme les amastigotes présents dans les macrophages, ces cultures amastigotes axéniques présentent les adaptations morphologiques, expriment les protéines spécifiques de ce stade de vie, inhibent celles qui sont spécifiques des promastigotes, et sont assez virulentes pour réussir l'infection de hamster et de lignées de macrophages. Les études comparatives de transcriptome entre amastigotes dérivés de macrophage et amastigotes axéniques montrent cependant que la régulation de l'expression génique diffère entre ces systèmes, principalement au niveau métabolique, protéolytique et dans le transport intracellulaire [83].

Les études menées sur la chronologie de la différenciation en milieu axénique ont fait ressortir que la transformation morphologique vers les amastigotes commençait 5 heures après le changement de conditions de culture et cette différenciation est synchrone à un arrêt du cycle cellulaire en phase G1. Le stress thermique est responsable de l'arrêt du cycle chez les promastigotes, alors que la modification de l'acidité permet de relancer le cycle et de diriger les parasites vers la différenciation. Globalement, l'étude conclue que les quatre premières heures correspondent à la réception des signaux de différenciation, les

parasites perdent leur motilité et forment des agrégats entre cinq et sept heures, atrophient leur flagelle entre neuf et douze heures, et la différenciation morphologique est effectuée entre 10 et 24 heures après le stress. Plusieurs heures à quelques jours sont ensuite nécessaires à la différenciation complète et à la stabilisation de l'expression des gènes spécifiques aux amastigotes. La différenciation inverse, d'amastigotes axéniques à promastigotes, dure de 24 à 48 heures [84].

### **2.3.5. Implication de virus de *Leishmania* dans la sévérité de l'infection.**

Les parasites *Leishmania* sont eux-même parasitables. Chez quelques parasites d'Amérique du sud (sous-genre *Viannia*), *Leishmania* contient un virus à double brin d'ARN de la famille des *Totiviridae* : LRV1 (pour *Leishmania RNA virus 1*) [85, 86]. Celui-ci modifie la virulence des parasites ainsi que la métastase parasitaire [87].

Le développement d'une leishmaniose mucocutanée est en général observé après la guérison d'une lésion primaire et est associé à une réponse inflammatoire incontrôlée. Il a été démontré que des parasites *L. guyanensis* infectés par LRV1 stimulaient la production de molécules proinflammatoires par les macrophages, tel que TNF- $\alpha$  par exemple. Cette réponse au virus est effectuée via le récepteur TLR3 (pour *Toll-like receptor 3*), un récepteur de l'ARN double-brin situé notamment sur la membrane de la vacuole parasitophore. Les auteurs émettent l'hypothèse que l'ARN viral est libéré par les parasites lysés rapidement après l'infection et que la liaison de l'ARN viral avec TLR3 induit la production de cytokines augmentant la réponse inflammatoire et exacerbant ainsi la maladie [88]. Cette récente découverte nécessite plus de données pour définir l'importance épidémiologique et clinique de ce phénomène.



## Chapitre 3 : Origine et organisation cellulaire des *Kinetoplastidae*.

*Leishmania infantum* fait partie de l'ordre des *Kinetoplastidae*, qui regroupe des protozoaires unicellulaires rapidement séparés des autres eucaryotes lors de l'évolution. Les *Kinetoplastidae* possèdent plusieurs particularités structurales et biologiques qui leurs sont propres. De ce fait, les mécanismes moléculaires de ces organismes sont parfois difficilement comparables aux mécanismes observés chez les modèles d'étude plus communs, que ce soit les bactéries, les levures, ou encore les cellules murines et humaines. Dans le cas de *Leishmania*, les données sont souvent comparées à celles obtenues chez d'autres *Kinetoplastidae* plus largement étudiés tel que *Trypanosoma brucei*. Suite à une brève présentation de *Trypanosoma brucei* et *Trypanosoma cruzi*, la divergence évolutive des *Kinetoplastidae* est abordée et une présentation de leur organisation cellulaire est décrite afin de mieux comprendre l'originalité de ces modèles d'étude. Les mécanismes atypiques d'expression et de régulation génique seront détaillés dans le chapitre 4.

Les *Kinetoplastidae* regroupent de nombreuses espèces de parasites, dont certains sont d'importants pathogènes humains. Les genres *Trypanosoma* et *Leishmania* sont responsables de millions d'infections dans les régions tropicales et subtropicales.

*Trypanosoma brucei*, transmis par piqûre de glossines, est notamment responsable de la maladie du sommeil chez l'homme. Un demi-million de personnes est infecté par *T. brucei* et environ 50 000 personnes en meurent chaque année [89]. Le parasite vit dans la circulation sanguine, puis le liquide céphalo-rachidien de son hôte mammifère. Il exprime une glycoprotéine de surface variable (VSG), lui permettant de constamment renouveler ses antigènes de surface et d'échapper ainsi au système immunitaire de son hôte.

*Trypanosoma cruzi* est le parasite responsable de la maladie de Chagas et est transmis *via* les déjections de punaises hématophages de la famille des *Reduviidae*. Environ 75 millions de personnes vivent dans des zones à risque pour cette maladie, et 11 à 18 millions sont actuellement infectées en Amérique du Sud et en Amérique centrale. Il est

estimé que 300 000 nouveaux cas surviennent et 13 000 personnes en succombent chaque année [90]. Chez son hôte mammifère, le parasite est phagocyté par le macrophage, puis migre se nicher dans le cytoplasme de la cellule immunitaire.

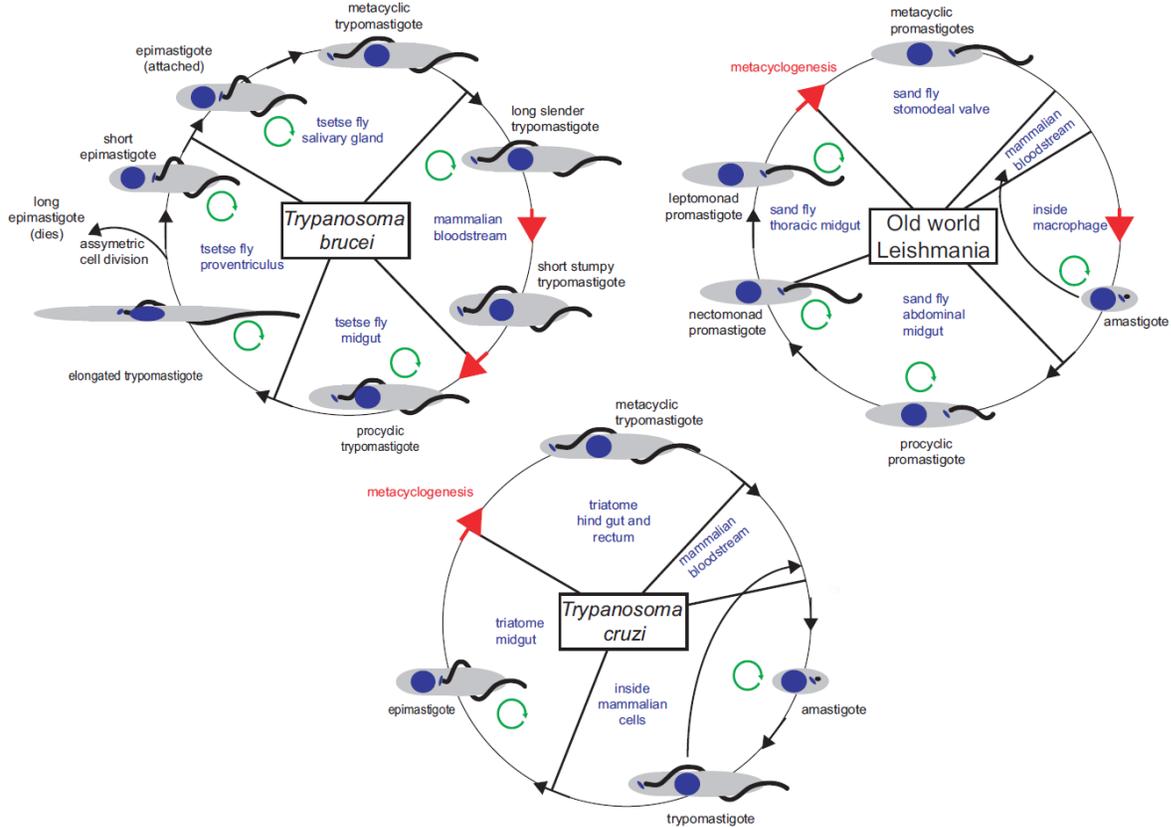


Figure 3-1 : Comparaison des cycles parasitaires des *Kinetoplastidae* infectieux *T. brucei*, *T. cruzi* et *Leishmania* spp.

Les flèches circulaires vertes représentent les stades parasitaires multiplicatifs, les flèches rouges indiquent les stades de différenciation les plus étudiés. La description du cycle de **Leishmania** a été effectuée dans le chapitre 2. **T. brucei** se divise sous une forme trypomastigote longue et effilée, extracellulaire dans la circulation sanguine de son hôte mammifère. Son mécanisme d'évasion immunitaire implique des glycoprotéines membranaires variables (VSG) qui provoquent des vagues parasitémiques. Lors d'un pic de parasitémie, les trypomastigotes se modifient en une forme plus courte qui est transmise lors d'un repas de glossine. Dans le mésentéron de l'insecte, les parasites se différencient en trypomastigotes procycliques prolifératifs. Ils migrent dans le proventricule, deviennent plus longs et une division asymétrique résulte à la formation d'épimastigotes longs et courts. Les longs épimastigotes ne sont pas capables de continuer leur cycle de vie. Les épimastigotes courts migrent dans la glande salivaire, s'attachent aux cellules épithéliales et continuent leur division. Des échanges génétiques par méiose sont effectuables à ce stade du parasite.

Le remaniement des protéines membranaires va définir le stade métacyclique infectieux et non réplicatif, libéré dans la circulation d'un nouvel hôte mammifère lors du prochain repas sanguin de la glossine. *T. brucei* n'a pas de stade intracellulaire, contrairement à *Leishmania* et *T. cruzi*. Voir [91] pour plus de détails sur *T. brucei*. **T. cruzi** sous forme promastigote métacyclique est présent dans les fèces de son insecte vecteur, le triatome. Ces fèces sont généralement proches du site de morsure par l'insecte sur la peau des mammifères. Les parasites entrent par la blessure ou *via* les muqueuses intactes de l'hôte puis envahissent les cellules proches du site d'inoculation : macrophages, fibroblastes ou cellules épithéliales par phagocytose ou endocytose. Une fois à l'intérieur de la cellule, les promastigotes métacycliques se différencient en amastigotes, se divisant dans la vacuole parasitophore. Le parasite peut éventuellement se différencier en trypomastigote dans le cytoplasme et infecter un nouveau macrophage suite à la lyse de la cellule phagocytaire initiale. Durant le repas d'un triatome, les trypomastigotes sont transférés dans l'estomac de l'insecte où certains d'entre eux se différencient en épimastigotes, migrent dans l'intestin et commencent leur multiplication. Ils s'attachent alors aux cellules intestinales, permettant leur différenciation en métacycliques infectieux à nouveau transmissibles par les fèces de l'insecte. Voir [92] pour plus de détails sur *T. cruzi*. Schémas des cycles adaptés de [93].

Il est estimé que plus de 12 millions de personnes souffrent de l'une des manifestations cliniques des leishmanioses. Deux millions de nouvelles infections et 70 000 décès sont enregistrés chaque année. Au total, ce sont 350 millions de personnes qui sont exposées à un risque d'infection à travers le monde [94].

Ces chiffres considérables s'expliquent par la distribution géographique de ces maladies, qui touche majoritairement les populations pauvres avec un faible accès aux médicaments, par l'absence de vaccin, mais aussi par les particularités biologiques de ces parasites. Les *Kinetoplastidae* sont passés maîtres dans l'art du parasitisme et de l'évasion du système immunitaire, et cela est sans doute lié à leur divergence rapide de la branche des eucaryotes.

### 3.1. Origine des *Kinetoplastidae*.

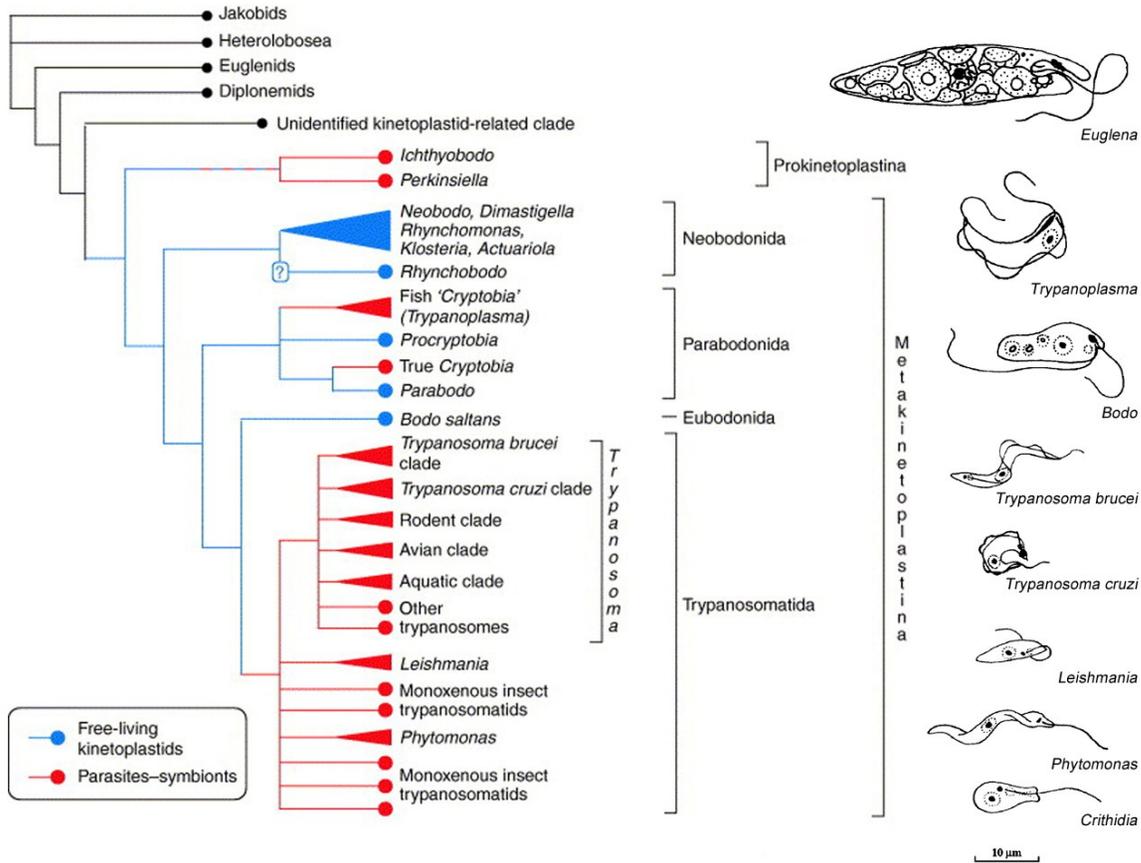


Figure 3-2 : Relations évolutives des *Kinetoplastidae*.

Relations estimées en fonction des arbres phylogéniques obtenus à partir des SSU rRNA et de la phylogénie des protéines. L'arbre est issu de [95], les schémas sont adaptés de [96].

La classification de l'ordre des *Kinetoplastidae* fait l'objet de nombreux débats aujourd'hui encore. Les premières relations phylogénétiques établies ne s'accordaient pas toujours entre l'acquisition d'un cycle dimorphique, la morphologie, les cycles cellulaires et la distribution des hôtes. Les études basées sur le séquençage de l'ARN de la petite sous-unité ribosomale (SSU rRNA) montraient un pourcentage de divergence au sein des *Kinetoplastidae* quasi similaire aux résultats observés entre vertébrés et invertébrés (13,1% de divergence contre 13,8%) [96]. L'évolution rapide et les nombreuses mutations observées sur les SSU rRNA créaient des incohérences et rendaient difficile l'estimation correcte des relations entre les *Kinetoplastidae*. Le séquençage de nombreux génomes (voir Chapitre 4) a permis d'optimiser les relations phylogéniques de cet ordre en comparant

directement certains gènes plus conservés. Les derniers résultats obtenus soutiennent actuellement la monophylie des *Kinetoplastidae* (Figure 3-2) [95, 97].

La classification actuelle montre que les *Kinetoplastidae* ont un ancêtre commun avec les Euglénidés, protozoaires connus pour avoir acquis un plastide d'une algue verte par endosymbiose secondaire. Ce lien de parenté, en plus du contenu enzymatique des *Kinetoplastidae*, est sujet à des spéculations sur la possession ou non par les parasites de ce genre de plastide lors de leur évolution (revu dans [98]). L'ordre des *Kinetoplastidae* contient cinq familles. Les quatre premières, *Prokinetoplastida*, *Neobodonida*, *Parabodonida* et *Eubodonida* représentent les *Bodonidae* (Figure 3-2). Ces protozoaires biflagellés sont pour la plupart monogéniques et non parasitaires. Parmi les exceptions, *Cryptobia* parasite des systèmes digestifs ou reproducteurs de quelques invertébrés (sangue, mollusque) et est commensal sur la peau ou dans l'intestin de poissons. *Ichthyodo* est un ectoparasite des poissons téléostéens et *Trypanoplasma* est un parasite sanguin de poissons et d'urodèles transmis par les sangsues [96]. Aucun cas de parasitisme par les *Bodonidae* n'a été observé sur les plantes et les vertébrés terrestres, mais les études phylogénétiques supportent l'idée que les *Trypanosomatidae* descendent des ces organismes [97]. La cinquième sous-classe, les *Trypanosomatidae*, descend des *Eubodonida* et regroupe des espèces parasitaires monoflagellées, dont *Trypanosoma* spp. et *Leishmania* spp. Ces parasites infectent une grande variété d'hôtes, notamment chez les mammifères (hommes, animaux domestiques, bétail, rongeurs, etc). Parallèlement à ces parasites digéniques, les *Trypanosomatidae* contiennent également des parasites monogéniques vivant dans les organes digestifs d'une grande variété d'insectes (*Crithidia*, *Herpetomonas*, *Leptomonas*). Les *Phytomonas* sont responsables du parasitisme des plantes à fleur, ciblant soit le latex, le phloème ou les fruits et graines. Ces parasites végétaux possèdent à eux seuls une grande diversité d'organisation de leurs génomes [99].

Parasitisme et commensalisme ont donc été acquis à différentes reprises au cours de l'évolution. Puisque la séparation des genres *Trypanosoma* et *Leishmania* est datée entre 400 et 600 millions d'années [100-102], leur divergence était bien antécédente à l'apparition des mammifères il y a 165 millions d'années [103], ainsi que celle des premiers insectes ailés il y a 300 millions d'années [1]. Les premiers cycles digéniques chez les

trypanosomes remontent peut-être à l'apparition des vertébrés, les sangsues ayant servi de vecteur pour infecter poissons et amphibiens [104]. L'arrivée des phlébotomes et la présence de mammifères laissent sous-entendre que les premières leishmanioses sont apparues il y a 50 millions d'années [1], infectant les primates depuis ces dernières 15 millions d'années. Les *Trypanosomatidae* ont donc eu le temps de développer une multitude de mécanismes de résistance et d'adaptation chez les mammifères.

## **3.2. Organisation cellulaire des *Trypanosomatidae*.**

### **3.2.1. Organisation générale et organites.**

Une cellule de *Trypanosomatidae* possède plusieurs particularités propres à cette famille (Figure 3-3). Cependant, une grande variabilité morphologique est observable entre les différentes espèces, et même durant le cycle de vie d'une même espèce. Le premier constat de l'organisation de ces cellules est l'abondance d'organites en copie unique. Les *Trypanosomatidae* ont un seul flagelle, une seule mitochondrie, un seul appareil de Golgi face à la poche flagellaire et un seul noyau dont la membrane externe est fusionnée au réticulum endoplasmique. Cette organisation sous-entend une régulation précise de la réplication des organelles lors de la division cellulaire.

Parmi les petits organites présents chez les *Trypanosomatidae*, on retrouve les peroxisomes, organites ubiquitaires des cellules eucaryotes et impliquées dans le catabolisme des acides gras. Les glycosomes sont spécifiques de cet ordre et leur évolution est liée à celle des peroxisomes. La taille et le contenu enzymatique des glycosomes varient en fonction des espèces. On y retrouve habituellement le contenu des peroxysomes, mais également les sept premières enzymes de la glycolyse et les deux premières étapes du métabolisme du glycérol. L'utilité de cet organite est principalement de concentrer les intermédiaires réactionnels de la glycolyse et d'en augmenter la productivité. Il est estimé que le volume cumulé des glycosomes dans les cellules représente 4% du volume total [105].

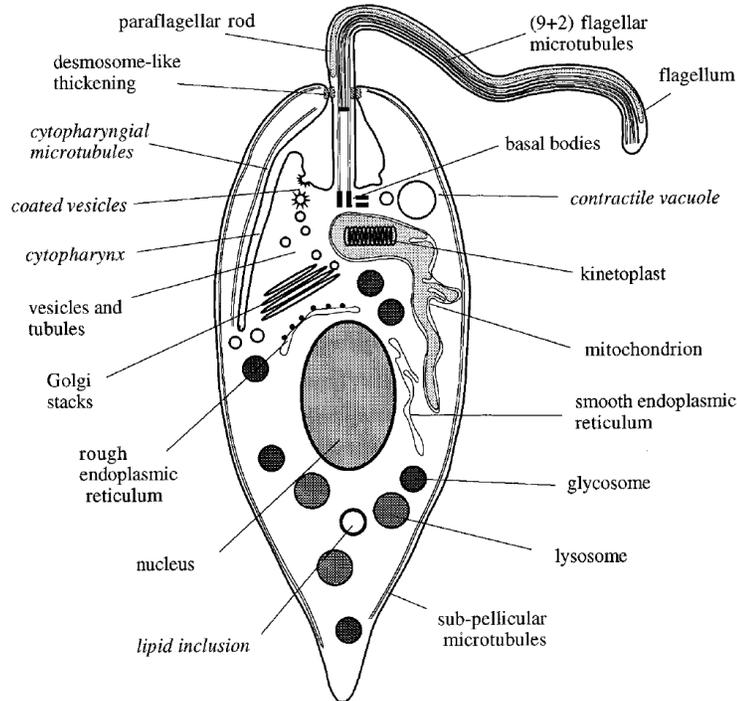


Figure 3-3 : Schéma général de la structure cellulaire des *Trypanosomatidae*.

Les légendes en italique sont spécifiques de certaines espèces (Adapté de [106]).

Les lysosomes sont des organites acides communs à tous les eucaryotes. Ils contiennent des enzymes de dégradation permettant d'assurer le renouvellement protéique. Les acidocalcisomes constituent une seconde classe de vacuoles acides chez les *Trypanosomatidae*. Contrairement aux lysosomes, les acidocalcisomes ne présentent pas d'activité protéase et semblent être un simple organite de stockage contenant les principales réserves de la cellule en phosphore, calcium, sodium, zinc ou magnésium [107]. Leur membrane possède plusieurs systèmes de pompes et d'efflux autorisant la circulation des ions, permettant sans doute aux microorganismes de s'adapter à des stress environnementaux [108].

### 3.2.2. Le noyau et le nucléole.

L'organisation du noyau des *Trypanosomatidae* a récemment fait l'objet d'une revue de littérature complète et bien illustrée [109]. Dans cet organite, on retrouve une compartimentation de la transcription et de la réplication, mais quelques différences sont observables chez les *Kinetoplastidae*.

Le nucléole est la principale structure identifiable dans les noyaux eucaryotes. Sa principale fonction est la transcription des ARN ribosomaux et l'assemblage des ribosomes, mais il intervient également dans d'autres processus cellulaires comme la régulation de la mitose, la progression du cycle cellulaire, la biogénèse de particules ribonucléoprotéiques et la réponse à différents types de stress tels que choc thermique, hypoxie ou dommages de l'ADN (revus dans [110]). Chez l'homme, la réponse au stress la mieux caractérisée impliquant cet organite passe par la stabilisation de la protéine p53 et ainsi l'arrêt du cycle cellulaire. Le nucléole agit ainsi tel un senseur de stress. D'autres mécanismes de réponse au stress via le nucléole ont été mis en évidence, notamment par la modification de l'expression des ARNr. Chez *Leishmania* et *T. cruzi*, il a été noté qu'un stress thermique ou une interruption de la transcription de l'ARN polymérase II par l'actinomycine D entraînent la relocalisation nucléolaire de protéines de liaison aux ARNm, pourtant cytoplasmiques dans des conditions physiologiques [111, 112]. Le rôle de cette relocalisation n'est pas encore bien défini, et celle-ci n'est pas conservée chez *T. brucei*.

Au moins chez les amniotes, le nucléole a une structure tripartite. Le centre fibrillaire semble conserver les ARNr inactifs. Il est entouré par un composant fibrillaire dense dû à l'accumulation de ribonucléoprotéines et jouant un rôle dans les premières étapes de maturation des ARNr. La finalisation de la maturation des ARNr et l'assemblage des ribosomes sont effectués dans le compartiment granulaire sous-jacent. Les eucaryotes non amniotiques ont dans leur vaste majorité un nucléole bipartite composé d'un corps fibrillaire et d'un compartiment granulaire. C'est une structure bipartite qui est identifiable chez *T. brucei* [109].

Au moins chez *T. brucei*, un corps nucléaire spécifique des trypanosomes est identifiable : l'ESB (pour *expression site body*). Il correspond au site de transcription des ARNm procycline par l'ARN polymérase I et diffère du nucléole par l'absence des fibrillarines, protéines généralement utilisées comme marqueur de cette structure.

Les sites de transcription et de maturation des SL-RNA (pour *splice leader RNA*, voir Chapitre 4) correspondent à une concentration des ARN polymérase II au niveau des *loci* des SL-RNA et se localisent également dans des complexes impliqués dans leur maturation. Le SL-RNA étant impliqué dans la maturation de tous les ARN messagers chez

les *Kinetoplastidae*, il est normal de retrouver une transcription très active, la possible existence d'une compartimentation de leur production semble logique.

Les besoins en épissage des ARNm diffèrent largement chez les trypanosomes (voir Chapitre 4). Les speckles, corps contenant les facteurs d'épissage, ne sont pas clairement caractérisés. Certains facteurs d'épissage semblent se concentrer dans des structures similaires aux speckles mais d'autres tels que les protéines Lsm entourent le nucléole, suggérant qu'il n'y a pas de compartimentation de l'épissage des transcrits chez les trypanosomes.

Les corps de Cajal impliqués dans la biogénèse de complexes nucléaires ribonucléoprotéiques et les corps PML (pour *Promyelocytic leukemia protein*) potentiellement impliqués dans une régulation de la transcription ne semblent pas être présents chez les trypanosomes.

### **3.2.3. La mitochondrie.**

Les *Kinetoplastidae* tiennent leur nom de la présence dans leur unique mitochondrie de l'ADN mitochondrial (ADNmt ou ADNk pour kinétoplastique) qui est centralisé dans une région appelée le kinétoplaste. Tandis que le corps de la mitochondrie s'étend tout au long de la cellule, le kinétoplaste est concentré et physiquement lié au corps basal flagellaire. Le réseau d'ADNk est en liaison avec le flagelle par des filaments qui traversent la membrane mitochondriale au niveau du complexe TAC (pour « *Tripartite attachment complex* ») [113].

Le réseau d'ADNk a une organisation et une architecture uniques. Deux types d'ADN circulaires contiennent l'information génétique pour la synthèse de sous-unités de la chaîne respiratoire de la mitochondrie. Entre 25 et 50 maxicercles contiennent une vingtaine de gènes [114] et 5 000 à 10 000 minicercles portent chacun 1 à 5 ARN guides impliqués dans l'édition des ARN [115]. Minicercles et maxicercles sont enchevêtrés entre eux, permettant de garder le réseau compact, et sont stabilisés par plusieurs homologues d'histones, appelés KAP (pour « *Kinetoplast associated proteins* »).

L'expression des gènes mitochondriaux est complexe et coûteuse en énergie (revu dans [116, 117]). L'édition des ARN concerne deux tiers des transcrits mitochondriaux. Son but est de créer des cadres de lecture exploitables pour la traduction, et 3 processus ressortent : le rétablissement du cadre de lecture par ajout ou délétion de nucléotides, la génération de codon de démarrage AUG et d'autres codons du côté 5' par ajout de triplets de nucléotides [118], et enfin la quasi création d'un cadre de lecture complet à partir d'un cryptogène, également appelé *pan-editing*. Un exemple spectaculaire de *pan-editing* est celui du gène *coxIII* chez *T. brucei* où 50% du transcrit est modifié, par ajouts localisés de 547 uridines et délétions de 41 autres [119]. Ce mécanisme est d'autant plus complexe qu'il est sujet à des phénomènes d'édition alternative, permettant l'obtention de nouveaux cadres de lecture et une régulation de l'expression des gènes [117].

Outre son modèle mécanistique, quelques particularités de la mitochondrie sont spécifiques des *Kinetoplastidae*. La dyskinetoplastie [120] correspond à une altération ou une perte d'ADNk, et donc provoque l'incapacité des cellules à effectuer les phosphorylations oxydatives. C'est observable chez *T. equiperdum* et *T. evani* qui ont d'ailleurs perdu la forme insecte de leurs cycles de vie. Les stades sanguins survivent puisqu'ils utilisent majoritairement la glycolyse comme source d'énergie, et ces formes se transmettent directement d'un hôte vertébré à un autre. La transkinetoplastie [121] définit les changements colossaux dans les populations de maxicercles et de minicercles induits par des drogues telles que les arsénites, la tunicamycine et la pentamidine. Enfin, une observation des plus inattendues a montré la présence de fragments d'ADNk de minicercles à l'intérieur de rétrotransposon LINE-1 ou dans le locus du gène de la  $\beta$ -globine lors d'infections par *T. cruzi* de patients humains, de poulets ou de lapins [122]. Aucun mécanisme n'a été décrit pour expliquer le passage d'ADN de la mitochondrie d'un parasite cytoplasmique vers le noyau des cellules de l'hôte. Enfin, l'étude comparative de l'édition des ARNm au cours des stades de vie de *T. brucei* a montré que l'édition de certains transcrits est régulée en fonction du stade de vie parasitaire [123].

### **3.2.4. Flagelle et poche flagellaire.**

#### **3.2.4.1. La poche flagellaire**

Le cytosquelette cellulaire est composé de microtubules sub-pelliculaires permettant de maintenir la forme du parasite. L'émergence de l'unique flagelle des *Trypanosomatidae* est effectuée au niveau d'une zone dépourvue de ces microtubules, et est accompagnée d'une invagination appelée poche flagellaire [124]. Tous les trafics vésiculaires entrants (endocytose) ou sortants (sécrétion ou exocytose) de la cellule sont effectués au niveau de la poche flagellaire qui est un véritable portail dynamique entre la cellule et l'environnement extérieur [125]. Quand l'organisme en est pourvu, le cytopharynx peut agir comme un espace d'échange alternatif, notamment chez *T. cruzi* [107]. Certains *Trypanosomatidae*, comme *Leishmania*, *T. cruzi* ou *Crithidia fasciculata*, possèdent une vacuole contractile qui est reliée au milieu extracellulaire par l'intermédiaire de cette poche flagellaire. La vacuole contractile permet au parasite de réguler sa diminution de volume lors d'un stress osmotique, comme par exemple au moment du changement d'hôtes [126], et elle utilise les réserves des acidocalcisomes pour fonctionner. Le protéome de cette vacuole a été récemment caractérisé et suggère notamment son rôle dans le métabolisme des phosphates [127].

#### **3.2.4.2. La membrane flagellaire**

En lui-même, le flagelle est composé de trois parties dont la composition est spécifique et diffère du reste de la cellule : la membrane flagellaire, le compartiment soluble (ou matrice) et le cytosquelette (l'axonème) [128]. Contrairement aux protéines de la matrice et celles de l'axonème qui sont importées par le transport intraflagellaire (IFT, voir ci-dessous), les protéines de la membrane flagellaire sont synthétisées dans le réticulum endoplasmique et directement adressées aux membranes par l'appareil de Golgi.

La présence du flagelle définit trois domaines membranaires, de composition et de rôle différents : la membrane plasmique pelliculaire (pour le corps cellulaire), la membrane flagellaire et la membrane de la poche flagellaire. La membrane flagellaire présente une composition unique, enrichie en stérols, glycolipides et sphingolipides, qui forment un radeau lipidique [129]. Si *T. brucei* est capable de se nourrir du cholestérol présent dans la circulation sanguine de son hôte mammifère, il est également capable de synthétiser son

propre ergosterol dans ses autres stades de vie, formant des radeaux plus larges et plus thermostables que ceux composés de cholestérol [130]. Un radeau agit comme une plateforme pour les signaux de transduction et recrute pour cela plusieurs types de protéines, notamment celles présentant une double acylation, une ancre glycosylphosphatidylinositol (GPI), ou des domaines transmembranaires. Par exemple, une protéine flagellaire de liaison au calcium chez *T. cruzi* [131] présente une double acétylation (myristilation et palmitoylation) et se localise au niveau de la membrane flagellaire. Chez *Leishmania*, les petites protéines myristilées (SMP) se localisent différemment en fonction de leurs modifications post-traductionnelles et de leur séquence d'acides aminés. SMP-1 est diacylée (myristilation et palmitoylation) et est l'un des constituants majeurs de la membrane flagellaire [132], tandis que SMP-2, monoacylée, réside dans la poche flagellaire et SMP-4 est assignée à la membrane pelliculaire par son contenu en acides aminés [133]. Toutes les protéines diacylées ne sont cependant pas toujours dirigées vers la membrane flagellaire. Plus récemment une étude menée sur les transporteurs de glucose chez *L. mexicana* (*LmexGT*) a mis en évidence l'importance de la séquence d'acide aminés NPM, ainsi que son contexte dans la protéine, pour permettre leur import dans la membrane flagellaire. *LmexGT1* est naturellement importée dans la membrane flagellaire, tandis que les paralogues *LmexGT2* et 3 restent au niveau de la membrane pelliculaire [134].

#### **3.2.4.3. Matrice et cytosquelette flagellaire**

Étant donné que cet organite ne dispose pas de sa propre machinerie de traduction, tous les composants protéiques de la matrice ou du cytosquelette sont synthétisés dans le cytoplasme, puis importés dans le flagelle. Pour y entrer, les protéines doivent d'abord passer à travers le complexe du pore flagellaire au niveau de la poche flagellaire. Ce complexe agit telle une barrière sélective pour le passage des protéines, créant un microenvironnement spécifique de l'organite [135]. En plus de cela, un mécanisme de contrôle de la qualité des protéines transitant vers l'organite a également été décrit au moins pour la tubuline chez *T. brucei*, évitant l'incorporation de protéines abérantes [136]. Les mécanismes impliqués dans le tri et le transport des protéines nouvellement synthétisées vers le flagelle sont encore très peu compris, aussi bien chez les *Kinetoplastidae* que pour les cils primaires des eucaryotes supérieurs. Quelques séquences peptidiques ont été identifiées

chez les eucaryotes pour adresser les protéines vers les cils primaires (revues dans [137]) et diffèrent de celles retrouvées chez les *Kinetoplastidae* pour le moment (revues dans [138]). Des protéines telles que *TbRab23* (impliquée dans le transport membranaire) [139], la tyrosine phosphatase *TrPTP2* [140], une petite protéine G atypique Rab-like 5 [141] ont été localisées dans le flagelle. De plus en plus de travaux de protéomique sont effectués sur flagelles et cils primaires, afin de mieux les comprendre [142, 143]. Une étude globale du protéome de la surface flagellaire et également de sa matrice chez la forme infectieuse de *T. brucei* a été réalisée [144], soulignant la diversité fonctionnelle des protéines retrouvées dans cet organelle (Tableau 1). Parmi les protéines identifiées, 16% des protéines sont spécifiques à *T. brucei* et 21% sont soumis à une régulation en fonction du stade de vie parasitaire, leurs ARNm étant au moins deux fois plus abondant dans la forme sanguine du parasite.

Tableau 1 : Type de protéines retrouvées dans la surface et la matrice flagellaire chez la forme sanguine de *T. brucei*.

Fonction putative	Surface flagellaire	Matrice flagellaire
Cyclases	17	0
Lyases	18	0
Transporteurs spécifiques de substrats	12	0
Transporteurs transmembranaires	15	0
Liaison aux nucléotides	31	117
Hydrolases	0	117
Liaisons aux protéines	0	54
Oxidoreductases	0	52
Liaisons de co-facteurs	0	20
Facteurs de traduction, liaison aux acides nucléiques	0	12
Liaisons aux acides carboxyliques	0	5
Peroxydases	0	4
Liaisons aux carbohydrates	0	4
Non annotées	83	350
Absentes de la banque de données DAVID	22	46
<b>Total</b>	<b>158</b>	<b>666</b>

Adapté de [144].

Le cytosquelette flagellaire contient un axonème composé d'une paire de microtubules centraux entourés de 9 doublés de microtubules [145] (axonème 9+2, Figures 3-3 et 3-4c-e). Le corps basal est situé à l'extrémité de l'axonème, il est composé de neuf triplets de microtubules et est analogue au centriole des mammifères [146]. La longueur du flagelle varie en fonction du stade de vie du parasite, il est notamment atrophié chez les formes amastigotes de *T. cruzi* et *Leishmania* (Figure 3-4b). Le *paraflagellar rod* (PFR, Figure 3-4d) est positionné le long de l'axonème. C'est un élément essentiel pour permettre le battement du flagelle, mais son mécanisme d'action exact est toujours discuté [147]. Son expression peut varier en fonction des stades de vie et chez *Leishmania*, au sein de l'hôte mammifère, le flagelle comporte uniquement l'axonème atrophié, le PFR n'étant pas exprimé.

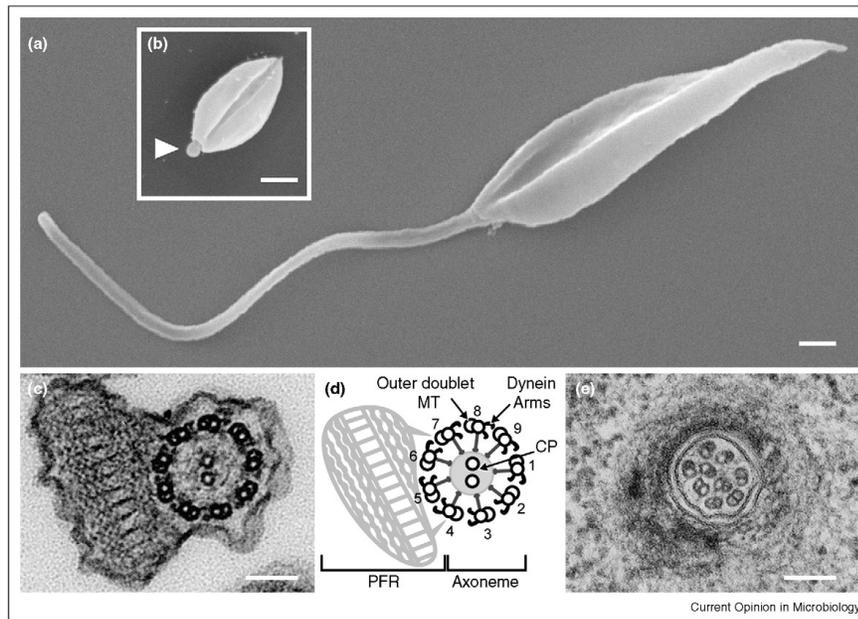


Figure 3-4 : Structure du flagelle de *Leishmania* chez les promastigotes et les amastigotes.

Micrographes électroniques montrant les différences de la taille flagellaire entre *L. mexicana* promastigote (a) et amastigote axénique (b). Les différences sont d'autant plus visibles sur des coupes transversales entre promastigote (c) et amastigote (e) où l'on note la présence du moteur PFR uniquement dans le stade promastigote, comme schématisé en (d). Échelle : 1  $\mu\text{m}$  (a, b) et 0,1  $\mu\text{m}$  (c-e) (Tiré de [148]).

Outre son rôle évident dans le déplacement, plusieurs fonctions sont décrites pour cet organelle en fonction des types cellulaires qui le portent (revu dans [137] pour les cils des eucaryotes). Chez les *Trypanoplatidae*, le flagelle permet la fixation des parasites dans l'appareil digestif des insectes, la chimiotaxie, la signalisation cellulaire, l'invasion des cellules hôtes et il est également suggéré qu'il possède un rôle sensoriel pour capter les modifications du milieu extérieur [138, 145].

Le transport intraflagellaire est un mécanisme décrit chez les ciliés et les flagellés, permettant le passage de cargos protéiques dans le compartiment soluble du flagelle. Ce transport est effectué par des kinésines (transport antérograde, du cytoplasme vers l'apex du flagelle) ou des dynéines cytoplasmiques (transport rétrograde) [138, 149, 150]. Il est suggéré que la régulation de la taille du flagelle est dépendante de l'équilibre existant entre transport antérograde et rétrograde des composants de l'axonème [148]. Les gènes codant pour les protéines moteurs et du IFT sont conservés dans la plupart des cellules eucaryotes ciliées, dont *Leishmania* et *Trypanosoma* [151, 152] et l'inhibition du transport perturbe la formation de l'organite. Une étude récente a permis une approche quantitative du IFT chez *T. brucei* [153]. Le transport rétrograde à une vitesse uniforme de 5µm/s, alors que deux types de transport antérograde sont différenciables en fonction de leur vitesse (autour de 1.5µm/s et 2.4µm/s). La température influence la vitesse de ce trafic, ce qui présente un aspect important pour les parasites digéniques passant d'un vecteur poïkilotherme à un hôte thermostable. Il est estimé que seul 45% des protéines du transport flagellaire circulent dans le flagelle, le reste étant cytoplasmique et fortement concentré au niveau de la base du flagelle, dû notamment au recyclage des protéines de l'IFT, et pouvant jouer des rôles annexes dans le contrôle du cycle cellulaire ou le positionnement des fuseaux mitotiques. Parmi les voies de signalisation transitant par les flagelles ou cils primaires des mammifères, la voie de signalisation *hedgehog* est sans doute la mieux décrite [154, 155], présentant un mécanisme d'import du facteur de transcription Gli dans les cils primaires, où il est phosphorylé, pour ensuite être dirigé vers le protéasome lors de son retour dans le cytoplasme. La digestion partielle qu'il y subit lui confère un rôle de répresseur de transcription. Ces structures jouent donc un rôle évident dans la signalisation de la régulation de l'expression génique.

Chez les *Trypanosomatidae*, les calflagines sont des protéines naturellement présentes dans la membrane flagellaire, liant le calcium et fonctionnant potentiellement comme des protéines signals. Chez *T. brucei*, le niveau d'expression de ces protéines est 10 fois plus élevé lorsque le parasite est dans la circulation sanguine de son hôte mammifère et leur inhibition est responsable d'une chute de la parasitémie, démontrant l'implication de ces protéines, et donc de la structure flagellaire dans la virulence parasitaire [156].

## **Chapitre 4 : Organisation du génome des *Trypanosomatidae* et régulation de l'expression génique.**

L'année 2005 a été un tournant pour la compréhension des génomes des *Trypanosomatidae* et leurs régulations, grâce à la publication des génomes de *Trypanosoma brucei* [157], *Trypanosoma cruzi* [158] et *Leishmania major* [159]. Depuis, d'autres organismes ont été séquencés, dont *L. infantum* et *L. braziliensis* [160], *L. mexicana* [161], ou le *sauroleishmania L. tarentolae* [162]. L'organisation du génome et la régulation de l'expression génique qui en découle sont conservées chez ces parasites et uniques à cette famille. Des études de séquençage ont été initiées chez le *Kinetoplastidae* non parasitaire *Bodo saltans* [163], la finalisation de ces travaux serait une information intéressante pour tenter de comprendre l'acquisition du parasitisme chez les *Trypanosomatidae*.

### **4.1. Le génome des *Trypanosomatidae*.**

#### **4.1.1. Organisation du génome et transcription polycistronique.**

Au sein même du genre *Leishmania*, une première différence est observable au niveau du nombre de chromosomes. Les espèces de l'Ancien Monde (*L. major* et *L. infantum*) comptent 36 chromosomes tandis que celles du nouveau monde telle que *L. braziliensis* en comptent 35. Cette différence est due à la fusion des chromosomes 20 et 34 [164]. Le génome des *Trypanosomatidae* présente une organisation unique où jusqu'à plusieurs centaines de gènes se succèdent dans la même orientation, définissant une unité polycistronique [159, 165]. Au total, le génome de *L. major* comporte 133 unités polycistroniques, dont la taille peut aller jusqu'à 1259 kb. Elles sont séparées par des régions de changement de brin, allant de 0.9 à 1.4 kb riches en A et T, et peuvent être convergentes ou divergentes en fonction du sens de lecture de l'ARN polymérase II (ARNP II) [166]. Le chromosome 1 de *L. major* est présenté en guise d'exemple sur la Figure 4-1. Il semble aujourd'hui accepté par la communauté scientifique que l'initiation de la transcription est effectuée de manière bidirectionnelle à partir des régions de changement de brin divergentes [167]. Elle se termine à l'arrivée des télomères, ou plus généralement au niveau des régions de changement de brin convergentes, qui contiennent souvent les

ARN de transfert (ARNt), des ARN ribosomiaux (ARNr) ou de petits ARN nucléaires (ARNsn) [168]. La transcription n'est cependant pas encore très bien comprise et définie, une activité transcriptionnelle étant également détectable dans le sens inverse du cadre de lecture [167].

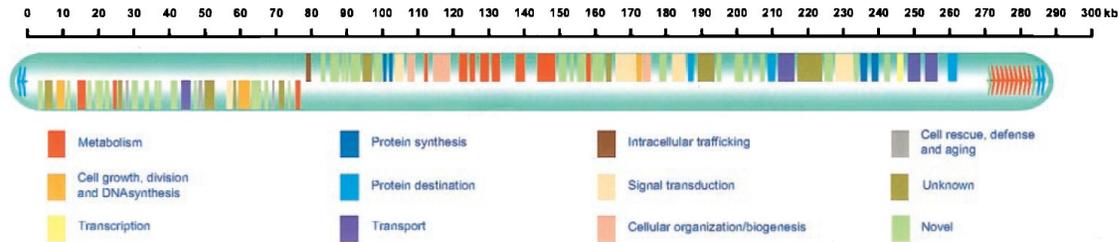


Figure 4-1 : Organisation des gènes sur le chromosome 1 de *Leishmania major*.

Le chromosome 1 de *L. major*, contient 79 gènes sur un total de 257 kb, une seule région de changement de brin divergente est présente. Les 50 gènes en aval de cette région sont tous sur le brin +, alors que les 29 gènes en amont sont tous sur le brin -. Un code couleur est attribué en fonction des annotations des séquences codantes. Tiré de [165].

Les polycistrons des *Kinetoplastidae* ne semblent pas grouper des protéines de fonction commune, comme décrit dans les opérons de bactéries ou de nématodes, ce qui exclut la régulation transcriptionnelle d'une voie métabolique particulière. Il a été proposé qu'un fort taux de recombinaison sur les chromosomes de ces organismes a engendré cette organisation qui apparaît aujourd'hui aléatoire [169]. Chez les *Trypanosomatidae*, les gènes des protéines fortement exprimées sont souvent présents en plusieurs copies répétées sur le génome. Cette plasticité génomique permet, à long terme, l'augmentation de la quantité de protéines par augmentation du nombre de gènes [169], mais ne permet pas la régulation ponctuelle de l'expression des gènes durant les cycles cellulaires ou parasitaires.

Chez les eucaryotes, les ARNP I, II et III contiennent respectivement 14, 12 et 17 sous-unités protéiques, avec chez la levure un corps commun de 5 sous-unités et de 6 polypeptides conservés. Ces trois classes d'ARN polymérase ont été identifiées chez les *Trypanosomatidae*, cependant leurs ARN cibles sont moins bien catégorisés que chez les autres eucaryotes. En effet, certains ARNm normalement transcrits par l'ARNP II, peuvent être transcrits par l'ARNP I chez ces organismes, dans le cas de l'ARNsn du miniexon (voir partie 4.1.3), celui-ci est transcrit par l'ARNP II (revu dans [170]) tandis que les autres

ARNsn sont traités par l'ARNP III. Les ARNP de ces parasites possèdent 9 des 11 sous-unités cœur de la levure, mais plusieurs homologues de composants spécifiques de certaines ARNP sont absents [159, 169, 171]. De même, peu d'homologues putatifs des facteurs de transcription identifiés chez les eucaryotes ont été retrouvés chez ces parasites, soulignant leur faible modularité de la transcription. Les mécanismes d'initiation de la transcription chez les *Trypanosomatidae* sont encore inconnus, seuls quelques promoteurs ont été décrits, presque exclusivement pour les ARNP I et III [172]. Le fonctionnement de l'ARNP II, qui traite les ARNm, reste encore un mystère. Le promoteur du minixon est l'unique cible caractérisée pour cette polymérase [173]. En outre, les expériences de transfection utilisant des plasmides prouvent que l'ARNP II est capable d'initier son activité en absence de promoteurs spécifiques chez les *Kinetoplastidae* [172].

L'initiation de la transcription est potentiellement modulable au niveau du remodelage de la chromatine, physiologiquement ou en réponse aux stressés (revu dans [174]). Les *Trypanosomatidae* ont de multiples copies des histones H2A, H2B, H3, H4 et H1. Ces protéines, impliquées dans l'enroulement de l'ADN, influencent l'accès à plusieurs complexes liant l'ADN, dont ceux de la transcription. En accord avec cette hypothèse, les génomes de ces parasites codent aussi de nombreuses enzymes impliquées dans la modification des histones, notamment deux familles d'acétyltransférases, trois familles de méthyltransférases et les trois classes connues d'histones déacétylases, deux d'entre elles étant essentielles chez *T. brucei* [159]. Chez *L. major*, la majorité des protéines H3 acétylées est retrouvée au niveau des zones de changement de brin, avec une densité plus importante lorsque les parasites sont en phase logarithmique de croissance (où l'activité de transcription est donc plus élevée). Ceci suggère que l'acétylation de ces protéines peut réguler l'initiation de la transcription [175].

Les génomes des *Kinetoplastidae* [176] et également celui de l'algue *Euglena gracilis* [177], ont pour autre particularité de contenir une base modifiée, une  $\beta$ -D-glucosyl-hydroxyméthyluracil, appelée « base J ». Cette base, dont la synthèse est effectuée à partir d'une thymine par les thymines hydroxylases JBP1 et JBP2 [178], est retrouvée au niveau des séquences répétées, tels que les télomères et les régions de changement de brin. Il est estimé que les bases J remplacent 1% des bases T. Si *T. brucei* survit sans base J [179], la

suppression de JBP2 est par contre létale pour *L. tarentolae* [180]. Les bases J correspondraient à un signal de terminaison de la transcription par l'ARNP II [180].

Si peu de facteurs de transcription sont annotés, les génomes de ces parasites contiennent en contrepartie un grand nombre de protéines avec un domaine de reconnaissance de l'ARN (RRM), ou des domaines CCCH. Ces données, cumulées à la transcription polycistronique, sont en accord avec des mécanismes de contrôle majoritairement post transcriptionnels de l'expression génique [159].

#### 4.1.2. Comparaison du contenu du génome des *Trypanosomatidae*.

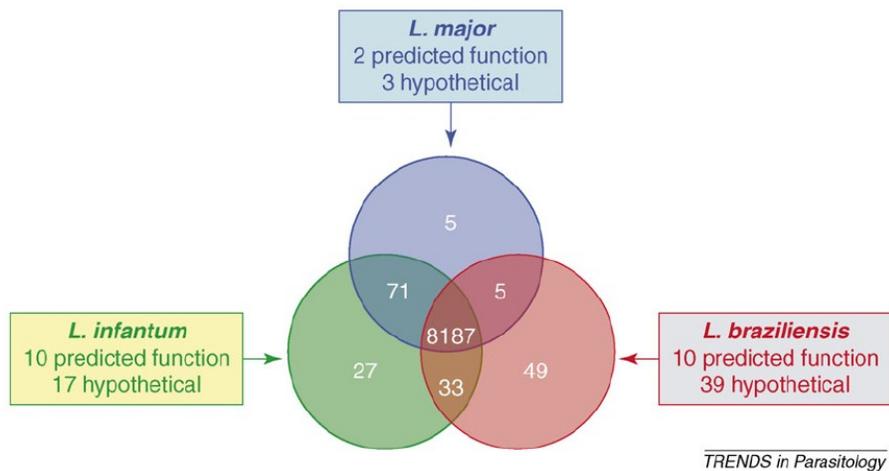


Figure 4-2 : Diagramme présentant la spécificité des gènes en fonction des espèces de *Leishmania*.

Tiré de [181].

Les *Trypanosomatidae* contiennent 8 000 à 12 000 gènes, avec un cœur commun d'environ 6 200 séquences codantes [160]. Les génomes ont une bonne conservation de la synténie malgré ces 200-500 millions d'années de séparation, puisque 68% du génome de *T. brucei* et 75% de celui de *L. major* restent dans le même contexte génétique [182]. Cependant, au niveau architectural, les chromosomes de *Leishmania* spp. diffèrent de ceux des *Trypanosoma* car ils n'ont pas de régions subtélomériques étendues contenant les gènes spécifiques de l'espèce [160]. Plus de 1 000 gènes sont pourtant uniques à *Leishmania* spp. et sont soit distribués « aléatoirement » dans le génome, soit introduits au niveau des régions de changement de brin et des points de rupture de la synténie [159]. Malgré la séparation des sous-genres *L. Viannia* et *L. Leishmania* datée entre 20 et 100 millions

d'années, la conservation de la synténie touche plus de 99% des gènes entre les espèces *L. major*, *L. infantum* et *L. braziliensis*. Très peu de gènes sont uniques à une espèce donnée de *Leishmania* (Figure 4-2) : 5 gènes pour *L. major*, 27 pour *L. infantum* et 49 pour *L. braziliensis*, des chiffres faibles comparativement aux 8 187 gènes communs aux trois espèces. Les différences dans le tropisme de la maladie et les diverses manifestations cliniques de leishmanioses ne s'expliquent donc pas par le contenu génétique en lui-même, mais plus au niveau de mutations ponctuelles dans les gènes ou au niveau de différences dans la régulation de l'expression génique.

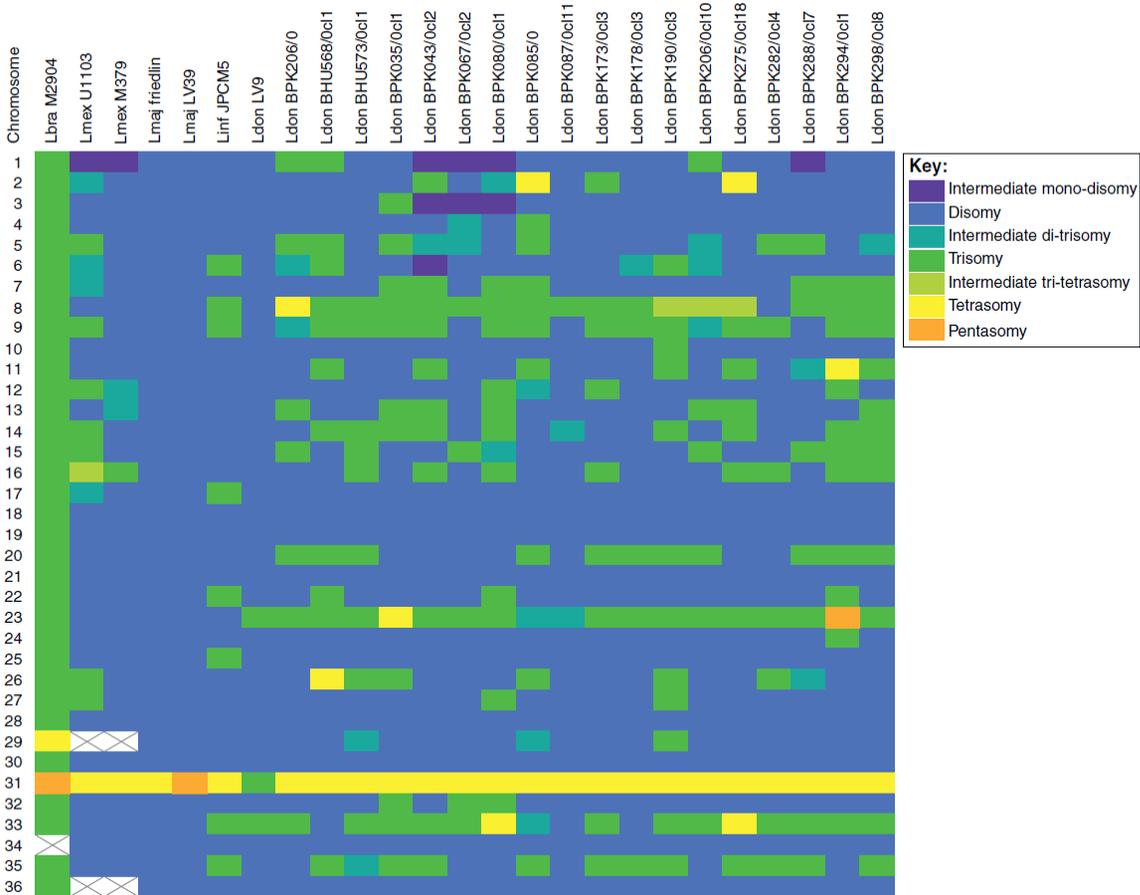
### **4.1.3. Plasticité génomique.**

#### **4.1.3.1. Amplification génique et aneuploïdie.**

Les études menées sur la résistance ou l'adaptation aux traitements contre *Leishmania* ont permis de mettre en évidence des mécanismes d'amplification génique spécifique de ce genre, permettant d'accroître le nombre de copies de gènes ou de groupes de gènes par la formation *de novo* de molécules d'ADN circulaires ou linéaires [183, 184]. Par exemple, l'ajout de méthotrexate, inhibiteur de la dihydrofolate réductase, stimule l'amplification d'un fragment d'ADN contenant le gène de cette enzyme. Si la surexpression du miniexon est associée à une perte de virulence chez *L. major* et *L. braziliensis* [185, 186], il a été aussi montré que l'amplification spontanée de la région contenant cette séquence présente des avantages de croissance [187]. L'amplification génique est possible grâce à la présence de séquences répétées de part et d'autres des gènes ciblés. Ceci est confirmé chez *L. tarentolae* puisque l'ajout de séquences répétées autour du gène *ptrI* (transporteur de ptéridines), et une pression de sélection au méthotrexate génère l'amplification du gène [188].

Les études de réarrangements chromosomiques et d'amplification ont été facilitées par les techniques de puces à ADN [189, 190]. Chez *L. major* et *L. infantum*, l'acquisition d'une résistance au méthotrexate est accompagnée d'une différence dans l'expression des gènes due à la modification du nombre de copies de certains gènes clés. Ces altérations génomiques sont variées, allant de la délétion de gènes à la formation d'ADN extrachromosomiques circulaires ou linéaires, voir même l'apparition d'une mono-, tri- ou tétrasomie. Globalement, lorsque des répétitions directes ou inversées, dont la séquence est

bien conservée entre ces deux espèces, sont présentes autour d'un gène subissant des pressions de sélection, une amplification est réalisée. Dans le cas contraire, une modification de la ploïdie est observable [190].



TRENDS in Parasitology

Figure 4-3 : Différence de ploïdie observée pour chaque chromosome en fonction des souches de *Leishmania* spp.

Tiré de [191].

La ploïdie varie beaucoup en fonction des espèces et des souches de *Leishmania* (Figure 4-3) [161]. Si peu de variations sont observables au niveau du contenu génique chez *Leishmania* spp., une grande hétérogénéité génétique peut provenir de l'amplification génique et des modifications de la ploïdie. La modification du nombre de copies d'un gène pourrait contribuer à la modulation de l'expression génique en réponse aux stress environnementaux rencontrés chez l'hôte, fournissant une nouvelle base pour élucider les différences de tropisme observables [161]. L'étude de ce phénomène est complexe, car

l'analyse du contenu génétique des parasites par hybridation *in situ* montre que dans une même population cellulaire, issu d'une même souche, chaque chromosome peut être présent à plusieurs états de ploïdie. Le contenu chromosomique varie donc grandement [192], générant une hétérogénéité génomique intra-souche constitutive, récemment nommée aneuploïdie en mosaïque [193].

#### **4.1.3.2. Les éléments transposables.**

Les éléments transposables sont des séquences d'ADN capables de se déplacer d'une fraction chromosomique à une autre, présents dans la plupart des organismes procaryotes et eucaryotes. Ils sont classés en deux catégories en fonction de leur séquence et de leur mode de transposition [194]. Les éléments transposables de classe I, ou rétroéléments ou rétrotransposons, se déplacent par un intermédiaire à ARN, la classe II se déplace par un intermédiaire à ADN. Les rétrotransposons sont subdivisés en deux groupes : les rétrotransposons qui disposent de *long terminal repeats* (LTR), dont la structure et le mécanisme de transposition sont équivalents à ceux des rétrovirus [195], et les rétrotransposons qui ne présentent pas de LTR (rétroposons), dont le modèle de transposition repose sur celui décrit pour l'élément R2 des insectes [196].

Ces éléments ne sont plus à considérer comme de l'ADN poubelle, égoïste ou parasite. Plusieurs études décrivent une domestication des éléments transposables, voir même une relation symbiotique avec les génomes dans lesquels ils nichent. Le mobilome permet à long terme la structuration des génomes [197], l'acquisition de nouvelles fonctions [198] ou la régulation de l'expression génique [199, 200]. Parmi ces éléments, les rétrovirus endogènes humains et leurs possibles participations dans la différenciation cellulaire, la régulation de l'immunité, la résistance aux infections ou la formation du syncytiotrophoblaste placentaire [201], sont une somme de données laissant place à plusieurs spéculations sur le rôle des éléments viraux dans l'évolution des cellules humaines [202]. Chez la grenouille des bois *Rana sylvatica*, le processus inhabituel d'« auto-congélation » est lié à la surexpression d'une protéine qui semble issue d'un rétrovirus endogène [203]. Ces évolutions semblent également mises en place chez les protozoaires, par exemple chez *Giardia intestinalis*, une famille de rétroposons localisée

dans les télomères produit des petits ARN sens et antisens, potentiellement impliqués dans l'interférence à l'ARN [204].

Chez les *Trypanosomatidae*, aucun élément de classe I n'a été identifié, mais ces génomes présentent plusieurs rétrotransposons de classe II [205]. VIPER est le seul rétrotransposon à LTR identifié chez *T. cruzi* et *T. brucei* [157, 206]. Les rétrotransposons constituent le groupe d'éléments transposables le plus important dans ces organismes (Figure 4-4), et correspondent à environ 3% du génome nucléaire chez *T. brucei*. Les rétrotransposons sont classifiés en fonction des espèces et de leur structure ou conservation. Tous les rétrotransposons se regroupent dans deux classes chez les *Trypanosomatidae* CRE et *ingi*.

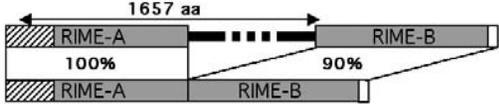
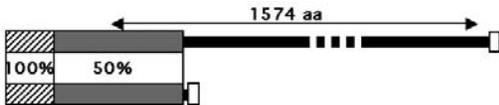
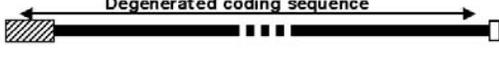
Name	Size (Kb)	Structure	Copy number per haploid genome		
			Tb	Tc	Lm
<i>ingi</i>	5.2		115 (3)	-	-
RIME	0.5		86	-	-
TbSIDER	0.57		26	-	-
L1Tc	4.9		-	320 (15)	-
NARTc	0.25		-	133	-
DIRE	~5		73 (0)	257 (0)	52 (0)
LmSIDER	~0.55		-	-	1858

Figure 4-4 : Types de rétrotransposons retrouvés chez les *Trypanosomatidae*.

Tiré de [205]

Le groupe CRE, présent chez *T. brucei* (SLAC), *T. cruzi* (CZAR) et *Crithidia fasciculata* (CRE1/CRE2), regroupe des rétrotransposons spécifiques d'un site et se retrouve au niveau du gène du SL-RNA [207-209]. Au sein du genre *Trypanosoma*, les rétrotransposons les plus nombreux constituent les *ingi* et RIME (Ribosomal Mobile Element, version tronquée des *ingi*) chez *T. brucei* et les L1Tc et NARTc (version tronquée de L1Tc) chez *T. cruzi* [210]. Ils sont dispersés dans le génome, mais semblent s'insérer au niveau de sites

spécifiques [211, 212]. *Ingi* et L1Tc codent pour une large protéine permettant la recombinaison [213-216]. Il a été démontré que les premiers 77 nt des rétroéléments de type L1Tc avaient une activité promotrice pour l'ARN polymérase II [217], permettant la synthèse de cette protéine, mais son activité dans la transposition des éléments mobiles n'a pas encore été démontrée. De plus, les L1Tc et NARTc possèdent une activité de type ribozyme, également dans les 77 premiers nucléotides de l'élément (revue dans [218]). Promoteur et activité ribozyme permettent ainsi la mobilité de ces éléments. Cette séquence de 77 nt constitue la région la plus conservée d'autres rétroposons éteints de *Trypanosomatidae* (voir ci-dessous), mais les L1Tc sont les seuls dont l'activité a été caractérisée à ce jour.

Les DIRE (pour *Degenerate ingi/L1Tc-Related Elements*) sont présents dans tous les génomes des *Trypanosomatidae*. Le fort degré de mutations a engendré des déplacements du cadre de lecture et l'apparition de codons stop prématurés, qualifiant ainsi ces éléments de dégénérés et inactifs [219].

*Ingi*, RIME, L1Tc, NARTc et DIRE partagent à leur extrémité 5' une séquence signature de 79 pb conservée [219]. La recherche d'autres éléments à l'aide de cette séquence signature a permis l'identification d'un autre groupe de rétroposons, appelés SIDER (pour *Short Interspersed degenerated retroposon*) [205]. Ils correspondent aux versions les plus dégénérées des éléments mobiles de *Trypanosomatidae* identifiés à ce jour. C'est dans le genre *Leishmania* que ces éléments sont le plus répandus (1858 *LmSIDER* contre 26 *TbSIDER*). L'analyse phylogénétique montre que ces éléments sont séparés en deux sous groupes [205], *LmSIDER1* (785 membres) et *LmSIDER2* (1073 membres) [205]. Les *LmSIDER2* contiennent une seconde séquence signature, juste en aval de la première, également de 79 pb. Le nombre élevé de ces séquences et le ratio SIDER1:SIDER2 sont également conservés chez *L. infantum* et *L. braziliensis* [220]. Entre 20 et 70 copies de SIDER1 de trypanosomes sont retrouvés uniquement chez les trypanosomes salivaires (*T. brucei*, *T. congolense* et *T. vivax*) mais ne sont pas retrouvés dans les autres espèces telles que *T. cruzi* [221].

Parmi les rétroposons, *ingi* et L1Tc ont toujours une activité mobile potentiellement active [219, 222], RIME et NARTc, aujourd'hui inactifs, semblent avoir eu des événements

de transposition récent. Par contre, les DIRE et SIDER sont eux éteints depuis longtemps [205]. Seuls certains sous-groupes de *TvSIDER1* présentent une bonne conservation de séquence, suggérant des transpositions récentes, sans doute effectuées à l'aide de la machinerie de transposition présente dans les *Tvingi* [221]. Si *Leishmania* n'a aujourd'hui plus aucun transposon actif, la présence de ces séquences dégénérées démontre que l'ancêtre de ce protozoaire hébergeait des rétroéléments actifs de type *ingi/LITc*, perdus au cours de l'évolution pour ne présenter aujourd'hui que des vestiges de ces éléments [219].

Les rétroposons se répartissent différemment dans le génome en fonction des espèces. *ingi*, RIME, *TbDIRE* et *TbSIDER* sont préférentiellement retrouvés dans les régions subtélomériques, puis dans les régions de changement de brin [205]. Chez *Leishmania*, les SIDER sont retrouvés dans les unités polycistroniques, et sont plus particulièrement prédits comme intégrés aux 3'UTR des ARNm [220]. L'inclusion de ces séquences spécifiquement dans le génome de *Leishmania* ne casse pas pour autant la synténie observée entre *L. major* et *T. brucei* [205]. Il a été démontré que la présence des *LmSIDER* dans le 3'UTR de certains gènes permettait de moduler l'expression génique, et la large expansion de ces séquences sous-entend une domestication des SIDER par *Leishmania* afin de réguler l'expression de ses gènes (voir partie 4.2.7.).

#### **4.1.4. Production d'ARNm matures par *trans*-splicing et polyadénylation.**

La production d'un transcrite polycistronique nécessite le clivage des différents cistrons pour obtenir des ARNm indépendants. Le clivage et/ou la fusion de fragments d'ARN est communément appelé l'épissage. On différencie l'épissage en *cis*, agissant sur un unique transcrite pour supprimer ses introns par exemple, et l'épissage en *trans*, permettant de lier deux molécules d'ARN d'origines différentes. Ces deux phénomènes coexistent chez les *Trypanosomatidae*.

Seuls deux gènes possèdent un intron chez ces parasites : une *poly(A) polymerase* [223] et une *ATP-dependent DEAD/H RNA helicase*. Les données obtenues par RNA-seq ont confirmé qu'il n'y avait pas d'autre transcrite contenant des introns chez *T. brucei* [224], limitant l'importance de l'épissage en *cis*. L'épissage en *trans* est lui obligatoire pour la maturation de tous les ARNm des *Trypanosomatidae* [225]. Il est utilisé pour ajouter la

coiffé du côté 5' du cadre de lecture, dans un mécanisme couplé à la polyadénylation du gène en amont (Figure 4-5), permettant l'obtention de transcrits individualisés et matures. La coiffe est portée par une séquence de 39 nucléotides appelée miniexon, encodée par un ARNsn appelé SL-RNA (pour *splice leader* RNA). Son ajout est réalisé par des réactions de *trans* estérification en deux étapes, similaires à celle de l'épissage en *cis*, mais passant par une structure en Y au lieu d'un intermédiaire en lasso [226-228]. L'épissage en *trans* avec le SL-RNA a été initialement découvert chez les trypanosomes, puis ensuite décrit dans plusieurs organismes incluant les *Euglenidae*, les nématodes, les trématodes jusqu'aux chordés les moins évolués, tels que les Urochordés [228]. Ce mécanisme n'est cependant pas décrit chez les hôtes mammifères de ces parasites, en faisant une cible de choix pour le développement de stratégies anti-parasitaires.

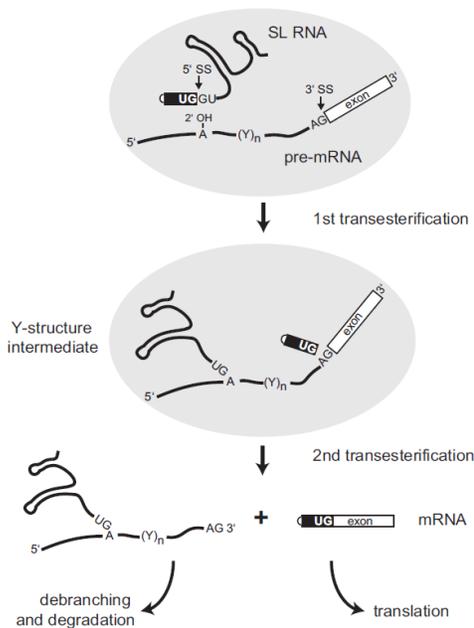


Figure 4-5 : Maturation des ARNm chez les *Trypanosomatidae* par épissage en *trans* et polyadénylation.

Tiré de [227].

Le SL-RNA est essentiel pour l'expression génique chez les *Trypanosomatidae* et l'inhibition de sa transcription est létale pour *Leishmania* [229]. Il est constitué d'une séquence en tige boucle de 39 nt chez *L. major*, suivi d'un intron supprimé lors du *trans* épissage. Le miniexon porte une coiffe de quatre nucléotides méthylés en aval du 7-

méthylguanosine donnant une structure appelée « cap4 », qui joue le même rôle que la coiffe des eucaryotes [226].

L'épissage fait intervenir le complexe protéique spliceosome. On retrouve ces complexes chez les *Trypanosomatidae*, composés autour des 5 petits ARN nucléaires riches en U (ARNsn U2 à 6). Si environ 170 facteurs additionnels sont caractérisés chez l'homme, peu de facteurs de l'épissage ont été identifiés chez les *Trypanosomatidae* à cause de la divergence de leur séquence nucléotidique. La constitution du spliceosome et sa séquence d'assemblage, ainsi que les protéines annexes primordiales telles que Sm et Lsm, ont été récemment revues dans [228]. Les grandes lignes de fonctionnement du spliceosome restent similaires à celles observées dans d'autres organismes. Curieusement, le complexe de l'ARNsn U1, impliqué généralement uniquement dans l'épissage en *cis* [230], est très fortement exprimé comparativement au nombre de gènes qui présentent des introns, suggérant une fonction annexe de ce complexe [231].

L'épissage en *trans* est couplé à la polyadénylation et les *Trypanosomatidae* partagent des séquences spécifiques au sein des régions intergéniques [232]. Une séquence étendue composée uniquement de pyrimidines (poly(Y), Figure 4-5) précède ou recoupe le site de clivage potentiel, constitué du dinucléotide AG [233, 234]. Les sites de polyadénylation sont plus difficiles à prédire car la séquence signal consensus AAUAA des eucaryotes est ici absente, et aucune autre séquence conservée n'a été mise en évidence pour cibler le clivage lors de la polyadénylation. Cette dernière est effectuée à une distance relativement fixe en aval du site d'épissage, de 100 à 400 nt en fonction des espèces [225].

Au final, l'épissage en *trans* permet l'obtention des ARNm matures, composée d'une coiffe en 5', d'un 5'UTR dont la taille moyenne est autour de 150 à 200 nt, du cadre de lecture, d'un 3'UTR de 1 500 nt en moyenne et d'une queue poly(A) [235]. Les séquences non traduites, majoritairement les 3'UTR, vont jouer un rôle majeur dans la régulation post-transcriptionnelle, permettant de palier l'absence de régulation transcriptionnelle de l'expression d'un gène. Si la transcription polycistronique est un choix coûteux en énergie pour le parasite, elle a pour avantage d'exprimer de manière constitutive l'ensemble du génome. Les adaptations du parasite aux stress environnementaux sont

supposément plus rapides, puisque la synthèse d'une protéine se fera directement en modifiant la stabilité des transcrits ou leur efficacité de traduction.

## **4.2. Régulation de l'expression génique au niveau de l'ARNm.**

### **4.2.1. Régulation de l'épissage des transcrits.**

L'efficacité de l'épissage dépend de la longueur et de la constitution de la séquence riche en polypyrimidine ainsi que celle présente dans le 5'UTR du gène épissé. Si la dégradation de l'ARN polycistronique précurseur est plus rapide que l'épissage, alors le transcrit sera faiblement exprimé [236]. L'arrangement des gènes sous forme de transcrits polycistroniques, et le couplage de la polyadénylation d'un gène avec l'ajout de la coiffe du suivant, engendre une quasi inexistence de régulation de la production individuelle d'ARNm. L'absence de maturation d'un transcrit se répercuterait sur le gène voisin, ou alors il nécessiterait un découplage de la polyadénylation / ajout de la coiffe, phénomène jamais décrit. La dispersion dans le génome des gènes codant une même fonction métabolique rend difficile la conception d'un système de régulation par non maturation des transcrits. Par contre les gènes de la cystéine peptidase B (CPB) font partie des exceptions, étant répétés en tandem dans le génome de *Leishmania*. Une étude menée sur la régulation de l'expression de CPB chez *L. mexicana* montre qu'une séquence riche en bases A et U réduit de 30 fois le niveau basal de l'ARNm chez les amastigotes [237]. Cette région est pourtant retrouvée uniquement sur le pré-ARN polycistronique, puisque elle est localisée en aval du site de polyadénylation, dans la région intergénique clivée lors de l'épissage en *trans*.

Les 18 paralogues du gène *HSP83* (pour *Heat shock protein*) sont également dupliqués en tandem dans le génome de *Leishmania*. L'absence des signaux de polyadénylation sur la dernière copie de ces gènes déstabilise le transcrit. Il a été montré que l'épissage de *HSP83* est plus efficace suite à un stress de température, voir même favorisé par rapport à d'autres transcrits [238]. Ce mécanisme diffère des observations chez d'autres eucaryotes, où un stress thermique provoque le désassemblage du spliceosome [239, 240], et inhibe donc l'épissage des transcrits. L'absence d'intron dans certaines HSP de *Drosophila melanogaster* permet de contourner ce problème. Puisque l'épissage en *cis*

n'est pas requis, l'expression des HSP peut être effectuée et favorisée lors d'une augmentation de la température.

Chez *T. cruzi*, il a été montré que les gènes en tandem codant pour les protéines UBP1 et UBP2, étaient maturés en deux étapes. Le premier passage du spliceosome omettait la maturation entre les deux gènes, formant ainsi un ARN bicistronique stable et gardé dans un stade de latence traductionnelle [241]. L'inhibition spécifique de la maturation de ce site est potentiellement effectuée par un homologue de la protéine de mammifère PTB (*polypyrimidine tract-binding protein*) [241], impliquée chez ces organismes dans la localisation des ARNm, la polyadénylation, la traduction et la régulation de l'épissage alternatif [242].

Une régulation de l'expression génique au niveau de l'épissage est donc possible, mais l'organisation du génome souligne que très peu de gènes peuvent être atteints par ce phénomène. L'épissage ne repose cependant pas sur une simple logique booléenne, et des diversités peuvent apparaître au niveau de la localisation des sites de maturation. Les récentes données obtenues par *RNA-seq* chez *T. brucei* ont montré une grande diversité de ces sites, puisque plus de 32 000 sites uniques de clivages en 3' ont été identifiés pour environ 8900 gènes [224, 243, 244]. Seuls 11% des transcrits ont finalement un site unique de polyadénylation [243]. La plupart des gènes a donc deux ou trois sites de clivage alternatifs, certains étant responsables parfois d'une perte du codon d'initiation de la traduction AUG, de l'inclusion ou l'exclusion de séquence signal ou de séquences régulatrices dans le 3'UTR, ou de la génération d'un cadre de lecture alternatif [227]. Ceci a également été observé dans la forme promastigote de *L. major*, où les résultats de *RNA-seq* indiquent qu'environ la moitié des transcrits existe avec des longueurs d'UTR variant parfois de centaines de nucléotides [245].

Chez *T. cruzi*, il a été démontré que l'ARN du gène *lyt1* (codant pour une protéine hémolytique) était différemment épissé en fonction du stade de vie du parasite [246]. L'initiation de la traduction doit alors être effectuée sur un second site AUG, tout en gardant le même cadre de lecture qu'avec le premier AUG, causant la perte des 28 premiers acides aminés. Cette différence provoque une modification de la localisation de la protéine

LYT1, la forme courte se retrouvant dans le kinetoplaste, la forme longue étant associée à la membrane plasmique et aux fonctions hémolytiques du parasite [247].

La polyadénylation alternative est un phénomène fréquemment observé chez les autres eucaryotes. La réduction du 3'UTR par polyadénylation alternative est associée à la prolifération et à la transformation, dues en partie à l'exclusion de site de liaison à des micros ARN (miRNA) dans les 3'UTR de proto-oncogènes [248, 249]. La réduction du 3'UTR est également décrite aux premières étapes du développement, pendant la reprogrammation de cellules souches pluripotentes, ou dans les tissus très prolifératifs comme dans les testicules [250, 251]. En contrepartie, les transcrits aux stades tardifs de développement et dans les cellules différenciées et non prolifératives tendent à inclure un 3'UTR plus long [252, 253]. Le rôle de la polyadénylation alternative dans l'inflammation, dans les neurones ou encore dans le développement, ainsi que les événements de régulation par modification de la taille de la queue poly(A) des ARNm ont fait l'objet d'une revue de littérature récente [254].

#### **4.2.2. Exportation des ARNm vers le cytoplasme.**

Le modèle habituel de l'exportation des ARN du noyau vers le cytoplasme implique les protéines exportines comme récepteurs, permettant le passage des transcrits par les pores nucléaires. Les exportines sont différentes en fonction des types d'ARN impliqués [255]. L'exportation des ARNm n'a pas encore été détaillée chez les *Trypanosomatidae*, mais l'existence d'une régulation à ce stade est fort probable [236]. Quelques études contradictoires ont été menées sur le rôle de l'exportine 1 dans la biogénèse du miniexon [256, 257]. Chez la levure, les ARNm sont pris en charge par l'hétérodimère Mex67/Mtr2, qui opère par son association avec le complexe TREX (Transcription / Export) permettant de coupler ces deux processus. TREX est notamment composé d'une ARN hélicase Sub2, récemment étudiée chez *T. cruzi* et *T. brucei* [258] où l'inhibition de l'expression de la protéine est létale et provoque une accumulation des ARNm dans le noyau. Chez *T. brucei*, *TbMex67* contient un motif en doigt de zinc essentiel pour la croissance du parasite, et interagit avec *TbMtr2* et une karyophérine, suggèrent la bonne conservation de ce système d'export des ARN chez les *Trypanosomatidae* [259].

### 4.2.3. La dégradation des ARNm.

La quantité de protéines produites ne dépend pas uniquement du taux de traduction de son ARNm, mais aussi de sa vitesse de synthèse et de dégradation. La vitesse de dégradation des ARNm varie souvent en réponse à un stimulus, permettant de rapidement adapter la quantité d'ARNm afin de corroborer aux besoins de la cellule. L'élimination des ARNm est cataloguée en deux classes : les mécanismes de contrôle de qualité, ou surveillance, qui éliminent la production de protéines non fonctionnelles et potentiellement toxiques, et les mécanismes augmentant ou diminuant la demi-vie des ARNm, affectant en conséquence l'abondance des protéines. Cinq types d'activités nucléolytiques participent à la suppression des ARNm : le retrait de la coiffe, la déadénylation, la dégradation de 5' vers 3', celle de 3' vers 5' et le clivage endoribonucléolytique. Leur recrutement et séquence d'action varient en fonction des ARNm ciblés et sont conditionnés par les facteurs régulateurs liés sur les ARNm. Les facteurs régulateurs se lient principalement au niveau de séquences spécifiques retrouvées dans les 3'UTR, et leur longueur imposante chez *Leishmania* et les autres *Trypanosomatidae* suggère la présence de multiples séquences régulatrices.

La coiffe en 5' et une longue queue poly(A) en 3' (de 250 à 300 nt) sont des éléments essentiels pour stabiliser les transcrits, ces structures se lient avec les protéines cytoplasmiques eIF4E (*eukaryotic initiation factor 4E*) et PABP (*poly(A) binding protein*) respectivement (voir la partie sur la traduction). Ces deux types protéiques interagissent avec le complexe eIF4F donnant un aspect circulaire à l'ARNm, le protégeant des activités exonucléolytiques et stimulant l'initiation de la traduction [260]. La dégradation exonucléolytique va devoir déstabiliser cette structure pour dépolymériser les transcrits. Si l'addition de la queue polyadénylée semble être effectuée par défaut lors de la maturation des transcrits, le contrôle de sa taille qui s'en suit est fortement régulé aussi bien dans le noyau que dans le cytoplasme, participant dans la modulation de la stabilité, le transport et la traduction des transcrits [254]. Il est généralement accepté qu'une longue queue polyadénylée est stabilisatrice des ARNm [261, 262], bien qu'une étude sur les ARNm de cellules de mammifères ait montré la présence d'une courte queue poly(A) d'une trentaine de résidus sur environ 25% des ARNm en cours d'expression [263]. Chez la levure et les métazoaires, plusieurs modèles ont été établis au regard de la dégradation des transcrits

(Figure 4-6), et ces modèles restent les références utilisées pour analyser les systèmes de dégradation des transcrits chez les *Trypanosomatidae*. La présentation qui va suivre suit donc ces modèles également, en mettant l'emphase sur les mécanismes étudiés chez les *Trypanosomatidae*.

#### 4.2.3.1. Mécanismes de dégradation des ARNm connus chez les eucaryotes.

Le mécanisme de dégradation le plus répandu reste dépendant de la déadénylation. Suite au désassemblage des structures eIF4E-PABP, une ou plusieurs déadénylases réduisent la queue poly(A) à environ 80 nt (Figure 4-6a). L'une des particularités de cette étape est sa réversibilité, une réadénylation permettant le retour des transcrits vers la traduction [260]. Deux voies de dégradation irréversibles sont ensuite possibles : soit la coiffe en 5' est retirée par un processus connu sous le nom anglais de *decapping*, permettant ensuite la dégradation du brin d'ARN par l'exonucléase 5' → 3' XRN1, soit le transcrit est dépolymérisé de 3' vers 5' par l'exosome, la coiffe étant hydrolysée par DcpS (*scavenger-decapping enzyme*) [264]. Ces deux voies peuvent agir ensemble, la contribution relative de chacune d'elles dans la dégradation d'un ARN est encore sujet à débat.

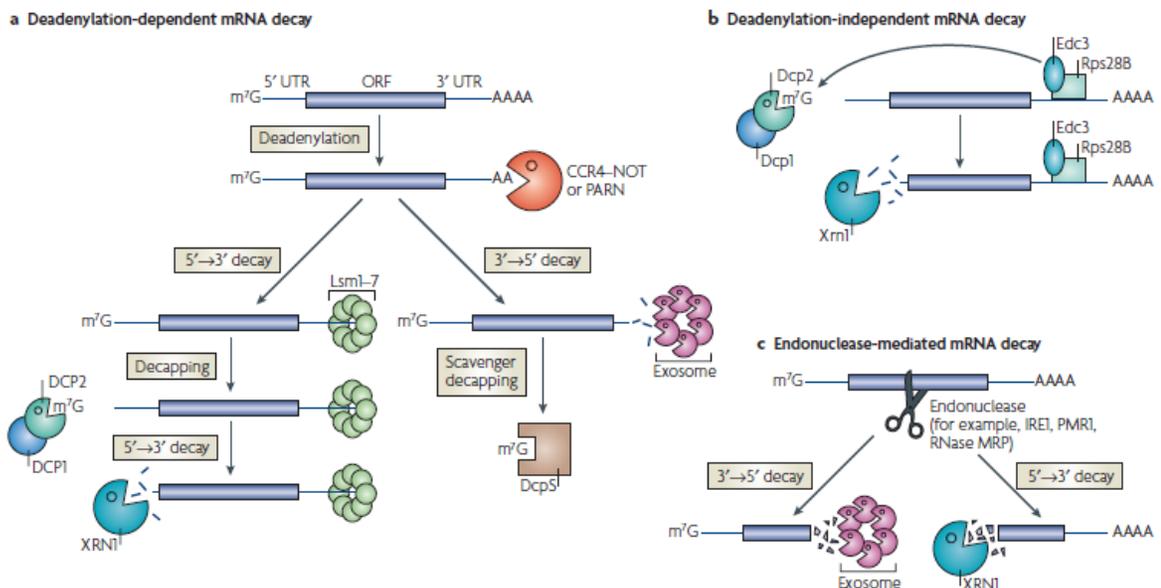


Figure 4-6 : Mécanismes de dégradation des ARNm dans les cellules eucaryotes.

Tiré de [260].

