



**HAL**  
open science

**Participation à l'étude de la régulation transcriptionnelle des gènes de Plasmodium falciparum lors du cycle érythrocytaire. Caractérisation de facteurs nucléaires des familles Myb et HMGB.**

Sylvie Briquet

► **To cite this version:**

Sylvie Briquet. Participation à l'étude de la régulation transcriptionnelle des gènes de Plasmodium falciparum lors du cycle érythrocytaire. Caractérisation de facteurs nucléaires des familles Myb et HMGB.. Biochimie [q-bio.BM]. Université de Versailles-Saint Quentin en Yvelines, 2006. Français. NNT: . tel-00086444

**HAL Id: tel-00086444**

**<https://theses.hal.science/tel-00086444>**

Submitted on 18 Jul 2006

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**THESE DE DOCTORAT  
DE L'ECOLE PRATIQUE DE HAUTES ETUDES  
ET DE L'UNIVERSITE DE VERSAILLES SAINT-QUENTIN**

**Sciences de la Vie et de la Terre**  
Spécialité : Génétique Moléculaire

Présentée par

**Sylvie BRIQUET**

Pour obtenir le grade de  
**DOCTEUR DE L'EPHE ET DE L'UVSQ**

**Participation à l'étude de la régulation transcriptionnelle des gènes de  
*Plasmodium falciparum* lors du cycle érythrocytaire.  
Caractérisation de facteurs nucléaires des familles Myb et HMGB.**

Date de soutenance : le 21 Juin 2006

Devant le jury composé de

<b>Pr. Bernard Mignotte</b>	<b>Président du jury</b>
<b>Pr. Michel Raymondjean</b>	<b>Rapporteur</b>
<b>Dr. Robert Ménard</b>	<b>Rapporteur</b>
<b>Dr. Odile Puijalon</b>	<b>Examineur</b>
<b>Dr. Catherine Vaquero</b>	<b>Directeur de thèse</b>

Laboratoire de Génétique Moléculaire et Physiologique de l'Ecole Pratique des Hautes Etudes, Université de Versailles/Saint-Quentin, LGBC, CNRS UPRES-A8087, Bâtiment Fermat - 45, avenue des Etats Unis 78035 Versailles. Directeur EPHE : B. Mignotte, e-mail : [bernard.mignotte@uvsq.fr](mailto:bernard.mignotte@uvsq.fr)

Laboratoire d'Immunobiologie Cellulaire et Moléculaire, INSERM, U511, CHU Pitié-Salpêtrière, 91 bd de l'Hôpital, 75013 Paris. Directeur de thèse : C. Vaquero, e-mail : [vaquero@ext.jussieu.fr](mailto:vaquero@ext.jussieu.fr)

*A mon grand-père, Avo Manuel,  
Pour l'amour dans son regard.*

## **Remerciements**

*Ce travail a été effectué à l'INSERM Unité U511 d'Immunobiologie Cellulaire et Moléculaire des Infections Parasitaires, dirigée par Dominique Mazier que je sais gré de m'avoir permis de préparer cette thèse. Merci pour la confiance que vous m'accordez depuis 8 ans dans votre laboratoire et pour votre soutien lors des moments difficiles.*

*Je remercie mes rapporteurs Messieurs Michel Raymondjean et Robert Ménard pour l'intérêt qu'ils ont porté à ce travail en acceptant de le juger. Je suis reconnaissante également à Madame Odile Puijalon qui a accepté d'être examinateur du jury de thèse.*

*Je tiens à remercier vivement Monsieur Bernard Mignotte, président du Jury, qui a suivi mon parcours à l'EPHE avec attention en contribuant à ma formation par la richesse de l'enseignement de l'Ecole Doctorale qu'il dirige à l'Université de Versailles-St-Quentin.*

*J'exprime toute ma gratitude à Catherine Vaquero, mon directeur de thèse d'avoir mener mes pas vers le chemin de la thèse et de m'avoir guidée avec rigueur, passion et confiance mais plus largement merci pour ton encadrement scientifique depuis plus de douze ans. Ce travail n'aurait pu aboutir sans tes précieux conseils et nos discussions animées, sans compter la liberté encadrée que tu m'as accordée et ton encouragement à aller toujours plus loin. Merci pour le temps que tu as consacré à redonner un peu de rigueur à ma plume et en général pour ta disponibilité et ta générosité.*

*Un grand merci,  
à notre équipe passée et actuelle qui a contribué à ce travail...j'ai profité pleinement des compétences et de la sympathique présence de Charlotte Boschet, Mathieu Gissot, Philippe Refour et puis de la participation récente mais non moins active d'Emilie Tissandié, Elisa Sevilla, Agostino Martinez, Nadou Lawson et Olivia Azlor,  
à tous les membres du laboratoire qui ont contribué de près ou de loin à ce projet, notamment Liliane Cicéron qui a assouvi maintes fois nos besoins gourmands en parasites, à Maurel Tefit et à Jean-François Franetich pour la production d'anticorps,  
à Abiba Doukani et à Jean-Pierre Lagarde qui ont partagé les efforts nécessaires au séquençage des régions complexes de *P. falciparum*,  
à nos collaboratrices de l'Institut Pasteur, Isabelle Thiery et Catherine Bourgoïn pour leur efficacité,  
à tous mes collègues amis qui ont partagé une pause de thé ou de chocolat, pour l'amitié et le soutien de Maryse, Nicolas, Carine, Eliane, Yvette, Saddi et biensûr de tous les étudiants*

*Enfin, j'adresse une pensée reconnaissante à ma famille pour m'avoir encouragée tout au long de ce projet et en particulier à mon époux Bruno et à ma fille Inès pour leur réconfort au cours de ces derniers mois.*

# TABLE DES MATIERES

TABLE DES MATIERES .....	1
TABLE DES ILLUSTRATIONS .....	4
FIGURES.....	4
TABLEAUX.....	4
LISTE DES ABREVIATIONS .....	5
INTRODUCTION.....	6
LE PALUDISME ET <i>PLASMODIUM FALCIPARUM</i> .....	7
I – Historique.....	7
II – Taxinomie et biodiversité.....	10
III - Le paludisme.....	11
III.1 - Symptômes .....	11
III.2 - Répartition des espèces.....	12
III.3 - Paludisme d'importation ou des voyageurs .....	12
III.4 – Transmission et population exposée.....	13
III.5 – Traitement et recherche d'un vaccin.....	15
IV - <i>Plasmodium falciparum</i> .....	16
IV. 1 – Origine.....	16
IV. 2 - Cycle de vie du parasite .....	17
IV.2.1 - Le cycle chez l'anophèle, l'hôte invertébré.....	19
IV.2.2 - Le cycle chez l'homme, l'hôte vertébré.....	20
IV.2.3 - Description cellulaire et moléculaire des stades sanguins de <i>P. falciparum</i> .....	23
REGULATION DE L'EXPRESSION DES GENES.....	29
I – Régulation des gènes chez les eucaryotes.....	29
I.1 – L'organisation de la chromatine .....	29
I.2 – La structure des gènes.....	30
I.3 – Les promoteurs et leurs éléments de régulation .....	30
I.4 – La transcription.....	31
I.4 - La régulation de la transcription.....	33
I.4.1 – Influence de la chromatine .....	33
a) Modifications des histones ou code des histones .....	33
b) Méthylation de l'ADN.....	34
c) Remodelage dépendant de l'ATP.....	34
d) Remodelage architectural : les facteurs de la famille HMGB.....	35
I.4.2 – Les facteurs de transcription .....	40
a) Protéines à domaines basiques .....	40
b) Protéines à motif en doigt de zinc .....	41
c) Protéines à motif hélice-tour-hélice : les facteurs de la famille Myb.....	41
d) Protéines à architecture $\beta$ .....	43
I.4.3 – Schéma de la régulation transcriptionnelle chez les eucaryotes .....	44
II- Le génome de <i>P. falciparum</i> .....	44
II.1 – Composition et organisation du génome.....	45
II.2 – Structure des chromosomes.....	46

III – La transcription chez <i>P. falciparum</i> .....	47
III.1 – Organisation des gènes .....	47
III.2 – Composition en protéines assignées à la régulation génétique.....	47
III.2.1 – La machinerie générale de la transcription .....	48
III.2.2 – Régulateurs de la structure de la chromatine.....	50
III.2.3 – Facteurs de transcription spécifiques de séquence.....	51
IV – La régulation génétique chez <i>P. falciparum</i> .....	53
IV.1 - L'expression des gènes plasmodiaux.....	53
IV.2 - La régulation transcriptionnelle .....	55
IV.2.2 - Les éléments de régulation.....	56
IV.2.1 – Les facteurs de transcription .....	59
IV.3 - La régulation post-transcriptionnelle .....	60
IV.3.1 – La régulation de l'épissage et de la stabilité des ARNm .....	60
IV.3.2 – Les ARN anti-sens.....	61
<b>PRESENTATION DES TRAVAUX DE THESE .....</b>	<b>63</b>
<b>MATERIELS ET METHODES .....</b>	<b>65</b>
<b>I – MATERIELS.....</b>	<b>66</b>
I.1 – Clones de <i>Plasmodium falciparum</i> .....	66
I.2 – Souches bactériennes et plasmides .....	66
I.3 – Anticorps .....	66
I.4 – Oligonucléotides.....	67
<b>II – METHODES .....</b>	<b>68</b>
II.1 – Culture cellulaire.....	68
II.1.1 - Culture continue .....	68
II.1.2 – Synchronisation au sorbitol .....	69
II.1.3 - Cryopréservation et décongélation .....	69
II.2 – Extraction des acides nucléiques et clonage.....	69
II.3 – Extraction et analyse des ARN totaux.....	70
II.4 – Expression et purification de protéines recombinantes.....	71
II.5 – Immunoprécipitation de protéines et western blot en série .....	72
II.6 – Localisation de protéines par immunofluorescence.....	72
II.7 – Retard sur gel ou EMSA (Electrophoretic Mobility Shift Assay) .....	73
II.7.1 – Expérience de liaison aux séquences MRE .....	73
II.7.2 – Expérience de liaison à l'ADN cruciforme .....	74
II.7.3 – Expérience de courbure de l'ADN .....	75
<b>RESULTATS.....</b>	<b>76</b>
<b>I – ANALYSE COMPARATIVE DU TRANSCRIPTOME DE DEUX CLONES 3D7 ET F12 DE <i>P. FALCIPARUM</i>, LORS DU CYCLE ERYTHROCYTAIRE. ....</b>	<b>77</b>
I.1 - Résumé .....	77
<b>ARTICLE 1:.....</b>	<b>79</b>
Transcriptome of 3D7 and its gametocyte-less derivative F12 <i>Plasmodium falciparum</i> clones during erythrocytic development using a gene-specific microarray assigned to gene regulation, cell cycle and transcription factors. ....	79
<b>II – CARACTERISATION FONCTIONNELLE ET ROLE DE PEMYB1, UN FACTEUR DE TRANSCRIPTION SPECIFIQUE DE SEQUENCE, DANS LA REGULATION DES GENES AU COURS DU CYCLE ERYTHROCYTAIRE DE DEUX CLONES DE <i>P. FALCIPARUM</i>, 3D7 ET F12. ....</b>	<b>91</b>