La majorité des transcrits subit une dégradation dépendante de la déadénylation, mais quelques ARNm spécifiques contournent ce mécanisme pour ajuster leur abondance. Le retrait de la coiffe indépendant de la déadénylation a été rapporté chez *S. cerevisiae* pour deux transcrits *RPS28B* et *EDC1* (qui code pour une protéine régulatrice du retrait de la coiffe, *Enhancer of decapping 1*). L'ARNm de *RPS28B* s'engage dans sa dégradation par un phénomène d'autorégulation. La protéine RPS28B lie directement une structure en tige boucle dans le 3'UTR de son ARN, puis interagit avec EDC3 qui semble recruter des facteurs impliqués dans le retrait de la coiffe (Figure 4-6b) [265]. L'extrémité non protégée est attaquée par XRN1. Dans le cas de l'ARNm de *EDC1*, la déadénylation semble être inhibée par l'appariement de la queue poly(A) avec une région poly(U) présente dans le 3'UTR du transcrit [266].

Le troisième mécanisme caractérisé est le clivage endonucléolytique (Figure 4-6c). Cette cassure crée des extrémités 5' et 3' non protégées, dégradées par XRN1 et l'exosome respectivement. Parmi les endoribonucléases les plus connues retrouvées chez les eucaryotes (revue dans [267]), on retrouve AGO2 qui intervient dans la voie des petits ARN interférents, PMR1 [268] et IRE2 [269] sont associés aux polysomes et clivent donc les ARNm en cours de traduction, le complexe MRP (voir Chapitre 6) impliqué dans la maturation des ARNr dans le nucléole et la maturation des ARNmt, mais également dans le clivage du 5'UTR du transcrit *CLB2* (Cyclin de type B) en fin de mitose chez *S. cerevisiae* [270].

#### **4.2.3.2. Les complexes enzymatiques impliqués dans la dégradation des ARNm.**

Trois complexes de déadénylases ont été caractérisés, chacun ayant des propriétés différentes. Le complexe PAN2-PAN3 (pour *PABP-dependent poly(A) nuclease*), stimulé par la présence des PABP, est impliqué dans la réduction de la queue poly(A) des transcrits nouvellement synthétisés pour atteindre une taille standard de 60-80 nucléotides chez *S. cerevisiae* [271]. Chez les mammifères, la queue poly(A) d'environ 200 nt de la  $\beta$ -globine est réduite par ce complexe à environ 80 nt [272], puis la suite est effectuée par le complexe CCR4-NOT. Il n'est pas certain que PAN2-PAN3 joue un rôle dans la déadénylation de tous les ARNm chez les eucaryotes [260]. CCR4-NOT est un complexe de 9 protéines. Seuls CCR4 et CAF1 ont une activité exonucléase, et CCR4 a par ailleurs un

rôle dans la transcription. CCR4-NOT est la principale déadénylase chez *S. cerevisiae* [273], et contrairement à PAN-PAN3, celle-ci est inhibée par la présence des PABP. Enfin, la troisième déadénylase caractérisée est PARN (*Poly(A)-specific ribonuclease*). Son activité est stimulée par la présence de la coiffe sur l'ARNm, mais inhibée par la présence des protéines de liaison à la coiffe [274].

Des homologues de ces trois complexes protéiques ont été identifiés dans les génomes des *Trypanosomatidae*, excepté la sous-unité CCR4. Si les composants de CCR4-NOT et PAN2-PAN3 sont retrouvés en copie unique, trois homologues de PARN ont pu être identifiés. Seuls CCR4-NOT et PARN-1 ont été étudiés chez *T. brucei* (Tableau 1), montrant un rôle essentiel de CAF1 pour la croissance cellulaire et son activité déadénylase au moins sur les ARNm exprimés de manière constitutive [275], et un rôle non essentiel de PARN-1 qui possède une activité déadénylase sur plusieurs ARNm, dont certaines procyclines, groupe de protéines du manteau exprimés en fonction du stade de vie du parasite [276]. CCR4-NOT reste la déadénylase majeure chez les trypanosomes.

La dégradation des ARNm de 5' vers 3' débute par le retrait de la coiffe, opération effectuée par le dimère protéique Dcp1 et Dcp2 chez *S. cerevisiae*, l'activité catalytique étant conduite par Dcp2 [277]. Un troisième composant existe chez les eucaryotes supérieur, Hedls (également nommé Ge-1) stimulant l'activité enzymatique [278]. D'autres facteurs sont requis pour une bonne efficacité du retrait de la coiffe. Parmi eux, les protéines Lsm1-7 (impliquées également dans l'épissage des transcrits) s'associent à l'extrémité 3' des ARN déadénylés pour induire le retrait de la coiffe. D'autres protéines accessoires sont mentionnées tels que la *DEAD/H-box RNA helicase* DHH1 [279], les protéines EDC [280, 281], ou encore RAP55 (*RNA associated protein of 55 kDa*, ou Lsm14) localisée dans les *P-bodies* et les granules de stress dans les cellules humaines [282] (voir la partie 4.2.4.).

Tableau 2 : Principaux composants des complexes impliqués dans la dégradation des ARNm chez les eucaryotes et comparaison avec ceux identifiés chez les *Trypanosomatidae*.

Complexe	Substrat	Composant	Fonction – Domaine	Orthologues <i>L. infantum</i>	Réf.	Commentaire			
<i>Déadénylases</i>									
CCR4- NOT	ARNm poly(A) <sup>+</sup>	<b>CCR4</b>	3'→5' exonuclease	Non trouvé	n.d.	NOT1 et CAF1 sont essentielles DHH1 est également retrouvé dans ce complexe			
		<b>CAF1</b>	3'→5' exonuclease	<i>LinJ.22.1480</i>	[275]				
		<b>NOT1</b>	Domaine de fonction inconnu	<i>LinJ.21.0880</i>	[275]				
		<b>NOT2</b>	Domaine de fonction inconnu commun (NOT2, 3, et 5)	<i>LinJ.12.0300</i>	[275]				
		<b>NOT5</b>	Domaine de fonction inconnu commun (NOT2, 3, et 5)	<i>LinJ.25.1920</i>	[275]				
PAN2- PAN3	ARNm poly(A) <sup>+</sup>	<b>PAN2</b>	Répétition WD40	<i>LinJ.10.0290</i>	DB	PAN2 est essentielle, mais pas PAN3			
		<b>PAN3</b>	3'→5' exonuclease	<i>LinJ.32.0760</i>	DB				
PARN	ARNm poly(A) <sup>+</sup>	<b>PARN</b>	3'→5' exonuclease	<i>LinJ.23.0660</i> (PARN-1) <i>LinJ.36.6630</i> (PARN-2) <i>LinJ.35.2390</i> (PARN-3)	[276]	PARN-1 est non essentielle, agit notamment sur les ARNm des procyclines chez <i>T. brucei</i>			
<i>Dégradation de 3' → 5'</i>									
Exosome	ARNm poly(A) <sup>+</sup> ; Produits de digestion endonucléolytique	<b>RRP41</b>	RNase PH	<i>LinJ.36.3170</i> (RRP41A)	[283] [284] [285]	La majorité des sous-unités est essentielle pour la croissance des parasites.			
		<b>RRP42</b>		<i>LinJ.02.0290</i> (RRP41B)					
		<b>RRP45</b>		<i>LinJ.22.1430</i> (RRP45)					
		<b>RRP46</b>		<i>LinJ.20.1400</i> (RRP42 =EAP1)					
		<b>MTR3</b>		<i>LinJ.32.3660</i> (EAP2)					
		<b>OIP2</b>		<i>LinJ.36.5120</i> (EAP4)					
		<b>CSL4</b>		S1 RNA binding; KH RNA binding			<i>LinJ.23.1820</i>	[283]	Associé également à la maturation de l'ARNr 5.8S
		<b>RRP4</b> <b>RRP40</b>					<i>LinJ.14.0010</i> <i>LinJ.04.0130</i>		
<b>RRP44</b>	<i>PIN</i> RNA binding; RNase II	<i>LinJ.28.0470</i>	[283] [285]	RRP44 et MRT4 ne semblent pas directement associées à l'exosome					
<b>RRP6</b>	RNase D	<i>LinJ.34.4330</i>	[283] [286]						
<b>n.d.</b>	n.d.	<i>LinJ.06.0870</i> (EAP3)	[285]						
<b>n.d.</b>	Helicase	<i>LinJ.36.3150</i> (MRT4)	[236]						
Scavenger decapping enzyme	7meGpppG	<b>DCPS</b>	HIT pyrophosphatase	Non trouvé	n.d.				
<i>Dégradation de 5' → 3'</i>									
Retrait de la coiffe	ARNm poly(A) <sup>+</sup>	<b>DCP1A</b> <b>DCP1B</b>	EVH1	Non trouvé	n.d.	RAP55 associée à DHH1 dans les P-bodies			
		<b>DCP2</b>	Nudix Mut T	<i>LinJ.36.0320</i> (?)	DB				
		<b>DHH1</b>	DExD/H-box RNA helicase	<i>LinJ.35.0370</i>	[275] [287] [288] [289]				
		<b>RAP55</b>	n.d.	<i>LinJ.25.0550</i>	[275]				
		<b>XRNI</b>	5'→3'exonuclease	XRNA : <i>LinJ.06.0260</i> XRNB : <i>LinJ.17.1250</i> XRNC : <i>LinJ.23.0600</i> XRND : <i>LinJ.36.1870</i>	[290] [291]		XRNA : Nucléaire, cytoplasmique, P-bodies, protéine essentielle XRNB-C : cytoplasmique, protéines non essentielle XRND : Nucléaire et essentielle		

Pour plus de clarté, uniquement les homologues de *L. infantum* (*LinJ*) (espèce étudiée dans cette thèse) sont mentionnés ici, mais la majorité des études fonctionnelles et structurales ont été effectuées chez *T. brucei* et *T. cruzi*. Voir les revues de littérature de la dégradation des ARNm chez les eucaryotes [260, 292] et les *Trypanosomatidae* [236] pour plus de détails. DB : Protéine non étudiée, annotée dans les bases de données [293]. n.d. : non déterminé.

Une activité de decapping est détectée chez les typanosomes [294], mais aucun orthologue des protéines DCP1 et DCP2 n'a pu être mis en évidence et caractérisé jusqu'à ce jour. L'homologue de DCP2 mentionné dans le Tableau 2 reste actuellement le meilleur candidat possible pour *T. brucei*, cependant son inhibition n'affecte pas la croissance des parasites, contrairement à ce que l'on pourrait s'attendre étant donné le caractère essentiel de cette voie métabolique (communication du Dre Christine Clayton sur TritrypDB.org). La présence des 39 nt du SL-RNA dans toutes les extrémités 5' des transcrits permet de concevoir un mécanisme de retrait de la coiffe par une activité endoribonucléolytique ciblée, mais ceci reste encore au stade de l'hypothèse [295]. Chez *T. brucei*, des mutations effectuées dans la séquence de DHH1 provoquent une perte de la croissance, une augmentation des *P-bodies* couplée à une diminution des polysomes et une diminution du niveau global des ARNm. DHH1 présente tout de même une sélectivité pour les ARNm régulés durant les stades de vie du parasite, puisque plusieurs ARNm normalement dégradés au stade insecte sont stabilisés suite à ces mutations, et inversement [289]. Tout comme chez les autres eucaryotes, les études menées sur DHH1 chez *T. brucei* ont montré son association avec RAP55 dans les *P-bodies*.

L'exonuclease 5' → 3' XRN1 agit sur les extrémités 5' non protégées par une coiffe, que ce soit lors de la synthèse de l'ARNm, suite au retrait de la coiffe ou à un clivage endoribonucléolytique. Les trypanosomes ont quatre homologues de ce gène (Tableau 2). Seuls XRNA, nucléaire et cytoplasmique, et XRND, uniquement détectée dans le noyau, sont essentiels à la croissance normale de *T. brucei*. Il est assumé que la majorité des ARNm est dégradée dans le cytoplasme, mais cette donnée n'a pas encore été prouvée. Les fonctions de XRNB et XRNC, toutes deux cytoplasmiques sont encore inconnues.

Enfin la dégradation de 3' vers 5' est assurée par un complexe protéique comptant 10 à 12 sous-unités, l'exosome. Les protéines formant le cœur du complexe ont toutes un domaine RNase PH (phosphorolytique de 3' vers 5'), et plusieurs protéines annexes incluent des facteurs 3' → 5' exonuclease hydrolytiques ou *RNA Helicase*. En plus de la dégradation des ARN cytoplasmiques, l'exosome est en charge de celle des ARNnc (pour non-codant) dans le noyau.

Les homologues de la majorité des protéines de l'exosome ont été caractérisés chez *T. brucei*, lui conférant une composition et une architecture similaire à celle retrouvée chez la levure et les mammifères (voir les références dans le Tableau 3). À l'exception de RRP6 et MTR4, toutes ces sous-unités sont essentielles pour une croissance normale du parasite. Les expériences de purification du complexe suggèrent que MTR4 et RRP44 ne sont pas associées à l'exosome, ou alors très faiblement comparativement aux autres interactions protéiques au sein du complexe. Chez *Leishmania*, nous avons noté que RRP44 présente un élément de type SIDER2 dans son 3'UTR, son expression n'est donc peut-être pas régulée comme celle des autres sous-unités, et pourrait expliquer sa faible association avec l'exosome.

#### 4.2.3.3. La surveillance des ARNm.

Le rôle des ARNm étant de servir de gabarit pour la synthèse protéique, il est logique de retrouver dans les cellules un mécanisme de contrôle qualité pour supprimer les ARNm comportant des erreurs. Trois mécanismes ont été décrits chez les eucaryotes pour réaliser cette surveillance : la dégradation des ARNm non-sens (NMD pour *nonsense-mediated mRNA decay*), des ARNm non traductibles (NGD pour *No-go mRNA decay*) et des ARNm sans stop (NSD pour *non-stop mRNA decay*), ces deux derniers étant moins bien référencés actuellement. Cette surveillance est effectuée dans le cadre d'un premier cycle de traduction, dit cycle pionnier, évitant la production élevée de protéines non fonctionnelles, récemment sujet d'une revue de littérature [296]. L'envoi des ARNm nouvellement synthétisés est initié par la liaison l'hétérodimère CBP80-CBP20 (pour *cap binding protein*) dans le noyau [297], ainsi que le complexe de jonction des exons (EJC) au moins pour les ARN ayant subi un *cis*-épissage, initiant la vérification de l'intégrité du cadre de lecture.

Le NMD est effectué lorsque le ribosome rencontre un codon stop prématuré, reconnu s'il est à moins de 50 nt d'un EJC et recrute rapidement l'endoribonuclease SMG6 et autres enzymes de dégradation de l'ARNm. Plusieurs études suggèrent qu'une longue distance entre le codon stop et le site de polyadénylation peut activer le NMD. S'il existe réellement, ce mécanisme d'activation doit être annulé ou adapté chez les *Trypanosomatidae* étant donné la longueur importante du 3'UTR de ces transcrits. Lorsque

l'ARNm ne possède pas de codon stop, la traduction est effectuée tout au long du 3'UTR. Le polypeptide n'est alors pas libéré du ribosome faisant intervenir des facteurs spécifiques du NSD recrutant l'exosome. Le NGD a été observé chez *S. cerevisiae*, lorsque la présence d'une forte structure secondaire dans l'ARNm fait barrière au passage du ribosome. L'endoribonuclase DOM34 intervient, puis XRN1 et l'exosome. Des descriptions plus approfondies de ces mécanismes sont disponibles dans de récentes revues de littérature, par exemple [298].

Une structure similaire au *cap-binding complex* de 300 kDa [299] ainsi qu'un complexe EJC [300] ont été décrits chez *T. brucei* et sont impliqués, comme chez les autres eucaryotes, dans l'exportation des ARNm vers le cytoplasme. Parmi les multiples facteurs essentiels impliqués dans la régulation du NMD, les protéines UFP1 et UPF2 (pour *up-frameshift*) ont également été étudiées chez *T. brucei* [301], suggérant l'existence de NMD sans pour autant apporter de fortes preuves. Actuellement, aucune étude ne mentionne la possible existence des NGD et NSD chez ces organismes.

#### **4.2.4. Les granules à ARNm.**

Les « corps de traitement » (ou *P-bodies* pour *processing bodies*) et les granules de stress font partie des granules à ARNm retrouvés dans le cytoplasme des cellules. Ils contiennent des transcrits exclus transitoirement de la machinerie de traduction. Ces granules sont capables de relarguer les ARNm qu'ils contiennent afin de les engager ou réengager vers la traduction. Les *P-bodies*, initialement décrits comme des agrégats contenant les enzymes de dégradation des ARNm, sont constitutifs et répondent à des stimuli qui affectent la destruction ou la traduction des transcrits, alors que les granules de stress sont eux spécifiquement induits lors de stress engendrant une inhibition globale de la traduction, par différentes voies de signalisation. Des *P-bodies* et des granules de stress présentant différentes compositions peuvent coexister au sein d'une même cellule. Ces agrégats se sont bien conservés au cours de l'évolution et sont retrouvés dans des organismes unicellulaires jusqu'à des structures plus spécialisées, telles que les neurones de vertébrés. Plusieurs protéines sont communes à ces structures, et d'autres permettent également de stimuler l'interaction entre *P-bodies* et granules de stress [302]. Avec la machinerie de traduction, ces trois composants cellulaires vont dicter le destin des ARNm

dans le cytoplasme, de manière plus ou moins spécifique en fonction des protéines qui sont liées à l'ARNm. Le lien étroit entre *P-bodies* et traduction est souligné par trois observations : (1) les taux de traduction et dégradation des ARNm sont influencés par un équilibre dynamique entre polysomes et les mRNP (complexes ribonucléoprotéiques) contenus dans les *P-bodies*; (2) la machinerie de dégradation des ARNm, qui s'accumule dans les *P-bodies*, est également impliquée dans la répression de la traduction chez plusieurs types cellulaires (cellules somatiques, neurones et oocytes); enfin (3) plusieurs régulateurs spécifiques d'ARNm contrôlent la traduction et / ou la dégradation des ARNm en recrutant les composants d'une structure à l'autre [303].

Plusieurs revues de littérature rapportent les différents composants retrouvés dans les *P-bodies* et / ou les granules de stress chez les eucaryotes [302-304]. Il en ressort que des éléments impliqués dans la dégradation des ARN (déadénylase, enzymes de retrait de la coiffe et ses activateurs, protéines du NMD ou la voie des miRNA) sont plus enrichis dans les *P-bodies*. À l'opposé les granules de stress contiennent plus spécifiquement des composants de la machinerie de traduction, incluant la sous-unité 40S, plusieurs protéines du complexe d'initiation de la traduction ( $eIF2\alpha$ , 3, 4A, 4B, 4G) et de répression de la traduction, ou encore les PABP et les ARNm poly(A)<sup>+</sup>. Les caractéristiques des *P-bodies* et granules de stress sont leur inhibition par blocage de l'élongation de la traduction (avec de la cycloheximide) et leur stimulation par arrêt de l'initiation de la traduction (avec de la puromycine). L'ajout d'inhibiteur de la transcription, tel que l'actinomycine D provoque une perte des *P-bodies*.

La présence de granules similaires a été décrite chez *T. brucei* et *T. cruzi* [287, 288, 291], aucune étude n'a été reportée chez *Leishmania*, mais leur absence serait étonnante. La formation des granules de stress nutritif chez *T. cruzi* (24h) ou *T. brucei* (3h) est inhibée par la cycloheximide et stimulée par la puromycine, l'actinomycine D n'ayant pas d'effets sur ces structures. Ces caractéristiques, en plus des constituants retrouvés, sont similaires à celles observées chez les eucaryotes et rapprochent ces structures des granules de stress. Cependant, l'absence de petite sous-unité ribosomale suggère l'absence de transcrite en provenance de la machinerie de traduction.

Tableau 3 : Contenu des *P-bodies* et granules de stress identifiés chez les trypanosomes.

Composant	<i>T. cruzi</i> (stress nutritif)	<i>T. brucei</i> (stress nutritif)	<i>T. brucei</i> (stress thermique 41°C)
ARNm poly(A) <sup>+</sup>	GS	GS	Peu d'ARN
DHH1	GS		PB et GT
SCD6			PB et GT
PABP1	GS	GS	GT
PABP2	GS	Non	GT
XRNA	GS		XB
eIF2A			GT
eIF3B			GT
eIF3Db	(GS)		
eIF4E1	GS		GT
eIF4E2			GT
eIF4E3			GT
eIF4E4			GT
RPL3	Non		
RPS19	Non		
UBP1	GS	GS	
UBP2	GS		
RBP3	(GS)		
RBP4	(GS)		
RBP5	(GS)		
RBP6	GS		
LA	Non		
CBP20	Non		

Tableau adapté de [295]. GS : Granule de stress; (GS) : co-localisation partielle; GT : Granule de choc thermique; PB : P-bodies; XB : Granules contenant XRNA.

Plusieurs protéines contenant un domaine de liaison à l'ARN ont été détectées (ex, UB1-2, RBP3-6) et sont donc potentiellement impliquées dans la localisation des transcrits. Par contre, *TcCAF1/NOT1* ne semble pas être associées à *TcDHH1* dans ces complexes, mais des immunoprécipitations effectuées avec cette protéine montrent son association avec des ARNm dégradés chez les épimastigotes [305]. Les granules de stress nutritif sont aujourd'hui difficilement classables comme *P-bodies* ou granules de stress, étant donné leur contenu et leurs propriétés. Suite à un stress thermique, le nombre de granules contenant DHH1 et SCD6 (homologue de RAP55) augmente, et une structure inhabituelle, contenant XRNA, mais sans les autres marqueurs communs des *P-bodies* se forme à l'extrémité postérieure de la cellule. Son rôle reste à définir.

Chez *T. brucei*, l'arrêt de l'épissage en *trans* par la sinefungin, un inhibiteur de la méthylation du SL-RNA, provoque l'apparition de plus d'une dizaine de granules périnucléaires. Ils contiennent les protéines retrouvées dans les *P-bodies*, avec en plus la PABP2 et l'homologue de la RNA hélicase VASA. La présence de VASA, en plus d'autres observations, suggère que ces structures sont similaires aux granules germinaux [306]. Ces derniers sont associés aux pores nucléaires et sont potentiellement impliqués dans la régulation de l'export des ARNm, dans le contrôle qualité ou dans la destinée des ARNm [306, 307].

#### **4.2.5. Régulation de l'expression génique par les petits ARN.**

La voie des microARN (miRNA) est un mécanisme retrouvé aussi bien chez les plantes que dans le règne animal (distribution phylogénique revue dans [308]). Elle correspond à une grande famille de régulateurs post-transcriptionnels d'une vingtaine de nucléotides contrôlant plusieurs processus cellulaires. Initialement décrite comme une voie impliquée dans le contrôle des infections virales, il est estimé que 50% des gènes codant pour des protéines sont contrôlés par la voie des miRNA chez les mammifères [309]. La synthèse d'un ARN précurseur est effectuée soit par transcription et clivage par l'enzyme nucléaire Drosha (RNase III), soit à partir d'un exon lors de l'épissage d'un gène par le spliceosome. Dans le cytoplasme, Dicer (RNase III) clive ce précurseur en un duplex d'environ 20 nt dont l'un des brins (miRNA mature, ou ARN guide) est incorporé au complexe miRISC (*miRNA-induced silencing complex*). miRISC est un complexe multiprotéique contenant notamment les protéines argonautes (AGO2, PIWI). Le miRNA s'hybride ensuite avec un ARNm présentant une séquence complémentaire, généralement dans son 3'UTR, et cause soit une répression de la traduction, soit une dégradation du transcrit initiée par le complexe CCR4-CAF, ou encore un clivage endoribonucléolytique effectué par AGO2 (chez les plantes et les mammifères). La régulation de l'expression des gènes par la voie des miRNA est complexifiée par les multiples mécanismes de régulation de la voie elle-même, revue dans [309].

Un autre mécanisme faisant intervenir des petits ARN endogènes, et des complexes enzymatiques quasi similaires, a été décrit et revu dans [310]. La voie des petits ARN interférants endogènes (endo-siRNA pour *small interfering ARN*) peut être initiée par les

éléments transposables, des transcrits naturels antisens, des pseudogènes, ou de longues structures tige-boucles. Cette voie est impliquée dans l'inhibition des protéines de transposon et le clivage d'ARNm ciblés. Les piRNA (pour *PIWI-interacting RNA*) sont des molécules un peu plus longues allant jusqu'à 30 nt et qui sont produites indépendamment de Dicer et Drosha [311]. Ils s'associent avec la protéine PIWI et ont un rôle dans le développement des cellules germinales chez les mammifères et l'inhibition des rétrovirus [312]. Enfin, la voie des exo-siRNA (siRNA exogène) est elle impliquée dans la destruction des ARN double brins d'origine exogène, tels que les virus à ARN double brin par exemple. Cette voie est utilisée en laboratoire pour inhiber l'expression ciblée d'une protéine, sous le nom « RNAi » pour interférence à l'ARN.

L'inhibition de l'expression génique par RNAi est fonctionnelle et largement utilisée chez *T. brucei* [313], et a été plus récemment démontrée chez *L. braziliensis* [314], mais l'activité RNAi a été perdue chez *T. cruzi* [315], *L. infantum* et *L. major* [316]. Dans le cas du genre *Leishmania*, la perte de cette activité est subséquente à la division entre les sous-genre *Leishmania* et *Vianna* [160, 314]. Chez *T. brucei*, l'abondance d'endo-siRNA issus des éléments transposables actifs *ingi* et SLACS comparée à la faible expression de ces transcrits suppose leur rôle important dans la défense de la cellule contre ces éléments [317]. Chez *L. braziliensis*, des petits ARNm en provenance de SLACS ont également été mis en évidence. De plus, de premières études effectuées en comparant l'activité du RNAi entre des souches infectées ou non par le virus LRV tendent à montrer l'implication du virus dans l'atténuation de la voie des endo-siRNA [314], résultat en accord avec l'hypothèse selon laquelle une pression virale avait agi comme force de sélection pour la perte de cette voie [318].

Chez *L. infantum*, malgré l'absence d'activité RNAi, le génome contient une protéine de type argonaute inhabituelle possédant un domaine PIWI. Cette protéine est mitochondriale et requise pour la croissance du parasite au stade amastigote spécifiquement. Elle ne semble pas impliquée dans la biogénèse ou la stabilité des ARNnc, sa suppression est associée à la perte de stabilité d'ARNm codant pour des protéines membranaires [319].

#### 4.2.6. La régulation des ARNm par les éléments riches en AU.

Les éléments riches en AU (ARE pour « *AU-rich elements* ») font partie des séquences non traduites impliqués dans la régulation les mieux caractérisés actuellement. Ils sont retrouvés dans plusieurs transcrits labiles et influencent la stabilité de l'ARNm et / ou l'efficacité de traduction en fonction des protéines de liaison aux ARE (AUBP pour *AU-rich element binding proteins*) qui interagissent avec l'élément. Les ARE ont une longueur allant généralement de 50 à 150 nt [320] et sont classés en trois catégories [321]. Les ARE de classe I contiennent une à trois copies du pentamère AUUUA au sein d'une région riche en uridines, ils sont par exemple retrouvés dans les ARNm de proto-oncogènes comme *FOS* et *MYC*. Les ARE de classe II présentent de multiples copies chevauchantes du pentamère, généralement 5 à 8, et sont retrouvés par exemple dans les ARNm des cytokines *GM-CSF* ou *TNF $\alpha$* . Enfin, les ARE de classe III n'ont pas le pentamère et correspondent à une séquence simplement riche en uridine (URE pour « *uridine-rich element* »), telle que retrouvée dans l'ARNm du proto-oncogènes *JUN*. Il est estimé qu'environ 8% des ARNm des mammifères contiennent un ARE [322], et ces éléments sont enrichis dans les transcrits codant pour des facteurs de croissance, des cytokines, des facteurs de transcription, des protéines impliquées dans le métabolisme des ARN et les gènes de réponse rapide [323, 324].

Une vingtaine d'AUBP a été identifiée jusqu'à aujourd'hui chez les eucaryotes supérieurs (Figure 4-7), incluant celles qui dirigent les ARNm vers la dégradation, telles que le facteur liant les éléments riches en AU (AUF1), les protéines à domaine en doigt de zinc tristetraproline (TTP) et butyrate-regulated factor (BRF1-2) ou encore la protéine régulatrice de l'épissage à domaine KH (KSRP). D'autres AUBP sont au contraire impliquées dans la stabilisation des transcrits telles que HuR, ou dans le contrôle de la traduction tel que TIA-1 (pour *T-cell intracellular antigen 1*) [298, 325]. Les AUBP exercent donc majoritairement un effet déstabilisateur sur les ARNm contenant un ARE, et ceci est effectué en recrutant les déadénylases. La stabilisation résulte d'interférence avec ce processus.

La dégradation rapide des ARNm par les ARE peut être vue comme un processus par défaut. La stabilisation est elle le résultat soit de la dissociation d'une AUBP, soit de

l'interférence de l'interaction entre une AUBP et les ribonucléases ou le remplacement d'une AUBP déstabilisatrice par une autre stabilisatrice. Deux mécanismes de régulation des AUBP ont été décrits, la modification post-traductionnelle par phosphorylation (Figure 4-7) et l'exportation du noyau.

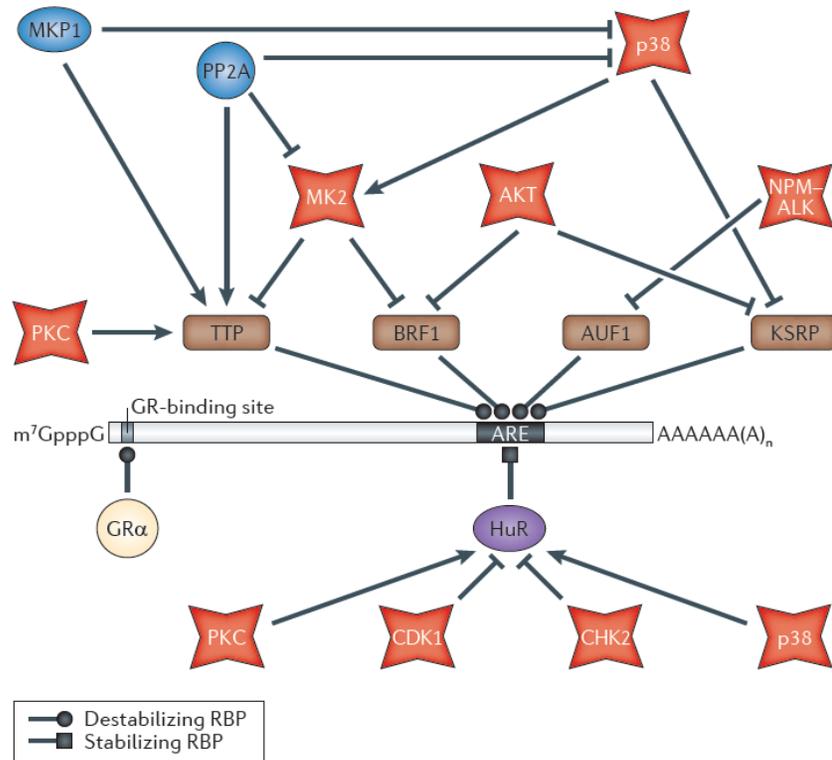


Figure 4-7 : Signalisation impliquée dans la régulation de la dégradation des ARNm contenant des ARE.

Les transcrits présentant un ARE sont stabilisés par (1) la phosphorylation de TTP et BRF1 effectuée par MK2 (protéine kinase activée par les MAPK 2), (2) la phosphorylation de KSRP par p38 (MAPK), (3) par la phosphorylation de BRF1 et KSRP via AKT et (4) la phosphorylation de médiée par NPM-ALK du facteur AUF1. Par contre, la phosphorylation de TTP par la protéine kinase C augmente son activité déstabilisatrice. Ces kinases sont elle-même régulées, notamment par les phosphatases MKP1 et la protéine phosphatase 2A, dont l'activité conduit au final à la stabilisation des ARNm contenant un ARE. Tiré de [325].

La protéine HuR (ou ELAVL1 pour *embryonic lethal abnormal vision-like 1*) a majoritairement un effet stabilisateur sur les ARNm contenant un ARE et sa liaison est régulée par phosphorylation. Si PKC et p38 ont un effet positif, HuR est par contre inhibée par les kinases CDK1 et CHK2.

La localisation des AUBP conditionne également leur activité et cette localisation varie en fonction de leur phosphorylation. HuR est principalement retrouvée dans le noyau. Il semble qu'elle y lie des ARNm et est cotransportée vers le cytoplasme, permettant initialement de protéger les ARNm de la dégradation par les autres AUBP. La phosphorylation des protéines TTP et BRF1 induit leur interaction avec les protéines adaptatrices 14-3-3 (protéines liant les phosphosérines ou phosphothréonines et jouant un rôle dans les interactions protéines-protéines) pour soit séquestrer ces AUBP dans le noyau ou les exclure des granules de stress. Les protéines AUF1 circulent entre le noyau et le cytoplasme. Des interférences dans ce transport, ou la dégradation des protéines AUF1 peuvent altérer son abondance dans le cytoplasme, et ainsi modifier le taux de dégradation des ARNm contenant des ARE. Également, durant la phase S du cycle cellulaire, une dégradation des protéines AUF1 médiée par le protéasome permet de stabiliser l'ARNm de DNMT1 (DNA methyltransferase 1, responsable de maintenance de la méthylation de l'ADN durant la division cellulaire). Le rythme circadien semble être l'un des paramètres modifiant le transfert d'AUF1 entre noyau et cytoplasme.

Chez les *Trypanosomatidae*, quelques ARE ont été caractérisés et font majoritairement partie des ARE de classe III (voir Tableau 4). Ils sont notamment présents dans les ARNm stables dans le stade procyclique de *T. brucei*, mais instables dans la forme sanguine du parasite. Plusieurs transcrits sont régulés par un URE chez *T. cruzi* (revu dans [326]) et nous avons identifié un URE responsable de la dégradation d'un transcrit spécifique du stade amastigote chez *L. infantum* (voir Chapitre 8). Si les AUBP précédemment citées n'ont pas été caractérisées ou étudiées chez les *Trypanosomatidae*, des protéines de liaison aux URE *TcUBP1* et *TcUBP1* ont été bien étudiées chez *T. cruzi*. Chez *T. brucei*, l'expression de la protéine humaine HuR a montré que l'activité stabilisatrice des ARNm par cette AUBP est tout de même conservée chez ces parasites. L'expression de HuR est également associée à un arrêt de la croissance chez la forme sanguine, réduit l'abondance de certains ARNm spécifiques de cette forme et augmente au contraire des transcrits spécifiques de la forme procyclique du parasite [327].

#### **4.2.7. Participation des rétroposons inactifs dans la régulation des ARNm endogènes.**

Tel que présenté précédemment, les rétroposons dégénérés et inactifs SIDER sont présents dans le 3'UTR de nombreux transcrits, spécifiquement chez *Leishmania* spp. L'étude de certains de ces éléments a montré leur implication dans la régulation de l'expression des gènes, et laisse entrevoir un mécanisme global d'expression génique régulé par domestication de ces rétroposons. Chez *L. infantum*, un rétroposon de type *Li*SIDER1 est présent dans le 3'UTR de l'un des gènes de l'amastine et est responsable de l'augmentation de la traduction spécifiquement au stade amastigote (voir Chapitres 5 et 11). Les études effectuées sur les SIDER2 de *Leishmania* démontrent un rôle plus spécifique dans la dégradation des ARNm, la plupart des ARNm contenant cette structure étant faiblement exprimée. Deux gènes ont été plus précisément étudiés à la fois chez *L. infantum* et *L. major*, et démontrent un modèle de dégradation différent entre ces deux espèces : si la dégradation par les *Lm*SIDER2 est constitutive dans les stades amastigote et promastigote, celle par les *Li*SIDER2 est uniquement active dans la forme promastigote du parasite [328]. La dégradation des ARNm par les SIDER2 est effectuée par un mécanisme indépendant de la déadénylation et est initiée par un clivage endonucléolytique au sein de la seconde séquence signature de 79 nt, spécifique des SIDER2 [329]. Chez *L. infantum*, l'accumulation spécifiquement au stade amastigote du transcrit de la protéine GCVT-2, élément du complexe mitochondrial impliqué dans le clivage de la glycine, est également médiée par un *Li*SIDER2 [330]. La caractérisation des protéines impliquées dans la reconnaissance des SIDER2 et le clivage subséquent des ARNm sont des travaux importants pour bien comprendre ce mécanisme et sa régulation chez *Leishmania*.

#### **4.2.8. Les interactions ARNm–protéines dans la régulation de l'expression génique chez les *Trypanosomatidae*.**

Le Tableau 4 présente un résumé de plusieurs travaux effectués chez les *Kinetoplastidae* au regard des interactions ARNm et protéines régulatrices. La caractérisation d'une protéine effectrice doit satisfaire certains critères : (1) la protéine doit lier sa séquence *in vivo* et ne doit pas reconnaître des versions mutées non fonctionnelles de cette séquence, (2) la protéine doit être requise pour la régulation de l'ARNm, et (3) les capacités de la protéine à moduler la stabilité ou l'expression d'un ARN qui ne fait pas

partie de ses cibles habituelles, mais auquel la séquence cible a été liée artificiellement, est un argument supplémentaire important pour prouver l'activité de la protéine régulatrice. Si aucune protéine identifiée jusqu'à maintenant ne satisfait tous ces critères, les études effectuées offrent de très bonnes indications des fonctions des protéines régulatrices, ou du moins des séquences ribonucléiques ciblées [295].

### **4.3. Régulation traductionnelle et post-traductionnelle.**

La régulation de la traduction et de la stabilité des protéines jouent des rôles importants dans l'expression différentielle des gènes durant le cycle parasitaire chez *Leishmania*. En effet, les analyses de protéomique montrent que peu de transcrits voient leur abondance modifiée entre les stades promastigote et amastigote dérivés de macrophages. Cette modification concerne 7% du transcriptome chez *L. infantum*, et 9.3% chez *L. major* [331]. Chez *L. infantum*, les études de protéomique entre les deux stades de vie montrent que 6.1% des protéines sont spécifiques du stade promastigote, alors que 12.4% se retrouvent préférentiellement au stade amastigote, soit un total de 18.5% de protéines différentiellement exprimées [332]. Ceci suggère que la régulation de la traduction correspond à une étape importante chez ces parasites pour subvenir aux besoins de la cellule.

#### **4.3.1. La traduction.**

La traduction d'un ARNm par les ribosomes et les protéines associées est un processus effectué en trois étapes : l'initiation qui constitue l'étape critique pour réguler la traduction, l'élongation et la terminaison. Élongation et terminaison sont des mécanismes conservés et peu modulables. Il a cependant été récemment montré que l'élongation était stoppée en réponse au stress thermique chez les cellules humaines et murines, et cet arrêt est généralement retrouvé autour du 65e codon [333], suggérant une fine régulation de cette observation. Un mécanisme similaire semble exister chez la levure [334]. De plus, la terminaison permet le recyclage des ribosomes pour effectuer de nouveaux cycles de traduction ce qui implique tout de même un rôle indirect dans la régulation de l'initiation de la traduction. Nous discuterons ici uniquement de l'initiation de la traduction. Pour une revue plus détaillée de la traduction chez les eucaryotes, voir [335].

Tableau 4 : Résumé des principales interactions ARNm-protéine étudiées chez les *Kinetoplastidae*.

Éléments	ARNm porteur de l'élément	Organisme	Protéine régulatrice	Régulation(s)	Réf.
<b>Études effectuées à partir d'un élément des transcrits</b>					
<i>Éléments riches en U ou AU (3'UTR)</i>					
<b>URE de 26 nt</b>	Procyclin EP1 (PARP)	<i>T. brucei</i>	n.d.	↘ stabilité des ARNm, ↘ traduction (forme sanguine)	[336]
	Phosphoglycerate kinase (PGKB)	<i>T. brucei</i>	n.d.	↘ stabilité des ARNm (forme sanguine)	[327]
	Pyruvate phosphate dikinase	<i>T. brucei</i>	n.d.	↘ stabilité des ARNm (forme sanguine)	[327]
	Sous-unités de la Cytochrome oxidase	<i>T. brucei</i>	n.d.	↘ stabilité des ARNm, ↘ traduction (forme sanguine)	[337]
<b>ARE</b>	Small mucin (SMUG)	<i>T. cruzi</i>	<i>TcUBP-1</i> <i>TcUBP-2</i>	<i>TcUBP-1</i> ↘ stabilité des ARNm (épi- et trypomastigote) <i>TcUBP-2</i> (exprimé chez épimastigote) inhibe l'action de <i>TcUBP1</i>	[338] [339] [326]
<b>URE (~100 nt)</b>	HSP70	<i>T. cruzi</i>	n.d.	↗ stabilité des ARNm lors d'un choc thermique (amastigote)	[340]
<b>URE de 43 nt</b>	Plus de 900 ARNm majoritairement enrichi chez les amastigotes	<i>T. cruzi</i>	<i>TcUBP1</i>	↗ stabilité des ARNm chez les amastigotes	[341]
<i>Retroéléments (3'UTR)</i>					
<b>SIDER1</b>	Amastine	<i>L. infantum</i>	n.d.	↗ traduction par choc thermique (amastigotes)	[342]
<b>SIDER2</b>	Protéine hypothétique	<i>L. major</i>	n.d.	↘ stabilité des ARNm (constitutivement)	[205]
		<i>L. infantum</i>	n.d.	↘ stabilité des ARNm (promastigotes)	[328]
	Aminomethyltransferase	<i>L. major</i>	n.d.	↘ stabilité des ARNm (constitutivement)	[205]
		<i>L. infantum</i>	n.d.	↘ stabilité des ARNm (promastigotes)	[328]
		<i>L. infantum</i>	n.d.	↘ stabilité des ARNm (promastigotes)	[330]
<i>Autres Éléments</i>					
<b>EP1 loop III (tige boucle de 16 nt)</b>	Procycline EP1 (PARP)	<i>T. brucei</i>	n.d.	↗ traduction (procyclique)	[343]

Éléments	ARNm porteur de l'élément	Organisme	Protéine régulatrice	Régulation(s)	Réf.
<b>Élément de réponse à un choc froid</b>	Procycline EP1 (PARP)	<i>T. brucei</i>	n.d.	↗ traduction (forme sanguine subissant $\Delta T = -15^{\circ}\text{C}$ )	[344]
<b>Indéfini de 418 nt (plus de 2 éléments)</b>	PGKC Glycosomale	<i>T. brucei</i>	n.d.	↘ stabilité des ARNm, peut-être ↘ traduction (procycliques)	[345]
<b>Glycine responsive element (dans GPEET loop II)</b>	procycline GPEET	<i>T. brucei</i>	TbAlba	↗ traduction (procycliques)	[346]
<b>Élément riche en G de 27 nt</b>	Small mucin (SMUG)	<i>T. cruzi</i>	35, 50, 66 et 80 kDa	↗ stabilité des ARNm (épimastigote ?)	[347]
<b>90 nt et 91 nt</b>	Trans-sialidase	<i>T. cruzi</i>	n.d. (UBP-1?)	↗ stabilité des ARNm (metacyclique)	[348]
<b>5'UTR</b>	HSP70	<i>T. cruzi</i>	n.d.	↗ stabilité des ARNm lors d'un choc thermique (amastigote)	[340]
<b>200 nt</b>	Amastine	<i>T. cruzi</i>	30+36kDa	↗ stabilité des ARNm (amastigote)	[349]
<b>300 nt</b>	Amastine	<i>L. infantum</i>	32 kDa	↗ traduction par choc thermique (amastigote)	[350]
<b>117 nt</b>	HSP70	<i>L. infantum</i>	n.d.	↗ stabilité des ARNm (amastigote)	[351]
<b>3'UTR 500 nt</b>	Adenosyltransferase (MAT2) et Lorien	<i>L. infantum</i>	n.d.	↘ stabilité des ARNm (promastigote)	[352]
<b>PRE (AUGUAnAGUn)</b>	Paraflagellar rod (PFR1, PFR2C, PFR4), kinesin II, protein antigenique	<i>L. mexicana</i>	n.d.	↘ stabilité des ARNm (amastigote)	[353] [354]
<b>250 nt</b>	A600-4	<i>L. mexicana</i>	n.d.	↗ stabilité des ARNm (amastigote)	[355]
<b>500 nt</b>	A600-4	<i>L. mexicana</i>	n.d.	↗ traduction (pro- et amastigotes)	[355]
<b>270 nt</b>	HSP83	<i>L. amazonensis</i>	n.d.	↗ épissage des ARNm, ↗ traduction par choc thermique (amastigote)	[238]
<b>202 nt</b>	GP63 major surface protease (MSPL)	<i>L. chagasis</i>	n.d.	↗ stabilité des ARNm suite au traitement à la cycloheximide (promastigote)	[356]
<b>Études effectuées à partir d'une protéine de liaison aux ARNm</b>					
URE	Groupe de protéines régulées en fonction des stades de vie	<i>T. brucei</i>	<b>TbDRBD3</b>	↗ la stabilité des ARNm (au moins chez procycliques)	[357]
n.d.	5'UTR avec polypyrimidines riches en C	<i>T. brucei</i>	<b>TbDRBD3 (PTB1)</b>	Affecte <i>trans</i> - et <i>cis</i> -splicing. ↗ la stabilité des ARNm (au moins chez procycliques)	[358]

Éléments	ARNm porteur de l'élément	Organisme	Protéine régulatrice	Régulation(s)	Réf.
n.d.	5'UTR avec polypyrimidines riches en C	<i>T. brucei</i>	<b>TbDRBD4 (PTB2)</b>	Affecte <i>trans</i> -splicing. ↗ la stabilité des ARNm (au moins chez procycliques)	[358]
UWUUUGUACC	LIGKA, PNT1, PNT2=CPC2, (Histone 4V)	<i>T. brucei</i>	<b>TbPUF9</b>	↗ la stabilité des ARNm (au moins chez forme sanguine)	[359]
n.d.	ZFP1; ZC3H11, CFB1 Tb927.4.1000 et Tb927.8.7820	<i>T. brucei</i>	<b>TbRBP3</b>	Pas d'effets significatifs	[360]
UBP1/2-responsive element	Cyclin F box (CFB1)	<i>T. brucei</i>	<b>TbUBP1</b> <b>TbUBP2</b>	↗ stabilité des ARNm (au moins chez la forme sanguine)	[361]
3'UTR	Famille de protéines transmembranaires (x5)	<i>T. brucei</i>	<b>TbUBP1</b> <b>TbUBP2</b>	↗ ou ↘ la stabilité des ARNm en fonction des transcrits (au moins chez la forme sanguine)	[362]
3'UTR	Trans-sialidase (TS-like E), Mitochondrial carrier (MCP12)	<i>T. brucei</i>	<b>TbZC3H20</b>	↗ stabilité des ARNm (procyclique)	[363]
Région codante	188 gènes, dont certains du métabolisme énergétique	<i>T. brucei</i>	<b>TbRBP42</b>	n.d. protéine enrichie dans les polysomes	[364]
EP1 Loops II et III	Procycline EP1 et GPEET	<i>T. brucei</i>	<b>TbZFP3</b>	TbZFP3 ↗ traduction (en compétition avec un régulateur négatif)	[365] [366]
Séquences riches en A	223 ARNm	<i>T. cruzi</i>	<b>TcZFP2</b>	↗ stabilité des ARNm pendant metacyclogenèse, ZFP2 étant sous exprimée	[367]
n.d.	9 protéines surexprimées chez les metacycliques	<i>T. cruzi</i>	<b>TcPUF6</b>	↘ stabilité des ARNm (procyclique)	[368]
Strucutre tige boucle ~30 nt	Protéines ribosomales	<i>T. cruzi</i>	<b>TcRBP3</b>	non déterminé	[369]
Élément riche en AG	148 ARNm, donc 46% de protéines avec des signaux transmembranaires	<i>T. cruzi</i>	<b>TcRBP40</b>	TcRBP40 exprimé préférentiellement chez les metacycliques	[370]
CAUAGAAG	TOP2, RPA1, KAP3, DHFR-TS	<i>C. fasciculata</i>	<b>CfCSBPA/B</b>	n.d.	[371] [372]

L'initiation permet l'assemblage du ribosome 80S, dans lequel le codon d'initiation AUG des ARNm s'apparie à l'ARNt de la méthionine. La synthèse des protéines est principalement régulée durant cette phase d'initiation, permettant un contrôle rapide et réversible de l'expression génique. Chez les eucaryotes, cette régulation est effectuée par la modification de l'activité des facteurs d'initiation (eIF) mais aussi par des protéines de liaison à l'ARN régulatrices ou par des microARN. L'initiation est divisée en plusieurs étapes : la formation du complexe de préinitiation 43S, l'attachement de ce complexe au 5'UTR de l'ARNm, le balayage du 5'UTR pour identifier le codon AUG et l'association de la grosse sous-unité 60S.

Un mécanisme d'initiation de la traduction indépendant de la coiffe, également appelé traduction alternative, est détaillé chez les eucaryotes. Ce mécanisme est effectué par des séquences d'ARN correspondant à des sites internes d'entrée des ribosomes (IRES) qui ont été décrits aussi bien dans les ARN viraux que dans les ARNm cellulaires. Ces mécanismes sont bien illustrés et détaillés dans les revues de littérature récentes [335, 373, 374], nous concentrerons cette partie sur le cas de la traduction cytoplasmique des *Trypanosomatidae* (réemment revu dans [375]).

Parallèlement à la régulation de l'initiation, l'enrichissement des codons présents dans les transcrits fortement exprimés a montré qu'une sélection de traduction existe chez les *Trypanosomatidae* [376].

#### **4.3.1.1. Les facteurs de liaison de la coiffe.**

Comme mentionné précédemment, la coiffe présente une structure inhabituelle chez les *Trypanosomatidae*, mais les protéines qui s'y lient sont similaires à celles retrouvées chez les autres eucaryotes. La protéine eIF4E lie directement la coiffe et son complexe (eIF4F) interagit avec les PABP résultant dans une circularisation des ARNm, permettant d'éviter la dégradation par les exonucléases et de rapprocher le 3'UTR et les protéines régulatrices qui s'y lient pour influencer l'efficacité de traduction.

eIF4E est généralement unique chez les eucaryotes, mais quatre paralogues ont été caractérisés dans les génomes de *Leishmania* et *Trypanosoma* (eIF4E1-4), présentant chacun des différences d'affinité pour des analogues de coiffe et des différences de leurs

complexes d'association [377, 378]. Ces protéines divergent de celles des autres eucaryotes puisque aucune provenant du génome de *Leishmania* (*LeishIF4E*) n'est capable de complémentarité chez des levures mutantes n'exprimant pas leur propre eIF4E [379]. *LeishIF4E1* est exprimé préférentiellement au stade amastigote et cette régulation est sans doute due à la présence d'un SIDER1 dans le 3'UTR de son transcrite. L'expression semble être plus constitutive pour les autres paralogues [378]. Par contre, l'expression de *LeishIF4E4* est réduite chez les amastigotes et la protéine ne lie alors plus la coiffe des ARNm [380]. *LeishIF4E1* est capable de lier la coiffe des ARNm, mais contrairement aux observations chez les autres modèles eucaryotes, cette protéine n'interagit pas directement avec eIF4G, excluant son implication dans un mécanisme de traduction dépendant de la coiffe. Une protéine spécifique des *Trypanosomatidae* a été identifiée comme partenaire de *LeishIF4E1*, spécifiquement au stade promastigote, qui pourrait correspondre à un régulateur traductionnel impliqué dans le changement entre traduction dépendante de la coiffe et traduction alternative [380]. Chez *T. brucei* les eIF4E (*TbIF4E*) sont groupés en deux sous-classes : les paralogues *TbIF4E1* et *TbIF4E4*, dont la séquence d'acide aminés a une taille similaire à celle des métazoaires, puis *TbIF4E3* et *TbIF4E4* présentant une longue extrémité N-terminale de 150 acides aminés. Toutes sont exprimées de manière constitutive chez ce parasite, mais des différences fonctionnelles sont suggérées par une localisation différente (*TbIF4E1* et *TbIF4E2* étant nucléaires et cytoplasmiques, *TbIF4E3* et *TbIF4E4* étant uniquement cytoplasmiques). De plus, l'inhibition de la production de ces protéines montre un rôle essentiel uniquement de *TbIF4E3* dans la forme procyclique, tandis que *TbIF4E1*, 3 et 4 sont nécessaires à la croissance des formes sanguines des parasites. D'autres résultats supportent que *TbIF4E3* et 4 ont des rôles importants, mais distincts, dans l'initiation de la traduction [381].

eIF4A est un autre facteur d'initiation de la traduction qui lie la coiffe. Chez l'homme, trois paralogues sont identifiés, deux d'entre eux étant fonctionnellement non distinguables, et le troisième étant un inhibiteur de la traduction. Deux paralogues sont présents dans les génomes des *Trypanosomatidae* (*TbIF4A1* et *TbIF4A3*) et les quelques études publiées sur ceux-ci s'accordent avec les autres modèles eucaryote : ils sont très abondants (*TbIF4A1* étant au moins 10 fois plus abondante que *TbIF4A3*), exprimées constitutivement au cours du cycle parasitaire, *TbIF4A1* est cytoplasmique alors que

*TbIF4A3* est nucléaire. Tout cela souligne une fonction distincte et un rôle prépondérant de *TbIF4A1* dans l'initiation de la traduction [382].

Le facteur d'initiation eIF4G est impliqué entre autre dans la stimulation de l'activité hélicase d'eIF4A. Si deux paralogues sont présents chez l'homme, les *Trypanosomatidae* en possèdent cinq de taille, composition et potentiel isoélectrique très différents [377].

#### **4.3.1.2. Les protéines de liaison aux queues poly(A).**

Les PABP sont impliquées dans la stabilisation des ARNm, leur exportation vers le cytoplasme, et leur interaction avec eIF4G en font également un régulateur traductionnel. Ces protéines existent en plusieurs versions chez les métazoaires, une nucléaire (PABPN) et au moins une cytoplasmique (PABPC). Elles sont généralement uniques chez les organismes unicellulaires, mais *Leishmania* possède en revanche trois paralogues des PABP (*LeishPABP1* à 3) [383, 384]. Le genre *Trypanosoma* a perdu PABP3 au cours de son évolution, et deux gènes sont donc retrouvés chez les trypanosomes [385]. Chez *Leishmania*, ces protéines sont majoritairement cytoplasmiques, mais se concentrent dans le noyau lors d'une inhibition de la traduction par l'actinomycine D. Si cette accumulation est totale pour *LeishPABP2* et *LeishPABP3*, elle reste minime pour *LeishPABP1*, suggérant des fonctionnalités différentes [384]. Une relocalisation des PABPC a également été reportée lors d'irradiations aux ultraviolets chez les métazoaires [386] ou suite à une protéolyse virale lors d'infection de cellules par ces agents (*Herpesviridae*, *Reoviridae* ou *Bunyaviridae*, revu dans [387]). L'utilité de ce rôle n'est pas encore définie, surtout au regard des conditions physiologiques. Les expériences de purification des complexes suggèrent que *LeishPABP2* et 3 interagissent ensemble, mais *LeishPABP1* semble faire partie d'une autre structure. Si des contacts de *LeishIF4G3* avec *LeishPABP2* et *LeishPABP3* sont détectés *in vitro*, ils n'ont pas été confirmés *in vivo*. *LeishPABP1* semble être l'orthologue le plus important dans la traduction [384]. Chez *T. brucei*, *TbPABP2* s'accumule dans le noyau suite à l'action de la sinefungin et d'un stress thermique. De plus, *TbPABP1* et *TbPABP2* s'associent différemment aux polysomes [385]. Tout cela suggère également que ces protéines ont des fonctionnalités différentes.

Les expériences conduites chez d'autres *Trypanosomatidae* ont notamment montré chez *T. cruzi* l'association de *TcPABP2* avec *TcUBP1*, précédemment décrite comme liant les URE de transcrits dégradés au stade trypomastigote [339], l'association chez *T. brucei* de *TbPABP2* avec le complexe ribonucléoprotéique de *cis*-épissage U1 [388], et enfin l'association de *CfPABP2* de *C. fasciculata* au niveau du 3'UTR de gènes régulés durant le cycle cellulaire [371]. Ces résultats suggèrent globalement une certaine affinité des PABP des *Trypanosomatidae* pour certains sous-groupes d'ARNm.

#### **4.3.1.3. De multiples possibilités de formation du complexe d'initiation de la traduction.**

eIF4F correspond au complexe de liaison à la coiffe et est composé des facteurs d'initiation précédemment décrits : eIF4E, eIF4A et eIF4G. Il est capable de lier les PABP et le 43S, la structure formée étant soit dans un état de latence traductionnelle, soit dirigée vers la traduction. La constitution d'eIF4F doit donc jouer un rôle important pour définir la synthèse protéique à partir d'un transcrit. La multiplication des paralogues de ces constituants spécifiquement chez les *Trypanosomatidae*, et la démonstration des différences fonctionnelles de certains d'entre eux, laissent imaginer de multiples agencement possibles de eIF4F et des PABP, potentiellement impliqués dans la reconnaissance de certains transcrits spécifiquement, ou des modifications traductionnelles en fonction du stade de vie du parasite (notamment avec *LeishIF4E1* et *LeishIF4E4*, régulés différemment pendant le cycle parasitaire). Cette multiplicité est en plus modulable par les différentes possibilités de modification post-traductionnelles, telles que des phosphorylations, de ces sous-unités, sans compter les protéines interagissant avec les 4E (4E-BP), elles-mêmes régulées par phosphorylation via la kinase mTOR, pouvant inhiber l'initiation de la traduction dépendante de la coiffe et modifier la localisation des protéines eIF4E durant des stress thermiques [389] (revue de littérature [390]). L'identification de différents complexes eIF4G n'est pas fréquente, un cas ayant tout de même été rapporté chez les plantes [391].

#### **4.3.1.4. Réponse au stress et différenciation en amastigote par phosphorylation d'eIF2 $\alpha$ .**

Le facteur d'initiation de la traduction eIF2 est l'un des premiers à entrer en jeu dans la chronologie de la traduction. Il forme un complexe ternaire avec GTP, méthionine et l'ARNt de la méthionine, puis se lie ensuite avec eIF5, eIF3 et la sous-unité 40S. Ce

complexe ainsi formé (43S) est recruté par le complexe eIF4F-ARNm. Suite à la reconnaissance du codon d'initiation, eIF5 hydrolyse le GTP contenu dans eIF2, ces facteurs étant recyclés pour initier d'autres cycles de traduction.

eIF2 est composé de trois sous-unités, dont eIF2 $\alpha$  est phosphorylable. Lorsqu'eIF2 $\alpha$  est phosphorylée, eIF2 conserve dans son complexe la molécule de GDP hydrolysée, interdisant le recyclage du complexe ternaire avec une molécule de GTP (médiée par l'échangeur de guanine eIF2B). La phosphorylation d'eIF2 $\alpha$  est un signal cellulaire important pour inhiber la traduction, mais il a été rapporté que dans certains cas, la traduction alternative pouvait être faite indépendamment de eIF2 [392]. La phosphorylation de ce facteur d'initiation joue un rôle important dans la différenciation des *Trypanosomatidae*, et également d'autres parasites tels que *Toxoplasma gondii* [393] ou *Plasmodium falciparum* [394]. Chez l'ensemble de ces parasites plus spécifiquement, une extension N-terminale est présente dans la séquence d'acides aminés d'eIF2 $\alpha$ . Elle peut atteindre autour de 115 résidus chez les *Trypanosomatidae*, lui conférant sans doute une ou plusieurs fonctions additionnelles. Une autre différence au niveau de la structure primaire de cette protéine chez les *Kinetoplastidae* est son site de phosphorylation, la sérine conservée chez les autres eucaryotes étant échangée pour une thréonine en position 169 pour *T. brucei* [395] et *Leishmania* [396].

Quatre kinases d'eIF2 $\alpha$  ont été caractérisées chez les eucaryotes : HRI (pour *heme-regulated inhibitor kinase*) qui répond lors de l'absence d'hème, GCN2 est stimulé par la déplétion d'acides aminés, PKR est une kinase dépendante des ARN double brin et PERK répond à l'accumulation de protéines non repliées dans le reticulum endoplasmique, (revues dans [397, 398]). *T. brucei* code trois kinases putatives d'eIF2 $\alpha$  (*TbIF2K1* à 3), dont *TbIF2K2* présente des homologies avec PERK, est capable de phosphoryler *TbIF2 $\alpha$* , mais sa localisation au niveau de la poche flagellaire et des endosomes diffère des autres organismes. Cette kinase est elle-même phosphorylée uniquement dans la forme sanguine, suggérant une activité différente pendant le cycle parasitaire [395]. Si *TbIF2K2* est capable de phosphoryler la protéine eIF2 $\alpha$  de la levure, la divergence évolutive de ce système de régulation est inspirée par l'incapacité de *TbIF2 $\alpha$*  à être phosphorylée par les kinases humaines.

Chez *L. infantum*, *LiPERK* réside dans le réticulum endoplasmique et phosphoryle spécifiquement *eIF2 $\alpha$* . *LiPERK* est régulée par autophosphorylation sur son domaine kinase et ce mécanisme est nécessaire pour la phosphorylation d'*eIF2 $\alpha$* . La surexpression d'une version tronquée, ne portant pas le domaine N terminal de *LiPERK*, présente un phénotype de mutant dominant négatif qui inhibite la phosphorylation d'*eIF2 $\alpha$* . Cette inhibition est associée à un retard de croissance lors de la différenciation des parasites de promastigote vers amastigote [399]. Dans des cellules *L. infantum* de type sauvage, les stress de température et d'acidité combinés provoquent une importante diminution de la traduction, exceptée pour certains gènes impliqués dans l'induction de la différenciation tels que le gène *A2*. Suite à sa différenciation complète en milieu axénique, les amastigotes conservent un faible taux de traduction comparativement aux promastigotes, et toutes ces données corrélerent avec une augmentation de la phosphorylation d'*eIF2 $\alpha$*  [396].

*eIF2 $\alpha$*  subit une modification post traductionnelle par clivage de sa région N-terminale accomplie par une aminopeptidase. Si chez les eucaryotes cette excision ne concerne que la première méthionine de la séquence d'acides aminés, ce clivage est en revanche plus varié chez *Leishmania*, sans doute en relation à sa longue extrémité N terminale. L'excision de la première méthionine est uniquement effectuée chez les promastigotes, et celle-ci est essentielle pour des clivages subséquents, permettant l'obtention de protéine de 42, 35 et 25 kDa (la taille initiale étant de 47 kDa). Les auteurs proposent un mécanisme de régulation traductionnelle dans lequel la protéolyse d'un faible pourcentage d'*eIF2 $\alpha$*  uniquement au stade promastigote prévient sa phosphorylation permettant de garder un niveau élevé de traduction (Cloutier C., Samant M. et Papadopoulou B., en cours de publication).

#### **4.3.2. Les ARNr antisens sont régulateurs de la traduction.**

Chez *Leishmania*, des ARN anti-sens complémentaires de tous les ARNr (28S, 18S et 5,8S) sont naturellement présents dans les deux stades de vie. L'une des particularités des *Trypanosomatidae* est que leur ARNr 28S est clivé en six fragments stables, appelés LSU- $\alpha$  à LSU- $\gamma$  (pour *large subunit*). Les ARNras (ARN ribosomal anti-sens) et ARNrs (ARN ribosomal sens) de LSU- $\gamma$  ont été récemment étudiés [400]. Si un faible niveau de fragmentation des ARNras et des ARNrs est visible en conditions physiologiques au sein

des ribosomes assemblés, les stress inducteurs d'apoptose provoquent une augmentation de ces fragmentations, conduisant à une perte des ARNr et un arrêt de la traduction. La position des sites de coupure sur ces deux brins suggère que la fragmentation est effectuée sur le dimère sens et anti-sens. Ceci a été également observé chez *T. brucei* et dans des cellules de mammifère (THP-1), suggérant un mécanisme conservé. Chez *Leishmania*, une *DEAD-box ATP-dependent RNA helicase* (LinJ.32.0410) a été identifiée comme liant les ARNras et ARNrs de LSU- $\gamma$ . Les auteurs proposent qu'en conditions physiologiques, l'hélicase lie les ARNras et ARNrs au sein des ribosomes en cours de traduction pour les protéger de la fragmentation. Suite à un stress sévère, l'interaction hélicase-ARNr est perturbée, ce qui favorise l'attaque par des nucléases et l'arrêt de la traduction.

### **4.3.3. Modifications post-traductionnelles.**

Les modifications post-traductionnelles jouent un rôle important pour rapidement modifier l'activité, la stabilité ou la localisation des protéines. Ces modifications peuvent être effectuées par différents stimuli, notamment lors de stress environnementaux ou autres signaux extérieurs, en encore durant la progression du cycle cellulaire. Ces modifications peuvent être de diverses natures telles que la phosphorylation, la méthylation, l'acétylation, la glycosylation, l'acylation ou encore l'ubiquitination pour exemple (revu dans [401]).

Chez les *Trypanosomatidae*, les études sont majoritairement effectuées lors du changement de stade parasitaire dans le but de mieux comprendre les mécanismes d'adaptation des parasites à leur nouvel environnement. Les études de protéomique comparant les stades promastigotes et amastigotes ont d'ailleurs montré que plusieurs protéines présentent un profil de migration différent entre les deux stades de vie, suggérant une modification post-traductionnelle différentielle [332]. Une autre étude a recherché les changements post traductionnels effectués pendant la différenciation de promastigotes vers amastigotes, au niveau de la phosphorylation, la glycosylation, l'acétylation et la méthylation chez *L. donovani*. Une vingtaine de protéines a été identifiée dans chaque cas, excepté pour l'acétylation qui semble moins importante à ce stade. Aucune phosphorylation sur une thréonine n'a été mise en évidence à ce stade, mais il a été recensé deux types de glycosylation plus rares : fucosylation et hexosylation. Au regard de l'abondance des protéines, la méthylation semble être concomitante à une stabilisation des protéines [402].

Des recherches bioinformatiques ont permis d'identifier près de 200 protéines kinases putatives par génome chez *T. brucei*, *T. cruzi* et *L. major*, représentant environ 30% du nombre de kinases retrouvées chez l'homme, mais aussi le double que celles dénombrées chez *Plasmodium* [403]. Chez *Leishmania*, outre les kinases d'eIF2 $\alpha$  présentées précédemment, plusieurs autres ont été clonées et caractérisées, montrant leur implication dans la régulation du cycle cellulaire et dans le développement parasitaire. Chez *L. amazonensis*, la kinase PKA, impliquée dans la voie de l'adenylate cyclase, est majoritairement active chez les parasites métacycliques [404]. Différentes MAPK (pour *mitogen-activated protein kinase*) ont été étudiées chez *Leishmania* et sont impliquées dans la survie intracellulaire des amastigotes [405], des activités de phosphorylation confinées au stade amastigote [406] ou encore dans le contrôle de la taille du flagelle [407, 408]. Ces données confèrent un rôle clef des MAPK dans la signalisation intracellulaire et le développement cellulaire. Le rôle du flagelle comme senseur impliqué dans de cette voie de régulation est possible et discuté [409]. Plusieurs études globales du phosphoprotéome ont été effectuées chez les *Trypanosomatidae* [410-412] montrant l'implication de ces modifications entre les stades de vie des parasites, notamment chez *Leishmania* [402]. Il a été démontré que des chocs d'acidité ou de température pouvaient induire la phosphorylation chez *Leishmania* [413, 414], également observée lors de la différenciation de souches virulentes chez *L. major* [414].

La méthylation est un mécanisme de régulation post traductionnelle affectant les histones, et la maturation des transcrits notamment. La méthylation des arginines est effectuée par les PRMT (pour *protein arginine methyltransferase*) et différents types sont rapportés en fonction du type de methylarginine généré. Si les PRMT sont au nombre de neuf chez l'homme, seuls cinq homologues sont rapportées chez les *Kinetoplastidae* (revu dans [415]). Deux d'entre elles, *TbPRMT3* et *TbPRMT7* sont significativement moins exprimées en phase stationnaire de croissance chez *T. brucei*. *TbPRMT1* est la seule dont l'interférence à l'ARN a montré une inhibition de croissance spécifiquement au stade sanguin du parasite, mais la plus importante pour la croissance est *TbPRMT6*.

Nous avons discuté précédemment du rôle des acylations dans la localisation flagellaire des protéines. Plusieurs enzymes sont retrouvées dans le génome de *Leishmania*

permettant la liaison de lipide, suggérant d'abondantes modifications post-traductionnelles de ce type. Il a été également démontré que la N-glycosylation de GP63 permettait de réguler son activité lors de différents stades infectieux du parasite [416].

#### 4.3.4. Dégradation des protéines.

Parmi les modifications post-traductionnelles, l'ubiquitination consiste en l'ajout d'un peptide de 76 acides aminés, majoritairement sur des lysines. Cette réaction est effectuée par plusieurs enzymes agissant en cascade : E1 (*ubiquitin-activating enzyme*), E2 (*ubiquitin-conjugated enzyme*) et E3 (*ubiquitin-protein ligase*), et l'ubiquitine peut-être recyclée et régulée par DUB (*deubiquitylating enzyme*) suite à la protéolyse. La monoubiquitination est un signal impliqué dans la régulation de la transcription, le trafic des protéines ou la réparation de l'ADN chez les eucaryotes, mais encore peu de choses sont connues sur ce mécanisme. En revanche, la polyubiquitination est une modification post-traductionnelle permettant de diriger les protéines ciblées vers le protéasome, un complexe de 26S hautement régulé contenant plus de 1 000 protéines (revu dans [417]). Ces cibles sont des protéines de processus cellulaires variés, tels que la transcription, la régulation du cycle cellulaire, les protéines de contrôle de qualité. Outre sa fonction dans la dégradation totale des protéines, le protéasome est également impliqué dans des protéolyses partielles et contrôlées, telles que le facteur de transcription Gli par exemple. Chez *L. infantum*, la présence de ces domaines suggère l'existence de trois protéines E1, sept E2 et treize E3 putatives. L'ubiquitination de la protéine CBP (*cyclin-binding protein*) a été mise en évidence [418] et le protéasome a été caractérisé [419, 420]. À la vue de sa surexpression pendant la différenciation en amastigote, le protéasome semble avoir un rôle important dans la modification de l'expression génique à ce stade [80].

Les *Trypanosomatidae* et d'autres rares eucaryotes unicellulaires tels que *Toxoplasma* et *Plasmodium* ont la particularité de posséder également des protéines de choc thermique bactérienne HslU et HslV [421], cette dernière étant considérée comme l'équivalent du protéasome eucaryote. Le complexe dodécamérique HslUV est inductible par choc thermique, et des mutants n'exprimant pas ce complexe chez *E. coli* deviennent sensibles à la température [422]. L'étude de ces protéines chez *L. donovani* a montré que seul HslV a un rôle essentiel dans la croissance des parasites, et une thermosensibilité est

suggérée suite à sa déplétion. En revanche, les deux homologues HslU présents chez *Leishmania* n'interagissent pas avec HslV, ils interagissent entre eux directement pour former un complexe dodécamérique [423]. De par son absence chez la majorité des eucaryotes, ce complexe est une nouvelle cible de choix pour le développement de molécules anti-parasitaires.

En plus de ces structures, *Leishmania* code plus de 150 peptidases, majoritairement des cystéines protéases, exprimées principalement chez les promastigotes [424] et des métalloprotéases, plus spécifiques du stade amastigote [425], certaines étant des facteurs de différenciation du parasite.

Enfin, l'autophagie (revue dans [426]) est un mécanisme de recyclage des protéines induit lors de stress cellulaires tels que l'absence de nutriments, et les protéines ATG (*autophagy-related proteins*) sont régulées par ubiquitination ou acétylation de leur lysine chez les eucaryotes. La présence de deux ATG4, six ATG8, de CPA (*Cysteine peptidase A*) et CPB et plusieurs autres composants de cette voie suggère son rôle dans le recyclage des protéines, hypothèse supportée par l'augmentation du nombre d'autophagosomes chez les parasites métacycliques et en cours de différenciation en amastigote [427, 428].



## Chapitre 5 : Les amastines.

Les amastines sont des glycoprotéines membranaires décrites uniquement chez les *Trypanosomatidae* et spécifiques à ce groupe. Leur origine semble coïncider avec la séparation entre *Trypanosomatidae* et *Bodo saltans* [429] (voir Chapitre 3). Elles sont exprimées préférentiellement au stade amastigote, ce qui suggère qu'elles jouent un rôle à l'interface hôte-parasite et qu'elles possèdent des fonctions essentielles dans l'adaptation ou la survie au sein des cellules phagocytaires.

Avec les protéines A2 [430, 431], la métallopeptidase majeure de surface GP63, la cystéine protéase et LPG3, les amastines sont considérées comme des facteurs de virulence [432]. Chez *Leishmania* spp., certaines amastines sont testées pour le développement de vaccins [433] et sont également utilisables comme biomarqueur de leishmaniose viscérale active chez l'homme [434]. Ces protéines font parti des séquences les plus immunogéniques parmi les antigènes de surface de *Leishmania* chez la souris [435] et sollicitent une forte réponse immunitaire chez l'homme, notamment dans le cas des leishmanioses viscérales [434].

L'élucidation de leur(s) rôle(s) et de leur régulation est importante afin de mieux comprendre la survie des parasites dans les macrophages, les mécanismes de différenciation de *Leishmania* entre les stades promastigote et amastigote, et enfin dans le but de trouver une cible thérapeutique contre ces infections. Comprendre le rôle des amastines chez *T. cruzi* et *Leishmania* spp. est rendu difficile par deux facteurs : (1) les amastines sont la plus grande famille de gènes exprimée de manière stade spécifique chez *Leishmania*, pouvant compter plus de 60 membres par génome, rendant impossible la génération d'une souche déficiente par recombinaison homologue et (2) l'absence d'interférence à l'ARN fonctionnelle chez ces parasites interdit les études de sous-expression par déstabilisation des transcrits. Ceci sera peut-être réalisable dans un avenir proche en effectuant de l'interférence à l'ARN chez *L. braziliensis*.

L'étude de plusieurs souches cliniques de *L. donovani* a montré de grandes différences de susceptibilité et résistance au traitement avec le stibogluconate de sodium, et la découverte de SNP dans certains membres de la sous famille  $\delta$ -amastine (voir ci-dessous) est l'un des éléments qui corrèle avec cette observation, suggérant également un rôle des amastines dans la résistance à ce traitement [436].

## 5.1. Structure et fonctions potentielles des amastines.

Les amastines sont des petites protéines, généralement prédites entre 20 et 25 kDa, et avec un point isoélectrique variant de 4 à 10. Les alignements des cadres de lecture à partir des séquences retrouvées chez *Leishmania* spp. et *Trypanosoma* spp. ont permis de mettre en évidence une structure bien conservée entre les membres de cette famille [437].

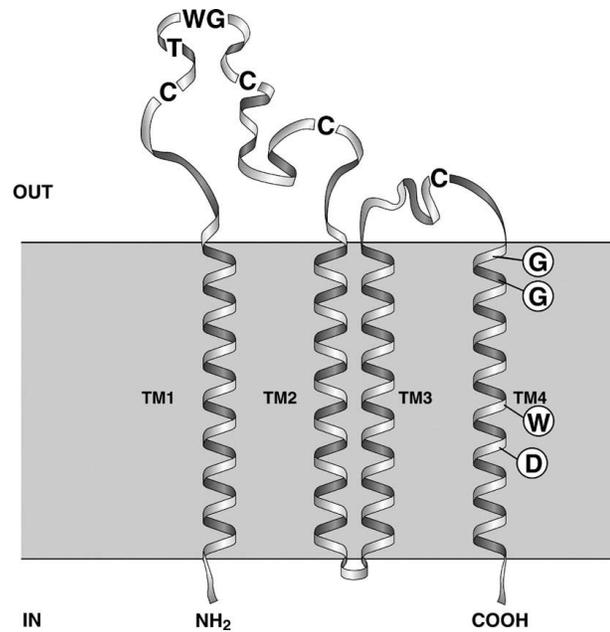


Figure 5-1 : Topologie prédite de la structure des protéines amastines.

Adapté de [437].

Les amastines ont généralement 4 domaines transmembranaires hydrophobes, avec de courtes extrémités N- et C-terminales, prédites du côté cytoplasmique, et deux domaines extracellulaires (Figure 5.1). Le premier est composé d'une soixantaine d'acides aminés, et contient une séquence signature conservée chez la moitié des protéines amastines putatives annotées, de 11 acides aminés, placée entre deux cystéines pouvant potentiellement stabiliser cette structure. La séquence signature constitue éventuellement le domaine

fonctionnel de ces protéines : C-[IVLYF]-[TS]-[LFV]-[WF]-G-X-[KRQ]-X-[DENT]-C. Le second domaine extracellulaire est plus court, et est variable au niveau de sa taille (autour de 20 acides aminés) et de sa séquence. Ce domaine pourrait intervenir dans la détermination de la spécificité des différents membres de la famille des amastines [437]. Chez différentes souches de *T. cruzi*, plus de variabilité est observable dans ces boucles extracellulaires [438].

Les séquences d'acides aminés de ces protéines ne partagent aucune similarité avec les protéines déjà caractérisées chez les autres organismes, et leur fonction est donc toujours inconnue. Ses domaines transmembranaires nous permettent de spéculer sur un rôle dans le transport d'ions, métaux, nutriments ou métabolites à travers la membrane plasmique. La plupart des amastines étant exprimées une fois que le parasite est différencié en amastigote, donc une fois le parasite à l'intérieur du phagolysosome, les amastines ont peut-être un rôle dans le maintien de l'homéostasie dans un environnement acide [437]. Ceci est également suggéré par la présence de deux amastines dans le protéome de la vacuole contractile chez *T. cruzi*, au moins au stade épimastigote [127]. Outre ce rôle de transporteur, les amastines peuvent également avoir un rôle dans la transduction de signaux offrant un senseur au parasite ou manipulant l'environnement de l'hôte [439]. Enfin, les gènes amastines sont fréquemment alternés avec les gènes tuzines dans le génome des parasites, suggérant une possible interaction entre ces deux protéines de fonctions inconnues. Chez *T. cruzi*, les quatre amastines sont présentes au moins sur la membrane plasmique pelliculaire et sur la membrane flagellaire [440].

## **5.2. Distribution et classification des amastines.**

Les amastines font partie d'une grande famille multigénique [437, 439]. Initialement, 54 gènes ont été retrouvés chez *L. major* et 37 chez *L. infantum*, leur séquence d'acides aminés partagent 24 à 98% d'identité [437], soulignant de possibles modifications fonctionnelles chez certains d'entre eux. Depuis la finalisation des séquencages des génomes de ces espèces, et en plus le séquencage d'autres souches de *Leishmania*, 28 à 54 amastines sont annotées par génome chez les *Leishmania* pathogènes de l'homme, un peu moins de la moitié de ces gènes est retrouvé chez le parasite non pathogène humain *L. tarentolae* et 2 à 6 gènes seulement dans les génomes de

*Trypanosoma* spp. (Tableau 5). La multiplication des gènes a donc été principalement effectuée chez *Leishmania* suite à sa séparation avec le genre *Trypanosoma*, et la multiplication semble corrélée avec la pathogénicité des espèces. Les amastines ont également été recherchées dans deux autres génomes de *Trypanosomatidae* infectant des artropodes : 8 sont présentes chez *Leptomonas seymouri*, infectant uniquement des insectes (*Dysdercus suturellus*) et 10 chez *Crithidia* spp. (Figures 1-1 et 3-1) [439]. *Leptomonas* et *Crithidia* sont phylogénétiquement plus proches du genre *Leishmania* que du genre *Trypanosoma* [439].

Tableau 5 : Nombre de gènes codant pour des protéines Amastines retrouvés par génome de *Trypanosomatidae*.

Amastines	<i>L. bra</i>	<i>L. inf</i>	<i>L. maj</i>	<i>L. mex</i>	<i>L. tar</i>	<i>T. bru</i>	<i>T. cru</i>	Total
$\alpha$	1	1	1	1	1	2	0	7
$\beta$	2	2	3	3	2	0	2	14
$\gamma$	6	4	4	4	3	0	0	21
$\delta$	38	33	46	20	6	0	4	147
Total	47	40	54	28	12	2	6	187

*L. bra* : *L. braziliensis*; *L. inf* : *L. infantum*; *L. maj* : *L. major*, *L. mex* : *L. mexicana*; *L. tar* : *L. tarentolae*; *T. bru* : *T. brucei* TREU927; *T. cru* : *T. cruzi* Esmeraldo-like. Données issues de [161, 162, 439].

La distribution des gènes amastines sur les chromosomes est assez équivalente entre les principales espèces de *Leishmania*. Dans le cas de *L. major*, les séquences sont dispersées sur 7 chromosomes, la principale concentration étant sur le chromosome 34 où 22 gènes sont répétés en tandem dans une région d'environ 157 kb, chacun de ces gènes étant séparé par celui de la tuzine. La seconde région présentant une concentration de ces gènes est sur le chromosome 8, où 17 amastines sont répétées quasiment successivement en tandem dans une région d'environ 100 kb.

Une large étude évolutive des amastines a été effectuée, prenant en considération leur phylogénie, leur emplacement dans le génome, leur séquence d'acides aminés, leur régulation en fonction des stades de vie des parasites ou encore la séquence du 3'UTR régulateur [439]. Cette étude a été effectuée avec les génomes alors disponibles *L. major*, *L. braziliensis*, *L. infantum*, *T. cruzi* et *T. brucei*, en plus des séquences identifiées chez *Leptomonas* et *Crithidia*. La phylogénie de ces gènes fait ressortir quatre sous-familles

d'amastines (nommées  $\alpha$ ,  $\beta$ ,  $\gamma$  et  $\delta$ , Figure 5-2), confirmant et enrichissant des résultats précédemment obtenus avec *L. major* et *L. infantum* [437].

Les  $\alpha$ -amastines sont conservées chez tous les *Trypanosomatidae* et correspondent au niveau des chromosomes à une répétition en tandem de deux isoformes structurellement distinctes. Les rares amastines de *T. brucei*, qui ne présente pas de stade intracellulaire amastigote, sont toutes des  $\alpha$ -amastines. *T. cruzi* ne possède pas cette sous-famille, sans doute perdue lors de son évolution. Le gène ancestral des amastines, présent dans l'ancêtre commun des *Trypanosomatidae*, devait être ou ressembler à une  $\alpha$ -amastine. Cette sous-famille regroupe des gènes un peu plus grands que les autres, avec 3 à 10 hélices transmembranaires additionnelles. Aucune séquence signal n'a pu être identifiée et les membres de cette sous famille sont généralement exprimés de manière constitutive.

Les  $\beta$ -amastines sont également conservées chez tous les *Trypanosomatidae* et sont organisées sur les chromosomes en deux à trois copies répétées en tandem de deux isoformes structurellement distinctes. L'apparition de cette classe marque sans doute le début de la diversification des amastines, initiée par la transposition d'une  $\alpha$ -amastine du chromosome 28 au chromosome 30 dans le cas de *L. major*. Les  $\beta$ -amastines ont été perdues chez *T. brucei*, mais sont présentes chez *T. cruzi* (à l'inverse des  $\alpha$ -amastines). Les ARNm des  $\beta$ -amastines sont généralement surexprimés dans les formes insectes de leur stade de vie. Ceci est notamment confirmé au niveau protéique pour la  $\beta$ -Amastine LinJ.30.0920, deux fois plus abondante dans des fractions membranaires de promastigotes que d'amastigotes [441]. Une étude comparative de transcriptome effectuée à partir de parasites *L. infantum* isolés de patients atteints de leishmaniose viscérale et de leishmaniose cutanée post kala-azar (voir Chapitre 1) montre la surexpression de près de deux fois d'un transcrit de  $\beta$ -amastine dans la forme cutanée, ce qui reste suggestif d'un rôle de ces gènes au sein de l'hôte mammifère [442], peut être pour la préparation au retour dans un insecte vecteur. Les ARNm des deux  $\beta$ -amastines de *T. cruzi* sont surexprimés au stade insecte épimastigote chez quatre souches testées [440].



Les  $\gamma$ -amastines sont absentes dans les génomes du genre *Trypanosoma*, sans doute suite à une perte de ces gènes, mais sont conservées chez *Leishmania* spp. et également retrouvées chez *Leptomonas* et *Crithidia*. Ici encore, 2 à 4 gènes codant pour deux isoformes structurellement distinctes sont dupliqués directement en tandem. Ces gènes sont légèrement plus abondants chez les amastigotes, notamment chez les amastigotes axéniques cultivés en laboratoire. Lorsque ces gènes sont plus abondants chez les amastigotes intracellulaires, l'amplification reste modeste, autour de 1.9 fois.

Enfin les  $\delta$ -amastines, qui semblent dériver des  $\beta$ -amastines, représentent la classe la plus amplifiée au niveau génomique, retrouvée dans plusieurs types de loci. Elles sont répétées en tandem, mais cette fois en association avec les tuzines. Une petite sous-famille de proto- $\delta$ -amastines conservée chez les *Trypanosomatidae* et contenant des gènes répétés en tandem se sépare rapidement des autres. La grande majorité des  $\delta$ -amastines est spécifique du genre *Leishmania*, n'a sans doute pas été acquise chez les autres organismes et est largement surexprimée au stade amastigote intracellulaire du parasite, en moyenne 9.4 fois.

Les rares  $\delta$ -amastines provenant d'autres organismes sont des proto- $\delta$ -amastines. Les fonctions spécifiques conférées aux amastigotes par des amastines sont possiblement limitées à ce groupe, puisque celui-ci est le seul à contenir des gènes surexprimés à ce stade de vie chez des parasites intracellulaires. Les données supportent fortement l'hypothèse selon laquelle les  $\delta$ -amastines sont une partie importante des modifications évolutives des amastigotes de *Leishmania* à leur nouvel environnement dans le phagolysosome des macrophages, et ces protéines ont subi des mécanismes d'adaptation associés à l'acquisition du parasitisme des vertébrés. Deux  $\delta$ -amastines ont également été annotées suite au séquençage du génome de *L. tarentolae*. L'absence de cette sous-famille chez *Crithidia* et *Leptomonas* mais sa présence dans tous les génomes de *Leishmania* suggère que les  $\delta$ -amastines ont dû se développer et évoluer avec, ou très peu de temps après, l'origine du genre *Leishmania*. Chaque espèce de *Leishmania* présente des variations dans ces groupes de gènes, certaines présentant plus de gènes ou des groupements uniques en fonction des espèces [161]. Chez *T. cruzi*, les transcrits des  $\delta$ -amastines sont surexprimés dans plusieurs souches au stade amastigote, sauf chez une souche G présentant

une faible capacité d'infection par des parasites métacycliques ou trypomastigotes, mais une bonne capacité d'infection par les amastigotes extracellulaires *in vitro* comparativement aux autres souches [440, 443]. De plus, l'exposition de cellules hôtes avant l'infection à des protéines  $\delta$ -Amastines purifiées inhibe l'invasion. La surexpression de  $\delta$ -Amastine stimule la métacyclogénèse, ainsi que le passage du parasite d'amastigote à trypomastigote, et accélère le tropisme du parasite vers le foie. Les auteurs suggèrent que la  $\delta$ -amastine est un inhibiteur de l'invasion des cellules par les amastigotes. En étant moins infectieux, les parasites évitent d'alerter le système immunitaire. Les  $\delta$ -Amastines permettent en contrepartie d'accélérer la différenciation en trypomastigote pour pouvoir retrouver une forme infectieuse. Cette balance entre coûts (être moins infectieux) et bénéfiques (faible réponse immunitaire) participerait dans la survie du parasite à son stade intracellulaire [443].

### **5.3. Les amastines de *Trypanosoma cruzi*.**

Les amastines ont été initialement décrites chez *T. cruzi* comme des glycoprotéines membranaires spécifiquement exprimées au stade amastigote du parasite, l'ARNm étant 50 à 100 fois plus abondant à ce stade (les chiffres varient en fonction des publications). La glycosylation des protéines amastines est suggérée par la taille plus grande que prédite de ces protéines et leur capacité à lier la concanavaleine A, bien qu'aucun signal de glycosylation ne soit identifiable sur la protéine [444]. Le couplage de gènes amastines avec ceux de la tuzine, dont l'ARNm est lui moins abondant chez les amastigotes, a ensuite été caractérisé. Si les tuzines ont un très court 3'UTR, voir même pas du tout de 3'UTR, l'amastine possède au contraire un 3'UTR de 630 nt, dont le clonage en aval du gène rapporteur luciférase a permis de mettre en évidence son rôle dans l'expression stade spécifique de cet ARNm. Le 3'UTR joue non seulement un rôle dans la stabilisation des transcrits chez les amastigotes, mais également dans leur déstabilisation dans les autres stades [445, 446], suggérant plusieurs mécanismes différents de régulation. La différence d'abondance observée est liée à la modification de la demi-vie du transcrit, qui est 7 fois plus longue chez les amastigotes. Le tiers central du 3'UTR de l'amastine est responsable de cette régulation et lie un facteur *trans* protéique labile de 36 kDa, identifié par

protection à la RNase et des expériences de liaison covalentes aux UV [349, 447] (voir Tableau 4).

Des études ont ensuite été effectuées sur une famille de six petites protéines possédant un domaine de reconnaissance de l'ARN (RRM). Parmi elles, la protéine *TcUBP1* (*T. cruzi* uridine binding protein 1) est la seule qui a permis de détecter une liaison à l'ARNm de l'amastine. Cela a été observé *in vitro* en fusionnant *TcUBP1* à l'étiquette GST, permettant sa purification à partir de surexpression chez *E. coli*, puis en incubant les protéines purifiées avec des ARN totaux de *T. cruzi*. L'interaction a également été confirmée *in vivo*, la production d'un anticorps spécifique à *TcUBP1* ayant permis d'effectuer des co-immunoprécipitations des ARNm liés au complexe protéique de *TcUBP1* [448]. L'ARNm de l'amastine co-immunoprécipite également avec *TcPABP1*, aussi bien au stade trypanostigote qu'au stade épimastigote. L'ARNm de l'amastine a également été purifié par co-immunoprécipitation du complexe nucléoprotéique *TcDHH1* chez les épimastigotes, alors que *TcUBP1* ne semble pas présente dans ce complexe [305]. DHH1 étant une hélicase retrouvée dans les mécanismes de retrait de la coiffe et de dégradation 5' vers 3' des ARNm, la présence de l'ARNm de l'amastine est en accord avec une dégradation des ARNm de l'amastine dans ce stade de vie du parasite. Le fait que les ARNm ont majoritairement une queue poly(A) encore intacte suggère cependant que les ARNm de l'amastine sont dans un état de conservation non traduit, ou que le complexe purifié est dans les étapes initiales de dégradation.

*TcUBP1* présente une structure secondaire de types  $\beta\alpha\beta\beta\alpha\beta$ , pour une taille de 27 kDa [448]. Elle est présente en petite quantité dans le noyau dans des conditions physiologiques, mais cette accumulation est accentuée lors de stress avec l'arsenite, molécule généralement utilisée pour induire la formation de granules de stress chez les mammifères. Cette accumulation est accompagnée d'une augmentation des ARN poly(A) au niveau nucléaire, dont l'ARNm de l'amastine [449]. L'observation d'une concentration des ARNm dans le noyau peut être le résultat de deux processus, soit l'importation des ARN cytoplasmiques vers le noyau, soit l'arrêt de l'exportation des ARNm nouvellement synthétisés vers le cytoplasme.

## 5.4. Les amastines chez *Leishmania*.

Les amastines ont par la suite été caractérisées chez *Leishmania*, montrant également une expression de l'ARNm spécifiquement au stade amastigote du parasite chez *L. infantum*, *L. major* et *L. mexicana*. Plus précisément, le premier gène étudié correspondait à la  $\delta$ -amastine LinJ.34.1020 (anciennement LinJ34.0840). Chez *L. infantum*, la localisation membranaire a été confirmée en surexprimant cette protéine fusionnée avec la protéine fluorescente GFP. Il a également été montré que la régulation stade spécifique de l'amastine était obtenue par son 3'UTR, notamment par une modification de la demi-vie de son ARNm, fortement augmentée chez les amastigotes. Le principal facteur influant la régulation stade spécifique du transcrite est la modification du pH du milieu de culture [450].

Une identification plus précise des séquences du 3'UTR de l'ARNm des amastines a ensuite été réalisée en utilisant le gène LinJ.34.1010. Dans le cas de *Leishmania*, les 3'UTR sont généralement plus longs que chez *Trypanosoma*, et le 3'UTR de ces gènes atteint 1,8 kb. Les clonages de différentes parties de cette structure en aval du gène rapporteur luciférase ont permis de mettre en évidence qu'une séquence de 450 nt, située dans le dernier tiers du 3'UTR, conférait une augmentation de la traduction, spécifiquement au stade amastigote, sans pour autant modifier la demi-vie des ARNm. Ceci suggère que plusieurs mécanismes de régulation interviennent pour cette famille de gènes [342]. Cette séquence de 450 nt, correspondant à un *LiSIDER1*, est retrouvée en plusieurs exemplaires bien conservés dans le génome de *Leishmania*, dont plusieurs autres transcrits amastines.

Une autre étude a par la suite montré qu'une séquence de 100 nt, légèrement en aval de la séquence conservée de 450 nt, était également responsable de la stimulation de la traduction spécifiquement chez les amastigotes. Il n'est pas clair si ses deux séquences agissent conjointement ou si ce sont deux mécanismes de régulation indépendants l'un de l'autre. Au final, il en ressort que l'augmentation de la traduction des transcrits est effectuée suite à un stress thermique par les éléments de 450 et 100 nt, et l'augmentation de la stabilité des ARNm est elle obtenue par un stress d'acidité, impliquant des séquences non décrites [350].

Une étude globale par Northern blot de la stabilité des transcrits de l'amastine chez *L. infantum* a montré que 17 des 20 transcrits testés étaient préférentiellement exprimés au stade amastigote complètement différencié. Chez *L. major*, 44 des 55 gènes de l'amastine alors annotés possèdent un *LmSIDER1* dans la région intergénique en 3', avec une homologie allant de 64 à 83%, et donc supérieure à l'homologie globale des régions intergéniques, inférieure à 40% [437].



## Chapitre 6 : Les protéines à domaine Alba.

Les protéines à domaine Alba (de l'anglais *Acetylation Lowers Binding Affinity*) sont des protéines retrouvées initialement chez les *Archaea* puis dont des homologues ont été décrits chez les eucaryotes. Dans cette thèse, les protéines Alba ont été étudiées dans le contexte de régulation de transcrits des amastines. L'unique revue de littérature de ces protéines remonte à 2003 [451], cependant la majorité des découvertes sur ces protéines ont été effectuée ces cinq dernières années, notamment chez les parasites protozoaires.

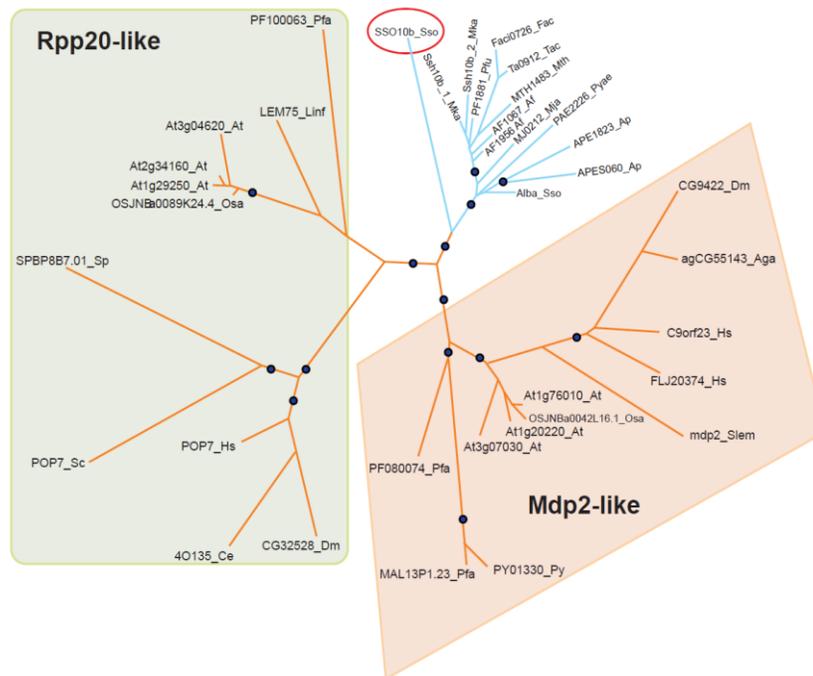


Figure 6-1 : Arbre phylogénique de type *maximum-likelihood* de la super famille Alba des *Archaea* jusqu'aux eucaryotes supérieurs.

Abréviations : Af, *Archaeoglobus fulgidus*; Aga, *Anopheles gambiae*; Ap, *Aeropyrum pernix*; At, *Arabidopsis thaliana*; Ce, *Caenorhabditis elegans*; Dm, *Drosophila melanogaster*; Fac, *Ferroplasma acidarmanus*; Hs, *Homo sapiens*; Linf, *Leishmania infantum*; Mja, *Methanococcus jannaschii*; Mka, *Methanopyrus kandleri*; Mta, *Methanothermobacter thermautotrophicus*; Osa, *Oryza sativa*; Pfa, *Plasmodium falciparum*; Pfu, *Pyrococcus furiosus*; Pyae, *Pyrobaculum aerophilum*; Py, *Plasmodium yoelii*; Sc, *Saccharomyces cerevisiae*; Slem, *Stylonychia lemnae*; Sp, *Schizosaccharomyces pombe*; Sso, *Sulfolobus solfataricus*; Tac, *Thermoplasma acidophilum*. Extrait de [451].

Le but de ce chapitre est donc de faire un bilan des connaissances actuelles sur ces protéines, qui ont acquis des rôles très diversifiés au cours de l'évolution. Phylogénétiquement, trois sous familles de protéines Alba ressortent : Celles des *Archaea*, puis deux groupes spécifiques des eucaryotes : les protéines Alba ressemblant aux protéines Rpp25 des RNase P / MRP, qui possèdent notamment un domaine RGG supplémentaire en C-terminal, et celles ressemblant aux protéines mdp2 ou Rpp20 des RNase P / MRP (Figure 6-1).

## 6.1. Les protéines Alba chez les *Archaea*.

Les protéines à domaine Alba ont été identifiées chez les *Archaea*, notamment chez le *Crenarchaea Sulfolobus acidocaldarius* (*SacAlba*). Ces organismes, vivant en conditions extrêmes dans les eaux thermales sulfurées, présentent un contenu en G+C d'environ 40% alors que leur température optimale de croissance est de 75°C. Les protéines *SacAlba* (initialement nommées HSNP pour *Helix-stabilising nucleoid protein*) ont été isolées à partir des nucléoïdes de ces organismes et il a été démontré que leur présence stabilisait la structure double brin de l'ADN [452]. Les histones n'étant pas ubiquitaires chez les *Crenarchaea* (revu dans [453]), les protéines Alba peuvent donc remplacer ou compléter ces structures dans un environnement où l'hybridation double brin de l'ADN semble difficile à maintenir. Alba est d'ailleurs retrouvée préférentiellement chez les *Archaea* thermophiles et hyperthermophiles [453]. Chez *Sulfolobus shibatae*, *SshAlba* constitue 4% des protéines cellulaires. *In vitro*, à 45°C et plus, cette protéine est capable d'augmenter le surenroulement négatif de l'ADN [454]. Ceci est habituellement effectué par des ADN gyrases pour ouvrir la double hélice et favoriser des processus tels que la transcription, la duplication de l'ADN ou pour permettre l'ancrage de protéines structurales. Des études de microscopie électronique ont permis de mettre en évidence qu'à une forte concentration, les protéines *SacAlba* sont capables de recouvrir l'ADN double brin et le protéger de digestions nucléolytiques *in vitro*, sans pour autant le compacter [455]. Les protéines Alba lient l'ADN double brin sans réelle spécificité pour une séquence [454, 456].

Chez les eucaryotes, les sirturines (famille de deacétylases apparentées à la protéine Sir2 des levures) deacétylent les histones afin de les séparer de l'ADN et d'entraîner le déploiement de l'hétérochromatine. Ceci stimule la transcription chez les métazoaires par

exemple (revu dans [457, 458]). Il a été démontré chez les *Archaea* que *SsoAlba* était déacétylée par *SsoSir2* au niveau de la lysine en position 16. *SsoAlba* acétylée inhibe la transcription, tandis que la mutation de la lysine 16, et donc l'impossibilité d'acétyler la protéine *SsoAlba*, stimule la transcription *in vitro* [456]. Cette observation a également été faite chez *Archaeoglobus fulgidus*, où il est proposé que la déacétylation d'*AfuAlba* par *AfuSir2* provoque l'oligomérisation des protéines *AfuAlba* sous forme de tétramère, et augmente son affinité pour l'ADN [459]. *SsoAlba* est acétylée par la protéine Pat [460].

Les génomes des *Archaea* ont deux copies du gène Alba, et les études de phylogénie ont permis de discriminer Alba1, gène le plus conservé entre les *Archaea* (60 à 90%), d'Alba2, présentant plus de divergence (60 à 70%) entre les espèces, Alba1 et Alba2 étant 30 à 40% identiques [461]. Alba1 hétérodimerise avec Alba2 *in vitro*, dans des conditions physiologiques. La protéine *SsoAlba2* est 20 fois moins exprimée que *SsoAlba1*. L'homodimère *SsoAlba2* lie l'ADN avec 10 à 20 fois moins d'affinité que l'homodimère *SsoAlba1* ou l'hétérodimère *SsoAlba1/SsoAlba2*. L'acétylation d'*SsoAlba1* n'affecte l'affinité de liaison que de 3 à 4 fois. Il est donc supposé par les auteurs que l'acétylation d'Alba1 sur la lysine 16 interviendrait plus dans une voie de signalisation entre protéines plutôt qu'une modulation directe de l'affinité de liaison à l'ADN. Des mesures de constante de dissociation (Kd) d'Alba1 ont démontré que cette protéine lie préférentiellement l'ADNbd (50 nM), alors qu'elle lie 5 fois moins l'ADNsb et 20 fois moins l'ARNsb. La mutation de la phénylalanine en position 60, permettant la dimérisation des dimères d'Alba, réduit l'affinité de *SsoAlba1* pour l'ADNbd, montrant l'importance de l'oligomérisation d'Alba pour la liaison à l'ADN [462]. Une autre étude a montré que *SshAlba* est elle capable de lier aussi bien l'ADNsb, l'ADNdb et l'ARN chez ces organismes, avec une préférence pour les ARNr. *SshAlba* a d'ailleurs été co-purifiée avec les ribosomes en utilisant une faible concentration saline. L'irradiation des cellules aux UV a permis de mettre en évidence la liaison de *SshAlba* aux ARNr 16S et 23S ainsi que quelques ARNm [463].

Deux protéines Alba1 ont été cristallisées chez ces organismes : *SsoAlba* (de *Sulfolobus solfataricus*) [464] et *PhoAlba* (de *Pyrococcus horikoshii*) [465]. Elles présentent une structure identique, présentant quatre feuillets  $\beta$  et deux hélices  $\alpha$ , ordonnés

en  $\beta 1-\alpha 1-\beta 2-\alpha 2-\beta 3-\beta 4$ . *PhoAlba* est structurellement similaire aux protéines de liaison à l'ARN telles que la partie C-terminale du facteur d'initiation 3 de *Bacillus stearothermophilus*, ou la protéine Yhhp d'*E. coli* impliquée dans la division cellulaire. Ces protéines sont très hydrophobes sur l'arrière pour permettre la dimérisation. *PhoAlba* est capable d'interagir à la fois avec les ARNm, l'ARNt de la tyrosine et son précurseur *in vitro*. Cette liaison ne semble cependant pas responsable d'une activité catalytique [465].

La purification du complexe RNase P chez l'*Archaea Methanothermobacter thermoautotrophicus* n'a pas permis l'identification de l'homologue Alba de cette espèce. De plus, la présence d'*AthAlba* n'augmente pas l'activité du complexe *in vitro* chez cet organisme [466], et cette protéine n'est pas requise pour reconstruire l'activité de ce complexe *in vitro* chez les *Archaea* [467-470].

## **6.2. Les protéines à domaine Alba chez les parasites protozoaires.**

### **6.2.1. *Toxoplasma gondii* (agent de la toxoplasmose)**

Chez l'*Apicomplexa Toxoplasma gondii*, deux protéines *TgAlba* sont généralement annotées. L'étude des protéines liant le promoteur du gène *ENO1*, dont l'expression est spécifique du stade bradyzoite, a permis l'identification de plusieurs facteurs nucléaires *in vitro*. Le principal facteur étudié est *TgNF3*, et les auteurs suggèrent qu'il agit tel un coordinateur des fonctions nucléaires et nucléolaires en modifiant les activités du nucléosome lors de la prolifération intracellulaire des tachyzoites virulents. Les protéines *TgAlba1* et *TgAlba2* (respectivement *TgNF5* et *TgNF6* dans l'étude) font partie des autres protéines identifiées dans ce complexe. Les immunolocalisations présentent des signaux principalement cytoplasmiques, mais une accumulation à la périphérie nucléaire suggère un rôle dans les deux compartiments. Un fractionnement sub-cellulaire semble avoir mis en évidence la présence de *TgAlba1* et 2 dans le noyau [471].

Tableau 6 : Les protéines à domaine Alba chez les *Apicomplexa* et les *Kinetoplastidae*

	Protéine	kDa	pI	Domaines	PTM	Localisation	Complexes	ARN / ADN cibles
<i>Toxoplasma gondii</i>	<i>TgAlba1</i> TGME49_021380	27.2	10.12	Alba (1.9E-26) RGG-box	Phospho. [472] 80S - 81T - 90T	Cytoplasmique Nucléaire [471, 473]	PABP, eIFs, DEAD-box RNA helicase, etc [473]	<b>ADN</b> <b>Promoteur <i>ENO1</i></b> [471]
	<i>TgAlba2</i> TGME49_018820	15.4	8.72	Alba (3.2E-11)	Phospho. [472, 474] 77T - 140S - 142S			<b>ARN</b> ~ 30 ARNm [473]
<i>Plasmodium berghei</i>	<i>PbAlba1</i> PBANKA_142330	26.9	10.87	Alba (3.9E-19) RGG-box	?	<b>Oocyte <i>P-granules</i></b> : [475]  <b>Complexe</b> <b>FACT-L</b> : [476] Nucléaire	<b>Oocyte <i>P-granules</i></b> : Dhh1, RAP55, eIF4E, PABP, <u>Alba1, 2 &amp; 3</u> Bruno, Musashi. [475]  <b>Complexe</b> <b>FACT-L</b> : [476] <u>Alba 3 &amp; 4;</u> Fact S; 40S & 60S, etc	<b>ARN</b> <b>Oocyte <i>P-granules</i></b> : <i>p25, p28, warp,</i> <i>rad51, gap45</i> , etc [475]
	<i>PbAlba2</i> PBANKA_135920	23.6	9.95	Alba (3.7E-11) RGG-box- Like	?			<b>ADN</b> <b>Complexe</b> <b>FACT-L</b> : [476]
	<i>PbAlba3</i> PBANKA_120440	12	10.17	Alba (1.8E-05)	?			
	<i>PbAlba4</i> PBANKA_136030	42.2	8.6	n.d.	?			
<i>Plasmodium falciparum</i>	<i>PfAlba1</i> PF3D7_0814200	27.3	11.13	Alba (3.2E-18) RGG-box	Phospho. [472, 477] 83S - 88S - 129T	Nucléaire (stade anneau) & cytoplasmique (Trophozoïtes & Schizontes) [478]	<b><i>PfAlba3</i> – <i>PfSir2A</i> deacetylation</b> [479]  <b>TARE6 IP</b> : [478] <u><i>PfAlba1, 2 &amp;</i></u> <u><i>4</i></u> eIF5A, histones, protéines ribosomales, Lsm2&4 etc  Voir aussi <b>Double</b> <b>hybride</b> [480]	
	<i>PfAlba2</i> PF3D7_1346300	25	8.32	Alba (1.9E-13) RGG-box- Like	Phospho. [472, 477] 4S - 7S - 98S	Nucléaire (stade anneau) & cytoplasmique (Trophozoïtes & Schizontes) [478]		<b>ADN (<i>PfAlba3</i>)</b> Promoter <i>var</i> [479] TARE1&2 [479] TARE6 [478, 479]
	<i>PfAlba3</i> PF3D7_1006200	12	9.94	Alba (5.3E-08)	Acetyl. [479] 22K Phospho. [472, 477] 3S – 4T – 15S – 25T – 36T	Nucléaire et périphérie nucléaire (stade anneau) [478, 479]  Nucléaire (Trophozoïtes & Schizontes) [479]		<b>ADN (<i>PfAlba4</i>)</b> Préférence pour la liaison à des structures d'ADN relâchées
	<i>PfAlba4</i> PF3D7_1347500	42.2	7.57	n.d.	Phospho. [472, 477] 119S	Nucléaire (stade anneau) & cytoplasmique (Trophozoïtes & Schizontes) [478]		<b><i>in vitro synth.</i></b> <b>ARN</b> <u><i>PfAlba1-4</i></u> poly(A) & poly(U) [478]
	( <i>PfAlba5</i> ) PF3D7_0216200	27.2	9.74	Alba (1.2E-05)	?	?		
	( <i>PfAlba6</i> ) PF3D7_1202800	11.4	9.33	Alba (1.1E-08)	?	?		

	Protéine	kDa	pI	Domaines	PTM	Localisation	Complexes	ARN / ADN cibles
<i>T. cruzi</i>	<i>TcAlba1</i> Tc00.104705350408 9.60	13.5	9.78	Alba (5.1E-17)	?	?	<i>TcPOP2</i> - <b><i>TcAlba3</i></b> & <i>TcDhh1</i> - <b><i>TcAlba3</i></b> (Ferrarini MG, Avila AR, non publié)	<b><i>TcAlba3&amp;4</i></b> Association aux polysomes chez épimastigotes et lors de stress nutritifs [481]
	<i>TcAlba2</i> Tc00.104705350408 9.70	13	8.48	Alba (8.5E-26)	Phospho. [482] 2S	?		
	<i>TcAlba3</i> (Tc00.10470535108 77.30)	25.1	10.34	Alba (3.4E-30) RGG-box	Phospho. [482] 144S – 157S	Cytoplasmique (polysome & <i>P-granules</i> )		
	<i>TcAlba4</i> (Tc00.10470535108 77.40)	24.5	10.24	Alba (3.4E-30) RGG-box	Phospho. [482] 144S – 157S	Cytoplasmique (polysome & <i>P-granules</i> )		
<i>T. brucei</i>	<i>TbAlba1</i> Tb11.02.2040	14	7.67	Alba (2.9E-21)	?	Cytoplasmique & Granules de stress [346, 483] Matrice flagellaire (forme sangu.) [144]	<i>TbDhh1</i> - <b><i>TcAlba3/4</i></b> [483]  <b>IP-<i>TbAlba</i>:</b> [346] eIF4E4 PABP	<b>ARN 3'UTR GPEET:</b> [346] <i>TbAlba1</i> à 4  <i>TbAlba3</i> & 4 <b>Co-localise with Poly(A)+RNA</b> [483]  <i>TbAlba2&amp;3</i> <b>Partiellement associées aux polysomes</b> [346]
	<i>TbAlba2</i> Tb11.02.2030	12.8	9.8	Alba (1.4E-22)	Phospho. [412] 24S	Cytoplasmique & Granules de stress [346, 483] Matrice flagellaire (forme sangu.) [144]		
	<i>TbAlba3</i> Tb927.4.2040	20.8	10.01	Alba (1.0E-28) RGG-box	Phospho. [484] 4Y	Cytoplasmique & Granules de stress [346, 483] Matrice flagellaire (forme sangu.) [144]		
	<i>TbAlba4</i> Tb927.4.2030	22.7	10.31	Alba (1.0E-28) RGG-box- Like	?	Cytoplasmique & Granules de stress [346, 483] Matrice flagellaire (forme sangu.) [144]		
<i>L. major</i>	<i>LmAlba13</i> LmjF.13.0450	13.3	5	Alba (5.4E-24)	?	?	IP <i>LmIF4E4</i> (Pro et Ama): Machinerie traduction [380]	?
	<i>LmAlba20</i> LmjF.34.2580	22.6	10.12	Alba (7.8E-27) RGG-box	?			
<i>L. infantum</i>	<i>LiAlba13</i> LinJ.13.0270	13.4	4.8	Alba (1.3E-24)	?	Voir Chapitre 10	Voir Chapitre 10	Voir Chapitre 9
	<i>LiAlba20</i> LinJ.34.2410	22	19.93	Alba (4.1E-26) RGG-box	?			

Une étude plus ciblée sur les protéines *TgAlba* a récemment été publiée [473]. Des délétions des gènes ont permis de suggérer l'implication des protéines *TgAlba* dans la régulation de la réponse au stress alcalin, et ainsi la différenciation du parasite sous forme bradyzoïte. Ces protéines interagissent avec une trentaine de transcrits sans fonction commune clairement identifiable, dont leurs propres ARNm. Au niveau protéique, les protéines *TgAlba* interagissent avec la machinerie de traduction, dont la PABP et des facteurs d'initiation de la traduction. Les études de délétions ont montré que *TgAlba1* (*Rpp25-like*) est nécessaire pour la traduction de *TgAlba2* (*Rpp20-like*), et cette action est effectuée par le 3'UTR de *TgAlba2*, démontrant à la fois une action des protéines *TgAlba* dans le contrôle de la traduction et une dépendance des protéines pour leur co-traduction.

### **6.2.2. *Plasmodium* spp (agent de la malaria).**

Parmi les recherches effectuées chez l'*Apicomplexa Plasmodium berghei*, une première étude analysait les *P-granules* d'oocytes, et a caractérisé les complexes mRNP contenant des ARNm en répression traductionnelle au sein de ce stade de vie. En effectuant des co-immunoprécipitations distinctes à partir de la DEAD-box RNA helicase DOZI (homologue de DHH1) et à partir de CITH (homologue de RAP55), il a été démontré que ces deux protéines interagissent ensemble dans un complexe contenant en plus eIF4E, PABP, des homologues de protéines de liaison à l'ARN Bruno et Musashi et également les protéines *PbAlba1*, 2 et 3. L'absence de protéines impliquées dans la dégradation des ARNm soutient le fait que les transcrits sont dans un stade de latence traductionnelle [475]. Une seconde étude s'est concentrée sur la transcription dans les gamétocytes mâles par le complexe FACT (pour *Facilitates chromatin transcription*) et a montré son importance pour permettre la génération de gamètes mâles fertiles. L'une des protéines de ce complexe, un homologue de SPT16, présente une localisation nucléaire, co-précipite avec plusieurs protéines ribosomales et également avec les protéines *PbAlba3* et 4 [476].

Chez *Plasmodium falciparum*, deux études impliquant les protéines *PfAlba* montrent leur implication dans la liaison à l'ADN. La première porte sur *PfAlba3* et son domaine RGG. Deux homologues de Sir2 (*PfSir2A* et B) sont présents chez cet organisme et impliqués dans l'inhibition de la transcription *via* l'hétérochromatine de différents gènes *var*. *PfAlba3* est exprimée dans tous les stades de développement asexués, et sa liaison à

l'ADN inhibe la transcription. Cette liaison à l'ADNdb semble non spécifique, mais l'acétylation de *PfAlba3* diminue l'affinité de la protéine pour l'ADN. Si la forme prédominante dans la cellule est acétylée (80%), *PfAlba3* interagit et est déacétylée par *PfSir2A*, nucléaire et retrouvée principalement au niveau des régions télomériques et subtélomériques. *PfAlba3* est localisée dans le noyau (à sa périphérie au stade anneau et a une localisation plus ponctuée chez les trophozoïtes et les schizontes) et se lie à ces régions, notamment au niveau du promoteur du gène *var*, TARE1 et TARE6 (pour *Telomere-Associated Repetitive Elements*) [479]. La seconde étude a démontré des protéines se liant justement aux éléments TARE6, modulant l'expression des gènes de virulence, chez ce même parasite. L'association directe des protéines *PfAlba1* à 4 a été démontrée, confirmant également la localisation nucléaire de ces protéines dans le stade anneau. Une localisation cytoplasmique dans les stades plus matures, où les protéines forment des *speckles*, est visualisable. Il semble que contrairement à l'étude précédente, une certaine spécificité de liaison est présente, notamment pour *PfAlba4*, qui *in vitro* se lie préférentiellement aux séquences d'ADN relâchées. La longueur minimale de l'ADN double brin pour permettre la liaison d'*PfAlba* est de 14 pb. Les études de localisation des protéines ont montré un fort signal cytoplasmique pour les protéines *PfAlba1*, 2 et 4 durant les stades trophozoïte et schizonte, similaire aux *P-granules* observés dans les gamétocytes de *P. berghei*. En revanche, *PfAlba4* est nucléaire et colocalise avec *PfSir2* pour le stade anneau du parasite. *PfAlba1*, 2 et 4 présentent une modification de leur localisation nucléaire ou cytoplasmique en fonction des stades de vie. Toutes trois sont nucléaires au stade anneau, et se concentrent dans le cytoplasme au stade trophozoïte et schizonte. Ce genre de modification est souvent associé à des événements de protéolyse ou de sumoylation. Enfin, des études sur la capacité des protéines *PfAlba* à lier des séquences d'ARN poly(A) et poly(U) montrent des différences d'affinité entre les paralogues. *PfAlba1* a deux fois plus d'affinité pour le Poly(A) que pour les Poly(U), alors que *PfAlba2* et 4 semblent dimeriser (confirmé dans une étude de double hybride [480]) pour lier les deux sondes avec la même efficacité et que *PfAlba3* a une forte affinité pour les Poly(A). Des études de compétition ont permis de mettre en évidence que *PfAlba1* utilise le même domaine pour lier l'ADN et l'ARN. Enfin, *PfAlba3* dimérise en solution [478].

L'expression des gènes *var* est conditionnée par leur localisation au sein du noyau, participant ainsi dans l'expression monoallélique, la variation phénotypique et la pathogénicité du parasite. La polymérisation de l'actine participe dans le déplacement de ces gènes d'un compartiment périnucléaire répressif à un compartiment périnucléaire où la traduction est active. Il a été démontré que ce processus de localisation est dépendant de la présence d'un élément dans un intron du gène *var*, permettant de recruter le complexe de l'actine. Parmi les protéines de liaison à cet élément, trois protéines Alba ont été identifiées [485]. Cependant leur rôle dans ce processus n'a pas été plus longement étudié, puisque les études précédemment décrites montrent que les protéines *PfAlba* lient l'ADN sans spécificité de séquence.

Sans pouvoir actuellement donner une signification biologique à cette observation, notons que 3 protéines Alba de *P. falciparum* ont été identifiées parmi le protéome des vacuoles alimentaires [486]. Ces structures, de type lysosomale, sont impliquées dans la digestion de protéines telles que l'hémoglobine afin de fournir au parasite les acides aminés requis pour sa propre traduction et la maintenance de la stabilité osmotique de l'érythrocyte, inhibant ainsi une hémolyse prématurée de la cellule hôte. Les protéines Alba sont généralement très abondantes, il est envisageable que cette observation soit simplement un cas de recyclage de la protéine.

### **6.2.3. *Trypanosoma* spp.**

Chez *Trypanosoma cruzi*, quelques études mentionnent la présence de *TcAlba* dans des activités de liaison à l'ARN. Des co-immunoprécipitations ont été effectuées avec *TcDHH1* et *TcPOP2* (déadénylase), montrant la liaison de ces protéines avec *TcAlba3* (selon le rapport de thèse de Ferrarini MG, équipe Avila AR, Curitiba, Brésil). Une analyse des protéines liées aux ARNm en cours de traduction, que ce soit dans le stade épimastigote ou lors de stress nutritionnel a été effectuée. Malgré une forte concentration en sel, *TcAlba3* (et/ou 4, les séquences sont très similaires) est retrouvée associée aux polysomes dans des conditions physiologiques et lors de stress. Une faible concentration en sel confirme ces résultats et permet de plus l'identification de *TcAlba3* (ou *TcAlba4*) dans les fractions post-polysomales [481]. Ces deux données suggèrent que *TcAlba3* et/ou 4 est capable de transiter avec des ARNm entre granules de stress et la machinerie de traduction.

Chez *T. brucei*, l'étude de la régulation du gène *GPEET*, spécifiquement exprimé au stade procyclique, a montré l'association des quatre protéines *TbAlba* avec le *glycerol responsive element* (GRE), riche en uridines et présent dans le 3'UTR du transcrit. Les quatre protéines *TbAlba* de *T. brucei* interagissent ensemble, tel que montré par co-immunoprécipitation et également suggéré par le fait que la déplétion de *TbAlba3* par RNAi réduit l'expression de *TbAlba1* et 2. L'inhibition de l'expression des 4 protéines *TbAlba* réduit la traductibilité d'un gène rapporteur CAT dont l'ARNm a été fusionné avec le 3'UTR de *GPEET*, mais étrangement ne semble pas affecter la traduction d'autres ARNm endogènes, dont *GPEET*. Des expériences de «*pull down*» ont montré l'association des protéines *TbAlba* avec les PABP et eIF4E4 (via un pont ARN pour cette dernière, non testé pour PABP). De plus, seules les protéines *TbAlba2* et *TbAlba3* co-sédimentent partiellement avec les polyribosomes. La déplétion de *TbAlba3* est responsable d'une réduction de la croissance des parasites, suggérant un rôle centrale. Les études de localisation montrent que les protéines *TbAlba* sont toutes cytoplasmiques, et forment des agrégats lors de stress nutritionnelles qui co-localisent avec les ARN poly(A)<sup>+</sup>. Enfin *TbAlba* ne semble pas jouer de rôle dans la maturation des ARNt [346]. La seconde étude a montré la colocalisation des protéines *TbAlba3* et *TbAlba4* avec *TbDhh1* et avec des ARNm poly(A)<sup>+</sup> dans des granules de stress nutritionnel. Comme observé dans l'étude précédente, la déplétion de *TbAlba3* et *TbAlba4* inhibe la croissance cellulaire, et plus de détails morphologiques sont donnés chez les procycliques. Les cellules n'exprimant pas la protéine sont près de deux fois plus allongées à cause d'une polymérisation active des microtubules à l'extrémité postérieure, et le noyau apparaît plus fréquemment en position postérieure par rapport au kinetoplaste. *TbAlba3* et *TbAlba4* sont présentes durant tous les stades de vie du parasite au sein de l'insecte vecteur, sauf lors du stade de transition dans la région proventriculaire (de trypomastigote vers épimastigote). L'expression forcée de *TbAlba3* à ce stade perturbe la migration du noyau vers l'extrémité postérieure de la cellule lors de la différenciation, montrant pour la première fois l'implication d'une protéine de liaison à l'ARN dans le développement des trypanosomes [483]. L'étude du protéome flagellaire de la forme sanguine de *T. brucei* a permis l'identification des 4 protéines *TbAlba* dans la matrice du flagelle [144].

Notons que le complexe RNase P nucléaire de *T. brucei* a récemment été caractérisé. Les protéines *TbAlba* ne sont pas présentes et une seule protéine, PRORP, semble faire partie de cette structure [487], tandis que son paralogue PROP2 se dirige vers la mitochondrie.

#### **6.2.4. *Leishmania* spp.**

La principale étude mentionnant les protéines Alba est chez *L. major*, où la caractérisation des partenaires de *LmIF4E1* (stades promastigotes et amastigotes) et de *LmIF4E4* (promastigotes seulement) ont identifié les deux protéines *LmAlba*, suggérant leur implication dans la traduction. Le nombre important de protéines identifiées dans l'étude ne permet pas d'avoir une idée précise des autres partenaires des *LmAlba* via ce complexe. Les études du phosphoprotéome de *Leishmania* n'ont pas caractérisé de protéines Alba [410, 488]. Les protéines à domaine Alba de *Leishmania* sont discutées plus en détail dans les Chapitres 9, 10 et 11 de cette thèse.

### **6.3. Les protéines à domaine Alba chez les eucaryotes supérieurs comme sous unités des RNase P / MRP.**

Les RNases P sont des ribozymes retrouvées dans tous les organismes, des bactéries jusqu'aux eucaryotes supérieurs. Elles sont constituées d'un ARN catalytique nommé RPR (H1 chez l'homme). Cependant, leur contenu protéique (les Rpp) se complexifie au cours de l'évolution : une seule Rpp est présente chez les bactéries, cinq sont identifiées chez les *Archaea*, neuf sont caractérisées chez la levure *S. cerevisiae*, et au moins dix sous-unités sont présentes dans le complexe RNase P de l'homme (Figure 6-2). Les RNase P ont une activité endoribonucléase permettant la maturation des ARNt à partir des ARNt précurseurs, mais sont également impliquées dans d'autres fonctions selon les organismes. Les autres substrats de ce complexe chez les bactéries sont les précurseurs d'ARNtm (ARN transfert-messenger), des ARNr, quelques ARNm et des *riboswitch*. Chez les eucaryotes ce complexe est également impliqué dans la transcription de petits ARNnc par l'ARNPIII, dans l'efficacité de transcription des ARNr par l'ARNPI (pour revue [489, 490]) et dans la dégradation de l'ARNm de Viperin, protéine antivirale [491].

Les protéines Rpp20 et Rpp25, appartenant à la super famille des protéines à domaine Alba, sont retrouvées uniquement dans les complexes RNase P / MRP des eucaryotes. Nous allons concentrer la suite de cette partie sur ces complexes chez les eucaryotes, plus précisément sur les RNase P nucléaires. En effet, d'autres complexes RNase P mitochondriaux et plastidiques ont été décrits, ils semblent être constitués d'une à trois protéines, sans ARN catalytique, à l'image de la RNase P nucléaire de *T. brucei* [487]. Aucune étude ne montre l'association des protéines Rpp20 et Rpp25 dans ces structures [492, 493].

Chez les eucaryotes, il est considéré que la duplication des gènes des RPR a permis la spécialisation de plusieurs formes de RNase P [494]. La RNase P classique, telle que précédemment citée, a ainsi donné naissance à la RNase MRP, nucléolaire, qui est spécialisée dans la maturation des ARNr. Celle-ci partage la plupart de ses sous-unités avec la RNase P (Figure 6-2). En plus de son activité sur les ARNr, il a été démontré que la RNase MRP coupait le 5'UTR du transcrit de CLB2 dans le cytoplasme, permettant ensuite sa dégradation par l'exoribonucléase Xrn1p [270]. L'association de la RNase MRP avec TRET (pour *telomerase reverse transcriptase*) définit la RNase MRP-TRET. TERT utilise l'ARN de la RNase MRP pour la génération d'ARN double brins, utilisés par Dicer pour générer des petits ARN interférants inhibant l'expression de l'ARN MRP spécifiquement [495]. À ce jour, RNase MRP et MRP-TERT n'ont pas été caractérisées chez les *Kinetoplastidae* [494].

Rpp20 et Rpp25 sont des sous-unités permanentes et fonctionnelles de la RNase P. Elles sont associées à la chromatine des ARNr activement transcrits. Rpp20 lie les gènes des ARNr en début de phase G1 tandis que Rpp25 s'associe uniquement en fin de phase G1, quand la transcription par les ARNPI et ARNPIII augmente, suggérant un rôle d'activateur transcriptionnel de Rpp25 [496]. Chez *S. cerevisiae*, Pop6 et Pop7, les homologues respectifs de Rpp25 et Rpp20, lient le domaine P3 de RPR uniquement sous forme d'hétérodimère [497], et ceci est également observé chez l'homme [498]. La double capacité de liaison de ces protéines pour l'ARN et l'ADN suggère leur rôle dans l'adressage de la RNase P vers la chromatine. Aucune spécificité de séquence n'a encore été démontrée pour l'ADN [489]. Chez l'homme, les niveaux d'expression de Rpp20 et

Rpp25 sont co-dépendants et une seule copie de ces deux sous unités est présente au sein des RNase P et MRP. Ces protéines s'accumulent dans les nucléoles, et il semble que la localisation de Rpp20 est dépendante de celle de Rpp25 [499]. Rpp20 semble avoir une activité ATPase, et lie notamment une protéine de choc thermique HSP27 [500].

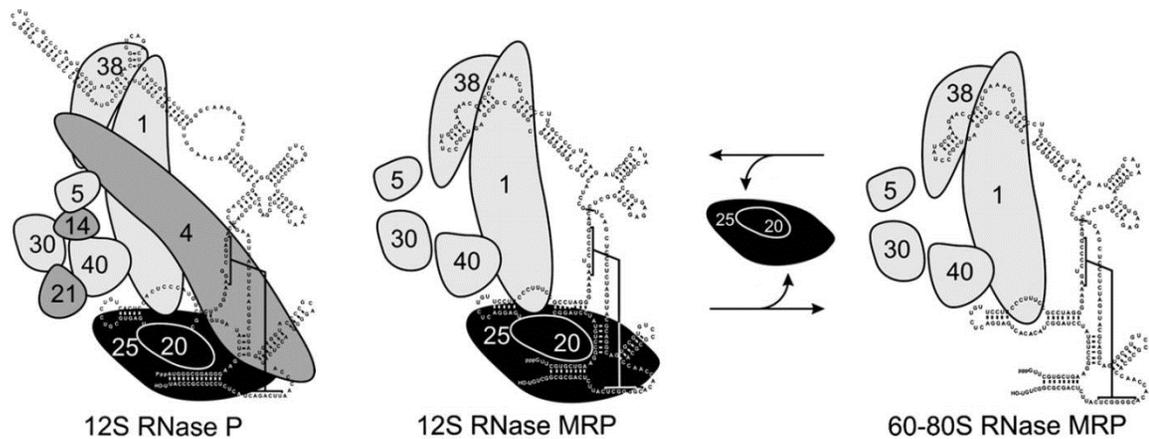


Figure 6-2 : Composition des principaux complexes des RNases P et MRP humaines.

Les protéines communes aux trois complexes sont en gris claire, les sous-unités plus spécifiques de la RNase P sont en gris foncé. Les protéines Rpp20 et Rpp25, en noir, sont uniquement associées aux complexes de 12S. 1, hPop1; 4, hPop4; 5, hPop5; 14, Rpp14; 30, Rpp30; 20, Rpp20; 21, Rpp21; 25, Rpp25; 38, Rpp38; and 40, Rpp40. hPop5 n'a pas été retrouvé dans les expérimentation et sa présence est uniquement suggestive. Adapté de [501].

Des expériences de sédimentation dans du glycérol montrent que la RNase P sédimente à 12S, tandis que la RNase MRP sédimente à 12S et 60-80S principalement. La RNase MRP étant impliquée dans la maturation des ARNr, il est estimé que la sédimentation autour de 60-80S représente des ribosomes en cours d'assemblage. Il a été montré que Rpp20 et Rpp25 sont retrouvées associées uniquement à la RNase MRP de 12S (Figure 6-2), suggérant une interaction transitoire de ces sous-unités, potentiellement dans un phénomène d'inhibition de l'activité de la RNase MRP [501].

Notons que chez les plantes, peu d'informations ont été trouvées sur les protéines Alba. L'étude du protéome nucléaire du riz en réponse à la déshydratation a permis l'identification de plusieurs protéines présentant au moins 2.5X de différence d'expression. Parmi les 109 protéines concernées, Alba est retrouvée en moindre quantité lors du stress,

supposant sa régulation d'expression ou de localisation en réponse au stress chez les plantes. Les autres protéines identifiées assurent différents processus cellulaires : régulation transcriptionnelle, remodelage de la chromatine, signalisation, défense cellulaire et dégradation protéique [502]. Une seconde étude s'est intéressée à l'analyse protéomique des protéines de liaison à l'ARN associées au cytosquelette. Ici encore, parmi les 148 protéines identifiées, 3 protéines à domaine Alba sont mentionnées [503].

## Chapitre 7 : Hypothèses et objectifs

Le genre *Leishmania*, ainsi que les autres *Trypanosomatidae*, contiennent plusieurs parasites protozoaires qui ont rapidement divergé des autres eucaryotes. Cette séparation précoce leur a permis des adaptations évolutives spécifiques. Par exemple, il ne semble pas y avoir de régulation de l'expression des gènes au niveau de l'initiation de transcription par l'ARN polymérase II, et les études du contenu protéique de leur génome montrent la présence de nombreuses protéines sans homologues avec des protéines connues chez d'autres eucaryotes (plus de 60% des gènes codant pour des protéines chez *L. infantum* sont encore annotés comme hypothétiques sur TritypDB.org, v4.0).

*Leishmania infantum* est un modèle d'étude de choix pour plusieurs raisons. En tant qu'agent de la leishmaniose viscérale, une maladie négligée responsable de 50 000 décès chaque année, il est urgent de mieux comprendre cet organisme afin de proposer de nouveaux traitements efficaces et à un coût acceptable pour les populations vivant dans des zones à risque. Le développement d'un vaccin serait également une avancée majeure contre les leishmanioses. Comme les autres *Trypanosomatidae*, l'absence de régulation transcriptionnelle fait de ces organismes un modèle d'étude important des mécanismes de régulation post-transcriptionnelle, incluant les mécanismes de stabilisation ou dégradation des ARNm, ou les mécanismes de régulation traductionnelle. Certaines découvertes chez les *Trypanosomatidae* ont par ailleurs été validées chez d'autres eucaryotes supérieurs, l'exemple le plus connu étant l'épissage en *trans*, identifié dès 1982 chez *T. brucei*, puis ensuite chez d'autres organismes. Ceci montre la pertinence de ce modèle d'étude à une échelle plus vaste que la découverte de mécanismes moléculaires spécifiques de cet ordre.

L'un des axes de recherche du laboratoire du Dre Papadopoulou s'intéresse à la régulation et au(x) rôle(x) des amastines. Ces protéines membranaires, retrouvées uniquement chez les *Kinetoplastidae* et principalement chez *Leishmania* spp. et *T. cruzi*, sont exprimées spécifiquement dans la forme amastigote des parasites (Chapitre 5). La détermination du rôle de ces protéines, ainsi que la confirmation de leur implication dans la virulence du parasite sont difficiles. En effet, il n'est pas possible d'effectuer d'étude de

type interférence à l'ARN pour estimer l'effet de l'absence de ces protéines, et leur large duplication dans les génomes de *Leishmania* spp. rend impossible la génération de souches déficientes par recombinaison homologue de tous les paralogues de cette famille. Puisque les amastines font partie de protéines des plus immunogéniques d'après des études chez la souris, et qu'elles sollicitent en plus une forte réponse immunitaire chez l'homme, elles restent une cible de choix pour la compréhension des mécanismes moléculaires de l'adaptation et la survie du parasite chez l'hôte mammifère. L'élucidation de leur mécanisme de régulation participerait à mieux comprendre les mécanismes d'adaptation du parasite dans le phagolysosome du macrophage et sa communication avec cet environnement. Les études précédemment effectuées au sein du laboratoire se sont concentrées sur l'un des 43 gènes de l'amastine, LinJ.34.1010, de la sous-famille des  $\delta$ -amastines, utilisé comme modèle d'étude pour l'expression stade-spécifique des protéines chez *Leishmania*.

La majorité des mécanismes de régulation post-transcriptionnelles chez *Leishmania* est effectuée par des séquences présentes dans le 3'UTR des transcrits. L'utilisation de constructions utilisant le gène rapporteur luciférase sous contrôle de différentes portions du 3'UTR de l'amastine a montré l'implication d'un rétroélément éteint de type SIDER1 dans la stimulation de la traduction du transcrit spécifiquement au stade amastigote. Ceci a permis d'expliquer un premier phénomène régulateur de ce transcrit, mais aussi d'offrir une première évidence expérimentale du rôle régulateur des SIDERs chez *Leishmania*. Cependant, cette première avancée dans la régulation de l'expression stade spécifique de la protéine n'explique pas comment le transcrit de l'amastine s'accumule 8 à 10 fois plus au stade amastigote qu'au stade promastigote.

Notre **hypothèse de départ** est donc que des mécanismes d'accumulation des amastines spécifiquement chez les amastigotes, et / ou des mécanismes de dégradation actifs au stade promastigote doivent pouvoir expliquer ce phénomène.

Chez les autres eucaryotes, les séquences les mieux caractérisées impliquées dans la dégradation des transcrits sont les éléments riches en bases AU (ARE), dont la classe III représente des éléments uniquement riches en uridine (URE, voir Chapitre 4). Des

exemples de dégradation des transcrits par des URE avaient été mis en évidence chez *Trypanosoma*, mais jamais chez *Leishmania*. En analysant la séquence du 3'UTR de l'amastine, nous avons identifié un URE d'environ 100 nt.

Le **premier objectif** de nos travaux est ainsi de déterminer si le URE présent dans le 3'UTR du transcrit de l'amastine peut avoir un rôle régulateur. Dans un premier temps (Chapitre 8), nous avons décidé d'utiliser comme moyen expérimental le clonage du 3'UTR de ce transcrit, ainsi qu'une version tronquée sans URE, en aval du gène rapporteur luciférase. Cette approche est couramment utilisée et a fait ses preuves dans le cas de *Leishmania* et bien d'autres organismes. Des expériences de type Northern blot et des quantifications de l'activité luciférase permettent, respectivement, l'étude de la stabilité des transcrits et celle de leur taux de traduction.

Après avoir confirmé l'intérêt de cette séquence, le **second objectif** est d'identifier quels sont les facteurs protéiques impliqués dans la régulation de l'expression de l'amastine (Chapitre 9). La caractérisation de protéines liant des motifs des ARNm est une étape délicate et peu de complexes ARNm-protéines ont été bien définis à ce jour chez les *Trypanosomatidae*. Cet objectif est réalisable en utilisant des techniques de chromatographie d'affinité à l'ARN, la plus couramment utilisée étant le couplage d'un ARN avec de la biotine, l'incubation avec des extraits de protéines et la purification de complexes grâce à l'interaction entre la biotine et une matrice de streptavidine. Nous avons décidé d'utiliser une autre approche, basée sur l'ajout d'une étiquette ARN à l'extrémité 3' de la séquence d'intérêt. Les avantages de cette approche sont sa simplicité, son coût réduit comparativement aux techniques couramment utilisées, le fait qu'elle n'introduit pas d'acides nucléiques modifiés au sein de la séquence, optimisant ainsi les interactions, et assurant une homogénéité de la structure secondaire de l'ARN. Après identification d'un facteur putatif interagissant avec l'ARNm de l'amastine, la création de souches mutantes n'exprimant pas ce facteur permettra d'évaluer son rôle dans la régulation du transcrit de l'amastine et son importance pour la survie du parasite aux deux stades de vie.

Compte tenu de l'expression stade-spécifique de l'ARNm de l'amastine, la ou les protéines régulatrices interagissant avec le transcrit peuvent être régulées entre les deux stades de vie du parasite. Le **troisième et dernier objectif** de cette thèse est donc

d'examiner le mode de régulation de la protéine de liaison à l'ARN (Chapitre 10) comme, par exemple, quelle est son expression lors du cycle parasitaire, quelles sont les autres facteurs qu'elle lie et quelle est sa localisation sub-cellulaire dans les différents stades du parasite ? L'élaboration de protéines de fusion à des étiquettes de type HA ou encore des fusions avec des protéines fluorescentes de type GFP, faciliteront ces recherches.

## **Chapitre 8 : Dégradation stade spécifique des ARNm par un mécanisme indépendant de la déadénylation chez *Leishmania*.**

### **Résumé**

Le cycle de vie de *Leishmania* alterne entre promastigote (insecte vecteur) et amastigote (hôte mammifère). Presque tous les contrôles de l'expression génique y sont effectués de manière post-transcriptionnelle par des séquences de l'extrémité 3' non traduite (3'UTR) des ARNm. Nous étudions l'ARNm de l'amastine, spécifique du stade amastigote. Un élément riche en uridine (URE) de ~100 nt participe à la régulation du transcrite au cours du cycle parasitaire. La majeure partie de la dégradation des ARNm de l'amastine au stade promastigote n'est pas initiée par la déadénylation. Cet ARNm chez les amastigotes est par contre déadénylé avant d'être dégradé, tout comme la majorité des ARNm eucaryotes. Le retrait de la coiffe ne semble pas impliqué dans cette dégradation par l'URE chez les promastigotes. Ceci suggère que la dégradation de l'ARNm de l'amastine est biphasique, la première phase étant stade spécifique et dépendante d'une voie inhabituelle de dégradation par un URE.

## Article

### Deadenylation-independent stage-specific mRNA degradation in *Leishmania*

Simon Haile, Aurélien Dupé, and Barbara Papadopoulou\*

Publié dans: *Nucleic Acids Res.* 2008 March; 36(5): 1634–1644.

Research Centre in Infectious Diseases, CHUL Research Centre and Department of Medical Biology, Faculty of Medicine, Laval University, Quebec, Canada

\*To whom correspondence should be addressed. Phone: +1 418 654 2705, Fax: +1 418 654 2715, ; Email: barbara.papadopoulou@crchul.ulaval.ca

#### Abstract

The life cycle of *Leishmania* alternates between developmental forms residing within the insect vector (e.g. promastigotes) and the mammalian host (amastigotes). In *Leishmania* nearly all control of gene expression is post-transcriptional and involves sequences in the 3'-untranslated regions (3'UTRs) of mRNAs. Very little is known as to how these *cis*-elements regulate RNA turnover and translation rates in trypanosomatids and nothing is known about mRNA degradation mechanisms in *Leishmania* in particular. Here, we use the amastin mRNA—an amastigote-specific transcript—as a model and show that a ~100 nt U-rich element (URE) within its 3'UTR significantly accounts for developmental regulation. RNase-H-RNA blot analysis revealed that a major part of the rapid promastigote-specific degradation of the amastin mRNA is not initiated by deadenylation. This is in contrast to the amastin mRNA in amastigotes and to reporter RNAs lacking the URE, which, in common with most eukaryotic mRNAs studied to-date, are deadenylated before being degraded. Moreover, our analysis did not reveal a role for decapping in the stage-specific degradation of the amastin mRNA. Overall, these results suggest that degradation of the amastin mRNA of *Leishmania* is likely to be bi-phasic, the first phase

being stage-specific and dependent on an unusual URE-mediated pathway of mRNA degradation.

## **Introduction**

Parasitic species belonging to the kinetoplastid group of protozoans are of high medical and veterinary importance mainly in tropical and subtropical countries. These include trypanosomatids such as *Leishmania* spp. (Leishmaniasis), *Trypanosoma brucei* (Sleeping sickness) and *Trypanosoma cruzi* (Chagas disease). Their life cycle involves a mammalian host and an arthropod vector. The parasites face disparate conditions during their life cycle, a drastic change in temperature between their insect (~20–25°C) and mammalian (37°C) hosts being typically a shared challenge. Other changes can be unique to the life style of a given parasite (e.g. the acidic pH that *Leishmania* spp. face within the phagolysosomes of host macrophages). The parasites employ various adaptive strategies to be able to survive and grow in their distinct niches. These stage- or species-specific structural and metabolic adaptations are associated with significant changes in gene expression as evidenced by recent functional genomic analyses [Rochette *et al.*, unpublished data; (1–4)].

Kinetoplastids branched very early from the eukaryotic lineage and perhaps because of these they display several unusual or unique biological features. Tens to hundreds of genes are collinearly transcribed into a polycistronic mRNAs (5–7). Individual mRNAs are resolved from nuclear pre-mRNAs via coupled co-transcriptional processes of *trans*-splicing and 3'-end cleavage/polyadenylation (8). The process of *trans*-splicing involves the addition of ~40-bases long capped mini-exon to the 5'-termini of all mRNAs from a separate RNA substrate (*SL-RNA*) [reviewed in (8)]. Transcription has been postulated to initiate at strand switch regions on each chromosome probably within the intercistronic regions (9,10). There is, however, no indication for specific RNA polymerase II promoters associated with protein-coding genes to-date and the genomes of the parasites are short of genes encoding major transcription factors and co-activators. Consequently, most of gene regulation in these parasites occurs at the post-transcriptional levels, mainly via mRNA stability and translation [reviewed in (11,12)].

Eukaryotic mRNAs are typically capped at their 5'-ends and possess poly(A) tails at their 3'-ends. Such modifications at the mRNA termini, plus the proteins they bind to, confer mRNA stability and translatability. The degradation of eukaryotic mRNA is typically triggered by poly(A) tail shortening [reviewed in (13)]. Deadenylation, in turn, invokes decapping and the cumulative outcome is unprotected ends that are readily attacked by 5'- and 3'-exonucleases [reviewed in (13,14)]. Two 5'-exonucleases have been identified in yeast, the nuclear Rat1p/Xrn2p and the cytoplasmic Xrn1p (15,16). More than 10 conserved 3'-exonucleases have been characterized and some, together with other protein factors, are found in a complex called the exosome. Besides their role in determining mRNA half-lives in the cytoplasm, the 3'- and 5'-exonucleases are also involved in pre-mRNA quality control and the maturation of ribosomal, nuclear and small nucleolar RNAs [for examples see (17–19)]. In yeast, degradation from the 5'-end seems to be the major determinant of mRNA degradation and the 3'-pathway is only detectable in mutants of the 5'-pathway [reviewed in (13)]. In mammals, there is a growing consensus that both pathways are probably significant determinants of mRNA degradation (20–22).

The genomes of trypanosomes harbor genes that encode for deadenylases [reviewed in (11)]. A clear orthologue of a decapping enzyme has so far proven to be elusive even though decapping activity has been detected in the related organism *Leptomonas seymouri* (23). Degradation of mRNAs in *T. brucei* involves both 5'- and 3'-pathways (24) and homologues of all major 3'-exonucleases (25–27), and 5'-exonucleases (28) are well-characterized. Trypanosome 3'-exonucleases are also found in a complex and display all major RNA processing and degradation functions associated with their counterparts in other eukaryotes (25,26,29,30). Trypanosomes harbor four homologues of Xrn1p and Xrn2p/Rat1p (28). One of these, termed XRNA, is cytoplasmic and has been shown to be critical in the degradation of both stable and unstable mRNAs in *T. brucei* (28).

mRNA degradation *cis* determinants are predominantly found in the 3'-untranslated regions (UTRs) in all three species of trypanomatids. In *Leishmania*, in contrast to other trypanosomatids, several hundreds of the 3'UTR *cis*-elements belong to two major classes of short interspersed degenerate retroposons, SIDER1 and SIDER2 (31), and several lines of evidence indicate that these elements modulate post-transcriptional gene expression in

this parasite (31–33). These elements are thought to have evolved specifically in the *Leishmania* lineage perhaps under the selective pressure presented by the parasite's unique niches. Whether or not the mechanisms through which these elements function are also unique remains to be determined. The degradation of several unstable mRNAs in *T. brucei* [reviewed in (11) and *T. cruzi* (34)] is dependent on U-rich elements (UREs). These elements seem to be structurally (35) and functionally (24,29,36–40) similar to the mammalian AU-rich elements (AREs) found in several proto-oncogene and inflammatory cytokine mRNAs (41). These elements confer instability in a number of mRNAs in procyclic trypanosomes, developmental forms of the parasite that reside in the mid-gut of the insect vector. In *T. cruzi*, protein factors (UBPs) that bind these elements with high affinity and avidity have been identified (34,42,43). Homologues of these RNA-binding proteins also exist in *T. brucei* even though these proteins seem to have broad RNA-binding specificity (44). No UREs have been reported in *Leishmania* so far and virtually nothing is known about mechanisms underlying mRNA degradation in these parasites. In this study, we demonstrate URE-mediated degradation of the stage-specific amastin mRNA in *Leishmania*, thereby embarking on attempts to gain insight into the mechanistic basis of mRNA degradation in this organism.

## Materials and Methods

### Plasmids

The parental plasmid used in this study, pSPBT1-YNEO $\alpha$ LUC, was previously described (33). This plasmid allows the integration of reporter constructs into the *BTL* genomic locus. Sequence corresponding to the amastin (LinJ34\_V3.1030) 3'UTR and downstream intergenic region was amplified by PCR from *L. infantum* LEM 1317 genomic DNA (gDNA) using the forward primer, amastin-3'UTR-F (5'-AAGCTACTTCGGATCCGCGG-3') and the reverse primer, amastin-IR-R (5'-GGATCCGGCTCGCCAGTGTAGCAGAA-3'). The URE deletion construct was made by fusion PCR as described (45). Primers amastin-3'UTR-F and amastin-Up-URE-R (5'-CGTCCATGCGGCTCTTCTCTGCGCGTGCGTGTGTGTCG-3') were used to amplify the upstream region from the URE. Primers amastin-Down-URE-F (5'-CGACACACACGCACGCGCAGAGAAGAGCCGCATGGACG-3') and amastin-IR-R,

was used to amplify the downstream region of the URE. Primer amastin-Up-URE-R is the reverse complement of primer amastin-Down-URE-F. One hundred nanograms of the resulting PCR products, together with primers Amastin-3'UTR-F and amastin-IR-R, were used for a fusion PCR to generate amastin 3'UTR $\Delta$ URE. The only difference between full-length 3'UTR and 3'UTR- $\Delta$ URE is, therefore, the deletion of the 100-nt long URE. The resulting PCR fragments were digested with *Bam*HI (primers amastin-3'UTR-F and amastin-IR-R contain this site) and cloned into the *Bam*HI site of the pSPBT1-YNEO $\alpha$ LUC plasmid, downstream of the firefly luciferase (*LUC*) open reading frame (Figure 2).

### **Parasite culture and transfections**

*Leishmania donovani infantum* MHOM/MA/67/ITMAP-263, the parental strain for all the parasite lines employed in this study, was cultured in SDM-79 medium (pH 7), supplemented with 10% heat-inactivated fetal calf serum (Multicell, Wisent Inc.) and 5 mg/ml hemin at 25°C as promastigotes and were switched to MAA/20 medium (pH 5.8) at 37°C with 5% CO<sub>2</sub> and passed ~2–3 times to generate axenic amastigotes (46). For genomic integration into the *BTI* locus, ~2.5  $\mu$ g of *Hpa*I-*Hind*III digests (these enzymes only cut on either side of the *BTI* targeting regions) were transfected into promastigote *L. infantum* by electroporation as described (47). Transfected cells were plated on SDM-79 medium with 1.5% agar and 0.01 mg/ml of G418 (Sigma), and individual clones were obtained after 2–3 weeks.

### **RNA analysis**

Parasites were treated with 0.5  $\mu$ g/ml Actinomycin-D (Sigma) with/without 2.5  $\mu$ M Sinefungin (Sigma) for mRNA decay rate determination. All time points include 5 min of centrifugation times. Total RNA was extracted from cultures with OD<sub>600nm</sub> = 0.3–0.5 after lysis with Trizol (Invitrogen) according to the manufacturer's instructions. Ten to forty micrograms RNA was used for standard RNA blotting. Hybridization was done at 42°C with 50% formamide solution and RNA gels were transferred into Hybond-N membranes (Amersham Biosciences) by upward capillary movement. RNase-H (Invitrogen) digestions were carried out as described with 20–60  $\mu$ g RNA as a starting material (48) in the presence of a specific oligonucleotide 300R (5'-TTGTCTCCGTTCCCTCC GGGATCCCG-3') with or without oligo dT (Invitrogen) and the reaction contained RNase inhibitor (Invitrogen).

RNase H samples were run typically on 25 cm × 20 cm, 2–2.4% agarose gels. Polyacrylamide gels were 5% and transfer was carried out using BioRad's submarine system. Blots were probed with a fragment corresponding the last 300 bases of the amastin mRNA. Ten to twenty micrograms RNA was used as a starting material for the Terminator (Epicentre Biotechnologies) treatments of RNA samples. The reactions were carried out according to manufacturer's instructions at 30°C for 1–1.5 h and an RNase inhibitor (Invitrogen) was included in the reaction. Minus Terminator controls were treated equally except that the enzyme was not added. The reaction was stopped by adding 5 mM EDTA, and was directly analyzed by an RNA blot after denaturation in standard MOPS loading buffer. All quantifications of RNA blots were done using the ImageQuant 5.2 software.

### **Luciferase assay**

10<sup>7</sup> parasites were resuspended in 100 µl of lysis buffer (25 mM Tris–phosphate at pH 7.8, 2 mM dithiothreitol, 1% Triton X-100, 10% glycerol) and incubated at room temperature for 30 min, then transferred at –80°C for 2 h or overnight. After thawing, 20 µl of the parasite lysates was added to 96-well plates. One hundred microliters of luciferase assay buffer [20 mM tricine, 1.07 mM (MgCO<sub>3</sub>)<sub>4</sub>.Mg(OH)<sub>2</sub>.5 H<sub>2</sub>O, 2.67 mM MgSO<sub>4</sub>, 0.1 mM EDTA, 270 mM coenzyme A, 470 mM luciferin, 530 mM adenosine triphosphate, 33.3 mM DTT] was added to each well. Bioluminescence was measured using a Dynex MLX luminometer.

## **Results**

### **Quantitative analysis of the stage-specific mRNA regulation of the amastin mRNA**

This study attempted to begin to elucidate mechanisms underlying mRNA degradation in *Leishmania* using the amastin mRNA as a model. The amastin gene encodes a putative surface antigen and is a member of a multi-gene family (49,50). What we refer to as amastin in this paper is the original gene that was first identified-LinJ34\_V3.1030 (51). We previously demonstrated that most of the amastin mRNAs accumulate preferentially in amastigotes due to increased mRNA stability (49,51).

We first wished to obtain a quantitative data on the amastin mRNA regulation by using an additional probe that is specific to the mRNA's 3'UTR, the last 300 bases (probe 300) to exclude possible cross-reactivity with other amastins. Overall the amastin mRNA is ~3.9 more abundant in axenic amastigotes than in promastigotes-the developmental forms of the parasite that reside in the mammalian host and the insect vector, respectively (Figure 1A). The values obtained using the amastin coding sequence as a probe and the probe 300 are comparable and Figure 1A is representative of both. The decay kinetics of mRNAs was followed upon inhibition of *de novo* transcription using Actinomycin D and in some cases together with the arrest of splicing using Sinefungin. Sinefungin is assumed to prevent the splicing of pre-made pre-mRNAs by inhibiting the *de novo* methylation of the cap structure of the Spliced-Leader RNA (*SL-RNA*) and/or mRNA maturation from incomplete transcriptional inhibition. The combination of the two drugs was previously suggested to be important in drawing a more accurate representation of decay rates (11,30). Based on an analysis using the probe 300, the amastin mRNA decays with an overall half-life of ~40 min in promastigotes versus >120 min in amastigotes (Table 1, Figure 1B).

### **Stage-specific mRNA accumulation of the amastin mRNA is mediated through its 3'UTR**

We asked if the stage-specific accumulation of the amastin mRNA is mediated through sequences in the 3'UTR, as is the case for most trypanosomatid mRNAs studied to date. For this and subsequent analysis, we made a bicistronic Neomycin-Luciferase (NEO-LUC) construct which contained the 3'UTR and the downstream intergenic region of the endogenous amastin locus directly after the LUC-coding sequence (Figure 2A). The construct also included additional intergenic and artificial sequences that allowed the proper maturation of the reporter LUC and the selectable marker NEO mRNAs. The NEO-LUC cassette is flanked by sequences that allowed the targeting of the construct into the Biopterin Transporter 1 (*BT1*) genomic locus by homologous recombination (data not shown). The protein encoded by the *BT1* gene is not required under culture conditions in biopterin-rich media but it is essential otherwise (52). The *BT1* locus is, therefore, used routinely for the integration of transgenic constructs, particularly when a physiologically relevant expression level is desirable. All the other sequences except the 3'UTR have been used routinely in several published [e.g. (33)] and unpublished reporter gene analysis and

they do not affect RNA levels between promastigotes and amastigotes. Any difference in the LUC-amastin 3'UTR mRNA (*LUC-3'UTR*) levels should therefore be due to the amastin 3'UTR. Figure 2B shows that *LUC-3'UTR* mRNA is ~3.2-fold higher in amastigotes than in promastigotes, comparable with the regulation of the endogenous amastin mRNA (Figure 1A). mRNA decay analysis of the *LUC-3'UTR* showed that its half life is ~30 min in promastigotes and ~80 min in amastigotes (Table 1, Figure 3). In promastigotes, ~50–60% the *LUC-3'UTR* mRNA population are degraded with a half-life of ~10–15 min (Figure 3A and B). The rest of the population (40–50%) is degraded with a slower kinetics, the half-life being ~45 min (Figure 3A and B, data not shown). These results suggest that the stage specificity of the steady-state levels and degradation kinetics of the amastin mRNA are mediated mainly through sequences in its 3'UTR and that degradation is probably bi-phasic.

**A U-rich ~100 bases region in the 3'UTR of the amastin mRNA contributes significantly to its stage-specific accumulation and its facilitated degradation in promastigotes**

Next, we assessed whether the amastin mRNA regulation could be attributed to a distinct element in its 1.8 kb long 3'UTR. We have previously shown that the last 770 bases region of the amastin 3'UTR, in-and-by itself, does not account for stage-specific accumulation of the amastin mRNA (33). Sequence scanning of the rest of the 3'UTR revealed that it contains a ~100 bases long U-rich sequence (URE) (Figure 2A). Given the role of similar elements in the degradation of unstable mRNAs in other eukaryotes including *T. brucei* and *T. cruzi* [reviewed in (11,34)], we deleted this region in a construct which is otherwise identical to the *LUC-3'UTR* plasmid. The stable transfectant of this plasmid (*LUC-3'UTR*ΔURE) was analyzed in parallel with that of the *LUC-3'UTR*. Copy number differences were ruled out by Southern blot analysis and comparison of *NEO* mRNA levels (data not shown). As shown in Figure 2B, the *LUC-3'UTR*ΔURE mRNA steady-state levels are ~2.2-fold higher than that of *LUC-3'UTR* in promastigotes whereas no significant difference could be observed in amastigotes, suggesting that the URE accounts significantly for the stage-specific accumulation of the amastin mRNA. Decay analysis showed that the difference in the steady-state levels between *LUC-3'UTR* and *LUC-3'UTR*ΔURE mRNA levels observed in promastigotes is

accompanied by an overall increase in mRNA half-life from ~30 to >60 min (Figure 3A and B, Table 1). No significant changes could be detected in amastigotes (Figure 3C, Table 1). A more detailed comparison revealed that differences in degradation kinetics are more apparent in the initial phase (the first 15–30 min) than the later phase (Figure 3A and B). Luciferase activity levels generally followed the mRNA trends (Figure 2C, data not shown) and the URE-mediated mRNA degradation is, therefore, unlikely to be due to differential translatability of the mRNA. Taken together, these results suggested that a significant part of the amastin mRNA degradation in promastigotes is URE-mediated and that this probably occurs during the initiation phase of the degradation.

### **Differential degradation of the amastin mRNA in promastigotes seems to be initiated via deadenylation-independent mechanisms**

mRNA degradation in eukaryotes is typically initiated and modulated via differential deadenylation [reviewed in (13)]. Thus, we asked if the promastigote-specific degradation of the amastin mRNA is due to rapid deadenylation. The endogenous amastin or *LUC-3'UTR* mRNAs are relatively long (~2.6 and ~3.5 kb, respectively) and changes in mRNA sizes due to deadenylation in ordinary RNA blot analysis would be difficult to visualize. We, therefore, like in many other previous studies by others [e.g. (24)], exploited the DNA–RNA hybrid-dependent ribonuclease activity of RNase H to cleave these RNAs at specific sites and generate shorter 3'-end fragments. This was achieved by incubating RNA samples from several time points after ACT-D ± SIN treatments with RNase H in the presence of anti-sense oligonucleotide targeted to a region of ~300 bases from the end of the amastin 3'UTR (Figure 4A). Oligo dT was added in some samples so that the poly(A) tail could be cleaved off and the resulting fragment served as a marker for a ~100% deadenylated mRNA species. This analysis in promastigotes showed that the poly(A) tail of the *LUC-3'UTR* mRNA is still intact (relative to time 0) long after a significant part of the mRNA body has been degraded (Figure 4B, left panel). In contrast, a significant deadenylation clearly precedes the degradation of the mRNA in amastigotes (Figure 4B, right panel). Interestingly, the deadenylation pattern of the relatively more stable *LUC-3'UTRAURE* mRNA was not different between promastigotes (Figure 4B, middle panel) and amastigotes (data not shown) and it remarkably resembles that of *LUC-3'UTR* in amastigotes (Figure 4B, right panel). Noteworthy, a minor population of the *LUC-*

*3'UTR* mRNA in promastigotes appears to be deadenylated at later time points (the last lanes in Figure 4B, left panel and 4C) and is probably coincident with the second and slower phase of the mRNA degradation (Figure 3A and B).

We carried out a similar analysis for the endogenous amastin mRNA both in promastigotes and amastigotes. In order to detect the deadenylation pattern of the endogenous amastin mRNA in promastigotes (expression of the amastin mRNA is much higher in amastigotes compared to promastigotes), we had to increase the sensitivity of the assay and to expose the Northern blots much longer. This analysis shows that consistent with the reporter mRNA results described earlier (Figure 4B, left panel), the amastin mRNA in promastigotes—in contrast to the same mRNA in amastigotes—is degraded without detectable deadenylation (Figure 5A and C upper panel). We next wanted to assess the deadenylation profile of the amastin mRNA during differentiation. To induce differentiation *in vitro*, we subject amastigotes grown at 37°C and pH 5.8 to promastigote conditions (25°C and pH 7.0) of growth. Approximately 5 h following the switch from amastigote to promastigote conditions, a significant decrease in amastin mRNA levels could be observed (Figure 5B). Interestingly, under these conditions, the deadenylation pattern of the endogenous amastin mRNA (Figure 5C lower panel) was comparable to that seen using adapted promastigote cultures (Figure 5A). Overall, these results suggest that the initial phase of the degradation of the amastin mRNA in promastigotes *per se* is likely to be deadenylation-independent and that the stage-specific regulation of the amastin mRNA is unlikely to be a function of differential deadenylation rate.

### **Is decapping the rate-limiting step for the promastigote-specific degradation of the amastin mRNA?**

Deadenylation-independent mechanisms of mRNA degradation are rare in eukaryotes and are mainly initiated via an endonuclease or deadenylation-independent decapping activities. A role for a decapping mechanism is usually demonstrated via genetic or RNA interference (RNAi) approaches. However, no clear homologue of a decapping enzyme could so far be identified in the genomes of trypanosomatids. We, therefore, resorted in this study into an indirect assessment of the decapping rate of the *LUC-3'UTR*mRNA versus the *LUC-3'UTRAURE* mRNA in promastigotes. To do this, we took

advantage of the commercially available 5'-phosphate-dependent riboexonuclease, Terminator, which has so far been used mainly to enrich mRNA preparation as the enzyme degrades all uncapped RNAs (e.g. decapped mRNAs, rRNAs, tRNAs) while capped mRNAs are protected. A recent study has made use of the enzyme to qualitatively assess the cap status of a deadenylated intermediate of the *hsp70* mRNA (53). Here, we anticipated that relative mRNA decapping rates could quantitatively be inferred by comparison of Terminator treated versus untreated RNA samples derived from various time points after transcriptional arrest. For this assay, uncapped rRNAs that are cleaved by the enzyme served as positive controls (with 50–80% efficiency in our case). Histone 4A (*Hist4A*) mRNA is very stable within the time points used in the assay and therefore served as a negative control and indicated that the enzyme is highly specific (Figure 6A). Tube-to-tube variations in the efficiency of the Terminator activity were corrected with the levels of 18S RNA after loading normalization with the *Hist4A* mRNA (Figure 6A). Values from Terminator-treated samples were then divided by those from Terminator-untreated samples. The resulting numbers for each of the various time points after ACT-D treatment were expressed as a function of time 0. Such analysis indicated that there is no significant difference between the percentage of capped *LUC-3'UTR* and *LUC-3'UTRAURE* mRNAs at various time points after transcriptional arrest (relative to time 0) (Figure 6B). These results, within the limits of the assay employed, suggest that no detectable difference in decapping rate can account for the URE-mediated facilitated degradation of the amastin mRNA in promastigotes.

## Discussion

In this study, we began a series of experiments aimed at dissecting the mechanisms of mRNA degradation responsible for the expression of stage-specific mRNAs in *Leishmania*. Specifically, we studied the degradation of the amastin mRNA, which was previously shown to be more abundant in amastigotes than in promastigotes (51). Further interest in undertaking this study stems from the observations that the amastin gene encodes a putative surface protein and is a member of multi-gene family with an expected role in the biology and/or pathobiology of the parasite (49,50).

In common with a number of other previously studied mRNAs, the stage-specific amastin mRNA is regulated via sequences in its 3'UTR at the level of mRNA stability. Sequence analysis of its 3'UTR revealed the presence of a ~100 bases long U-rich sequence (URE) and our reporter mRNA analysis showed that this element contributes significantly to the stage-specific accumulation and degradation of the amastin mRNA. A similar role of UREs was previously shown for *T. brucei* and *T. cruzi* mRNAs [reviewed in (11,34)], suggesting that this mechanism is probably conserved in trypanosomatids. Noteworthy, the designation of this element as a URE, and its reference relative to elements in other eukaryotes, is solely based on the relatively high (50%) U-content of the region. Also, the specific regulatory region within the element, if distinct at all, might not necessarily be as U-rich.

The URE does not seem to account for all of the stage-specific regulation of the amastin mRNA. If any, the rest of the amastin mRNA regulation might be explained by other elements in its 1.8 kb 3'UTR. The 3'UTR of the amastin mRNA contains a sequence related to the *Leishmania* specific SIDER1 retroelements (31). We have previously shown that the amastin SIDER1 alone does not alter reporter RNA levels (33). However, the possibility that it might have a context dependent function, perhaps in synergy with the URE, remains to be tested.

mRNA decay analysis of the amastin mRNA revealed a remarkable resemblance of degradation kinetics with that of unstable mRNAs in *T. brucei*. Unlike the degradation of the stable mRNAs, at least the initial phase of the degradation of unstable URE-containing mRNAs appears to be deadenylation-independent (24,29). At least at first glance, an alternative interpretation of the deadenylation analysis we employed in this study can be associated with the possibility that deadenylated species of the endogenous amastin or *LUC-3'UTR* mRNAs in promastigotes might be degraded too fast to be detectable. Our data do not rule out the singular possibility that deadenylated amastin mRNA species might be degraded faster in promastigotes. However, whether the differential degradation of the amastin mRNA is mediated through and/or preceded by undetectable deadenylation is altogether a different question. Several indications seem to go against this possibility. First, there are no appreciable levels of deadenylated species at time 0 and at early time points

following actinomycin D treatment. Secondly, deadenylated species are actually detectable in the later time points (Figure 4B left panel and Figure 4C). Thirdly, as mRNAs get deadenylated, there must be a proportional decrease of the poly(A) plus mRNAs. This would imply that the remaining polyadenylated portion of the 3'-end fragments should disappear faster relative to the full-length mRNA. However, our results did not indicate that this is the case (Figures 3 and 4). In fact, the decay kinetics of the non-deadenylated 3'-end fragments was not distinguishable from those of the full-length, RNase-H-untreated mRNAs (Figure 4D). Fourthly, we did not observe detectable differences when we compared degradation from the 5'- and 3'-ends in RNase H-RNA blots using an oligonucleotide that allowed cutting at ~800 bases upstream of the 3'-end and subsequent detection with region-specific probes (see Supplementary Data in Figure S1). Indeed, RNAi-mediated depletion of a homologue of a yeast deadenylase in *T. brucei* has minimal effects on the degradation kinetics of unstable mRNAs (including URE-mRNAs) while that of moderately and highly stable mRNAs is clearly altered (personal communication by Dr Clayton, ZMBH, Germany). These observations collectively compelled us to conclude that the URE-mediated promastigote-specific degradation of the amastin mRNA is deadenylation-independent. It is, therefore, tempting to suggest that the likely possibility that deadenylated species might be degraded faster in promastigotes can instead be viewed as further evidence to our conclusion; namely that URE amastin mRNAs are degraded faster in promastigotes regardless of whether they contain poly(A) or not (i.e. degradation is independent of deadenylation status).

Significant levels of deadenylated species are detectable at steady-state in both stages of the parasite. This is a surprising observation because deadenylated mRNAs are considered to be unstable. Interestingly, a recent global analysis of poly(A) length of mammalian mRNAs has revealed a surprisingly high proportion of the mRNAs (~25%) having oligoadenylated species with only <30 length of poly(A) (54). Even a more surprising finding of the study by Meijer *et al.* is that some stable mRNAs fall among these mRNAs. For example, 25% of the beta-actin mRNA, an mRNA known to be very stable, contained a short poly(A) tail. These observations collectively challenge the current notion that deadenylated mRNA species are unstable. A related observation in this study is that a higher proportion of deadenylated and oligoadenylated species are detected for stable

mRNAs (the amastin or *LUC-3'UTR* mRNAs in amastigotes and the *LUC-3'UTRAURE* mRNA in both stages) than unstable mRNAs (the amastin and *LUC-3'UTR* mRNAs in promastigotes). This can be because the deadenylated species are degraded faster in promastigotes. Alternatively, this can be viewed as an indication for the possibility that the URE-mediated regulation of the amastin mRNA is not only deadenylation-independent but also polyadenylation-dependent. In other words, the poly(A) tail might have a destabilizing role rather than a stabilizing one. In bacteria, poly(A) tail serves as a 'holding toe' for 3'-exonucleases thereby facilitating degradation (55). However, this is not the case for almost all functional eukaryotic mRNAs studied so far. The few cases where poly(A) tail has a destabilizing function in eukaryotes fall within the realms of nuclear RNA degradation of certain pre-mRNAs and/or aberrant mRNAs as part of quality-control mechanisms (56–58).

Mechanisms of deadenylation-independent mRNA degradation, albeit rare, include deadenylation-independent decapping mechanisms such as the one associated with at least one pathway of the non-sense-mediated decay (59,60). Unfortunately, no clear homologue to any of the decapping enzymes from other eukaryotes has so far been identified in the genomes of trypanosomatids even though *in vitro* decapping activity was biochemically detected in a related species (23). In this study, we attempted to indirectly compare decapping rates by taking advantage of a 5'-phosphate-dependent enzyme, Terminator. This analysis did not show a significant difference between *LUC-3'UTR* and *LUC-3'UTRAURE* mRNAs in promastigotes suggesting the differential mRNA degradation is unlikely to be a function of an altered decapping rate. We, therefore, favor the scenario that an endonuclease-mediated cleavage might be the rate-limiting step, a well-studied example of which can be found in the degradation of the mammalian alpha-globin mRNA [reviewed in (61)]. Alternatively, a novel pathway might be awaiting a discovery.

Our current model of the amastin mRNA degradation based on this study, and in analogy with the situation in *T. brucei*, is that two mechanisms are likely to be involved: one that appears to be deadenylation- and decapping-independent, URE-mediated rapid degradation that is developmentally regulated, and the other is constitutive, which is likely to be slower and deadenylation-dependent (Figure 7). This study is the first not only in

reporting a URE-mediated degradation in *Leishmania* but also in attempting to draw the mechanistic basis of differential mRNA degradation in this parasite. The extent to which URE-mediated degradation differs from mammalian ARE-mediated mRNA instability and whether there are species-specific aspects of this mechanism among trypanosomatids remain to be determined.

## **Acknowledgments**

This work was supported by the Canadian Institutes of Health Research (CIHR) MOP-12182 operating grant to B.P. S.H. is supported by a postdoctoral fellowship from the CIHR. A.D. is a recipient of a fellowship from the CIHR STP-53924 Strategic Training Program. We thank all Papadopoulou lab members for discussions, and Annie Rochette for sharing some unpublished data. We are also indebted to Dr Christine Clayton (ZMBH, University of Heidelberg) for critical reading of the manuscript and for communicating unpublished data. Funding to pay the Open Access publication charges for this article was provided by the CIHR MOP-12182 grant.

## References

1. Cohen-Freue G, Holzer TR, Forney JD, McMaster WR. Global gene expression in *Leishmania*. *Int. J. Parasitol.* 2007; **37**:1077–1086.
2. Brems S, Guilbride DL, Gundlesdodjir-Planck D, Busold C, Luu VD, Schanne M, Hoheisel J, Clayton C. The transcriptomes of *Trypanosoma brucei* Lister 427 and TREU927 bloodstream and procyclic trypomastigotes. *Mol. Biochem. Parasitol.* 2005; **139**:163–172.
3. Minning TA, Bua J, Garcia GA, McGraw RA, Tarleton RL. Microarray profiling of gene expression during trypomastigote to amastigote transition in *Trypanosoma cruzi*. *Mol. Biochem. Parasitol.* 2003; **131**:55–64.
4. Atwood JA, 3rd, Weatherly DB, Minning TA, Bundy B, Cavola C, Opperdoes FR, Orlando R, Tarleton RL. The *Trypanosoma cruzi* proteome. *Science.* 2005; **309**:473–476.
5. Berriman M, Ghedin E, Hertz-Fowler C, Blandin G, Renauld H, Bartholomeu DC, Lennard NJ, Caler E, Hamlin NE, *et al.* The genome of the African trypanosome *Trypanosoma brucei*. *Science.* 2005; **309**:416–422.
6. El-Sayed NM, Myler PJ, Blandin G, Berriman M, Crabtree J, Aggarwal G, Caler E, Renauld H, Worthey EA, *et al.* Comparative genomics of trypanosomatid parasitic protozoa. *Science.* 2005; **309**:404–409.
7. Ivens AC, Peacock CS, Worthey EA, Murphy L, Aggarwal G, Berriman M, Sisk E, Rajandream MA, Adlem E, *et al.* The genome of the kinetoplastid parasite, *Leishmania major*. *Science.* 2005; **309**:436–442.
8. Liang XH, Haritan A, Uliel S, Michaeli S. *trans* and *cis* splicing in trypanosomatids: mechanism, factors, and regulation. *Eukaryot. Cell.* 2003; **2**:830–840.
9. Martinez-Calvillo S, Yan S, Nguyen D, Fox M, Stuart K, Myler PJ. Transcription of *Leishmania major* Friedlin chromosome 1 initiates in both directions within a single region. *Mol. Cell.* 2003; **11**:1291–1299.
10. Martinez-Calvillo S, Nguyen D, Stuart K, Myler PJ. Transcription initiation and termination on *Leishmania major* chromosome 3. *Eukaryot. Cell.* 2004; **3**:506–517.
11. Clayton C, Shapira M. Post-transcriptional regulation of gene expression in trypanosomes and leishmanias. *Mol. Biochem. Parasitol.* 2007; **156**:93–101.
12. Haile S, Papadopoulou B. Developmental regulation gene expression in trypanosomatid parasitic protozoa. *Curr. Opin. Microbiol.* 2007; **10**:569–577.
13. Parker R, Song H. The enzymes and control of eukaryotic mRNA turnover. *Nat. Struct. Mol. Biol.* 2004; **11**:121–127.
14. Mitchell P, Tollervey D. mRNA stability in eukaryotes. *Curr. Opin. Genet. Dev.* 2000; **10**:193–198.
15. Heyer WD, Johnson AW, Reinhart U, Kolodner RD. Regulation and intracellular localization of *Saccharomyces cerevisiae* strand exchange protein 1 (Sep1/Xrn1/Kem1), a multifunctional exonuclease. *Mol. Cell. Biol.* 1995; **15**:2728–2736.
16. Johnson AW. Rat1p and Xrn1p are functionally interchangeable exoribonucleases that are restricted to and required in the nucleus and cytoplasm, respectively. *Mol. Cell. Biol.* 1997; **17**:6122–6130.