II.1 – Caractérisation fonctionnelle de PfMyb1.....	91
II.1.1 Résumé .....	91
<b>ARTICLE 2:</b> .....	<b>93</b>
Characterization of PfMyb1 transcription factor during erythrocytic development of 3D7 and F12 <i>Plasmodium falciparum</i> clones.....	93
II.1.2 – Résultats complémentaires .....	98
a) Expression de la protéine PfMyb1 recombinante.....	98
b) Etude de la cinétique d’expression du transcrit <i>pfmyb1</i> .....	99
c) Etude de la cinétique d’expression de la protéine PfMyb1 .....	101
II.2 – Rôle de PfMyb1 lors du cycle érythrocytaire et début de caractérisation de son réseau de régulation.....	103
II.2.1 Résumé .....	103
<b>ARTICLE 3:</b> .....	<b>105</b>
PfMyb1, a <i>Plasmodium falciparum</i> transcription factor, is required for intra-erythrocytic growth and controls key genes for cell cycle regulation.....	105
<b>III – CARACTERISATION FONCTIONNELLE DE DEUX FACTEURS NUCLEAIRES ARCHITECTURAUX DE <i>P. FALCIPARUM</i> PFHMGB1 ET PFHMGB2.</b> .....	<b>119</b>
III.1 – Résumé.....	119
<b>ARTICLE 4 :</b> .....	<b>121</b>
The High Mobility Group Box (HMGB) nuclear factors of <i>Plasmodium falciparum</i> . .....	121
<b>DISCUSSION</b> .....	<b>132</b>
<b>REFERENCES BIBLIOGRAPHIQUES</b> .....	<b>147</b>
<b>ANNEXES</b> .....	<b>166</b>
<b>I - NUMEROS D’ACCESSION DES PROTEINES LIEES A LA REGULATION DE L’EXPRESSION DES GENES CHEZ <i>P. FALCIPARUM</i>.</b> .....	<b>166</b>
I.1 – Protéines identifiées par nos soins .....	166
I.2 – Protéines identifiées par Coulson et collaborateurs (Coulson et al. 2004).....	169
<b>II - DONNEES SUPPLEMENTAIRES DE L’ARTICLE 1.</b> .....	<b>170</b>
<b>III – DONNEES SUPPLEMENTAIRES DE L’ARTICLE 4.</b> .....	<b>173</b>
III.1 - Donnée supplémentaire 1. Arbre phylogénétique de 159 domaines “HMG-box”. .....	173
III.1 - Donnée supplémentaire 2. Arbre phylogénétique de 34 boîtes A et boîtes B. ....	175
III.3 - Donnée supplémentaire 3. Modèles de structures des protéines PfHMGB1 et PfHMGB2 obtenus par homologie avec la protéine HMG1 du hamster chinois <i>Cricetulus griseus</i> . ....	176
<b>LISTE DES PUBLICATIONS.</b> .....	<b>177</b>

# TABLE DES ILLUSTRATIONS

## FIGURES

Figure 1. Le vecteur du paludisme: <i>Anopheles gambiae</i> lors d'un repas sanguin. ....	13
Figure 2. Distribution globale de l'endémie palustre due à <i>P. falciparum</i> . ....	14
Figure 3. Arbre phylogénétique construit à partir des séquences du gène SSU rRNA de onze espèces de <i>Plasmodium</i> . ....	16
Figure 4. Cycle de vie de <i>Plasmodium falciparum</i> . ....	18
Figure 5. Gamétocytes puis exflagellation de <i>P. falciparum</i> visualisés après coloration au Giemsa d'un frottis sanguin, par microscopie optique (x 1000). ....	19
Figure 6. Représentation du cycle érythrocytaire détaillé de <i>P. falciparum</i> . ....	22
Figure 7. Formes cellulaires érythrocytaires de <i>P. falciparum</i> visualisées après coloration au Giemsa d'un frottis sanguin, par microscopie optique (x 1000), (Silamut et al. 1999). ....	23
Figure 8. Nucléosome et chromatine. ....	29
Figure 9. Assemblage de complexes nucléoprotéiques par l'intermédiaire de la protéine HMGB. ....	36
Figure 10. Structure du domaine "HMG box" du facteur HMG1 de <i>Cricetulus griseus</i> , le hamster chinois. ....	37
Figure 11. Représentation schématique de l'alignement multiple des séquences de facteurs architecturaux. ....	38
Figure 12. Représentation schématique des trois domaines fonctionnels des protéines. ....	43
Figure 13. Machinerie basale de la transcription, éléments de régulation et facteurs de transcription (Tjian 1995). ....	44
Figure 14. Cartes du vecteur de clonage pCR II-TOPO et du vecteur d'expression pQE30. ....	66
Figure 15. Analyse de la protéine PfMyb1 recombinante. ....	99
Figure 16. Analyse de l'expression temporelle des ARNm de <i>pfmyb1</i> . ....	100
Figure 17. Analyse de l'expression temporelle de la protéine PfMyb1. ....	102

## TABLEAUX

Tableau 1. Classification des Apicomplexa. ....	10
Tableau 2. Données générales sur les génomes de <i>P. falciparum</i> et de cinq autres eucaryotes unicellulaires (Abrahamsen et al. 2004). ....	45
Tableau 3. Pourcentage de protéines assignées à la régulation de la transcription pour 8 génomes d'eucaryotes (Coulson et al. 2004). ....	47



## LISTE DES ABREVIATIONS

A	adénine
A <sub>260 nm</sub>	absorbance à 260 nm
ADN	acide désoxyribonucléique
ADNc	ADN complémentaire
ARN	acide ribonucléique
ARNdb	ARN double brin
ARNm	ARN messenger
ARNr	ARN ribosomique
ARNt	ARN de transfert
ATP	adénosine tri-phosphate
BSA	"bovine serum albumine", albumine de serum bovin
cpm	coup par minute
DAPI	4', 6-diamidino-2-phenylindole
DDT	dichlorodiphényltrichloroéthane
DEPC	diéthylpyrocarbonate
DNase	désoxyribonucléase
dNTP	désoxyribonucléotide tri-phosphate
DTT	dithiotreitol
EDTA	éthylène diamine tétra-acétate
EMSA	"electrophoretic mobility shift assay", expérience de retard sur gel
FITC	fluorescein isothiocyanate
FT	facteur de transcription
h	heure
HCA	hydrophobic cluster analysis
His	histidine
HMG	high mobility group
HSP	heat shock protein
HTH	hélice-tour-hélice
Ig	immunoglobuline
IPTG	isopropyl-1-thio-β-D-galactopyranoside
kb	kilobase
kDa	kiloDalton
LB	Luria Bertani
Mb	megabase
μCi	micro Curie
μg	microgramme
μl	microlitre
μmole	micromole
min	minute
ml	millilitre
MRE	"Myb regulatory element", élément de régulation Myb
NAP	"nucleosome assembly protein", protéine d'assemblage du nucléosome
nm	nanomètre
p.	page
PAGE	"polyacrylamide gel electrophoresis", électrophorèse sur gel de polyacrylamide
pb	paire de base
PBS	"phosphate buffer saline", tampon phosphate salin
PCR	"polymerase chain reaction", réaction de polymérisation en chaîne
p.i	post-invasion
PMSF	phénylméthylsulfonyl fluoride
PVDF	polyvinylidifluoride
rpm	rotation par minute
RPMI	milieu "Roswell Park Memorial Institute"
RT	"reverse transcription", transcription inverse
SDS	sodium dodécyl sulfate
T	thymine
TAF	"TBP associated factors", facteurs associés à la TBP
TBE	Tris borate EDTA
TBP	"TATA-binding protein", protéine liant la boîte TATA
UTR	"untranslated region", région non traduite
UV	ultraviolet
V	volt

# **INTRODUCTION**

## LE PALUDISME ET *PLASMODIUM FALCIPARUM*

Le paludisme est la maladie infectieuse la plus importante affectant l'humanité, après la tuberculose. Près de 40% de la population mondiale vit dans une zone à risque et 1,5 à 2,7 millions de personnes meurent chaque année, dont 1 million d'enfants de moins de 5 ans soit 1 enfant toutes les 30 secondes (<http://www.who.int/malaria>). Les autres personnes à haut risque d'infection sont les femmes enceintes et les voyageurs non-immunisés.

Parmi les quatre espèces de *Plasmodium* parasitant l'homme, *P. falciparum* est de loin l'espèce la plus pathogène, induisant des atteintes neurologiques (neuropaludisme ou malaria cérébrale), hématologiques (anémie sévère) et / ou des complications durant la grossesse (paludisme gestationnel).

La situation est d'autant plus préoccupante que depuis plusieurs années, il y a coïncidence entre le développement grandissant de résistances des parasites face aux médicaments et des moustiques face aux insecticides, du changement global du climat, de la pauvreté continue et de l'instabilité politique, principalement en Afrique. Aucun vaccin n'est aujourd'hui disponible.

### I – Historique

Le paludisme est une maladie transmise par un moustique appelé "anophèle" et est causé par un petit parasite protozoaire du genre *Plasmodium* qui infecte alternativement l'homme et le moustique. C'est une maladie très ancienne et on pense que l'homme préhistorique a dû en souffrir (Desowitz 1991).

On rapporte l'existence de fièvres mortelles, que l'on pense causées par le paludisme, depuis que l'écriture existe (soit vers 6000 ou 5500 av. J.-C.). Les textes védiques datant de 1600 av. J.-C., en Inde, et ceux d'Hippocrate, du Ve siècle av. J.-C., en font mention.

La maladie est probablement originaire d'Afrique et a suivi les migrations humaines vers les côtes de la Méditerranée, jusqu'en Inde et en Asie du Sud-Est. Dans le passé, le paludisme était fréquent dans les marais Pontins, autour de Rome et son nom a été tiré de l'italien (mal-aria ou "mauvais air"). Il était aussi connu sous le nom de fièvre romaine.

Les "traités de médecine" des Mayas et des Aztèques ne mentionnent pas le paludisme. Il semble que les colonisateurs européens et les esclaves aient apporté le paludisme dans le Nouveau Monde, où l'anophèle existait déjà.

Il y a plus de 350 ans, les Amérindiens employaient déjà le quinquina ou *Cinchona*, un arbre d'Amérique du Sud, sous forme de décoction d'écorce pour traiter le paludisme. Ayant constaté les propriétés antipaludiques de l'écorce de quinquina, les missionnaires jésuites venus en Amérique du Sud les ont fait connaître en Europe dès 1630, et en Inde, dès 1657 (Talisuna et al. 2004). En 1820, les pharmaciens parisiens Pelletier et Caventou ont extrait et identifié le principe actif du quinquina : la quinine, un alcaloïde végétal toxique.

En 1880, Charles-Louis-Alphonse Laveran, un médecin de l'armée française (prix Nobel de médecine en 1907), a découvert le parasite responsable du paludisme en examinant des échantillons de sang au microscope, en Algérie. Sa découverte a alors été rejetée par le milieu médical et ce n'est qu'en 1886 qu'elle fut acceptée par les scientifiques italiens, les chefs de file dans le domaine à cette époque (Nosny 1980).