17. Amberg DC, Goldstein AL, Cole CN. Isolation and characterization of RAT1: an essential gene of *Saccharomyces cerevisiae* required for the efficient nucleocytoplasmic trafficking of mRNA. *Genes Dev.* 1992; **6**:1173–1189.
18. Allmang C, Kufel J, Chanfreau G, Mitchell P, Petfalski E, Tollervey D. Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *Embo J.* 1999; **18**:5399–5410.
19. Bousquet-Antonelli C, Presutti C, Tollervey D. Identification of a regulated pathway for nuclear pre-mRNA turnover. *Cell.* 2000; **102**:765–775.
20. Chen CY, Gherzi R, Ong SE, Chan EL, Raijmakers R, Pruijn GJ, Stoecklin G, Moroni C, Mann M, *et al.* AU binding proteins recruit the exosome to degrade ARE-containing mRNAs. *Cell.* 2001; **107**:451–464.
21. Mukherjee D, Gao M, O'Connor JP, Raijmakers R, Pruijn G, Lutz CS, Wilusz J. The mammalian exosome mediates the efficient degradation of mRNAs that contain AU-rich elements. *Embo J.* 2002; **21**:165–174.
22. Stoecklin G, Mayo T, Anderson P. ARE-mRNA degradation requires the 5'-3' decay pathway. *EMBO Rep.* 2006; **7**:72–77.
23. Milone J, Wilusz J, Bellofatto V. Identification of mRNA decapping activities and an ARE-regulated 3' to 5' exonuclease activity in trypanosome extracts. *Nucleic Acids Res.* 2002; **30**:4040–4050.
24. Irmer H, Clayton C. Degradation of the unstable EP1 mRNA in *Trypanosoma brucei* involves initial destruction of the 3'-untranslated region. *Nucleic Acids Res.* 2001; **29**:4707–4715.
25. Estevez AM, Kempf T, Clayton C. The exosome of *Trypanosoma brucei*. *Embo J.* 2001; **20**:3831–3839.
26. Estevez AM, Lehner B, Sanderson CM, Ruppert T, Clayton C. The roles of intersubunit interactions in exosome stability. *J. Biol. Chem.* 2003; **278**:34943–34951.
27. Haile S, Cristodero M, Clayton C, Estevez AM. The subcellular localisation of trypanosome RRP6 and its association with the exosome. *Mol. Biochem. Parasitol.* 2007; **151**:52–58.
28. Li CH, Irmer H, Gudjonsdottir-Planck D, Freese S, Salm H, Haile S, Estevez AM, Clayton C. Roles of a *Trypanosoma brucei* 5'->3' exoribonuclease homolog in mRNA degradation. *RNA.* 2006; **12**:2171–2186.
29. Haile S, Estevez AM, Clayton C. A role for the exosome in the *in vivo* degradation of unstable mRNAs. *RNA.* 2003; **9**:1491–1501.
30. Colasante C, Robles A, Li CH, Schwede A, Benz C, Voncken F, Guilbride DL, Clayton C. Regulated expression of glycosomal phosphoglycerate kinase in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* 2007; **151**:193–204.
31. Bringaud F, Muller M, Cerqueira GC, Smith M, Rochette A, El-Sayed NM, Papadopoulou B, Ghedin E. Members of a large retroposon family are determinants of post-transcriptional gene expression in *Leishmania*. *PLoS Pathog.* 2007; **3**(9):1291–1307.
32. Boucher N, Wu Y, Dumas C, Dubé M, Sereno D, Breton M, Papadopoulou B. A common mechanism of stage-regulated gene expression in *Leishmania* mediated by a conserved 3'-untranslated region element. *J. Biol. Chem.* 2002; **277**:19511–19520.
33. McNicoll F, Muller M, Cloutier S, Boilard N, Rochette A, Dube M, Papadopoulou B. Distinct 3'-untranslated region elements regulate stage-specific mRNA accumulation and translation in *Leishmania*. *J. Biol. Chem.* 2005; **280**:35238–35246.

34. D'Orso I, De Gaudenzi JG, Frasch AC. RNA-binding proteins and mRNA turnover in trypanosomes. *Trends Parasitol.* 2003; **19**:151–155.
35. Drozd M, Clayton C. Structure of a regulatory 3' untranslated region from *Trypanosoma brucei*. *RNA.* 1999; **5**:1632–1644.
36. Hotz HR, Hartmann C, Huober K, Hug M, Clayton C. Mechanisms of developmental regulation in *Trypanosoma brucei*: a polypyrimidine tract in the 3'-untranslated region of a surface protein mRNA affects RNA abundance and translation. *Nucleic Acids Res.* 1997; **25**:3017–3026.
37. Di Noia JM, D'Orso I, Sanchez DO, Frasch AC. 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.* 2000; **275**:10218–10227.
38. D'Orso I, Frasch AC. 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.* 2001; **276**:15783–15793.
39. Quijada L, Guerra-Giraldez C, Drozd M, Hartmann C, Irmer H, Ben-Dov C, Cristodero M, Ding M, Clayton C. Expression of the human RNA-binding protein HuR in *Trypanosoma brucei* increases the abundance of mRNAs containing AU-rich regulatory elements. *Nucleic Acids Res.* 2002; **30**:4414–4424.
40. Mayho M, Fenn K, Craddy P, Crosthwaite S, Matthews K. Post-transcriptional control of nuclear-encoded cytochrome oxidase subunits in *Trypanosoma brucei*: evidence for genome-wide conservation of life-cycle stage-specific regulatory elements. *Nucleic Acids Res.* 2006; **34**:5312–5324.
41. Chen CY, Xu N, Shyu AB. mRNA decay mediated by two distinct AU-rich elements from *c-fos* and granulocyte-macrophage colony-stimulating factor transcripts: different deadenylation kinetics and uncoupling from translation. *Mol. Cell. Biol.* 1995; **15**:5777–5788.
42. D'Orso I, Frasch AC. TcUBP-1, a developmentally regulated U-rich RNA-binding protein involved in selective mRNA destabilization in trypanosomes. *J. Biol. Chem.* 2001; **276**:34801–34809.
43. D'Orso I, Frasch AC. 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.* 2002; **277**:50520–50528.
44. Hartmann C, Benz C, Brems S, Ellis L, Luu VD, Stewart M, D'Orso I, Busold C, Fellenberg K, *et al.* The small trypanosome RNA-binding proteins TbUBP1 and TbUBP2 influence expression of F-box protein mRNAs in bloodstream trypanosomes. *Eukaryot. Cell.* 2007; **6**(11):1964–1978.
45. Derbise A, Lesic B, Dacheux D, Ghigo JM, Carniel E. A rapid and simple method for inactivating chromosomal genes in *Yersinia*. *FEMS Immunol. Med. Microbiol.* 2003; **38**:113–116.
46. Sereno D, Lemesre JL. Axenically cultured amastigote forms as an *in vitro* model for investigation of antileishmanial agents. *Antimicrob. Agents Chemother.* 1997; **41**:972–976.
47. Papadopoulou B, Roy G, Ouellette M. A novel antifolate resistance gene on the amplified H circle of *Leishmania*. *Embo J.* 1992; **11**:3601–3608.

48. Salles FJ, Richards WG, Strickland S. Assaying the polyadenylation state of mRNAs. *Methods*. 1999; **17**:38–45.
49. Rochette A, McNicoll F, Girard J, Breton M, Leblanc E, Bergeron MG, Papadopoulou B. Characterization and developmental gene regulation of a large gene family encoding amastin surface proteins in *Leishmania* spp. *Mol. Biochem. Parasitol.* 2005; **140**:205–220.
50. Rafati S, Hassani N, Taslimi Y, Movassagh H, Rochette A, Papadopoulou B. Amastin peptide-binding antibodies as biomarkers of active human visceral leishmaniasis. *Clin. Vaccine Immunol.* 2006; **13**:1104–1110.
51. Wu Y, El Fakhry Y, Sereno D, Tamar S, Papadopoulou B. A new developmentally regulated gene family in *Leishmania* amastigotes encoding a homolog of amastin surface proteins. *Mol. Biochem. Parasitol.* 2000; **110**:345–357.
52. Kundig C, Haimeur A, Legare D, Papadopoulou B, Ouellette M. Increased transport of pteridines compensates for mutations in the high affinity folate transporter and contributes to methotrexate resistance in the protozoan parasite *Leishmania tarentolae*. *Embo J.* 1999; **18**:2342–2351.
53. Bonisch C, Temme C, Moritz B, Wahle E. Degradation of hsp70 and other mRNAs in *Drosophila* via the 5' 3' pathway and its regulation by heat shock. *J. Biol. Chem.* 2007; **282**:21818–21828.
54. Meijer HA, Bushell M, Hill K, Gant TW, Willis AE, Jones P, de Moor CH. A novel method for poly(A) fractionation reveals a large population of mRNAs with a short poly(A) tail in mammalian cells. *Nucleic Acids Res.* 2007; **35**:e132.
55. Kushner SR. mRNA decay in *Escherichia coli* comes of age. *J. Bacteriol.* 2002; **184**:4658–4665. discussion 4657.
56. LaCava J, Houseley J, Saveanu C, Petfalski E, Thompson E, Jacquier A, Tollervey D. RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. *Cell.* 2005; **121**:713–724.
57. Wyers F, Rougemaille M, Badis G, Rousselle JC, Dufour ME, Boulay J, Regnault B, Devaux F, Namane A, *et al.* Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. *Cell.* 2005; **121**:725–737.
58. West S, Gromak N, Norbury CJ, Proudfoot NJ. Adenylation and exosome-mediated degradation of cotranscriptionally cleaved pre-messenger RNA in human cells. *Mol. Cell.* 2006; **21**:437–443.
59. Hagan KW, Ruiz-Echevarria MJ, Quan Y, Peltz SW. Characterization of *cis*-acting sequences and decay intermediates involved in nonsense-mediated mRNA turnover. *Mol. Cell. Biol.* 1995; **15**:809–823.
60. Muhlrاد D, Parker R. Premature translational termination triggers mRNA decapping. *Nature.* 1994; **370**:578–581.
61. Liu H, Kiledjian M. An erythroid-enriched endoribonuclease (ErEN) involved in alpha-globin mRNA turnover. *Protein Pept. Lett.* 2007; **14**:131–136.

## Figure Legends

### Figure 1. Endogenous amastin mRNA steady-state levels and decay kinetics.

(A) RNA blot analysis of the amastin mRNA levels in promastigotes (Pro) and amastigotes (Ama). 18S was used as a loading control. Graphic representation of the data is shown in the right panel and values are relative to promastigotes. Blots were probed with one fragment which is complementary specifically to the last 300 bases of the amastin 3'UTR (probe 300, an average of three experiments), and another probe corresponding to the amastin-coding region (five experiments). (B) Decay kinetics of the amastin mRNA in promastigotes and amastigotes. 18S was used as a loading control. Blots were hybridized with probe 300. Note that the promastigote blot is over-exposed and comparison of steady-state levels with the amastigote blot is not advisable in this case. mRNA levels from different time points after transcriptional arrest with actinomycin-D (ACT-D) and inhibition of splicing with Sinefungin (SIN) were expressed relative to time 0 and data are representative of two experiments. '-15' denotes pre-treatment with SIN (including centrifugation time) before ACT-D is added at time 0. Graphic representations of the experiments are shown in the bottom panel. The absolute mRNA half-life values ( $t_{1/2}$ ) in promastigotes and amastigotes are indicated. For mRNA half-life values derived from ACT-D treatment alone see Table 1.

### Figure 2. The use of reporter mRNAs to demonstrate the role of specific 3'UTR sequences in the stage-specific accumulation of the amastin mRNA.

(A) Schematic representation of the reporter constructs, LUC-3'UTR and LUC-3'UTR $\Delta$ URE. The exact sequence of the U-rich element (URE) that is deleted in LUC-3'UTR $\Delta$ URE is indicated. The NEO-LUC bicistron is flanked with sequences that allow construct targeting to the endogenous *BTI* locus via homologous recombination. Y is a polypyrimidine stretch that allows 5'-end splicing of the Neomycin phosphotransferase (*NEO*) mRNA.  $\alpha$ Tub IR stands for the intergenic region of the alpha-tubulin gene and allows the 3'-end maturation of the *NEO* mRNA and the 5'-maturation of the Luciferase (*LUC*) mRNA. Ama IR denotes the intergenic region of the amastin gene, which allows the 3'-end maturation of the *LUC* mRNA. (B) RNA blot analysis of LUC-3'UTR and LUC-3'UTR $\Delta$ URE mRNA levels in promastigotes (Pro) and amastigotes (Ama). 18S was used as

a loading control. Graphic representation of the steady-state reporter mRNA levels is indicated in the right panel. Values are expressed relative to promastigote *LUC-3'UTR* mRNA levels. Data are derived from two to three experiments but 11 experiments for the *LUC-3'UTR* versus *LUC-3'UTRAURE* comparisons in promastigotes. The *P*-value for the *LUC-3'UTR* versus *LUC-3'UTRAURE* in promastigotes is shown and was calculated using Excel's Type 3 two-tailed *T*-test. (C) Luciferase activity in promastigotes. Luciferase (LUC) protein levels were inferred by measuring LUC enzymatic activity. Data are derived from three experiments.

**Figure 3. Decay kinetics of reporter mRNAs. mRNA decay rate was followed after inhibition of transcription alone with Actinomycin D (ACT-D) for *LUC-3'UTR* versus *LUC-3'UTRAURE* in promastigotes (Pro)**

(A), ACT-D plus Sinefungin (SIN) for *LUC-3'UTR* versus *LUC-3'UTRAURE* in Promastigotes (B) and ACT-D alone for *LUC-3'UTR* versus *LUC-3'UTRAURE* in amastigotes (Ama) (C). '-15' time point in (B) represents a pre-treatment with SIN (including centrifugation time) before the addition of ACT-D at time 0. Histone 4A (Hist4A) was used as a loading control. Graphic representations of mRNA decay rates are indicated on the right panels where mRNA levels from different time points (after ACT-D with or without SIN treatments) were expressed relative to time 0. Data are derived from two experiments for the amastigote data and four experiments for the promastigote data. The absolute mRNA half-life values ( $t_{1/2}$ ) are indicated. See Table 1 for the entire complement of mRNA half-lives calculated in this study and their corresponding standard errors.

**Figure 4. Determination of poly(A) tail status during reporter mRNA decay.**

(A) Schematic representation of the deadenylation assay. RNase-H digests DNA/RNA hybrid. In this case the DNA is an oligonucleotide (300R), which is reverse complementary to a region of ~300 bases upstream of the poly(A) tail of the reporter mRNA. The presence of oligo dT in the reaction allows the trimming of the poly(A) tails (see later). The vertical darker arrows indicate the sequence to which the oligonucleotide was targeted into. The last 300 bases of the 3'UTR is underlined and is where the probe for the RNA blots was targeted. (B) Deadenylation of reporter mRNAs in promastigotes (Pro) and amastigotes

(Ama). RNA samples were collected from different time points after treatment with ACT-D and subjected to RNase-H treatment. 3' fragments were analysed by RNA blot. At time 0, RNA samples were also treated with RNase-H digestion in the presence of oligo dT (in addition to the specific oligonucleotide) to generate poly(A) minus 3'-end markers ('0 + dT' lanes). Hist 4A is used as a loading control. (C) As in (B) but RNA was derived from ACT-D and SIN treatment of *LUC-3'UTR* promastigotes. Data are representative of three experiments. (D) Quantitative representation of the decay rate of *LUC-3'UTR* polyadenylated 3'-end fragments in promastigotes (B, left panel) is compared to that of the full-length *LUC-3'UTR* mRNA in promastigotes (Figure 3A). The lateral bracket indicated in Figure 4B left panel shows the portion of the gel used for quantification of the largely poly(A)+ 3'-end fragments. Note that the full-length *LUC-3'UTR* mRNA used for this quantification was probed by the luciferase coding sequence, which is ~1.8 kb upstream of the polyadenylation site while the 3'-end products of the RNase H were probed by the last 300 terminal sequence of the 3'-end.

**Figure 5. Deadenylation analysis of the endogenous amastin mRNA.**

(A) RNase H-RNA blot deadenylation assay for assessing deadenylation of the endogenous amastin mRNA in promastigotes. RNA samples were collected at 0, 15 and 30 min (half-life of the endogenous mRNA in promastigotes, see Figure 1B) following transcriptional arrest with Actinomycin-D (ACT-D) (left panel) and treatment with ACT-D plus Sinefungin (SIN) to inhibit splicing (right panel) and subjected to RNase-H treatment as detailed in Figure 4 and under 'Materials and Methods' section. (B) Time-point analysis of the amastin mRNA during a switch from axenic amastigote (Ama) (pH 5.8, 37°C) to promastigote (Pro) (pH 7.0, 25°C) conditions. Blots were probed with the probe 300 fragment. (C) RNase H-RNA blot deadenylation assay. The 5 h post-switch was selected based on the data presented in (B), a time point where a detectable decrease in mRNA level was observed and yet not down to an extent below the detection limit of the deadenylation assay. RNA samples were run on 5% polyacrylamide gels. Other methodological details for the deadenylation analysis are as in Figure 4. The data are representative of two experiments.

**Figure 6. Indirect measurement of decapping rates.**

(A) RNA samples were derived from LUC-3'UTR and LUC-3'UTR $\Delta$ URE lines treated with actinomycin-D (ACT-D) for different time points. An equal volume of RNA was separated into two different tubes. One set of the tubes contained the reaction mix containing the Terminator enzyme (+ lanes) and the other contained the same mix but without the enzyme (– lanes). Both Terminator – and + reactions were incubated for 1–1.5 h according to manufacturer's instructions (Epicentre Biotechnologies). Data are representative of two experiments. Similar results were obtained in two other experiments involving ACT-D and SIN (data not shown). 18S RNA signals were used to control for tube-to-tube variation of Terminator degradation efficiency after normalization of loading with histone 4A (*Hist4A*) mRNA. (B) Graphic representation of the ratio of + Terminator signals to – Terminator signals relative to time 0.

**Figure 7. A proposed model for the stage-specific mRNA degradation of the amastin mRNA.**

A significant part of the promastigote-specific facilitated degradation of the amastin mRNA (and the *LUC-3'UTR* mRNA) is URE-dependent, is mostly deadenylation-independent and probably decapping-independent (A). An additional unknown mechanism that is URE-independent (B) might be involved as well because deletion of the URE did not account for all of the stage-specific mRNA regulation at the steady-state levels as indicated in Figure 2. A constitutive deadenylation-dependent pathway (C) is responsible for the decay of more stable mRNAs (*LUC-3'UTR $\Delta$ URE* in both developmental forms, and the amastin and *LUC-3'UTR* mRNAs in amastigotes). A minor sub-population of the amastin and *LUC-3'UTR* mRNAs in promastigotes is probably degraded via a deadenylation-dependent constitutive pathway.

## Tables

**Table 1. mRNA half-lives for the endogenous amastin and reporter mRNAs analysed in this study**

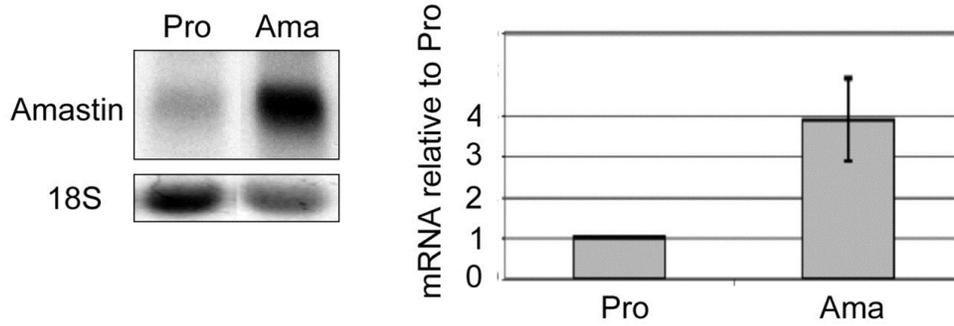
mRNA	mRNA half-lives (min)			
	ACT-D		ACT-D + SIN	
	Pro	Ama	Pro	Ama
Amastin	45 ± 20	>120	45 ± 15	>180
LUC-3'UTR	31.3 ± 19.6	82.5 ± 10.6	28 ± 15.5 <sup>a</sup>	115 ± 10
LUC-3'UTRΔURE	>120	76 ± 1.4	63 ± 21.3 <sup>a</sup>	137 ± 5

Decay rates for the endogenous amastin, *LUC-3'UTR* and *LUC-3'UTRΔURE* mRNAs in promastigotes (Pro) and axenic amastigotes (Ama) were determined after transcriptional inhibition using Actinomycin D (ACT-D) with or without splicing arrest using Sinefungin (SIN).

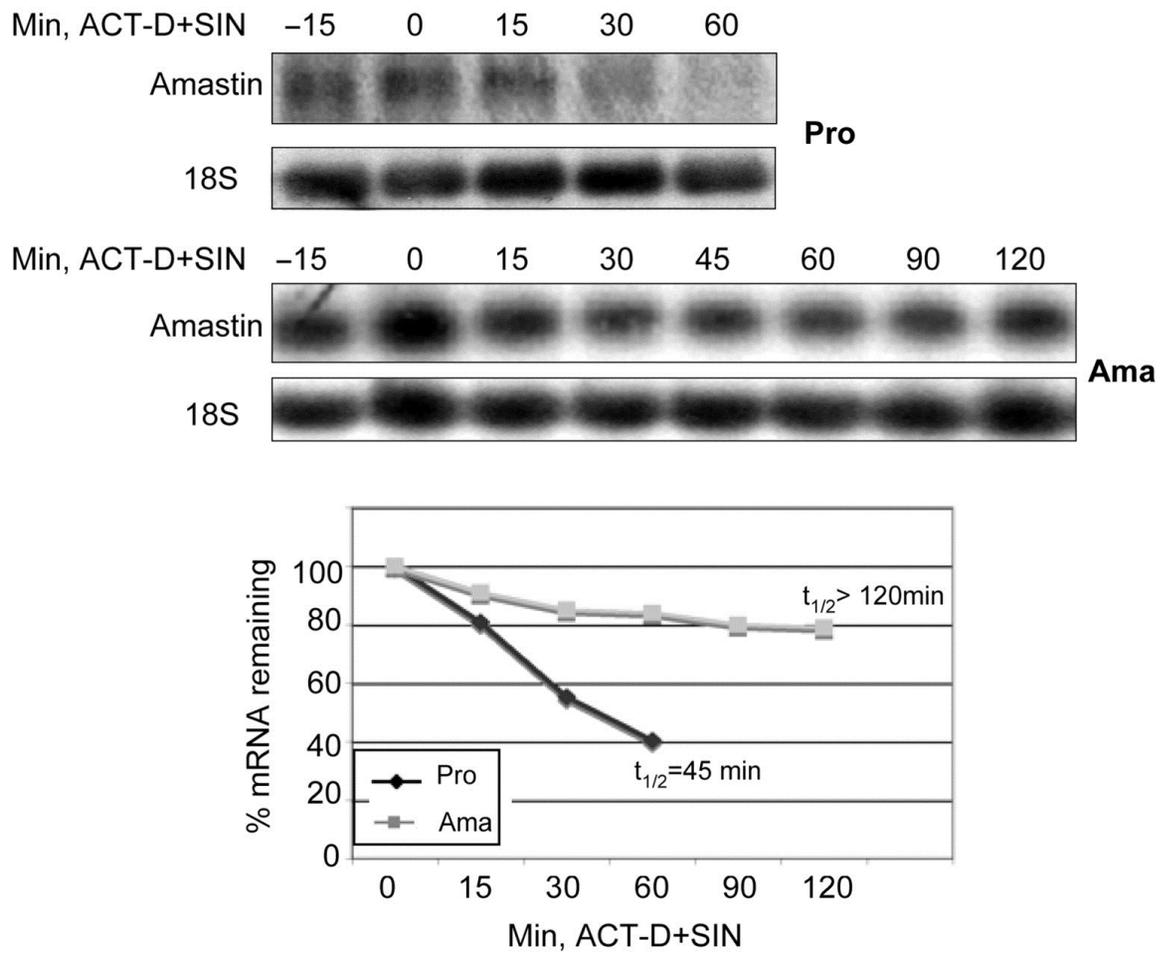
<sup>a</sup>Despite the variation of the absolute half-life values which are directly reflected in this Table, *LUC-3'UTRΔURE* mRNA was consistently more stable than *LUC-3'UTR* mRNA in promastigotes in each of the four independent experiments we carried out whereas no significant differences were observed between these two RNAs in amastigotes. Numerically, this parameter is reflected by the average of the relative differences (the half-life of *LUC-3'UTRΔURE* mRNA minus that of the *LUC-3'UTR* mRNA), which is 34 min ± an SD of 8 min.

# Figures

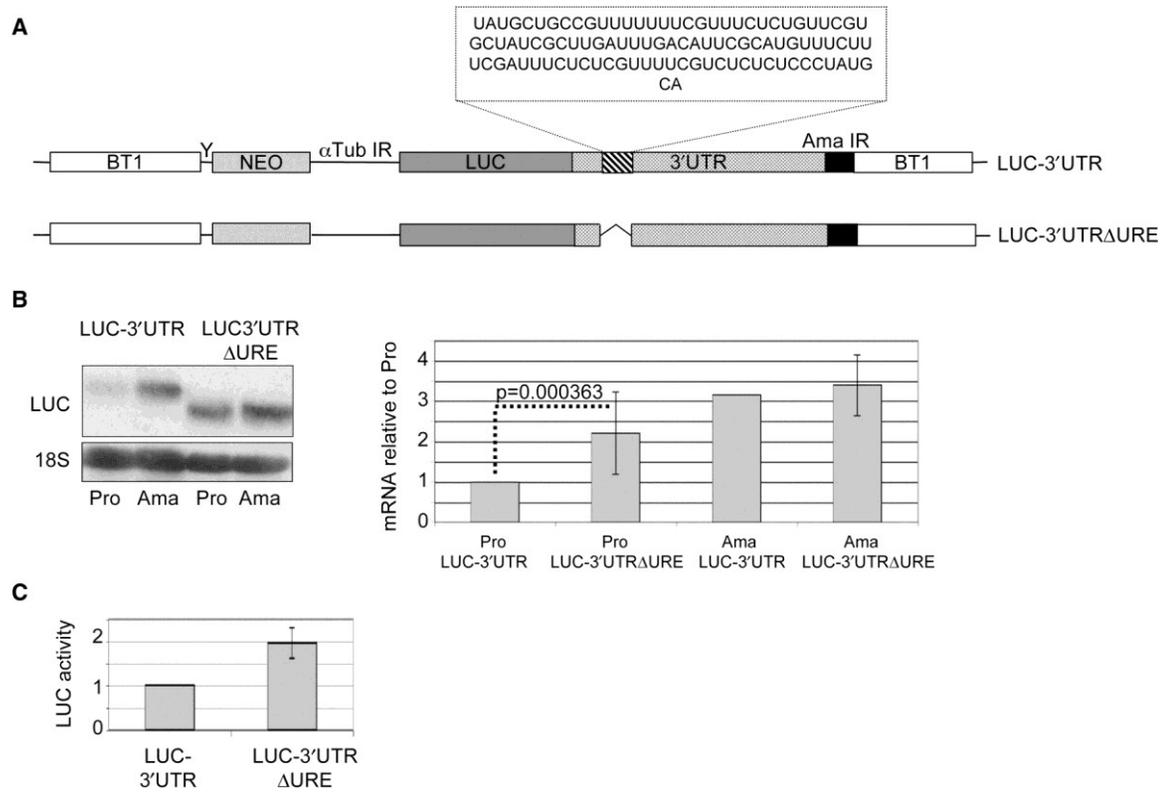
**A**



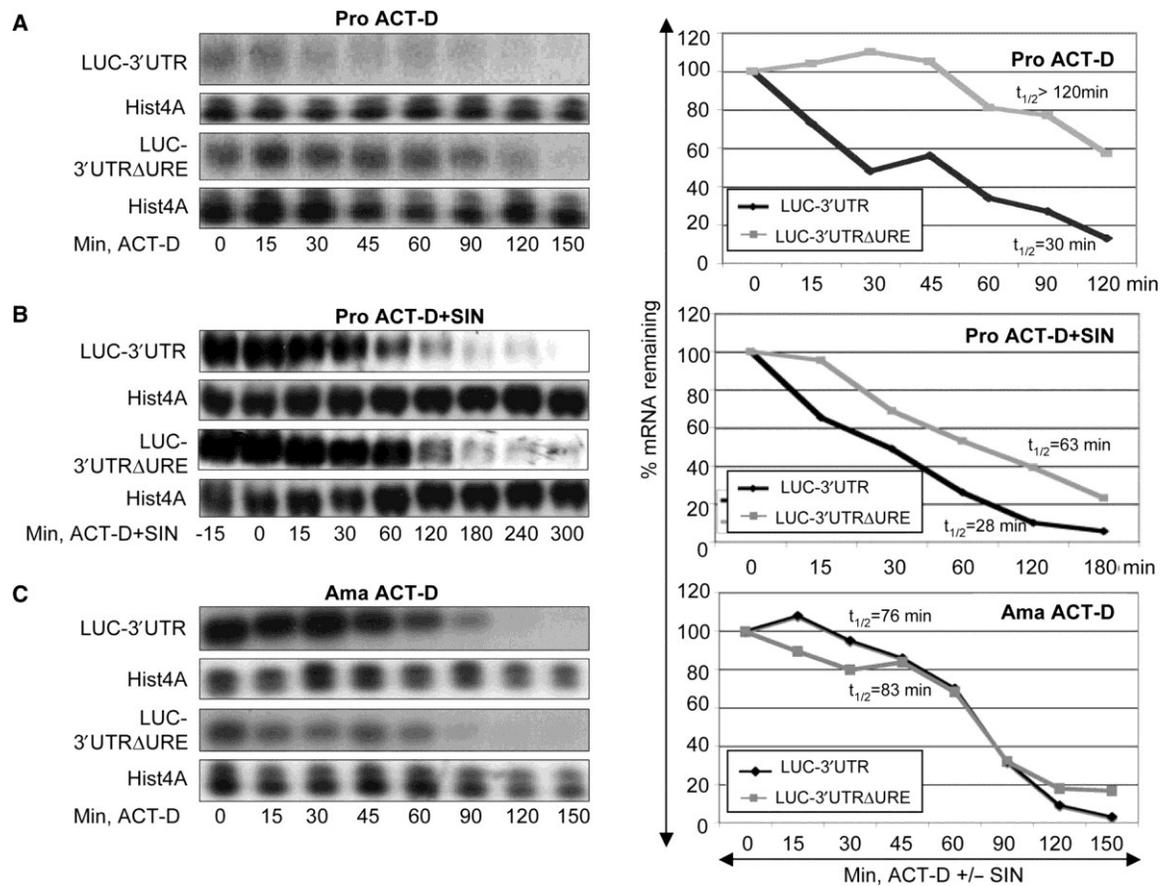
**B**



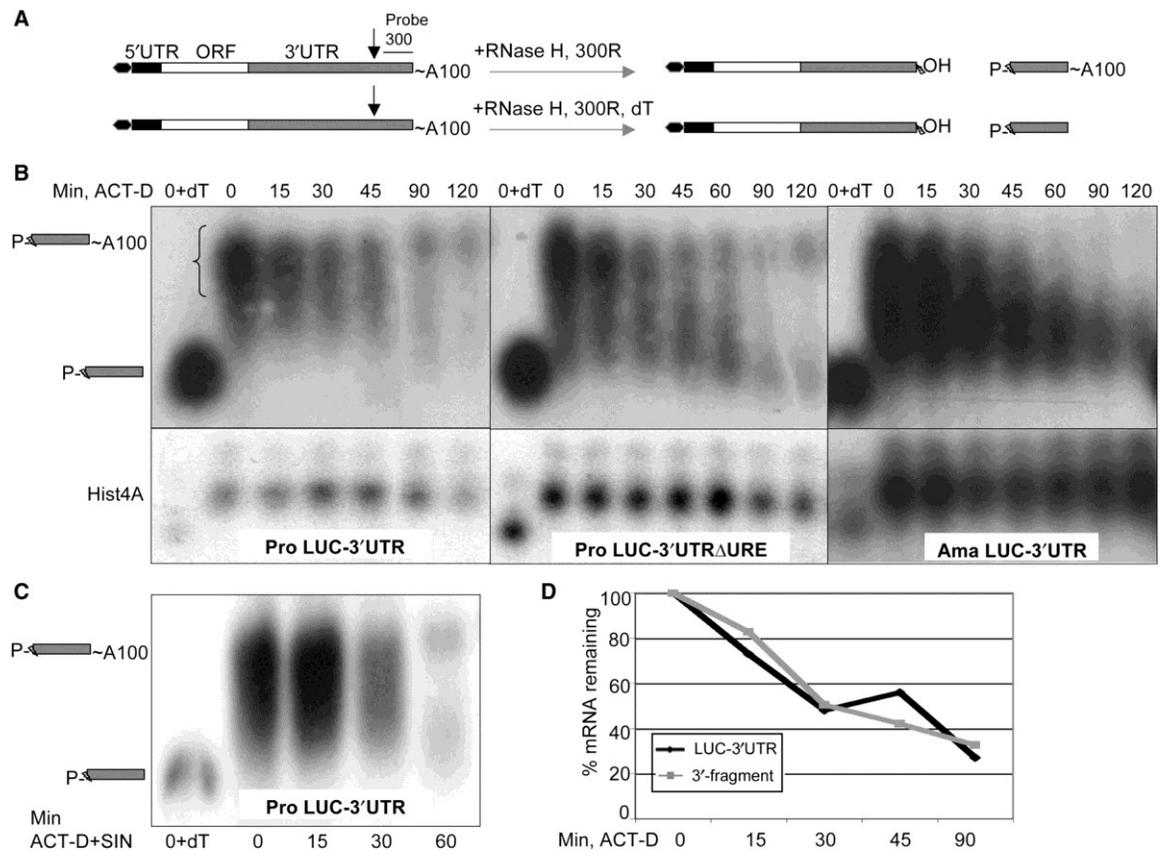
**Figure 1.**



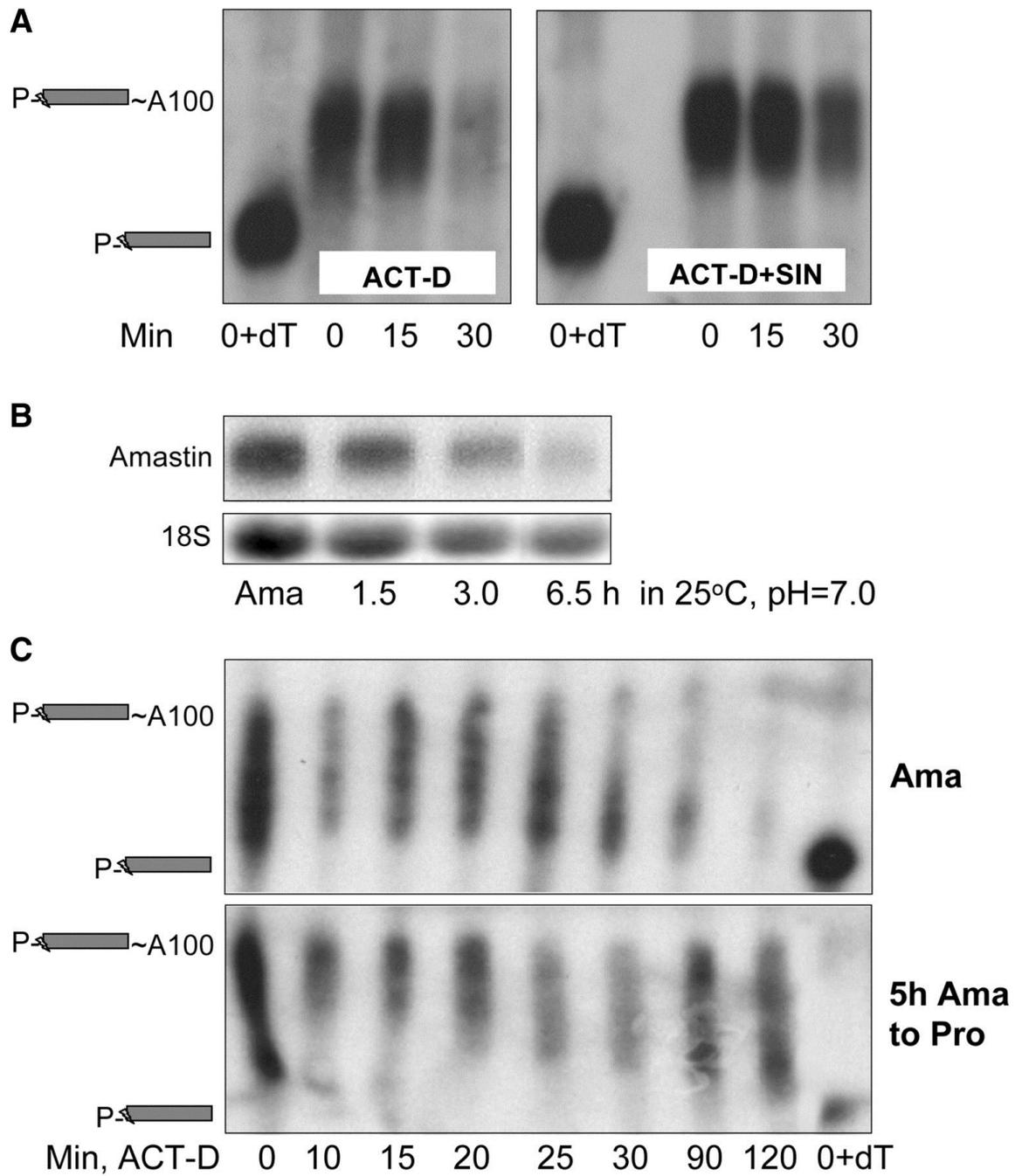
**Figure 2.**



**Figure 3.**



**Figure 4.**



**Figure 5.**

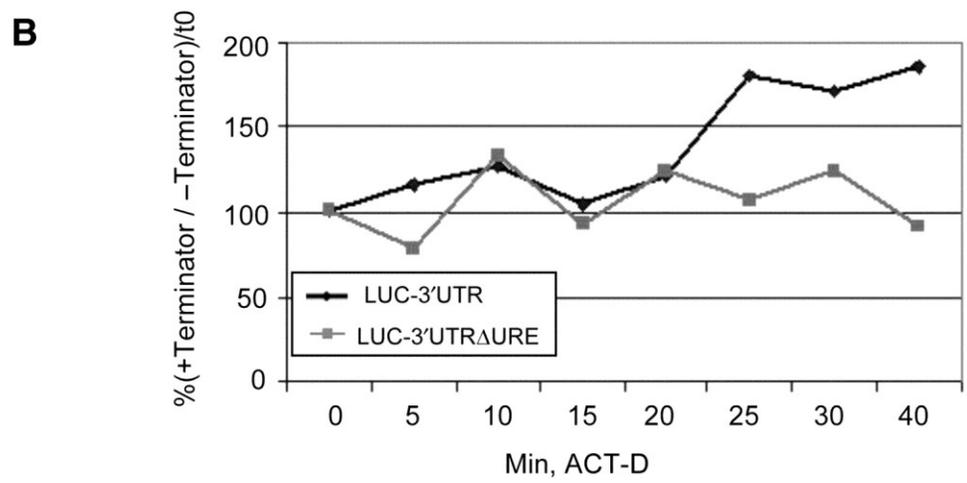
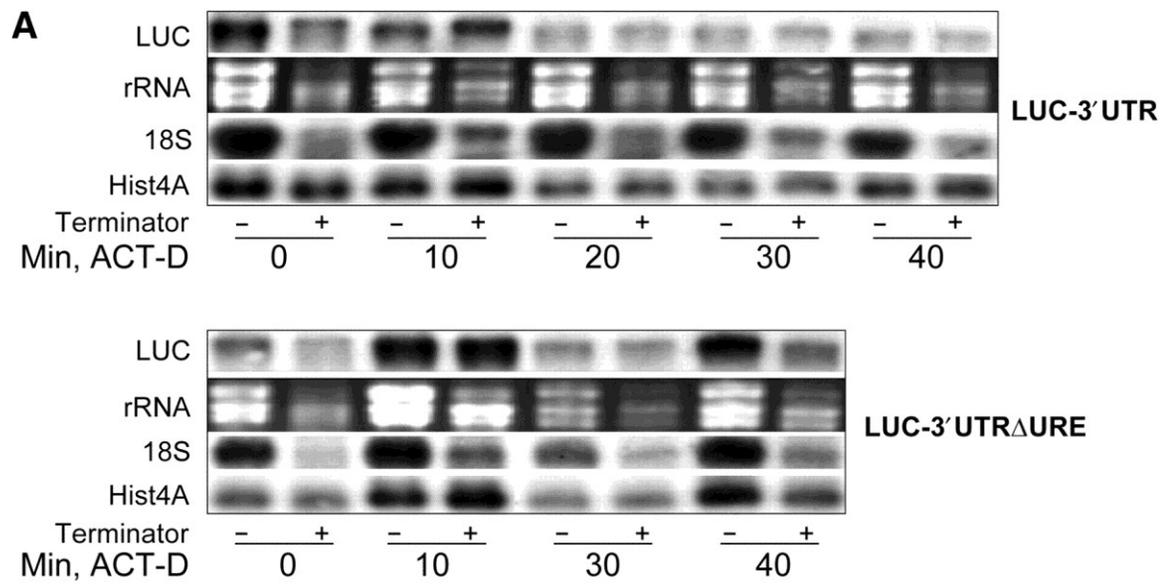
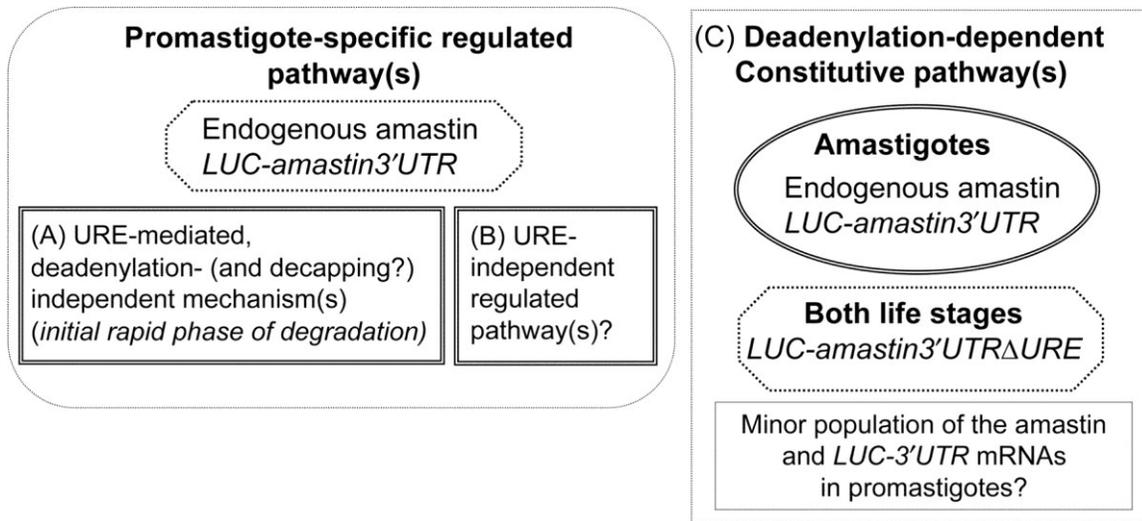


Figure 6.



**Figure 7.**

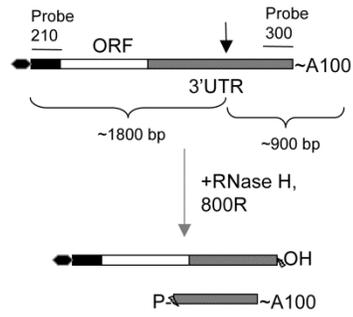
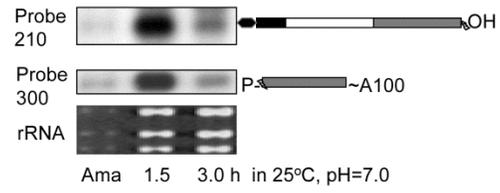
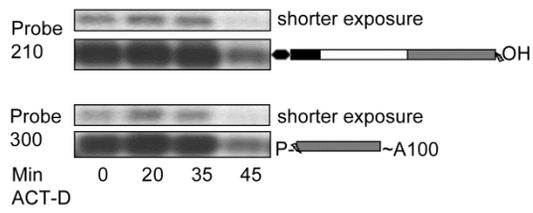
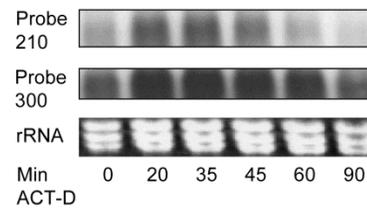
## Supplementary Figures

### Figure S1.

Decay kinetics of 5' and 3' ends of the endogenous amastin mRNA. (A) Diagram describing the RNase H digestion. RNase H digests DNA/RNA hybrid. In this case the DNA is an oligonucleotide (800R), which is reverse complementary to the 5'-end region of ~800 bases upstream of the poly(A) tail of the amastin mRNA. The vertical darker arrows indicate the sequence to which the oligonucleotide was targeted into. Probes used for region specific detection are indicated; probe 210 corresponds to the first 210 bases of the mRNA which is the 5'UTR of the amastin mRNA and probe 300 corresponds to the last 300 bases of the amastin 3'UTR. (B) Region specific analysis of 5'- and 3'-fragments generated by RNase H treatment of samples from amastigotes that have been cultured in promastigote (pH 7.0, 25° C) conditions for 0, 1.5 and 3 hours, respectively. Blots were probed with the probe 210 and probe 300, respectively. rRNAs are used as loading controls. (C) Region specific decay analysis of 5'- and 3'-fragments in amastigotes that have been cultured in promastigote (pH 7.0, 25° C) conditions for 5 hours. RNA samples were analyzed from different time points following actinomycin-D (ACT-D) treatment. (D) Region specific decay analysis of intact (RNase H untreated samples) in amastigotes that have been cultured in promastigote (pH 7.0, 25° C) conditions for 5 hours.

### Figure S2.

Overexposed version of Figure 5A.

**A****B****C****D****Figure S1.**

Overexposed version of Figure 5A

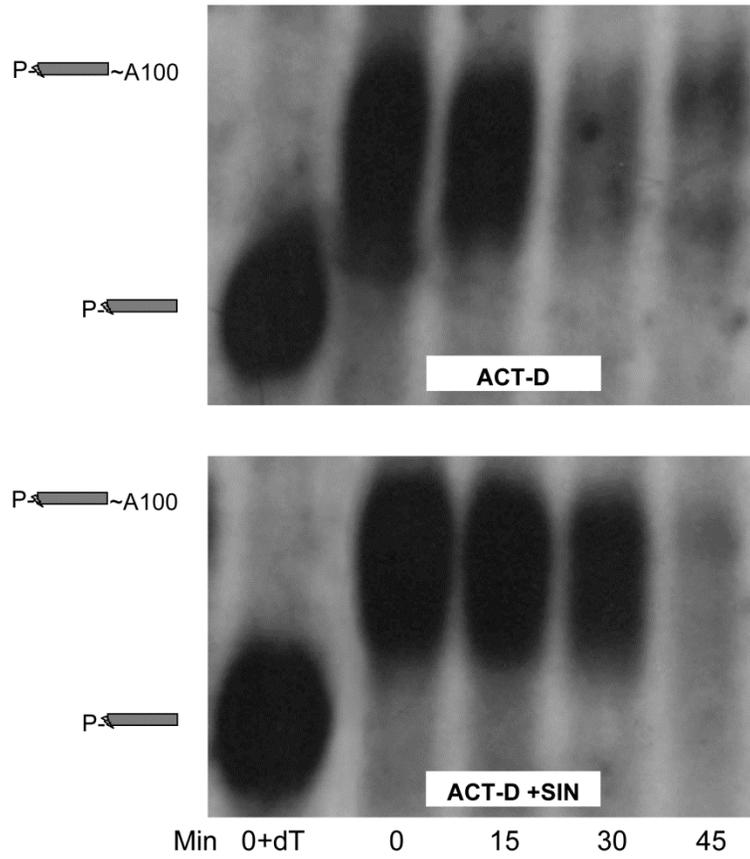


Figure S2.



## **Chapitre 9 : Une protéine à domaine Alba est impliquée dans la stabilité stade spécifique des transcrits des Amastines chez *Leishmania infantum*.**

### **Résumé**

Les delta-amastines font partie d'une famille de protéines de surface spécifiquement exprimée au stade amastigote chez *Leishmania*. Nous avons précédemment identifié un élément riche en U (URE) dans le 3'UTR de ce transcrit qui participe à la régulation de la stabilité du transcrit en fonction du cycle parasitaire. Ici, une chromatographie d'affinité utilisant une région contenant l'URE de l'amastine a identifié une protéine à domaine Alba (*LiAlba20*) comme liant l'ARNm. La suppression du gène *LiAlba20* provoque une déstabilisation de l'ARNm de l'amastine spécifiquement au stade amastigote. Des analyses transcriptomiques montrent que d'autres transcrits de delta-amastines sont différemment exprimés en absence de *LiAlba20*. La suppression du second gène codant pour une protéine Alba, *LiAlba13*, n'a pas d'effet sur l'accumulation de ce transcrit, aussi bien chez les promastigotes que les amastigotes. Nous montrons donc un rôle spécifique de *LiAlba20* dans la stabilité de transcrits différemment exprimés durant le développement chez *Leishmania*.

## Article

### **An Alba domain protein contributes to the stage-regulated stability of amastin transcripts in *Leishmania***

Aurélien Dupé, Carole Dumas and Barbara Papadopoulou\*

Pour publication dans *Molecular Microbiology*

Research Center in Infectious Disease, CHU de Quebec Research Center (CHUL),  
Department of Microbiology-Infectious Disease and Immunology, Laval University,  
Quebec, QC, Canada G1V 4G2

\*Corresponding author: Research Center in Infectious Disease, CHU de Quebec Research Center (CHUL), RC-709, 2705 Laurier Blvd, Quebec, QC, Canada, G1V 4G2. Tel: +1 (418) 525-4444, ext: 47608; Fax: +1 (418) 654-2715; E-mail: barbara.papadopoulou@crchul.ulaval.ca

**Running title:** Amastin gene regulation by Alba proteins

**Keywords:** *Leishmania*, amastins, Alba domain proteins, mRNA stability.

## Summary

*Leishmania infantum* promastigotes differentiate into amastigote forms within the phagolysosome of mammalian macrophages causing visceral leishmaniasis. Delta-amastins belong to a multigenic surface protein family of potential virulence factors that are specifically expressed in the amastigote life cycle stage by distinct regulatory elements in the 3'UTR controlling either mRNA stability or translation. Here, we provide novel insights on *trans*-acting factors regulating amastin developmental gene expression. Using RNA affinity chromatography with a 300-nt regulatory region within the amastin 3'UTR as bait, we identified an Alba domain protein of 25 kDa (*LiAlba20*) as a specific amastin mRNA-binding partner. Genomic depletion of *LiAlba20* results in amastin mRNA

destabilization specifically in amastigotes, supporting a role of *LiAlba20* in amastin gene regulation. Furthermore, as shown by comparative DNA microarray analysis, many delta-amastin transcripts but also other known developmentally regulated transcripts were downregulated in *LiAlba20*<sup>-/-</sup> knockout parasites. Inactivation of the second Alba domain gene, *LiAlba13*, does not seem to affect amastin mRNA stability in either life stage of the parasite. These data indicate an important role of Alba domain proteins in the regulation of *Leishmania* differentially expressed transcripts and open a new field of investigation for better understanding mechanisms contributing to post-transcriptional control in these parasites.

## Introduction

Protozoa parasites of the genus *Leishmania* are the causative agent of a broad range of diseases of high medical and veterinary importance, mainly in tropical and subtropical countries, known as leishmaniasis. *Leishmania infantum* species is responsible for visceral leishmaniasis in the “Old World”, the most severe form of the disease which is lethal, if left untreated. Each year, 200,000 to 400,000 cases of visceral leishmaniasis are estimated (Alvar *et al.*, 2012). *Leishmania* flagellated metacyclic promastigotes are transmitted in the mammalian host through the bite of a phlebotomine sand fly where they develop as non-motile amastigotes within the phagolysosome of macrophages (Bates, 2007). Elevated temperature and acidic pH constitute the two main environmental factors that trigger amastigote differentiation (Zilberstein & Shapira, 1994). *Leishmania* proteins are differentially regulated both as a survival strategy and in response to changes in nutrient availability (Brotherton *et al.*, 2010, Lahav *et al.*, 2011, McNicoll *et al.*, 2006).

*Leishmania*, together with other trypanosomatids such as *Trypanosoma brucei* (sleeping sickness) and *T. cruzi* (Chaga’s disease) share unusual mechanisms of regulation of gene expression (Kramer, 2012). Unrelated genes are organized as large, directional, polycistronic clusters on the same DNA strand of the chromosome. RNA polymerase II-mediated polycistronic transcription initiates at strand-switch regions in the absence of typical promoter sequences (Ivens *et al.*, 2005, Myler *et al.*, 1999). Mature individual mRNAs are obtained by two coupled cleavage reactions; namely *trans*-splicing that adds the splice leader RNA at the 5’ end of all *Leishmania* transcripts, and polyadenylation (Gunzl, 2010). As a result, regulation of gene expression in these parasites occurs

predominantly at the post-transcriptional level and involves changes in mRNA stability and translation efficiency, which are mainly mediated by sequences in the 3'untranslated regions (3'UTRs) (Clayton & Shapira, 2007, Haile & Papadopoulou, 2007). Thus, elucidating the function of 3'UTR regulatory sequences and RNA-binding proteins is crucial to our understanding of how genes are regulated throughout the life cycle of this parasite. *Leishmania* harbors a large family of extinct retroposons, SIDERs (Short Interspersed DEgenerated Retroposons), that are widely distributed within the 3'UTRs of unrelated transcripts and participate in post-transcriptional regulation (Bringaud *et al.*, 2007, Smith *et al.*, 2009). Uridine-rich elements (UREs), initially reported as destabilization factors in proto-oncogenes and cytokine mRNAs in mammalian cells (Chen & Shyu, 1995, Schoenberg & Maquat, 2012), and URE-binding proteins have been described in *T. brucei* (Hotz *et al.*, 1997, Mayho *et al.*, 2006, Quijada *et al.*, 2002) and *T. cruzi* (Di Noia *et al.*, 2000, Li *et al.*, 2012, Rodrigues *et al.*, 2010, D'Orso *et al.*, 2003, D'Orso & Frasch, 2002) and more recently in *Leishmania* (Haile *et al.*, 2008). Other non-categorized elements within 3'UTRs involved in stage-specific mRNA regulation, such as the PRE element AUGUAnAGUn are responsible for the degradation of transcripts encoding paraflagellar rod subunits or an antigenic protein in *L. mexicana* (Holzer *et al.*, 2008, Mishra *et al.*, 2003).

*Amastin* genes encode surface antigens putatively related to virulence and are members of a multi-gene family comprising more than 40 genes in *L. infantum*. Amastins have been described initially in *T. cruzi* (Teixeira *et al.*, 1994) and also in *Leishmania* (Wu *et al.*, 2000) and most of them are specific to the amastigote stage of the parasitic life cycle (Rochette *et al.*, 2005, Rochette *et al.*, 2009). A large phylogenetic study (Jackson, 2010) has reported that the *amastin* gene family is sub-divided into four groups ( $\alpha$ - to  $\delta$ -amastins) and that members of each group share common regulation patterns. The  $\delta$ -amastins are more specific to *Leishmania* and show the most drastic mRNA upregulation in amastigotes. Within this group, LinJ.34.1010 has been extensively studied in our laboratory and is referred to as "the amastin" in this article. The regulation of amastins in *Leishmania* involves distinct elements within their long 3'UTR. We have shown previously that the last ~800 nt of the amastin 3'UTR contains a 450 nt SIDER1 element and a 300 nt element, which both stimulate translation specifically in the amastigote stage (Boucher *et al.*, 2002,

McNicoll *et al.*, 2005). An additional 100 nt URE located upstream of the above regulatory elements was shown to be responsible for amastin mRNA degradation, specifically in the promastigote life cycle stage by a deadenylation-independent pathway (Haile *et al.*, 2008). However, the mechanism of amastin developmental gene regulation remains poorly understood, as no *trans*-acting factors regulating these mRNAs have been identified thus far.

In this study, we identified a 25 kDa Alba domain protein (*LiAlba20*) as a specific RNA-binding factor to the amastin 3'UTR and showed that genetic depletion of *LiAlba20* results in amastin mRNA destabilization, specifically in the amastigote life cycle stage of the parasite. We further showed that several members of  $\delta$ -amastins but also of other developmentally regulated transcripts were downregulated in the absence of *Alba20*. These data demonstrate for the first time a role of Alba domain proteins in mRNA stability in protozoan parasites.

## Results

### **An Alba domain protein, *LiAlba20*, binds specifically to the amastin 3'UTR**

To get a better insight into the mechanism of amastin (LinJ.34.1010) developmental gene regulation, first we looked for protein factors interacting with the 3'UTR U-rich element (URE) shown previously to play a role in the degradation of amastin mRNA in the promastigote life stage (Haile *et al.*, 2008). Initial UV-cross linking experiments with the 100 nt URE did not reveal any specific protein binding *in vitro*. However, using a larger RNA fragment of 300 nt harboring the URE and 100 nt flanking sequences at either side, we detected a ~20 kDa protein bound specifically to this RNA (data not shown). In order to identify this protein, we used RNA affinity chromatography with the 300 nt RNA region fused at the 3'end to a ~30 nt D8-tag sequence (Fig. 1A). The D8-tag allows the 300 nt RNA to bind to a Sephadex column, which could then be used to assess protein binding to this RNA. Eluted proteins were loaded on acrylamide gel (Fig. 1B) and bands between ~15-25 kDa specifically bound to the 300 nt-D8 RNA were sent for MS/MS peptide identification. MS/MS analysis identified a 25 kDa Alba domain protein harboring also an RGG-box (LinJ.34.2410, named *LiAlba20* in this study) as the protein with the higher number of unique peptides (Fig. 1C) bound to the 300 nt amastin 3'UTR region.

Interestingly, Alba and RGG domains are both known to be involved in RNA and/or DNA binding in *Archaea* and eukaryotes (Aravind *et al.*, 2003). RNA affinity chromatography was repeated under the same conditions and the eluted proteins were used for Western blot analysis using a specific antibody produced against the *T. brucei* ortholog *TbAlba3* (Fig. 1D). A single band at 25 kDa, corresponding to *LiAlba20* was obtained, hence confirming the mass spectrometry data.

To assess the specificity of *LiAlba20* protein binding to the 300 nt region within the amastin 3'UTR, we carried out RNA co-immunoprecipitation (RIP) assays using an anti-HA antibody and promastigote cell extracts. The *TbAlba3* antibody used in Figure 1D was not adequate for immunoprecipitation. For this reason, we made a HA-*LiAlba20* fusion protein where the HA tag was cloned in frame at the N-terminus of *LiAlba20* to avoid any interference with the RGG-box domain present at the C-terminus. Episomal expression of the HA-*LiAlba20* protein was confirmed by western blot (Fig. 2A). RNAs co-eluted with HA-*LiAlba20* were reverse-transcribed and tested for gene specific RT-PCR. Three couples of primers were tested and all confirmed the RNA affinity chromatography results shown that the 300 nt region within the amastin 3'UTR interacts with *LiAlba20* protein (Fig. 2B). This interaction seems specific, as control transcripts such as *BT1* or *GAPDH* were not detectable by this approach. The amastin transcript was not detected in the absence of the RT enzyme, strongly suggesting that *LiAlba20* binds to the *amastin* transcript *in vivo* but not to its DNA. The *LiAlba20*-amastin 3'UTR interaction was also investigated using amastigote cell extracts. The increased association of amastin mRNA with the HA-*LiAlba20* protein in amastigotes, as determined by RT-PCR, is explained by the fact the amastin transcript is on average 7- to 10-fold more abundant in this life cycle stage (Fig. 2C and Haile *et al.*, 2008, Wu *et al.*, 2000). However, using this approach we cannot conclude if there is a difference in the affinity of *LiAlba20* protein binding to the amastin transcript between the two life cycle stages of the parasite.

The RIP experiment suggests that episomally transfected HA-*LiAlba20* protein is able to interact with *amastin* mRNA in both life cycle stages of the parasite. To further investigate these interactions, we examined the expression of *LiAlba20* mRNA and protein throughout the parasite developmental stages. In *L. infantum*, previous microarray analysis (Rochette *et al.*, 2009) and proteomic studies (Brotherton *et al.*, 2010) indicated that the

*LiAlba20* mRNA and protein are constitutively expressed. Northern blot hybridization with a *LiAlba20* specific probe (Fig. 3A) and Western blot using the *TbAlba3* antibody (Fig. 3B) showed that endogenous *LiAlba20* is indeed constitutively expressed and could therefore interact with the amastin mRNA in both life cycle stages.

### ***LiAlba20* does not bind to the URE within the amastin 3'UTR**

The RNA affinity experiment in Figure 1 was carried out with a D8-tagged RNA fragment of 300 nt harboring the 100 nt URE and flanking sequences on either side. First, we investigated whether the *LiAlba20* protein binds the URE. For that, we repeated the RNA affinity experiment with D8-tagged RNA of 200 nt lacking the URE ( $\Delta$ URE-D8). MS/MS analysis of the protein elution fraction identified Alba20 as the most prominent interacting factor of this truncated RNA, hence suggesting that Alba20 does not interact with the URE on the *amastin* 3'UTR (data not shown). We therefore looked for the presence of other conserved sequence motifs in the amastin (LinJ.34.1010) URE flanking regions. To increase the statistical power of this analysis we extended our study to all  $\delta$ -*amastin* transcripts. Using the MEME/MAST application (Bailey *et al.*, 2009), we compared the ~2 kb 3'UTR sequences of 30  $\delta$ -*amastins* with 50 ~2 kb 3'UTR control sequences from constitutively expressed transcripts. A MEME MAST motif search revealed four motifs within the 300 nt region, all outside the 100 nt URE, that were conserved in the 3'UTR of other  $\delta$ -*amastins* (Fig. S1A). Interestingly, the 5'-GUGCGUGYGCGUGC-3' motif was found in all  $\delta$ -*amastins* tested (Fig. S1B). Even more interestingly, this motif overlaps with the LeishCICS-c1-3338 motif, predicted *in silico* in several amastigote-specific transcripts (Vasconcelos *et al.*, 2012). In amastin LinJ.34.1010, this motif is present at ~20 nt downstream of the URE and is therefore included in the 300 nt RNA used as bait in RNA affinity chromatography experiments (Fig. 1). This conserved motif could be an interesting candidate to investigate further as a putative *cis*-acting sequence within  $\delta$ -*amastin* 3'UTRs for interaction with Alba20 protein.

### **Genetic inactivation of *LiAlba20* results in a decreased stability of the amastin mRNA specifically in amastigotes**

To further investigate the potential role of *LiAlba20* in *amastin* gene regulation, we generated an Alba20 null mutant in *L. infantum* by replacing the two endogenous *LiAlba20*

alleles by the hygromycin and puromycin resistance genes, respectively (Fig. 4A). Correct integration of the targeting cassettes was verified by Southern blot (Fig. 4B) and PCR amplification (data not shown). The absence of the *LiAlba20* transcript and protein was also confirmed by northern and Western blots (data not shown). As a control, we generated an add-back strain by expressing the pSP $\alpha$ NEO $\alpha$ -HA-*LiAlba20* plasmid into the *LiAlba20*<sup>-/-</sup> background. To assess the role of the *LiAlba20* protein in amastin mRNA developmental regulation, we carried out northern blot analysis of promastigote and amastigote total RNA from wild type and *LiAlba20*<sup>-/-</sup> null mutant and add-back strains. Genetic depletion of *LiAlba20* does not seem to affect significantly amastin gene expression in promastigotes (Fig. 5A), and this was reproducible in seven independent northern blot experiments (data not shown). Interestingly, a 3.4-fold decrease in amastin mRNA accumulation was observed in *LiAlba20*-depleted parasites, (Fig. 5A). This effect is specific to the loss of Alba20 as it was fully reverted in the add-back mutant strain in six reproducible northern blot experiments (Fig. 5A and data not shown).

To determine whether decrease in amastin mRNA levels in *LiAlba20*<sup>-/-</sup> amastigotes is due to mRNA destabilization, the half-life of the endogenous amastin transcript was measured both in wild type and *LiAlba20*<sup>-/-</sup> cells following inhibition of mRNA synthesis and maturation with actinomycin-D and sinefungin, respectively (Fig. 5B, normalized data are shown in Fig. 5C). As published previously (Haile *et al.*, 2008), the endogenous amastin mRNA has a half-life of >150 min in amastigotes. Here, we show that half-life of the amastin mRNA in *LiAlba20*<sup>-/-</sup> amastigotes is reduced to ~80 min. These results suggest that Alba20 is involved in the stabilization of amastin mRNA specifically in the amastigote stage. However, this stabilization effect is only partial (~3-fold compared to 6- to 8-fold in wild type cells), suggesting that other factors are also contributing to the increased stability of amastin mRNA in amastigotes.

### ***LiAlba20* depletion affects regulation of several $\delta$ -amastin members and other amastigote-specific transcripts**

To assess global transcriptome changes in parasites lacking Alba20, we carried out DNA microarray experiments to compare wild type and *LiAlba20*<sup>-/-</sup> both in promastigotes and axenic amastigotes. The whole list of genes whose expression is modulated in *LiAlba20*<sup>-/-</sup> is presented in Tables S2 and S3. In total, 217 protein coding transcripts were

differentially regulated in the absence of *LiAlba20*, corresponding to 2.6% of the *L. infantum* genome (TritrypDB version 4.2, Aslett *et al.*, 2010). In *LiAlba20*<sup>-/-</sup> promastigotes, 35 genes were found to be upregulated, comprising four promastigote-specific genes and 17 (49%) amastigote-specific genes (Fig. 6A). The latest category includes the amastigote-specific *p1/s1 nuclease* (Farajnia *et al.*, 2004), the virulence factor *A2* (Zhang *et al.*, 2003), members of the hydrophilic acylated surface HASP proteins (Alce *et al.*, 1999), and two members of the  $\delta$ -*amastin* family (LinJ.34.1700 and LinJ.34.1730; 1.8-fold variation). Among the 32 downregulated genes in *LiAlba20*<sup>-/-</sup> promastigotes, two were previously reported as being amastigote-specific genes and six (18.7%) as being promastigote-specific genes (Fig. 6A and Table S2). In axenic amastigotes lacking *LiAlba20*, 77 transcripts were found upregulated from which 16 (20.8%) were previously reported as being promastigote-specific and five as being amastigote-specific. The 96 downregulated genes in *LiAlba20*<sup>-/-</sup> amastigotes comprise 30 (31%) known amastigote-specific transcripts (Fig. 6A and Tables 1 and S3), including *HASPB*, the mitochondrial membrane protein *p27* (Sharma *et al.*, 2010), *A2*, *META1* (Puri *et al.*, 2011) and ten  $\delta$ -*amastin* genes (LinJ.08.0680, LinJ.08.0690, LinJ.08.0720, LinJ.08.0780, LinJ.08.0790, LinJ.08.1320, LinJ.34.1010, LinJ.34.1030, LinJ.34.1730, LinJ.34.4370) with a 1.7- to 2.6-fold variation.

Taken together, these observations suggest that depletion of *LiAlba20* affects preferentially the expression of amastigote-specific rather than promastigote-specific transcripts, among them some important *Leishmania* virulence factors, such as  $\delta$ -*amastins* and *A2*. *Alba20* genetic inactivation alters also the expression of other cellular processes, such as transport, metabolic and biosynthetic processes, and motility (Fig. 6B). Fold variation for most of these transcripts in the absence of *LiAlba20* is  $\sim 2$ , suggesting that *LiAlba20* is not the only protein involved in their stage-specific accumulation. The most highly modulated *amastin* transcripts belong to the  $\delta$ -*amastin* subfamily and their accumulation is significantly decreased in amastigotes lacking *LiAlba20* (Table 1). Decreased accumulation of these  $\delta$ -*amastin* transcripts is most likely due to a decrease in their stability as also seen for LinJ.34.1010 (Figs. 5B and 5C).

### ***LiAlba13*, the second Alba domain protein in *Leishmania*, is not involved in amastin mRNA stability**

*Leishmania* species have only two Alba domain proteins; *LiAlba20* (Rpp25-like, ortholog of *TbAlba3* and *TbAlba4*) and *LiAlba13* (LinJ.13.0270, Rpp20-like, ortholog of *TbAlba1* and *TbAlba2*). Although RNA affinity experiments revealed that only *LiAlba20* binds to the amastin mRNA (Figs. 1C and 1D), co-immunoprecipitation experiments (Dupé A. *et al.*, manuscript in preparation) indicated that these two Alba domain proteins interact together in *L. infantum*. To assess potential interactions between *LiAlba13* and amastin mRNA, the *LiAlba13* protein tagged with HA at the C-terminus was expressed in *L. infantum* and RIP-RT-PCR experiments were carried out as previously described with HA-*LiAlba20*. *LiAlba13*-HA could pull down the LinJ.34.1010 amastin transcript both in promastigotes and amastigotes (Fig. 7A). To assess the potential implication of this interaction in amastin mRNA regulation, we next generated a null mutant lacking *LiAlba13* as well as an add-back mutant strain (Figs. 7B and 7C). Northern blot analysis indicated that in contrast to *LiAlba20*, *LiAlba13* does not affect amastin transcript expression in either life cycle stage of the parasite (Fig. 7D), suggesting that the two *Leishmania* Alba domain proteins have possibly different regulatory roles in this organism.

### **Discussion**

In this study, we show that the *L. infantum* Alba20 domain protein binds to sequence motifs in the 3'UTR of several  $\delta$ -amastin transcripts to regulate their expression in a stage-specific manner. In the case of the  $\delta$ -amastin LinJ.34.1010, *LiAlba20* stabilizes the transcript specifically in the amastigote life cycle stage of the parasite. In contrast to *LiAlba20*, *LiAlba13*, the other Alba domain protein in *Leishmania* has no effect on amastin transcript accumulation. Thus far, very few RNA-binding proteins with a known regulatory function have been identified in *Leishmania* or in other kinetoplastids, and to our knowledge this is the first report on the implication of an Alba domain protein in mRNA stability in these parasites.

Alba domain proteins are found only in *Archaea* and eukaryotes and have different affinities for both DNA and RNA. In *Archaea*, they can interact with both RNA (Guo *et al.*, 2003, Jelinska *et al.*, 2010) and DNA (Laurens *et al.*, 2012, Sandman & Reeve, 2005),

being part of the nucleosome in *Chrenarchaea*. In higher eukaryotes, they constitute the Rpp20 and Rpp25 subunits of RNase P /MRP involved in rRNA and tRNA maturation (Aravind *et al.*, 2003, Welting *et al.*, 2007). In protozoan parasites such as *Plasmodium*, Alba domain proteins possess both abilities, being implicated in untranslated mRNA storage in oocyte P-granules (Mair *et al.*, 2010), but interact also with DNA through the FACT complex (*Facilitates Chromatin Transcription*) in male gametocytes (Laurentino *et al.*, 2011), TARE (*Telomere-Associated Repetitive Elements*) and the *var* promoter region regulating the expression of virulence genes (Chene *et al.*, 2012, Goyal *et al.*, 2011). In *Toxoplasma gondii*, Alba proteins interact with the translation machinery and seem to play a role in the response to stress (Gissot *et al.*, 2013). They also co-immunoprecipitate with TgNF3, a histone-interacting protein implicated in stage-specific gene silencing (Olguin-Lamas *et al.*, 2011). In *Trypanosoma brucei*, TbAlba interacts with the translation machinery and appears to inhibit the procyclic-specific GPEET membrane protein expression at the level of translation initiation (Mani *et al.*, 2011). Furthermore, TbAlba3 and TbAlba4 co-localize with the stress granule marker TbDHH1 after starvation stress (Subota *et al.*, 2011). Here, our studies in *L. infantum* also demonstrate a link between Alba domain proteins and stage-specific gene expression. Based on comparative DNA microarray analysis, LiAlba20 seems to be involved not only in the regulation of several  $\delta$ -*amastin* transcripts but also of other amastigote-specific putative virulence factors in *Leishmania* such as the *A2* and *HASPB* genes.

In the case of the  $\delta$ -*amastin* LinJ.34.1010 transcript, it becomes now clear that its developmental gene regulation is rather complex and involves distinct regulatory elements in the 3'UTR, which potentially interact with different RNA-binding proteins. Amastin gene expression is regulated at the level of translation initiation both by a conserved 450-nt SIDER1 retroposon element in the 3'UTR (Boucher *et al.*, 2002) and a second downstream region of 100 nucleotides located at the 3'-end of several amastin transcripts in response to heat shock (McNicoll *et al.*, 2005). In addition, we have shown that acidic pH encountered in the phagolysosomes of macrophages, the location of amastigote differentiation, triggers the accumulation of amastin transcripts by a distinct mechanism that is independent of the 450-nt and 100-nt elements (McNicoll *et al.*, 2005). In the promastigote life stage of the parasite, we have reported previously that amastin mRNA degradation is partly mediated

by a U-rich element within the 3'UTR in a deadenylation-independent manner (Haile *et al.*, 2008). Here, we show that Alba20 promotes a partial stabilization of the amastin transcript, specifically in amastigotes. Our data indicate a ~3.4-fold decrease in amastin mRNA steady-state levels in *L. infantum* amastigotes lacking Alba20 as compared to a ~6-8-fold increased accumulation of the amastin transcript in wild type cells (Rochette *et al.*, 2005, Wu *et al.*, 2000), suggesting that Alba20 is not the only *trans*-acting factor involved in amastin mRNA preferential accumulation in amastigotes. Binding of *LiAlba20* to the amastin 3'UTR does not involve the URE, as RNA affinity chromatography using a 200 nt RNA lacking the URE revealed also the association of *LiAlba20* protein with this truncated RNA. By MEME/MAST motif search, we identified a 14-mer sequence motif downstream of the URE, which is conserved in the majority of  $\delta$ -*amastin* 3'UTRs and could be a putative candidate for Alba20 protein binding. However, additional experiments are required to confirm the *cis*-acting sequence requirements for  $\delta$ -*amastin*-Alba20 interactions.

Although *LiAlba20* and *LiAlba13* proteins were found associated with the amastin transcript in both life cycle stages of the parasite, only *LiAlba20* regulates amastin mRNA stability, specifically in amastigotes. As shown in this study, *LiAlba20* protein is constitutively expressed in both life cycle stages and based on recent iTRAQ analysis, *LiAlba13* is also not strongly regulated during the parasite's differentiation (Lahav *et al.*, 2011). It is worth mentioning that only *LiAlba20* possesses a RGG domain known to be involved in RNA binding both in *Archaea* and eukaryotes (Aravind *et al.*, 2003). Thus, the interaction of *LiAlba13* with the amastin 3'UTR is most likely indirect and could be explained by the fact *LiAlba20* forms a complex with *LiAlba13*, as suggested by co-immunoprecipitation experiments (Dupé A. *et al.*, manuscript in preparation). Usually, the correct expression of Alba domain proteins depends on the presence of other Alba proteins. This has been reported in *T. gondii*, *T. brucei* and in the case of Rpp20 / Rpp25 (Gissot *et al.*, 2013, Mani *et al.*, 2011, Welting *et al.*, 2007). However, expression of the endogenous *LiAlba20* seems to be equal between *L. infantum* wild type and parasites lacking *LiAlba13* (data not shown). This suggests that there was no compensation in the absence of *LiAlba13* and that in some cases these proteins may act in an independent manner. *LiAlba13* might play a more important role in translation as suggested by the overexpression of *LiAlba13*

protein but not of *LiAlba20* in *L. infantum* treated with gentamicin (Daneshvar *et al.*, 2012), a drug that binds to the 30S subunit of the bacterial ribosome. In promastigotes, neither *LiAlba20* nor *LiAlba13* proteins have an effect on amastin gene regulation. This could be due either to the absence of additional protein partners or to differential post-translational modifications (PTMs) of these proteins, and most specifically of *LiAlba20*, rendering them hence unable to regulate amastin transcript stability. It is also possible that the fast URE-mediated amastin mRNA degradation in promastigotes somehow masks the effect of the Alba domain proteins. On the other hand, in amastigotes, Alba20 could be post-translationally modified and able to recruit other *trans*-acting factors that together contribute to the stabilization of  $\delta$ -amastins, although this remains to be validated experimentally.

Altogether, these data contribute to our better understanding of how amastin transcripts are regulated throughout the developmental stages of *Leishmania*. Our findings that Alba domain proteins contribute to the stability not only of several  $\delta$ -amastin transcripts but also possibly of other developmentally regulated transcripts provide the first insight on *trans*-acting factors implicated in stage-specific mRNA regulation in *Leishmania*. Future studies would be required to understand the role of Alba domain proteins and the RGG-box more specifically in mRNA binding and post-transcriptional regulation in protozoan parasites.

## Experimental Procedures

### Cell lines and culture

The parental strain *Leishmania donovani infantum* MHOM/MA/67/ITMAP-263 was used in this study. Promastigotes were cultured at 25°C in SDM-79 medium (pH 7.0), supplemented with 5 mg ml<sup>-1</sup> hemin and 10% heat-inactivated fetal calf serum (Multicell, Wisent Inc.). Exponential phase promastigotes were obtained after 3 days of growth (OD<sub>600nm</sub> = ~0.3). To obtain well differentiated axenic amastigotes, 2 to 3 passages were carried out in MAA/20 medium (pH 5.8) at 37°C with 5% CO<sub>2</sub> as previously described (Sereno & Lemesre, 1997). When needed, G418 (30 µg ml<sup>-1</sup> for episomal transfection), hygromycin B (160 µg ml<sup>-1</sup> for genomic integration) and/or puromycin (70 µg ml<sup>-1</sup> for genomic integration) were added for selection.

### **DNA manipulations, transfections and Southern blot**

Genomic DNA used for PCR amplification is from *L. infantum* MHOM/MA/67/ITMAP-263. The primers (ordered at Integrated DNA Technologies) are listed in Table S1. The knockouts of *Alba20* and *Alba13* genes were done by fusing a 500 nt PCR fragment of both 5' and 3' flanking regions of the targeted ORF to the hygromycin and puromycin resistance genes. For episomal transfections of HA-tagged proteins, we used the pSP $\alpha$ NEO $\alpha$  plasmid already described in (Padmanabhan *et al.*, 2012). The ORF sequences were inserted into the XbaI (5'end) and NdeI (3'end) sites of the vector. Restriction sites and HA tag were incorporated within primers and all inserts were first cloned in pGEM-T easy vector (Promega) and sequenced to verify that no mutations were present. *L. infantum* transfected cells were cloned using serial half dilution in SDM-79 medium and individual clones were obtained after 2–3 weeks. All southern blot experiments were run as previously described. The digested DNA was denatured on the gel, transferred on Hybond-N membranes (Amersham Biosciences) overnight, and nucleic acids were cross-linked to the membrane by UV-cross linking exposure with a Stratalinker. 100  $\mu$ g of PCR products corresponding to the ORF of the targeted genes or selectable marker genes were marked with  $^{32}$ P-dCTP (Invitrogen). Probes were purified using the mini-elute columns (Qiagen), denatured and incubated overnight. After washings, exposure of films (Kodak) on the membranes was performed.

### **Protein extraction and Western blotting**

Total protein extracts were obtained by pelleting 1 ml of logarithmic grown cells ( $\sim 10^6$  cells ml $^{-1}$ ). Pellets were resuspended in 100  $\mu$ l of 2x Laemmli Sample Buffer and boiled before loading 5  $\mu$ l onto gel. For Western blotting, SDS-PAGE gels were transferred on Immobilon-P polyvinylidene difluoride membranes (Millipore). Anti-HA (100 ng ml $^{-1}$ , Applied Biological Materials Inc.) or anti-*TbAlba2040/2* [(GeneID: Tb927.4.2040; a kind gift from Dr Philippe Bastin (Subota *et al.*, 2011), Pasteur Institute, Paris, France)] and HRP-linked anti-mouse (NEB) antibodies were used for hybridization 1 hour at room temperature. Chemiluminescence analysis was done using the Amersham Hyperfilm<sup>TM</sup>/ECL<sup>TM</sup> kit (GE HealthCare) and exposure of films (Kodak) on the membrane. For ARN affinity purification and RNA immunoprecipitation,  $\sim 5 \cdot 10^9$  cells were resuspended in 1 ml of cytoplasmic extraction buffer (CEB: 10 mM Hepes, 3 mM MgCl $_2$ ,

14 mM KCl, pH 7.5, 5% glycerol and 0.2% NP-40, supplemented with 1 mM PMSF, 5 µl RNase OUT (Invitrogen) and 1X cOmplete ULTRA Tablets, EDTA-free EASYpack Roche) and lysed with 15 to 20 strokes using a Douncer. Insoluble material was removed by a 10 min centrifugation at 13 000 rpm and proteins were quantified using the Bradford method (BioRad Protein assay).

### **RNA affinity chromatography using Sephadex and *in vitro*-transcribed D8-tagged RNA**

Amastin 3'UTR sequences were amplified by PCR using the Phusion enzyme (Thermo Scientific) and primers containing the promoter sequence for T7 RNA polymerase at the 5' end and the D8-RNA tag at the 3' end (see Table S1). The D8-RNA tag has binding affinity for Sephadex matrix, as described by Dr Engelke's lab (Srisawat & Engelke, 2002, Walker *et al.*, 2008). PCR products were transcribed *in vitro* using the MEGAscript transcription kit (Ambion). RNA probes were purified with NucAway spin columns (Ambion) and quantified with Nanodrop. 1.5 ml of Sephadex G-100 (GE Healthcare) diluted in binding buffer (50 mM Hepes, pH 7.4, 10 mM MgCl<sub>2</sub>, 100 mM NaCl) was loaded on Ultrafree-MC columns with a cut off of 5.0 µm (Millipore). 100 µg of RNA and 5 µl RNase OUT (Invitrogen) were added and incubated for 1h30 at 4°C in a rotator. Five washing steps with 5 ml binding buffer were done before the addition of ~8 µg of proteins. After 20 min of incubation at 4°C in a rotator, columns were washed 5 times with 5 ml binding buffer and elution was done in 300 µl buffer (8 M Urea; 5 mM EDTA). Samples were TCA-precipitated, resuspended in 2x Laemmli Sample Buffer and loaded on a 15% SDS-PAGE for Sypro staining and MS/MS peptide identification.

### **RNA analysis**

~10<sup>7</sup> *Leishmania* cells were pelleted and RNA was purified using 1 ml of Trizol (Invitrogen) according to the manufacturer's instructions. mRNA half-lives were determined by treating the cells with 0.5 µg ml<sup>-1</sup> actinomycin-D (Sigma) and 2.5 µM sinefungin (Sigma). Time points include centrifugation times. Purified RNAs were quantified using a Nanodrop. 15 to 25 ng were used in a standard northern blot protocol, with Hybond-XL membranes (Amersham Biosciences), as previously described (Haile *et*

*al.*, 2008). Transcript relative expression was determined by quantification of the signals using ImageQuant.

### **RNA Immunoprecipitation (RIP) and RT-PCR**

Total proteins from  $5 \cdot 10^9$  *Leishmania* cells were used. RIP was performed as described (Peritz *et al.*, 2006) using an anti-HA antibody (10  $\mu$ g, from Applied Biological Materials Inc.) and Dynabeads Protein G (60  $\mu$ l from Invitrogen) in 200  $\mu$ l PSB-0.1% Tween. After washing with 1x PBS-0.1% Tween supplemented with protease and RNase inhibitors, RNA was eluted using 300  $\mu$ l of Proteinase K buffer (30 mM Tris-HCl pH 7.5), 15  $\mu$ l Proteinase K (1 mg ml<sup>-1</sup> final, Roche) and 2  $\mu$ l RNase OUT. RNA purification was performed with Phenol-Chloro-Isoamyl protocol and the pellet was finally resuspended in 20  $\mu$ l of DEPC-treated water. After DNase treatment and inactivation, 5  $\mu$ l of eluted RNA were reverse-transcribed using the SuperScript Reverse transcriptase III kit following manufacturer's instructions in a 20  $\mu$ l final reaction volume. Another 5  $\mu$ l of RNA was used as a control without reverse transcriptase. After the first strand synthesis, 2  $\mu$ l of RNA were used in a usual PCR reaction in a 25  $\mu$ l final volume.

### **DNA microarray data processing and statistical analysis**

To account for sample heterogeneity, variation between slides and variations due to hybridization, cDNA microarray experiments were run in triplicate for *L. infantum* axenic amastigotes and promastigotes using a full-genome array as previously described (Rochette *et al.*, 2009). Briefly, total RNA from mid-log wild type and *LiAlba20*<sup>-/-</sup> strains was prepared using RNeasy Plus (Qiagen) and assessed using RNA 6000 Nano Assay Chips (Agilent Technologies). cDNA was obtained starting with 5  $\mu$ g of total RNA using random primer hexamer (GE Healthcare) and aminoallyl-dUTP mix (Sigma) following the protocol for Superscript III (Invitrogen) and purified using MinElute PCR purification columns (Qiagen). cDNA was coupled with Alexa 555 or Alexa 647 dyes (Invitrogen) and purified with MinElute PCR purification columns. After pre-hybridization of the array, probes were hybridized overnight at 42°C (2.5X modified Denhardt, 30% deionized formamide, 6X SSPE, 0.5% SDS, 0.1 mg ml<sup>-1</sup> ssDNA and 0.75 mg ml<sup>-1</sup> yeast tRNA). The fluorescence signal intensities were measured using the Perkin Elmer ScanArray Express scanner and statistical analysis was performed in R 2.2.1 using the false discovery rate (FDR) method

with a threshold p-value of 0.05. Only genes with a fold difference expression greater than 1.68 (upregulated in *LiAlba20<sup>-/-</sup>*) or smaller than 0.59 (downregulated in *LiAlba20<sup>-/-</sup>*) between wild type and *LiAlba20<sup>-/-</sup>* were considered as differentially expressed, as previously done (Rochette *et al.*, 2009). All microarray data will be freely available on the GEO NCBI database in the MIAME format; <http://www.ncbi.nlm.nih.gov/geo/>.

## **Acknowledgements**

We would like to thank Dr Philippe Bastin (Pasteur Institute, Paris, France) for providing us with the *TbAlba3* antibody, Dr Osvaldo de Melo Neto (Centro de Pesquisas Aggeu Magalhães/Fiocruz, Recife, PE, Brazil) for scientific discussions about this work, and Drs Eric Madore and Frédéric Raymond for their help in DNA microarray experiments and analysis. AD was a recipient of a doctoral fellowship from the Canadian Institutes of Health Research (CIHR) Strategic Training Program STP-53924 and the Centre for Host-Parasite Interactions (CHPI). This study was supported by the CIHR MOP-12182 grant awarded to BP. BP is a member of the Centre for Host-Parasite Interactions 'Programme Regroupements Stratégiques' of the Fonds de Québec pour la Recherche sur la Nature et les Technologies.

## References

- Alce, T.M., S. Gokool, D. McGhie, S. Stager & D.F. Smith, (1999) Expression of hydrophilic surface proteins in infective stages of *Leishmania donovani*. *Mol Biochem Parasitol* **102**: 191-196.
- Alvar, J., I.D. Velez, C. Bern, M. Herrero, P. Desjeux, J. Cano, J. Jannin & M. den Boer, (2012) Leishmaniasis worldwide and global estimates of its incidence. *PLoS One* **7**: e35671.
- Aravind, L., L.M. Iyer & V. Anantharaman, (2003) The two faces of Alba: the evolutionary connection between proteins participating in chromatin structure and RNA metabolism. *Genome Biol* **4**: R64.
- Aslett, M., C. Aurrecochea, M. Berriman, J. Brestelli, B.P. Brunk, M. Carrington, D.P. Depledge, S. Fischer, B. Gajria, X. Gao, M.J. Gardner, A. Gingle, G. Grant, O.S. Harb, M. Heiges, C. Hertz-Fowler, R. Houston, F. Innamorato, J. Iodice, J.C. Kissinger, E. Kraemer, W. Li, F.J. Logan, J.A. Miller, S. Mitra, P.J. Myler, V. Nayak, C. Pennington, I. Phan, D.F. Pinney, G. Ramasamy, M.B. Rogers, D.S. Roos, C. Ross, D. Sivam, D.F. Smith, G. Srinivasamoorthy, C.J. Stoeckert, Jr., S. Subramanian, R. Thibodeau, A. Tivey, C. Treatman, G. Velarde & H. Wang, (2010) TriTrypDB: a functional genomic resource for the *Trypanosomatidae*. *Nucleic Acids Res* **38**: D457-462.
- Bailey, T.L., M. Boden, F.A. Buske, M. Frith, C.E. Grant, L. Clementi, J. Ren, W.W. Li & W.S. Noble, (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res* **37**: W202-208.
- Bates, P.A., (2007) Transmission of *Leishmania* metacyclic promastigotes by phlebotomine sand flies. *Int J Parasitol* **37**: 1097-1106.
- Boucher, N., Y. Wu, C. Dumas, M. Dube, D. Sereno, M. Breton & B. Papadopoulou, (2002) A common mechanism of stage-regulated gene expression in *Leishmania* mediated by a conserved 3'-untranslated region element. *J Biol Chem* **277**: 19511-19520.
- Bringaud, F., M. Muller, G.C. Cerqueira, M. Smith, A. Rochette, N.M. El-Sayed, B. Papadopoulou & E. Ghedin, (2007) Members of a large retroposon family are determinants of post-transcriptional gene expression in *Leishmania*. *PLoS Pathog* **3**: 1291-1307.
- Brotherton, M.C., G. Racine, A.L. Foucher, J. Drummelsmith, B. Papadopoulou & M. Ouellette, (2010) Analysis of stage-specific expression of basic proteins in *Leishmania infantum*. *J Proteome Res* **9**: 3842-3853.
- Chen, C.Y. & A.B. Shyu, (1995) AU-rich elements: characterization and importance in mRNA degradation. *Trends Biochem Sci* **20**: 465-470.
- Chene, A., S.S. Vembar, L. Riviere, J.J. Lopez-Rubio, A. Claes, T.N. Siegel, H. Sakamoto, C. Scheidig-Benatar, R. Hernandez-Rivas & A. Scherf, (2012) PfAlbas constitute a new eukaryotic DNA/RNA-binding protein family in malaria parasites. *Nucleic Acids Res* **40**: 3066-3077.
- Clayton, C. & M. Shapira, (2007) Post-transcriptional regulation of gene expression in trypanosomes and leishmanias. *Mol Biochem Parasitol* **156**: 93-101.
- D'Orso, I., J.G. De Gaudenzi & A.C. Frasch, (2003) RNA-binding proteins and mRNA turnover in trypanosomes. *Trends Parasitol* **19**: 151-155.

- D'Orso, I. & A.C. Frasch, (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.
- Daneshvar, H., S. Wyllie, S. Phillips, P. Hagan & R. Burchmore, (2012) Comparative proteomics profiling of a gentamicin-attenuated *Leishmania infantum* cell line identifies key changes in parasite thiol-redox metabolism. *J Proteomics* **75**: 1463-1471.
- Di Noia, J.M., I. D'Orso, D.O. Sanchez & A.C. Frasch, (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.
- Farajnia, S., M.H. Alimohammadian, N.E. Reiner, M. Karimi, S. Ajdari & F. Mahboudi, (2004) Molecular characterization of a novel amastigote stage specific Class I nuclease from *Leishmania major*. *Int J Parasitol* **34**: 899-908.
- Gissot, M., R. Walker, S. Delhaye, T.D. Alayi, L. Huot, D. Hot, I. Callebaut, C. Schaeffer-Reiss, A.V. Dorsselaer & S. Tomavo, (2013) *Toxoplasma gondii* Alba Proteins Are Involved in Translational Control of Gene Expression. *J Mol Biol* **425**: 1287-1301.
- Goyal, M., A. Alam, M.S. Iqbal, S. Dey, S. Bindu, C. Pal, A. Banerjee, S. Chakrabarti & U. Bandyopadhyay, (2011) Identification and molecular characterization of an Alba-family protein from human malaria parasite *Plasmodium falciparum*. *Nucleic Acids Res* **40**: 1174-1190.
- Gunzl, A., (2010) The pre-mRNA splicing machinery of trypanosomes: complex or simplified? *Eukaryot Cell* **9**: 1159-1170.
- Guo, R., H. Xue & L. Huang, (2003) Ssh10b, a conserved thermophilic archaeal protein, binds RNA *in vivo*. *Mol Microbiol* **50**: 1605-1615.
- Haile, S., A. Dupé & B. Papadopoulou, (2008) Deadenylation-independent stage-specific mRNA degradation in *Leishmania*. *Nucleic Acids Res* **36**: 1634-1644.
- Haile, S. & B. Papadopoulou, (2007) Developmental regulation of gene expression in trypanosomatid parasitic protozoa. *Curr Opin Microbiol* **10**: 569-577.
- Holzer, T.R., K.K. Mishra, J.H. LeBowitz & J.D. Forney, (2008) Coordinate regulation of a family of promastigote-enriched mRNAs by the 3'UTR PRE element in *Leishmania mexicana*. *Mol Biochem Parasitol* **157**: 54-64.
- Hotz, H.R., C. Hartmann, K. Huober, M. Hug & C. Clayton, (1997) Mechanisms of developmental regulation in *Trypanosoma brucei*: a polypyrimidine tract in the 3'-untranslated region of a surface protein mRNA affects RNA abundance and translation. *Nucleic Acids Res* **25**: 3017-3026.
- Ivens, A.C., C.S. Peacock, E.A. Worthey, L. Murphy, G. Aggarwal, M. Berriman, E. Sisk, M.A. Rajandream, E. Adlem, R. Aert, A. Anupama, Z. Apostolou, P. Attipoe, N. Bason, C. Bauser, A. Beck, S.M. Beverley, G. Bianchetti, K. Borzym, G. Bothe, C.V. Bruschi, M. Collins, E. Cadag, L. Ciarloni, C. Clayton, R.M. Coulson, A. Cronin, A.K. Cruz, R.M. Davies, J. De Gaudenzi, D.E. Dobson, A. Duesterhoeft, G. Fazelina, N. Fosker, A.C. Frasch, A. Fraser, M. Fuchs, C. Gabel, A. Goble, A. Goffeau, D. Harris, C. Hertz-Fowler, H. Hilbert, D. Horn, Y. Huang, S. Klages, A. Knights, M. Kube, N. Larke, L. Litvin, A. Lord, T. Louie, M. Marra, D. Masuy, K. Matthews, S. Michaeli, J.C. Mottram, S. Muller-Auer, H. Munden, S. Nelson, H. Norbertczak, K. Oliver, S. O'Neil, M. Pentony, T.M. Pohl, C. Price, B. Purnelle,

- M.A. Quail, E. Rabbinowitsch, R. Reinhardt, M. Rieger, J. Rinta, J. Robben, L. Robertson, J.C. Ruiz, S. Rutter, D. Saunders, M. Schafer, J. Schein, D.C. Schwartz, K. Seeger, A. Seyler, S. Sharp, H. Shin, D. Sivam, R. Squares, S. Squares, V. Tosato, C. Vogt, G. Volckaert, R. Wambutt, T. Warren, H. Wedler, J. Woodward, S. Zhou, W. Zimmermann, D.F. Smith, J.M. Blackwell, K.D. Stuart, B. Barrell, *et al.*, (2005) The genome of the kinetoplastid parasite, *Leishmania major*. *Science* **309**: 436-442.
- Jackson, A.P., (2010) The evolution of amastin surface glycoproteins in trypanosomatid parasites. *Mol Biol Evol* **27**: 33-45.
- Jelinska, C., B. Petrovic-Stojanovska, W.J. Ingledew & M.F. White, (2010) Dimer-dimer stacking interactions are important for nucleic acid binding by the archaeal chromatin protein Alba. *Biochem J* **427**: 49-55.
- Kramer, S., (2012) Developmental regulation of gene expression in the absence of transcriptional control: the case of kinetoplastids. *Mol Biochem Parasitol* **181**: 61-72.
- Lahav, T., D. Sivam, H. Volpin, M. Ronen, P. Tsigankov, A. Green, N. Holland, M. Kuzyk, C. Borchers, D. Zilberstein & P.J. Myler, (2011) Multiple levels of gene regulation mediate differentiation of the intracellular pathogen *Leishmania*. *Faseb J* **25**: 515-525.
- Laurens, N., R.P. Driessen, I. Heller, D. Vorselen, M.C. Noom, F.J. Hol, M.F. White, R.T. Dame & G.J. Wuite, (2012) Alba shapes the archaeal genome using a delicate balance of bridging and stiffening the DNA. *Nat Commun* **3**: 1328.
- Laurentino, E.C., S. Taylor, G.R. Mair, E. Lasonder, R. Bartfai, H.G. Stunnenberg, H. Kroeze, J. Ramesar, B. Franke-Fayard, S.M. Khan, C.J. Janse & A.P. Waters, (2011) Experimentally controlled downregulation of the histone chaperone FACT in *Plasmodium berghei* reveals that it is critical to male gamete fertility. *Cell Microbiol* **13**: 1956-1974.
- Li, Z.H., J.G. De Gaudenzi, V.E. Alvarez, N. Mendiondo, H. Wang, J.C. Kissinger, A.C. Frasch & R. Docampo, (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.
- Mair, G.R., E. Lasonder, L.S. Garver, B.M. Franke-Fayard, C.K. Carret, J.C. Wiegant, R.W. Dirks, G. Dimopoulos, C.J. Janse & A.P. Waters, (2010) Universal features of post-transcriptional gene regulation are critical for *Plasmodium* zygote development. *PLoS Pathog* **6**: e1000767.
- Mani, J., A. Guttinger, B. Schimanski, M. Heller, A. Acosta-Serrano, P. Pescher, G. Spath & I. Roditi, (2011) Alba-domain proteins of *Trypanosoma brucei* are cytoplasmic RNA-binding proteins that interact with the translation machinery. *PLoS One* **6**: e22463.
- Mayho, M., K. Fenn, P. Craddy, S. Crosthwaite & K. Matthews, (2006) Post-transcriptional control of nuclear-encoded cytochrome oxidase subunits in *Trypanosoma brucei*: evidence for genome-wide conservation of life-cycle stage-specific regulatory elements. *Nucleic Acids Res* **34**: 5312-5324.
- McNicoll, F., J. Drummelsmith, M. Muller, E. Madore, N. Boilard, M. Ouellette & B. Papadopoulou, (2006) A combined proteomic and transcriptomic approach to the study of stage differentiation in *Leishmania infantum*. *Proteomics* **6**: 3567-3581.

- McNicoll, F., M. Muller, S. Cloutier, N. Boilard, A. Rochette, M. Dube & B. Papadopoulou, (2005) Distinct 3'-untranslated region elements regulate stage-specific mRNA accumulation and translation in *Leishmania*. *J Biol Chem* **280**: 35238-35246.
- Mishra, K.K., T.R. Holzer, L.L. Moore & J.H. LeBowitz, (2003) A negative regulatory element controls mRNA abundance of the *Leishmania mexicana* Paraflagellar rod gene PFR2. *Eukaryot Cell* **2**: 1009-1017.
- Myler, P.J., L. Audleman, T. deVos, G. Hixson, P. Kiser, C. Lemley, C. Magness, E. Rickel, E. Sisk, S. Sunkin, S. Swartzell, T. Westlake, P. Bastien, G. Fu, A. Ivens & K. Stuart, (1999) *Leishmania major* Friedlin chromosome 1 has an unusual distribution of protein-coding genes. *Proc Natl Acad Sci U S A* **96**: 2902-2906.
- Olguin-Lamas, A., E. Madec, A. Hovasse, E. Werkmeister, I. Callebaut, C. Slomianny, S. Delhay, T. Mouveaux, C. Schaeffer-Reiss, A. Van Dorsselaer & S. Tomavo, (2011) A novel *Toxoplasma gondii* nuclear factor TgNF3 is a dynamic chromatin-associated component, modulator of nucleolar architecture and parasite virulence. *PLoS Pathog* **7**: e1001328.
- Padmanabhan, P.K., C. Dumas, M. Samant, A. Rochette, M.J. Simard & B. Papadopoulou, (2012) Novel Features of a PIWI-Like Protein Homolog in the Parasitic Protozoan *Leishmania*. *PLoS One* **7**: e52612.
- Peritz, T., F. Zeng, T.J. Kannanayakal, K. Kilk, E. Eiriksdottir, U. Langel & J. Eberwine, (2006) Immunoprecipitation of mRNA-protein complexes. *Nat Protoc* **1**: 577-580.
- Puri, V., A. Goyal, R. Sankaranarayanan, A.J. Enright & T. Vaidya, (2011) Evolutionary and functional insights into *Leishmania* META1: evidence for lateral gene transfer and a role for META1 in secretion. *BMC Evol Biol* **11**: 334.
- Quijada, L., C. Guerra-Giraldez, M. Drozd, C. Hartmann, H. Irmer, C. Ben-Dov, M. Cristodero, M. Ding & C. Clayton, (2002) Expression of the human RNA-binding protein HuR in *Trypanosoma brucei* increases the abundance of mRNAs containing AU-rich regulatory elements. *Nucleic Acids Res* **30**: 4414-4424.
- Rochette, A., F. McNicoll, J. Girard, M. Breton, E. Leblanc, M.G. Bergeron & B. Papadopoulou, (2005) Characterization and developmental gene regulation of a large gene family encoding amastin surface proteins in *Leishmania* spp. *Mol Biochem Parasitol* **140**: 205-220.
- Rochette, A., F. Raymond, J. Corbeil, M. Ouellette & B. Papadopoulou, (2009) Whole-genome comparative RNA expression profiling of axenic and intracellular amastigote forms of *Leishmania infantum*. *Mol Biochem Parasitol* **165**: 32-47.
- Rodrigues, D.C., R. Silva, E. Rondinelli & T.P. Urmenyi, (2010) *Trypanosoma cruzi*: modulation of HSP70 mRNA stability by untranslated regions during heat shock. *Exp Parasitol* **126**: 245-253.
- Sandman, K. & J.N. Reeve, (2005) Archaeal chromatin proteins: different structures but common function? *Curr Opin Microbiol* **8**: 656-661.
- Schoenberg, D.R. & L.E. Maquat, (2012) Regulation of cytoplasmic mRNA decay. *Nat Rev Genet* **13**: 246-259.
- Sereno, D. & J.L. Lemesre, (1997) Axenically cultured amastigote forms as an *in vitro* model for investigation of antileishmanial agents. *Antimicrob Agents Chemother* **41**: 972-976.

- Sharma, P., S. Gurumurthy, R. Duncan, H.L. Nakhasi & P. Salotra, (2010) Comparative *in vivo* expression of amastigote up regulated *Leishmania* genes in three different forms of Leishmaniasis. *Parasitol Int* **59**: 262-264.
- Smith, M., F. Bringaud & B. Papadopoulou, (2009) Organization and evolution of two SIDER retroposon subfamilies and their impact on the *Leishmania* genome. *BMC Genomics* **10**: 240.
- Srisawat, C. & D.R. Engelke, (2002) RNA affinity tags for purification of RNAs and ribonucleoprotein complexes. *Methods* **26**: 156-161.
- Subota, I., B. Rotureau, T. Blisnick, S. Ngwabyt, M. Durand-Dubief, M. Engstler & P. Bastin, (2011) ALBA proteins are stage regulated during trypanosome development in the tsetse fly and participate in differentiation. *Mol Biol Cell* **22**: 4205-4219.
- Teixeira, S.M., D.G. Russell, L.V. Kirchhoff & J.E. Donelson, (1994) A differentially expressed gene family encoding "amastin," a surface protein of *Trypanosoma cruzi* amastigotes. *J Biol Chem* **269**: 20509-20516.
- Vasconcelos, E.J., M.C. Terraio, J.C. Ruiz, R.Z. Vencio & A.K. Cruz, (2012) In silico identification of conserved intercoding sequences in *Leishmania* genomes: unraveling putative *cis*-regulatory elements. *Mol Biochem Parasitol* **183**: 140-150.
- Walker, S.C., F.H. Scott, C. Srisawat & D.R. Engelke, (2008) RNA affinity tags for the rapid purification and investigation of RNAs and RNA-protein complexes. *Methods Mol Biol* **488**: 23-40.
- Welting, T.J., F.M. Peters, S.M. Hensen, N.L. van Doorn, B.J. Kikkert, J.M. Raats, W.J. van Venrooij & G.J. Pruijn, (2007) Heterodimerization regulates RNase MRP/RNase P association, localization, and expression of Rpp20 and Rpp25. *Rna* **13**: 65-75.
- Wu, Y., Y. El Fakhry, D. Sereno, S. Tamar & B. Papadopoulou, (2000) A new developmentally regulated gene family in *Leishmania* amastigotes encoding a homolog of amastin surface proteins. *Mol Biochem Parasitol* **110**: 345-357.
- Zhang, W.W., S. Mendez, A. Ghosh, P. Myler, A. Ivens, J. Clos, D.L. Sacks & G. Matlashewski, (2003) Comparison of the A2 gene locus in *Leishmania donovani* and *Leishmania major* and its control over cutaneous infection. *J Biol Chem* **278**: 35508-35515.
- Zilberstein, D. & M. Shapira, (1994) The role of pH and temperature in the development of *Leishmania* parasites. *Annu Rev Microbiol* **48**: 449-470.

## Figure Legends

### **Fig. 1. RNA affinity chromatography identified *LiAlba20* as an interacting protein factor with the 300 nt URE-containing region of the amastin (LinJ.34.1010) 3'UTR.**

A. Schematic representation of the LinJ.34.1010 3'UTR. The 100 nt URE starts at 597 nt after the ORF codon stop. PCR amplification of the 300 nt region including the URE and 5' and 3' flanking sequences (300nt URE-D8) was carried out to add a T7 promoter sequence and a D8-tag. B. RNA affinity chromatography using a Sephadex column and D8-tagged *in vitro*-transcribed RNA. Chromatography was performed as described in the Experimental Procedures using the *in vitro* transcribed 300 nt URE-D8 or only the Sephadex matrix as a negative control. After elution, protein samples were resolved on a 12% PAGE and stained with Sypro. The bands of interest were cut and sent for MS/MS peptide identification. The results of MS/MS analysis are shown in C. D. Western blot using the elution material from RNA affinity chromatography to confirm the interaction between the amastin 300 nt URE-D8 RNA and *LiAlba20* protein. RNA affinity chromatography was reproduced and transferred on a membrane after resolution on a 12% PAGE. A *T. brucei* Alba3 specific antibody (*TbAlba2040/2*) produced in mouse was used here. This antibody specifically recognizes *LiAlba20*.

### **Fig. 2. RNA co-immunoprecipitation to assess HA-*LiAlba20*-amastin mRNA specific interaction.**

A. Confirmation of episomal expression of HA-*LiAlba20* protein in *L. infantum* by western blot analysis (promastigote cell extracts were used here) using an anti-HA antibody. B. Reverse-transcribed PCR data from RNA co-immunoprecipitation experiments (RIP RT-PCR). RIP was performed using promastigote cell extracts of wild type (WT) or HA-*LiAlba20* transfected cells and an anti-HA antibody. cDNA synthesis was performed using random hexamer primers in the absence (-) or the presence (+) of reverse transcriptase (RT). PCR amplification of ~100 bp within the ORF of *amastin* (LinJ.34.1010), *GAPDH* (LinJ.36.2480) and *BTI* (LinJ.35.5120). C. RIP RT-PCR using extracts from *L. infantum* axenic amastigotes. P: Promastigotes; A: Axenic amastigotes.

**Fig. 3. Regulation of the endogenous Alba20 in *L. infantum*.**

A. Northern blot hybridization from total RNA of wild type promastigotes (P) and axenic amastigotes (A). *LiAlba20* and  $\alpha$ -tubulin ORFs were radiolabeled with  $^{32}\text{P}$ -dCTP and used as probes for hybridization. B. Western blot from promastigote, axenic amastigote and differentiating amastigote extracts (promastigote to amastigote differentiation). Endogenous *LiAlba20* was detected using the *TbAlba2040/2* antibody.

**Fig. 4. Generation of a diploid *LiAlba20* knockout mutant in *L. infantum* by gene replacement.**

A. Schematic representation of the *LiAlba20* locus (LinJ.34.2410) genomic organization showing the positions of NdeI and HindIII restriction sites used for southern blot confirmation. B. Southern blot hybridization to confirm genomic integration of the hygromycin phosphotransferase (HYG) and puromycin (PUR) expression targeting cassettes into the *LiAlba20* locus. All probes used here correspond to the ORF radiolabelled with  $^{32}\text{P}$ -dCTP.

**Fig. 5. Effect of *LiAlba20* depletion on *amastin* mRNA stability.**

A. Northern blot hybridization was performed with total RNA extracted from promastigotes (P) and axenic amastigotes (A) of WT, *LiAlba20*<sup>-/-</sup> and the HA-*LiAlba20* add-back strains. Data normalization was performed using the 18S rRNA as a control. Data shown here are representative of 3 independent experiments with similar results. B. Comparison of *amastin* mRNA half-life in axenic amastigotes between WT and *LiAlba20*<sup>-/-</sup> strains. mRNA half-lives were measured following inhibition of transcription with actinomycin D (added at  $t_{0\text{min}}$ ) and *trans*-splicing with sinefungin (added at  $t_{.15\text{min}}$ ). C. Graphical representation of *amastin* mRNA half-lives in *L. infantum* amastigotes from results shown in B. Data were normalized using the 18S rRNA signal and are representative of three independent experiments.

**Fig. 6. Summary of DNA microarray analysis data comparing wild type cells with *LiAlba20*<sup>-/-</sup> in both life stages of *L. infantum*.**

A. Pie graph of the stage specificity of the transcripts identified as upregulated or downregulated in *LiAlba20*<sup>-/-</sup> as revealed by microarray assay. Our microarray data were compared to the microarray data of wild type *L. infantum* promastigotes versus axenic amastigotes previously published by Rochette *et al.*, 2009. For example, among the 96 downregulated genes in the strain *LiAlba20*<sup>-/-</sup> during the amastigote life stage, 66 are constitutively expressed, whereas 30 are amastigote specific in wild type cells. B. Gene ontology annotation of identified up- or down-regulated genes in *LiAlba20*<sup>-/-</sup>.

**Fig. 7. Depletion of *LiAlba13* does not affect *amastin* mRNA stability.**

A. RIP RT-PCR experiment using wild type cells (WT) and parasites episomally expressing *LiAlba13*-HA as shown in Figure 2B. B. Schematic representation of the *LiAlba13* locus (LinJ.13.0270) genomic organization showing the positions of PstI restriction sites used for southern blot confirmation. C. Southern blot hybridization to confirm genomic integration of the HYG and PUR expression cassettes into the *LiAlba13* locus and minimal dilution cloning. Depletion of *LiAlba13* was also confirmed by northern blot hybridization (data not shown). D. Northern blot hybridization to assess the role of *LiAlba13* protein on amastin mRNA fate. Northern blot was performed with total RNA extracted from WT, *LiAlba13*<sup>-/-</sup> and the *LiAlba13*-HA add-back strains grown as promastigotes (P) and axenic amastigotes (A). Results shown here are representative of three independent experiments.



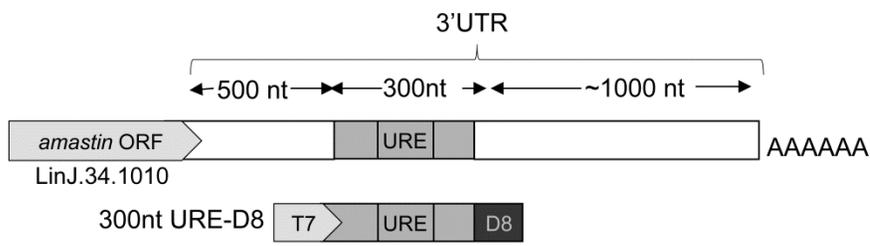
## Tables

**Table 1.** Selected amastigote-specific genes downregulated in the *L. infantum* Alba20<sup>-/-</sup> mutant.

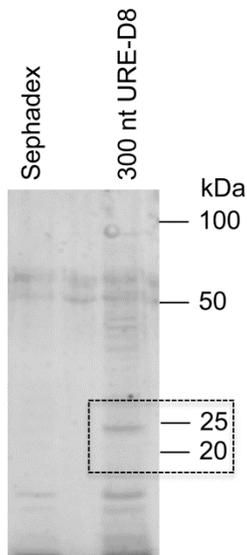
<i>Gene ID</i>	<i>Description</i>	<i>Fold difference (LiAlba20<sup>-/-</sup>/WT)</i>
LinJ.08.0680	<b>δ-amastin</b>	0,38
LinJ.08.0690	<b>δ-amastin</b>	0,38
LinJ.08.0720	<b>δ-amastin</b>	0,38
LinJ.08.1320	<b>δ-amastin</b>	0,38
LinJ.34.4370	<b>δ-amastin</b>	0,53
LinJ.34.1730	<b>δ-amastin</b>	0,58
LinJ.34.1030	<b>δ-amastin</b>	0,59
LinJ.34.1010	<b>δ-amastin</b>	0,59
LinJ.08.0790	<b>δ-amastin</b>	0,59
LinJ.10.0250	<b>hypothetical protein, conserved (<i>Amastin-like domain</i>)</b>	0,60
LinJ.17.0990	META1	0,42
LinJ.22.0670	A2 protein	0,52
LinJ.28.1070	P27 protein	0,53
LinJ.23.1220	hydrophilic acylated surface protein b (HASPb)	0,56
LinJ.10.0360	folate/biopterin transporter, putative	0,41
LinJ.31.0370	amino acid transporter aATP11, putative	0,50
LinJ.03.0480	Phosphate-repressible phosphate permease-like protein	0,54
LinJ.27.1680	casein kinase I-like protein	0,58
LinJ.27.0310	methylmalonyl-coenzyme a mutase, putative	0,53
LinJ.14.0730	fatty acid elongase, putative	0,37
LinJ.23.0300	argininosuccinate synthase, putative	0,44
LinJ.34.3430	hypothetical protein ( <i>pseudogene</i> )	0,27
LinJ.08.0650	hypothetical protein, unknown function	0,47
LinJ.30.2210	hypothetical protein, conserved	0,49
LinJ.34.1500	hypothetical protein, conserved	0,50
LinJ.33.1440	hypothetical protein, conserved	0,51
LinJ.03.0700	hypothetical protein, conserved	0,54
LinJ.16.1460	hypothetical protein, conserved	0,55
LinJ.09.0290	hypothetical protein, conserved	0,56
LinJ.11.1070	hypothetical protein	0,57
LinJ.08.0800	hypothetical protein, unknown function	0,59

# Figures

**A**



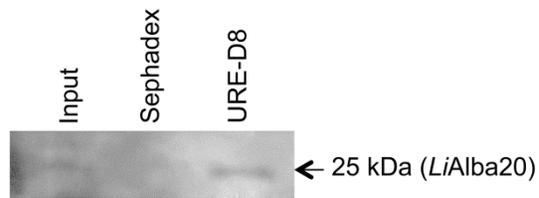
**B**



**C**

Gene ID	Description	M. W.	Peptides
LinJ.34.2410	Hypothetical protein, conserved (Alba20)	22 kDa	16
LinJ.36.4100	S-adenosylhomocysteine hydrolase	48 kDa	3
LinJ.17.0090	Elongation factor 1 alpha	49 kDa	3
LinJ.22.0430	Kinesin, putative	97 kDa	2
LinJ.25.0940	CYPA cyclophilin a	19 kDa	2

**D**



**Figure 1.**

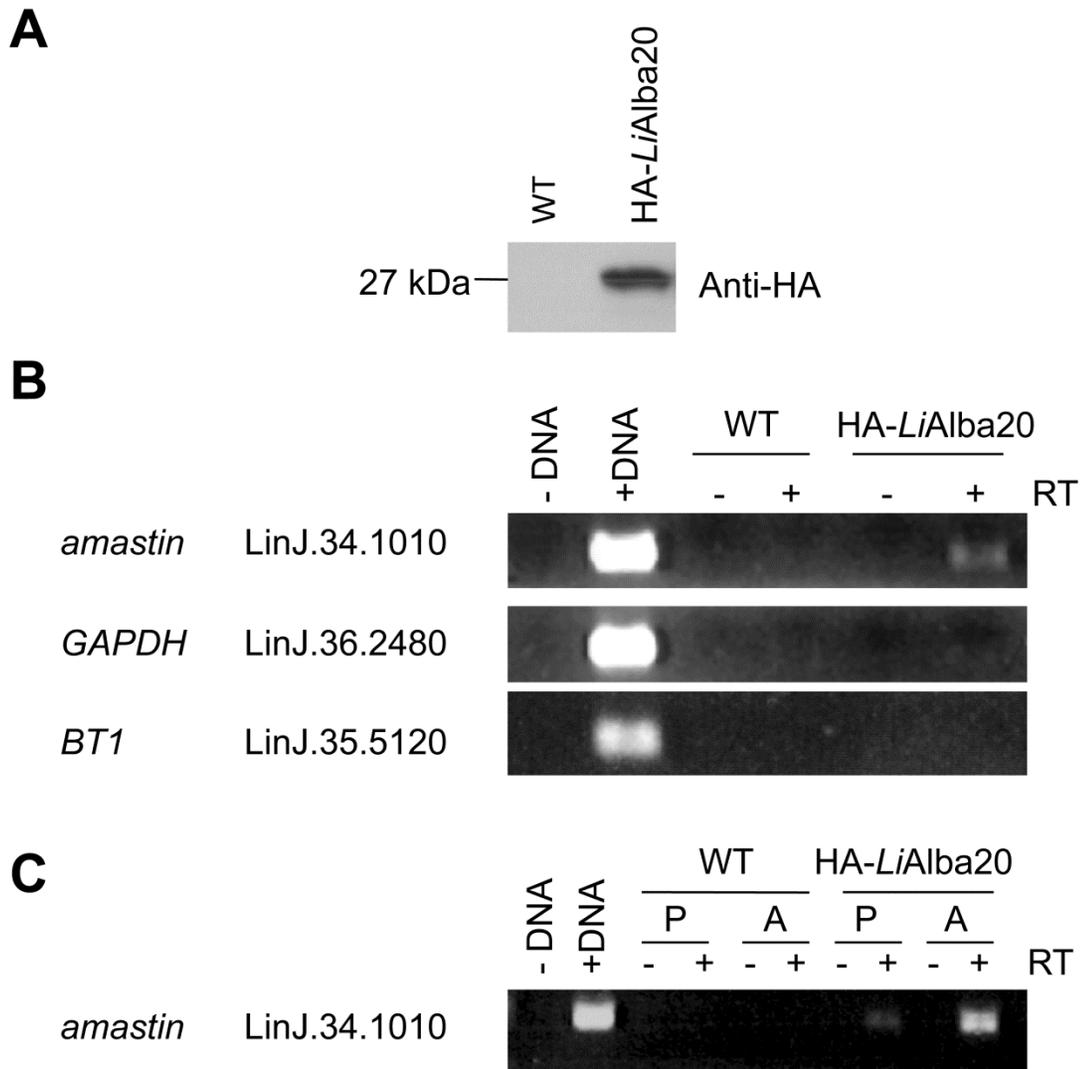
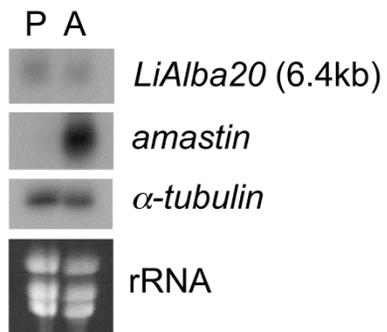
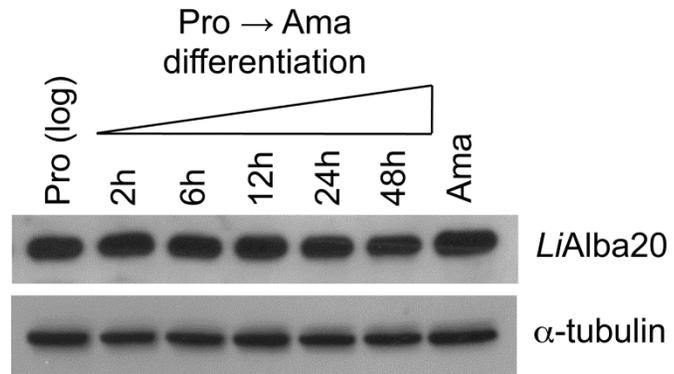
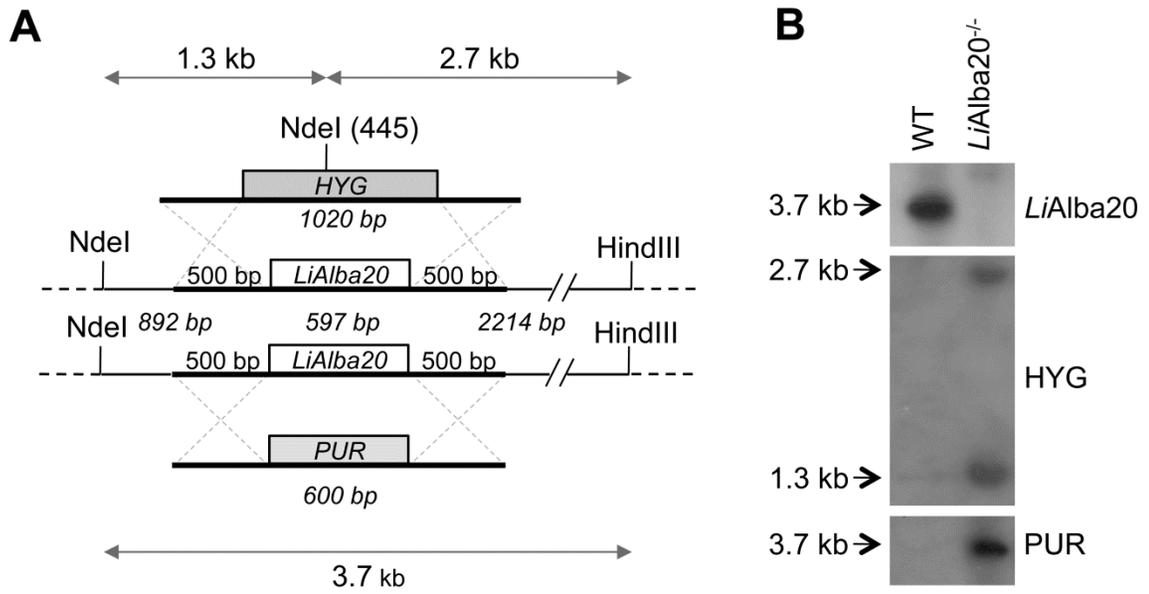
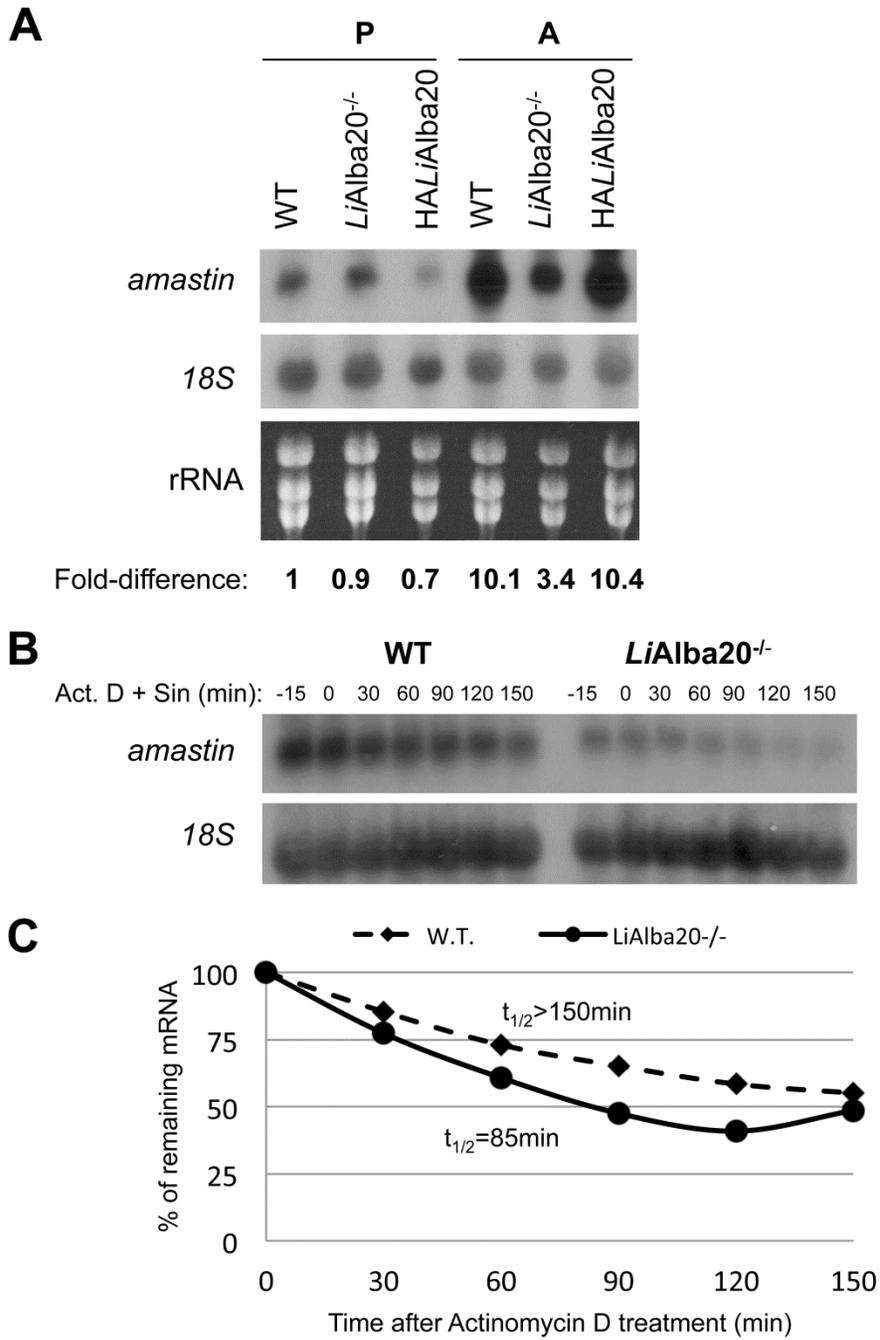


Figure 2.

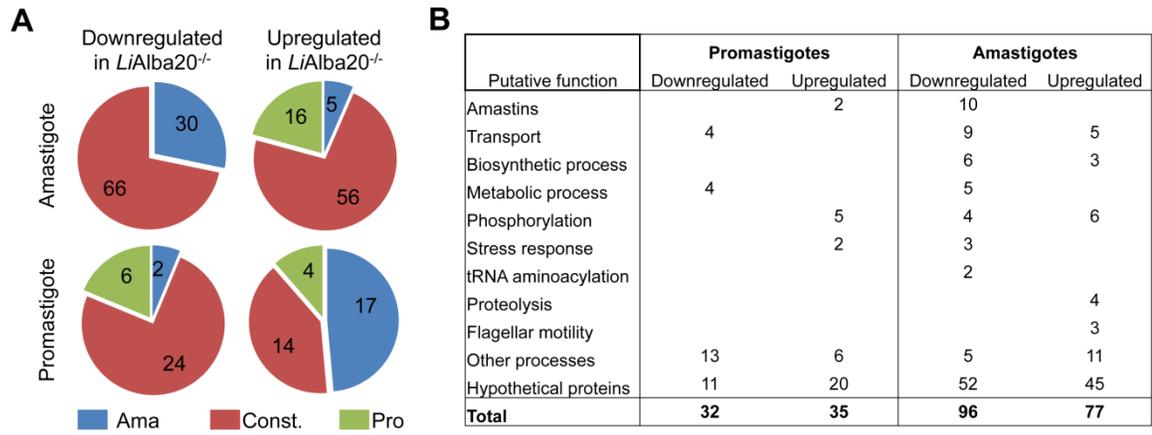
**A****B****Figure 3.**



**Figure 4.**



**Figure 5.**



**Figure 6.**

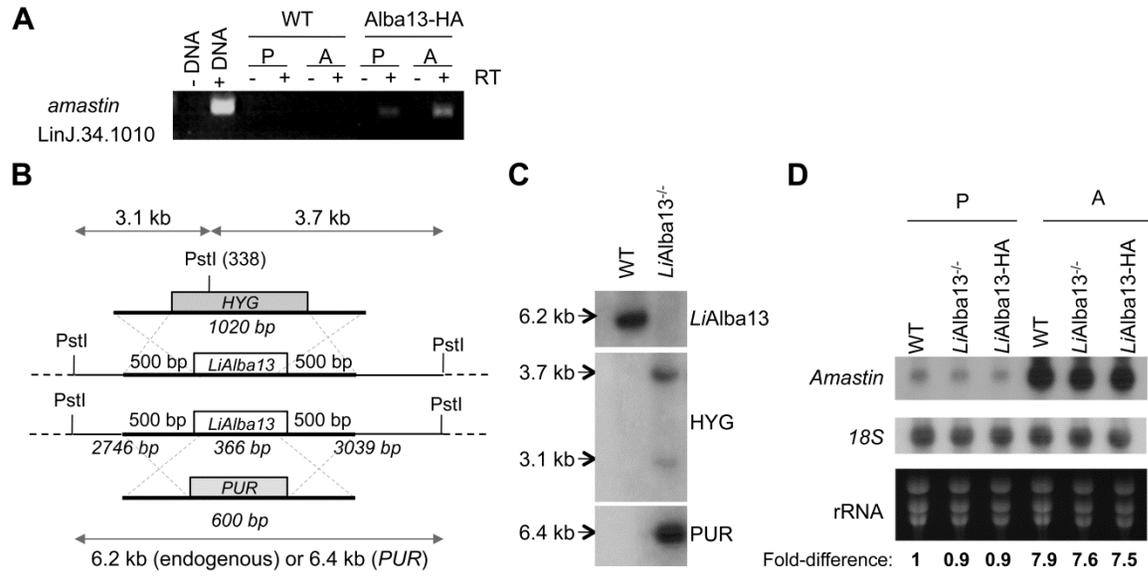


Figure 7.

## Supporting Information

**Table S1. List of primers used in this study**

### D8-tagged amastin 3'UTR RNAs

Name	Sequence	Gene ID (distance from codon stop)	PCR product
T7-300nt URE F	TAATACGACTCACTATAGGGAAACGGACGCCCTCGCACGG	LinJ.34.1010 & 1030 (493)	353bp
300nt URE D8 R	TCCGCAACCGTATCAAAACGTAAATTACTCGGAGGGGGCGGGGAAGACGA		

### Primers for generating targeting cassettes for *Alba* gene knockouts

Name	Sequence	Gene ID	PCR product
5'UTR LiAlba20 F	GCACGTCCTTCCTGTCCCCCATGC	LinJ.34.2410	525bp (Hygro) 521bp (Puro)
5'UTR LiAlba20 R HYG	TCCGGTGAGTTCAGGCTTTTTCATGCTGCTTAGCGGGTTTTCC		
5'UTR LiAlba20 R PUR	GTGGGCTTGTA CT CGGTCATGCTGCTTAGCGGGTTTTCC		
3'UTR LiAlba20 F HYG	CACTCGTCCGAGGGCAAAGGAATAGAAACGCGGGGAAAGTAAATAGCG		
3'UTR LiAlba20 F PUR	GCATGACCCGCAAGCCCGGTGCCTGAAAACGCGGGGAAAGTAAATAGCG		
3'UTR LiAlba20 R	AAGACACACATTGATACGGAAGC		
5'UTR LiAlba13 F	CCATTTATCGAAGGGGTTC	LinJ.13.0270	520bp
5'UTR LiAlba13 R PUR	GTGGGCTTGTA CT CGGTCATCTTGGTACAAAGGAAAAGGT		
5'UTR LiAlba13 R HYG	GTGAGTTCAGGCTTTTTCATCTTGGTACAAAGGAAAAGGT		
3'UTR LiAlba13 F PUR	CCCGCAAGCCCGGTGCCTGACTCAGCTGCCTGGGTGGCGC		
3'UTR LiAlba13 F HYG	GTCCGAGGGCAAAGGAATAGCTCAGCTGCCTGGGTGGCGC		
3'UTR LiAlba13 R	GTGTAGCATTTCGCTTTTCA		
HYG F	ATGAAAAAGCCTGAACTCAC	Hygromycin phosphotransferase	1026bp
HYG R	CTATTCCTTTGCCCTCGGAC	Puromycin	600bp
PUR F	ATGACCGAGTACAAGCCACG		
PUR R	TCAGGCACCGGGCTTGC		

### Cloning of HA-tagged Alba proteins into pSP $\alpha$ .NEO $\alpha$ expression vector

Name	Sequence	Gene ID	PCR product
XbaI ORF LiAlba20 HA F	TCTAGAAATGTACCCTTACGACGTGCCAGACTACGCTCCGCCACGTCCGTCTCG	LinJ.34.2410	636bp
NdeI ORF LiAlba20 R	CATATGCTAGTTCTCGGGTCATCGCGGC		
XbaI ORF LiAlba13 F	TCTAGAAATGTCAGCTGAAAACCTGCGG	LinJ.13.0270	405bp
NdeI ORF LiAlba13 HA R	CATATGCTAAGCGTAGTCTGGCACGTCGTAAGGGTACTCATCATCGGCCTCATCC		

**Real Time PCR (For RNA immunoprecipitation; primers were selected within ORFs)**

<b>Name</b>	<b>Sequence</b>	<b>Gene ID</b>	<b>PCR product</b>
Amastin F	ACATAGCGAGCACTGTTTTTC	LinJ.34.1010 & 1030	134bp
Amastin R	CCATAAGGTGATGCAGAAAT		
GAPDH F	GATCCAGGAGATAGACAAG	LinJ.30.3000	106bp
GAPDH R	GTCGTTGATGAAGTCAGAG		
BT1 F	GACAAGCTGTTGCTTATTAT	LinJ.35.5160	127bp
BT1 R	GCACAATCTGGTAGAATAGA		

**Table S2** : DNA microarray data: *L. infantum* transcripts differentially expressed between wild type and *LiAlba20*<sup>-/-</sup> promastigotes.

<i>Fold difference</i>	<i>Coeff</i>	<i>p.value</i>	<i>Gene ID</i>	<i>Putative biological function</i>	<i>Stage-specific expression<sup>a</sup></i>
<b>Downregulated genes in <i>LiAlba20</i><sup>-/-</sup></b>					
<b>0,04</b>	-4,61	1,75E-05	<b>LinJ.34.2410</b>	hypothetical protein, conserved (Alba20)	
<b>0,31</b>	-1,70	3,37E-04	<b>LinJ.32.2800</b>	L-Lysine transport protein (LiAAP7)	
<b>0,33</b>	-1,60	2,96E-04	<b>LinJ.12.0510</b>	hypothetical protein, unknown function (Spectrin repeat-like)	
<b>0,38</b>	-1,38	3,37E-04	<b>LinJ.36.2720</b>	membrane-bound acid phosphatase 2, putative	
<b>0,40</b>	-1,34	3,37E-04	<b>LinJ.30.2210</b>	hypothetical protein, conserved	<b>Ama</b>
<b>0,43</b>	-1,23	3,37E-04	<b>LinJ.10.0280</b>	hypothetical protein, conserved	
<b>0,46</b>	-1,11	2,08E-03	<b>LinJ.10.0360</b>	folate/biopterin transporter, putative	<b>Ama</b>
<b>0,48</b>	-1,06	8,37E-04	<b>LinJ.20.1670</b>	hypothetical protein, conserved	
<b>0,50</b>	-1,01	8,37E-04	<b>LinJ.11.0120</b>	hypothetical protein, conserved	
<b>0,51</b>	-0,97	8,37E-04	<b>LinJ.27.2490</b>	calpain-like cysteine peptidase	
<b>0,51</b>	-0,96	1,64E-03	<b>LinJ.24.1130</b>	formin, putative	
<b>0,53</b>	-0,92	8,37E-04	<b>LinJ.28.2250</b>	glycerophosphoryl diester phosphodiesterase, putative	
<b>0,54</b>	-0,89	1,26E-03	<b>LinJ.07.0280</b>	tRNA-methyl transferase, putative	
<b>0,54</b>	-0,88	1,77E-03	<b>LinJ.04.0300</b>	beta-fructofuranosidase, putative	<b>Pro</b>
<b>0,55</b>	-0,87	1,92E-03	<b>LinJ.32.1760</b>	hypothetical protein, conserved	
<b>0,55</b>	-0,86	2,57E-03	<b>LinJ.15.1260</b>	nucleoside transporter 1, putative	<b>Pro</b>
<b>0,56</b>	-0,85	1,77E-03	<b>LinJ.19.0520</b>	hypothetical protein, conserved	<b>Pro</b>
<b>0,56</b>	-0,83	8,37E-04	<b>LinJ.28.2200</b>	DNA-directed RNA polymerase-like protein	
<b>0,57</b>	-0,81	1,98E-03	<b>LinJ.11.0620</b>	hypothetical protein, conserved	
<b>0,57</b>	-0,81	8,48E-04	<b>LinJ.02.0140</b>	phosphoglycan beta 1,3 galactosyltransferase (SCGR6)	
<b>0,57</b>	-0,80	1,00E-03	<b>LinJ.26.0120</b>	adenine phosphoribosyltransferase (APRT)	
<b>0,57</b>	-0,80	1,24E-03	<b>LinJ.28.2220</b>	mitochondrial DEAD box protein, putative	
<b>0,58</b>	-0,80	3,36E-03	<b>LinJ.23.1830</b>	hypothetical protein, unknown function	
<b>0,58</b>	-0,79	1,50E-03	<b>LinJ.29.0260</b>	thymine-7-hydroxylase, putative	
<b>0,58</b>	-0,78	8,48E-04	<b>LinJ.32.0680</b>	hypothetical protein, conserved (LigT-like)	
<b>0,58</b>	-0,78	3,95E-03	<b>LinJ.31.1240</b>	vacuolar-type proton translocating pyrophosphatase 1, putative	
<b>0,59</b>	-0,77	1,82E-03	<b>LinJ.08.0180</b>	hypothetical protein, unknown function (Leucine-rich repeat)	<b>Pro</b>
<b>0,59</b>	-0,77	1,26E-03	<b>LinJ.24.0650</b>	MCAK-like kinesin, putative	
<b>0,59</b>	-0,76	1,20E-03	<b>LinJ.23.0880</b>	acetyl-CoA synthetase, putative	
<b>0,59</b>	-0,76	1,77E-03	<b>LinJ.28.1520</b>	hypothetical protein (CCCH zinc finger protein)	
<b>0,59</b>	-0,76	2,13E-03	<b>LinJ.07.1140</b>	nucleolar RNA-binding protein, putative	
<b>0,59</b>	-0,76	1,30E-03	<b>LinJ.29.1890</b>	paraflagellar rod protein 1D, putative	<b>Pro</b>
<b>0,60</b>	-0,75	4,26E-03	<b>LinJ.14.0700</b>	fatty acid elongase, putative	<b>Pro</b>
<b>Upregulated genes in <i>LiAlba20</i><sup>-/-</sup></b>					

2,60	1,38	1,10E-02	LinJ.29.1850	histone H2A, putative	
2,57	1,36	3,37E-04	LinJ.30.1520	p1/s1 nuclease	Ama
2,54	1,34	3,37E-04	LinJ.30.2340	hypothetical protein, unknown function	Ama
2,48	1,31	1,54E-03	LinJ.17.0020	hypothetical protein, conserved	Ama
2,44	1,28	3,37E-04	LinJ.34.1380	hypothetical protein, conserved	
2,14	1,10	6,80E-04	LinJ.06.1360	hypothetical protein, conserved (MFS general substrate transporter)	Ama
2,09	1,07	1,20E-03	LinJ.36.4050	similar to leishmania major. l411.4-like protein	
2,08	1,06	1,69E-03	LinJ.04.1230	protein kinase, putative,casein kinase I, putative	Ama
2,07	1,05	6,52E-04	LinJ.22.0670	A2 protein	Ama
2,00	1,00	1,20E-03	LinJ.35.5210	hypothetical protein, unknown function	Ama
2,00	1,00	7,60E-04	LinJ.27.1680	casein kinase I-like protein	Ama
1,99	0,99	1,77E-03	LinJ.20.1370	kinase-like protein	
1,98	0,98	8,37E-04	LinJ.23.1230	small hydrophilic endoplasmic reticulum-associated protein (sherp) (SHERP)	Ama
1,93	0,95	7,78E-02	LinJ.35.2270	kinetoplastid membrane protein-11 (KMP11-2)	
1,88	0,91	1,27E-03	LinJ.20.1420	pumilio protein 9, putative (PUF9B)	
1,88	0,91	8,37E-04	LinJ.36.3790	hypothetical protein, unknown function (MFS general substrate transporter)	
1,87	0,90	4,06E-03	LinJ.23.1220	hydrophilic acylated surface protein b (HASPB)	Ama
1,86	0,89	8,37E-04	LinJ.23.1200	hydrophilic acylated surface protein a (HASPA1)	Ama
1,85	0,89	1,34E-01	LinJ.28.2950	heat-shock protein hsp70, putative	
1,81	0,86	8,37E-04	LinJ.13.0190	hypothetical protein, unknown function	Ama
1,81	0,85	8,37E-04	LinJ.30.3690	CYC2-like protein, putative (CYC7)	
1,80	0,85	1,00E-03	LinJ.34.1700	amastin-like surface protein, putative	Ama
1,80	0,85	1,69E-03	LinJ.36.7360	hypothetical protein, conserved (PSP1 C-terminal conserved region)	Pro
1,77	0,82	1,92E-03	LinJ.34.1730	amastin-like surface protein, putative	Ama
1,76	0,82	1,77E-03	LinJ.35.2200	adenine aminohydrolase (AAH)	Pro
1,75	0,81	8,48E-04	LinJ.16.0500	hypothetical protein, conserved (Prokaryotic membrane lipoprotein lipid attachment site profile)	
1,75	0,81	8,48E-04	LinJ.19.1290	hypothetical protein, unknown function (Protein kinase domain)	Ama
1,75	0,80	2,71E-03	LinJ.35.4820	hypothetical protein, conserved	Pro
1,74	0,80	8,48E-04	LinJ.03.0700	hypothetical protein, conserved	Ama
1,72	0,78	9,00E-04	LinJ.29.2240	hypothetical protein, conserved (Leucine-rich repeat)	
1,70	0,76	4,36E-03	LinJ.09.0630	surface antigen-like protein	Ama
1,69	0,76	1,27E-03	LinJ.20.0030	histone-lysine N-methyltransferase, putative (DOT1)	Pro
1,69	0,76	1,69E-03	LinJ.35.5040	hypothetical protein, conserved	
1,69	0,76	1,00E-03	LinJ.19.1670	hypothetical protein, unknown function (Leucine-rich repeat; RING/U-box)	
1,69	0,76	2,47E-03	LinJ.23.1240	hydrophilic acylated surface protein a (HASPA2)	Ama

<sup>a</sup>Ama (amastigotes) and Pro (promastigotes) represent differentially expressed genes in both *L. infantum* life stages according to previously published DNA microarray analysis by Rochette A, *et al.* (2009). The same statistical analysis was applied in both studies.

**Table S3** : DNA microarray data: *L. infantum* transcripts differentially expressed between wild type and *LiAlba20*<sup>-/-</sup> amastigotes.

<i>Fold difference</i>	<i>Coeff</i>	<i>p.value</i>	<i>Gene ID</i>	<i>Putative biological function</i>	<i>Stage-specific expression<sup>a</sup></i>
<b>Downregulated genes in <i>LiAlba20</i><sup>-/-</sup></b>					
<b>0,06</b>	-4,18	7,63E-04	<b>LinJ.34.2410</b>	hypothetical protein, conserved (Alba20)	
<b>0,25</b>	-1,99	2,76E-03	<b>LinJ.28.2950</b>	heat-shock protein hsp70, putative	
<b>0,27</b>	-1,87	3,11E-03	<b>LinJ.34.3430</b>	hypothetical protein (pseudogene)	<b>Ama</b>
<b>0,35</b>	-1,51	7,63E-04	<b>LinJ.28.3000</b>	heat-shock protein hsp70, putative	
<b>0,36</b>	-1,48	1,38E-04	<b>LinJ.32.2800</b>	amino acid transporter, putative, amino acid permease, putative (AAT16)	
<b>0,36</b>	-1,48	1,60E-04	<b>LinJ.12.0510</b>	hypothetical protein, unknown function (Spectrin repeat)	
<b>0,37</b>	-1,45	1,38E-04	<b>LinJ.14.0730</b>	fatty acid elongase, putative	<b>Ama</b>
<b>0,37</b>	-1,43	1,63E-03	<b>LinJ.09.0930</b>	histone h1-like protein	
<b>0,38</b>	-1,39	1,13E-03	<b>LinJ.08.1320</b>	amastin-like protein	<b>Ama</b>
<b>0,39</b>	-1,36	5,80E-04	<b>LinJ.17.1110</b>	hydrolase, alpha/beta fold family-like protein	
<b>0,41</b>	-1,30	1,51E-03	<b>LinJ.10.0360</b>	folate/biopterin transporter, putative	<b>Ama</b>
<b>0,42</b>	-1,25	2,36E-04	<b>LinJ.17.0990</b>	hypothetical protein, conserved (META1)	<b>Ama</b>
<b>0,44</b>	-1,19	8,41E-04	<b>LinJ.23.0300</b>	argininosuccinate synthase, putative	<b>Ama</b>
<b>0,45</b>	-1,16	3,08E-03	<b>LinJ.22.0930</b>	dynein heavy chain, cytosolic, putative	
<b>0,46</b>	-1,13	8,05E-04	<b>LinJ.31.0750</b>	hypothetical protein, conserved (MFS general substrate transporter)	
<b>0,46</b>	-1,13	1,80E-03	<b>LinJ.28.2870</b>	protein phosphatase 4 catalytic subunit, putative	
<b>0,46</b>	-1,12	8,41E-04	<b>LinJ.34.2420</b>	adaptor complex subunit medium chain 3, putative	
<b>0,46</b>	-1,11	1,17E-03	<b>LinJ.34.3960</b>	hypothetical protein, conserved	
<b>0,47</b>	-1,10	1,96E-03	<b>LinJ.08.0650</b>	hypothetical protein, unknown function	<b>Ama</b>
<b>0,47</b>	-1,10	2,24E-03	<b>LinJ.30.0650</b>	histidyl-tRNA synthetase, putative	
<b>0,47</b>	-1,10	5,94E-03	<b>LinJ.08.0160</b>	hypothetical protein, conserved (GPI-GlcNAc transferase complex, PIG-H component)	
<b>0,47</b>	-1,09	8,41E-04	<b>LinJ.09.0370</b>	hypothetical protein, conserved	
<b>0,47</b>	-1,09	3,67E-03	<b>LinJ.10.0260</b>	hypothetical protein, conserved	
<b>0,48</b>	-1,07	4,02E-03	<b>LinJ.16.0060</b>	hypothetical protein, conserved (P-loop containing nucleoside triphosphate hydrolases)	
<b>0,49</b>	-1,03	7,64E-04	<b>LinJ.30.2210</b>	hypothetical protein, conserved (CCCH zinc finger)	<b>Ama</b>
<b>0,49</b>	-1,02	6,11E-03	<b>LinJ.26.2440</b>	hypothetical protein, conserved	
<b>0,50</b>	-1,01	1,11E-03	<b>LinJ.21.0860</b>	hypothetical protein, conserved	
<b>0,50</b>	-1,01	2,82E-03	<b>LinJ.34.1500</b>	hypothetical protein, conserved	<b>Ama</b>
<b>0,50</b>	-1,01	7,31E-04	<b>LinJ.36.1680</b>	universal minicircle sequence binding protein, putative (UMSBP1)	
<b>0,50</b>	-1,00	6,63E-04	<b>LinJ.14.0710</b>	fatty acid elongase, putative	
<b>0,50</b>	-1,00	4,16E-03	<b>LinJ.30.1650</b>	hypothetical protein, conserved	
<b>0,50</b>	-0,99	7,63E-04	<b>LinJ.31.0370</b>	amino acid transporter aATP11, putative (AAT1.3)	<b>Ama</b>
<b>0,51</b>	-0,98	2,76E-03	<b>LinJ.19.1250</b>	hypothetical protein, conserved	

0,51	-0,97	3,00E-03	LinJ.19.0430	hypothetical protein, conserved (Spectrin & Ankyrin repeats)	
0,51	-0,96	3,93E-03	LinJ.23.0070	hypothetical protein, conserved	
0,51	-0,96	2,37E-03	LinJ.33.1440	hypothetical protein, conserved	Ama
0,52	-0,95	3,44E-03	LinJ.33.3030	hypothetical protein, conserved	
0,52	-0,94	8,05E-04	LinJ.11.0120	hypothetical protein, conserved	
0,52	-0,93	1,17E-03	LinJ.22.0670	A2 protein	Ama
0,52	-0,93	7,32E-03	LinJ.09.0580	leucine-rich repeat protein, putative	
0,53	-0,93	7,63E-04	LinJ.27.0310	methylmalonyl-coenzyme a mutase, putative	Ama
0,53	-0,92	1,13E-03	LinJ.34.4370	amastin-like surface protein, putative	Ama
0,53	-0,92	7,63E-04	LinJ.13.0200	hypothetical protein, unknown function (FYVE/PHD zinc finger & alpha/beta-Hydrolases)	
0,53	-0,91	5,51E-03	LinJ.27.1980	hypothetical protein, conserved	
0,53	-0,91	1,77E-03	LinJ.26.1770	hypothetical protein, unknown function	
0,53	-0,91	1,13E-03	LinJ.28.1070	P27 protein (P27)	Ama
0,53	-0,90	1,17E-03	LinJ.10.0710	hypothetical protein, unknown function	
0,54	-0,90	3,40E-02	LinJ.05.0380	microtubule-associated protein, putative	
0,54	-0,90	1,61E-03	LinJ.13.1090	adenylosuccinate synthetase, putative	
0,54	-0,90	1,17E-03	LinJ.30.0830	mitochondrial oligo_U binding protein TBRGG1, putative	
0,54	-0,89	3,14E-03	LinJ.03.0480	phosphate-Repressible Phosphate Permease-like protein	Ama
0,54	-0,89	3,24E-03	LinJ.27.1690	hypothetical protein, conserved	
0,54	-0,89	1,17E-03	LinJ.20.1290	hypothetical protein, conserved (Domain of unknown function)	
0,54	-0,89	8,91E-03	LinJ.26.1370	hypothetical protein, conserved	
0,54	-0,89	1,18E-02	LinJ.32.0390	hypothetical protein, conserved	
0,54	-0,88	2,26E-03	LinJ.03.0700	hypothetical protein, conserved	Ama
0,54	-0,88	3,68E-03	LinJ.35.0040	phosphatidylinositol-specific phospholipase-like protein	
0,55	-0,88	2,30E-02	LinJ.36.6790	hypothetical protein, conserved	
0,55	-0,87	9,15E-03	LinJ.23.1330	hypothetical protein, unknown function	
0,55	-0,87	1,17E-03	LinJ.14.0300	hypothetical protein, conserved	
0,55	-0,87	1,51E-03	LinJ.36.0710	GPI transamidase component Tta2, putative;with=GeneDB:Tb927.10.5080 (TTA2)	
0,55	-0,87	8,41E-04	LinJ.29.1100	hypothetical protein, conserved (Quinone oxidoreductase PIG3)	
0,55	-0,86	2,23E-03	LinJ.18.1270	hypothetical protein, conserved (Mitochondrial carrier)	
0,55	-0,86	2,12E-03	LinJ.35.1070	protein kinase, putative	
0,55	-0,86	1,07E-02	LinJ.35.3410	hypothetical protein, conserved (Putative death-receptor fusion protein)	
0,55	-0,86	1,13E-03	LinJ.16.1460	hypothetical protein, conserved	Ama
0,55	-0,85	1,17E-03	LinJ.24.1680	ring-box protein, putative	
0,56	-0,83	3,47E-03	LinJ.09.0290	hypothetical protein, conserved (CHASE domain profile)	Ama
0,56	-0,83	8,71E-03	LinJ.33.2350	hypothetical protein, conserved	

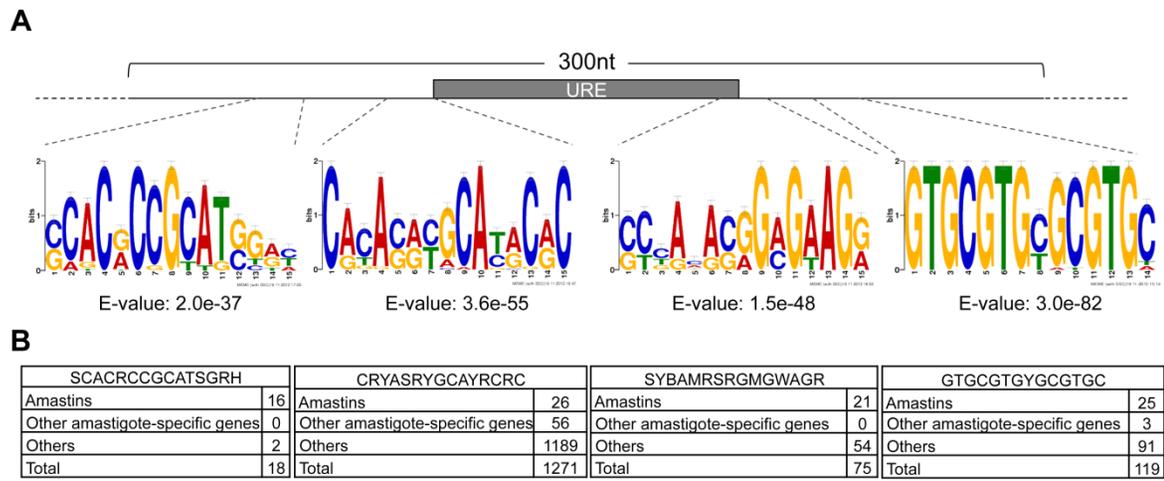
0,56	-0,83	4,80E-02	LinJ.03.0860	hypothetical protein, conserved (Pectin lyase-like)	
0,56	-0,83	1,57E-03	LinJ.23.1220	hydrophilic acylated surface protein b (HASPB)	Ama
0,56	-0,82	3,53E-03	LinJ.31.0520	hypothetical protein, conserved	
0,57	-0,82	5,35E-03	LinJ.36.5870	isoleucyl-tRNA synthetase, putative (ILERS)	
0,57	-0,81	1,20E-02	LinJ.11.1070	hypothetical protein	Ama
0,57	-0,80	8,08E-03	LinJ.27.2050	hypothetical protein, conserved	
0,58	-0,79	1,38E-02	LinJ.14.0600	hypothetical protein, conserved	
0,58	-0,79	7,57E-03	LinJ.18.0840	hypothetical protein, conserved	
0,58	-0,79	2,96E-03	LinJ.08.0080	ATP-dependent DEAD/H RNA helicase, putative	
0,58	-0,79	1,17E-03	LinJ.27.1680	casein kinase I-like protein	Ama
0,58	-0,78	3,26E-03	LinJ.19.1120	hypothetical protein, unlikely	
0,58	-0,78	1,37E-02	LinJ.34.1730	amastin-like surface protein, putative	Ama
0,58	-0,78	3,13E-03	LinJ.24.2400	hypothetical protein, conserved	
0,59	-0,77	3,12E-03	LinJ.08.0780	amastin-like protein	
0,59	-0,77	3,37E-03	LinJ.35.2720	hypothetical protein, conserved	
0,59	-0,77	1,85E-02	LinJ.31.0130	hypothetical protein, conserved (Heat shock factor binding protein 1)	
0,59	-0,76	6,54E-03	LinJ.25.2390	hypothetical protein, conserved (Zinc finger DHHC-type, Homologue of TbPAT10, palmitoyl acyltransferase by homology)	
0,59	-0,76	2,81E-03	LinJ.03.0890	inorganic pyrophosphatase, putative	
0,59	-0,76	6,51E-03	LinJ.33.2490	hypothetical protein, conserved (ATP synthase regulation protein NCA2)	
0,59	-0,76	2,98E-02	LinJ.34.1010	amastin-like surface protein, putative	Ama
0,59	-0,75	8,97E-03	LinJ.03.0320	hypothetical protein, conserved (DDRKG domain)	
0,59	-0,75	1,17E-03	LinJ.08.0800	hypothetical protein, unknown function	Ama
0,59	-0,75	1,94E-03	LinJ.08.0790	amastin-like protein	Ama
0,60	-0,75	1,17E-03	LinJ.10.0250	hypothetical protein, conserved (Amastin surface glycoprotein)	
<b>Upregulated genes in <i>LiAlba20</i><sup>-/-</sup></b>					
6,53	2,71	5,95E-05	LinJ.23.0700	hypothetical protein	
4,53	2,18	5,95E-05	LinJ.26.2710	hypothetical protein, unknown function	
3,11	1,64	1,38E-04	LinJ.35.1000	aldose 1-epimerase-like protein	
2,51	1,33	4,56E-04	LinJ.04.0200	surface antigen-like protein	Ama
2,51	1,33	4,56E-04	LinJ.17.0970	META domain containing protein, putative;with=GeneDB:LmjF17.0870 (META2)	Pro
2,51	1,33	7,31E-04	LinJ.17.0960	hypothetical protein, unknown function	
2,50	1,32	2,36E-04	LinJ.33.0660	hypothetical protein, conserved	Pro
2,49	1,31	7,31E-04	LinJ.32.0530	hypothetical protein, conserved	
2,39	1,26	1,13E-03	LinJ.09.0080	RNA-binding protein 5, putative	
2,37	1,24	2,36E-04	LinJ.35.0450	hypothetical protein, conserved (Male gamete fusion factor)	
2,36	1,24	6,31E-04	LinJ.02.0520	hypothetical protein, unknown function	Pro
2,34	1,22	1,11E-03	LinJ.31.1480	hypothetical protein, unknown function (Leucine-rich repeat)	Ama

2,26	1,17	8,41E-04	LinJ.36.4050	similar to leishmania major. l411.4-like protein	
2,22	1,15	4,75E-04	LinJ.16.0930	flagellar calcium-binding protein, putative	
2,19	1,13	9,06E-04	LinJ.16.0920	flagellar calcium-binding protein, putative	
2,10	1,07	1,74E-03	LinJ.30.2870	hypothetical protein, conserved	Pro
2,09	1,07	8,41E-04	LinJ.20.1220	calpain-like cysteine peptidase, putative,cysteine peptidase, Clan CA, family C2, putative	
2,09	1,06	1,11E-03	LinJ.13.0350	hypothetical protein, conserved (C2 domain [Calcium/lipid-binding domain, CaLB])	
2,09	1,06	8,05E-04	LinJ.31.2380	3'-nucleotidase/nuclease precursor, putative	
2,08	1,06	1,10E-03	LinJ.26.1980	hypothetical protein, conserved	
2,01	1,01	7,58E-04	LinJ.18.1280	hypothetical protein, conserved	Pro
1,98	0,99	8,41E-04	LinJ.09.0630	surface antigen-like protein	Ama
1,95	0,96	2,18E-03	LinJ.35.2200	adenine aminohydrolase (AAH)	Pro
1,94	0,96	8,41E-04	LinJ.36.4440	hypothetical protein, conserved	
1,94	0,95	2,14E-03	LinJ.36.2060	hypothetical protein, conserved (domain in TBC and LysM domain containing proteins)	
1,93	0,95	7,63E-04	LinJ.31.1130	N-acyl-L-amino acid amidohydrolase, putative	
1,92	0,94	1,47E-01	LinJ.29.1850	histone H2A, putative	
1,92	0,94	1,61E-02	LinJ.34.1830	hypothetical protein, conserved	
1,92	0,94	2,85E-03	LinJ.18.1070	calpain-like cysteine peptidase, putative,cysteine peptidase, Clan CA, family C2, putative	
1,90	0,93	1,51E-03	LinJ.20.1370	kinase-like protein	
1,89	0,92	1,17E-03	LinJ.04.1030	hypothetical protein, conserved	
1,89	0,92	9,54E-04	LinJ.19.1670	hypothetical protein, unknown function (Leucine-rich repeat; RING/U-box)	
1,86	0,89	1,11E-03	LinJ.23.1430	membrane-bound acid phosphatase 2 (MBAP2)	
1,85	0,89	1,60E-03	LinJ.35.5310	hypothetical protein, conserved (Ankyrin repeat)	Pro
1,85	0,89	1,98E-03	LinJ.06.1350	hypothetical protein, unknown function (Cytochrome b5-like heme/steroid binding domain)	Pro
1,84	0,88	4,55E-03	LinJ.26.1480	hypothetical protein, unknown function	
1,84	0,88	4,27E-03	LinJ.36.4340	hypothetical protein, unknown function	
1,83	0,88	1,17E-03	LinJ.32.0500	hypothetical protein, conserved	Pro
1,83	0,87	8,05E-04	LinJ.28.2850	membrane-bound acid phosphatase, putative	Ama
1,83	0,87	1,09E-03	LinJ.07.0980	hypothetical protein, conserved	Pro
1,83	0,87	1,42E-03	LinJ.27.1160	cation transporter, putative	
1,82	0,87	1,14E-03	LinJ.34.0210	hypothetical protein, conserved	
1,82	0,86	1,38E-03	LinJ.10.0520	GP63, leishmanolysin,metallo-peptidase, Clan MA(M), Family M8 (GP63-3)	
1,82	0,86	1,17E-03	LinJ.33.0310	glucose transporter/membrane transporter D2, putative	Ama
1,82	0,86	1,17E-03	LinJ.35.4280	hypothetical protein, unknown function	
1,82	0,86	1,78E-03	LinJ.36.6770	histidine secretory acid phosphatase, putative	Pro
1,81	0,86	8,41E-04	LinJ.35.0520	proteophosphoglycan ppg4	
1,81	0,86	1,17E-03	LinJ.27.1710	glycosomal phosphoenolpyruvate carboxykinase, putative	
1,81	0,86	1,17E-03	LinJ.19.0520	hypothetical protein, conserved (P-loop containing nucleoside triphosphate hydrolases)	Pro

1,81	0,85	1,17E-03	LinJ.34.0480	hypothetical protein, unknown function (Concanavalin A-like lectins/glucanases)	
1,80	0,85	1,21E-03	LinJ.25.1530	cyclin (CYCA)	
1,80	0,85	1,42E-03	LinJ.24.1990	hypothetical protein, conserved	
1,80	0,85	8,41E-04	LinJ.35.0540	proteophosphoglycan 5	
1,79	0,84	6,05E-03	LinJ.36.3000	hypothetical protein, conserved	
1,78	0,83	3,20E-03	LinJ.10.0590	hypothetical protein, unknown function	
1,77	0,83	1,05E-03	LinJ.34.4290	lipophosphoglycan biosynthetic protein (lpg2) (LPG2)	<b>Pro</b>
1,77	0,82	8,97E-03	LinJ.27.0130	hypothetical protein, conserved (WW domain & CCCH zinc finger)	
1,77	0,82	3,30E-03	LinJ.20.1190	endo-1,4-beta-xylanase z precursor-like protein	
1,76	0,82	2,95E-03	LinJ.10.1370	hypothetical protein, conserved	
1,76	0,82	1,36E-02	LinJ.06.0250	hypothetical protein, conserved	
1,76	0,81	1,72E-03	LinJ.27.0250	kinetoplast-associated protein-like protein	
1,75	0,81	1,46E-03	LinJ.32.2700	GPI transamidase component Tta1, putative (TTA1)	
1,75	0,81	1,46E-03	LinJ.02.0500	hypothetical protein, conserved	
1,75	0,81	1,17E-03	LinJ.23.1560	lathosterol oxidase-like protein	
1,75	0,80	1,17E-03	LinJ.16.0500	hypothetical protein, conserved (Prokaryotic membrane lipoprotein lipid attachment site profile)	
1,74	0,80	2,18E-03	LinJ.31.1870	protein kinase-like protein	
1,72	0,78	1,17E-03	LinJ.06.1340	protoporphyrinogen oxidase-like protein	
1,72	0,78	1,13E-03	LinJ.31.0430	calpain-like cysteine peptidase, putative, cysteine peptidase, Clan CA, family C2, putative	
1,71	0,78	1,11E-03	LinJ.09.1580	cytochrome b5-like protein	
1,71	0,78	3,94E-03	LinJ.22.0970	hypothetical protein, conserved (Protein kinase domain)	<b>Pro</b>
1,71	0,78	1,13E-03	LinJ.34.1060	p-glycoprotein (MDR1)	
1,71	0,77	1,94E-03	LinJ.18.1640	hypothetical protein, conserved	<b>Pro</b>
1,70	0,77	7,85E-03	LinJ.05.0450	hypothetical protein, conserved	
1,70	0,76	3,77E-03	LinJ.13.1150	hypothetical protein, unknown function	
1,69	0,76	6,61E-03	LinJ.20.0170	hypothetical protein (TPR-like)	<b>Pro</b>
1,69	0,75	1,17E-03	LinJ.26.0010	Lsm7p protein, putative	
1,68	0,75	2,79E-03	LinJ.35.0530	proteophosphoglycan 5	

<sup>a</sup>Ama (amastigotes) and Pro (promastigotes) represent differentially expressed genes in both *L. infantum* life stages according to previously published DNA microarray analysis by Rochette A, *et al.* (2009). The same statistical analysis was applied in both studies

**Fig. S1. Putative Alba20 protein binding sequence motifs common to 3'UTRs of  $\delta$ -amastin transcripts in *L. infantum*.** A. Schematic representation of the 300 nt region within the amastin (LinJ.34.1010) 3'UTR used in RNA affinity chromatography experiments (see Fig. 1). A MEME MAST motif search revealed four motifs within the 300 nt region, all outside the 100 nt URE, that were conserved in the 3'UTR of other  $\delta$ -amastins. For MEME/MAST motif search, we used the 2 kb intergenic region downstream of each  $\delta$ -amastin gene and BLAST it to TriTrypDB database. To perform a discriminative search, we also selected 50 similar regions located downstream of constitutively expressed small membrane proteins. Search was done only for the given strand and different attempts revealed no identification of motifs longer than 15 nt. Among the 25 best motifs identified, only the ones present in the 300 nt 3'UTR region of the amastin transcript are shown here. B. To confirm the specificity of those motifs, a full search was performed on the genome of *L. infantum* using TritypDB (<http://tritrypdb.org>) search tools. Briefly, all gene-coding sequence positions were downloaded to define a list of intergenic regions with respect to the orientation of the polycistronic units. Each motif was then searched in a strand-specific manner to identify genes harboring the motif in the 3'intercoding region. The consensus sequence, shown on the top of the tables in IUPAC notation, was used as a DNA motif pattern search in TriTrypDB.



**Fig. S1**



## **Chapitre 10 : Analyses structurelle et fonctionnelle des protéines à domaine Alba chez *Leishmania*.**

### **Résumé**

Les protéines à domaine Alba sont présentes des *Archaea* jusqu'aux eucaryotes supérieurs. Chez les parasites protozoaires plus spécifiquement, elles régulent l'expression de protéines membranaires et facteurs de virulence différemment exprimés durant les cycles parasitaires. Nous montrons que les protéines Alba de *Leishmania* présentent des mutations au niveau d'acides aminés impliqués dans des modifications post-traductionnelles. Bien que constitutivement exprimées, elles sont plus importantes pour la croissance du parasite au stade amastigote (dans les macrophages d'un hôte mammifère) qu'au stade promastigote (insecte vecteur). Ces protéines sont cytoplasmiques, mais elles se concentrent dans le nucléole et le flagelle durant la différenciation du parasite de promastigote à amastigote. Nous montrons que les protéines Alba forment un complexe qui interagit avec d'autres protéines de liaison à l'ARN et la machinerie de traduction. Cependant, elles ne s'associent pas fortement aux polysomes en cours de traduction, suggérant un rôle de ces protéines dans l'initiation de la traduction.

## Article

### Structural and functional analysis of Alba domain proteins in *Leishmania*

Aurélien Dupé, Carole Dumas and Barbara Papadopoulou\*

Pour publication dans *Plos One*

Research Center in Infectious Disease, CHU de Quebec Research Center (CHUL),  
Department of Microbiology, Infectious Disease and Immunology, Laval University,  
Quebec, QC, Canada G1V 4G2

\*Corresponding author:

Research Center in Infectious Disease, CHUL RC-709

2705 Laurier Blvd, Quebec, QC, Canada, G1V 4G2

Tel: +1 418 656-4141, ext: 47608; Fax: +1 418 654-2715

E-mail: barbara.papadopoulou@crchul.ulaval.ca

**Keywords:** *Leishmania*, Alba domain proteins, differentiation, heat stress, differential protein localization, flagellum, nucleolus

**Abbreviations:** Alba: Acetylation lowers binding affinity; PABP: Poly(A) binding protein; PTM: Posttranslational modification; mRNA: messenger RNA; tRNA: transfer RNA; rRNA: ribosomal RNA

## Abstract

Alba domain proteins (for Acetylation Lowers Binding Affinity) are found in *Archaea* and eukaryotes and more recently have been characterized in protozoan parasites where they play a role in the regulation of expression of various virulence factors or stage-specific membrane proteins. Here we provide structural and functional insights on the Alba domain proteins in the parasitic protozoan *Leishmania*. We show that the *Leishmania*

genome codes for two Alba domain proteins, Alba13 and Alba20, that, despite their conserved sequence and structure with other trypanosomatid Alba proteins, have specific mutations in key regulatory sequences implicated in posttranslational modifications. Interestingly, the *Leishmania* Alba domain proteins are important for amastigote growth and display a differential subcellular localization throughout the parasite life cycle. They are indeed localized in the cytosol and the perinuclear space in the insect promastigote stage but they translocate to the nucleolus and also to the flagellum during the promastigote to mammalian amastigote stage differentiation process or under heat stress, which triggers differentiation. Furthermore, we show that Alba13 and Alba20 proteins form a complex interacting with other RNA-binding proteins, Poly(A) Binding Proteins, and components of the 40S and 60S ribosomal subunits but that they only associate with ribosomal subunits and not with translating polysomes, hence suggesting a role of these proteins in translation initiation. Taken together, these data provide new knowledge on the function of Alba domain proteins in *Leishmania*.

## **Introduction**

Trypanosomatids are a group of kinetoplastid protozoa that are responsible for a variety of diseases in tropical and subtropical regions around the world. These include sleeping sickness caused by *Trypanosoma brucei*, the ‘Chagas’ disease caused by *Trypanosoma cruzi*, and different forms of leishmaniasis caused by *Leishmania* spp. *Leishmania infantum* is the pathogenic agent of visceral leishmaniasis, the most severe form of the disease, which is responsible for more than 500 000 new cases and around 50 000 deaths annually [1]. *Leishmania* metacyclic promastigotes are injected into the skin of the mammalian host by the sandfly’s bite and internalized in the phagolysosome of macrophages where they differentiate into round-shape amastigotes, a form particularly adapted to replicate within the harsh environment of phagolysosomal vacuoles [2, 3]. The modification of the temperature between the poikilothermic insect vector and the thermostable mammalian host and the acidic environment of the phagolysosomal vacuoles trigger amastigote differentiation. Moreover, the differences in nutrient availability in the phagolysosome and the adaptive strategies of the parasite to resist the host defense mechanisms are responsible for the differential regulation and expression of many sets of genes [4].

Trypanosomatids share an unusual mechanism of regulation of gene expression. Polymerase II-mediated transcription produces long polycistronic units containing up to hundreds genes of non-related functions. Processing to individual and mature transcripts is obtained by coupled 5' *trans*-splicing and 3' polyadenylation reactions (reviewed in [5]). The 3'-untranslated regions (3'UTR) in *Leishmania* reach an average of ~1.5 kb and sequence motifs within the 3'UTRs are often responsible for posttranscriptional regulation through interactions with RNA-binding proteins [6]. mRNA metabolism is associated with various physiological or stress responsive granules. If none of them has been reported in *Leishmania*, more work has been done in *Trypanosoma* species where P-bodies, heat stress granules, starvation granules or perinuclear granules have been described [7, 8].

Alba domain proteins (for Acetylation Lowers Binding Affinity) are found in *Archaea* and eukaryotes and display an unusual evolutionary history. Being nucleosome constituents and having RNA binding affinities in *Archaea*, they are found as RNase P/MRP subunits in higher eukaryotes mainly implicated in tRNA and rRNA maturation [9]. More recently, Alba domain proteins have been characterized in protozoan parasites where they play a role in the regulation of expression of various virulence factors or stage-specific membrane proteins. In *Plasmodium falciparum*, the RGG-box-containing *PfAlba3* binds to DNA in telomeric and subtelomeric regions in an acetylation-dependent manner and modulates the expression of the virulent *var* genes. Interestingly, the four *PfAlba* proteins re-localize from the cytoplasm to the nucleus in the trophozoite and schizont stages where they bind to the Telomere-Associated Repetitive Elements (TARE1 and TARE6), also associated to the regulation of *var* genes [10, 11]. In the closely related species *Plasmodium berghei*, Alba proteins have an RNA-binding activity and *PbAlba1-3* are part of RNP complexes of translationally silent transcripts in oocyte P-granules [12]. In *T. cruzi*, the association of some Alba domain proteins with polysomal-enriched fractions suggests their interaction with the translation machinery [13]. In *T. brucei*, the four *TbAlba* proteins interact with the procyclic-specific *GPEET* transcript and seem to play a role in its translation inhibition [14]. Interestingly, the RGG-box-containing *TbAlba* proteins are present in starvation and heat stress granules and seem to play a more global role in cell organization, involving the aborted nucleus migration during the *T. brucei* differentiation [15].

In this study, we provide data on the structural and functional characterization of the Alba domain proteins in the parasitic protozoan *Leishmania*. We show that the *Leishmania* genome codes for two Alba domain proteins, Alba13 and Alba20, that are localized in the cytosol but interestingly they translocate to the nucleolus and the flagellum specifically during the promastigote to amastigote differentiation process or under heat stress. Furthermore, we show that Alba13 and Alba20 proteins form a complex interacting with other RNA-binding proteins and components of the 40S and 60S ribosomal subunits.

## Results

### Alba domain proteins of *Leishmania*

Alba domain proteins are nucleosome constituent and RNA binding proteins in *Archaea* and are globally defined as small basic proteins. This family was later phylogenetically subdivided into three groups that include *Archaea* proteins and eukaryotic proteins sharing sequence homologies to RNase P subunits, Rpp20/Pop7 and Rpp25/Pop6. A global inter-species alignment published previously highlights the two features of Alba domain proteins in DNA and/or RNA binding [9]. In *Trypanosoma brucei* and *T. cruzi*, there are four tandemly repeated Alba domain proteins corresponding to two proteins of each eukaryotic subgroups. Interestingly, all *Leishmania* species sequenced to date (*L. infantum*, *L. major*, *L. braziliensis*, *L. mexicana* and *L. tarentolae* [16-18]) possess only two Alba domain proteins (LinJ.13.0270 and LinJ.34.2410 in the case of *L. infantum*), corresponding to one of each eukaryotic subgroups, suggesting that duplication events of *Alba* genes occurred after *Leishmania* and *Trypanosoma* speciation, as illustrated in the Neighbor-joining tree presented in Figure 1A.