En 1882, on a émis l'hypothèse de la transmission de la maladie par un moustique.

Le 18 décembre 1897, le British Medical Journal signalait que le docteur Ronald Ross avait découvert des kystes paludéens dans les parois de l'estomac d'anophèles ayant piqué un malade atteint de paludisme (Dobson 1999).

En juillet 1898, Giovanni Batista Grassi, un scientifique italien, a décrit les transformations du parasite dans le moustique et prouvé que le paludisme observé chez l'homme était transmis par des moustiques du genre *Anopheles* (Fantini 1999).

La quinine devint introuvable lorsque Java tomba aux mains des Japonais durant la Deuxième Guerre mondiale. Il fallait donc mettre au point de toute urgence un antipaludique synthétique. Contrairement à la quinine, tirée de l'écorce d'un arbre, la chloroquine est un produit synthétique qui appartient aux composés dits amino-4-quinoléines. La chloroquine a été mise au point par une société pharmaceutique allemande, en 1934.

En 1950, des projets pilotes de pulvérisation de DDT pour lutter contre le paludisme ont pris forme. L'Organisation Mondiale de la Santé (OMS) a lancé des programmes d'éradication du paludisme à l'échelle mondiale au milieu des années 1950 (Giglioli 1951).

Dans les années 1960, des souches de *Plasmodium falciparum* résistantes à la chloroquine sont apparues à cause de son utilisation excessive et, probablement, de doses insuffisantes. À ce moment-là, il n'y avait pas de médicament susceptible de traiter ces formes de paludisme résistantes à la chloroquine sauf l'antipaludique le plus ancien, la quinine. Un dérivé de quinoléine a été synthétisé, il s'agit de la méfloquine dont la structure chimique ressemble à celle de la quinine. Le qinghaosu, dérivé d'une plante (*Artemisia annua* L), employé depuis plus de deux mille ans en Chine pour traiter les fièvres associées au paludisme a donné naissance à un « nouvel » antipaludique : l'artémisine. On a démontré l'efficacité de ce médicament contre les formes mortelles de paludisme à *Plasmodium falciparum* et contre les souches de *P. falciparum* résistantes à la chloroquine (Pe Than and Tin 1986).

De 1955 jusqu'aux années 1970, l'U.S. Agency for International Aid (USAID) a versé environ un milliard de dollars à l'OMS et à différents programmes nationaux d'éradication du paludisme. En 1967, ayant réalisé qu'il était impossible d'éradiquer le paludisme dans le monde pour différentes raisons, l'OMS a mis en oeuvre des moyens afin de limiter l'ampleur de la maladie. En 1972, devant son échec, l'OMS a mis fin au programme mondial d'éradication du paludisme.

Les premières tentatives de développer un vaccin datent des années 70. Le vaccin SPf 66 a été le premier à être testé sur le terrain en 1988 (Millet et al. 1993). Il a produit une faible réduction des cas cliniques en Amérique latine mais n'a fait preuve d'aucune efficacité en Afrique. Depuis, les essais vaccinaux se sont multipliés mais font face à la complexité du développement du parasite chez ses deux hôtes successifs, l'homme et le moustique ainsi qu'à un mécanisme d'échappement au système immunitaire extrêmement compliqué.

En 2001, l'OMS proclame la période 2001-2010, la décennie des Nations Unies pour faire reculer le paludisme dans les pays en voie de développement, particulièrement en Afrique (Decade to Roll Back Malaria). Le programme Roll Back Malaria constitue la seule tentative coordonnée contre la maladie (WHO 2000).

Les projets de séquençage du génome de *Plasmodium falciparum* et d'*Anopheles gambiae* (principal moustique vecteur du paludisme) ont été initiés en 1996 et 1998, respectivement, par un Consortium International des Centres de Génome et la séquence des génomes complets ont été publiés en 2002 (Gardner et al. 2002; Holt et al. 2002).

Depuis 100 ans, tous les efforts visant à maîtriser le paludisme ont échoué. La séquence des génomes du *Plasmodium* et de l'anophèle marque un tournant et permettra d'explorer des aspects encore inconnus de la biologie du parasite et de son vecteur afin de promouvoir le développement de nouvelles stratégies destinées à contrôler le paludisme, y compris des diagnostics, vaccins et médicaments.

## II – Taxinomie et biodiversité.

Laveran appela l'organisme qu'il observa le 6 Novembre 1880, *Laverania falcipara*, qui est devenu plus tard *Plasmodium falciparum*. Ce sont surtout des caractères morphologiques et des particularités du cycle biologique qui ont été utilisés comme critères taxinomiques dans la classification classique de Garnham. Aujourd'hui, *Plasmodium falciparum* prend place dans le règne des protistes (sous Alveolata sous Eukaryota) et appartient au Phylum des organismes parasites obligatoires, les Apicomplexa (anciennement Sporozoa, "capables de germer d'autres cellules") (Vivier 1989) et à l'ordre des Haemosporida. En terme morphologique, les Apicomplexa présentent un complexe apical, outil qui leur permet de pénétrer à l'intérieur de la cellule hôte.

**Tableau 1. Classification des Apicomplexa.**

<b>Classification des Apicomplexa (Sporozoa), Selon Karapelou 1987.</b>			
<b>Classe</b>	<b>Ordre</b>	<b>Famille</b>	<b>Genre</b>
Haemospridea	Haemosporida	Plasmodiidae	Plasmodium
Coccidea	Eimeriida	Eimeriidae	Eimeria
		Cryptosporiidae	Cryptosporidium
		Sarcocystiidae	Sarcocystis
Piroplasmidea	Piroplasmida	Babesiidae	Toxoplasma
		Theileriidae	Babesia
			Theileria

Le genre *Plasmodium* comprend près de 200 espèces constituant un groupe de 9 sous-genres de parasites intracellulaires obligatoires de vertébrés, capables de se développer uniquement chez les mammifères (3 sous-genres), les oiseaux (4 sous-genres) et les lézards (2 sous-genres) (Rich and Ayala 2003). Pour tous les *Plasmodium* parasites de primates et de rongeurs, les moustiques vecteurs appartiennent exclusivement au genre *Anopheles*. Les parasites des humains et des primates font tous partie soit du genre *Plasmodium* (*Plasmodium*) soit du sous-genre *P. (Laverania)*, alors que toutes les autres espèces infectant les mammifères font partie du sous-genre hétérogène *P. (Vinckeia)*.

Toutes les hémospories se trouvent dans le sang de vertébrés, comme le suggère leur nom. Elles ont toutes deux hôtes : un hôte vertébré chez lequel se passe la reproduction asexuée (l'hôte intermédiaire) et un insecte piqueur chez lequel se passe la reproduction sexuée (l'hôte définitif).

### **III - Le paludisme**

#### **III.1 - Symptômes**

Les manifestations cliniques du paludisme sont très diverses. Le paludisme débute par une fièvre 8 à 30 jours après l'infection, qui peut s'accompagner - ou non - de maux de tête, de douleurs musculaires, d'un affaiblissement, de vomissements, de diarrhées, de toux. Des cycles typiques alternant fièvre, tremblements avec sueurs froides et transpiration intense, peuvent alors survenir : c'est "l'accès palustre".

La périodicité de ces cycles dépend de l'espèce de parasite en cause, et coïncide avec la multiplication des parasites et l'éclatement des globules rouges, qui conduit également à l'anémie. Classiquement, *P. falciparum* provoque la fièvre tierce maligne (survenant toutes les 48 h), *P. ovale* et *P. vivax* provoquent la fièvre tierce bénigne, enfin, *P. malariae* provoque la fièvre quarte (survenant toutes les 72 h). Le paludisme à *P. falciparum* peut être exempt de ces accès, mais peut être fatal s'il n'est pas traité. Dans certains cas, les globules rouges infectés peuvent bloquer les vaisseaux sanguins irriguant le cerveau : c'est l'accès pernicieux ou neuropaludisme, souvent mortel. Dans les régions où le paludisme est hautement endémique, les personnes régulièrement piquées par des moustiques infectés deviennent

naturellement immunisées ("immunité acquise") et sont alors des porteurs asymptomatiques de l'infection.

### III.2 - Répartition des espèces

*Plasmodium vivax* est le moins exigeant en température et sa présence est la plus étendue. L'éradication est pratiquement achevée en Europe. Il est présent dans le Bassin méditerranéen (Turquie, Moyen Orient, Afrique du Nord), dans toute la partie tropicale de l'Asie, peu important en Afrique tropicale, absent de l'Afrique de l'Ouest, présent à Madagascar, Ile Maurice et Comores et dans les régions de basse altitude de l'Amérique Centrale et du Sud.

*Plasmodium ovale* est l'espèce la plus rare. Il est présent surtout en Afrique Centrale, ne tue pas mais peut entraîner des rechutes 4 à 5 ans après la primo infection (risque de reviviscence).

*Plasmodium malariae* est le plus fréquent en Afrique tropicale, présente quelques foyers en Afrique du Nord mais aussi en Amérique Centrale et du Sud (mer des Caraïbes et Golfe du Mexique) et en Asie (Iran). Cette espèce n'est pas meurtrière mais peut entraîner des rechutes jusqu'à 20 ans après la primo infection.

*Plasmodium falciparum*, l'espèce la plus pathogène et la seule responsable des cas mortels est prépondérante dans les régions tropicales d'Afrique, d'Amérique Latine et d'Asie.

### III.3 - Paludisme d'importation ou des voyageurs

La France est la nation européenne la plus touchée par le paludisme d'importation, avec plus de 7000 cas estimés en 2001, dont 90 % en provenance d'Afrique (Centre National de Référence de l'Epidémiologie du Paludisme Importé et Autochtone). 80 % des cas enregistrés en France sont dus à *P. falciparum*. La mortalité est de l'ordre de 1 % à cause du neuropaludisme. Malgré tout, *P. falciparum* ne s'implante pas car la température est insuffisante et faute de vecteur local compatible alors que *P. vivax* supporte des températures plus basses et peut être transmis par un vecteur européen potentiel tel que *Anopheles labranchiae* comme cela a été décrit en Italie (Baldari et al. 1998).



### III.4 – Transmission et population exposée

Le paludisme est transmis à l'homme par la piqûre d'un moustique femelle, elle-même infectée après avoir piqué un homme impaludé : la femelle, en prenant le repas de sang nécessaire à la complétion de son cycle ovogénique, injecte le parasite à son hôte. La transmission de *Plasmodium* d'un homme à un autre se fait donc par l'intermédiaire du moustique, le principal en cause étant *Anopheles gambiae*. Il existe un seul cas de contamination inter-humaine directe, lorsqu'une femme enceinte infectée contamine son enfant par voie transplacentaire. Le paludisme transfusionnel est rare, mais il demeure une complication grave potentielle à cause de la persistance asymptomatique du parasite de certains donneurs. De la même manière, une contamination lors d'une greffe d'organe reste possible (Fischer et al. 1999).



**Figure 1. Le vecteur du paludisme: *Anopheles gambiae* lors d'un repas sanguin.**

Le rôle des vecteurs est très important dans l'épidémiologie des maladies, dans lesquelles ils interviennent. Généralement, ils focaliseront la maladie, c'est-à-dire que leur présence, indispensable à la transmission, décidera de sa présence ou de son absence dans une zone géographique.