The *L. infantum* LinJ.13.0270 (named *LiAlba13*) is part of the Rpp20-like subgroup. Proteins of this subgroup have a molecular weight of 12 to 14 kDa and interestingly, although they have a basic *pI* in *Trypanosoma* spp., in *Leishmania* spp. their *pI* is acidic. This difference suggests that *Leishmania* proteins have different charged amino acids, underlining differences in their abilities to bind their substrates. The charge differences are mainly observed in the C- and N-terminal regions (Fig. 1B). Within *Leishmania* species, Rpp20-like proteins are well conserved sharing amino acid sequence identities up to 88-95% and similarities up to 91-98%, with *L. braziliensis* being the most divergent. However,

*LiAlba13* protein identity and similarity with *Trypanosoma* spp. Rpp20-like proteins drop to 48-54% and 58-65%, respectively. Neighbor-joining tree suggests that duplication events of Rpp20-like proteins occurred before the *T. brucei* and *T. cruzi* speciation. *T. brucei* paralogues share only 54% of identity. *LiAlba13* shares few more similarities with *TbAlba2* (Tb11.02.2030, 61%) than *TbAlba1* (Tb11.02.2040, 58%).

Given the lower primary sequence similarity between the *Leishmania* and *Trypanosoma* Alba proteins and particularly *LiAlba13* and *TbAlba2*, we carried out a structural comparative analysis using the Phyre algorithm. The best score was obtained with the Alba protein from *Sulfolobus solfataricus*, *Sso10b1* [19], with an E-value of 8.2e-09. The  $\beta\alpha\beta\alpha\beta\beta$  topology is conserved in *LiAlba13* with an additional  $\alpha$  helix at the C-terminal region (Fig. 1B), corresponding to the previously described synapomorphic C-terminal extension containing the FDxh signature motif, which is specific of the Rpp20-like subgroup [9]. The coil regions are the most divergent in trypanosomatids but are well conserved in *Leishmania* spp. (positions 3-13, 17-22, 80-90 and 116-131 of the alignment in Fig. 1B). The lysine residue usually deacetylated by Sir2 (reviewed for parasites in [20, 21]) is not conserved in trypanosomes (position 19 of the alignment) but we cannot rule out the possible acetylation on neighboring lysine residues. While the acetylation of Alba domain proteins by Sir2 has been characterized in *Archaea* [22] and *Plasmodium falciparum* [11], there are no published reports in trypanosomatids to date. Another possible well-known posttranslational modification (PTM) is protein phosphorylation. It has been reported that *TbAlba2* in *T. brucei* bloodstream forms is phosphorylated at Serine 24 [23], an amino acid that is not conserved in *Leishmania* spp. Nevertheless, in *T. cruzi* metacyclics, Ser 2 of Tc00.1047053504089.70 is phosphorylated and this amino acid is well conserved in *Leishmania* [24]. Although none of the Alba domain proteins from *L. infantum* has been identified in phosphoproteome studies so far, the conservation of Ser 2 suggests that Alba proteins might be phosphorylated in *Leishmania* [25-27]. Phosphorylation of Alba domain proteins on Ser/Thr residues has also been reported in *Toxoplasma gondii* and *P. falciparum* [28-30]. In *P. falciparum*, phosphorylation by protein kinase CK2 [31] has been characterized for *PfAlba1* and 2, belonging to the Rpp20-like subgroup.

LinJ.34.2410 (*LiAlba20*) is part of the Rpp25-like subgroup. In trypanosomatids, all Rpp25-like proteins have a molecular weight ranging from 20 to 25 kDa and a *pI* ranging from 9.9 to 10.3. Within *Leishmania* spp., Rpp25-like proteins share a 88 to 95 % identity and 92 to 96% similarity, with *L. tarentolae* being the most divergent. Sequence comparison of *LiAlba20* with *Trypanosoma* spp. Rpp25-like proteins shows a drop in identities and similarities, ranging from 48-59% and 59-68%, respectively as also observed for *LiAlba13*. As opposed to Rpp20-like protein evolution, a neighbor-joining tree suggests that duplication events of Rpp25-like proteins occurred after the *Trypanosoma* species separation, resulting in better conserved paralogues than orthologues. *T. brucei* paralogues share 77% of sequence identity. The sequence of *LiAlba20* is more similar to *TbAlba3* (Tb927.4.2040, 68%) than to *TbAlba4* (Tb927.4.2030, 59%).

Structure prediction of *LiAlba20* using Phyre outputs *Alba2* from *Aeropyrum pernix* K1 [32] as the best match (E-value 1.2e-07). The  $\beta\alpha\beta\alpha\beta\beta$  secondary structure seems conserved in *LiAlba20*, with an additional  $\alpha$  helix in the N-terminal region (Fig. 1C) and a longer C-terminal extension containing the RGG-box domain, which is specific of Rpp25-like proteins [9] together with the signature sequence GYQxP (position 146-150 of the alignment). As for *LiAlba13*, the coil regions are the ones with the highest divergence between *Leishmania* spp. and *Trypanosoma* spp. (positions 15-23 and 107-118 of the alignment Fig. 1C, see also the C-terminal extension ending with the DRRDDREN motif specific to *Leishmania* spp). Here, the lysine residue deacetylated by Sir2 is replaced by a threonine in *Leishmania* (position 38 of the alignment). In *P. falciparum*, deacetylation of *PfAlba3* (Rpp25-like subgroup) has been reported [10]. In *T. brucei* procyclic forms, *TbAlba3* is phosphorylated at tyrosine 4 [33]. Tyrosine phosphorylation is rare in trypanosomatids, and this conserved amino acid in *Trypanosoma* spp. is replaced by an arginine in all *Leishmania* species. In *T. cruzi* metacyclics, Ser residues 144 and 157 are phosphorylated [24], and only Ser144 is conserved in trypanosomatids, suggesting a potential phosphorylation of *LiAlba20* on this residue.

Taken together, these sequence predictions indicate that Alba domain proteins have a conserved structure in trypanosomatids, but *Leishmania* shares specific mutations in key regulatory sequences implicated for example in protein phosphorylation. RGG-box

methylation remains also a good candidate for PTM. Immunoprecipitation experiments in wild type cells using mono- or dimethyl-arginine antibody did not suggest any stage-specific modulation of *LiAlba20* methylation (data not shown). In *L. major* and *L. infantum*, the RGG-box presents a good conservation of the direct repeat RGGRGGS, resulting in 3 SR sites. In *Saccharomyces cerevisiae*, Npl3p, implicated in mRNA splicing and transport, possesses a RGG and SR-rich domains, and phosphorylation of SR modulates the interaction of this protein with the nuclear import receptor Mtr10pm, regulating the nuclear import of Npl3p. A possible explanation of the absence of *LiAlba* proteins in large-scale phosphoproteomic studies might be related to their fast degradation pattern during the protein extraction procedure in *Leishmania*. We have observed two bands of ~20 and ~25 kDa by Western blotting when proteins were extracted with non denaturing buffers as CEB, even in the presence of various protease inhibitor cocktails, whereas direct resuspension of parasite pellets in denaturing Laemmli Sample Buffer showed only the 25 kDa band (data not shown).

#### **Alba domain protein partners identified by co-immunoprecipitation**

To examine the functional role of *LiAlba20* and *LiAlba13* proteins, we carried out immunoprecipitation studies coupled with MS/MS peptide identification to identify putative Alba domain interacting proteins. For this purpose, we used *L. infantum* recombinant strains expressing either the HA-*LiAlba20* or *LiAlba13*-HA proteins and wild-type cells as negative controls. The MS/MS results from both *L. infantum* promastigote and axenic amastigote forms, representative of three independent experiments for each construct, are shown in Table 1. Alba20 and Alba13 are known to interact together [10, 14, 34] and this interaction is also revealed in *L. infantum* both by immunoprecipitation studies (Table 1) and by co-localization (see below). This explains why most of the identified proteins were found associated with both HA-*LiAlba20* or *LiAlba13*-HA constructs. The three poly(A) binding proteins (*LiPABP1-3*) are amongst the most abundant proteins found to interact with *LiAlba13* and *LiAlba20* in both life stages, based on the number of identified peptides (Table 1). This observation correlates with the implication of *LiAlba* in RNA metabolism and confirms the results observed in *T. brucei* [14], *P. berghei* [12] and more recently in *T. gondii* [35]. Nevertheless, association of *LiAlba13* and *LiAlba20* with PABPs might not occur by a direct protein-protein interaction as prior treatment with

RNase and co-immunoprecipitation did not identify *LiPABP2* by western blot analysis (data not shown), as it was also shown in *T. brucei* in the case of eIF4E4 [14]. *LiAlba13* and *LiAlba20* also interact with the ATP-dependent RNA helicase (LinJ.32.0410) an homologue of the yeast Dbp1 implicated in 43S complex ribosome scanning and in the distribution of mRNAs between P-bodies and polysomes [36]. We have previously reported that this RNA helicase protects rRNA from degradation by preventing rRNA antisense fragmentation and translation inhibition [37]. Interestingly, immunoprecipitation with this RNA helicase also pulls down *LiAlba20*, hence confirming their interaction (M. Samant *et al.*, manuscript in preparation). A 25 kDa RNA binding protein (LinJ.32.0790) specific to trypanosomatids with a nucleolar RNA binding domain, and a ZC3H41 helicase homologue (LinJ.27.1220 [38]) are some of the other proteins implicated in RNA metabolism that seem to interact with *LiAlba13* and *LiAlba20* domain proteins in both developmental stages of the parasite. To our knowledge, these proteins have not been studied yet. LinJ.21.0490 is an NTF2-like protein (for Nuclear Transport Factor 2) that interacts with both Alba domain proteins. In other organisms, NTF2 is involved in RanGTP transport and cargo shuttling between the cytoplasm and nucleoplasm (reviewed in [39]). Proteins with this domain are not well studied yet in trypanosomes, but a recent study shows that the RNA-binding protein DRBD3 from *T. brucei* localizes to the nucleus upon oxidative stress and interacts with a NTF2 domain protein [40]. The mitochondrial RNA proteins MRP1 and MRP2 were also identified to interact with *LiAlba13* and *LiAlba20* proteins both in promastigotes and amastigotes and were absent from our negative controls. Even if these proteins are known to be contaminants in immunoprecipitation experiments, it is nevertheless intriguing that both *LiAlba* proteins and the four *TbAlba* retrieve those proteins [14]. Nevertheless, Alba domain proteins are not expected to be mitochondrial (see below), and their interaction with MRP1 and MRP2 must have occurred during cell lysis. Another explanation would be that part of MRP1 and MRP2 reside in the cytosol. Some of the ribosomal proteins of the 40S (S2, S3A, S6, S9, S14, S24e) or 60S (L6, L7a) subunits and the ribosomal protein L1a were also found associated with *LiAlba13* and *LiAlba20* proteins in both parasite life stages.

In amastigotes, *LiAlba13* and *LiAlba20* seem to interact with several ribosomal proteins from the 40S (L14, S3, S4, S5, S7, S11, S15a, S17, S18, S19, S20, S23) and 60S

(L2, L5, L7, L9, L10, L10a, L11, L13, L21, L24, L23a, L26, L27) subunits. These data suggest that *LiAlba* domain proteins interact with the translation machinery. In addition, DRBD2 (LinJ.35.2240, a possible homologue of yeast Gbp2p involved in mRNA export), a homologue to Rap55 (LinJ.25.0550, a well-known interactor of DHH1 in other organisms), a putative vacuolar protein (LinJ.19.0020 [41]), and a second NTF-2 like protein (LinJ.18.0300) were found to interact with *LiAlba13* and *LiAlba20* only in amastigotes. In *T. brucei* under starvation stress, *TbAlba3* and *TbAlba4* are co-localizing in starvation granules with the RNA helicase *TbDHH1* [15]. *TcAlba3* and *TcDHH1* interaction seems to occur in *T. cruzi* (MG Ferrarini, FB. Holetz, AR Ávila, unpublished data). In our experiments, the *LiDHH1* orthologue (LinJ.35.0370) was not identified as a partner of *LiAlba* proteins despite the fact that Rap55 was present. Several other proteins like a 44 kDa RNA binding protein specific to *Leishmania* (LinJ.18.0590), an ATP-dependent phosphofructokinase, a seryl-tRNA synthetase, and an acotinase were found to interact with either *LiAlba13* or *LiAlba20*, specifically in promastigotes.

In the closely related species *T. brucei*, Mani *et al.* proposed that *TbAlba* proteins regulate the procyclic specific GPEET protein expression, most likely at the level of translation initiation. In their experiments, interaction of some *TbAlba* proteins with the translation machinery is clearly visible on polysome fractionations [14]. In *T. cruzi*, orthologues of *LiAlba20* are also enriched in polysomal fractions [13]. In *L. infantum*, polysome profiling studies confirm that both *LiAlba20* and *LiAlba13* are enriched in the 40S and 60S ribosomal subunits but not (or much less) in the polysomes, suggesting that they might play a role in translation initiation (Figure 2). These findings are in line with immunoprecipitation studies where several ribosomal proteins of the 40S and 60S subunits were found to interact with *LiAlba20* and *LiAlba13* in promastigotes but to a greater extent in amastigotes (Table 1). Following EDTA treatment *LiAlba20* is only found in the 3 first light fractions of the sucrose gradient. Thus, in contrast to the results obtained in *T. brucei*, *LiAlba20* seems to bind to mRNAs that are not actively translated.

### **Differential subcellular localization of Alba domain proteins in *Leishmania***

Next, we investigated the cellular localization of the *Leishmania* Alba domain proteins. For these studies, we first analyzed recombinant parasites episomally expressing

HA-*LiAlba20* or *LiAlba13*-HA. In *L. infantum* promastigotes, both Alba domain proteins showed a cytoplasmic localization (Fig. 3A-3B, upper panels), as also observed in *T. brucei* [14, 15] and consistent with a role of these proteins in RNA metabolism. In some promastigote forms, we observed a perinuclear localization of Alba proteins (Fig. 3A-3B, upper panels). However, during promastigote to amastigote differentiation triggered by increased temperature and acidic stress, the localization of *LiAlba* proteins changed. From six hours to overnight exposure to the differentiation signals, both Alba proteins concentrate in cytoplasmic granules, and the most visible effect is an enrichment in the flagellum and inside the nucleus (Fig. 3A-3B, lower panels). In the case of *LiAlba20*, fusing the HA tag at the C-terminal region resulted in similar localization (data not shown).

To test if the unusual localization observed during differentiation was not driven by the presence of the HA tag, we generated new constructs based on fluorescent protein fusions. All constructs were done by fusing proteins with eYFP or mCherry in the N-terminal region (Fig. S1) and transfected into *L. infantum* wild type cells. Over-expression of the *LiAlba* proteins should not create false localization, because *LiAlba* proteins seem to be very abundant and are constitutively expressed in both life stages (our immunoprecipitations results and data not shown). The first strain generated was an episomal co-transfection of eYFP-*LiAlba13* and mCh-*LiAlba20* in the wild type background. The fluorescent signals observed reproduce the immunolocalization experiments, showing cytoplasmic localization in promastigotes and a concentration of the protein in the nucleus during amastigote differentiation (Fig. 4). Flagellar localization is also visible but the signal is less intense. This might be explained by the competition of fluorescent Alba versus endogenous Alba proteins for the flagellar internalization. A clear association of *LiAlba* proteins is also visible around the plasma membrane during differentiation. Using fluorescent proteins, signals were observed in amastigotes and show that *LiAlba* proteins display a cytoplasmic localization a little more punctuated than in promastigotes and without perinuclear accumulation. We observed that Alba proteins in *L. infantum* always co-localize during the life cycle and differentiation.

Interestingly, within the nucleus, *LiAlba* proteins accumulate in a region with a weaker DAPI staining, usually corresponding to the nucleolus, as described previously for

other proteins [42, 43]. As a control for nucleolar localization, we used the Nop10 protein [44] (nuclear localization of *LiNop10* in promastigotes, amastigotes and during differentiation is confirmed on Supplementary Fig. 2). Our second strain was generated by episomal transfections of the eYFP-*LiAlba13* and mCh-*LiNop10* to confirm the nucleolar localization during the promastigote to differentiation. A co-localization signal with Nop10 confirms the presence of *LiAlba* proteins in the nucleolar compartment (Fig. 5A). It is noticeable that during the differentiation the Nop10 signal is less focussed and this correlates with the nucleolar rearrangement observed during stress in eukaryotes (reviewed in [45, 46]).

To evaluate if other proteins interact with the *LiAlba* protein complex during the differentiation of the parasite, immunoprecipitation experiments were carried out after overnight incubation at 37°C in MAA pH 5.8, the usual medium used for axenic amastigote development and growth. Just the two *LiAlba* proteins and a HSP70 (LinJ.28.2960) were identified (data not shown). We have previously shown that this HSP70 protein interacts with the ATP-dependent RNA helicase and has a cytoplasmic localization based on digitonin fractionation data [47]. This low output result might be explained by various reasons: (1) many parasites do not survive to an overnight differentiation stress, and cell death results in the dissociation of complexes; (2) during stress, *LiAlba* proteins are released from the mRNA and therefore identification of RNA metabolism related partners failed; and (3) a cleavage occurs within *LiAlba* proteins, removing the HA-tag. Nevertheless, the identification of the HSP70 protein suggests that *LiAlba* proteins may be heat responsive and cause dissociation from their target mRNAs.

### **Heat shock stress induces the flagellar and nucleolar accumulation of *LiAlba* proteins**

To further investigate the signals necessary for the relocalization of the *LiAlba* proteins during the differentiation, we first stressed the cells only with temperature (overnight switch of exponentially grown parasites from 25° to 37°C in SDM pH 7.0) or pH variations (exponentially grown parasites were resuspended in HCl acidified SDM pH 5.8 and kept at 25°C). The temperature was sufficient for relocalizing *LiAlba* proteins in the nucleolus and flagellum (Fig. 5B). In contrast, pH stress did not provoke a differential

localization (Fig. 5B). Temperature stress is responsible for translation inhibition in *Leishmania* mediated by phosphorylation of the eIF2 $\alpha$  subunit [48]. Therefore, we decided to verify if *LiAlba* relocalization was mediated directly by the temperature stress or if translation inhibition with cycloheximide was sufficient for a nucleolar / flagellar localization of Alba domain proteins. Our data did not show any differences between treated and untreated parasites (Fig. S3).

Inhibition of transcription using Actinomycin D is another type of stress already reported in *Leishmania* for the nucleolar accumulation of cytoplasmic RNA-binding proteins [43]. Here, Actinomycin D treatment did not result in the accumulation of *LiAlba* proteins in the nucleus or in the flagellum (Fig. S3). The result presents a totally diffused signal all over the cell body, including the nucleus but excluding the flagellum, suggesting that the perinuclear accumulation of *LiAlba* proteins in promastigotes is dependent on the presence of mRNA.

#### **Depletion of Alba domain proteins in *Leishmania* decreases parasite growth in the amastigote stage**

We have previously generated knockout mutants for both *LiAlba13* and *LiAlba20* proteins of *L. infantum* (Dupé *et al.*, submitted). Here, we examined the effect of Alba domain protein depletion in the parasite growth in both developmental stages. No differences in promastigote growth were observed between either *LiAlba13*<sup>-/-</sup> or *LiAlba20*<sup>-/-</sup> null mutants and the wild type and add-back mutant strains (Fig. 6A). However, we observed a drastic decrease in amastigote growth in the absence of either *LiAlba13* or *LiAlba20* (Fig. 6B). This defect in amastigote growth was more pronounced in the *LiAlba13*<sup>-/-</sup> mutant and the phenotype was due to the depletion of Alba domain proteins as partial or total reversion was observed with the add-back mutants (Fig. 6B). *LiAlba13* seems more important for the amastigote growth than *LiAlba20*. Despite this observation, these two strains are able to differentiate in axenic amastigotes based on their ability to acquire round shapes and lose the flagellum (data not shown) but they grow much slower than wild type cells under those conditions. These data suggest that although Alba domain proteins are constitutively expressed in both life stages of the parasite (Dupé *et al.*, submitted), their function is more important for amastigote growth and development, may

be due to their putative role in mRNA regulation and/or interaction with other protein factors, specifically in this life stage.

## Discussion

Here, we characterized the Alba domain proteins and their partners in both life stages of *L. infantum*. If Alba domain proteins are known to be part of the nucleosome and to interact with RNA in *Archaea*, they are nevertheless constituents of RNase P / MRP in higher eukaryotes, implicated in the maturation of rRNAs and tRNAs. Interestingly, recent studies in protozoan parasites like *Trypanosoma* [14], *Plasmodium* [12], *Toxoplasma* [35] and *Leishmania* (Dupé *et al.*, submitted) highlight another aspect of Alba domain proteins that is the interaction with mRNA and the translation machinery. We do not have any direct experimental evidence on the potential role of Alba domain proteins in translation regulation in *Leishmania*, but these proteins interact by immunoprecipitation with a large number of ribosomal proteins part of the 40S and 60S subunits and all three poly(A)-binding proteins. These interactions and their high enrichment in the 40S and 60S ribosomal subunits, but not in translating polysomes, as shown by polysome profiling suggest that Alba domain proteins may be involved in the translation initiation process. It seems that at least three different pathways existed for the evolution of those particular proteins. Nevertheless, the involvement of Alba domain proteins as DNA binding proteins in *Apicomplexa* still exists [10, 49, 50]. In another study, we demonstrated that *LiAlba20* participates in the stabilization of the amastigote specific amastin transcript (Dupé *et al.*, submitted).

Another interesting aspect of our study concentrates on the subcellular localization of Alba domain proteins in *L. infantum*. Under conditions of promastigote and axenic amastigote growth, *LiAlba* proteins are cytoplasmic, and a more pronounced perinuclear accumulation is visible in promastigotes. The granules are indeed not always visible, and might exist under specific factors, like cell density or cell cycle progression. Under mRNA processing inhibition, a formation of perinuclear granules was described in *T. brucei*, as a compartment that can determine the fate of newly transcribed mRNAs and containing P-bodies proteins [8]. In *T. brucei*, *TbAlba3* and *TbAlba4* colocalize with *TbDHH1* in granules under starvation stress [15], and the Rap55 homologue, a well known partner of

DHH1, seems to interact with *LiAlba* proteins in our experiment. Co-localization experiments of *LiAlba* proteins and amastin mRNA, or with enzymes implicated in mRNA metabolism as DHH1, or degradation enzymes are necessary to have a better interpretation of the role of these structures in normal conditions in *Leishmania*. In *T. gondii*, it has been shown that both *TgAlba* proteins are mainly cytoplasmic in intracellular parasites but an accumulation of the proteins in the nucleus and in perinuclear spots was visible in extracellular parasites [35]. In *T. brucei* procyclics, the four *TbAlba* are cytoplasmic, and are part of stress granules following nutriment starvation [14].

Interestingly, following 6 hours of combined elevated temperature and acidic pH stress triggering *L. infantum* promastigote to amastigote differentiation, *Alba* domain proteins concentrate in the flagellum and the nucleolus. Actually, heat shock alone is able to trigger the flagellar and nucleolar localization. Heat shock stress is accompanied by a decrease in global translation [48, 51] but inhibition of translation with cycloheximide did not promote the relocalization of *LiAlba* proteins. These results suggest that the signal mediating the relocalization of *LiAlba* proteins during the stress is not due to translational arrest but to a direct adaptation to temperature stress. This can be mediated by the cytoplasmic HSP70 protein, which seems to associate with *LiAlba* proteins during the stress, and/or unidentified PTMs like phosphorylation or methylation during differentiation. The absence of PABPs and other RNA-binding proteins during the differentiation as revealed by immunoprecipitation might indicate that *LiAlba20* proteins do not bind anymore to mRNA, but additional experimental evidence is required to confirm whether or not this occurs. Flagellar and nucleolar localization might represent a sequestration of the *LiAlba* proteins to allow differential regulation of their target mRNAs. The flagellum does not possess its own translational machinery, making obligatory the synthesis and transport of its constituents from the cytosol [52]. It is unexpected that mRNA storage or degradation occurs in this very specialized organelle. RNA FISH would be required to confirm the absence of *Alba*-bound mRNAs in this compartment.

In eukaryotes, the primary function of the nucleolus is the biogenesis of ribosome subunits including the processing of the rRNA (reviewed in [53]). The nucleolus has also revealed other roles and is implicated in stress response, including cell-cycle regulation via

p53 activity regulation during stress response or the regulation of RNA polymerase I by the transcription factor TIF-IA (reviewed in [45]). The fact that many RNA-binding proteins relocate to the nucleolus in trypanosomatids underlines that this organelle might have kept this stress response ability. In *L. mexicana* and *T. cruzi*, Actinomycin D treatment or heat stress provoke the nucleolar accumulation in a phosphorylation- and ATP-dependant manner of poly(A)<sup>+</sup> mRNAs and various proteins such as *TcSR62* (mRNA stability and processing), *TcPABP1* and *LmexPABP2*, *TcPTB2* (homologue of polypyrimidine tract binding protein), *TcSF3b155* (splicing factor) and *TcFIP1* (involved in polyadenylation), but this pathway is absent from *T. brucei* [43, 54]. If not nucleolar, response to stress also relocates cytoplasmic proteins to the nucleus. For example, transcription silencing with Actinomycin D causes the accumulation of *LmPABP2* and *LmPABP3* in the nucleoplasm [55], and a motif of *TcUBP-1* protein is responsible for the nuclear accumulation of this U-rich element binding protein under arsenite stress [56].

Like other members of the Rpp25-like Alba proteins, *LiAlba20* possesses an RGG box in its N-terminal part. In order to evaluate the role of this domain, we searched for proteins in the *Leishmania* proteome having this domain. In the human proteome, Corley *et al.* defined a RGG domain protein as any protein containing 3 "RGG" separated with 0 to 6 amino acids each or any protein with 2 "RGG" separated by 5 to 9 amino acids, containing at least one "RG" between the 2 "RGG" [57]. Searching for this protein motif pattern in *L. infantum* database using TrityDB.org [58] retrieves 27 putative proteins (Supplementary Table S2), including nucleolar-related proteins as fibrillarin (LinJ19.0090), nucleolar protein family A (LinJ.34.4120), nucleolar RNA helicase II (LinJ.05.0140) and RNA binding protein (LinJ.21.0600). These findings suggest a putative role of the RGG domain in the nucleolar localization of *LiAlba20* and *LiAlba13*. Because Alba domain proteins are constituents of the RNase P / MRP in higher eukaryotes, *LiAlba* proteins might play a role in tRNA or rRNA processing during differentiation. It has been shown that nuclear RNase P in *T. brucei* is constituted of a unique protein PROPR1 [59], and that downregulation of *TbAlba* proteins does not accumulate methionine or isoleucine tRNA precursors [14], hence excluding *LiAlba* proteins of being implicated in this process. The identification of subunits of the NTF-2 nuclear transport factor is nevertheless a good candidate for the nuclear import of *LiAlba* proteins. The fact that these proteins are identified as Alba partners in

both promastigotes and amastigotes suggest that shuttling of *LiAlba* proteins between the nucleolus and the cytoplasm might also occur in these life stages, but faster mRNA export to the cytoplasm could explain the lack of *LiAlba* detection in the nucleus. mRNA-dependent export of *LiAlba* proteins is also suggested by the diffused signal observed in the nucleus and the cytoplasm under conditions of translation inhibition.

The flagellum is another organelle shown to play an important role as sensor in stress response [60, 61]. In *T. brucei* bloodstream form, the four *TbAlba* domain proteins were identified as part of the flagellum-specific matrix proteome [62]. Given that RNA metabolism is unlikely to occur within the flagellum, two other hypotheses could be drawn. *LiAlba* proteins translocate to this organelle to be sequestered in order to allow differential regulation of cytoplasmic target mRNAs. Alternatively, Alba domain proteins acquire PTMs within the flagellum. In mammals, primary cilium is known to play an important role in the PTM of Gli, where phosphorylation of this factor provokes a downstream specific cleavage by the proteasome, leading to transcription inhibition (Reviewed in [63]). Interestingly, in *L. infantum* promastigotes, proteasome constitution analysis by Blue Native gel and MS/MS identified *LiAlba20* as part of the complex, suggesting a potential proteasome targeting of *LiAlba20* (Racine G. and Ouellette M., personal communication). In *T. gondii*, the Rpp20-like *TgAlba2* co-immunoprecipitates with various proteasome subunits, and comparison of the interacting partners with the yeast “translatome” structure, linking protein synthesis and degradation [64], implies a potential role of *TgAlba* in regulating translation [35]. Furthermore, in *Trypanosoma rangeli*, tyrosine phosphatase *TrPTP2* localizes in the flagellum [65]. This is not the case in *T. brucei* [66], but nevertheless, tyrosine-phosphorylated proteins seem to be enriched in the flagellum, basal body and nucleolus [33]. If tyrosine phosphorylation remains a rare event in *Trypanosoma*, it has been reported at least for the *T. cruzi* Alba domain proteins, allowing us to hypothesize that transient flagellar localization of *LiAlba* proteins permit phosphorylation of the protein.

In conclusion, the results presented here are in accordance with recently published data in other protozoan parasites where Alba domain proteins seem to play an important role in mRNA storage and translation initiation interface. Despite their constitutive

expression in *Leishmania*, Alba domain proteins may play distinct roles in promastigotes and amastigotes as also supported by the important differences in growth phenotype observed between wild type and *LiAlba20*<sup>-/-</sup> and *LiAlba13*<sup>-/-</sup> mutants, specifically in the amastigote stage. Translocation of *LiAlba* proteins from the cytoplasm to the flagellum and the nucleolus upon heat stress and during promastigote to amastigote differentiation underlines an important role, yet to be elucidated, that these proteins may play in regulating gene expression under those conditions. The example of nuclear accumulation of the Alba domain proteins in plants upon dehydration [67], implies that the role of Alba domain proteins in stress response may be conserved throughout evolution.

## **Materials and Methods**

### **Cell lines and culture**

*Leishmania donovani infantum* MHOM/MA/67/ITMAP-263 promastigotes were cultured at 25°C in SDM-79 medium (pH 7.0) supplemented with 10% heat-inactivated fetal calf serum (Multicell, Wisent Inc.) and 5 mg/ml hemin. Axenic amastigotes were grown in MAA/20 medium (pH 5.8) at 37°C with 5% CO<sub>2</sub>, and used after two to three passages, when fully differentiated. For transfectant selection, 30 µg/ml G418, 160 µg/ml hygromycin B and/or 70 µg/ml puromycin were used. Strains *LiAlba20*<sup>-/-</sup> (LinJ.34.2410) and *LiAlba13*<sup>-/-</sup> (LinJ.13.0270) and add-back mutants were described previously (Dupé A. *et al.*, 2013). Briefly, open reading frames of both *LiAlba13* and *LiAlba20* alleles were replaced with the coding sequences of puromycin and hygromycin. Add back control strains were obtained by episomal transfection of plasmids pSPαNEOα-HA-*LiAlba20* or pSPαNEOα-*LiAlba13*-HA into the corresponding null mutant background (described in Dupé A. *et al.*, 2013). For growth phenotype measurements, promastigote stationary phase parasites were counted on a Malassay. 10<sup>6</sup> parasites were then inoculated in 5 ml of fresh SDM media and O.D.<sub>600nm</sub> was checked each 24 hours. For axenic amastigotes, a similar approach was used but with 10<sup>7</sup> parasites as a starter. Experiments were done in triplicate, using passages 5 to 7 for promastigotes and passage 3 for amastigotes.

### **DNA manipulations and plasmid preparation**

All DNA manipulations were performed using the genomic DNA of the cultivated strain. For PCR amplifications, Phusion enzyme (Thermo Scientific) was used according to

the manufacturer instructions. Primers (from Integrated DNA Technologies) are listed in Table S1. PCR products were cloned in pGEM-T Easy vector (Promega) for sequencing and sub-cloning strategies. For fluorescent protein preparation, we developed a strategy to target any gene of the genome. Our strategy allows us to transfect episomal expression vectors or to integrate the cassettes of interest into endogenous locus to allow expression and regulation at levels similar to the native protein. For this, the fluorescent coding sequence (e.g. mCherry, yellow fluorescent protein (eYFP)) was fused at the N-terminal part of *LiAlba13*, *LiAlba20* or *LiNop10* to keep intact the 3'UTR and therefore regulation of the transcripts. Briefly, a first cassette was prepared by fusion PCR, containing a resistance gene, the  $\alpha$ -tubulin 3'UTR for optimal processing and expression of the marker and the fluorescent protein without the stop codon. NEO $\alpha$ -eYFP and HYG $\alpha$ -mCherry cassettes were cloned into pGEM-T Easy. The second step consists in the amplification of 500 nt of the 5' intergenic region of the targeted gene, and the amplification of at least 500 nt of the ORF (+ 3' intergenic region if required) without the ATG start codon (to avoid potential cleavage of the fluorescent protein by methionine aminopeptidases for example). These two fragments were fused by PCR and an *EcoRV* site was added at the junction. Cloning in *EcoRV* digested pSP72 and digestion using *EcoRV* allowed us to perform a blunt insertion of the cassette to finally obtain the pSPNEO $\alpha$ -eYFP-*LiAlba13*, pSPHYG $\alpha$ -mCh-*LiAlba20* and pSPHYG $\alpha$ -mCh-*LiNop10* vectors. Transfections were performed by electroporation as previously described [68].

### **Sequence alignments**

All the Alba protein sequences from trypanosomatids were downloaded from TritypDB.org (v4.0). For both *LiAlba* sequences, protein secondary structure prediction was performed using the Phyre webserver [69]. TriTryps Alba domain protein alignments were carried out with MEGA5 software using ClustalW (Gonnet matrix and 10 gaps allowed). Neighbor-joining tree was build using MEGA5 bootstrap method and Poisson model for substitutions. Phyre results and ClustalW alignment were merged using the BioEdit software.

### **Indirect immunofluorescence assay slide preparation**

Using episomal expression of pSP- $\alpha$ NEO $\alpha$ -*LiAlba13*-HA and pSP- $\alpha$ NEO $\alpha$ -HA-

*LiAlba20* in the null mutant background of the respective genes,  $\sim 10^6$  parasites were washed 3 times with PBS kept on ice before attachment on microscopy slides during 10 min at room temperature. Cold 100 % methanol was used for cell fixation 10 min at  $-20^\circ\text{C}$  and 3 x 5 min washes with PBS were performed prior to a 30 min permeabilization at room temperature in the following permeabilization solution (1X PBS, 0.2 % Triton X-100, 5 % Fetal Calf Serum). Blocking buffer (1X PBS; 2 % BSA; 0.02 % Tween-20; 0.1% Triton X-100) was added for overnight incubation at  $4^\circ\text{C}$ . Anti-HA (Applied Biological Materials Inc.) diluted 1:400 in blocking buffer was used as primary antibody in an 1 hour incubation at room temperature, and unbound antibodies were removed by 5 x 5 min washes at room temperature in PBS. 2  $\mu\text{g}$  of secondary antibody (Alexa Fluor 488 goat anti-mouse, LiveTech) were used in similar conditions in the dark and washed as described for the anti-HA antibody. For DNA staining, DAPI solution (Sigma) diluted to 0.5  $\mu\text{g}/\text{ml}$  in 1X PBS was added and incubated 5 min at room temperature and then washed 2 x 5min with 1X PBS. After drying the slides, a drop of mounting solution (Mobi Glow) and a cover slide was added and sealed with nail polish.

### **Fluorescent protein imaging and slide preparation**

Wild type cells were episomally transfected or co-transfected with pSPHYG $\alpha$ -mCh-*LiAlba20*, pSPNEO $\alpha$ -eYFP-*LiAlba13* and/or pSPHYG $\alpha$ -mCh-*LiNop10*. About  $10^6$  parasites were washed in PBS and incubated in a PBS + 2 % paraformaldehyde (Sigma) 10 min at room temperature for fixation. Cells were dropped on slide and dried prior DAPI coloration as described above. After slide drying, Mobi Glow and cover slide were added and sealed with nail polish. For protein expression inhibition studies, cells were treated for 1 hour or overnight with 50  $\mu\text{g}/\text{ml}$  cycloheximide or 25  $\mu\text{g}/\text{ml}$  Actinomycin D. All slides were analyzed by epifluorescence microscope using a Nikon Eclipse TE300 with a 100X objective and oil immersion. Images acquisition was performed with ImagePro Plus 5 software, and ImageJ or Photoshop were used for picture merging and montage.

### **Immunoprecipitation and MS/MS peptide identification**

For identifying interacting partners of the Alba domain proteins, we used wild type cells episomally transfected with pSP- $\alpha$ NEO $\alpha$ -*LiAlba13*-HA and pSP- $\alpha$ NEO $\alpha$ -HA-*LiAlba20*. For amastigotes and promastigotes,  $5 \cdot 10^9$  cells in exponential phase were

pelleted and resuspended in 1 ml of cytoplasmic extraction buffer (CEB: 10 mM Hepes, 3 mM MgCl<sub>2</sub>, 14 mM KCl, pH 7.5, 5% glycerol and 0.2% NP-40), supplemented with 1 mM PMSF and 2X cOmplete ULTRA Tablets Mini, EDTA-free EASYpack (Roche). Cells were lysed with 15 to 20 strokes using a Douncer and quality was assessed under microscope before centrifugation for 10 min at 4°C and 13 000 rpm to remove insoluble materials. To avoid further identification of proteins binding to the beads, protein extracts were washed by 1 hour incubation at 4°C on a rotor with 60 µl of Dynabeads Protein G (Invitrogen), resuspended in 200 µl PSB-0.1% Tween. In parallel, another 60 µl of beads were resuspended in 200 µl PSB-0.1% Tween, incubated with 10 µg of anti-HA antibody (Applied Biological Materials Inc.) during 1 hour at 4°C on a rotor. Unbound antibody was removed by three 5 min washes in 500 µl of PSB-0.1% Tween and the “washed“ proteins were added to the bead and antibody mix. Incubation was performed for 90 min at 4°C on a rotor and unbound material was removed by 5 washes using PSB-0.1% Tween supplemented with 1X cOmplete ULTRA Tablets Mini, EDTA-free EASYpack. The mix was transferred in a new Eppendorff tube and elution was performed by direct addition of 60 µl of 2x Laemmli Sample Buffer and incubation for 5 min at 100°C. Elutions were loaded on 12% SDS-PAGE and migration was stopped as soon as the bromophenol blue marker entered to the resolving part of the gel. After Coomassie staining and destaining (to remove traces of SDS), bands were excised and conserved in 1% acetic acid for MS/MS peptide identification. The results presented in Table 1 are only pulled down peptides with the HA-tagged proteins, and absent from the wild type negative control (data not shown). Immunoprecipitations from wild type, HA-*LiAlba20* and *LiAlba13*-HA were all run in triplicates.

### **Polysome Profiling Analysis**

Polysome profiling was done as previously described [37]. Briefly, 3×10<sup>9</sup> wild type parasites were treated with 100 µg/ml cycloheximide (Sigma) for 10 min and washed with cycloheximide-containing PBS buffer. Cell lysis was performed with a Dounce homogenizer in lysis buffer [10 mM Tris-HCl pH 7.4, 10 mM MgCl<sub>2</sub>, 150 mM NaCl, 0.5% IGEPAL, 100 µg/ml cycloheximide, 1 mM PMSF, 100 U/ml RNaseOUT (Amersham), 2X final concentration of protease inhibitor tablet (Roche)]. Lysates were divided into two tubes, and one of them was treated 10 min at room temperature with 50 mM EDTA as

control. All the lysates were pelleted by centrifugation and the supernatant (40 OD<sub>260 nm</sub> units) was layered on top of a 15% to 45% linear sucrose gradient (10 ml) prepared in gradient buffer (50 mM Tris-HCl pH 7.4, 10 mM MgCl<sub>2</sub>, 50 mM KCl, 100 U/ml RNaseOUT). Ultracentrifugation (35,000 rpm for 2.15 hours at 4°C) was performed in a Beckman SW40 Ti rotor and fractions were collected with an ISCO Density Gradient Fractionation System under constant monitoring of the absorbance at 254 nm. Proteins from the various fractions were concentrated using the TCA method, and finally resuspended in 50µl of 2X Laemmli buffer. 10µl of each fraction were loaded on 12% SDS-PAGE and transferred for further Western blot analysis. Anti-HA (100 ng/ml, Applied Biological Materials Inc.) or anti-*TbAlba2040/2* (GeneID: Tb927.4.2040, a kind gift from Dr Bastin [15], Pasteur Institute, Paris, France) and HRP-linked anti-mouse (NEB) antibodies were used for 1 hour at room temperature.

### **Acknowledgements**

AD was a recipient of a doctoral fellowship from the Canadian Institutes of Health Research (CIHR) Strategic Training Program STP-53924 and the Centre for Host-Parasite Interactions (CHPI). This study was supported by the CIHR operating grant MOP-12182 awarded to BP. BP is a member of a CIHR Group on Host-Pathogen Interactions and of the Centre for Host-Parasite Interactions 'Programme Regroupements Stratégiques' of the Fonds de Québec pour la Recherche sur la Nature et les Technologies.

## References

1. Herwaldt, B.L., *Leishmaniasis*. Lancet, 1999. **354**(9185): p. 1191-9.
2. Bates, P.A., *Transmission of Leishmania metacyclic promastigotes by phlebotomine sand flies*. Int J Parasitol, 2007. **37**(10): p. 1097-106.
3. Kaye, P. and P. Scott, *Leishmaniasis: complexity at the host-pathogen interface*. Nat Rev Microbiol, 2011. **9**(8): p. 604-15.
4. Zilberstein, D. and M. Shapira, *The role of pH and temperature in the development of Leishmania parasites*. Annu Rev Microbiol, 1994. **48**: p. 449-70.
5. Kramer, S., *Developmental regulation of gene expression in the absence of transcriptional control: the case of kinetoplastids*. Mol Biochem Parasitol, 2012. **181**(2): p. 61-72.
6. Haile, S. and B. Papadopoulou, *Developmental regulation of gene expression in trypanosomatid parasitic protozoa*. Curr Opin Microbiol, 2007. **10**(6): p. 569-77.
7. Clayton, C., *mRNA Turnover in Trypanosomes*. RNA Metabolism in Trypanosomes, 2012. **28**: p. 79.
8. Kramer, S., et al., *Inhibition of mRNA maturation in trypanosomes causes the formation of novel foci at the nuclear periphery containing cytoplasmic regulators of mRNA fate*. J Cell Sci, 2012. **125**(Pt 12): p. 2896-909.
9. Aravind, L., L.M. Iyer, and V. Anantharaman, *The two faces of Alba: the evolutionary connection between proteins participating in chromatin structure and RNA metabolism*. Genome Biol, 2003. **4**(10): p. R64.
10. Chene, A., et al., *PfAlbas constitute a new eukaryotic DNA/RNA-binding protein family in malaria parasites*. Nucleic Acids Res, 2011. **40**(7): p. 3066-77.
11. Goyal, M., et al., *Identification and molecular characterization of an Alba-family protein from human malaria parasite Plasmodium falciparum*. Nucleic Acids Res, 2011. **40**(3): p. 1174-90.
12. Mair, G.R., et al., *Universal features of post-transcriptional gene regulation are critical for Plasmodium zygote development*. PLoS Pathog, 2010. **6**(2): p. e1000767.
13. Alves, L.R., et al., *Proteomic analysis reveals the dynamic association of proteins with translated mRNAs in Trypanosoma cruzi*. Gene, 2010. **452**(2): p. 72-8.
14. Mani, J., et al., *Alba-domain proteins of Trypanosoma brucei are cytoplasmic RNA-binding proteins that interact with the translation machinery*. PLoS One, 2011. **6**(7): p. e22463.
15. Subota, I., et al., *ALBA proteins are stage regulated during trypanosome development in the tsetse fly and participate in differentiation*. Mol Biol Cell, 2011. **22**(22): p. 4205-19.
16. Ivens, A.C., et al., *The genome of the kinetoplastid parasite, Leishmania major*. Science, 2005. **309**(5733): p. 436-42.
17. Peacock, C.S., et al., *Comparative genomic analysis of three Leishmania species that cause diverse human disease*. Nat Genet, 2007. **39**(7): p. 839-47.
18. Raymond, F., et al., *Genome sequencing of the lizard parasite Leishmania tarentolae reveals loss of genes associated to the intracellular stage of human pathogenic species*. Nucleic Acids Res, 2012. **40**(3): p. 1131-47.
19. Chou, C.C., et al., *Crystal structure of the hyperthermophilic archaeal DNA-binding protein Sso10b2 at a resolution of 1.85 Angstroms*. J Bacteriol, 2003. **185**(14): p. 4066-73.

20. Zheng, W., *Sirtuins as emerging anti-parasitic targets*. Eur J Med Chem, 2013. **59**: p. 132-40.
21. Sereno, D., *et al.*, *Looking for putative functions of the Leishmania cytosolic SIR2 deacetylase*. Parasitol Res, 2006. **100**(1): p. 1-9.
22. Bell, S.D., *et al.*, *The interaction of Alba, a conserved archaeal chromatin protein, with Sir2 and its regulation by acetylation*. Science, 2002. **296**(5565): p. 148-51.
23. Nett, I.R., *et al.*, *The phosphoproteome of bloodstream form Trypanosoma brucei, causative agent of African sleeping sickness*. Mol Cell Proteomics, 2009. **8**(7): p. 1527-38.
24. Marchini, F.K., *et al.*, *Profiling the Trypanosoma cruzi phosphoproteome*. PLoS One, 2011. **6**(9): p. e25381.
25. Hem, S., *et al.*, *Identification of Leishmania-specific protein phosphorylation sites by LC-ESI-MS/MS and comparative genomics analyses*. Proteomics, 2010. **10**(21): p. 3868-83.
26. Morales, M.A., *et al.*, *Phosphoproteome dynamics reveal heat-shock protein complexes specific to the Leishmania donovani infectious stage*. Proc Natl Acad Sci U S A, 2010. **107**(18): p. 8381-6.
27. Schmidt-Arras, D., *et al.*, *Adaptation of a 2D in-gel kinase assay to trace phosphotransferase activities in the human pathogen Leishmania donovani*. J Proteomics, 2011. **74**(9): p. 1644-51.
28. Nebl, T., *et al.*, *Quantitative in vivo analyses reveal calcium-dependent phosphorylation sites and identifies a novel component of the Toxoplasma invasion motor complex*. PLoS Pathog, 2011. **7**(9): p. e1002222.
29. Solyakov, L., *et al.*, *Global kinomic and phospho-proteomic analyses of the human malaria parasite Plasmodium falciparum*. Nat Commun, 2011. **2**: p. 565.
30. Trecek, M., *et al.*, *The phosphoproteomes of Plasmodium falciparum and Toxoplasma gondii reveal unusual adaptations within and beyond the parasites' boundaries*. Cell Host Microbe, 2011. **10**(4): p. 410-9.
31. Dastidar, E.G., *et al.*, *Involvement of Plasmodium falciparum protein kinase CK2 in the chromatin assembly pathway*. BMC Biol, 2012. **10**: p. 5.
32. Kumarevel, T., *et al.*, *Crystal structure of an archaeal specific DNA-binding protein (Ape10b2) from Aeropyrum pernix K1*. Proteins, 2008. **71**(3): p. 1156-62.
33. Nett, I.R., *et al.*, *Identification and specific localization of tyrosine-phosphorylated proteins in Trypanosoma brucei*. Eukaryot Cell, 2009. **8**(4): p. 617-26.
34. Welting, T.J., *et al.*, *Heterodimerization regulates RNase MRP/RNase P association, localization, and expression of Rpp20 and Rpp25*. Rna, 2007. **13**(1): p. 65-75.
35. Gissot, M., *et al.*, *Toxoplasma gondii Alba Proteins Are Involved in Translational Control of Gene Expression*. J Mol Biol, 2013. **425**(8): p. 1287-301.
36. Beckham, C., *et al.*, *The DEAD-box RNA helicase Ded1p affects and accumulates in Saccharomyces cerevisiae P-bodies*. Mol Biol Cell, 2008. **19**(3): p. 984-93.
37. Padmanabhan, P.K., *et al.*, *Apoptosis-like programmed cell death induces antisense ribosomal RNA (rRNA) fragmentation and rRNA degradation in Leishmania*. Cell Death Differ, 2012. **19**(12): p. 1972-82.
38. Kramer, S., N.C. Kimblin, and M. Carrington, *Genome-wide in silico screen for CCCH-type zinc finger proteins of Trypanosoma brucei, Trypanosoma cruzi and Leishmania major*. BMC Genomics, 2010. **11**: p. 283.
39. Stewart, M., *et al.*, *Molecular mechanism of translocation through nuclear pore*

- complexes during nuclear protein import*. FEBS Lett, 2001. **498**(2-3): p. 145-9.
40. Fernandez-Moya, S.M., *et al.*, *Alterations in DRBD3 ribonucleoprotein complexes in response to stress in Trypanosoma brucei*. PLoS One, 2012. **7**(11): p. e48870.
  41. Robello, C., *et al.*, *Trypanosoma cruzi: molecular cloning of a gene coding for a putative vacuolar protein*. Exp Parasitol, 2000. **94**(2): p. 129-31.
  42. Daniels, J.P., K. Gull, and B. Wickstead, *Cell biology of the trypanosome genome*. Microbiol Mol Biol Rev, 2010. **74**(4): p. 552-69.
  43. Nazer, E. and D.O. Sanchez, *Nucleolar accumulation of RNA binding proteins induced by Actinomycin D is functional in Trypanosoma cruzi and Leishmania mexicana but not in T. brucei*. PLoS One, 2011. **6**(8): p. e24184.
  44. Ruan, J.P., E. Ullu, and C. Tschudi, *Characterization of the Trypanosoma brucei cap hypermethylase Tgs1*. Mol Biochem Parasitol, 2007. **155**(1): p. 66-9.
  45. Boisvert, F.M., *et al.*, *The multifunctional nucleolus*. Nat Rev Mol Cell Biol, 2007. **8**(7): p. 574-85.
  46. Olson, M.O., *Sensing cellular stress: another new function for the nucleolus?* Sci STKE, 2004. **2004**(224): p. pe10.
  47. Padmanabhan, P.K., *et al.*, *Novel features of a PIWI-like protein homolog in the parasitic protozoan Leishmania*. PLoS One, 2012. **7**(12): p. e52612.
  48. Cloutier, S., *et al.*, *Translational control through eIF2alpha phosphorylation during the Leishmania differentiation process*. PLoS One, 2012. **7**(5): p. e35085.
  49. Olguin-Lamas, A., *et al.*, *A novel Toxoplasma gondii nuclear factor TgNF3 is a dynamic chromatin-associated component, modulator of nucleolar architecture and parasite virulence*. PLoS Pathog, 2010. **7**(3): p. e1001328.
  50. Zhang, Q., *et al.*, *A critical role of perinuclear filamentous actin in spatial repositioning and mutually exclusive expression of virulence genes in malaria parasites*. Cell Host Microbe, 2011. **10**(5): p. 451-63.
  51. Kramer, S., *et al.*, *Heat shock causes a decrease in polysomes and the appearance of stress granules in trypanosomes independently of eIF2(alpha) phosphorylation at Thr169*. J Cell Sci, 2008. **121**(Pt 18): p. 3002-14.
  52. Kohl, L. and P. Bastin, *The flagellum of trypanosomes*. Int Rev Cytol, 2005. **244**: p. 227-85.
  53. Tschochner, H. and E. Hurt, *Pre-ribosomes on the road from the nucleolus to the cytoplasm*. Trends Cell Biol, 2003. **13**(5): p. 255-63.
  54. Nazer, E., R.E. Verdun, and D.O. Sanchez, *Nucleolar localization of RNA binding proteins induced by actinomycin D and heat shock in Trypanosoma cruzi*. PLoS One, 2011. **6**(5): p. e19920.
  55. da Costa Lima, T.D., *et al.*, *Functional characterization of three leishmania poly(a) binding protein homologues with distinct binding properties to RNA and protein partners*. Eukaryot Cell, 2010. **9**(10): p. 1484-94.
  56. Cassola, A. and A.C. Frasch, *An RNA recognition motif mediates the nucleocytoplasmic transport of a trypanosome RNA-binding protein*. J Biol Chem, 2009. **284**(50): p. 35015-28.
  57. Corley, S.M. and J.E. Gready, *Identification of the RGG box motif in Shadoo: RNA-binding and signaling roles?* Bioinform Biol Insights, 2008. **2**: p. 383-400.
  58. Aslett, M., *et al.*, *TriTrypDB: a functional genomic resource for the Trypanosomatidae*. Nucleic Acids Res, 2010. **38**(Database issue): p. D457-62.
  59. Taschner, A., *et al.*, *Nuclear RNase P of Trypanosoma brucei: a single protein in*

- place of the multicomponent RNA-protein complex. Cell Rep, 2012. 2(1): p. 19-25.*
60. Gluenz, E., M.L. Ginger, and P.G. McKean, *Flagellum assembly and function during the Leishmania life cycle. Curr Opin Microbiol, 2010. 13(4): p. 473-9.*
  61. Maric, D., C.L. Epting, and D.M. Engman, *Composition and sensory function of the trypanosome flagellar membrane. Curr Opin Microbiol, 2010. 13(4): p. 466-72.*
  62. Oberholzer, M., et al., *Independent analysis of the flagellum surface and matrix proteomes provides insight into flagellum signaling in mammalian-infectious Trypanosoma brucei. Mol Cell Proteomics, 2011. 10(10): p. M111 010538.*
  63. Goetz, S.C., P.J. Ocbina, and K.V. Anderson, *The primary cilium as a Hedgehog signal transduction machine. Methods Cell Biol, 2009. 94: p. 199-222.*
  64. Sha, Z., et al., *The eIF3 interactome reveals the translasome, a supercomplex linking protein synthesis and degradation machineries. Mol Cell, 2009. 36(1): p. 141-52.*
  65. Prestes, E.B., et al., *Trypanosoma rangeli protein tyrosine phosphatase is associated with the parasite's flagellum. Mem Inst Oswaldo Cruz, 2012. 107(6): p. 713-9.*
  66. Szoor, B., et al., *A novel phosphatase cascade regulates differentiation in Trypanosoma brucei via a glycosomal signaling pathway. Genes Dev, 2010. 24(12): p. 1306-16.*
  67. Choudhary, M.K., et al., *Dehydration-responsive nuclear proteome of rice (Oryza sativa L.) illustrates protein network, novel regulators of cellular adaptation, and evolutionary perspective. Mol Cell Proteomics, 2009. 8(7): p. 1579-98.*
  68. Papadopoulou, B., G. Roy, and M. Ouellette, *A novel antifolate resistance gene on the amplified H circle of Leishmania. Embo J, 1992. 11(10): p. 3601-8.*
  69. Kelley, L.A. and M.J. Sternberg, *Protein structure prediction on the Web: a case study using the Phyre server. Nat Protoc, 2009. 4(3): p. 363-71.*

## Figure Legends

### Figure 1. Alignments and phylogeny of Alba domain proteins from *Leishmania* spp., *T. brucei* and *T. cruzi*.

(A) Neighbour-joining tree showing the phylogenetic relationship between the Alba domain proteins of tritryps. Evolutionary distances (scale) were estimated as the number of amino acid substitutions per site, considering Poisson correction. The two subgroups Rpp20-like and Rpp25-like are marked. (B) ClustalW alignment of Rpp20-like Alba domain proteins merged with the secondary structure prediction of LinJ.34.2410 by Phyre. ss pred: Secondary structure *in silico* prediction. C: coil; H: helix; E: Sheet. Red squares indicate amino acids known to be phosphorylated. Black star shows the expected position for Sir2 acetylation and the red stars underline the signature motif of the subgroup. Sequence variations in coiled regions between *Trypanosoma* spp and *Leishmania* spp. are underlined with a black bar. LinJ.13.0270 was used for Phyre structure prediction. (C) As in B for the Rpp25-like Alba domain proteins. LinJ.34.2410 was used for Phyre structure prediction. Accession numbers define the strains: LinJ\*: *L. infantum*; LmjF\*: *L. major*; LtaP\*: *L. tarentolae*; LbrM\*: *L. braziliensis*; Tb\*: *T. brucei*; Tc\*: *T. cruzi*.

### Figure 2. Polysome profiling analysis revealed enrichment of the *Leishmania* Alba domain proteins in ribosomal subunits.

Polysome profile analysis of *L. infantum* wild type cells were performed using logarithmic promastigotes and axenic amastigote cells as described in Materials and Methods. Graphical representations present the RNA content of each collected fractions after ultracentrifugation on 15-45% sucrose gradient. F: Free RNA; 40S: 40S ribosomal subunit; 60S: 60S ribosomal subunit; M: monosome, and polysomes. Each fraction was loaded on 12% SDS-PAGE and transferred on membrane for Western blot analysis to detect the endogenous *LiAlba20* and *LiAlba13-HA*. Anti-HA or anti-*TbAlba2040/2* antibodies were used to detect *LiAlba13* and *LiAlba20* proteins, respectively. For each growth phase, half of the protein extracts were incubated with EDTA before ultracentrifugation to disrupt mRNA and ribosomes as a control.

**Figure 3. Immunolocalization of *L. infantum* recombinant Alba domain proteins in promastigotes and during promastigote to amastigote differentiation.**

*LiAlba20*-HA (A) and HA-*LiAlba13* (B) proteins were episomally transfected in *L. infantum* and immunolocalization was performed with an anti-HA antibody as described in Materials and Methods. For (A) and (B), the two upper panels represent the localization of *LiAlba* proteins in exponentially grown promastigotes, and the three lower panels represent their localization after overnight differentiation from promastigotes to amastigotes. The differentiation was induced by transferring parasites from SDM media pH 7.0 at 25°C to MAA media pH 5.8 at 37°C. DAPI staining presents kinetoplastid DNA (small and intense spot) and nucleus (larger and more diffuse area).

**Figure 4. Alba domain proteins systematically co-localize during the *Leishmania* lifecycle and during differentiation.**

*LiAlba* proteins co-localize during the parasite life cycle. Wild type cells were episomally co-transfected with pSPNEO $\alpha$ -eYFP-*LiAlba13* and pSPHYG $\alpha$ -mCh-*LiAlba20* (see Supplementary Figure S1). Exponentially grown promastigotes, overnight differentiated parasites to axenic amastigotes (as described in Figure 3) and fully differentiated amastigotes at passage 4 (~12 days of growth) are compared here.

**Figure 5. Nucleolar and flagellar localization during amastigote differentiation is induced upon heat stress.**

Wild type cells were co-transfected with pSPNEO $\alpha$ -eYFP-*LiAlba13* and pSPHYG $\alpha$ -mCh-*LiNop10*. *LiNop10* is used as a nucleolar marker (see Supplementary Figure S2). (A) Exponentially grown promastigotes, overnight differentiating parasites and amastigotes at passage 4 are compared, as described in the previous figures. (B) Localization experiments of the same strain under independent acidic and heat stress. Exponential promastigotes grown in SDM+10% SFB at 25°C were submitted to an overnight temperature stress (37°C), or an acidification of the media (SDM+10% SFB acidified to pH 5.8 with HCl).

**Figure 6. Genetic depletion of *LiAlba20* or *LiAlba13* proteins affects significantly parasite growth in the amastigote stage.**

For promastigotes (A), 10<sup>6</sup> promastigotes (passage 5-7) were inoculated in 5 ml of SDM at

25°C and O.D.<sub>600nm</sub> was measured each 24 hours. For axenic amastigotes (B), 10<sup>7</sup> parasites at passage 2 were inoculated in fresh media at 37°C. To avoid any side effect of puromycin and hygromycin used for the gene depletion, stable null mutant clones were passed without these drugs. Neomycin was kept at usual concentration for stabilization of the episomal add-back strains.

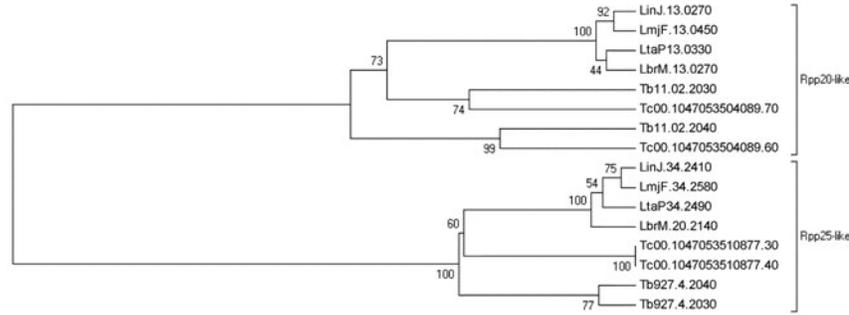
## Tables

**Table 1. *LiAlba* domain protein partners identified by co-immunoprecipitation and MS/MS peptide identification.**

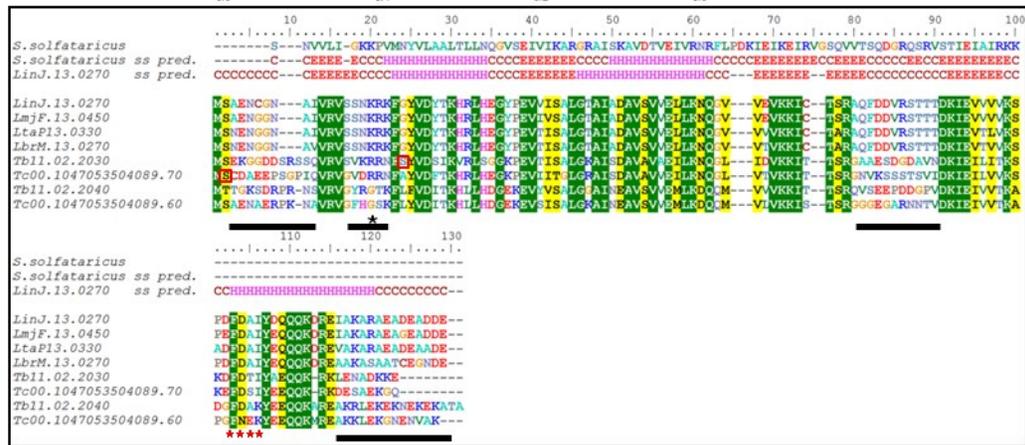
<i>Gene ID</i>	<i>Identified Proteins</i>	<i>M.W.</i>	<i>Promastigote</i>		<i>Amastigote</i>	
			<i>Alba13-HA</i>	<i>HA-Alba20</i>	<i>Alba13-HA</i>	<i>HA-Alba20</i>
LinJ.34.2410	hypothetical protein, conserved   <b>Alba20</b>	22 kDa	14	15	15	16
LinJ.13.0270	hypothetical protein, conserved   <b>Alba13</b>	13 kDa	8	7	8	6
LinJ.32.0410	ATP-dependent RNA helicase, putative   <b>Homologue of Dbp1</b>	67 kDa	3	4	12	1
LinJ.35.5360	polyadenylate-binding protein 1, putative   <b>PABP1</b>	63 kDa	4	3	7	3
LinJ.35.4200	poly(A)-binding protein 2,poly(a) binding protein, putative   <b>PABP2</b>	65 kDa	12	15	17	12
LinJ.25.0080	poly(A)-binding protein 3   <b>PABP3</b>	61 kDa	2	6	7	7
LinJ.32.0790	RNA binding protein, putative   <b>NRBD -&gt; 5S rRNA</b> (Tryp. specific)	25 kDa	4	3	4	6
LinJ.21.0490	hypothetical protein, conserved   <b>NTF2-like</b>	47 kDa	1	3	9	9
LinJ.27.1220	hypothetical protein, conserved   <b>ZC3H41</b>	60 kDa		2	13	9
LinJ.26.1220	heat shock protein 70-related protein	70 kDa	2	2	9	4
LinJ.36.3010	40S ribosomal protein S24e	16 kDa	4		3	3
LinJ.35.0400	40S ribosomal protein S3A, putative	30 kDa	2	1	3	2
LinJ.36.5240	40S ribosomal protein SA, putative	28 kDa	4	1	1	2
LinJ.28.1050	40S ribosomal protein S14	16 kDa	2		3	3
LinJ.21.2150	40S ribosomal protein S6, putative	28 kDa	3	1	2	3
LinJ.07.0760	40S ribosomal protein S9, putative	22 kDa	1	2	1	2
LinJ.19.0050	40S ribosomal protein S2	29 kDa	2		2	1
LinJ.15.1060	60S ribosomal protein L6, putative	21 kDa	2	1	2	1
LinJ.07.0550	60S ribosomal protein L7a, putative	39 kDa	3	2	2	3
LinJ.29.1160	ribosomal protein L1a, putative	41 kDa	2	2	3	3
LinJ.09.1180	mitochondrial RNA binding protein 2,MRP2, gBP25	27 kDa	2	1	6	4
LinJ.27.0980	mitochondrial RNA binding protein 1,gBP21, MRP1	27 kDa	2	2	4	4
LinJ.35.2240	RNA-binding protein, putative   <b>DRBD2</b>	30 kDa			7	4
LinJ.18.0300	hypothetical protein, conserved   <b>NTF2-like</b>	40 kDa			2	
LinJ.19.0020	hypothetical protein, conserved	28 kDa			2	1
LinJ.25.0550	hypothetical protein SCD6.10   <b>Rap-55</b>	32 kDa			2	1
LinJ.22.1370	40S ribosomal protein L14, putative	26 kDa			3	
LinJ.11.1180	40S ribosomal protein S15A, putative	15 kDa			1	2
LinJ.28.2750	40S ribosomal protein S17, putative	16 kDa			3	3
LinJ.36.0990	40S ribosomal protein S18, putative	19 kDa			1	2
LinJ.34.2620	40S ribosomal protein S19 protein, putative	20 kDa			2	2
LinJ.21.1300	40S ribosomal protein S23, putative	16 kDa			1	2
LinJ.15.1010	40S ribosomal protein S3, putative	24 kDa			6	6
LinJ.13.1120	40S ribosomal protein S4, putative	31 kDa			7	4
LinJ.11.0960	40S ribosomal protein S5	21 kDa			3	1
LinJ.04.0750	60S ribosomal protein L10, putative	25 kDa			1	2
LinJ.18.0630	60S ribosomal protein L10a, putative	25 kDa			1	2
LinJ.04.0460	60S ribosomal protein L11 (L5, L16)	22 kDa			3	1
LinJ.29.2570	60S ribosomal protein L13, putative	25 kDa			1	3
LinJ.32.4050	60S ribosomal protein L2, putative	28 kDa			1	4
LinJ.16.0470	60S ribosomal protein L21, putative	18 kDa			1	2
LinJ.06.0590	60S ribosomal protein L23a, putative	16 kDa			2	2
LinJ.24.2140	60S ribosomal protein L26, putative	16 kDa			1	3
LinJ.35.3810	60S ribosomal protein L27A/L29, putative	16 kDa			1	1
LinJ.35.2040	60S ribosomal protein L32	15 kDa				2
LinJ.35.1870	60S ribosomal protein L5, putative	34 kDa			3	2
LinJ.26.0150	60S ribosomal protein L7, putative	31 kDa			1	1
LinJ.30.3390	60S ribosomal protein L9, putative	21 kDa			4	4
LinJ.36.1130	ribosomal protein L24, putative	15 kDa			1	2
LinJ.20.1620	ribosomal protein S11 homolog	16 kDa			1	2
LinJ.28.1100	ribosomal protein S20, putative	13 kDa			1	2
LinJ.01.0430	ribosomal protein S7, putative	24 kDa			4	3
LinJ.25.1850	3-oxo-5-alpha-steroid 4-dehydrogenase, putative	34 kDa			2	1
LinJ.30.2570	reticulon domain protein, 22 kDa potentially aggravating protein	22 kDa			2	1
LinJ.29.2620	ATP-dependent phosphofructokinase	54 kDa	7	1		
LinJ.19.1590	inosine-5'-monophosphate dehydrogenase	56 kDa	2	2		
LinJ.23.0860	3-ketoacyl-coa thiolase-like protein	47 kDa		2		
LinJ.18.0510	aconitase, putative	97 kDa		2		
LinJ.05.0960	dipeptidyl-peptidase III, putative,metallo-peptidase, Clan M-, Family M49	76 kDa		3		
LinJ.24.2150	transketolase	72 kDa		2		
LinJ.11.0100	seryl-tRNA synthetase, putative	53 kDa		2		
LinJ.18.0590	RNA binding protein, putative   Leish. Specific	44 kDa	2			
LinJ.32.0930	60S ribosomal protein L18a, putative	21 kDa	2			
LinJ.36.2130	chaperonin Hsp60, mitochondrial precursor	61 kDa		3		

# Figures

**A**



**B**



**C**

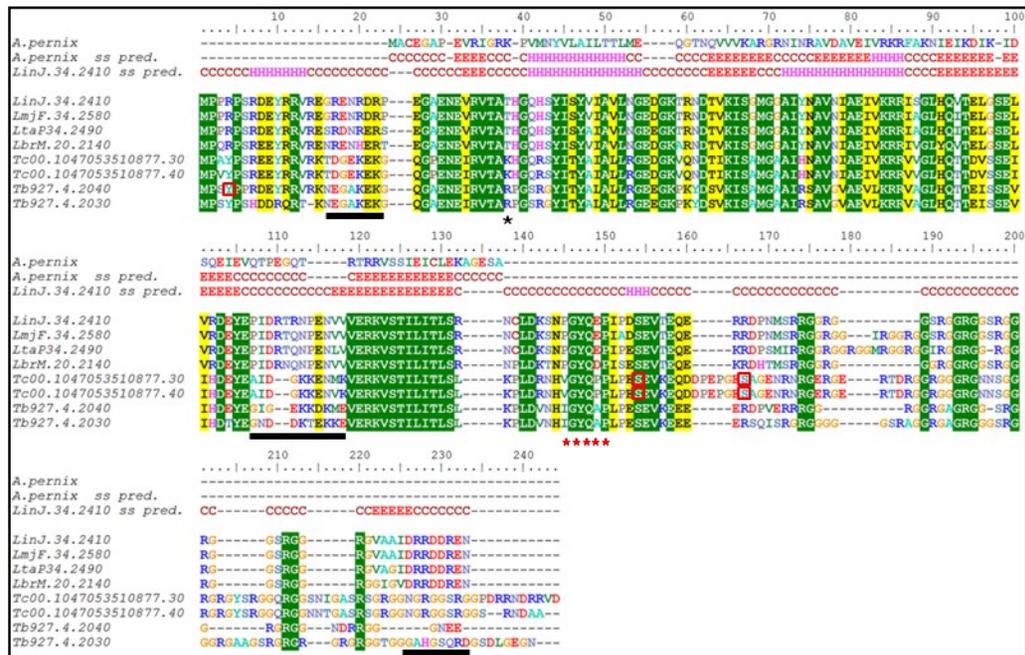
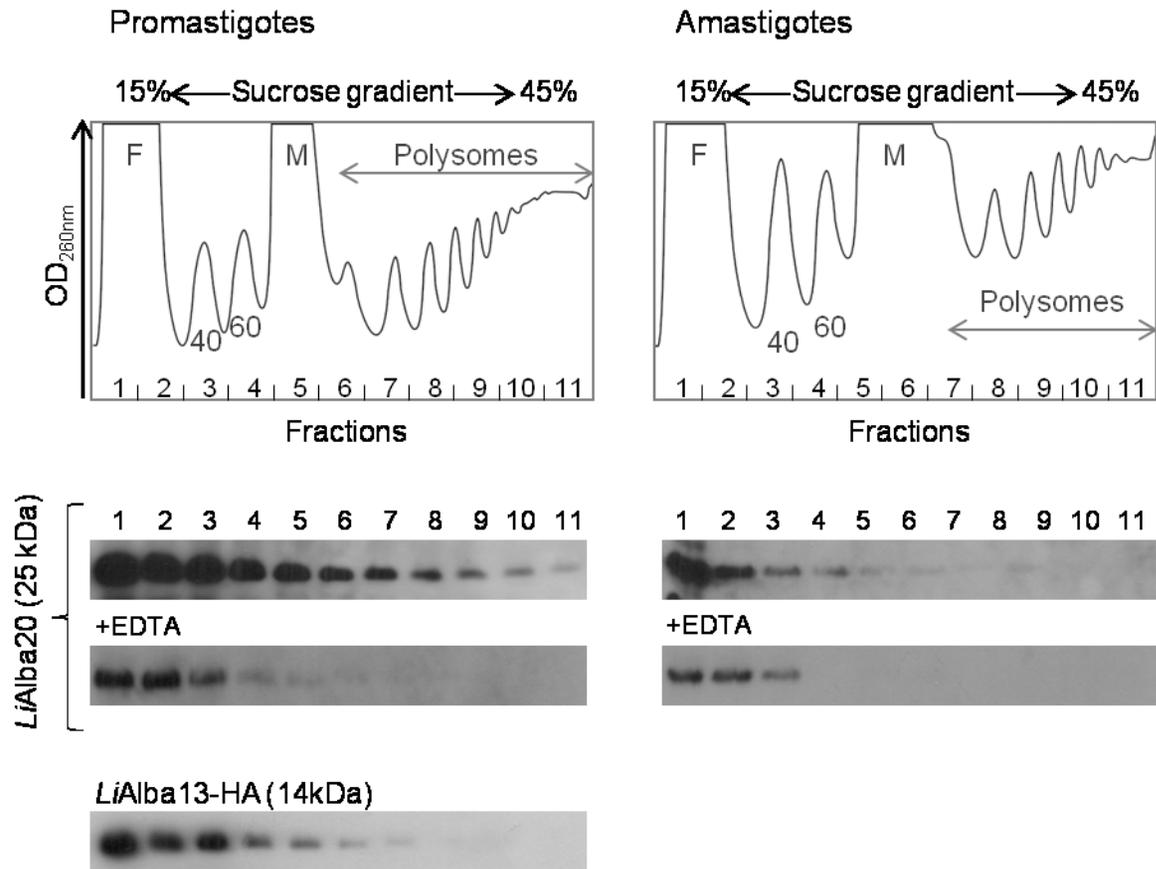


Figure 1.



**Figure 2.**

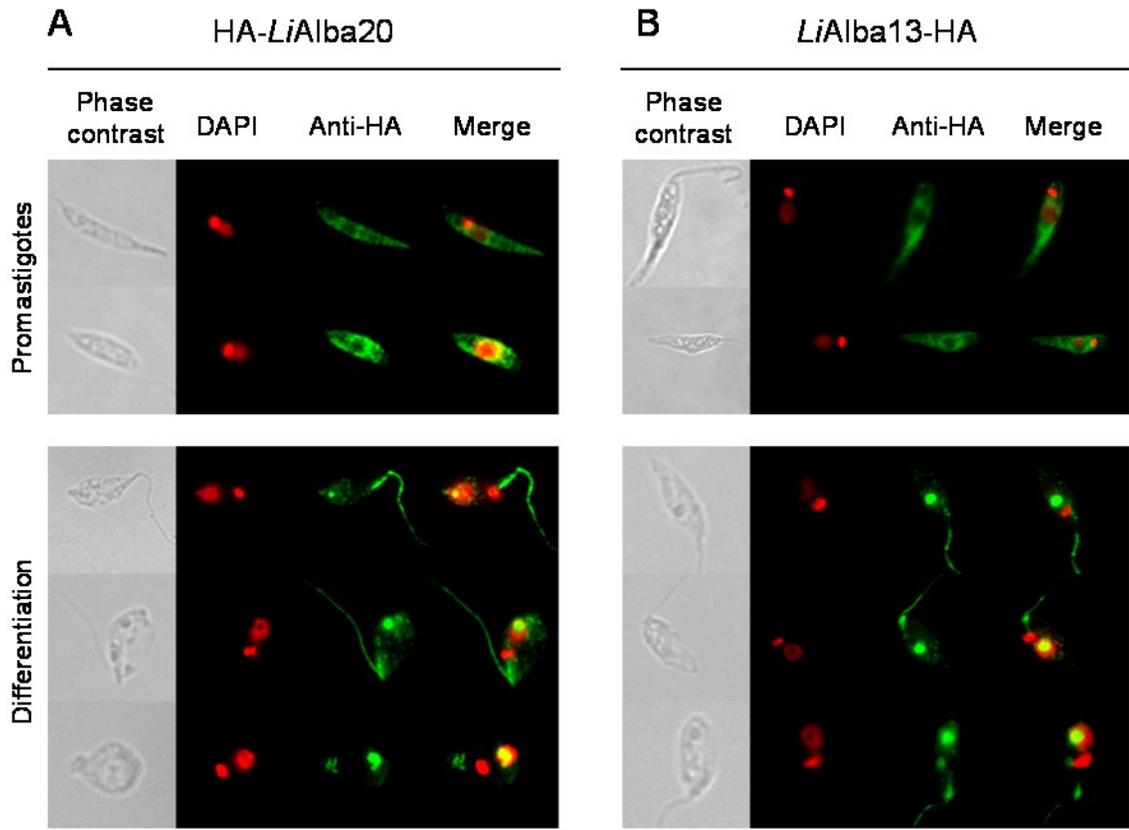
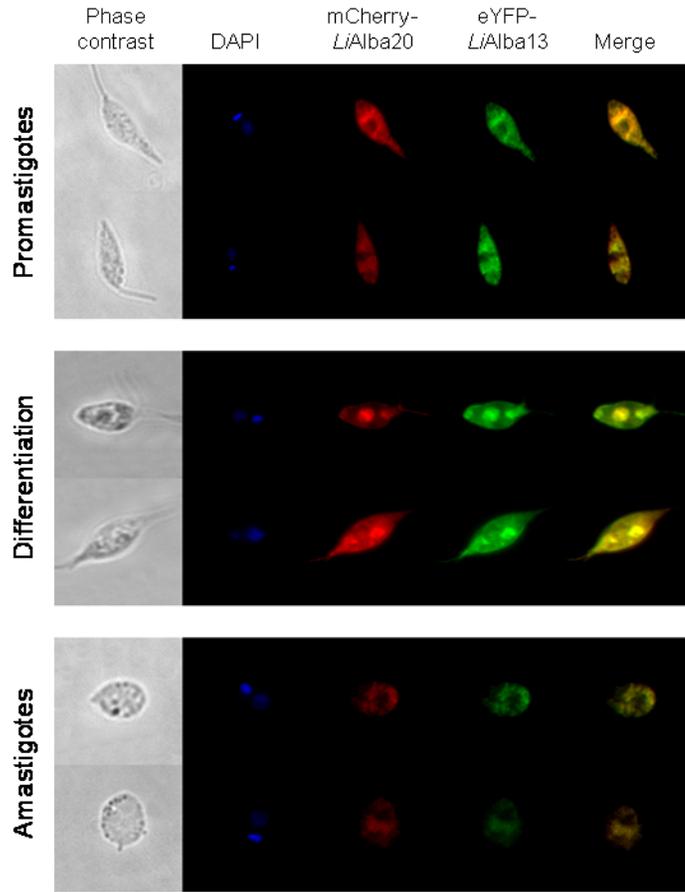
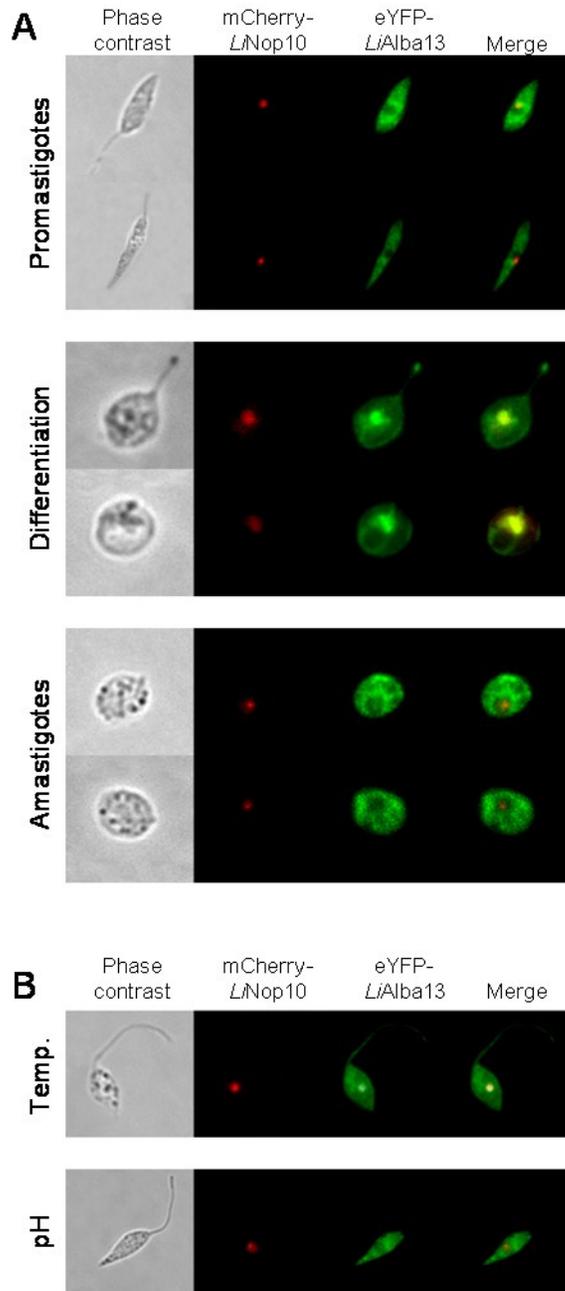


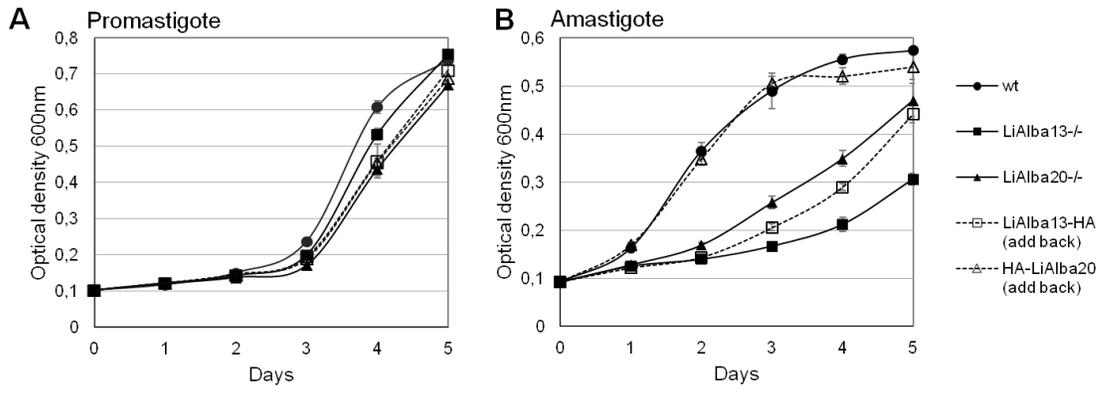
Figure 3.