Partout où les conditions de milieu permettent l'implantation de l'anophèle et l'accomplissement de son cycle de reproduction, le paludisme sévit à l'état endémique et atteint souvent des populations rurales et pauvres des pays en développement. Les zones tropicales et subtropicales sont particulièrement favorables à sa transmission. Les risques ne sont pas également répartis. Les facteurs climatiques (température et pluviométrie),

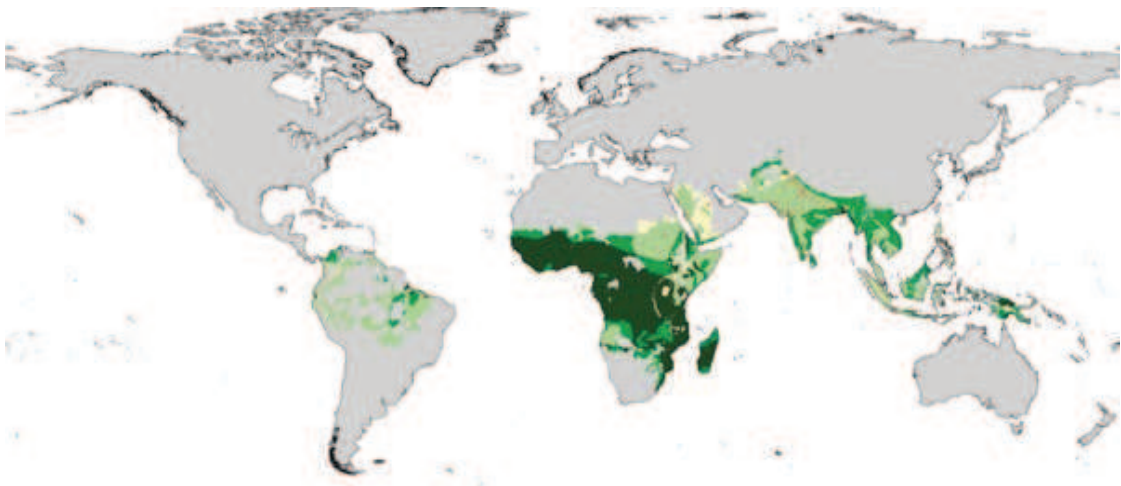
l'environnement et la biogéographie conditionnent la distribution des espèces d'anophèles et modulent l'intensité de la transmission. Dans les zones intertropicales chaudes et humides, les anophèles abondent en permanence. En zones subtropicales ou tempérées chaudes la transmission du paludisme n'est possible qu'à la belle saison, époque favorable au développement du moustique. Le risque varie, dans une même région, selon les époques et le lieu où l'on se trouve :

. la nuit : le risque est à son maximum. L'activité maximale des moustiques se situe entre 22 h et 4 h du matin.

. les zones rurales : le risque est bien plus important en milieu rural qu'en ville. On ne rencontre pas de paludisme dans les villes d'Asie et d'Amérique du Sud (contrairement à l'Afrique subsaharienne).

. la saison des pluies : Le risque s'accroît à la fin de la saison des pluies, lorsque le taux d'humidité élevé favorise la prolifération des moustiques.

. l'altitude: le risque change aussi selon l'altitude, l'anophèle ne peut guère se reproduire au-dessus de 1500 m en Afrique et au-dessus de 2500 m en Asie et en Amérique.



**Figure 2. Distribution globale de l'endémie palustre due à *P. falciparum*.**

Classes d'endémie : en vert clair, zones où la prévalence de l'infection infantile < 10 %, en vert, zones où la prévalence de l'infection infantile est comprise entre 11 % et 50 % et en vert foncé, zones où la prévalence de l'infection infantile > 50 %. En jaune : ces zones non classées représentant 6 % de la population à risque sont dues à une contradiction avec le tracé des limites de risque d'endémie établi en 2002. En gris, zones en dehors des limites de transmission et zones dont la densité de population est inférieure à 1 personne/km<sup>2</sup> (Snow et al. 2005).

Bien entendu, dans le cas du paludisme, et plus largement lorsqu'il s'agit de zoonoses (maladie commune à l'homme et à l'animal), des paramètres tenant à l'homme lui-même vont s'ajouter à ceux qu'imposent le cycle et ses protagonistes naturels. Parmi ces éléments, les conditions de vie des populations sont extrêmement importantes : en effet, un grand nombre de parasitoses sont des maladies de la pauvreté. En particulier, parmi les parasitoses qui sévissent dans les régions chaudes, leur répartition n'a pas toujours été limitée à ces zones, mais elles ont disparu des pays développés sous l'effet de l'amélioration des conditions de vie et d'hygiène.

### **III.5 – Traitement et recherche d'un vaccin**

La quinine est restée pendant longtemps le traitement de choix de l'accès palustre par sa rapidité d'action. Lors du traitement par la quinine ou ses dérivés, ces produits se concentrent à l'intérieur des érythrocytes et tuent le parasite à son stade de développement schizonte.

L'apparition de la résistance des anophèles aux insecticides suite aux méthodes d'éradication de masse avec notamment l'utilisation du dichlorodiphényltrichloroéthane (DDT, 1955-1969) et la chimiorésistance de *P. falciparum* ont contribué à la détérioration de la situation du paludisme ces trente dernières années. La résistance de *P. falciparum* à la chloroquine est apparue vers 1957 à la frontière entre la Thaïlande et le Cambodge et en Colombie (Payne 1987) puis s'est étendue dans toutes les directions de l'Asie du Sud-Est, Centrale et Occidentale et de l'Amérique du Sud. A la fin des années 70, la chloroquinorésistance s'est identifiée en Afrique de l'Est et se propage depuis lors d'Est en Ouest à travers le continent (Wernsdorfer 1991).

Face à la chimiorésistance grandissante de *P. falciparum*, on ne dispose que de peu de composés anti-paludiques : chloroquine, amodiaquine, pyriméthamine-sulfadoxine, quinine, méfloquine, halofantrine et artémisine. Aucun de ces médicaments n'égalent la chloroquine en coût, sécurité et disponibilité sur le marché.

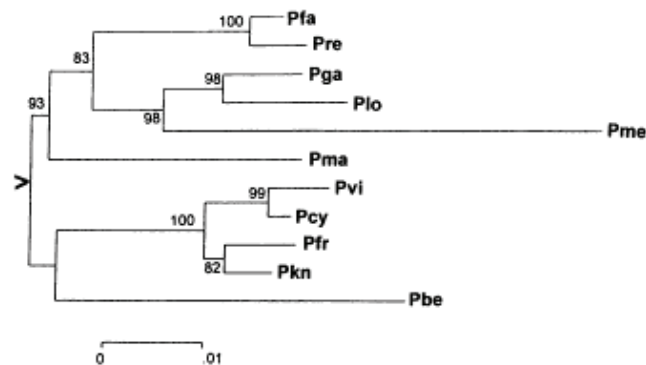
Par ailleurs, la lutte anti-vectorielle a échoué dans la plupart des pays tropicaux et est aujourd'hui pratiquement abandonnée. Enfin, une des difficultés majeures dans la mise au point d'un vaccin contre *Plasmodium* est, qu'au cours de sa vie, le parasite passe successivement par plusieurs stades avec des phases d'intense multiplication asexuée chez

l'homme (dans les cellules du foie -phase hépatique - puis dans les globules rouges du sang - phase érythrocytaire) et une phase de reproduction sexuée suivie de multiplication, chez l'insecte. Chaque cycle se termine par la libération d'un parasite d'une forme différente, donc porteur d'antigènes différents et induisant des réponses immunitaires différentes, ce qui complique d'autant la recherche d'un vaccin. Un des freins majeurs à la recherche d'un vaccin antipaludique reste l'absence de modèle animal : les rongeurs ne sont pas sensibles aux mêmes souches que l'homme, et les singes ne développent pas de pathologie cérébrale comme l'espèce humaine.

## IV - *Plasmodium falciparum*

### IV. 1 – Origine

*P. falciparum* a pour ancêtre d'origine *P. reichenowi*, un parasite de chimpanzé décrit entre 1917-1920 par Reichenow (Reichenow 1920). On ne peut distinguer ces 2 espèces morphologiquement mais elles présentent cependant une forte spécificité d'hôte, elles auraient divergé il y a 5-10 millions d'années en même temps que les chimpanzés et les hominidés. Le clade *falciparum-reichenowi* est éloigné de *P. malariae* et *P. vivax*, autres espèces de parasites humains, elles aussi éloignées l'une de l'autre mais il est monophylétique avec les parasites aviaires, *P. gallinaceum* et *P. lophurae*.



**Figure 3. Arbre phylogénétique construit à partir des séquences du gène SSU rRNA de onze espèces de *Plasmodium*.**

Pfa: *P. falciparum*; Pre: *P. reichenowi*; Pga: *P. gallinaceum*; Plo: *P. lophurae*; Pme: *P. mexicanum*; Pma: *P. malariae*; Pvi: *P. vivax*; Pcy: *P. cynomolgi*; Pfr: *P. fragile*; Pkn: *P. knowlesi*; Pbe: *P. berghei* (Escalante and Ayala 1994).

## IV. 2 - Cycle de vie du parasite

Le cycle des *Plasmodium* est complexe et comporte deux phases essentielles (Figure 4)

:

une phase de différenciation sexuée chez le moustique, l'hôte définitif qui héberge la forme adulte. Cette phase sexuée, associée à un arrêt de prolifération et initiée chez l'homme aboutit à la fécondation ; elle est responsable de la dissémination du parasite par le moustique,

une phase de prolifération asexuée chez le moustique puis chez l'homme, l'hôte intermédiaire qui héberge la forme immature. Cette phase asexuée comprend 3 étapes de multiplication intense : la sporogonie, la schizogonie hépatique et la schizogonie érythrocytaire. La schizogonie érythrocytaire est responsable des manifestations cliniques du paludisme (fièvre et anémie).

L'homme constitue le réservoir du parasite.