**Figure 4.**



**Figure 5.**



**Figure 6.**

## Supplementary data

**Table S1. List of primers used in this study**

Name	Sequence 5'→3'	Target plasmid or gene	Product size	Fusion size
<b>For pGEM-NEOalpha-YFP preparation</b>				
Neo F	ATGATTGAACAAGATGGATT	pSPaNEOa	795 bp	2261 bp
Neo R	TCAGAAGAACTCGTCAAGAA			
alpha F (+Neo)	TTCTTGACGAGTCTTCTGATCGAGGGAGGAGGAGGAGGC	pSPaNEOa (LinJ.13.0330 3'IR)	789 bp	
alpha R (+GFP/YFP/CFP)	TCCTCGCCCTTGCTCACCATTCCAAGGAAGAGAGAAAGGT			
YFP F / GFP F / CFP F	ATGGTGAGCAAGGGCGAGGA	peYFP-N1 (without stop codon) (or pSPaNEOa-eGFP)	717 bp	
YFP R / GFP R / CFP R	CTTGACAGCTCGTCCATGC			
<b>For pGEM-HYGalpha-mCh preparation</b>				
Hyg F	ATGAAAAGCCTGAACTCAC	pSPaHYGa	1026 bp	2483 bp
Hyg R	CTATTCCCTTGCCTCGGAC			
alpha F (+Hyg)	GTCCGAGGGCAAAGGAATAGTCGAGGGAGGAGGAGGAGGC	pSPaNEOa (LinJ.13.0330 3'IR)	789 bp	
alpha R (+mCh)	TCCTCGCCCTTGCTCACCATAAGGAAGAGAGAAAGGTGAT			
mCh F	ATGGTGAGCAAGGGCGAGGA	pC1-mCherry (without stop codon)	708 bp	
mCh R	CTTGACAGCTCGTCCATGC			
<b>For pGEM-PURalpha-CFP preparation (prepared, but not used in this study)</b>				
Pur F	ATGACCAGTACAAGCCACGGTGCCTCGCCAC	pSPaPURA	600 bp	2066 bp
Pur R	TCAGGCACCGGGCTTGGGG			
alpha F (+Pur)	CCCACAAGCCCGGTGCCTGATCGAGGGAGGAGGAGGAGGC	pSPaNEOa (LinJ.13.0330 3'IR)	789 bp	
alpha R (+GFP/YFP/CFP)	TCCTCGCCCTTGCTCACCATTCCAAGGAAGAGAGAAAGGT			
YFP F / GFP F / CFP F	ATGGTGAGCAAGGGCGAGGA	peCFP-N1 (without stop codon)	717 bp	
YFP R / GFP R / CFP R	CTTGACAGCTCGTCCATGC			
<b>For LiAlba20 targetting (pSP-5'IR-EcoRV-ORF+3'IR)</b>				
5'Alba20 F	GCACGTCCTTCTGTTCCCCCATGC	LinJ.34.2410 (5'IR)	506 bp	1600 bp
5'Alba20 R EcoRV	GATATCTGCTGCTTAGCGGGTTTCC			
ORF Alba20 F EcoRV 5'IR	GGAAAACCCGCTAAGCAGCAGATATCCCGCCACGTCGGTCTGTGA	LinJ.34.2410 ORF (-ATG, 594 bp) + 3'IR (500 bp)	1120 bp	
3'Alba20 R	AAGACACACATTGATACGGAAGC			
<b>For LiAlba13 targetting (pSP-5'IR-EcoRV-ORF+3'IR)</b>				
5'Alba13 F	CCATTTATCGAAGGGTTCC	LinJ.13.0270 (5'IR)	506 bp	1369 bp
5'Alba13 R EcoRV	GATATCCTTGGTACAAGGAAAAGGTAG			
ORF Alba13 F EcoRV 5'IR	CTACCTTTTCCTTTGTACCAAGGATATCTCAGCTGAAAACGCGGAAACG	LinJ.13.0270 ORF (-ATG, 363 bp) + 3'IR (500 bp)	889 bp	
3'Alba13 R	GTGTAGCATTGCTTTTTTCA			
<b>For Nop10 targetting (pSP-5'IR-EcoRV-ORF+3'IR)</b>				
5'Nop10 F	TGCGACCTGCTGCTGCTTGG	LinJ.36.0370 (5'IR)	506 bp	1006 bp
5'Nop10 R EcoRV	GATATCGCTTGGACTGAGTATATGTG			
ORF Nop10 F EcoRV 5'IR	CACATATACTCAGTCCAAGCGATATCCACCTTCGAGTGATACATGGT	LinJ.36.0370 ORF (-ATG, 189 bp) + 3'IR (311 bp)	526 bp	
3'Nop10 R	TGTGGCACCGGTGCAAAAG			

**Table S2. List of RGG-box-containing proteins in *L. infantum*.**

Gene ID	Description	RGGbox Count	RGGbox Locations	Length (a.a)	M.W. (kDa)	pI
LinJ.05.0140	nucleolar RNA helicase II, putative	1	(658-672)	2037	73811	8.05
LinJ.16.1110	hypothetical protein, conserved	1	(451-467)	1551	56373	8.8
LinJ.19.0050	40S ribosomal protein S2	1	(26-38)	798	28760	10.06
LinJ.19.0090	fibrillarin, putative	3	(2-18), (21-35), (41-50)	885	30959	10.24
LinJ.19.0440	nucleosome assembly protein, putative	1	(335-347)	1053	39792	4.28
LinJ.21.0600	Ia RNA binding protein, putative	1	(318-333)	1011	36869	9.37
LinJ.24.0800	hypothetical protein, conserved	1	(259-270)	933	32837	10.26
LinJ.25.1120	hypothetical protein, conserved	1	(391-400)	1347	48287	6.29
LinJ.25.1600	hypothetical protein, conserved	1	(6-19)	2241	80271	7.5
LinJ.25.2060	hypothetical protein, conserved	1	(1356-1367)	4425	156045	10.38
LinJ.27.1220	hypothetical protein, conserved	1	(96-116)	1590	59994	8.37
LinJ.28.0900	RNA binding protein rbp16, putative	1	(104-117)	432	15241	10.55
LinJ.30.0810	hypothetical protein, conserved	1	(877-892)	3177	116804	7.25
LinJ.30.2600	RNA-binding protein, putative	1	(203-215)	651	23110	7.6
LinJ.31.0090	hypothetical protein, conserved	1	(97-106)	696	25600	9.07
LinJ.31.1770	nucleosome assembly protein-like	1	(375-389)	1212	44789	4.25
LinJ.31.2190	hypothetical protein, conserved	1	(11-27)	2790	103180	7.94
LinJ.32.0460	40S ribosomal protein S2	1	(25-37)	795	28644	10.06
LinJ.32.0890	hypothetical protein, conserved	2	(4-12), (248-260)	1536	56062	6.92
LinJ.32.1880	hypothetical protein, conserved	1	(188-202)	735	26202	11.07
LinJ.34.2410	hypothetical protein, conserved	1	(161-177)	597	21993	9.93
LinJ.34.4120	nucleolar protein family a, putative	3	(13-26), (27-38), (144-160)	636	21078	12.01
LinJ.35.3100	hypothetical protein, conserved	1	(143-158)	1224	43067	6.02
LinJ.35.3150	ATP-dependent RNA helicase, putative	2	(71-86), (264-275)	2775	100539	6.54
LinJ.35.4450	hypothetical protein, conserved	5	(436-453), (455-472), (478-495), (497-515), (517-534)	1701	59511	10.15
LinJ.35.5000	hypothetical protein, conserved	1	(983-994)	3069	106413	9.13
LinJ.36.5330	hypothetical protein, conserved	1	(211-224)	2889	103465	7.24

**RGG box definition:** Corley & Gready, 2008 *Identification of the RGG Box Motif in Shadoo: RNA-Binding and Signaling Roles?* Bioinformatics and Biology Insights

**Motif search** (Perl regular expression):

RGG.{0,6}RGG.{0,6}RGG|RGGRG.{3,7}RGG|RGG.RG.{2,6}RGG|RGG..RG.{1,5}RGG|RGG...RG.{0,4}RGG|RGG....RG.{0,3}RGG|RGG.....RG.{0,2}RGG|RGG.....RG.{0,1}RGG|RGG.....RGRGG

*That means: Any protein containing 3 "RGG" separated with 0 to 6 a.a. each OR any protein with 2 "RGG" separated by 5 to 9 a.a., containing at least one "RG" between the 2 "RGG"*

**Database:** Tritypdb.org v4.0

## Supplementary Figure Legends

### **Supplementary Figure S1. Schematic representation of the plasmids used for fluorescent protein transfections.**

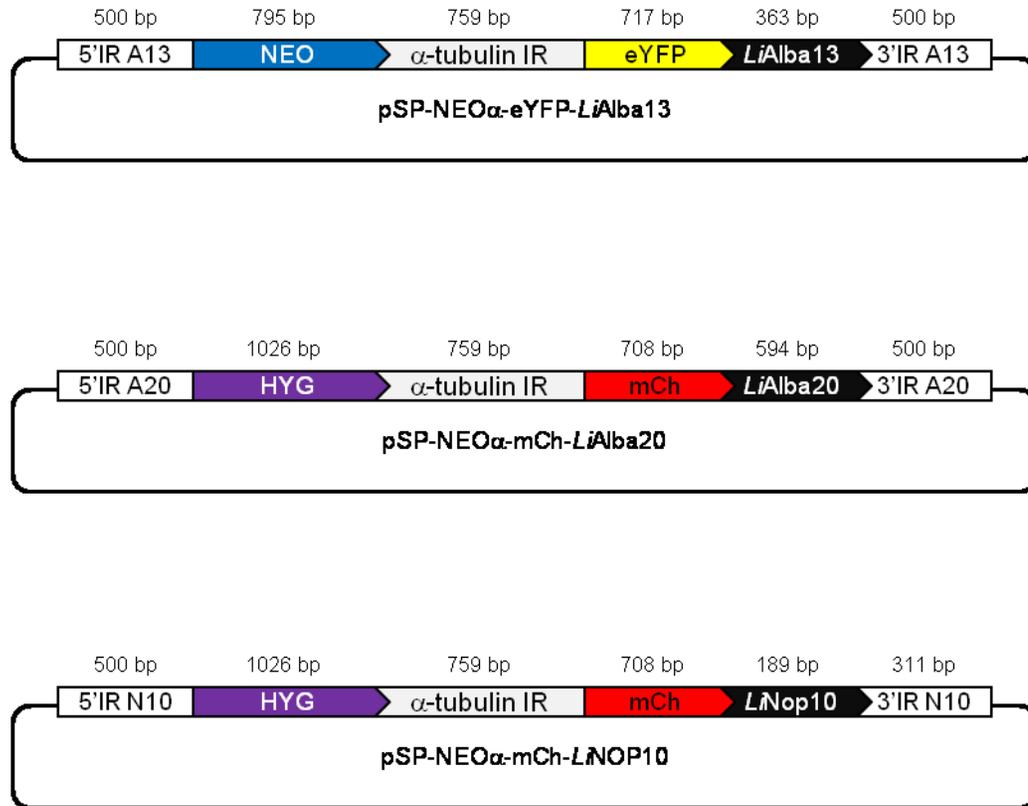
The primers used for plasmid preparation are listed in Table S1 and the cloning strategy used is the same for the 3 vectors. Briefly, 500nt from the 5' intergenic region (5'IR) of the target gene was PCR-amplified and fused to the ORF together with 300 to 500 additional nucleotides of the 3'IR. The fused product contains an *EcoRV* site that replaces the ATG start codon of the gene. This sequence was cloned into the *EcoRV* site of a native pSP72 vector. In the resulted plasmid, *EcoRV* digestion allows a single cleavage just upstream of the ORF. The cassette containing the selection marker, the alpha tubulin IR and a fluorescent protein without the stop codon was prepared by fusion PCR (i.e. HYG $\alpha$ -mCh or NEO $\alpha$ -eYFP) and cloned into the *EcoRV* site in front of the ORF.

### **Supplementary Figure S2. Confirmation of the nucleolar localization of mCh-LiNop10 in promastigotes, amastigotes and during differentiation.**

pSPHYG $\alpha$ -mCh-LiNop10 was episomally transfected in *L. infantum* wild type cells and the signal was compared to the DAPI staining. LiNop10 always localizes as a single dot within the nucleus, which often overlays with a region with weaker DAPI staining corresponding to the nucleolus, mainly observable in promastigotes.

### **Supplementary Figure S3. Effects of transcription or translation inhibition on LiAlba13 localization.**

Exponentially grown promastigotes expressing mCh-LiNop10 and eYFP-LiAlba13 were treated with actinomycin-D (transcription inhibition, middle panels) or cycloheximide (translation elongation inhibition, lower panels). The results presented here were obtained after overnight drug treatments. Similar localizations were visible after 6h of treatment. One hour of treatment did not show any variation compared to promastigote cells.



**Figure S1.**

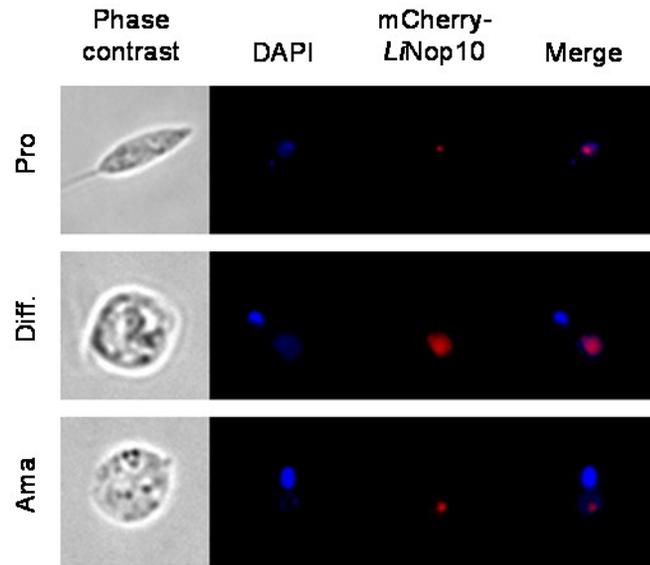
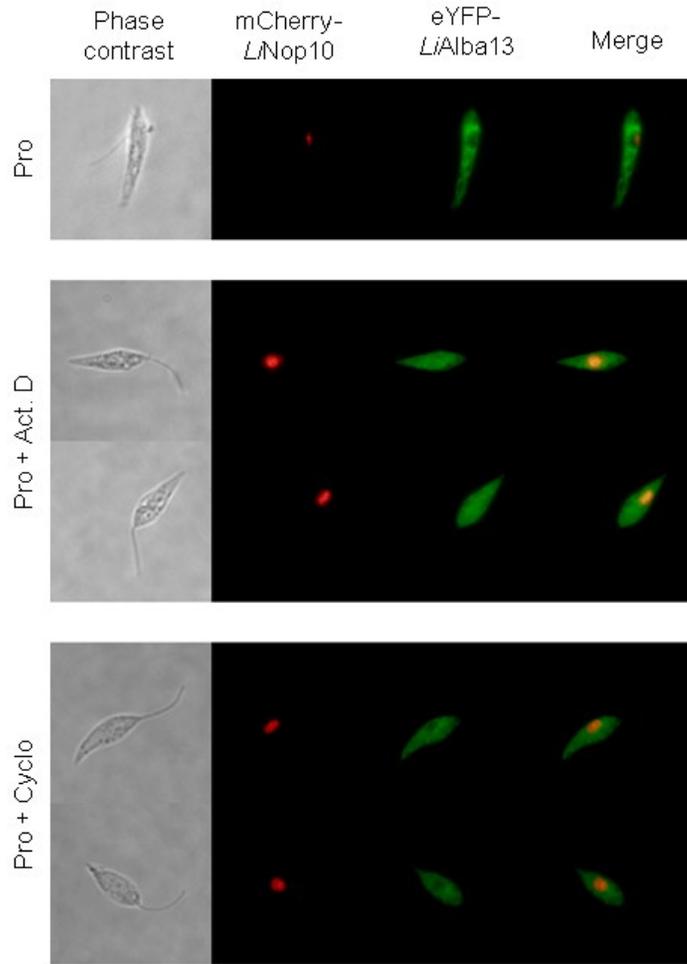


Figure S2.



**Figure S3.**

## Chapitre 11 : Discussion générale

Nos travaux portent sur la régulation d'une  $\delta$ -amastine, un facteur de virulence exprimé spécifiquement au stade amastigote du parasite *Leishmania*. Chez les *Trypanosomatidae*, le mécanisme de la transcription en absence de séquences promotrices typiques de l'ARN polymérase II est encore peu compris, et la transcription de type polycistronique de gènes codant pour des fonctions distinctes exclut la régulation de l'expression génique à cette étape (voir Chapitre 4). Nous décrivons ici la présence d'une séquence de 100 nt riche en uridines (URE), présente dans le 3'UTR du transcrit d'une  $\delta$ -amastine (LinJ.34.1010, voir Chapitre 5) qui est responsable de la dégradation du transcrit spécifiquement au stade promastigote et fait intervenir un mécanisme de dégradation indépendant de la déadénylation (Chapitre 8). Nous avons également identifié la protéine *LiAlba20* (voir Chapitre 6) comme facteur interagissant avec le 3'UTR de l'amastine LinJ.34.1010 dans une région proche de l'URE. L'étude de *LiAlba20* a montré qu'elle est impliquée dans la stabilisation du transcrit de l'amastine spécifiquement au stade amastigote. De plus, une approche d'hybridation d'ADNc sur puce à ADN a mis en évidence une diminution significative de l'accumulation d'autres  $\delta$ -amastines, chez les amastigotes, en absence de la protéine *LiAlba20*. Ceci suggère un phénomène de régulation commun à cette famille (Chapitre 9). Si l'activité de la protéine *LiAlba20* sur les transcrit des  $\delta$ -amastines diffère entre les stades amastigotes et promastigotes, elle est néanmoins exprimée constitutivement dans ces deux stades. Afin de compléter cette étude, nous avons donc vérifié la localisation de *LiAlba20* ainsi que ses partenaires protéiques durant le cycle parasitaire. Les protéines qui interagissent avec *LiAlba20* sont majoritairement les mêmes et la protéine est cytoplasmique dans les deux stades de vie. Cependant, lors de la différenciation du parasite du stade promastigote vers le stade amastigote, la localisation de *LiAlba20* diffère car elle présente une accumulation flagellaire et nucléolaire (Chapitre 10). Deux protéines à domaine Alba sont présentes dans le génome de *Leishmania*. Nous avons donc également analysé le rôle de *LiAlba13* dans la régulation du transcrit de l'amastine, ainsi que ses complexes protéiques et sa localisation durant le cycle cellulaire. Les deux protéines à domaine Alba co-localisent en tout temps dans la cellule, et sont retrouvées

dans des complexes identiques, cependant nos travaux montrent que *LiAlba13* n'intervient pas dans la régulation de la stabilité du transcrit de la  $\delta$ -amastine LinJ.34.1010, suggérant que ces protéines interagissent ensemble, mais ont des rôles différents.

Plusieurs points sont à souligner dans ces travaux :

- Dans le cas du transcrit de l'amastine, la présence du URE provoque une dégradation du transcrit qui est indépendante de la déadénylation. Chez les autres organismes, les ARE et URE dirigent les transcrits vers des processus de dégradation dépendants de la déadénylation. Le mécanisme décrit pour l'amastine chez *Leishmania* est donc atypique.
- Aujourd'hui, les études effectuées sur les protéines à domaine Alba font ressortir trois principales fonctions : (1) la liaison et stabilisation de l'ADN chez les *Archaea*; (2) l'appartenance aux complexes RNase P / MRP chez les eucaryotes supérieurs suggérant leur rôle dans la maturation des ARNt et/ou des ARNr; et (3) leur rôle dans la liaison des ARNm et la modulation de leur traduction chez les parasites protozoaires. Notre étude montre pour la première fois que des ARNm peuvent être stabilisés par une protéine à domaine Alba.
- Ces travaux constituent la première mise en évidence de la circulation de protéines de liaison à l'ARNm au sein du flagelle, ponctuellement lors de la différenciation, suggérant la participation de cet organelle dans la réponse au stress pour réguler des ARNm différemment régulés pendant le cycle parasitaire.
- D'un point de vue plus technique, des optimisations ont été apportées dans les manipulations génétiques chez *Leishmania*. L'utilisation de PCR chevauchante et de cycle d'amplification de type *Touch down PCR* ont remplacé les précédentes approches longues basées sur les plasmides. Nous avons montré l'intérêt des tags ARN dans des colonnes d'affinité comme méthode simple, rapide et peu coûteuse pour l'identification de protéines de liaison à l'ARN. Enfin, une nouvelle approche a été ici développée pour pouvoir intégrer dans le locus endogène des protéines recombinantes, comme par exemple des fusions avec des protéines fluorescentes. Cette approche est aujourd'hui la plus optimisée chez les *Trypanosomatidae* pour préserver l'intégrité du

locus ciblé et vérifier la localisation de protéines cibles avec un niveau d'expression et une régulation conservée.

## 11.1. Les mécanismes impliqués dans la régulation de l'expression de l'amastine spécifiquement au stade amastigote.

Nos travaux soulignent et complètent la complexité de la régulation de l'amastine chez le parasite *L. infantum*. Ce transcrite comporte un long 3'UTR, atteignant 1.8 kb, pour un cadre de lecture de 636 pb. Ceci suggère la présence de multiples motifs et séquences impliqués dans sa régulation, et plusieurs d'entre eux sont maintenant identifiés (Figure 11-1).

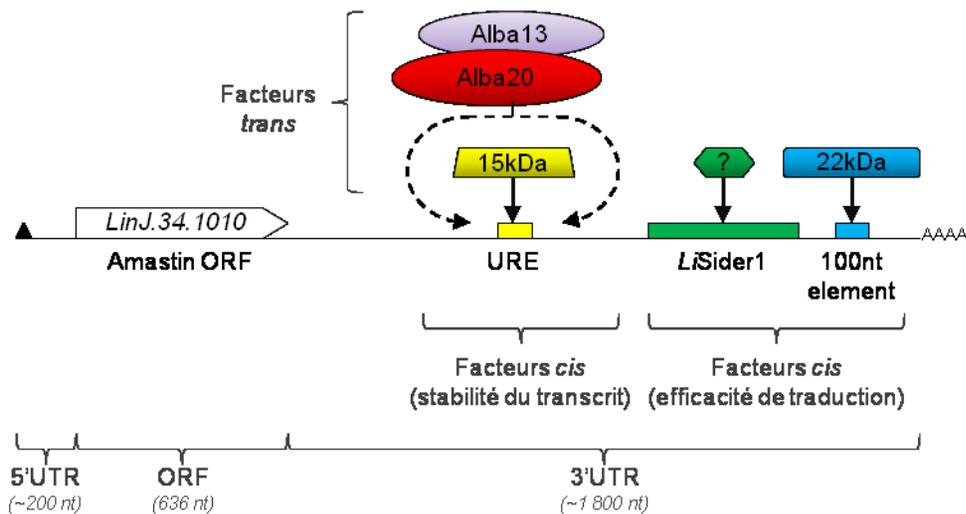


Figure 11-1 : Représentation du transcrite de la  $\delta$ -amastine LinJ.34.1010 et des facteurs *cis* et *trans* impliqués dans la régulation de l'expression génique.

### 11.1.1. Implication du *LiSIDER1* et d'une seconde séquence de 100 nt dans la traduction préférentielle au stade amastigote.

Les travaux précédemment effectués pour comprendre le mécanisme d'expression de la  $\delta$ -amastine LinJ.34.1010 ont mis en évidence la présence d'un rétroélément éteint de 450 nt, nommé *LiSIDER1*, dans le 3'UTR du transcrite. Cet élément est responsable d'une augmentation de sa traduction spécifiquement au stade amastigote [342]. Il stimule l'association de l'ARNm avec les polysomes en réponse à un stress de température [350].

Il a été suggéré que le *LiSIDER1*, présent dans le 3'UTR de nombreux transcrits, dont plusieurs amastines, constitue un mécanisme commun de régulation pour plusieurs

gènes spécifiques du stade amastigote. Actuellement, le *LiSIDER1* de cette amastine est le seul qui a été étudié, il est donc encore impossible d'affirmer son rôle à grande échelle. Aucune protéine de liaison au *LiSIDER1* n'a été identifiée à ce jour.

Un second élément long de 100 nt, placé 100 nt en aval du *LiSIDER1*, présente une activité similaire sur la stimulation de la traduction spécifiquement au stade amastigote. Ces deux éléments distincts stimulent la traduction, par un mécanisme encore non décrit [350]. Des expériences de liaison covalente aux UV (technique détaillée en [504]) ont montré que les 150 derniers nucléotides du transcrit, possédant la dernière moitié de l'élément de 100 nt, lient une protéine autour de 22 kDa. Cependant, des chromatographies d'affinité à l'ARN basées sur l'interaction streptavidine-biotine ont uniquement permis l'identification de MRP1 et MRP2, des protéines de liaison et d'édition de l'ARN mitochondrial, de 27 kDa (François McNicoll, résultats non publiés). La spécificité de cette liaison est à confirmer.

Ces études suggèrent que lorsque les parasites *Leishmania* sont transmis à un hôte mammifère par piqûre d'un phlébotome, la synthèse des protéines Amastines commence rapidement suite à la modification de la température ambiante, permise par la présence du *LiSIDER1* et d'un second élément de 100 nt. *Leishmania* modifie ainsi ses protéines de surface avant même la fusion des phagosomes avec les lysosomes dans le macrophage (mimée par un stress d'acidité). Cependant, cette activité traductionnelle précoce est effectuée sur un nombre limité de transcrits puisque leur stabilisation est activée par un stress acide [350]. L'augmentation de la quantité de transcrit dans la cellule semble donc s'effectuer pendant la différenciation morphologique du parasite de promastigote vers amastigote, un évènement qui prend place dans le phagolysosome. Notons qu'un choc thermique est pourtant responsable de l'inhibition globale de la traduction par phosphorylation du facteur d'initiation eIF2 $\alpha$ , un mécanisme également décrit chez *Leishmania* [396]. Le *LiSIDER1* est donc potentiellement impliqué dans un mécanisme de sélection pour une traduction préférentielle de l'amastine, ou bien cette augmentation de la traduction est le résultat d'un mécanisme indépendant du facteur eIF2 $\alpha$  (revue dans [505]). Des mécanismes de passage entre traduction dépendante de la coiffe et par une voie alternative ont d'ailleurs été suggérés chez *Leishmania* [380].

### **11.1.2. L'élément riche en uridines dégrade l'ARNm de l'amastine spécifiquement chez les promastigotes.**

Dans le Chapitre 8, nous avons mis en évidence l'existence de deux mécanismes distincts de dégradation du transcrit de l'amastine. Le premier correspond à une dégradation constitutive, avec une cinétique lente, retrouvée aux stades amastigote et promastigote. Ce mécanisme de dégradation est dépendant de la déadénylation, et doit correspondre à une réponse normale au renouvellement des ARNm. Le second mécanisme de dégradation, présent uniquement au stade promastigote, possède une cinétique rapide, est permis par la présence du URE et semble indépendant de la déadénylation. Notons que chez les mammifères, la plupart des ARNm, y compris ceux portant un ARE, sont dégradés par un mécanisme dépendant de la déadénylation. Les AUBP déstabilisent les ARNm cibles en recrutant des déadénylases, puis les transcrits sont digérés par des exonucléases (revue dans [325]). Comme mentionné dans la discussion de l'article, une autre conclusion de nos travaux pourrait être que la dégradation rapide du transcrit par le URE est dépendante de la présence de la queue polyadénylée, un mécanisme décrit chez les bactéries mais quasi absent chez les eucaryotes.

Pour compléter l'étude de la dégradation du transcrit en présence ou en absence du URE, nous avons développé une méthode de mesure indirecte pour estimer la dégradation de 5' vers 3' du transcrit. La comparaison de la demi-vie des ARNm totaux avec la demi-vie uniquement des ARNm présentant une coiffe ne suggère pas de différence en présence ou en absence de l'URE. Puisque la déadénylation n'est pas détectable en présence du URE, et que la présence de la coiffe ne semble pas impliquée dans la dégradation, il est peu probable qu'une activité uniquement de type exonucléase soit présente. D'après les mécanismes connus de dégradation des ARNm, il semble que l'URE de l'amastine soit responsable d'une dégradation effectuée par une activité endoribonucléolytique, ce qui reste à prouver étant donné qu'aucun produit de dégradation n'est visible par Northern blot. En utilisant des parasites *L. infantum* de type sauvage, l'extension d'amorces radiomarquées sur des extraits d'ARN totaux (tel que décrit en [400]) a permis de mettre en évidence un arrêt de la polymérisation, et donc un site potentiel de clivage de l'ARNm de l'amastine (voir Annexe 1). Ce site est situé 36 nucléotides en aval du URE. Cependant, cet arrêt de la polymérisation est obtenu aussi bien avec des extraits d'ARNm de promastigote

que des extraits d'ARNm d'amastigotes, ce qui ne corrèle pas avec un mécanisme de dégradation stade-spécifique. Des analyses de clivage par d'autres approches permettraient de confirmer ce résultat. Le URE permet donc de supprimer rapidement les transcrits chez les promastigotes, réduisant les possibilités de synthèse protéique dans un stade où les amastines ne sont pas requises. L'efficacité de ~70% du URE dans un système rapporteur suggère cependant qu'il n'est pas le seul élément impliqué dans la stabilisation des transcrits.

Il a été précédemment démontré que l'accumulation de l'ARNm de cet amastine au stade amastigote était une réponse à un stress d'acidité [350]. Il est donc possible que l'inactivation de la dégradation par l'URE soit une réponse à la variation du pH, un événement rencontré biologiquement lors de la fusion entre phagosomes et lysosomes au sein du macrophage. Des expériences supplémentaires de l'URE sous différents stress éclairciraient cette possibilité. Une autre interprétation serait que la dégradation par le URE et la transcription sont inhibées suite à un stress de température, stabilisant un faible niveau de transcrit en cours de traduction avant la pénétration dans le phagolysosome.

### **11.1.3. Recherche de protéines liant l'URE du transcrit de l'amastine.**

Suite à la publication de ces travaux, des expériences ont été effectuées dans le but d'identifier des protéines de liaison à l'URE de l'amastine. L'approche sélectionnée était la liaison covalente aux UV, telle que précédemment effectuée pour l'élément de 100 nt. Nous avons comparé les protéines liant un fragment de 300 nt correspondant aux 100 nt du URE plus 100 nt des séquences avoisinantes en 5' et en 3' (ce fragment correspond exactement à celui utilisé dans la chromatographie d'affinité à l'ARN présenté dans le Chapitre 9, sans l'étiquette D8). Comme contrôle négatif, nous avons utilisé un fragment comprenant uniquement les régions avoisinantes sans l'URE (en utilisant le même couple d'amorces que pour le fragment de 300 nt, mais sur la construction LUC-3'UTR $\Delta$ URE décrite dans le Chapitre 8). Cette approche nous a permis de mettre en évidence un signal autour de 17 kDa spécifiquement obtenu en présence de l'URE (voir Annexe 2). Cependant, il est estimé que la présence d'acides nucléiques augmente le poids moléculaire de quelques kilodaltons, la protéine identifiée a donc une taille suggérée de 15 kDa. L'interaction est visible aussi bien avec des extraits protéiques de souches sauvages de *L. infantum* promastigote et

amastigote, ainsi qu'avec des extraits protéiques de souche sauvage de *L. major*. La spécificité de cette liaison URE-protéine a été estimée en effectuant des expériences de compétition entre ARN radiomarqué contenant l'URE et des ARN froids ne le contenant pas. Les résultats suggèrent que la protéine caractérisée a une affinité ~ 10 fois supérieure pour la séquence riche en uridines comparativement à la séquence contrôle de 200 nt, tandis qu'un ARN portant le gène de la GFP n'est pas du tout capable d'entrer en compétition avec le URE.

Par la suite, nous avons tenté d'estimer plus précisément quelle portion de l'URE de 100 nt était impliquée dans la liaison avec la protéine de 15 kDa. Pour cela, trois délétions de ~30 nt ont été réalisées successivement à l'intérieur du URE (Annexe 3). Chacune de ces délétions permettait l'interaction avec la protéine, suggérant soit la présence d'au moins deux motifs distincts dans les 100 nt du URE, soit que cette protéine lie les séquences riches en uridines sans réelle spécificité. Quelques hypothèses, basées sur l'identification de protéines interagissant avec l'amastine chez *T. cruzi*, pourraient être émises pour tenter de caractériser cette protéine de 15 kDa. Par exemple, les protéines *TcUBP1* et *TcUBP2* lient les séquences ARE et des séquences riches en GU. L'interaction de *TcUBP1* avec le transcrit d'une amastine a par ailleurs été démontrée chez cet organisme [339, 448, 506]. Chez *L. infantum*, ces orthologues ont cependant un poids moléculaire théorique trop élevé pour que ces protéines complètes soient celles identifiées (*LiUBP1* : LinJ.25.0500, 27 kDa, *LiUBP2* : LinJ.25.0510, 19 kDa). La reproduction de ces expériences avec des extraits protéiques des souches *LiAlba20<sup>-/-</sup>* et *LiAlba13<sup>-/-</sup>* montre que ce ne sont pas les protéines *LiAlba* qui sont identifiées par cette approche expérimentale (Annexe 4). De même la génération d'une souche n'exprimant pas l'endoribonucléase L-PSP (LinJ.23.0220, 17 kDa), *Lilpsp<sup>-/-</sup>*, n'affecte pas la stabilité de l'amastine (Aurélien Dupé, résultat non publié). L'optimisation des conditions expérimentales des chromatographies d'affinité pourrait aider à caractériser cette protéine de 15 kDa liant des séquences riches en U.

#### **11.1.4. L'implication des protéines à domaine Alba dans la régulation du transcrit de l'amastine.**

##### **11.1.4.1. La protéine Rpp25-like : *LiAlba20***

La recherche de protéines liant le URE de l'amastine a été effectuée par une chromatographie d'affinité à l'ARN basée sur l'insertion d'une séquence d'ARN d'environ 40 nt, appelée D8, liant les billes de Sephadex. En utilisant cette approche, nous avons identifié la protéine *LiAlba20* comme partenaire du transcrit de l'amastine, et cette interaction a été confirmée par immunoprécipitation et RT-PCR. Cependant, l'utilisation de la séquence contrôle de 200 nt qui ne porte pas l'URE (voir ci-dessus) permet également l'identification de *LiAlba20* par spectrométrie de masse, avec un nombre de peptides similaire (Aurélien Dupé, résultat non publié). *LiAlba20* est donc différente de la protéine de 15 kDa suggérée par les expériences de liaison aux UV, et semble se fixer dans les régions avoisinant le URE. La liaison entre le transcrit de l'amastine et la protéine *LiAlba20* semble spécifique puisque d'autres transcrits endogènes ne sont pas co-élués par immunoprécipitation.

Nos expériences montrent que l'absence de *LiAlba20* provoque une diminution de la demi-vie des transcrits de l'amastine spécifiquement au stade amastigote, provoquant ainsi une diminution de leur accumulation. *LiAlba20* n'explique que partiellement la différence d'abondance de l'ARNm entre promastigotes et amastigotes chez les parasites sauvages. Nous avons donc identifié ici un second phénomène, indépendant du URE impliqué dans l'expression du transcrit de l'amastine, spécifiquement au stade amastigote. L'addition de l'effet accumulateur de *LiAlba20* chez les amastigotes et de l'effet déstabilisateur du URE chez les promastigotes explique peut-être intégralement l'expression stade-spécifique du transcrit de l'amastine chez *Leishmania*.

La protéine *LiAlba20* n'est pas régulée pendant le cycle parasitaire, et elle interagit avec le transcrit de l'amastine aussi bien au stade promastigote qu'au stade amastigote. Plusieurs hypothèses pourraient expliquer pourquoi elle agit sur le transcrit uniquement chez les amastigotes. La première est que la dégradation rapide du transcrit chez les promastigotes masque l'activité de *LiAlba20* à ce stade. La seconde est que *LiAlba20* subit une ou plusieurs modifications post-traductionnelles uniquement dans l'un des stades de vie

(telles que la phosphorylation ou la méthylation du domaine RGG de la protéine), permettant ainsi de moduler son activité. L'activation de *LiAlba20* pourrait alors recruter des partenaires actifs ou exprimés seulement à ce stade. Nos études des partenaires protéiques de *LiAlba20* indiquent que la grande majorité des protéines identifiées le sont dans les deux stades de vie du parasite et font partie de la machinerie de traduction. Les candidats potentiels retrouvés uniquement chez les amastigotes sont (1) l'orthologue de DRBD2 (LinJ.35.2240) portant un domaine de liaison aux ARN double brins et impliqué dans l'export des ARNm, (2) l'homologue de Rap-55 (SCD6, LinJ.25.0550) connu comme répresseur traductionnel et partenaire de la *RNA helicase* DHH1 ou (3) une protéine inconnue, potentiellement vacuolaire (LinJ.19.0020). Cependant, la liaison de ces facteurs uniquement au stade amastigote doit être vérifiée en utilisant des anticorps spécifiques à ces protéines.

Chez les *Apicomplexa* et les *Trypanosomatidae*, les protéines Alba interagissent avec la machinerie de traduction. De plus, les études effectuées chez *T. brucei* [346] et *Plasmodium* [475] montrent respectivement que ces protéines semblent être impliquées dans l'inhibition de la traduction et se retrouvent dans des granules en latence traductionnelle. Alba interagit ou localise avec le répresseur traductionnel DHH1 chez ces deux organismes [475, 483]. Dans le cas de *Leishmania*, les protéines qui interagissent avec *LiAlba* sont très similaires, cependant l'ARNm de l'amastine est supposément traduit chez les amastigotes, et *LiAlba20* interagit très faiblement avec les fractions polysomales, plus spécifiquement chez les amastigotes. L'utilisation d'un anticorps spécifique de l'amastine LinJ.34.1010 permettrait de confirmer son expression. En cumulant ces informations, une hypothèse serait donc que deux lots d'ARNm de l'amastine sont présents au stade amastigote. Le premier est activement traduit et ne lierait donc pas la protéine *LiAlba20*, tandis que le second, interagissant avec *LiAlba20*, serait conservé dans un état non traduit et non dégradé pour servir la machinerie traductionnelle en cas de besoin (Figure 11-2).

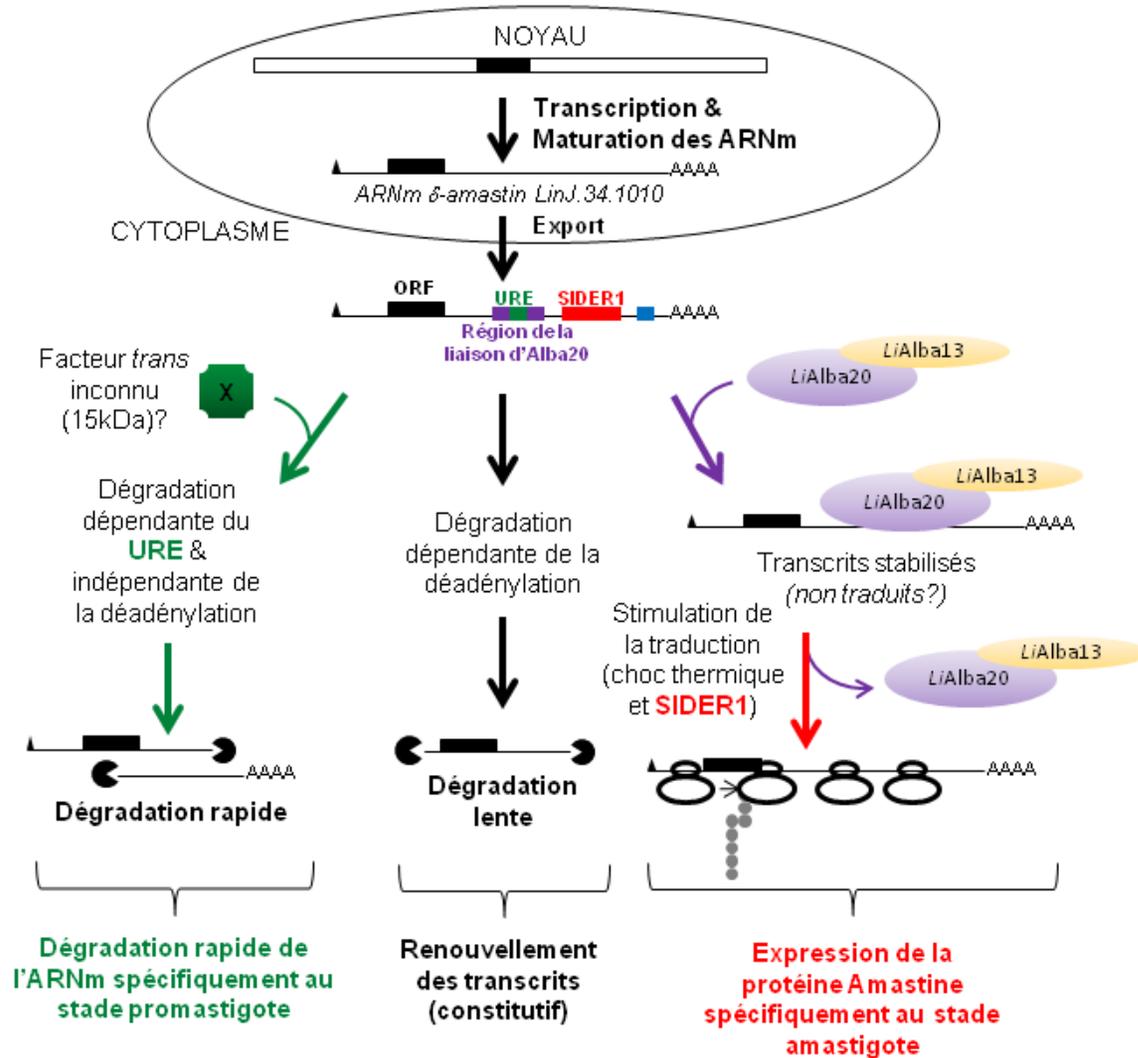


Figure 11-2 : Modèle mécanistique suggéré de l'expression de l'amastine LinJ.34.1010 en fonction du stade de vie chez *Leishmania*.

Ce stock d'ARNm spécifiquement au stade amastigote pourrait sans doute permettre une adaptation plus rapide aux besoins de la cellule suite à différents stress rencontrés dans le phagolysosome (stress acide, oxydatif ou nutritif par exemple). Ceci reste en accord avec le fait que l'absence de *LiAlba20* ne permet plus le stockage des transcrits de l'amastine, et favorise ainsi la dégradation du transcrit, tel qu'observé par Northern blot chez les amastigotes. La prise en compte de la localisation des protéines Alba sera détaillée ci-dessous.

#### 11.1.4.2. La protéine Rpp20-like : *LiAlba13*

Seules deux protéines à domaine Alba sont retrouvées dans les génomes de *Leishmania* spp. D'après nos résultats, *LiAlba20* et *LiAlba13* colocalisent et immunoprécipitent ensemble aux deux stades du parasite et lors de différents stress. Les expériences d'immunoprécipitation effectuées à partir de *LiAlba20* dans les deux stades de vie du parasite suggèrent une expression constitutive de *LiAlba13*. Notons enfin que si *LiAlba13* n'a pas été identifiée comme liant l'URE de l'amastine dans notre expérience de chromatographie d'affinité à l'ARN (Chapitre 9), cela est sans doute dû à la petite taille de cette protéine (13 kDa) qui était exclue des fragments envoyés pour identification par spectrométrie de masse (de 17 à 27 kDa).

Des immunoprécipitations effectuées à partir de *LiAlba13*-HA permettent de co-éluer l'ARNm de l'amastine, et ce dans les deux stades de vie du parasite. Cependant, aucune différence d'abondance du transcrite de l'amastine n'a été observé entre une souche sauvage et la souche *LiAlba13*<sup>-/-</sup>. Ainsi, cette protéine ne semble pas moduler l'accumulation stade-spécifique de l'ARNm de l'amastine, ni au stade promastigote, ni au stade amastigote. Cette protéine joue peut-être un rôle plus important dans la modulation de la traduction des transcrits. L'implication de *LiAlba13* dans la traduction chez *Leishmania* est suggérée par sa surexpression dans une souche atténuée à la gentamicine [507]. Le mécanisme d'action de cette aminoglycoside n'est pas bien établi, mais semble perturber l'activité ribosomale par perte des polysomes ou en causant des erreurs de lecture pendant la synthèse protéique. *LiAlba13*-HA ne semble pas interagir avec les polysomes, tout comme *LiAlba20*. Nous avons entamé des expériences avec le gène rapporteur luciférase sous contrôle du 3'UTR de l'amastine, intégré dans le génome des souches sauvages, *LiAlba20*<sup>-/-</sup> et *LiAlba13*<sup>-/-</sup>. La finalisation de ces expériences permettra peut-être de mettre en évidence le rôle de *LiAlba13* dans la traduction. L'état actuel de nos connaissances sur cette protéine rend difficile toute interprétation supplémentaire.

#### 11.1.5. Autres éléments *cis* communs aux 3'UTR des $\delta$ -amastines.

Afin de compléter notre étude sur les éléments *cis* impliqués dans la régulation de l'amastine, nous avons également effectué une recherche de motifs dans le 3'UTR de cet ARN en utilisant l'algorithme MEME/MAST. Les différentes sous-familles des amastines

sont régulées différemment et les  $\delta$ -amastines sont celles qui présentent la plus forte surexpression au stade amastigote [439]. Notre amastine de référence fait partie de cette sous-famille. Nous avons donc utilisé comme référence le 3'UTR des 30  $\delta$ -amastines de *L. infantum* (estimé ici comme inclus dans les 2 kb en aval du codon stop) et pris comme contrôles négatifs des séquences similaires issues de 50 autres transcrits codant pour des protéines membranaires, exprimés de manière constitutive d'après les résultats de transcriptomique [83]. L'annexe 5 liste les 25 principaux motifs identifiés et l'annexe 6 précise la position de ces motifs dans le 3'UTR de l'amastine LinJ.34.1010, comparativement à la position des éléments déjà caractérisés et impliqués dans la régulation du transcrit.

Le motif 1, dont la séquence consensus correspond à GTGCGTGYGCGTGC est le plus conservé et est retrouvé dans tous les 3'UTR des  $\delta$ -amastines (E-value :  $3 \times 10^{-82}$ ). Dans le cas de LinJ.34.1010, il est situé 20 nt en aval du URE, et est également présent à ~80 nt avant le site de polyadénylation de l'ARN, ces deux positions étant différentes de tous les éléments régulateurs connus. Un motif similaire, nommé LeishCICS-c1-3338 [508], a déjà été prédit *in silico* comme surreprésenté dans les ARNm spécifiques du stade amastigote chez *Leishmania*. Notons que le potentiel site de clivage décrit dans l'annexe 1 est situé 2 nt en aval de ce motif. Des constructions ont été effectuées de manière similaire à la délétion du URE présenté au chapitre 8, et les résultats préliminaires suggèrent l'intérêt de cet élément dans l'expression du transcrit de l'amastine (Aurélien Dupé, résultats non publiés). Selon la prédiction de la structure secondaire de l'ARNm de l'amastine (présentée en Annexe 7 uniquement pour la région de 300 nt entourant le URE), ce motif est impliqué dans une structure de type double brin, et le site de clivage potentiel est situé dans un environnement simple brin. La recherche de la séquence consensus du motif 1 dans le génome de *L. infantum* a donné 240 segments, dont 107 sont en position sens dans les 2kb en aval des cadres de lecture. Les amastines sont cependant les seules de ce groupe à être clairement spécifiques du stage amastigote axénique selon les études de transcriptomique [83]. Ce motif est un point de départ intéressant pour la suite des études de la régulation de l'amastine.

Parmi les autres motifs identifiés, les motifs 22 et 23 sont tous les deux riches en uridines, situés à l'intérieur de l'URE de 100 nt étudié et sont majoritairement prédits comme étant simple brin. Ces motifs sont des candidats intéressants pour cibler plus précisément des séquences impliquées dans la dégradation des transcrits spécifiquement au stade promastigote. Au sein de la séquence de 300 nt utilisés pour l'identification de *LiAlba20*, plusieurs motifs caractérisés en dehors de l'URE pourraient être la cible de cette protéine (motifs 1, 19, 20 et 28).

Le *LiSIDER1* étant un élément assez bien conservé, plusieurs motifs ont été caractérisés à son niveau, plus particulièrement dans sa séquence signature. En revanche, cette analyse n'a pas permis d'identifier de motifs conservés dans l'élément de 100 nt, à la fin du transcrit, impliqué dans la traduction de l'amastine au stade amastigote. Il est possible que la structure secondaire joue un rôle plus important que la structure primaire à ce niveau.

Parmi les motifs communs aux amastines qui sont dans des régions du 3'UTR encore non étudiées, les séquences consensus des motifs 9, 10, 11, 12 et 16 sont prédites comme appartenant à des 3'UTR de  $\delta$ -amastine dans plus de 90% des cas (sur recherche du génome entier), démontrant ainsi leur spécificité pour cette sous-famille et, de ce fait, ouvrant la voie à de nouvelles possibilités de recherche sur la régulation de ces transcrits.

## **11.2. La famille des protéines Alba chez les parasites protozoaires**

Les protéines à domaine Alba ont été décrites dans le Chapitre 6. Cette famille de protéines est impliquée dans différents processus cellulaires en fonction des organismes, allant de la stabilisation de structures double brins de l'ADN chez les *Archaea* jusqu'à l'appartenance aux RNase P / MRP chez les eucaryotes supérieurs. Nos travaux, ainsi que ceux publiés par d'autres équipes, montrent que chez les protozoaires, cette famille de protéines a un rôle dans la régulation de l'expression des ARNm. Notons qu'à l'heure actuelle, aucune activité de liaison à l'ADN des protéines Alba n'a été décrite chez les *Trypanosomatidae*, et les études d'identification de complexes liant les protéines Alba n'ont pas décrit de protéines nucléosomales. Chez les *Apicomplexa*, l'implication des

protéines Alba dans la liaison à l'ADN est documentée chez *Plasmodium* spp. [478, 479] et est suggérée chez *T. gondii* [471].

Les protéines Alba ne semblent pas essentielles pour la survie du parasite, puisque la génération de souches de *L. infantum* ne possédant pas *LiAlba20* ou *LiAlba13* sont viables. Cependant, un défaut de croissance est observé en leur absence chez *Leishmania*, tout comme chez *T. brucei* [346]. Chez ce dernier, Alba semble également assurer l'intégrité de l'organisation cellulaire [483]. Des délétions des gènes codant pour les protéines Alba ont également pu être réalisées chez *T. gondii*, indiquant qu'elles ne sont pas non plus essentielles chez cet *Apicomplexa* [473].

Les protéines Alba sont subdivisées en 3 sous-familles, et les parasites ont tous au moins un membre de la sous famille *Rpp20-like* et *Rpp25-like*. Chez les *Archaea* et les eucaryotes, ces sous-unités interagissent ensemble [346, 461, 478, 499]. Cela suppose que l'hétérodimerisation de ces protéines est la base du complexe Alba, surtout pour les organismes n'ayant qu'une seule copie de chaque sous-famille, comme c'est le cas chez *Leishmania* spp. Malgré les expériences menées chez les protozoaires, il est encore difficile de définir quelles sont les protéines qui interagissent directement avec le complexe Alba, la majorité des protéines identifiées comme partenaires potentiels sont des protéines de liaison à l'ARNm, telles que les PABP ou eIF4E, et cette interaction semble être effectuée par un pont ARN (Aurélien Dupé, résultats non publiés et [346]). Des expériences de type double hybride chez la levure (comme effectuées pour les protéines totales, dont Alba, chez *Plasmodium* [480]), ou la synthèse d'anticorps spécifiques aux protéines identifiées par immunoprécipitation, devraient permettre de mieux définir les partenaires protéiques directs des protéines Alba. Une estimation plus rapide pourrait être la reproduction des immunoprécipitations suite au traitement des extraits par des RNases, afin de supprimer les identifications contaminantes par les ponts ARN. Cependant, chez les parasites protozoaires, un consensus de protéines partenaires d'Alba existe : les PABP (suggérant que Alba lie au moins les ARNm polyadénylés), le facteur d'initiation de la traduction eIF4E, une ou plusieurs DEAD-box RNA hélicase, dont DHH11 au moins chez *Plasmodium berghei* [475], *T. brucei* [483], et *T. cruzi* (selon le rapport de thèse de Ferrarini MG, équipe Avila AR, Curitiba, Brésil).

Chez *T. brucei*, les protéines Alba sont très abondantes et ne doivent donc pas lier uniquement des ARNm tels que GPEET. De plus, leurs effets aussi bien sur le cycle cellulaire que sur la morphologie du parasite suggèrent que ces protéines ont des rôles multiples. Il serait donc intéressant de regarder plus précisément quels sont les ARNm qui sont liés par les protéines Alba. Dans notre étude, l'utilisation de puces à ADN a permis d'estimer que l'abondance de plus de 200 transcrits diffère en absence de *LiAlba20*. Cette première analyse montre bien que cette protéine est impliquée dans la régulation des transcrits, mais cette approche ne permet pas de différencier quels sont les transcrits qui lient directement les protéines Alba d'autres événements secondaires. La seule étude effectuée dans le but de bien caractériser l'ensemble des transcrits liés par Alba est chez *T. gondii*, où des résultats de co-immunoprécipitation ont été déposés sur puce à ADN, seule une trentaine de transcrits a été identifiée, et il ne semble pas y avoir de fonction commune des protéines codées par ces transcrits [473]. Une approche similaire chez *Leishmania*, comparée à nos résultats de puce à ADN, permettrait peut être de mieux caractériser si cette protéine est bien impliquée dans le stockage des ARNm de la famille des  $\delta$ -amastines au stade amastigote.

Chez *T. gondii*, la présence de *TgAlba1* (*Rpp25-like*) est requise pour la traduction du transcrit *TgAlba2* (*Rpp20-like*), et ceci très certainement par la liaison de *TgAlba1* sur le 3'UTR de l'ARNm de *TgAlba2* [473]. Chez *T. brucei*, la présence des protéines *TbAlba3* et *TbAlba4* (*Rpp25-like*) est requise pour la traduction de *TbAlba1* et *TbAlba2* (*Rpp20-like*), sans pour autant affecter l'abondance des ARNm de ces derniers [346]. Chez l'homme, les niveaux d'expression de *Rpp20* et *Rpp25* sont également co-dépendants [499]. Une activation traductionnelle des *Rpp20-like* par les *Rpp25-like* semble être conservée, sans doute pour s'assurer du maintien de la stœchiométrie des protéines Alba dans leur complexe. Nous avons montré que les deux protéines *LiAlba* lient les deux transcrits (résultats non publiés). Si l'absence de *LiAlba13* (*Rpp20-like*) n'influence pas la traduction de *LiAlba20* (*Rpp25-like*), nous ne savons cependant pas l'importance de *LiAlba20* dans la traduction de *LiAlba13*.

Enfin, les protéines Alba semblent impliquées dans l'expression stade spécifique de gènes, que ce soit l'ARNm de l'amastine chez *Leishmania*, l'ARNm de GPEET (spécifique

du stade procyclique) chez *T. brucei*, ou la liaison aux éléments TARE et au promoteur du gène *var* chez *Plasmodium*. Malgré cette tendance, il faut tout de même rappeler que les transcrits impliqués dans la virulence ou exprimés différemment en fonction des stades de vie sont les plus étudiés chez ces parasites, et il est donc encore trop tôt pour pouvoir affirmer que les protéines Alba sont uniquement impliquées dans la régulation de ce type de transcrits.

### **11.3. La localisation des protéines Alba lors de la différenciation du parasite et en réponse au stress.**

Notre étude montre que les protéines *LiAlba20* et *LiAlba13* sont cytoplasmiques chez les promastigotes et les amastigotes, ce qui appuie un rôle de ces protéines dans le métabolisme des ARNm. Cependant, le signal est plus périnucléaire chez les promastigotes suggérant que des petites différences de localisation existent entre les stades de vie de *Leishmania*.

Les études effectuées chez les autres parasites protozoaires montrent également des localisations différentes de ces protéines. En effet, dans la forme intracellulaire de *T. gondii*, les protéines *TgAlba* sont majoritairement cytoplasmiques tandis qu'elles s'accumulent clairement dans des structures périnucléaires, estimées comme étant des granules d'ARN, dans la forme extracellulaire [473]. Chez *P. falciparum*, les protéines *PfAlba* se concentrent au sein du noyau, dans des granules périnucléaires, visuellement similaires à ceux de *T. gondii*, et précipitent dans des fractions nucléaires spécifiquement dans le stade anneau. Cependant, elles sont cytoplasmiques dans les autres stades de développement [478, 479]. Chez cet organisme, des études des protéines de liaison à l'intron des gènes *var* ont permis d'identifier les protéines Alba dans des structures périnucléaires inhibant la transcription de ces gènes (malgré le fait que les auteurs soulignent que ces protéines ne semblent pas avoir de spécificité de liaison pour l'ADN) [485]. Chez *T. brucei*, les protéines sont cytoplasmiques dans la forme procyclique et s'accumulent dans des granules d'ARN suite à un stress nutritif [346], mais dans des structures différentes suite à un stress thermique (basé sur la non colocalisation avec *TbDHH1*) [483].

Notre étude s'accorde donc avec celles décrites ci-dessus sur une localisation différente des protéines Alba en fonction de stades de vie parasitaire, et ceci est lié à la régulation stade spécifique de gènes codant pour des protéines membranaires tels que *var*, *gpeet* ou *amastine*. De plus, nous montrons que lors de la différenciation du parasite *L. infantum* de promastigote à amastigote axénique, la localisation des protéines est fortement affectée puisqu'elles s'accumulent dans le flagelle et le nucléole. Cette observation est reproductible avec un simple stress thermique représentatif des conditions environnementales rencontrées par le parasite dans le macrophage. Ceci suggère un rôle des protéines Alba dans la réponse au stress thermique chez *Leishmania*. Des expériences de type hybridation *in situ* de l'ARNm sont nécessaires pour savoir où se localise le transcrite de l'amastine durant la différenciation. Des essais de co-immunoprécipitation des protéines interagissant avec *LiAlba20* lors de la différenciation ont permis d'identifier uniquement une HSP70 cytoplasmique, malgré un rendement d'élution similaire à celui obtenu dans les autres stades de vie. Il est envisageable que *LiAlba20* n'interagisse pas avec les ARNm durant la différenciation, ce qui expliquerait l'absence d'identification de PABP et autres protéines impliquées dans le métabolisme des ARNm. De plus, le flagelle ne possède pas de machinerie de traduction, et il n'est pas connu pour intervenir dans d'autres processus relatifs aux ARNm. Comme précédemment mentionné, l'élévation de la température permet d'augmenter le taux de traduction des ARNm de l'amastine, sans pour autant modifier son abondance. Si l'on suit notre hypothèse précédente, sur l'existence de deux lots différents d'ARNm de l'amastine chez les amastigotes, cette localisation transitoire pourrait être un mécanisme de séquestration. Il permettrait aux rares ARNm d'être directement traduits plutôt que d'être stockés. L'inhibition du transport flagellaire serait une approche intéressante pour vérifier son implication dans l'expression de l'amastine. La localisation nucléolaire est plus difficilement interprétable, la présence de *LiAlba20* dans le noyau est soit un site de séquestration secondaire, soit *LiAlba20* auraient un rôle plus globale que la régulation de l'amastine à ce stade. Ces deux structures sont de plus en plus décrites comme des senseurs et éléments de réponse aux stress environnementaux.

### **11.3.1. Localisation flagellaire**

Chez *T. brucei*, si les études de localisation n'ont pas été menées pendant la différenciation du parasite, des études de protéomique ont tout de même identifié les

protéines *TbAlba* dans la matrice flagellaire de formes sanguines du parasite, suggérant que leur circulation dans cet organelle est permise chez les *Trypanosomatidae* [144]. Notons que les études de localisation suite à un stress thermique ont été effectuées avec une incubation de 2h [483], et dans le cas de *Leishmania*, un minimum de 6h d'incubation est requis pour voir la protéine *LiAlba20* dans le flagelle et le nucléole. Il serait intéressant de voir si cette accumulation flagellaire est conservée chez les *Trypanosomatidae*. Les *Apicomplexa* ayant un flagelle uniquement aux stades gamètes, et les protéines Alba étant impliquées dans la liaison à l'ADN, une analogie est difficile à mettre en place.

Nous n'avons aucune donnée permettant d'expliquer le mécanisme et le but de la localisation flagellaire. Dans l'état actuel des connaissances, une première hypothèse serait que le flagelle séquestre les protéines pour permettre de libérer les transcrits cibles, qui seraient alors traduits ou régulés en conséquence dans le cytoplasme. Une seconde hypothèse pourrait être l'acquisition de modifications post-traductionnelles. Chez les vertébrés, le facteur modulateur de transcription Gli, impliqué dans la voie de signalisation *Hedgehog*, transite par le cil primaire pour être phosphorylé, puis spécifiquement clivé par le protéasome. Ceci permet sa maturation sous forme de répresseur traductionnel. Si aucune phosphorylation des protéines à domaine Alba n'a, à ce jour, été décrite chez *Leishmania*, elles ont par contre été montrées chez d'autres *Trypanosomatidae* au niveau de séquences conservées entre ces organismes. De plus, *LiAlba20* co purifie chez *L. infantum* avec le protéasome (Racine G., Ouellette M., résultats non publiés). Notons que chez *Trypanosoma rangeli*, la tyrosine phosphatase *TrPTP2* est localisée dans le flagelle [140]. Cette observation n'est pas reproduite chez *T. brucei* [509], mais les protéines portant des tyrosines phosphorylées semblent enrichies dans le flagelle, le corps basal et le nucléole [484]. Ces observations suggèrent que ces organites peuvent servir de centres de contrôle de modifications post-traductionnelles.

### 11.3.2. Localisation nucléolaire

D'après nos expériences d'immunoprécipitation, les protéines *LiAlba13* et *LiAlba20* pourraient transiter entre le cytoplasme et le compartiment nucléaire chez *Leishmania* par un complexe de type NTF2 (*Nuclear Transport Factor 2*, revu dans [510]). Chez *T. brucei*, DRBD3 est cytoplasmique et périnucléaire en conditions normales, elle immunoprécipite

avec des homologues de NTF2 et s'accumule dans le noyau suite à un stress à l'arsénite [511]. Ces études indiquent que des protéines de liaison à l'ARN peuvent transiter à travers la membrane nucléaire par ce type de transporteur. NTF2 n'est pas plus étudié à l'heure actuelle chez les *Trypanosomatidae*, mais semble donc être impliqué dans le transit de protéines de liaison à l'ARNm en réponse au stress, en faisant une cible de choix pour l'étude du rôle de ce type de transport.

Plusieurs études ont montré la translocation de protéines de liaison à l'ARNm entre le cytoplasme et le noyau suite à différents types de stress, mais la concentration nucléolaire est plus rarement mentionnée. Une étude a montré que des protéines telles que des PABP, *TcSR62* (impliquée dans la maturation et la stabilité des ARNm), *TcPTB2* (homologue de *Polypyrimidine tract binding protein*), *TcSF3b155* (facteur d'épissage) ou *TcFIP1* (impliqué dans la polyadénylation des transcrits) passaient du cytoplasme vers le nucléole suite à un stress thermique ou à une inhibition de la transcription. Ce mécanisme, actif et dépendant de la phosphorylation, est conservé chez *Leishmania* et *T. cruzi*, mais absent chez *T. brucei* [111, 112]. Dans le cas des protéines *LiAlba* chez *Leishmania*, seul le stress thermique est responsable de l'accumulation nucléolaire, tandis que l'inhibition de la transcription par l'actinomycine D semble provoquer une complète diffusion de *LiAlba20* et *LiAlba13* dans la cellule (résultat non publié).

L'appartenance des protéines à domaine Alba dans les complexes RNase P et MRP chez les eucaryotes supérieurs peut nous faire penser que *LiAlba20* et *LiAlba13* jouent un rôle dans la maturation des ARNt ou des ARNr. La RNase P a été caractérisée chez *T. brucei*, et est cependant très différente de celle rencontrée chez les autres eucaryotes, puisqu'une seule protéine semble suffisante pour la maturation des ARNt [487]. De plus, chez *T. brucei*, les protéines à domaine Alba ne semblent pas capables d'effectuer la maturation des ARNt *in vitro* [346]. Nos immunoprécipitations ont révélé que les protéines *LiAlba* de *Leishmania* interagissent avec une *DEAD-box RNA helicase* de 67 kDa. Cette protéine lie les ARNr LSU- $\gamma$  afin de les protéger d'une dégradation par des ARN anti-sens, et empêche ainsi l'apoptose suite à différents stress [400]. Il est donc envisageable que les protéines *LiAlba* participent dans le métabolisme des ARNr, ce qui signifie que *LiAlba* agirait à la fois sur les ARNr et les ARNm, deux processus bien distincts. Chez *T. brucei*,

une étude vient de montrer que l'ARNm *GPEET*, transcrit par l'ARN polymérase I dans le nucléole, liait un facteur NRG1 (pour *Negative regulator of GPEET*). L'inhibition de cette protéine ou d'autres facteurs nucléolaires associés est à la fois responsable de l'accumulation de *GPEET* et de la diminution de l'ARNr 5.8S et de ses précurseurs [512], soulignant la pluripotence des protéines de liaison à l'ARN chez cet organisme. Dans notre cas, l'amastine est transcrite par l'ARN polymérase II, et est donc indépendante du nucléole, et les ARNr LSU- $\gamma$  semblent s'accumuler normalement en l'absence des protéines *LiAlba* (résultats non publiés).

Outre son implication dans la biogénèse des ribosomes, le nucléole est également impliqué dans la réponse au stress chez les eucaryotes supérieurs (revu dans [110]). En réponse à différents stress, dont le choc thermique, le nucléole peut inhiber l'adressage de p53 vers le protéasome, et donc stopper la progression du cycle cellulaire. Les stress inhibent aussi la phosphorylation du facteur de transcription TIF-IA permettant l'arrêt de la transcription par l'ARN polymérase I. Le fait que plusieurs protéines s'accumulent dans cet organite lors de stress chez *Leishmania* suggère la conservation du rôle coordinateur du nucléole. Actuellement, aucune étude ne définit encore le rôle de la localisation nucléolaire de protéines suite à un stress chez les *Trypanosomatidae*.

Enfin, la translocation de protéines à domaine Alba entre le cytoplasme et le noyau est potentiellement un mécanisme de réponse au stress conservé au cours de l'évolution, puisque ces protéines s'accumulent dans le noyau en réponse à un stress de déshydratation chez les plantes [502].

## Conclusions.

Les travaux effectués durant cette thèse ont permis de mieux comprendre les mécanismes permettant l'expression d'une  $\delta$ -amastine, un facteur de virulence exprimé spécifiquement au stade amastigote chez *Leishmania*. Nous avons mis en évidence deux nouveaux mécanismes distincts qui permettent une bonne expression du transcrit au stade intracellulaire du parasite. Le premier est spécifiquement observé au stade promastigote et permet une dégradation rapide du transcrit de l'amastine par la présence d'un URE dans son 3'UTR, faisant intervenir un mécanisme atypique de dégradation indépendante de la déadénylation. Le second mécanisme opère plus spécifiquement au stade amastigote à l'aide d'une protéine à domaine Alba qui favorise la stabilisation du transcrit. Si séparément ces deux mécanismes ont un effet partiel, leur cumul explique sans doute intégralement la différence d'abondance des transcrits de l'amastine entre les deux stades de vie du parasite. Ces travaux complètent ceux précédemment obtenus, montrant que la traduction de ces transcrits est spécifiquement observée au stade amastigote. La régulation de l'amastine est donc complexe et est effectuée à plusieurs niveaux pour permettre une expression précise de ce facteur de virulence.

Nos résultats sur les protéines à domaine Alba sont en accord avec les travaux effectués par d'autres groupes. Il semble que chez les parasites protozoaires plus spécifiquement, ces protéines se sont spécialisées dans la régulation de l'expression génique. Leur effet est notamment visible à l'interface entre stabilisation des ARNm et inhibition de la traduction.

Les protéines à domaine Alba semblent constitutivement exprimées, mais leur localisation cytoplasmique diffère légèrement entre les stades de vie. Chez *Leishmania* plus particulièrement, lors de la différenciation du parasite vers le stade amastigote nous observons une accumulation des protéines à domaine Alba dans le nucléole et le flagelle. Chez les eucaryotes, quelques exemples suggèrent que ces organites sont coordonnateurs de la réponse au stress et impliqués dans la modulation de l'expression génique respectivement.

Notre étude constitue la première évidence de la circulation de protéines de liaison à l'ARN dans le flagelle. Cet organite, tout comme le nucléole, est sans doute impliqué dans la modulation de la traduction de certains transcrits lors de la différenciation. Si aucune explication n'est clairement établie suite à cette observation, elle suggère fortement que les protéines Alba sont des protéines impliquées dans la réponse au stress pour permettre le remaniement fonctionnel du parasite entre ses stades de vie.

## Bibliographie

1. Tuon, F.F., V.A. Neto, and V.S. Amato, *Leishmania: origin, evolution and future since the Precambrian*. FEMS Immunol Med Microbiol, 2008. **54**(2): p. 158-66.
2. Leblois, R., et al., *Guns, germs and dogs: On the origin of Leishmania chagasi*. Infect Genet Evol, 2011. **11**(5): p. 1091-5.
3. Banuls, A.L., M. Hide, and F. Prugnolle, *Leishmania and the leishmaniasis: a parasite genetic update and advances in taxonomy, epidemiology and pathogenicity in humans*. Adv Parasitol, 2007. **64**: p. 1-109.
4. Bates, P.A., *Transmission of Leishmania metacyclic promastigotes by phlebotomine sand flies*. Int J Parasitol, 2007. **37**(10): p. 1097-106.
5. Rogers, M.E. and P.A. Bates, *Leishmania manipulation of sand fly feeding behavior results in enhanced transmission*. PLoS Pathog, 2007. **3**(6): p. e91.
6. Claborn, D.M., *The biology and control of leishmaniasis vectors*. J Glob Infect Dis, 2010. **2**(2): p. 127-34.
7. Feliciangeli, M.D., *Natural breeding places of phlebotomine sandflies*. Med Vet Entomol, 2004. **18**(1): p. 71-80.
8. Coulibaly, E., et al., *Enquête sur les pratiques diagnostiques et thérapeutiques de la leishmaniose chez les vétérinaires praticiens en 2000*. Epidemiol Sante Anim, 2004. **45**: p. 33-44.
9. Murray, H.W., et al., *Advances in leishmaniasis*. Lancet, 2005. **366**(9496): p. 1561-77.
10. David, C.V. and N. Craft, *Cutaneous and mucocutaneous leishmaniasis*. Dermatol Ther, 2009. **22**(6): p. 491-502.
11. Akilov, O.E., A. Khachemoune, and T. Hasan, *Clinical manifestations and classification of Old World cutaneous leishmaniasis*. Int J Dermatol, 2007. **46**(2): p. 132-42.
12. Goto, H. and J.A. Lindoso, *Current diagnosis and treatment of cutaneous and mucocutaneous leishmaniasis*. Expert Rev Anti Infect Ther, 2010. **8**(4): p. 419-33.
13. Faucher, B. and R. Piarroux, *[Visceral leishmaniasis: An update.]*. Rev Med Interne, 2010.
14. Herwaldt, B.L., *Leishmaniasis*. Lancet, 1999. **354**(9185): p. 1191-9.
15. Alvar, J., et al., *The relationship between leishmaniasis and AIDS: the second 10 years*. Clin Microbiol Rev, 2008. **21**(2): p. 334-59, table of contents.
16. Anabwani, G.M., et al., *Comparison of two dosage schedules of sodium stibogluconate in the treatment of visceral leishmaniasis in Kenya*. Lancet, 1983. **1**(8318): p. 210-3.
17. Sundar, S., et al., *Failure of pentavalent antimony in visceral leishmaniasis in India: report from the center of the Indian epidemic*. Clin Infect Dis, 2000. **31**(4): p. 1104-7.
18. Baiocco, P., et al., *Molecular basis of antimony treatment in leishmaniasis*. J Med Chem, 2009. **52**(8): p. 2603-12.
19. Croft, S.L., S. Sundar, and A.H. Fairlamb, *Drug resistance in leishmaniasis*. Clin Microbiol Rev, 2006. **19**(1): p. 111-26.

20. Bern, C., *et al.*, *Liposomal amphotericin B for the treatment of visceral leishmaniasis*. Clin Infect Dis, 2006. **43**(7): p. 917-24.
21. Lachaud, L., *et al.*, *Parasite susceptibility to amphotericin B in failures of treatment for visceral leishmaniasis in patients coinfecting with HIV type 1 and Leishmania infantum*. Clin Infect Dis, 2009. **48**(2): p. e16-22.
22. Bhattacharya, S.K., *et al.*, *Phase 4 trial of miltefosine for the treatment of Indian visceral leishmaniasis*. J Infect Dis, 2007. **196**(4): p. 591-8.
23. Mishra, M., *et al.*, *Amphotericin versus pentamidine in antimony-unresponsive kala-azar*. Lancet, 1992. **340**(8830): p. 1256-7.
24. Sundar, S., *et al.*, *Injectable paromomycin for Visceral leishmaniasis in India*. N Engl J Med, 2007. **356**(25): p. 2571-81.
25. Kedzierski, L., Y. Zhu, and E. Handman, *Leishmania vaccines: progress and problems*. Parasitology, 2006. **133 Suppl**: p. S87-112.
26. Kaye, P. and P. Scott, *Leishmaniasis: complexity at the host-pathogen interface*. Nat Rev Microbiol, 2011. **9**(8): p. 604-15.
27. Elias, M.C., *et al.*, *Morphological events during the Trypanosoma cruzi cell cycle*. Protist, 2007. **158**(2): p. 147-57.
28. Woodward, R. and K. Gull, *Timing of nuclear and kinetoplast DNA replication and early morphological events in the cell cycle of Trypanosoma brucei*. J Cell Sci, 1990. **95 ( Pt 1)**: p. 49-57.
29. Wheeler, R.J., E. Gluenz, and K. Gull, *The cell cycle of Leishmania: morphogenetic events and their implications for parasite biology*. Mol Microbiol, 2010. **79**(3): p. 647-62.
30. Ambit, A., *et al.*, *Morphological events during the cell cycle of Leishmania major*. Eukaryot Cell, 2011. **10**(11): p. 1429-38.
31. Akopyants, N.S., *et al.*, *Demonstration of genetic exchange during cyclical development of Leishmania in the sand fly vector*. Science, 2009. **324**(5924): p. 265-8.
32. Mougneau, E., F. Bihl, and N. Glaichenhaus, *Cell biology and immunology of Leishmania*. Immunol Rev, 2011. **240**(1): p. 286-96.
33. Kamhawi, S., *Phlebotomine sand flies and Leishmania parasites: friends or foes?* Trends Parasitol, 2006. **22**(9): p. 439-45.
34. Sacks, D. and S. Kamhawi, *Molecular aspects of parasite-vector and vector-host interactions in leishmaniasis*. Annu Rev Microbiol, 2001. **55**: p. 453-83.
35. Schlein, Y., R.L. Jacobson, and J. Shlomai, *Chitinase secreted by Leishmania functions in the sandfly vector*. Proc Biol Sci, 1991. **245**(1313): p. 121-6.
36. Gossage, S.M., M.E. Rogers, and P.A. Bates, *Two separate growth phases during the development of Leishmania in sand flies: implications for understanding the life cycle*. Int J Parasitol, 2003. **33**(10): p. 1027-34.
37. Kamhawi, S., *et al.*, *A role for insect galectins in parasite survival*. Cell, 2004. **119**(3): p. 329-41.
38. Rogers, M.E., M.L. Chance, and P.A. Bates, *The role of promastigote secretory gel in the origin and transmission of the infective stage of Leishmania mexicana by the sandfly Lutzomyia longipalpis*. Parasitology, 2002. **124**(Pt 5): p. 495-507.
39. Killick-Kendrick, R., D.H. Molyneux, and R.W. Ashford, *Leishmania in phlebotomid sandflies. I. Modifications of the flagellum associated with attachment*

- to the mid-gut and oesophageal valve of the sandfly. Proc R Soc Lond B Biol Sci, 1974. **187**(1089): p. 409-19.
40. Sacks, D.L., *Metacyclogenesis in Leishmania promastigotes*. Exp Parasitol, 1989. **69**(1): p. 100-3.
  41. Mojtahedi, Z., J. Clos, and E. Kamali-Sarvestani, *Leishmania major: identification of developmentally regulated proteins in procyclic and metacyclic promastigotes*. Exp Parasitol, 2008. **119**(3): p. 422-9.
  42. Nugent, P.G., et al., *Proteomic analysis of Leishmania mexicana differentiation*. Mol Biochem Parasitol, 2004. **136**(1): p. 51-62.
  43. Yao, C., et al., *Proteomic examination of Leishmania chagasi plasma membrane proteins: Contrast between avirulent and virulent (metacyclic) parasite forms*. Proteomics Clin Appl, 2010. **4**(1): p. 4-16.
  44. Thomas, F., S. Adamo, and J. Moore, *Parasitic manipulation: where are we and where should we go?* Behav Processes, 2005. **68**(3): p. 185-99.
  45. House, P.K., A. Vyas, and R. Sapolsky, *Predator cat odors activate sexual arousal pathways in brains of Toxoplasma gondii infected rats*. PLoS One, 2011. **6**(8): p. e23277.
  46. Filardy, A.A., D.R. Pires, and G.A. DosReis, *Macrophages and neutrophils cooperate in immune responses to Leishmania infection*. Cell Mol Life Sci, 2011. **68**(11): p. 1863-70.
  47. Rittig, M.G. and C. Bogdan, *Leishmania-host-cell interaction: complexities and alternative views*. Parasitol Today, 2000. **16**(7): p. 292-7.
  48. Rogers, M.E., et al., *Transmission of cutaneous leishmaniasis by sand flies is enhanced by regurgitation of fPPG*. Nature, 2004. **430**(6998): p. 463-7.
  49. Kimblin, N., et al., *Quantification of the infectious dose of Leishmania major transmitted to the skin by single sand flies*. Proc Natl Acad Sci U S A, 2008. **105**(29): p. 10125-30.
  50. Lambris, J.D., D. Ricklin, and B.V. Geisbrecht, *Complement evasion by human pathogens*. Nat Rev Microbiol, 2008. **6**(2): p. 132-42.
  51. Brown, E.J., *Complement receptors and phagocytosis*. Curr Opin Immunol, 1991. **3**(1): p. 76-82.
  52. Sacks, D.L., *The structure and function of the surface lipophosphoglycan on different developmental stages of Leishmania promastigotes*. Infect Agents Dis, 1992. **1**(4): p. 200-6.
  53. Puentes, S.M., et al., *Serum resistance of metacyclic stage Leishmania major promastigotes is due to release of C5b-9*. J Immunol, 1990. **145**(12): p. 4311-6.
  54. Puentes, S.M., et al., *Binding and release of C3 from Leishmania donovani promastigotes during incubation in normal human serum*. J Immunol, 1989. **143**(11): p. 3743-9.
  55. Bogdan, C. and M. Rollinghoff, *The immune response to Leishmania: mechanisms of parasite control and evasion*. Int J Parasitol, 1998. **28**(1): p. 121-34.
  56. Dominguez, M., et al., *Early mechanisms of Leishmania infection in human blood*. Microbes Infect, 2003. **5**(6): p. 507-13.
  57. Beattie, L. and P.M. Kaye, *Leishmania-host interactions: what has imaging taught us?* Cell Microbiol, 2011. **13**(11): p. 1659-67.

58. Laufs, H., *et al.*, *Intracellular survival of Leishmania major in neutrophil granulocytes after uptake in the absence of heat-labile serum factors*. Infect Immun, 2002. **70**(2): p. 826-35.
59. Mollinedo, F., *et al.*, *Selective fusion of azurophilic granules with Leishmania-containing phagosomes in human neutrophils*. J Biol Chem, 2010. **285**(45): p. 34528-36.
60. Guimaraes-Costa, A.B., *et al.*, *Leishmania amazonensis promastigotes induce and are killed by neutrophil extracellular traps*. Proc Natl Acad Sci U S A, 2009. **106**(16): p. 6748-53.
61. Ritter, U., F. Frischknecht, and G. van Zandbergen, *Are neutrophils important host cells for Leishmania parasites?* Trends Parasitol, 2009. **25**(11): p. 505-10.
62. Peters, N.C., *et al.*, *In vivo imaging reveals an essential role for neutrophils in leishmaniasis transmitted by sand flies*. Science, 2008. **321**(5891): p. 970-4.
63. Castro, R., *et al.*, *The ultrastructure of the parasitophorous vacuole formed by Leishmania major*. J Parasitol, 2006. **92**(6): p. 1162-70.
64. Maurer, M., B. Dondji, and E. von Stebut, *What determines the success or failure of intracellular cutaneous parasites? Lessons learned from leishmaniasis*. Med Microbiol Immunol, 2009. **198**(3): p. 137-46.
65. Liu, D. and J.E. Uzonna, *The early interaction of Leishmania with macrophages and dendritic cells and its influence on the host immune response*. Front Cell Infect Microbiol, 2012. **2**: p. 83.
66. Mosser, D.M. and J.P. Edwards, *Exploring the full spectrum of macrophage activation*. Nat Rev Immunol, 2008. **8**(12): p. 958-69.
67. Mosser, D.M. and P.J. Edelson, *The mouse macrophage receptor for C3bi (CR3) is a major mechanism in the phagocytosis of Leishmania promastigotes*. J Immunol, 1985. **135**(4): p. 2785-9.
68. Ueno, N. and M.E. Wilson, *Receptor-mediated phagocytosis of Leishmania: implications for intracellular survival*. Trends Parasitol, 2012. **28**(8): p. 335-44.
69. Guy, R.A. and M. Belosevic, *Comparison of receptors required for entry of Leishmania major amastigotes into macrophages*. Infect Immun, 1993. **61**(4): p. 1553-8.
70. Cunningham, A.C., *Parasitic adaptive mechanisms in infection by leishmania*. Exp Mol Pathol, 2002. **72**(2): p. 132-41.
71. Silverman, J.M. and N.E. Reiner, *Exosomes and other microvesicles in infection biology: organelles with unanticipated phenotypes*. Cell Microbiol, 2010. **13**(1): p. 1-9.
72. Silverman, J.M., *et al.*, *An exosome-based secretion pathway is responsible for protein export from Leishmania and communication with macrophages*. J Cell Sci, 2010. **123**(Pt 6): p. 842-52.
73. Silverman, J.M., *et al.*, *Leishmania exosomes modulate innate and adaptive immune responses through effects on monocytes and dendritic cells*. J Immunol, 2010. **185**(9): p. 5011-22.
74. Osorio y Fortea, J., *et al.*, *Transcriptional signatures of BALB/c mouse macrophages housing multiplying Leishmania amazonensis amastigotes*. BMC Genomics, 2009. **10**: p. 119.
75. Chakraborty, D., *et al.*, *Leishmania donovani affects antigen presentation of macrophage by disrupting lipid rafts*. J Immunol, 2005. **175**(5): p. 3214-24.

76. Winberg, M.E., *et al.*, *Leishmania donovani* lipophosphoglycan inhibits phagosomal maturation via action on membrane rafts. *Microbes Infect*, 2009. **11**(2): p. 215-22.
77. Desjardins, M. and A. Descoteaux, *Inhibition of phagolysosomal biogenesis by the Leishmania lipophosphoglycan*. *J Exp Med*, 1997. **185**(12): p. 2061-8.
78. Bringaud, F., L. Riviere, and V. Coustou, *Energy metabolism of trypanosomatids: adaptation to available carbon sources*. *Mol Biochem Parasitol*, 2006. **149**(1): p. 1-9.
79. Naderer, T. and M.J. McConville, *The Leishmania-macrophage interaction: a metabolic perspective*. *Cell Microbiol*, 2008. **10**(2): p. 301-8.
80. Rosenzweig, D., *et al.*, *Retooling Leishmania metabolism: from sand fly gut to human macrophage*. *Faseb J*, 2008. **22**(2): p. 590-602.
81. Zilberstein, D. and M. Shapira, *The role of pH and temperature in the development of Leishmania parasites*. *Annu Rev Microbiol*, 1994. **48**: p. 449-70.
82. Bates, P.A., *Axenic culture of Leishmania amastigotes*. *Parasitol Today*, 1993. **9**(4): p. 143-6.
83. Rochette, A., *et al.*, *Whole-genome comparative RNA expression profiling of axenic and intracellular amastigote forms of Leishmania infantum*. *Mol Biochem Parasitol*, 2009. **165**(1): p. 32-47.
84. Barak, E., *et al.*, *Differentiation of Leishmania donovani in host-free system: analysis of signal perception and response*. *Mol Biochem Parasitol*, 2005. **141**(1): p. 99-108.
85. Salinas, G., *et al.*, *Leishmania RNA viruses in Leishmania of the Viannia subgenus*. *Am J Trop Med Hyg*, 1996. **54**(4): p. 425-9.
86. Stuart, K.D., *et al.*, *Molecular organization of Leishmania RNA virus 1*. *Proc Natl Acad Sci U S A*, 1992. **89**(18): p. 8596-600.
87. Hartley, M.A., *et al.*, *Leishmania RNA virus: when the host pays the toll*. *Front Cell Infect Microbiol*, 2012. **2**: p. 99.
88. Ives, A., *et al.*, *Leishmania RNA virus controls the severity of mucocutaneous leishmaniasis*. *Science*, 2011. **331**(6018): p. 775-8.
89. Cattand, P., J. Jannin, and P. Lucas, *Sleeping sickness surveillance: an essential step towards elimination*. *Trop Med Int Health*, 2001. **6**(5): p. 348-61.
90. WHO. *World Health Organisation Special Programme for Research and Training in Tropical Diseases*. 2005. Geneva, Switzerland.
91. Fenn, K. and K.R. Matthews, *The cell biology of Trypanosoma brucei differentiation*. *Curr Opin Microbiol*, 2007. **10**(6): p. 539-46.
92. de Souza, W., T.M. de Carvalho, and E.S. Barrias, *Review on Trypanosoma cruzi: Host Cell Interaction*. *Int J Cell Biol*, 2010. **2010**.
93. Kramer, S., *Developmental regulation of gene expression in the absence of transcriptional control: the case of kinetoplastids*. *Mol Biochem Parasitol*, 2012. **181**(2): p. 61-72.
94. Haag, J., C. O'HUigin, and P. Overath, *The molecular phylogeny of trypanosomes: evidence for an early divergence of the Salivaria*. *Mol Biochem Parasitol*, 1998. **91**(1): p. 37-49.
95. Simpson, A.G., J.R. Stevens, and J. Lukes, *The evolution and diversity of kinetoplastid flagellates*. *Trends Parasitol*, 2006. **22**(4): p. 168-74.

96. Vickerman, K., *The evolutionary expansion of the trypanosomatid flagellates*. Int J Parasitol, 1994. **24**(8): p. 1317-31.
97. Deschamps, P., *et al.*, *Phylogenomic analysis of kinetoplastids supports that trypanosomatids arose from within bodonids*. Mol Biol Evol, 2010. **28**(1): p. 53-8.
98. Bodyl, A., P. Mackiewicz, and R. Milanowski, *Did trypanosomatid parasites contain a eukaryotic alga-derived plastid in their evolutionary past?* J Parasitol, 2010. **96**(2): p. 465-75.
99. Marin, C., *et al.*, *Large differences in the genome organization of different plant Trypanosomatid parasites (Phytomonas spp.) reveal wide evolutionary divergences between taxa*. Infect Genet Evol, 2009. **9**(2): p. 235-40.
100. Overath, P., *et al.*, *The surface structure of trypanosomes in relation to their molecular phylogeny*. Int J Parasitol, 2001. **31**(5-6): p. 468-71.
101. Stevens, J.R. and W.C. Gibson, *The evolution of pathogenic trypanosomes*. Cad Saude Publica, 1999. **15**(4): p. 673-84.
102. Stevens, J.R., *et al.*, *The molecular evolution of Trypanosomatidae*. Adv Parasitol, 2001. **48**: p. 1-56.
103. O'Brien, S.J., *et al.*, *The promise of comparative genomics in mammals*. Science, 1999. **286**(5439): p. 458-62, 479-81.
104. Kerr, S.F., *Molecular trees of trypanosomes incongruent with fossil records of hosts*. Mem Inst Oswaldo Cruz, 2006. **101**(1): p. 25-30.
105. Hannaert, V., *et al.*, *Evolution of energy metabolism and its compartmentation in Kinetoplastida*. Kinetoplastid Biol Dis, 2003. **2**(1): p. 11.
106. Clayton, C., T. Hausler, and J. Blattner, *Protein trafficking in kinetoplastid protozoa*. Microbiol Rev, 1995. **59**(3): p. 325-44.
107. McConville, M.J., *et al.*, *Secretory pathway of trypanosomatid parasites*. Microbiol Mol Biol Rev, 2002. **66**(1): p. 122-54; table of contents.
108. Docampo, R. and S.N. Moreno, *The acidocalcisome*. Mol Biochem Parasitol, 2001. **114**(2): p. 151-9.
109. Daniels, J.P., K. Gull, and B. Wickstead, *Cell biology of the trypanosome genome*. Microbiol Mol Biol Rev, 2010. **74**(4): p. 552-69.
110. Boisvert, F.M., *et al.*, *The multifunctional nucleolus*. Nat Rev Mol Cell Biol, 2007. **8**(7): p. 574-85.
111. Nazer, E. and D.O. Sanchez, *Nucleolar accumulation of RNA binding proteins induced by Actinomycin D is functional in Trypanosoma cruzi and Leishmania mexicana but not in T. brucei*. PLoS One, 2011. **6**(8): p. e24184.
112. Nazer, E., R.E. Verdun, and D.O. Sanchez, *Nucleolar localization of RNA binding proteins induced by actinomycin D and heat shock in Trypanosoma cruzi*. PLoS One, 2011. **6**(5): p. e19920.
113. Ogbadoyi, E.O., D.R. Robinson, and K. Gull, *A high-order trans-membrane structural linkage is responsible for mitochondrial genome positioning and segregation by flagellar basal bodies in trypanosomes*. Mol Biol Cell, 2003. **14**(5): p. 1769-79.
114. Flegontov, P.N. and A.A. Kolesnikov, *Radically different maxicircle classes within the same kinetoplast: an artefact or a novel feature of the kinetoplast genome?* Kinetoplastid Biol Dis, 2006. **5**: p. 5.
115. Ochsenreiter, T., M. Cipriano, and S.L. Hajduk, *KISS: the kinetoplastid RNA editing sequence search tool*. Rna, 2007. **13**(1): p. 1-4.

116. Aphasizhev, R. and I. Aphasizheva, *Mitochondrial RNA processing in trypanosomes*. Res Microbiol, 2011.
117. Hajduk, S. and T. Ochsenreiter, *RNA editing in kinetoplastids*. RNA Biol, 2010. **7**(2): p. 229-36.
118. Feagin, J.E., *et al.*, *Creation of AUG initiation codons by addition of uridines within cytochrome b transcripts of kinetoplastids*. Proc Natl Acad Sci U S A, 1988. **85**(2): p. 539-43.
119. Feagin, J.E., J.M. Abraham, and K. Stuart, *Extensive editing of the cytochrome c oxidase III transcript in Trypanosoma brucei*. Cell, 1988. **53**(3): p. 413-22.
120. Schnauffer, A., G.J. Domingo, and K. Stuart, *Natural and induced dyskinetoplastic trypanosomatids: how to live without mitochondrial DNA*. Int J Parasitol, 2002. **32**(9): p. 1071-84.
121. Basselin, M., M.A. Badet-Denisot, and M. Robert-Gero, *Modification of kinetoplast DNA minicircle composition in pentamidine-resistant Leishmania*. Acta Trop, 1998. **70**(1): p. 43-61.
122. Nitz, N., *et al.*, *Heritable integration of kDNA minicircle sequences from Trypanosoma cruzi into the avian genome: insights into human Chagas disease*. Cell, 2004. **118**(2): p. 175-86.
123. Feagin, J.E., D.P. Jasmer, and K. Stuart, *Developmentally regulated addition of nucleotides within apocytochrome b transcripts in Trypanosoma brucei*. Cell, 1987. **49**(3): p. 337-45.
124. Field, M.C. and M. Carrington, *The trypanosome flagellar pocket*. Nat Rev Microbiol, 2009. **7**(11): p. 775-86.
125. Gull, K., *Host-parasite interactions and trypanosome morphogenesis: a flagellar pocketful of goodies*. Curr Opin Microbiol, 2003. **6**(4): p. 365-70.
126. Rohloff, P., A. Montalvetti, and R. Docampo, *Acidocalcisomes and the contractile vacuole complex are involved in osmoregulation in Trypanosoma cruzi*. J Biol Chem, 2004. **279**(50): p. 52270-81.
127. Ulrich, P.N., *et al.*, *Identification of contractile vacuole proteins in Trypanosoma cruzi*. PLoS One, 2011. **6**(3): p. e18013.
128. Bloodgood, R.A., *Protein targeting to flagella of trypanosomatid protozoa*. Cell Biol Int, 2000. **24**(12): p. 857-62.
129. Denny, P.W., M.C. Field, and D.F. Smith, *GPI-anchored proteins and glycoconjugates segregate into lipid rafts in Kinetoplastida*. FEBS Lett, 2001. **491**(1-2): p. 148-53.
130. Coppens, I. and P.J. Courtoy, *The adaptative mechanisms of Trypanosoma brucei for sterol homeostasis in its different life-cycle environments*. Annu Rev Microbiol, 2000. **54**: p. 129-56.
131. Maric, D., *et al.*, *Molecular determinants of ciliary membrane localization of Trypanosoma cruzi flagellar calcium-binding protein*. J Biol Chem, 2011. **286**(38): p. 33109-17.
132. Tull, D., *et al.*, *SMP-1, a member of a new family of small myristoylated proteins in kinetoplastid parasites, is targeted to the flagellum membrane in Leishmania*. Mol Biol Cell, 2004. **15**(11): p. 4775-86.
133. Tull, D., *et al.*, *Acylation-dependent and-independent membrane targeting and distinct functions of small myristoylated proteins (SMPs) in Leishmania major*. Int J Parasitol, 2012. **42**(3): p. 239-47.