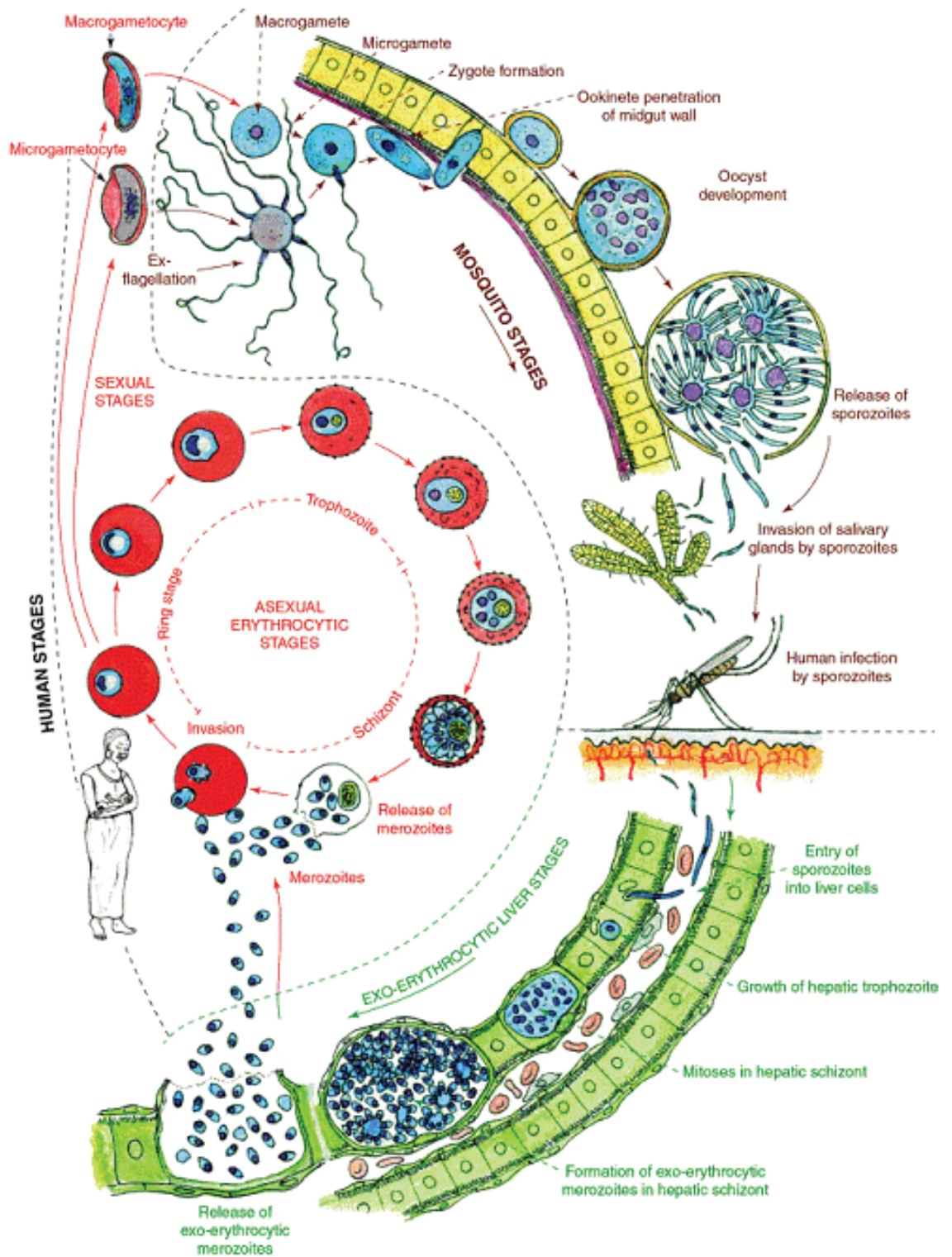


Figure 4. Cycle de vie de *Plasmodium falciparum*.

(Bannister and Mitchell 2003).

#### IV.2.1 - Le cycle chez l'anophèle, l'hôte invertébré

Chez l'anophèle femelle s'effectue le cycle sexué ou fécondation. En prenant son repas sanguin sur l'homme impaludé, l'anophèle absorbe des gamétocytes qui assurent la poursuite du cycle (Figure 4 et 5).

Dans l'estomac du moustique, les gamétocytes mâles et femelles se transforment très rapidement en microgamètes mâles et macrogamètes femelles, c'est la **gamétogénèse**. L'**exflagellation** a lieu entre 7 et 10 min après le début du repas sanguin. Pendant cette période incroyablement courte, trois divisions chromosomiques ont lieu dans le noyau, en même temps que l'assemblage de novo de 8 axonèmes typiques conduisant à la formation de huit gamètes flagellés par gamétocyte. La diminution de température, l'élévation de pH et l'exposition à des molécules du moustique tel l'acide xanthurénique, sont autant de facteurs impliqués dans l'induction de l'exflagellation (Billker et al. 1998; Garcia et al. 1998; Nijhout and Carter 1978; Sinden 1983).

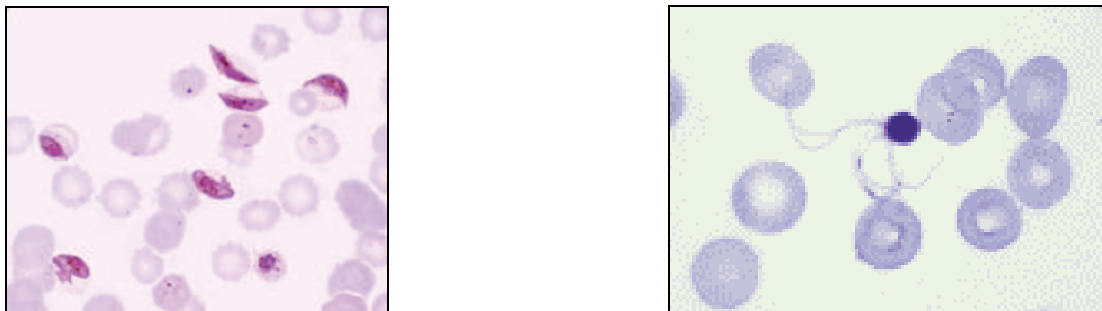


Figure 5. Gamétocytes puis exflagellation de *P. falciparum* visualisés après coloration au Giemsa d'un frottis sanguin, par microscopie optique (x 1000).

La **fécondation** a lieu lorsque l'un des microgamètes fusionne avec un macrogamète pour former d'abord un zygote, puis un **ookinète** très mobile. Suite à la fusion du petit noyau mâle et du gros noyau femelle, le génome diploïde est immédiatement réduit au stade haploïde au moyen d'une méiose (Janse et al. 1986; Sinden and Hartley 1985; Sinden et al. 1985). Au cours de cette méiose, les deux paires de chromosomes plasmodiaux viennent en contact, donnant lieu à des recombinaisons et des réarrangements inter- et intra-chromosomiques.

L'ookinète, forme invasive asexuée, traverse activement la paroi de l'estomac de l'anophèle et se fixe au niveau de sa face externe formant l'oocyste, dans lequel s'individualisent les **sporozoïtes** fusiformes (8 à 12 µm de long) : c'est la sporogonie, première étape de multiplication asexuée. Libérés par l'éclatement de l'oocyste, ces derniers gagnent avec prédilection les glandes salivaires de l'anophèle. Après maturation des sporozoïtes par l'expression différenciée de certains gènes, le moustique devient alors infectieux et le reste pendant environ un mois (Kaiser et al. 2004).

Au final, le repas sanguin (1 à 2 µl) ingéré par la femelle anophèle contenant 1 à 10<sup>5</sup> gamétocytes conduira à la formation de 12 macrogamètes, dont 5-6 deviendront des ookinètes, et 2 se développeront en oocystes 2-7 jours plus tard. A partir de 16000 sporozoïtes issus de ces 2 oocystes, seulement 10-20 seront inoculés par l'anophèle infectée lors de ses repas suivants. Cet important goulet d'étranglement suggère que la différenciation du parasite est sévèrement contrainte par l'environnement du moustique et donc par les interactions hôte-parasite.

#### IV.2.2 - Le cycle chez l'homme, l'hôte vertébré.

L'anophèle femelle injecte à l'homme, lors d'un nouveau repas sanguin, les sporozoïtes qui en une demi-heure migrent, *via* la circulation sanguine, vers le foie pour entamer la **schizogonie hépatique** ou **schizogonie exo-érythrocytaire**. Après plusieurs migrations à travers l'épithélium, les sporozoïtes migrent le long de l'endothélium pour traverser les cellules macrophagiques de Kupffer. Ils pénètrent ensuite plusieurs hépatocytes avant de trouver l'hépatocyte où il démarrera son développement (Frevert 2004). Ils envahissent ensuite les cellules hépatiques, où ils se "cachent" sous le nom de cryptozoïtes. Après 6-7 jours, ils se différencient en **schizontes** hépatiques matures (40 à 100 µ) et se divisent très activement pour donner naissance, en quelques jours, à des dizaines de milliers de nouveaux parasites : les "**mérozoïtes**", forme infectieuse pour les érythrocytes. Les cellules du foie éclatent en libérant ces parasites dans le sang : là, ils pénètrent à l'intérieur des globules rouges et débutent le cycle érythrocytaire.

Les rechutes tardives de paludisme observées lors d'infections par *P. vivax* et *P. ovale* sont dues à la possibilité pour les stades cryptozoïtes de subsister sous une forme latente ("**hypnozoïte**") dans la cellule hépatique de l'homme. Les rechutes tardives des personnes



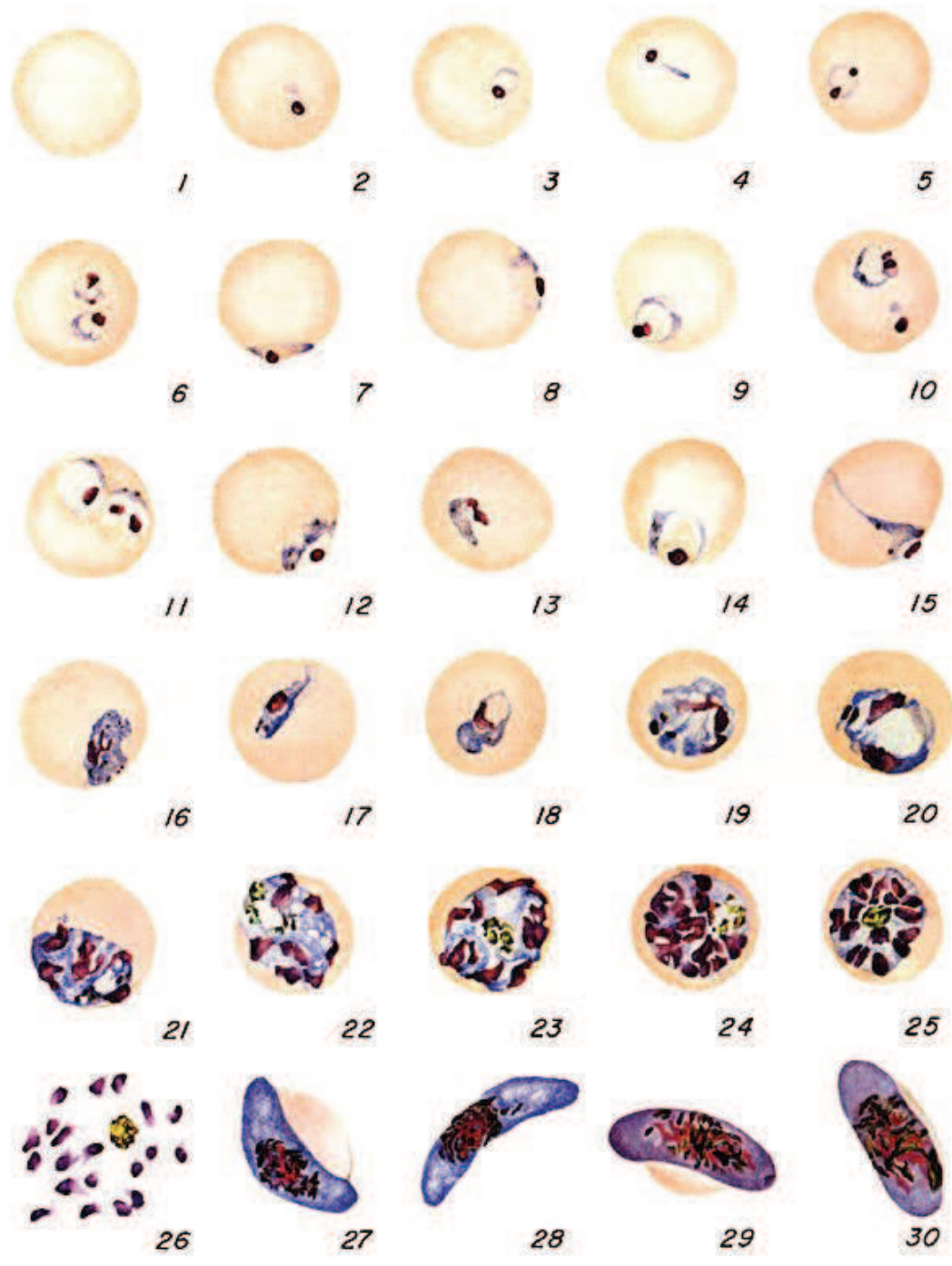
infectées par *P. malariae* résultent, quant à elles, de la persistance de formes "quiescentes" dans le réseau lymphatique humain. *P. falciparum* ne présente ni hypnozoïtes ni schizogonie tissulaire secondaire.

Lorsque les globules rouges parasités éclatent à leur tour, les mérozoïtes libérés dans la circulation sanguine infectent de nouveaux globules rouges ou érythrocytes. C'est le début du cycle asexué érythrocytaire ou **schizogonie érythrocytaire**.

Après l'invasion des érythrocytes par les mérozoïtes, le développement du parasite commence par le stade **anneau**, qui tient son nom de sa morphologie après coloration au Giemsa (Figure 6, stades 2-10). Il évolue en forme **trophozoïte** mesurant 2 à 3 µm et possède une volumineuse vacuole nutritive qui refoule en périphérie son cytoplasme et son noyau (Figure 6, stades 11-18). Il grossit, et son noyau se divise c'est alors un **schizonte** qui se charge de pigment malarique brun ou hemozoïne résultant de la digestion des constituants de l'érythrocyte, notamment l'hémoglobine. La multiplication des noyaux dont chacun s'entoure d'une plage cytoplasmique forme un schizonte mûr ou un corps en rosace (Figure 6, stades 19-26). Le corps en rosace dilaté et mûr éclate, puis libère 16 à 32 mérozoïtes qui vont parasiter de nouveaux érythrocytes et effectuer de nouveaux cycles schizogoniques érythrocytaires.

La lyse des érythrocytes est relativement synchronisée et chaque cycle érythrocytaire dure 48 heures pour *P. falciparum* (Figure 7), rythmant ainsi les accès thermiques périodiques (fièvres tierces).

Après plusieurs cycles schizogoniques, certains trophozoïtes vont être détournés du cycle érythrocytaire pour former des éléments à potentiel sexué (mâle et femelle), et se différencier en **gamétocytes**, première étape de la phase sexuée de l'hématozoaire ou **gamétocytogénèse** (Figure 6, stades 27-30). A maturité, les gamétocytes restent dans le sang périphérique de l'homme et ils ne vont continuer leur développement que s'ils sont absorbés par un anophèle ; c'est donc en changeant d'hôte à partir de ce stade bloqué que *Plasmodium* peut continuer son cycle et assurer sa dissémination.

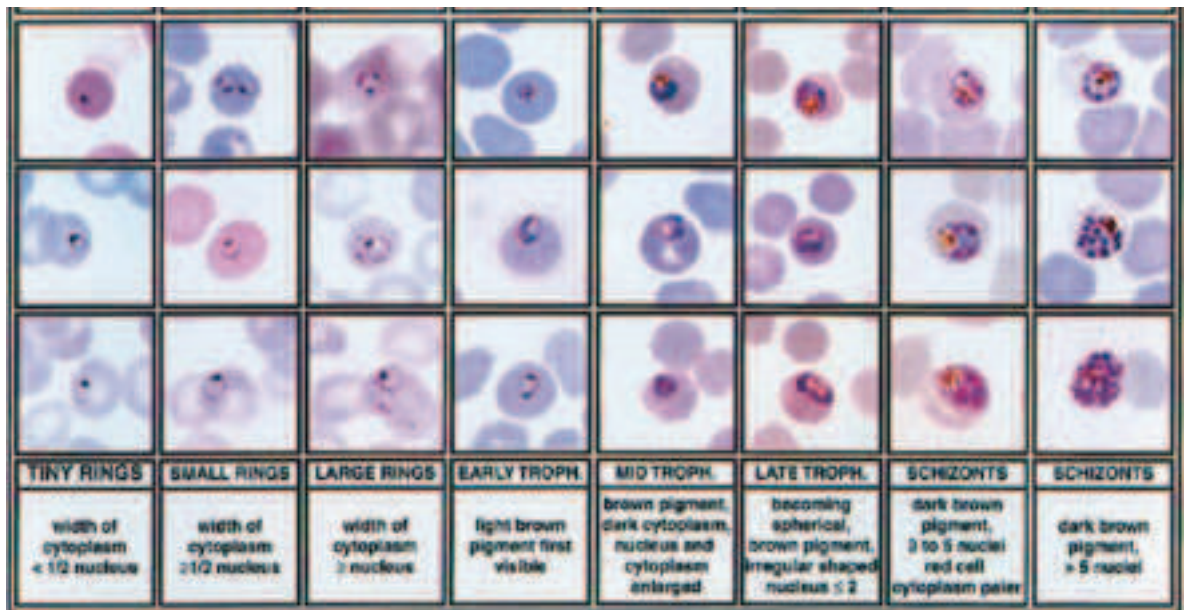


0 10 μ

PLASMODIUM FALCIPARUM

J.H. Nicholson

Figure 6. Représentation du cycle érythrocytaire détaillé de *P. falciparum*.  
 D'après (Coatney et al. 1971).



**Figure 7. Formes cellulaires érythrocytaires de *P. falciparum* visualisées après coloration au Giemsa d'un frottis sanguin, par microscopie optique (x 1000), (Silamut et al. 1999).**

Chacune des étapes du développement de *Plasmodium* comprend une séquence complexe d'événements moléculaires. Tout au long de son cycle cellulaire, le parasite présente des formes et des tailles différentes, qui reflètent des fonctions très différentes (Figure 5- 7).

#### **IV.2.3 - Description cellulaire et moléculaire des stades sanguins de *P. falciparum*.**

*P. falciparum* se développe dans le globule rouge en 48 heures. Au cours de son cycle, il modifie la cellule érythrocytaire de sorte à en tirer ses nutriments et combattre les défenses de l'hôte, avant de s'échapper et d'envahir un nouveau globule rouge par un processus à multiples étapes. Ces événements sont reflétés par la morphologie constamment changeante du parasite (Bannister and Mitchell 2003).

Le **mérozoïte** est une cellule adaptée morphologiquement et fonctionnellement à l'invasion de l'érythrocyte et à la formation de la vacuole parasitophore. Comme les autres stades invasifs (sporozoïte et ookinète), il a une courte durée de vie, il doit envahir un érythrocyte en moins de 10 minutes après avoir été libéré du schizonte. Brièvement extracellulaire, il est exposé au système immunitaire de l'hôte, c'est donc un stade important

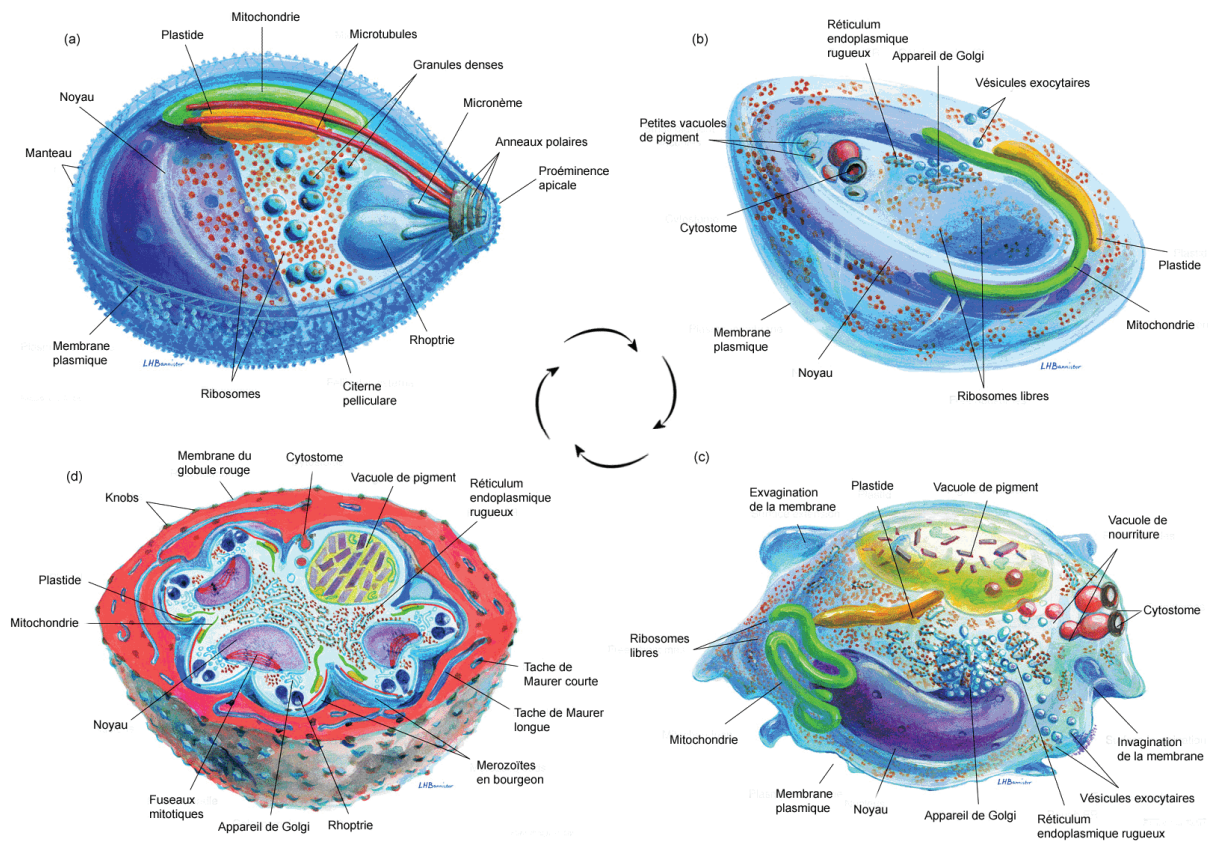
en terme de recherche immunologique. La surface du mérozoïte est recouverte d'un manteau poilu, qui lui permet d'adhérer aléatoirement puis de s'attacher au globule rouge. L'invasion est initiée par une interaction avec la cellule hôte par l'intermédiaire des protéines de surface du mérozoïte comme les MSP (Chitnis and Blackman 2000). Au pôle apical, se situe la machinerie nécessaire à l'invasion (d'où son appartenance aux Apicomplexa) c'est-à-dire les rhoptries, les micronèmes et les granules denses. Si la proéminence apicale touche l'érythrocyte, le mérozoïte se réoriente très rapidement et une jonction irréversible se forme impliquant l'intervention de protéines micronémales telles que AMA1 (Narum and Thomas 1994). Les 3 organelles apicales déversent alors leur contenu séquentiellement lors de l'invasion et entraînent alors un changement de forme et de composition de la membranaire érythrocytaire ainsi que la formation de la vacuole parasitophore (Chitnis and Blackman 2000). Au pôle basal, le noyau se place à l'opposé de l'apex et est accompagné de la mitochondrie et du plastide qui se colocalisent avec un cytosquelette constitué de 2 à 3 microtubules adossés à 3 anneaux apicaux conférant au mérozoïte sa forme ovoïde. Entre les rhoptries et le noyau, une mine de ribosomes libres constitue une ressource pour la synthèse protéique accélérée au stade anneau.