134. Tran, K.D., *et al.*, *Both sequence and context are important for flagellar targeting of a glucose transporter*. J Cell Sci, 2012. **125**(Pt 14): p. 3293-8.
135. Landfear, S.M. and M. Ignatushchenko, *The flagellum and flagellar pocket of trypanosomatids*. Mol Biochem Parasitol, 2001. **115**(1): p. 1-17.
136. Stephan, A., *et al.*, *An essential quality control mechanism at the eukaryotic basal body prior to intraflagellar transport*. Traffic, 2007. **8**(10): p. 1323-30.
137. Pazour, G.J. and R.A. Bloodgood, *Targeting proteins to the ciliary membrane*. Curr Top Dev Biol, 2008. **85**: p. 115-49.
138. Fridberg, A., K.T. Buchanan, and D.M. Engman, *Flagellar membrane trafficking in kinetoplastids*. Parasitol Res, 2007. **100**(2): p. 205-12.
139. Lumb, J.H. and M.C. Field, *Rab23 is a flagellar protein in Trypanosoma brucei*. BMC Res Notes, 2011. **4**: p. 190.
140. Prestes, E.B., *et al.*, *Trypanosoma rangeli protein tyrosine phosphatase is associated with the parasite's flagellum*. Mem Inst Oswaldo Cruz, 2012. **107**(6): p. 713-9.
141. Adhiambo, C., *et al.*, *A novel function for the atypical small G protein Rab-like 5 in the assembly of the trypanosome flagellum*. J Cell Sci, 2009. **122**(Pt 6): p. 834-41.
142. Ostrowski, L.E., *et al.*, *A proteomic analysis of human cilia: identification of novel components*. Mol Cell Proteomics, 2002. **1**(6): p. 451-65.
143. Pazour, G.J., *et al.*, *Proteomic analysis of a eukaryotic cilium*. J Cell Biol, 2005. **170**(1): p. 103-13.
144. Oberholzer, M., *et al.*, *Independent analysis of the flagellum surface and matrix proteomes provides insight into flagellum signaling in mammalian-infectious Trypanosoma brucei*. Mol Cell Proteomics, 2011. **10**(10): p. M111 010538.
145. Maric, D., C.L. Epting, and D.M. Engman, *Composition and sensory function of the trypanosome flagellar membrane*. Curr Opin Microbiol, 2010. **13**(4): p. 466-72.
146. Ralston, K.S., *et al.*, *The Trypanosoma brucei flagellum: moving parasites in new directions*. Annu Rev Microbiol, 2009. **63**: p. 335-62.
147. Portman, N. and K. Gull, *The paraflagellar rod of kinetoplastid parasites: from structure to components and function*. Int J Parasitol, 2010. **40**(2): p. 135-48.
148. Gluenz, E., M.L. Ginger, and P.G. McKean, *Flagellum assembly and function during the Leishmania life cycle*. Curr Opin Microbiol, 2010. **13**(4): p. 473-9.
149. Gerdes, J.M., E.E. Davis, and N. Katsanis, *The vertebrate primary cilium in development, homeostasis, and disease*. Cell, 2009. **137**(1): p. 32-45.
150. Rosenbaum, J., *Intraflagellar transport*. Curr Biol, 2002. **12**(4): p. R125.
151. Julkowska, D. and P. Bastin, *Tools for analyzing intraflagellar transport in trypanosomes*. Methods Cell Biol, 2009. **93**: p. 59-80.
152. Gouveia, J.J., *et al.*, *Intraflagellar transport complex in Leishmania spp. In silico genome-wide screening and annotation of gene function*. Genet Mol Res, 2007. **6**(4): p. 766-98.
153. Buisson, J., *et al.*, *Intraflagellar transport proteins cycle between the flagellum and its base*. J Cell Sci, 2012.
154. Huangfu, D., *et al.*, *Hedgehog signalling in the mouse requires intraflagellar transport proteins*. Nature, 2003. **426**(6962): p. 83-7.
155. Goetz, S.C., P.J. Ocbina, and K.V. Anderson, *The primary cilium as a Hedgehog signal transduction machine*. Methods Cell Biol, 2009. **94**: p. 199-222.

156. Emmer, B.T., *et al.*, *Calflagin inhibition prolongs host survival and suppresses parasitemia in Trypanosoma brucei infection*. Eukaryot Cell, 2010. **9**(6): p. 934-42.
157. Berriman, M., *et al.*, *The genome of the African trypanosome Trypanosoma brucei*. Science, 2005. **309**(5733): p. 416-22.
158. El-Sayed, N.M., *et al.*, *The genome sequence of Trypanosoma cruzi, etiologic agent of Chagas disease*. Science, 2005. **309**(5733): p. 409-15.
159. Ivens, A.C., *et al.*, *The genome of the kinetoplastid parasite, Leishmania major*. Science, 2005. **309**(5733): p. 436-42.
160. Peacock, C.S., *et al.*, *Comparative genomic analysis of three Leishmania species that cause diverse human disease*. Nat Genet, 2007. **39**(7): p. 839-47.
161. Rogers, M.B., *et al.*, *Chromosome and gene copy number variation allow major structural change between species and strains of Leishmania*. Genome Res, 2011. **21**(12): p. 2129-42.
162. Raymond, F., *et al.*, *Genome sequencing of the lizard parasite Leishmania tarentolae reveals loss of genes associated to the intracellular stage of human pathogenic species*. Nucleic Acids Res, 2011. **40**(3): p. 1131-47.
163. Jackson, A.P., M.A. Quail, and M. Berriman, *Insights into the genome sequence of a free-living Kinetoplastid: Bodo saltans (Kinetoplastida: Euglenozoa)*. BMC Genomics, 2008. **9**: p. 594.
164. Britto, C., *et al.*, *Conserved linkage groups associated with large-scale chromosomal rearrangements between Old World and New World Leishmania genomes*. Gene, 1998. **222**(1): p. 107-17.
165. Myler, P.J., *et al.*, *Leishmania major Friedlin chromosome I has an unusual distribution of protein-coding genes*. Proc Natl Acad Sci U S A, 1999. **96**(6): p. 2902-6.
166. Tosato, V., *et al.*, *Secondary DNA structure analysis of the coding strand switch regions of five Leishmania major Friedlin chromosomes*. Curr Genet, 2001. **40**(3): p. 186-94.
167. Martinez-Calvillo, S., *et al.*, *Transcription of Leishmania major Friedlin chromosome I initiates in both directions within a single region*. Mol Cell, 2003. **11**(5): p. 1291-9.
168. Martinez-Calvillo, S., *et al.*, *Transcription initiation and termination on Leishmania major chromosome 3*. Eukaryot Cell, 2004. **3**(2): p. 506-17.
169. Campbell, D.A., S. Thomas, and N.R. Sturm, *Transcription in kinetoplastid protozoa: why be normal?* Microbes Infect, 2003. **5**(13): p. 1231-40.
170. Jess, W., *et al.*, *Phylogenetic analysis of the RNA polymerases of Trypanosoma brucei, with special reference to class-specific transcription*. Curr Genet, 1990. **18**(6): p. 547-51.
171. Martinez-Calvillo, S., *et al.*, *Characterization of the RNA polymerase II and III complexes in Leishmania major*. Int J Parasitol, 2007. **37**(5): p. 491-502.
172. Clayton, C.E., *Life without transcriptional control? From fly to man and back again*. Embo J, 2002. **21**(8): p. 1881-8.
173. Luo, H., *et al.*, *Transcription initiation at the TATA-less spliced leader RNA gene promoter requires at least two DNA-binding proteins and a tripartite architecture that includes an initiator element*. J Biol Chem, 1999. **274**(45): p. 31947-54.
174. Moretti, N.S. and S. Schenkman, *Chromatin modifications in trypanosomes due to stress*. Cell Microbiol, 2013. **15**(5): p. 709-17.

175. Thomas, S., *et al.*, *Histone acetylations mark origins of polycistronic transcription in Leishmania major*. BMC Genomics, 2009. **10**: p. 152.
176. van Leeuwen, F., *et al.*, *beta-D-glucosyl-hydroxymethyluracil is a conserved DNA modification in kinetoplastid protozoans and is abundant in their telomeres*. Proc Natl Acad Sci U S A, 1998. **95**(5): p. 2366-71.
177. Dooijes, D., *et al.*, *Base J originally found in kinetoplastida is also a minor constituent of nuclear DNA of Euglena gracilis*. Nucleic Acids Res, 2000. **28**(16): p. 3017-21.
178. Borst, P. and R. Sabatini, *Base J: discovery, biosynthesis, and possible functions*. Annu Rev Microbiol, 2008. **62**: p. 235-51.
179. Cliffe, L.J., *et al.*, *JBP1 and JBP2 are two distinct thymidine hydroxylases involved in J biosynthesis in genomic DNA of African trypanosomes*. Nucleic Acids Res, 2009. **37**(5): p. 1452-62.
180. van Luenen, H.G., *et al.*, *Glucosylated hydroxymethyluracil, DNA base j, prevents transcriptional readthrough in leishmania*. Cell, 2012. **150**(5): p. 909-21.
181. Lynn, M.A. and W.R. McMaster, *Leishmania: conserved evolution--diverse diseases*. Trends Parasitol, 2008. **24**(3): p. 103-5.
182. El-Sayed, N.M., *et al.*, *Comparative genomics of trypanosomatid parasitic protozoa*. Science, 2005. **309**(5733): p. 404-9.
183. Beverley, S.M., *et al.*, *Unstable DNA amplifications in methotrexate-resistant Leishmania consist of extrachromosomal circles which relocalize during stabilization*. Cell, 1984. **38**(2): p. 431-9.
184. Beverley, S.M., *Gene amplification in Leishmania*. Annu Rev Microbiol, 1991. **45**: p. 417-44.
185. Antoniazzi, S., H.C. Lima, and A.K. Cruz, *Overexpression of miniexon gene decreases virulence of Leishmania major in BALB/c mice in vivo*. Mol Biochem Parasitol, 2000. **107**(1): p. 57-69.
186. de Toledo, J.S., *et al.*, *Leishmania (Viannia) braziliensis transfectants overexpressing the miniexon gene lose virulence in vivo*. Parasitol Int, 2009. **58**(1): p. 45-50.
187. Iovannisci, D.M. and S.M. Beverley, *Structural alterations of chromosome 2 in Leishmania major as evidence for diploidy, including spontaneous amplification of the mini-exon array*. Mol Biochem Parasitol, 1989. **34**(2): p. 177-88.
188. Grondin, K., G. Roy, and M. Ouellette, *Formation of extrachromosomal circular amplicons with direct or inverted duplications in drug-resistant Leishmania tarentolae*. Mol Cell Biol, 1996. **16**(7): p. 3587-95.
189. Leprohon, P., *et al.*, *Gene expression modulation is associated with gene amplification, supernumerary chromosomes and chromosome loss in antimony-resistant Leishmania infantum*. Nucleic Acids Res, 2009. **37**(5): p. 1387-99.
190. Ubeda, J.M., *et al.*, *Modulation of gene expression in drug resistant Leishmania is associated with gene amplification, gene deletion and chromosome aneuploidy*. Genome Biol, 2008. **9**(7): p. R115.
191. Mannaert, A., *et al.*, *Adaptive mechanisms in pathogens: universal aneuploidy in Leishmania*. Trends Parasitol, 2012. **28**(9): p. 370-6.
192. Sterkers, Y., *et al.*, *FISH analysis reveals aneuploidy and continual generation of chromosomal mosaicism in Leishmania major*. Cell Microbiol, 2011. **13**(2): p. 274-83.

193. Sterkers, Y., *et al.*, *Novel insights into genome plasticity in Eukaryotes: mosaic aneuploidy in Leishmania*. Mol Microbiol, 2012.
194. Capy, P., *et al.*, *Dynamics and evolution of transposable elements*, ed. Springer. 1997, Austin TX: Landes Bioscience. 197.
195. Whitcomb, J.M. and S.H. Hughes, *Retroviral reverse transcription and integration: progress and problems*. Annu Rev Cell Biol, 1992. **8**: p. 275-306.
196. Luan, D.D., *et al.*, *Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition*. Cell, 1993. **72**(4): p. 595-605.
197. Burns, K.H. and J.D. Boeke, *Human transposon tectonics*. Cell, 2012. **149**(4): p. 740-52.
198. Smit, A.F., *Interspersed repeats and other mementos of transposable elements in mammalian genomes*. Curr Opin Genet Dev, 1999. **9**(6): p. 657-63.
199. Shapiro, J.A., *Retrotransposons and regulatory suites*. Bioessays, 2005. **27**(2): p. 122-5.
200. Jordan, I.K., *et al.*, *Origin of a substantial fraction of human regulatory sequences from transposable elements*. Trends Genet, 2003. **19**(2): p. 68-72.
201. Blond, J.L., V. Cheynet, and F. Mallet, *Signification biologique des rétrovirus endogènes humains*. Virologie, 2001. **5**(2): p. 91-111.
202. Villarreal, L.P., *Can viruses make us human?* Proceedings of the American Philosophical Society, 2004. **148**(3): p. 296-323.
203. Sinzelle, L., *et al.*, *Characterization of a Xenopus tropicalis endogenous retrovirus with developmental and stress-dependent expression*. J Virol, 2011. **85**(5): p. 2167-79.
204. Ullu, E., H.D. Lujan, and C. Tschudi, *Small sense and antisense RNAs derived from a telomeric retroposon family in Giardia intestinalis*. Eukaryot Cell, 2005. **4**(6): p. 1155-7.
205. Bringaud, F., *et al.*, *Members of a large retroposon family are determinants of post-transcriptional gene expression in Leishmania*. PLoS Pathog, 2007. **3**(9): p. 1291-307.
206. Vazquez, M., *et al.*, *The short interspersed repetitive element of Trypanosoma cruzi, SIRE, is part of VIPER, an unusual retroelement related to long terminal repeat retrotransposons*. Proc Natl Acad Sci U S A, 2000. **97**(5): p. 2128-33.
207. Aksoy, S., *et al.*, *Multiple copies of a retroposon interrupt spliced leader RNA genes in the African trypanosome, Trypanosoma gambiense*. Embo J, 1987. **6**(12): p. 3819-26.
208. Gabriel, A., *et al.*, *A rapidly rearranging retrotransposon within the miniexon gene locus of Crithidia fasciculata*. Mol Cell Biol, 1990. **10**(2): p. 615-24.
209. Villanueva, M.S., *et al.*, *A new member of a family of site-specific retrotransposons is present in the spliced leader RNA genes of Trypanosoma cruzi*. Mol Cell Biol, 1991. **11**(12): p. 6139-48.
210. Bringaud, F., M. Berriman, and C. Hertz-Fowler, *Trypanosomatid genomes contain several subfamilies of ingi-related retroposons*. Eukaryot Cell, 2009. **8**(10): p. 1532-42.
211. Bringaud, F., *et al.*, *The Trypanosoma cruzi LITc and NARTc non-LTR retrotransposons show relative site specificity for insertion*. Mol Biol Evol, 2006. **23**(2): p. 411-20.

212. Bringaud, F., *et al.*, *The ingi and RIME non-LTR retrotransposons are not randomly distributed in the genome of Trypanosoma brucei*. Mol Biol Evol, 2004. **21**(3): p. 520-8.
213. Garcia-Salcedo, J.A., *et al.*, *Searching for promoter activity in RIME/Ingi retrotransposons from Trypanosoma brucei: binding of a nuclear protein to their 5' extremity*. Exp Parasitol, 2003. **104**(3-4): p. 140-8.
214. Olivares, M., C. Alonso, and M.C. Lopez, *The open reading frame 1 of the LITc retrotransposon of Trypanosoma cruzi codes for a protein with apurinic-apyrimidinic nuclease activity*. J Biol Chem, 1997. **272**(40): p. 25224-8.
215. Olivares, M., *et al.*, *The non-LTR (long terminal repeat) retrotransposon LITc from Trypanosoma cruzi codes for a protein with RNase H activity*. J Biol Chem, 2002. **277**(31): p. 28025-30.
216. Pays, E. and N.B. Murphy, *DNA-binding fingers encoded by a trypanosome retroposon*. J Mol Biol, 1987. **197**(1): p. 147-8.
217. Heras, S.R., *et al.*, *The LITc non-LTR retrotransposon of Trypanosoma cruzi contains an internal RNA-pol II-dependent promoter that strongly activates gene transcription and generates unspliced transcripts*. Nucleic Acids Res, 2007. **35**(7): p. 2199-214.
218. Sanchez-Luque, F., *et al.*, *Pr77 and LITcRz: A dual system within the 5'-end of LITc retrotransposon, internal promoter and HDV-like ribozyme*. Mob Genet Elements, 2012. **2**(1): p. 1-7.
219. Bringaud, F., *et al.*, *Evolution of non-LTR retrotransposons in the trypanosomatid genomes: Leishmania major has lost the active elements*. Mol Biochem Parasitol, 2006. **145**(2): p. 158-70.
220. Smith, M., F. Bringaud, and B. Papadopoulou, *Organization and evolution of two SIDER retroposon subfamilies and their impact on the Leishmania genome*. BMC Genomics, 2009. **10**: p. 240.
221. Bringaud, F., M. Berriman, and C. Hertz-Fowler, *TSIDER1, a short and non-autonomous Salivarian trypanosome-specific retroposon related to the ingi6 subclade*. Mol Biochem Parasitol, 2011. **179**(1): p. 30-6.
222. Bhattacharya, S., A. Bakre, and A. Bhattacharya, *Mobile genetic elements in protozoan parasites*. J Genet, 2002. **81**(2): p. 73-86.
223. Mair, G., *et al.*, *A new twist in trypanosome RNA metabolism: cis-splicing of pre-mRNA*. Rna, 2000. **6**(2): p. 163-9.
224. Siegel, T.N., *et al.*, *Genome-wide analysis of mRNA abundance in two life-cycle stages of Trypanosoma brucei and identification of splicing and polyadenylation sites*. Nucleic Acids Res, 2010. **38**(15): p. 4946-57.
225. Agabian, N., *Trans splicing of nuclear pre-mRNAs*. Cell, 1990. **61**(7): p. 1157-60.
226. Liang, X.H., *et al.*, *trans and cis splicing in trypanosomatids: mechanism, factors, and regulation*. Eukaryot Cell, 2003. **2**(5): p. 830-40.
227. Preusser, C., N. Jae, and A. Bindereif, *mRNA splicing in trypanosomes*. Int J Med Microbiol, 2012. **302**(4-5): p. 221-4.
228. Gunzl, A., *The pre-mRNA splicing machinery of trypanosomes: complex or simplified?* Eukaryot Cell, 2010. **9**(8): p. 1159-70.
229. Das, A. and V. Bellofatto, *RNA polymerase II-dependent transcription in trypanosomes is associated with a SNAP complex-like transcription factor*. Proc Natl Acad Sci U S A, 2003. **100**(1): p. 80-5.

230. Hannon, G.J., P.A. Maroney, and T.W. Nilsen, *U small nuclear ribonucleoprotein requirements for nematode cis- and trans-splicing in vitro*. J Biol Chem, 1991. **266**(34): p. 22792-5.
231. Luz Ambrosio, D., *et al.*, *Spliceosomal proteomics in Trypanosoma brucei reveal new RNA splicing factors*. Eukaryot Cell, 2009. **8**(7): p. 990-1000.
232. LeBowitz, J.H., *et al.*, *Coupling of poly(A) site selection and trans-splicing in Leishmania*. Genes Dev, 1993. **7**(6): p. 996-1007.
233. Lamontagne, J. and B. Papadopoulou, *Developmental regulation of spliced leader RNA gene in Leishmania donovani amastigotes is mediated by specific polyadenylation*. J Biol Chem, 1999. **274**(10): p. 6602-9.
234. Papadopoulou, B., G. Roy, and M. Ouellette, *Autonomous replication of bacterial DNA plasmid oligomers in Leishmania*. Mol Biochem Parasitol, 1994. **65**(1): p. 39-49.
235. Stiles, J.K., *et al.*, *Genomic organization, transcription, splicing and gene regulation in Leishmania*. Ann Trop Med Parasitol, 1999. **93**(8): p. 781-807.
236. Clayton, C. and M. Shapira, *Post-transcriptional regulation of gene expression in trypanosomes and leishmanias*. Mol Biochem Parasitol, 2007. **156**(2): p. 93-101.
237. Brooks, D.R., *et al.*, *The stage-regulated expression of Leishmania mexicana CPB cysteine proteases is mediated by an intercistronic sequence element*. J Biol Chem, 2001. **276**(50): p. 47061-9.
238. Zilka, A., *et al.*, *Developmental regulation of heat shock protein 83 in Leishmania. 3' processing and mRNA stability control transcript abundance, and translation is directed by a determinant in the 3'-untranslated region*. J Biol Chem, 2001. **276**(51): p. 47922-9.
239. Bond, U., *Heat shock but not other stress inducers leads to the disruption of a subset of snRNPs and inhibition of in vitro splicing in HeLa cells*. Embo J, 1988. **7**(11): p. 3509-18.
240. Yost, H.J. and S. Lindquist, *Heat shock proteins affect RNA processing during the heat shock response of Saccharomyces cerevisiae*. Mol Cell Biol, 1991. **11**(2): p. 1062-8.
241. Jager, A.V., *et al.*, *mRNA maturation by two-step trans-splicing/polyadenylation processing in trypanosomes*. Proc Natl Acad Sci U S A, 2007. **104**(7): p. 2035-42.
242. Wagner, E.J. and M.A. Garcia-Blanco, *Polypyrimidine tract binding protein antagonizes exon definition*. Mol Cell Biol, 2001. **21**(10): p. 3281-8.
243. Kolev, N.G., *et al.*, *The transcriptome of the human pathogen Trypanosoma brucei at single-nucleotide resolution*. PLoS Pathog, 2010. **6**(9): p. e1001090.
244. Nilsson, D., *et al.*, *Spliced leader trapping reveals widespread alternative splicing patterns in the highly dynamic transcriptome of Trypanosoma brucei*. PLoS Pathog, 2010. **6**(8): p. e1001037.
245. Rastrojo, A., *et al.*, *The transcriptome of Leishmania major in the axenic promastigote stage: transcript annotation and relative expression levels by RNA-seq*. BMC Genomics, 2013. **14**(1): p. 223.
246. Manning-Cela, R., A. Gonzalez, and J. Swindle, *Alternative splicing of LYT1 transcripts in Trypanosoma cruzi*. Infect Immun, 2002. **70**(8): p. 4726-8.
247. Benabdellah, K., E. Gonzalez-Rey, and A. Gonzalez, *Alternative trans-splicing of the Trypanosoma cruzi LYT1 gene transcript results in compartmental and functional switch for the encoded protein*. Mol Microbiol, 2007. **65**(6): p. 1559-67.

248. Mayr, C. and D.P. Bartel, *Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells*. Cell, 2009. **138**(4): p. 673-84.
249. Sandberg, R., et al., *Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites*. Science, 2008. **320**(5883): p. 1643-7.
250. Ji, Z., et al., *Progressive lengthening of 3' untranslated regions of mRNAs by alternative polyadenylation during mouse embryonic development*. Proc Natl Acad Sci U S A, 2009. **106**(17): p. 7028-33.
251. Ji, Z. and B. Tian, *Reprogramming of 3' untranslated regions of mRNAs by alternative polyadenylation in generation of pluripotent stem cells from different cell types*. PLoS One, 2009. **4**(12): p. e8419.
252. Licatalosi, D.D., et al., *HITS-CLIP yields genome-wide insights into brain alternative RNA processing*. Nature, 2008. **456**(7221): p. 464-9.
253. Mangone, M., et al., *The landscape of C. elegans 3'UTRs*. Science, 2010. **329**(5990): p. 432-5.
254. Weill, L., et al., *Translational control by changes in poly(A) tail length: recycling mRNAs*. Nat Struct Mol Biol, 2012. **19**(6): p. 577-85.
255. Kohler, A. and E. Hurt, *Exporting RNA from the nucleus to the cytoplasm*. Nat Rev Mol Cell Biol, 2007. **8**(10): p. 761-73.
256. Biton, M., et al., *RNAi interference of XPO1 and Sm genes and their effect on the spliced leader RNA in Trypanosoma brucei*. Mol Biochem Parasitol, 2006. **150**(2): p. 132-43.
257. Zeiner, G.M., N.R. Sturm, and D.A. Campbell, *Exportin 1 mediates nuclear export of the kinetoplastid spliced leader RNA*. Eukaryot Cell, 2003. **2**(2): p. 222-30.
258. Serpeloni, M., et al., *An essential nuclear protein in trypanosomes is a component of mRNA transcription/export pathway*. PLoS One, 2011. **6**(6): p. e20730.
259. Dostalova, A., et al., *The nuclear mRNA export receptor Mex67-Mtr2 of Trypanosoma brucei contains a unique and essential zinc finger motif*. Mol Microbiol, 2013. **88**(4): p. 728-39.
260. Garneau, N.L., J. Wilusz, and C.J. Wilusz, *The highways and byways of mRNA decay*. Nat Rev Mol Cell Biol, 2007. **8**(2): p. 113-26.
261. Munroe, D. and A. Jacobson, *mRNA poly(A) tail, a 3' enhancer of translational initiation*. Mol Cell Biol, 1990. **10**(7): p. 3441-55.
262. Munroe, D. and A. Jacobson, *Tales of poly(A): a review*. Gene, 1990. **91**(2): p. 151-8.
263. Meijer, H.A., et al., *A novel method for poly(A) fractionation reveals a large population of mRNAs with a short poly(A) tail in mammalian cells*. Nucleic Acids Res, 2007. **35**(19): p. e132.
264. Liu, H., et al., *The scavenger mRNA decapping enzyme DcpS is a member of the HIT family of pyrophosphatases*. Embo J, 2002. **21**(17): p. 4699-708.
265. Badis, G., et al., *Targeted mRNA degradation by deadenylation-independent decapping*. Mol Cell, 2004. **15**(1): p. 5-15.
266. Muhrad, D. and R. Parker, *The yeast EDC1 mRNA undergoes deadenylation-independent decapping stimulated by Not2p, Not4p, and Not5p*. Embo J, 2005. **24**(5): p. 1033-45.

267. Tomecki, R. and A. Dziembowski, *Novel endoribonucleases as central players in various pathways of eukaryotic RNA metabolism*. *Rna*, 2010. **16**(9): p. 1692-724.
268. Chernokalskaya, E., *et al.*, *A polysomal ribonuclease involved in the destabilization of albumin mRNA is a novel member of the peroxidase gene family*. *Rna*, 1998. **4**(12): p. 1537-48.
269. Yoshida, H., *et al.*, *XBPI mRNA is induced by ATF6 and spliced by IRE1 in response to ER stress to produce a highly active transcription factor*. *Cell*, 2001. **107**(7): p. 881-91.
270. Gill, T., *et al.*, *RNase MRP cleaves the CLB2 mRNA to promote cell cycle progression: novel method of mRNA degradation*. *Mol Cell Biol*, 2004. **24**(3): p. 945-53.
271. Brown, C.E., *et al.*, *PAN3 encodes a subunit of the Pab1p-dependent poly(A) nuclease in Saccharomyces cerevisiae*. *Mol Cell Biol*, 1996. **16**(10): p. 5744-53.
272. Yamashita, A., *et al.*, *Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover*. *Nat Struct Mol Biol*, 2005. **12**(12): p. 1054-63.
273. Tucker, M., *et al.*, *The transcription factor associated Ccr4 and Caf1 proteins are components of the major cytoplasmic mRNA deadenylase in Saccharomyces cerevisiae*. *Cell*, 2001. **104**(3): p. 377-86.
274. Martinez, J., *et al.*, *The mRNA cap structure stimulates rate of poly(A) removal and amplifies processivity of degradation*. *J Biol Chem*, 2001. **276**(30): p. 27923-9.
275. Schwede, A., *et al.*, *A role for Caf1 in mRNA deadenylation and decay in trypanosomes and human cells*. *Nucleic Acids Res*, 2008. **36**(10): p. 3374-88.
276. Utter, C.J., *et al.*, *PolyA-specific ribonuclease (PARN-1) function in stage-specific mRNA turnover in Trypanosoma brucei*. *Eukaryot Cell*, 2011. **10**(9): p. 1230-40.
277. Steiger, M., *et al.*, *Analysis of recombinant yeast decapping enzyme*. *Rna*, 2003. **9**(2): p. 231-8.
278. Yu, J.H., *et al.*, *Ge-1 is a central component of the mammalian cytoplasmic mRNA processing body*. *Rna*, 2005. **11**(12): p. 1795-802.
279. Coller, J.M., *et al.*, *The DEAD box helicase, Dhh1p, functions in mRNA decapping and interacts with both the decapping and deadenylase complexes*. *Rna*, 2001. **7**(12): p. 1717-27.
280. Schwartz, D., C.J. Decker, and R. Parker, *The enhancer of decapping proteins, Edc1p and Edc2p, bind RNA and stimulate the activity of the decapping enzyme*. *Rna*, 2003. **9**(2): p. 239-51.
281. Kshirsagar, M. and R. Parker, *Identification of Edc3p as an enhancer of mRNA decapping in Saccharomyces cerevisiae*. *Genetics*, 2004. **166**(2): p. 729-39.
282. Yang, W.H., *et al.*, *RNA-associated protein 55 (RAP55) localizes to mRNA processing bodies and stress granules*. *Rna*, 2006. **12**(4): p. 547-54.
283. Estevez, A.M., T. Kempf, and C. Clayton, *The exosome of Trypanosoma brucei*. *Embo J*, 2001. **20**(14): p. 3831-9.
284. Haile, S., A.M. Estevez, and C. Clayton, *A role for the exosome in the in vivo degradation of unstable mRNAs*. *Rna*, 2003. **9**(12): p. 1491-501.
285. Estevez, A.M., *et al.*, *The roles of intersubunit interactions in exosome stability*. *J Biol Chem*, 2003. **278**(37): p. 34943-51.
286. Haile, S., *et al.*, *The subcellular localisation of trypanosome RRP6 and its association with the exosome*. *Mol Biochem Parasitol*, 2007. **151**(1): p. 52-8.

287. Cassola, A., J.G. De Gaudenzi, and A.C. Frasch, *Recruitment of mRNAs to cytoplasmic ribonucleoprotein granules in trypanosomes*. Mol Microbiol, 2007. **65**(3): p. 655-70.
288. Holetz, F.B., *et al.*, *Evidence of P-body-like structures in Trypanosoma cruzi*. Biochem Biophys Res Commun, 2007. **356**(4): p. 1062-7.
289. Kramer, S., *et al.*, *The RNA helicase DHH1 is central to the correct expression of many developmentally regulated mRNAs in trypanosomes*. J Cell Sci, 2010. **123**(Pt 5): p. 699-711.
290. Li, C.H., *et al.*, *Roles of a Trypanosoma brucei 5'->3' exoribonuclease homolog in mRNA degradation*. Rna, 2006. **12**(12): p. 2171-86.
291. Kramer, S., *et al.*, *Heat shock causes a decrease in polysomes and the appearance of stress granules in trypanosomes independently of eIF2(alpha) phosphorylation at Thr169*. J Cell Sci, 2008. **121**(Pt 18): p. 3002-14.
292. Houseley, J. and D. Tollervey, *The many pathways of RNA degradation*. Cell, 2009. **136**(4): p. 763-76.
293. Aslett, M., *et al.*, *TriTrypDB: a functional genomic resource for the Trypanosomatidae*. Nucleic Acids Res, 2010. **38**(Database issue): p. D457-62.
294. Milone, J., J. Wilusz, and V. Bellofatto, *Identification of mRNA decapping activities and an ARE-regulated 3' to 5' exonuclease activity in trypanosome extracts*. Nucleic Acids Res, 2002. **30**(18): p. 4040-50.
295. Clayton, C., *mRNA Turnover in Trypanosomes*. RNA Metabolism in Trypanosomes, 2012: p. 79-97.
296. Maquat, L.E., W.Y. Tarn, and O. Isken, *The pioneer round of translation: features and functions*. Cell, 2010. **142**(3): p. 368-74.
297. Hwang, J., *et al.*, *UPF1 association with the cap-binding protein, CBP80, promotes nonsense-mediated mRNA decay at two distinct steps*. Mol Cell, 2010. **39**(3): p. 396-409.
298. Wu, X. and G. Brewer, *The regulation of mRNA stability in mammalian cells: 2.0*. Gene, 2012. **500**(1): p. 10-21.
299. Li, H. and C. Tschudi, *Novel and essential subunits in the 300-kilodalton nuclear cap binding complex of Trypanosoma brucei*. Mol Cell Biol, 2005. **25**(6): p. 2216-26.
300. Bercovich, N., *et al.*, *Identification of core components of the exon junction complex in trypanosomes*. Mol Biochem Parasitol, 2009. **166**(2): p. 190-3.
301. Delhi, P., *et al.*, *Is there a classical nonsense-mediated decay pathway in trypanosomes?* PLoS One, 2011. **6**(9): p. e25112.
302. Thomas, M.G., *et al.*, *RNA granules: the good, the bad and the ugly*. Cell Signal, 2011. **23**(2): p. 324-34.
303. Parker, R. and U. Sheth, *P bodies and the control of mRNA translation and degradation*. Mol Cell, 2007. **25**(5): p. 635-46.
304. Decker, C.J. and R. Parker, *P-bodies and stress granules: possible roles in the control of translation and mRNA degradation*. Cold Spring Harb Perspect Biol, 2012. **4**(9).
305. Holetz, F.B., *et al.*, *Protein and mRNA content of TcDHH1-containing mRNPs in Trypanosoma cruzi*. Febs J, 2010. **277**(16): p. 3415-26.

306. Kramer, S., *et al.*, *Inhibition of mRNA maturation in trypanosomes causes the formation of novel foci at the nuclear periphery containing cytoplasmic regulators of mRNA fate*. J Cell Sci, 2012. **125**(Pt 12): p. 2896-909.
307. Sheth, U., *et al.*, *Perinuclear P granules are the principal sites of mRNA export in adult C. elegans germ cells*. Development, 2010. **137**(8): p. 1305-14.
308. Berezikov, E., *Evolution of microRNA diversity and regulation in animals*. Nat Rev Genet, 2011. **12**(12): p. 846-60.
309. Krol, J., I. Loedige, and W. Filipowicz, *The widespread regulation of microRNA biogenesis, function and decay*. Nat Rev Genet, 2010. **11**(9): p. 597-610.
310. Okamura, K. and E.C. Lai, *Endogenous small interfering RNAs in animals*. Nat Rev Mol Cell Biol, 2008. **9**(9): p. 673-8.
311. Stefani, G. and F.J. Slack, *Small non-coding RNAs in animal development*. Nat Rev Mol Cell Biol, 2008. **9**(3): p. 219-30.
312. Vagin, V.V., *et al.*, *A distinct small RNA pathway silences selfish genetic elements in the germline*. Science, 2006. **313**(5785): p. 320-4.
313. Ngo, H., *et al.*, *Double-stranded RNA induces mRNA degradation in Trypanosoma brucei*. Proc Natl Acad Sci U S A, 1998. **95**(25): p. 14687-92.
314. Lye, L.F., *et al.*, *Retention and loss of RNA interference pathways in trypanosomatid protozoans*. PLoS Pathog, 2010. **6**(10): p. e1001161.
315. DaRocha, W.D., *et al.*, *Tests of cytoplasmic RNA interference (RNAi) and construction of a tetracycline-inducible T7 promoter system in Trypanosoma cruzi*. Mol Biochem Parasitol, 2004. **133**(2): p. 175-86.
316. Robinson, K.A. and S.M. Beverley, *Improvements in transfection efficiency and tests of RNA interference (RNAi) approaches in the protozoan parasite Leishmania*. Mol Biochem Parasitol, 2003. **128**(2): p. 217-28.
317. Djikeng, A., *et al.*, *RNA interference in Trypanosoma brucei: cloning of small interfering RNAs provides evidence for retroposon-derived 24-26-nucleotide RNAs*. Rna, 2001. **7**(11): p. 1522-30.
318. Beverley, S.M., *Protozoemics: trypanosomatid parasite genetics comes of age*. Nat Rev Genet, 2003. **4**(1): p. 11-9.
319. Padmanabhan, P.K., *et al.*, *Novel features of a PIWI-like protein homolog in the parasitic protozoan Leishmania*. PLoS One, 2012. **7**(12): p. e52612.
320. Bevilacqua, A., *et al.*, *Post-transcriptional regulation of gene expression by degradation of messenger RNAs*. J Cell Physiol, 2003. **195**(3): p. 356-72.
321. Chen, C.Y. and A.B. Shyu, *AU-rich elements: characterization and importance in mRNA degradation*. Trends Biochem Sci, 1995. **20**(11): p. 465-70.
322. Bakheet, T., B.R. Williams, and K.S. Khabar, *ARED 3.0: the large and diverse AU-rich transcriptome*. Nucleic Acids Res, 2006. **34**(Database issue): p. D111-4.
323. Barreau, C., L. Paillard, and H.B. Osborne, *AU-rich elements and associated factors: are there unifying principles?* Nucleic Acids Res, 2005. **33**(22): p. 7138-50.
324. Gingerich, T.J., J.J. Feige, and J. LaMarre, *AU-rich elements and the control of gene expression through regulated mRNA stability*. Anim Health Res Rev, 2004. **5**(1): p. 49-63.
325. Schoenberg, D.R. and L.E. Maquat, *Regulation of cytoplasmic mRNA decay*. Nat Rev Genet, 2012. **13**(4): p. 246-59.
326. D'Orso, I., J.G. De Gaudenzi, and A.C. Frasch, *RNA-binding proteins and mRNA turnover in trypanosomes*. Trends Parasitol, 2003. **19**(4): p. 151-5.

327. Quijada, L., *et al.*, *Expression of the human RNA-binding protein HuR in Trypanosoma brucei increases the abundance of mRNAs containing AU-rich regulatory elements*. *Nucleic Acids Res*, 2002. **30**(20): p. 4414-24.
328. Muller, M., P.K. Padmanabhan, and B. Papadopoulou, *Selective inactivation of SIDER2 retroposon-mediated mRNA decay contributes to stage- and species-specific gene expression in Leishmania*. *Mol Microbiol*, 2010. **77**(2): p. 471-91.
329. Muller, M., *et al.*, *Rapid decay of unstable Leishmania mRNAs bearing a conserved retroposon signature 3'-UTR motif is initiated by a site-specific endonucleolytic cleavage without prior deadenylation*. *Nucleic Acids Res*, 2010. **38**(17): p. 5867-83.
330. Muller, M. and B. Papadopoulou, *Stage-specific expression of the glycine cleavage complex subunits in Leishmania infantum*. *Mol Biochem Parasitol*, 2010. **170**(1): p. 17-27.
331. Rochette, A., *et al.*, *Genome-wide gene expression profiling analysis of Leishmania major and Leishmania infantum developmental stages reveals substantial differences between the two species*. *BMC Genomics*, 2008. **9**: p. 255.
332. McNicoll, F., *et al.*, *A combined proteomic and transcriptomic approach to the study of stage differentiation in Leishmania infantum*. *Proteomics*, 2006. **6**(12): p. 3567-81.
333. Shalgi, R., *et al.*, *Widespread regulation of translation by elongation pausing in heat shock*. *Mol Cell*, 2013. **49**(3): p. 439-52.
334. Grousl, T., *et al.*, *Heat shock-induced accumulation of translation elongation and termination factors precedes assembly of stress granules in S. cerevisiae*. *PLoS One*, 2013. **8**(2): p. e57083.
335. Jackson, R.J., C.U. Hellen, and T.V. Pestova, *The mechanism of eukaryotic translation initiation and principles of its regulation*. *Nat Rev Mol Cell Biol*, 2010. **11**(2): p. 113-27.
336. Hotz, H.R., *et al.*, *Mechanisms of developmental regulation in Trypanosoma brucei: a polypyrimidine tract in the 3'-untranslated region of a surface protein mRNA affects RNA abundance and translation*. *Nucleic Acids Res*, 1997. **25**(15): p. 3017-26.
337. Mayho, M., *et al.*, *Post-transcriptional control of nuclear-encoded cytochrome oxidase subunits in Trypanosoma brucei: evidence for genome-wide conservation of life-cycle stage-specific regulatory elements*. *Nucleic Acids Res*, 2006. **34**(18): p. 5312-24.
338. Di Noia, J.M., *et al.*, *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*, 2000. **275**(14): p. 10218-27.
339. D'Orso, I. and A.C. Frasch, *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*, 2002. **277**(52): p. 50520-8.
340. Rodrigues, D.C., *et al.*, *Trypanosoma cruzi: modulation of HSP70 mRNA stability by untranslated regions during heat shock*. *Exp Parasitol*, 2010. **126**(2): p. 245-53.
341. Li, Z.H., *et al.*, *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*, 2012. **287**(23): p. 19058-69.

342. Boucher, N., *et al.*, *A common mechanism of stage-regulated gene expression in Leishmania mediated by a conserved 3'-untranslated region element*. J Biol Chem, 2002. **277**(22): p. 19511-20.
343. Hehl, A., *et al.*, *A conserved stem-loop structure in the 3' untranslated region of procyclin mRNAs regulates expression in Trypanosoma brucei*. Proc Natl Acad Sci U S A, 1994. **91**(1): p. 370-4.
344. Engstler, M. and M. Boshart, *Cold shock and regulation of surface protein trafficking convey sensitization to inducers of stage differentiation in Trypanosoma brucei*. Genes Dev, 2004. **18**(22): p. 2798-811.
345. Colasante, C., *et al.*, *Regulated expression of glycosomal phosphoglycerate kinase in Trypanosoma brucei*. Mol Biochem Parasitol, 2007. **151**(2): p. 193-204.
346. Mani, J., *et al.*, *Alba-domain proteins of Trypanosoma brucei are cytoplasmic RNA-binding proteins that interact with the translation machinery*. PLoS One, 2011. **6**(7): p. e22463.
347. D'Orso, I. and A.C. Frasch, *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, 2001. **276**(19): p. 15783-93.
348. Bayer-Santos, E., *et al.*, *Regulatory elements in the 3' untranslated region of the GP82 glycoprotein are responsible for its stage-specific expression in Trypanosoma cruzi metacyclic trypomastigotes*. Acta Trop, 2012. **123**(3): p. 230-3.
349. Coughlin, B.C., *et al.*, *Amastin mRNA abundance in Trypanosoma cruzi is controlled by a 3'-untranslated region position-dependent cis-element and an untranslated region-binding protein*. J Biol Chem, 2000. **275**(16): p. 12051-60.
350. McNicoll, F., *et al.*, *Distinct 3'-untranslated region elements regulate stage-specific mRNA accumulation and translation in Leishmania*. J Biol Chem, 2005. **280**(42): p. 35238-46.
351. Quijada, L., *et al.*, *Identification of a putative regulatory element in the 3'-untranslated region that controls expression of HSP70 in Leishmania infantum*. Mol Biochem Parasitol, 2000. **110**(1): p. 79-91.
352. Garcia-Estrada, C., *et al.*, *Analysis of genetic elements regulating the methionine adenosyltransferase gene in Leishmania infantum*. Gene, 2007. **389**(2): p. 163-73.
353. Mishra, K.K., *et al.*, *A negative regulatory element controls mRNA abundance of the Leishmania mexicana Paraflagellar rod gene PFR2*. Eukaryot Cell, 2003. **2**(5): p. 1009-17.
354. Holzer, T.R., *et al.*, *Coordinate regulation of a family of promastigote-enriched mRNAs by the 3'UTR PRE element in Leishmania mexicana*. Mol Biochem Parasitol, 2008. **157**(1): p. 54-64.
355. Murray, A., *et al.*, *Regions in the 3' untranslated region confer stage-specific expression to the Leishmania mexicana a600-4 gene*. Mol Biochem Parasitol, 2007. **153**(2): p. 125-32.
356. Purdy, J.E., J.E. Donelson, and M.E. Wilson, *Regulation of genes encoding the major surface protease of Leishmania chagasi via mRNA stability*. Mol Biochem Parasitol, 2005. **142**(1): p. 88-97.
357. Estevez, A.M., *The RNA-binding protein TbDRBD3 regulates the stability of a specific subset of mRNAs in trypanosomes*. Nucleic Acids Res, 2008. **36**(14): p. 4573-86.

358. Stern, M.Z., *et al.*, *Multiple roles for polypyrimidine tract binding (PTB) proteins in trypanosome RNA metabolism.* Rna, 2009. **15**(4): p. 648-65.
359. Archer, S.K., *et al.*, *Trypanosoma brucei PUF9 regulates mRNAs for proteins involved in replicative processes over the cell cycle.* PLoS Pathog, 2009. **5**(8): p. e1000565.
360. Wurst, M., *et al.*, *An RNAi screen of the RRM-domain proteins of Trypanosoma brucei.* Mol Biochem Parasitol, 2009. **163**(1): p. 61-5.
361. Hartmann, C., *et al.*, *Small trypanosome RNA-binding proteins TbUBP1 and TbUBP2 influence expression of F-box protein mRNAs in bloodstream trypanosomes.* Eukaryot Cell, 2007. **6**(11): p. 1964-78.
362. Hartmann, C. and C. Clayton, *Regulation of a transmembrane protein gene family by the small RNA-binding proteins TbUBP1 and TbUBP2.* Mol Biochem Parasitol, 2008. **157**(1): p. 112-5.
363. Ling, A.S., J.R. Trotter, and E.F. Hendriks, *A zinc finger protein, TbZC3H20, stabilizes two developmentally regulated mRNAs in trypanosomes.* J Biol Chem, 2011. **286**(23): p. 20152-62.
364. Das, A., *et al.*, *The essential polysome-associated RNA-binding protein RBP42 targets mRNAs involved in Trypanosoma brucei energy metabolism.* Rna, 2012. **18**(11): p. 1968-83.
365. Walrad, P., *et al.*, *Differential trypanosome surface coat regulation by a CCCH protein that co-associates with procyclin mRNA cis-elements.* PLoS Pathog, 2009. **5**(2): p. e1000317.
366. Walrad, P.B., *et al.*, *The post-transcriptional trans-acting regulator, TbZFP3, coordinates transmission-stage enriched mRNAs in Trypanosoma brucei.* Nucleic Acids Res, 2011. **40**(7): p. 2869-83.
367. Morking, P.A., *et al.*, *The zinc finger protein TcZFP2 binds target mRNAs enriched during Trypanosoma cruzi metacyclogenesis.* Mem Inst Oswaldo Cruz, 2012. **107**(6): p. 790-9.
368. Dallagiovanna, B., *et al.*, *Functional genomic characterization of mRNAs associated with TcPUF6, a pumilio-like protein from Trypanosoma cruzi.* J Biol Chem, 2008. **283**(13): p. 8266-73.
369. Noe, G., J.G. De Gaudenzi, and A.C. Frasch, *Functionally related transcripts have common RNA motifs for specific RNA-binding proteins in trypanosomes.* BMC Mol Biol, 2008. **9**: p. 107.
370. Guerra-Slompo, E.P., *et al.*, *Molecular characterization of the Trypanosoma cruzi specific RNA binding protein TcRBP40 and its associated mRNAs.* Biochem Biophys Res Commun, 2012. **420**(2): p. 302-7.
371. Mitra, B. and D.S. Ray, *Presence of a poly(A) binding protein and two proteins with cell cycle-dependent phosphorylation in Crithidia fasciculata mRNA cycling sequence binding protein II.* Eukaryot Cell, 2004. **3**(5): p. 1185-97.
372. Mahmood, R. and D.S. Ray, *Nuclear extracts of Crithidia fasciculata contain a factor(s) that binds to the 5'-untranslated regions of TOP2 and RPA1 mRNAs containing sequences required for their cell cycle regulation.* J Biol Chem, 1998. **273**(37): p. 23729-34.
373. Kong, J. and P. Lasko, *Translational control in cellular and developmental processes.* Nat Rev Genet, 2012. **13**(6): p. 383-94.

374. Shatsky, I.N., *et al.*, *Cap- and IRES-independent scanning mechanism of translation initiation as an alternative to the concept of cellular IRESs*. *Mol Cells*, 2010. **30**(4): p. 285-93.
375. Shapira, M. and A. Zinoviev, *Translation in Trypanosomatids*. *RNA Metabolism in Trypanosomes*, 2012: p. 187-214.
376. Horn, D., *Codon usage suggests that translational selection has a major impact on protein expression in trypanosomatids*. *BMC Genomics*, 2008. **9**: p. 2.
377. Dhalia, R., *et al.*, *Translation initiation in Leishmania major: characterisation of multiple eIF4F subunit homologues*. *Mol Biochem Parasitol*, 2005. **140**(1): p. 23-41.
378. Yoffe, Y., *et al.*, *Binding specificities and potential roles of isoforms of eukaryotic initiation factor 4E in Leishmania*. *Eukaryot Cell*, 2006. **5**(12): p. 1969-79.
379. Yoffe, Y., *et al.*, *Cap-binding activity of an eIF4E homolog from Leishmania*. *Rna*, 2004. **10**(11): p. 1764-75.
380. Zinoviev, A., *et al.*, *A novel 4E-interacting protein in Leishmania is involved in stage-specific translation pathways*. *Nucleic Acids Res*, 2011. **39**(19): p. 8404-15.
381. Freire, E.R., *et al.*, *The four trypanosomatid eIF4E homologues fall into two separate groups, with distinct features in primary sequence and biological properties*. *Mol Biochem Parasitol*, 2011. **176**(1): p. 25-36.
382. Dhalia, R., *et al.*, *The two eIF4A helicases in Trypanosoma brucei are functionally distinct*. *Nucleic Acids Res*, 2006. **34**(9): p. 2495-507.
383. Bates, E.J., E. Knuepfer, and D.F. Smith, *Poly(A)-binding protein I of Leishmania: functional analysis and localisation in trypanosomatid parasites*. *Nucleic Acids Res*, 2000. **28**(5): p. 1211-20.
384. da Costa Lima, T.D., *et al.*, *Functional characterization of three leishmania poly(a) binding protein homologues with distinct binding properties to RNA and protein partners*. *Eukaryot Cell*, 2010. **9**(10): p. 1484-94.
385. Kramer, S., *et al.*, *Differential localization of the two T. brucei poly(A) binding proteins to the nucleus and RNP granules suggests binding to distinct mRNA pools*. *PLoS One*, 2013. **8**(1): p. e54004.
386. Burgess, H.M., *et al.*, *Nuclear relocalisation of cytoplasmic poly(A)-binding proteins PABP1 and PABP4 in response to UV irradiation reveals mRNA-dependent export of metazoan PABPs*. *J Cell Sci*, 2012. **124**(Pt 19): p. 3344-55.
387. Smith, R.W. and N.K. Gray, *Poly(A)-binding protein (PABP): a common viral target*. *Biochem J*, 2010. **426**(1): p. 1-12.
388. Palfi, Z., W.S. Lane, and A. Bindereif, *Biochemical and functional characterization of the cis-spliceosomal U1 small nuclear RNP from Trypanosoma brucei*. *Mol Biochem Parasitol*, 2002. **121**(2): p. 233-43.
389. Sukarieh, R., N. Sonenberg, and J. Pelletier, *The eIF4E-binding proteins are modifiers of cytoplasmic eIF4E relocalization during the heat shock response*. *Am J Physiol Cell Physiol*, 2009. **296**(5): p. C1207-17.
390. Besse, F. and A. Ephrussi, *Translational control of localized mRNAs: restricting protein synthesis in space and time*. *Nat Rev Mol Cell Biol*, 2008. **9**(12): p. 971-80.
391. Browning, K.S., *et al.*, *Identification of an isozyme form of protein synthesis initiation factor 4F in plants*. *J Biol Chem*, 1992. **267**(14): p. 10096-100.
392. Pestova, T.V., *et al.*, *eIF2-dependent and eIF2-independent modes of initiation on the CSFV IRES: a common role of domain II*. *Embo J*, 2008. **27**(7): p. 1060-72.

393. Joyce, B.R., *et al.*, *Phosphorylation of eukaryotic initiation factor-2{alpha} promotes the extracellular survival of obligate intracellular parasite Toxoplasma gondii*. Proc Natl Acad Sci U S A, 2010. **107**(40): p. 17200-5.
394. Fennell, C., *et al.*, *PfeIK1, a eukaryotic initiation factor 2alpha kinase of the human malaria parasite Plasmodium falciparum, regulates stress-response to amino-acid starvation*. Malar J, 2009. **8**: p. 99.
395. Moraes, M.C., *et al.*, *Novel membrane-bound eIF2alpha kinase in the flagellar pocket of Trypanosoma brucei*. Eukaryot Cell, 2007. **6**(11): p. 1979-91.
396. Cloutier, S., *et al.*, *Translational control through eIF2alpha phosphorylation during the Leishmania differentiation process*. PLoS One, 2012. **7**(5): p. e35085.
397. Raven, J.F. and A.E. Koromilas, *PERK and PKR: old kinases learn new tricks*. Cell Cycle, 2008. **7**(9): p. 1146-50.
398. Tong, L., R.A. Heim, and S. Wu, *Nitric oxide: a regulator of eukaryotic initiation factor 2 kinases*. Free Radic Biol Med, 2011. **50**(12): p. 1717-25.
399. Chow, C., *et al.*, *Promastigote to amastigote differentiation of Leishmania is markedly delayed in the absence of PERK eIF2alpha kinase-dependent eIF2alpha phosphorylation*. Cell Microbiol, 2011. **13**(7): p. 1059-77.
400. Padmanabhan, P.K., *et al.*, *Apoptosis-like programmed cell death induces antisense ribosomal RNA (rRNA) fragmentation and rRNA degradation in Leishmania*. Cell Death Differ, 2012. **19**(12): p. 1972-82.
401. Deribe, Y.L., T. Pawson, and I. Dikic, *Post-translational modifications in signal integration*. Nat Struct Mol Biol, 2010. **17**(6): p. 666-72.
402. Rosenzweig, D., *et al.*, *Post-translational modification of cellular proteins during Leishmania donovani differentiation*. Proteomics, 2008. **8**(9): p. 1843-50.
403. Parsons, M., *et al.*, *Comparative analysis of the kinomes of three pathogenic trypanosomatids: Leishmania major, Trypanosoma brucei and Trypanosoma cruzi*. BMC Genomics, 2005. **6**: p. 127.
404. Genestra, M., L. Cysne-Finkelstein, and L. Leon, *Protein kinase A activity is associated with metacyclogenesis in Leishmania amazonensis*. Cell Biochem Funct, 2004. **22**(5): p. 315-20.
405. Wiese, M., *A mitogen-activated protein (MAP) kinase homologue of Leishmania mexicana is essential for parasite survival in the infected host*. Embo J, 1998. **17**(9): p. 2619-28.
406. Morales, M.A., *et al.*, *Over-expression of Leishmania major MAP kinases reveals stage-specific induction of phosphotransferase activity*. Int J Parasitol, 2007. **37**(11): p. 1187-99.
407. Bengs, F., *et al.*, *LmxMPK9, a mitogen-activated protein kinase homologue affects flagellar length in Leishmania mexicana*. Mol Microbiol, 2005. **55**(5): p. 1606-15.
408. Erdmann, M., *et al.*, *Interacting protein kinases involved in the regulation of flagellar length*. Mol Biol Cell, 2006. **17**(4): p. 2035-45.
409. Rotureau, B., *et al.*, *The flagellum-mitogen-activated protein kinase connection in Trypanosomatids: a key sensory role in parasite signalling and development?* Cell Microbiol, 2009. **11**(5): p. 710-8.
410. Morales, M.A., *et al.*, *Phosphoproteomic analysis of Leishmania donovani pro- and amastigote stages*. Proteomics, 2008. **8**(2): p. 350-63.
411. Nakayasu, E.S., *et al.*, *Phosphoproteomic analysis of the human pathogen Trypanosoma cruzi at the epimastigote stage*. Proteomics, 2009. **9**(13): p. 3489-506.

412. Nett, I.R., *et al.*, *The phosphoproteome of bloodstream form Trypanosoma brucei, causative agent of African sleeping sickness*. Mol Cell Proteomics, 2009. **8**(7): p. 1527-38.
413. Salotra, P., R. Ralhan, and G. Sreenivas, *Heat-stress induced modulation of protein phosphorylation in virulent promastigotes of Leishmania donovani*. Int J Biochem Cell Biol, 2000. **32**(3): p. 309-16.
414. Rivero-Lezcano, O.M., C. Chicharro, and L. Rivas, *Acidic pH stress induces protein tyrosine phosphorylation in Leishmania pifanoi*. Mol Biochem Parasitol, 1997. **84**(1): p. 123-9.
415. Fisk, J.C. and L.K. Read, *Protein arginine methylation in parasitic protozoa*. Eukaryot Cell, 2011. **10**(8): p. 1013-22.
416. McGwire, B.S. and K.P. Chang, *Posttranslational regulation of a Leishmania HEXXH metalloprotease (gp63). The effects of site-specific mutagenesis of catalytic, zinc binding, N-glycosylation, and glycosyl phosphatidylinositol addition sites on N-terminal end cleavage, intracellular stability, and extracellular exit*. J Biol Chem, 1996. **271**(14): p. 7903-9.
417. Weissman, A.M., N. Shabek, and A. Ciechanover, *The predator becomes the prey: regulating the ubiquitin system by ubiquitylation and degradation*. Nat Rev Mol Cell Biol, 2011. **12**(9): p. 605-20.
418. Bhandari, D. and P. Saha, *mRNA cycling sequence binding protein from Leishmania donovani (LdCSBP) is covalently modified by ubiquitination*. FEMS Microbiol Lett, 2007. **273**(2): p. 206-13.
419. Robertson, C.D., *The Leishmania mexicana proteasome*. Mol Biochem Parasitol, 1999. **103**(1): p. 49-60.
420. Silva-Jardim, I., M.F. Horta, and F.J. Ramalho-Pinto, *The Leishmania chagasi proteasome: role in promastigotes growth and amastigotes survival within murine macrophages*. Acta Trop, 2004. **91**(2): p. 121-30.
421. Gille, C., *et al.*, *A comprehensive view on proteasomal sequences: implications for the evolution of the proteasome*. J Mol Biol, 2003. **326**(5): p. 1437-48.
422. Kanemori, M., *et al.*, *Synergistic roles of HslVU and other ATP-dependent proteases in controlling in vivo turnover of sigma32 and abnormal proteins in Escherichia coli*. J Bacteriol, 1997. **179**(23): p. 7219-25.
423. Chrobak, M., *et al.*, *Leishmania donovani HslV does not interact stably with HslU proteins*. Int J Parasitol, 2012. **42**(4): p. 329-39.
424. Mottram, J.C., *et al.*, *The multiple cpb cysteine proteinase genes of Leishmania mexicana encode isoenzymes that differ in their stage regulation and substrate preferences*. J Biol Chem, 1997. **272**(22): p. 14285-93.
425. Schneider, P., *et al.*, *Leishmania major: differential regulation of the surface metalloprotease in amastigote and promastigote stages*. Exp Parasitol, 1992. **75**(2): p. 196-206.
426. Codogno, P., M. Mehrpour, and T. Proikas-Cezanne, *Canonical and non-canonical autophagy: variations on a common theme of self-eating?* Nat Rev Mol Cell Biol, 2011. **13**(1): p. 7-12.
427. Williams, R.A., *et al.*, *Cysteine peptidases CPA and CPB are vital for autophagy and differentiation in Leishmania mexicana*. Mol Microbiol, 2006. **61**(3): p. 655-74.

428. Besteiro, S., *et al.*, *Endosome sorting and autophagy are essential for differentiation and virulence of Leishmania major*. J Biol Chem, 2006. **281**(16): p. 11384-96.
429. Manna, P.T., S. Kelly, and M.C. Field, *Adaptin evolution in kinetoplastids and emergence of the variant surface glycoprotein coat in African trypanosomatids*. Mol Phylogenet Evol, 2013. **67**(1): p. 123-8.
430. Oliveira, T.M., *et al.*, *A novel A2 allele found in Leishmania (Leishmania) infantum chagasi*. Rev Bras Parasitol Vet, 2011. **20**(1): p. 42-8.
431. Zhang, W.W., *et al.*, *Comparison of the A2 gene locus in Leishmania donovani and Leishmania major and its control over cutaneous infection*. J Biol Chem, 2003. **278**(37): p. 35508-15.
432. Azizi, H., *et al.*, *Searching for virulence factors in the non-pathogenic parasite to humans Leishmania tarentolae*. Parasitology, 2009. **136**(7): p. 723-35.
433. Bolhassani, A., *et al.*, *Leishmania major: Protective capacity of DNA vaccine using amastin fused to HSV-1 VP22 and EGFP in BALB/c mice model*. Exp Parasitol, 2011. **128**(1): p. 9-17.
434. Rafati, S., *et al.*, *Amastin peptide-binding antibodies as biomarkers of active human visceral leishmaniasis*. Clin Vaccine Immunol, 2006. **13**(10): p. 1104-10.
435. Stober, C.B., *et al.*, *From genome to vaccines for leishmaniasis: screening 100 novel vaccine candidates against murine Leishmania major infection*. Vaccine, 2006. **24**(14): p. 2602-16.
436. Downing, T., *et al.*, *Whole genome sequencing of multiple Leishmania donovani clinical isolates provides insights into population structure and mechanisms of drug resistance*. Genome Res, 2011. **21**(12): p. 2143-56.
437. Rochette, A., *et al.*, *Characterization and developmental gene regulation of a large gene family encoding amastin surface proteins in Leishmania spp*. Mol Biochem Parasitol, 2005. **140**(2): p. 205-20.
438. Cerqueira, G.C., *et al.*, *Sequence diversity and evolution of multigene families in Trypanosoma cruzi*. Mol Biochem Parasitol, 2008. **157**(1): p. 65-72.
439. Jackson, A.P., *The evolution of amastin surface glycoproteins in trypanosomatid parasites*. Mol Biol Evol, 2009. **27**(1): p. 33-45.
440. Kangussu-Marcolino, M.M., *et al.*, *Distinct genomic organization, mRNA expression and cellular localization of members of two amastin sub-families present in Trypanosoma cruzi*. BMC Microbiol, 2013. **13**: p. 10.
441. Lynn, M.A., A.K. Marr, and W.R. McMaster, *Differential quantitative proteomic profiling of Leishmania infantum and Leishmania mexicana density gradient separated membranous fractions*. J Proteomics, 2013. **82**: p. 179-92.
442. Salotra, P., *et al.*, *Upregulation of surface proteins in Leishmania donovani isolated from patients of post kala-azar dermal leishmaniasis*. Microbes Infect, 2006. **8**(3): p. 637-44.
443. Cruz, M.C., *et al.*, *Trypanosoma cruzi: role of delta-amastin on extracellular amastigote cell invasion and differentiation*. PLoS One, 2012. **7**(12): p. e51804.
444. Teixeira, S.M., *et al.*, *A differentially expressed gene family encoding "amastin," a surface protein of Trypanosoma cruzi amastigotes*. J Biol Chem, 1994. **269**(32): p. 20509-16.

445. Teixeira, S.M., L.V. Kirchhoff, and J.E. Donelson, *Post-transcriptional elements regulating expression of mRNAs from the amastin/tuzin gene cluster of Trypanosoma cruzi*. J Biol Chem, 1995. **270**(38): p. 22586-94.
446. Nozaki, T. and G.A. Cross, *Effects of 3' untranslated and intergenic regions on gene expression in Trypanosoma cruzi*. Mol Biochem Parasitol, 1995. **75**(1): p. 55-67.
447. Abuin, G., et al., *Expression of trans-sialidase and 85-kDa glycoprotein genes in Trypanosoma cruzi is differentially regulated at the post-transcriptional level by labile protein factors*. J Biol Chem, 1999. **274**(19): p. 13041-7.
448. De Gaudenzi, J.G., I. D'Orso, and A.C. Frasch, *RNA recognition motif-type RNA-binding proteins in Trypanosoma cruzi form a family involved in the interaction with specific transcripts in vivo*. J Biol Chem, 2003. **278**(21): p. 18884-94.
449. Cassola, A. and A.C. Frasch, *An RNA recognition motif mediates the nucleocytoplasmic transport of a trypanosome RNA-binding protein*. J Biol Chem, 2009. **284**(50): p. 35015-28.
450. Wu, Y., et al., *A new developmentally regulated gene family in Leishmania amastigotes encoding a homolog of amastin surface proteins*. Mol Biochem Parasitol, 2000. **110**(2): p. 345-57.
451. Aravind, L., L.M. Iyer, and V. Anantharaman, *The two faces of Alba: the evolutionary connection between proteins participating in chromatin structure and RNA metabolism*. Genome Biol, 2003. **4**(10): p. R64.
452. Reddy, T.R. and T. Suryanarayana, *Novel histone-like DNA-binding proteins in the nucleoid from the acidothermophilic archaeobacterium Sulfolobus acidocaldarius that protect DNA against thermal denaturation*. Biochimica et Biophysica Acta, 1988. **949**: p. 87-96.
453. White, M.F. and S.D. Bell, *Holding it together: chromatin in the Archaea*. Trends Genet, 2002. **18**(12): p. 621-6.
454. Xue, H., et al., *An abundant DNA binding protein from the hyperthermophilic archaeon Sulfolobus shibatae affects DNA supercoiling in a temperature-dependent fashion*. J Bacteriol, 2000. **182**(14): p. 3929-33.
455. Lurz, R., et al., *Electron microscopic study of DNA complexes with proteins from the Archaeobacterium Sulfolobus acidocaldarius*. Embo J, 1986. **5**(13): p. 3715-21.
456. Bell, S.D., et al., *The interaction of Alba, a conserved archaeal chromatin protein, with Sir2 and its regulation by acetylation*. Science, 2002. **296**(5565): p. 148-51.
457. Suganuma, T. and J.L. Workman, *Signals and combinatorial functions of histone modifications*. Annu Rev Biochem. **80**(2011): p. 473-99.
458. Yuan, H. and R. Marmorstein, *Structural Basis for Sirtuin Activity and Inhibition*. J Biol Chem, 2012. **287**(51): p. 42428-35.
459. Zhao, K., X. Chai, and R. Marmorstein, *Structure of a Sir2 substrate, Alba, reveals a mechanism for deacetylation-induced enhancement of DNA binding*. J Biol Chem, 2003. **278**(28): p. 26071-7.
460. Marsh, V.L., S.Y. Peak-Chew, and S.D. Bell, *Sir2 and the acetyltransferase, Pat, regulate the archaeal chromatin protein, Alba*. J Biol Chem, 2005. **280**(22): p. 21122-8.
461. Jelinska, C., et al., *Obligate heterodimerization of the archaeal Alba2 protein with Alba1 provides a mechanism for control of DNA packaging*. Structure, 2005. **13**(7): p. 963-71.

462. Jelinska, C., *et al.*, *Dimer-dimer stacking interactions are important for nucleic acid binding by the archaeal chromatin protein Alba*. *Biochem J*, 2010. **427**(1): p. 49-55.
463. Guo, R., H. Xue, and L. Huang, *Ssh10b, a conserved thermophilic archaeal protein, binds RNA in vivo*. *Mol Microbiol*, 2003. **50**(5): p. 1605-15.
464. Wardleworth, B.N., *et al.*, *Structure of Alba: an archaeal chromatin protein modulated by acetylation*. *Embo J*, 2002. **21**(17): p. 4654-62.
465. Hada, K., *et al.*, *Crystal structure and functional analysis of an archaeal chromatin protein Alba from the hyperthermophilic archaeon Pyrococcus horikoshii OT3*. *Biosci Biotechnol Biochem*, 2008. **72**(3): p. 749-58.
466. Ellis, J.C., J. Barnes, and J.W. Brown, *Is Alba an RNase P subunit?* *RNA Biol*, 2007. **4**(3): p. 169-72.
467. Kouzuma, Y., *et al.*, *Reconstitution of archaeal ribonuclease P from RNA and four protein components*. *Biochem Biophys Res Commun*, 2003. **306**(3): p. 666-73.
468. Boomershine, W.P., *et al.*, *Structure of Mth11/Mth Rpp29, an essential protein subunit of archaeal and eukaryotic RNase P*. *Proc Natl Acad Sci U S A*, 2003. **100**(26): p. 15398-403.
469. Tsai, H.Y., *et al.*, *Functional reconstitution and characterization of Pyrococcus furiosus RNase P*. *Proc Natl Acad Sci U S A*, 2006. **103**(44): p. 16147-52.
470. Fukuhara, H., *et al.*, *A fifth protein subunit Ph1496p elevates the optimum temperature for the ribonuclease P activity from Pyrococcus horikoshii OT3*. *Biochem Biophys Res Commun*, 2006. **343**(3): p. 956-64.
471. Olguin-Lamas, A., *et al.*, *A novel Toxoplasma gondii nuclear factor TgNF3 is a dynamic chromatin-associated component, modulator of nucleolar architecture and parasite virulence*. *PLoS Pathog*, 2011. **7**(3): p. e1001328.
472. Treeck, M., *et al.*, *The phosphoproteomes of Plasmodium falciparum and Toxoplasma gondii reveal unusual adaptations within and beyond the parasites' boundaries*. *Cell Host Microbe*, 2011. **10**(4): p. 410-9.
473. Gissot, M., *et al.*, *Toxoplasma gondii Alba Proteins Are Involved in Translational Control of Gene Expression*. *J Mol Biol*, 2013. **425**(8): p. 1287-301.
474. Nebl, T., *et al.*, *Quantitative in vivo analyses reveal calcium-dependent phosphorylation sites and identifies a novel component of the Toxoplasma invasion motor complex*. *PLoS Pathog*, 2011. **7**(9): p. e1002222.
475. Mair, G.R., *et al.*, *Universal features of post-transcriptional gene regulation are critical for Plasmodium zygote development*. *PLoS Pathog*, 2010. **6**(2): p. e1000767.
476. Laurentino, E.C., *et al.*, *Experimentally controlled downregulation of the histone chaperone FACT in Plasmodium berghei reveals that it is critical to male gamete fertility*. *Cell Microbiol*, 2011. **13**(12): p. 1956-74.
477. Solyakov, L., *et al.*, *Global kinomic and phospho-proteomic analyses of the human malaria parasite Plasmodium falciparum*. *Nat Commun*, 2011. **2**: p. 565.
478. Chene, A., *et al.*, *PfAlbas constitute a new eukaryotic DNA/RNA-binding protein family in malaria parasites*. *Nucleic Acids Res*, 2011. **40**(7): p. 3066-77.
479. Goyal, M., *et al.*, *Identification and molecular characterization of an Alba-family protein from human malaria parasite Plasmodium falciparum*. *Nucleic Acids Res*, 2011. **40**(3): p. 1174-90.
480. LaCount, D.J., *et al.*, *A protein interaction network of the malaria parasite Plasmodium falciparum*. *Nature*, 2005. **438**(7064): p. 103-7.

481. Alves, L.R., *et al.*, *Proteomic analysis reveals the dynamic association of proteins with translated mRNAs in Trypanosoma cruzi*. *Gene*, 2010. **452**(2): p. 72-8.
482. Marchini, F.K., *et al.*, *Profiling the Trypanosoma cruzi phosphoproteome*. *PLoS One*, 2011. **6**(9): p. e25381.
483. Subota, I., *et al.*, *ALBA proteins are stage regulated during trypanosome development in the tsetse fly and participate in differentiation*. *Mol Biol Cell*, 2011. **22**(22): p. 4205-19.
484. Nett, I.R., *et al.*, *Identification and specific localization of tyrosine-phosphorylated proteins in Trypanosoma brucei*. *Eukaryot Cell*, 2009. **8**(4): p. 617-26.
485. Zhang, Q., *et al.*, *A critical role of perinuclear filamentous actin in spatial repositioning and mutually exclusive expression of virulence genes in malaria parasites*. *Cell Host Microbe*, 2011. **10**(5): p. 451-63.
486. Lamarque, M., *et al.*, *Food vacuole proteome of the malarial parasite Plasmodium falciparum*. *Proteomics Clin Appl*, 2008. **2**(9): p. 1361-74.
487. Taschner, A., *et al.*, *Nuclear RNase P of Trypanosoma brucei: a single protein in place of the multicomponent RNA-protein complex*. *Cell Rep*, 2012. **2**(1): p. 19-25.
488. Morales, M.A., *et al.*, *Phosphoproteome dynamics reveal heat-shock protein complexes specific to the Leishmania donovani infectious stage*. *Proc Natl Acad Sci U S A*, 2010. **107**(18): p. 8381-6.
489. Jarrous, N. and V. Gopalan, *Archaeal/eukaryal RNase P: subunits, functions and RNA diversification*. *Nucleic Acids Res*, 2010. **38**(22): p. 7885-94.
490. Lai, L.B., *et al.*, *Unexpected diversity of RNase P, an ancient tRNA processing enzyme: challenges and prospects*. *FEBS Lett*, 2010. **584**(2): p. 287-96.
491. Mattijssen, S., *et al.*, *Viperin mRNA is a novel target for the human RNase MRP/RNase P endoribonuclease*. *Cell Mol Life Sci*, 2011. **68**(14): p. 2469-80.
492. Holzmann, J., *et al.*, *RNase P without RNA: identification and functional reconstitution of the human mitochondrial tRNA processing enzyme*. *Cell*, 2008. **135**(3): p. 462-74.
493. Gobert, A., *et al.*, *A single Arabidopsis organellar protein has RNase P activity*. *Nat Struct Mol Biol*, 2010. **17**(6): p. 740-4.
494. Woodhams, M.D., *et al.*, *RNase MRP and the RNA processing cascade in the eukaryotic ancestor*. *BMC Evol Biol*, 2007. **7 Suppl 1**: p. S13.
495. Maida, Y., *et al.*, *An RNA-dependent RNA polymerase formed by TERT and the RMRP RNA*. *Nature*, 2009. **461**(7261): p. 230-5.
496. Reiner, R., *et al.*, *Function and assembly of a chromatin-associated RNase P that is required for efficient transcription by RNA polymerase I*. *PLoS One*, 2008. **3**(12): p. e4072.
497. Perederina, A., *et al.*, *Specific binding of a Pop6/Pop7 heterodimer to the P3 stem of the yeast RNase MRP and RNase P RNAs*. *Rna*, 2007. **13**(10): p. 1648-55.
498. Hands-Taylor, K.L., *et al.*, *Heterodimerization of the human RNase P/MRP subunits Rpp20 and Rpp25 is a prerequisite for interaction with the P3 arm of RNase MRP RNA*. *Nucleic Acids Res*, 2010. **38**(12): p. 4052-66.
499. Welting, T.J., *et al.*, *Heterodimerization regulates RNase MRP/RNase P association, localization, and expression of Rpp20 and Rpp25*. *Rna*, 2007. **13**(1): p. 65-75.
500. Jiang, T. and S. Altman, *Protein-protein interactions with subunits of human nuclear RNase P*. *Proc Natl Acad Sci U S A*, 2001. **98**(3): p. 920-5.

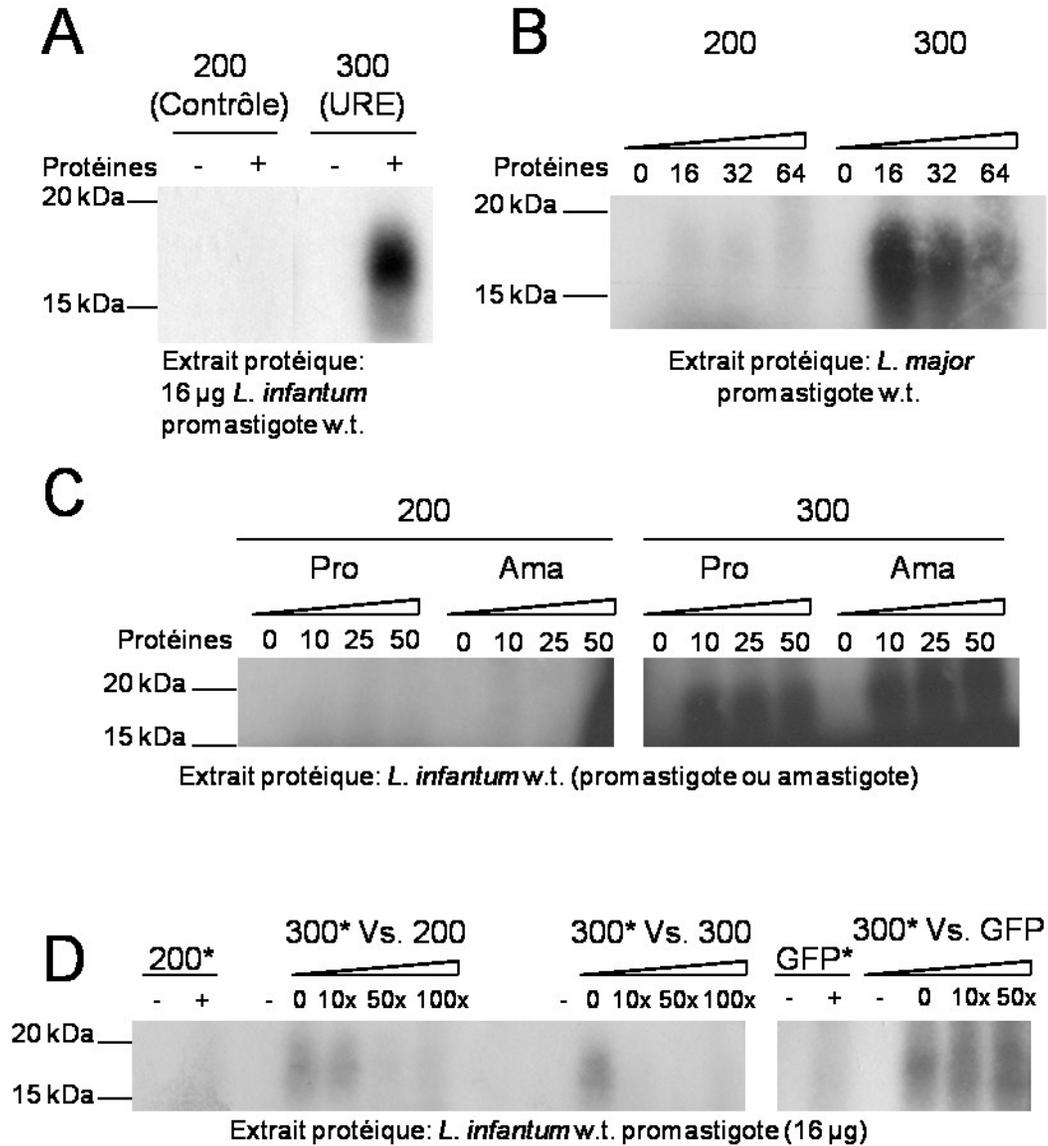
501. Welting, T.J., *et al.*, *Differential association of protein subunits with the human RNase MRP and RNase P complexes*. *Rna*, 2006. **12**(7): p. 1373-82.
502. Choudhary, M.K., *et al.*, *Dehydration-responsive nuclear proteome of rice (*Oryza sativa* L.) illustrates protein network, novel regulators of cellular adaptation, and evolutionary perspective*. *Mol Cell Proteomics*, 2009. **8**(7): p. 1579-98.
503. Doroshenk, K.A., *et al.*, *Proteomic analysis of cytoskeleton-associated RNA binding proteins in developing rice seed*. *J Proteome Res*, 2009. **8**(10): p. 4641-53.
504. Thomson, A.M., *et al.*, *Optimized RNA gel-shift and UV cross-linking assays for characterization of cytoplasmic RNA-protein interactions*. *Biotechniques*, 1999. **27**(5): p. 1032-9, 1042.
505. Komar, A.A., B. Mazumder, and W.C. Merrick, *A new framework for understanding IRES-mediated translation*. *Gene*, 2012. **502**(2): p. 75-86.
506. D'Orso, I. and A.C. Frasch, *TcUBP-I, a developmentally regulated U-rich RNA-binding protein involved in selective mRNA destabilization in trypanosomes*. *J Biol Chem*, 2001. **276**(37): p. 34801-9.
507. Daneshvar, H., *et al.*, *Comparative proteomics profiling of a gentamicin-attenuated *Leishmania infantum* cell line identifies key changes in parasite thiol-redox metabolism*. *J Proteomics*, 2012. **75**(5): p. 1463-71.
508. Vasconcelos, E.J., *et al.*, *In silico identification of conserved intercoding sequences in *Leishmania* genomes: unraveling putative cis-regulatory elements*. *Mol Biochem Parasitol*, 2012. **183**(2): p. 140-50.
509. Szoor, B., *et al.*, *A novel phosphatase cascade regulates differentiation in *Trypanosoma brucei* via a glycosomal signaling pathway*. *Genes Dev*, 2010. **24**(12): p. 1306-16.
510. Stewart, M., *et al.*, *Molecular mechanism of translocation through nuclear pore complexes during nuclear protein import*. *FEBS Lett*, 2001. **498**(2-3): p. 145-9.
511. Fernandez-Moya, S.M., *et al.*, *Alterations in DRBD3 ribonucleoprotein complexes in response to stress in *Trypanosoma brucei**. *PLoS One*, 2012. **7**(11): p. e48870.
512. Schumann Burkard, G., *et al.*, *Nucleolar proteins regulate stage-specific gene expression and ribosomal RNA maturation in *Trypanosoma brucei**. *Mol Microbiol*, 2013. **88**(4): p. 827-40.





## Annexe 2.

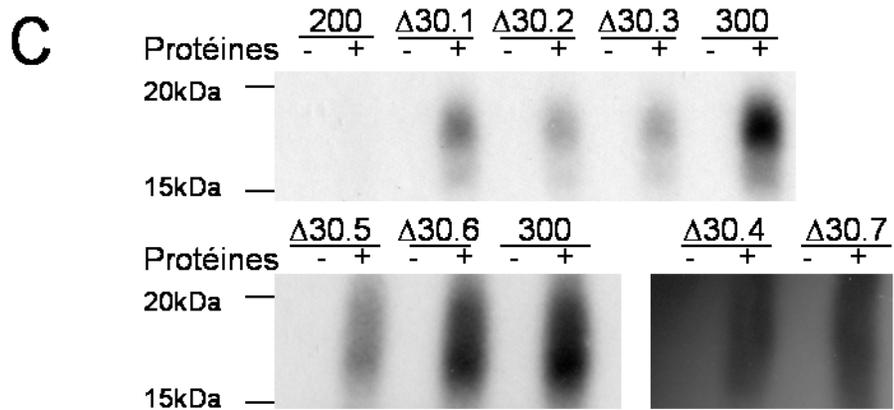
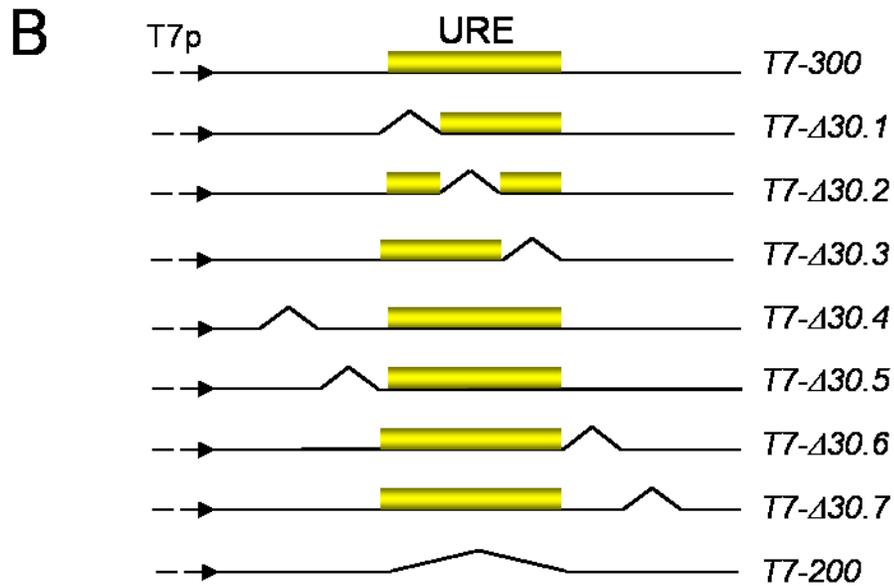
La liaison covalente aux UV montre l'interaction spécifique de l'URE de l'amastine avec une protéine de 17 kDa.





## Annexe 3.

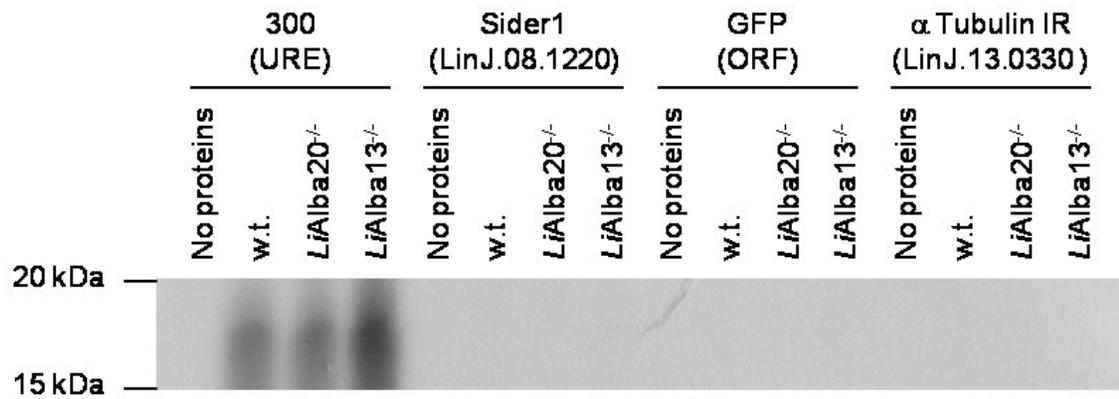
Des délétions supplémentaires dans la région de 300 nt incluant l'URE suggèrent que la protéine lie des séquences riches en uridines.





## Annexe 4.

La protéine visualisée par liaison covalente aux UV comme partenaire du URE n'est pas une protéine à domaine Alba.





## Annexe 5.

Motifs communs des 3'UTR des 30  $\delta$ -amastines obtenus par analyse MEME/MAST.

	Sites	E-value		Sites	E-value
#1 	30	3.0e-082	#14 	30	1.3e-048
#2 	29	5.5e-054	#15 	30	1.5e-048
#3 	22	3.3e-053	#16 	22	3.7e-043
#4 	29	2.0e-069	#17 	24	2.1e-047
#5 	25	3.2e-054	#18 	18	6.0e-041
#6 	30	1.3e-051	#19 	30	2.0e-037
#7 	28	7.0e-048	#20 	30	3.6e-055
#8 	30	2.0e-058	#21 	30	1.6e-057
#9 	21	1.4e-045	#22 	30	3.4e-049
#10 	26	2.6e-060	#23 	30	2.3e-041
#11 	29	1.7e-059	#24 	30	9.2e-033
#12 	22	8.4e-043	#25 	19	4.9e-035
#13 	30	7.9e-037			







## Annexe 7.

Structure secondaire prédite de l'URE (en bleu) présent dans le 3'UTR de l'amastine LinJ.34.1010 et position des motifs avoisinants.