**0 à 18 h post-invasion** : Après l'invasion de l'érythrocyte, le parasite s'aplatit en un fin disque concave, dense en périphérie où se localise le noyau qui s'allonge, conférant au parasite une forme **d'anneau** visualisé en coloration au Giemsa (Figure 6 et 7). La vacuole parasitophore contenant le parasite s'ajuste alors parfaitement à la physionomie du globule rouge et le parasite se nourrit alors des constituants érythrocytaires par le cytostome. Simultanément, il digère l'hémoglobine, sa principale source en acides aminés, formant un pigment brun caractéristique et exporte des protéines parasitaires vers l'érythrocyte altérant ainsi sa structure membranaire ; altérations dont il tire moult avantages, dont l'augmentation de la perméabilité membranaire pour la pénétration sélective des nutriments et l'expression de molécules cytoadhérentes déterminante dans la physiopathologie de neuropaludisme. L'anneau évolue vers une forme plus ronde mais aussi irrégulière, le stade trophozoïte.

**18 à 36 h post-invasion** : le **trophozoïte** se distingue de l'anneau plus par sa taille et sa forme que par les changements qui s'y opèrent. Il poursuit activement sa croissance et l'exportation de protéines parasitaires. Lors de cette phase de synthèse protéique accrue, le parasite met en place un réseau de membranes assurant les connexions entre la vacuole

parasitophore et la membrane érythrocytaire. Ces invaginations visibles sur frottis coloré et appelées "taches de Maurer" sont le siège de la maturation des molécules néo-synthétisées et exportées vers le globule rouge. L'agrégation de protéines parasites à la surface de l'érythrocyte produit de petites protubérances appelées "knobs", qui concentrent des protéines interagissant avec des protéines érythrocytaires. Ces complexes présentent en surface du globule rouge une protéine impliquée dans la cytoadhérence du parasite, PfEMP1 (Leech et al. 1984). Cette propriété à cytoadhérer entraîne une séquestration des érythrocytes infectés à l'endothélium des micro-vaisseaux périphériques et dans certains organes comme la rate, le poumon et le cerveau (MacPherson et al. 1985). Ceci explique l'absence de parasite du stade trophozoïte mature au stade schizonte dans la circulation lors d'une infection à *P. falciparum*.

**36 à 48 h post-invasion** : Le **schizonte** croît jusqu'à épuiser quasiment la réserve d'hémoglobine accumulant le pigment parasite compacté alors en cristaux d'hémozoïne dans la vacuole digestive. Le parasite subit une série de divisions nucléaires répétitives et de nombreuses modifications cytoplasmiques ; il synthétise et assemble tous les éléments nécessaires à la formation des 16 à 32 futurs **mérozoïtes** en commençant par les organelles apicaux accompagnée par la prolifération du réticulum endoplasmique et des ribosomes et la multiplication des mitochondries et des plastides. Au cours de la dernière division nucléaire, les corpuscules ovoïdes s'individualisent régulièrement espacés en périphérie du parasite. Les mérozoïtes naissants, les membranes de l'érythrocyte et de la vacuole parasitophore sont protéolysées et un nouveau cycle érythrocytaire est amorcé.



**Figure 9. Schémas de l'organisation tridimensionnelle de la cellule parasitaire aux différents stades du cycle érythrocytaire.**

(a) schéma d'un mérozoïte avec une partie de la membrane plasmique retirée pour montrer la structure interne, (b) organisation d'un jeune anneau coupé en deux, (c) vue globale d'un trophozoïte, (d) coupe transversale d'un globule rouge infecté par un schizonte. Les mérozoïtes bourgeonnent à la surface du schizonte (Bannister et al. 2000).

Pour parvenir à sa transmission, une sous-population de stades sanguins asexués du parasite va se retirer du cycle cellulaire et se développer en **gamétocytes**, cellules spécialisées pour la reproduction sexuée et l'invasion du moustique vecteur. Les facteurs qui influencent cette différenciation sexuée du mérozoïte ne sont pas connus ; le pourcentage de conversion des formes asexuées en gamétocytes n'est pas constant pendant le cycle et semble être sous le contrôle à la fois de facteurs innés et de facteurs environnementaux :

- il est établi que tous les mérozoïtes émergeant d'un même schizonte partagent tous le même destin ou sont "engagés", vers la prolifération asexuée ou la différenciation sexuée (Bruce et al. 1990). De plus, dans un même schizonte, chaque mérozoïte devient gamétocyte soit mâle soit femelle (Silvestrini et al. 2000).

- l'induction vers une voie de différenciation en gamétocytes est consécutive à un ou plusieurs stimuli tels que, pression immunitaire de l'hôte, pression médicamenteuse, hormones circulantes, effet d'une co-infection, contrôle autocrine de croissance, âge et état intracellulaire des érythrocytes (Dyer and Day 2000). De plus, les facteurs qui influencent l'engagement vers la gamétocytogénèse semblent stimuler directement les voies de transduction du signal du parasite.

Environ 24 h après l'invasion de l'érythrocyte par le mérozoïte "engagé", le développement du gamétocyte commence, le parasite prend une forme de cigare, maintenu par un cytosquelette de microtubules et la membrane étirée du globule rouge. La maturation dure 10 jours et consiste en une séquence ordonnée de 5 stades morphologiquement distincts accompagnée d'une expression séquentielle de protéines spécifiques, dont la protéine cytoplasmique Pfg27 et la protéine membranaire de la vacuole parasitophore Pfs16. L'abolition de l'expression des gènes *pfg27* et *pfs16* conduit à une incapacité totale de produire des gamétocytes (Lobo et al. 1999) ou à une production amoindrie (Kongkasuriyachai et al. 2004). Les microtubules disparaissant, le parasite se relâche et prend alors sa forme caractéristique en croissant. Le gamétocyte mâle est presque dépourvu de ribosomes et possède un matériel nucléaire diffus tandis que le gamétocyte femelle, dont le noyau est condensé, est riches en ribosomes et RE/Golgi, prêts pour la réinitiation de la traduction lors de la gamétogénèse ; ils apparaissent respectivement rose et bleu après coloration au Giemsa. Les gamétocytes matures, dont le métabolisme est réduit, sont des formes bloquées en phase G0 du cycle cellulaire.

*P. falciparum* est un organisme unicellulaire très petit, haploïde mais néanmoins un eucaryote compliqué, capable constamment de varier son expression génétique pour produire des formes phénotypiques différentes. Les caractéristiques structurales des différents stades de développement du *Plasmodium* constamment changeantes au cours de son cycle de vie sont autant d'indices révélateurs de la myriade de processus moléculaires subits au sein de l'organisme, reflétés par sa morphologie.

L'étude sur la biologie du parasite n'a réellement été développée que lorsque la culture *in vitro* des stades érythrocytaires a pu être établie en 1976 par Trager et Jensen (Trager and Jensen 1976). Il n'en reste pas moins que seule l'espèce *P. falciparum* peut être cultivée et que l'obtention des stades hépatocytaires, la production de sporozoïtes demeure difficile et laborieuse.

L'état des connaissances sur la biologie du parasite et la régulation de son cycle cellulaire reste très limité et le séquençage du génome récemment achevé constitue une avancée majeure pour l'étude et la lutte contre le parasite.

Au sein de ses deux hôtes, *P. falciparum* suit un programme de développement complexe et finement régulé pour générer une séquence de formes morphologiquement distinctes et exploiter les environnements très différents dans lesquels il réside. Ce cycle compliqué nécessite de manière évidente une régulation de l'expression des gènes coordonnée.

Les caractéristiques du parasite, tel que le biais des codons par la richesse en A+T, les difficultés de clonage, les conditions laborieuses de culture intensive et la purification de stades sanguins spécifiques contribuent au défi d'élucider divers aspects de la biologie du parasite. En dépit de cela, quelques études ont révélé un parallèle certain avec l'expression génétique chez les eucaryotes.



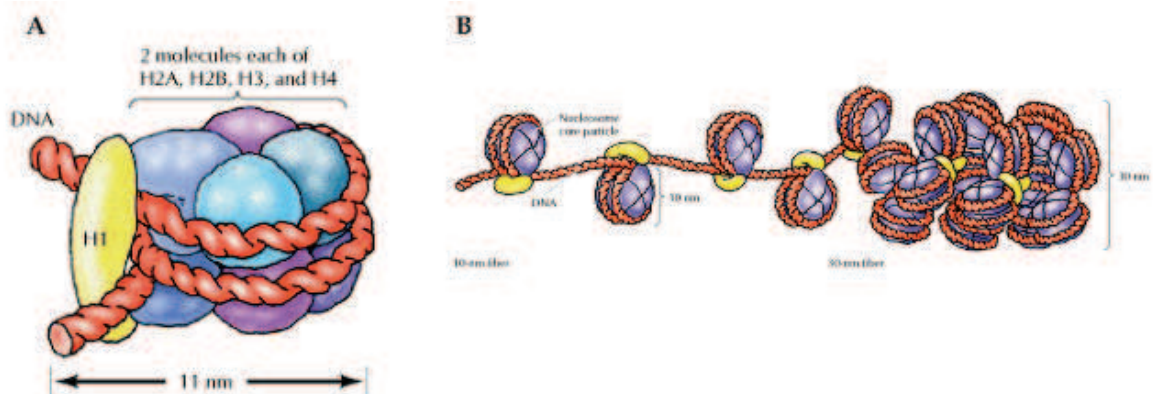
# REGULATION DE L'EXPRESSION DES GENES

## I – Régulation des gènes chez les eucaryotes

L'expression des gènes est régulée par de nombreux contrôles de la cellule intervenant au niveau de la transcription, la maturation, la stabilité et le transport des ARNm, la traduction ainsi que par des modifications post-traductionnelles des protéines codées par les messagers. Dans ce travail, je me suis particulièrement intéressée à la régulation transcriptionnelle.

### I.1 – L'organisation de la chromatine

La compression de l'ADN en chromosomes est indispensable à l'architecture de la cellule et à la régulation transcriptionnelle des gènes, l'ADN est alors associé à des protéines pour former la chromatine. La chromatine est l'organisation structurale basale du matériel génétique chez les eucaryotes. L'unité fondamentale de la chromatine est le nucléosome, qui est constitué par un cœur protéique formé par les histones H2A, H2B, H3 et H4 organisés en octamère sur lequel s'enroule l'ADN à peu près deux fois. Ce module de base se répète régulièrement pour former un nucléofilament qui se condense lui-même en répétitions de nucléosomes liés par un ADN nu et stabilisés par un cinquième type d'histone, l'histone H1. Cette forme se spiralise pour donner les nucléofilaments qui s'enroulent à nouveau pour constituer les fibres chromatiennes organisées en chromosomes



**Figure 8. Nucléosome et chromatine.**

A : le nucléosome avec l'ADN en rouge et les histones H2A en bleu, H2B en vert, H3 en violet, H4 en bleu ciel et H1 en jaune. B : Compaction de la chromatine (Cooper 2000).

L'organisation de l'ADN à l'intérieur de la chromatine et des chromosomes joue un rôle central dans la régulation des processus cellulaires. La structure de la chromatine n'est pas figée : elle est influencée par de nombreux facteurs qui peuvent la modifier. La chromatine des régions chromosomiques non transcrites ou inactives présente une forme condensée, c'est l'hétérochromatine; tandis que les régions en cours de transcription ou actives sont décondensées, c'est l'euchromatine. La chromatine est localement fluidifiée notamment par des protéines non-histones qui seront décrites dans le paragraphe I.4.1.

## **I.2 – La structure des gènes**

Un gène est défini comme la séquence nucléotidique qui donne naissance à une molécule d'ARN fonctionnel dont l'ARN messager codera pour une protéine. Chez les eucaryotes, les gènes sont monocistroniques et morcelés en régions codantes (exons) et en régions non codantes (introns). Dans le messager, deux régions transcrites mais non traduites encadrent la phase ouverte de lecture : la région 5'UTR, site d'ancrage des ribosomes pour la traduction et la région 3'UTR contenant un signal de polyadénylation permettant à l'enzyme poly-A polymérase d'ajouter une queue poly-A lors de la maturation du transcrit.

## **I.3 – Les promoteurs et leurs éléments de régulation**

Juste en amont du gène se situe le promoteur basal, séquence d'ADN minimale suffisante pour induire une transcription. Cette région interagit avec les composants de la machinerie basale de transcription. Le promoteur basal contient différentes séquences importantes (Smale and Kadonaga 2003), dont le site d'initiation de la transcription en aval d'une séquence appelée "boîte TATA" situé à -25/-30 chez les eucaryotes supérieurs et de -40/-120 chez la levure (Struhl 1995). La boîte TATA est le site de fixation de la "TATA binding protein" (TBP), une des protéines composant la machinerie basale de transcription (Nakajima et al. 1988). Dans certains gènes, le site d'initiation de la transcription fait partie d'un élément Initiateur (Inr) et le promoteur basal peut alors contenir soit une boîte TATA et un Initiateur, soit un seul de ces éléments, soit aucun. Les promoteurs n'ayant aucun de ces éléments présentent de multiples sites d'initiation de la transcription donnant naissance à des ARNm à extrémité 5' variable. Juste en amont de la boîte TATA peut se situer la séquence BRE (TFIIB-recognition element), reconnue par le facteur TFIIB et une boîte

CCAAT située à -70, reconnue par les facteurs CTF (CCAAT-binding transcription factor) et CBF (CCAAT-box-binding factor).

D'autres séquences d'ADN, généralement constituées de quelques nucléotides et dispersées tout au long de la région située en 5' du gène, appelées éléments de régulation agissant en *cis* ou "éléments *cis*-régulateurs" prennent part à la régulation transcriptionnelle. Les éléments de régulation sont situés en amont du promoteur basal ; entre -200 et -50 par rapport au site d'initiation, ils appartiennent au promoteur proximal et encore plus en amont ou en aval du point +1, ils appartiennent au promoteur distal. Sur ces éléments *cis*-régulateurs viennent se fixer des facteurs agissant en *trans* ou "facteurs *trans*-régulateurs" ou facteurs de transcription, protéines nucléaires dont l'abondance, la disponibilité et l'accessibilité à leurs séquences cibles vont réguler le niveau de transcription. Les éléments de régulation augmentent ou répriment la transcription d'un gène indépendamment de leur orientation et de leur distance par rapport au site d'initiation de la transcription. Ainsi, situés à des milliers de paires de bases du promoteur basal, ils peuvent agir sur la transcription ou perdre leur pouvoir dès qu'on les éloigne de 15 à 20 bases (McKnight 1982). Chaque gène comprend dans son promoteur plusieurs éléments de régulation contribuant de manière cumulative à la régulation de son expression et ces éléments peuvent être regroupés physiquement pour former des modules de régulation.

#### **I.4 – La transcription**

La transcription des gènes en ARNm est liée à l'activité de l'ARN polymérase II tandis que l'ARN polymérase I sert à la synthèse des ARN ribosomiques (ARNr 5,8S, 18S et 28S) et l'ARN polymérase III transcrit les gènes d'ARN de transfert (ARNt).

Au niveau du promoteur basal, a lieu le recrutement de la machinerie basale de transcription où l'ARN polymérase II s'assemble de concert avec des facteurs d'amorçage pour former le très volumineux complexe de pré-initiation (Cosma 2002; Ptashne 2002). Ces facteurs généraux de la transcription interviennent dans un ordre bien établi (Buratowski et al. 1989; Conaway and Conaway 1993; Zawel and Reinberg 1993).

Le premier facteur à rejoindre le promoteur basal est le facteur général TFIID (Davison et al. 1983; Fire et al. 1984), constitué de la TBP associée à des co-facteurs nommés TAF (TBP-associated factors) (Pugh and Tjian 1991; Tanese et al. 1991) dont certains se lient à

l'Inr (Chalkley and Verrijzer 1999; Oelgeschlager et al. 1996). La TBP, en se fixant à l'ADN dans son petit sillon, impose une courbure considérable à l'ADN (Kim et al. 1993). Le facteur TFIIA se joint au complexe ADN-TFIID en interagissant directement avec la TBP (Geiger et al. 1996; Tan et al. 1996) permettant ainsi le maintien du complexe déjà établi et la poursuite du processus d'assemblage (Buratowski 1994; Buratowski et al. 1989; Weideman et al. 1997). Le facteur TFIIB reconnaît alors la TBP et la distorsion de l'ADN (Nikolov et al. 1995; Reinberg and Roeder 1987; Zheng et al. 1987) et se fixe sur la séquence BRE avec une grande affinité. TFIIB est indispensable pour le recrutement ensuite de TFIIF-ARN polymérase II (Buratowski et al. 1989; Flores et al. 1991; Sopta et al. 1985). Cependant, la polymérase n'initiera la transcription qu'après fixation des facteurs TFIIE (Reinberg and Roeder 1987) puis TFIIH (Coin and Egly 1998; Takagi et al. 2003). Un dernier complexe protéique appelé médiateur vient compléter le recrutement de la machinerie basale de transcription qui pourra seulement alors répondre aux facteurs de transcription (Lewis and Reinberg 2003). La phosphorylation du domaine C-terminal de l'ARN polymérase II par TFIIH libère l'énergie nécessaire à la séparation des deux brins d'ADN, le complexe est dit "ouvert" et la transcription peut commencer (Wang et al. 1992). La polymérase interrompt ses contacts avec le promoteur basal après avoir synthétisé environ 30 bases d'ARNm et le reste de la machinerie de transcription poursuit par l'étape d'élongation. Les facteurs impliqués dans la maturation et l'export des ARNm ainsi que dans la modification de la chromatine sont alors recrutés par l'ARN polymérase (Bentley 2002). Enfin, la terminaison de la transcription, processus compliqué et encore mal connu, permet la libération du transcrit primaire du site de transcription ainsi que la libération de l'ARN polymérase II.

En cours de transcription, l'ARNm subit plusieurs étapes de maturation. A l'extrémité 5', vient s'ajouter une coiffe méthylée, censée protéger les ARNm des dégradations enzymatiques et à l'extrémité 3', le messenger est sectionné une vingtaine de bases après le signal de polyadénylation et allongé d'une queue poly-A (100-200 nt) par une enzyme poly-A polymérase. Les introns sont ensuite épissés et l'ARNm mature est alors transporté dans le cytoplasme où il pourra être traduit.

## **I.4 - La régulation de la transcription**

La régulation de la transcription des gènes eucaryotes intervient à différents niveaux : le remodelage de la chromatine, les modifications des histones, le recrutement de la machinerie basale de transcription et l'action concertée des facteurs de transcription au niveau du promoteur.

### **I.4.1 – Influence de la chromatine**

La chromatine, de par sa structure condensée, représente une barrière à l'accessibilité de l'ADN. Elle est en perpétuel équilibre entre état relâché et état condensé suivant les effets antagonistes de multiples complexes protéiques et il ne fait guère de doute que l'état condensé est défavorable à la transcription. Le remodelage de la chromatine est donc une étape cruciale dans la régulation de la transcription (Workman and Kingston 1998). Parmi les complexes de remodelage de la chromatine, on distingue ceux qui modifient les histones, ceux qui influent sur la méthylation de l'ADN et enfin ceux qui utilisent l'énergie de l'ATP pour modifier la structure du nucléosome (Kornberg and Lorch 1999; Taddei 2000).

#### ***a) Modifications des histones ou code des histones***

La plupart des modifications des histones ont lieu au niveau de l'extrémité N-terminale et sont généralement l'acétylation et/ou la méthylation de certains acides aminés mais peuvent aussi intervenir la phosphorylation, et l'ubiquitination (He and Lehming 2003; Lee and Young 2000). Plusieurs lysines sur la queue N-terminale de chaque histone, dont la position est définie et importante, peuvent être acétylées par des histones acétyltransférases (HAT) et désacétylées par des histones désacétylases (HDAC) (Legube and Trouche 2003). Les modifications des histones affectent directement la structure de la chromatine, en la condensant ou en la relâchant, contrôlant ainsi son accès à d'autres protéines. Les histones hyperacétylées sont associées aux régions transcriptionnellement actives et à une structure chromatinienne plus accessible, alors que les histones hypoacétylées se trouvent préférentiellement dans les régions transcriptionnellement silencieuses (Hebbes et al. 1994; Hebbes et al. 1988). L'acétylation perturbe la structure du nucléosome en neutralisant les charges positives des lysines, diminuant ainsi son affinité pour l'ADN chargé négativement (Grunstein 1997; Luger and Richmond 1998). Elle influence aussi l'interaction avec des facteurs de transcription spécifique (Workman and Kingston 1998). Les histones H2B, H3 et

H4 peuvent être méthylées et la méthylation peut être associée avec les formes acétylées de H3 et H4, suggérant alors que méthylation, acétylation et transcription sont corrélées. La phosphorylation des histones H1 et H3 est quant à elle, impliquée dans la condensation des chromosomes lors de la mitose et celle de H3 est liée à une activité transcriptionnelle augmentée (Koshland and Strunnikov 1996). Les histones H2A, H2B et H3 peuvent subir une ubiquitination et cette modification est associée à un ADN transcriptionnellement actif. Le rôle de l'ubiquitination et de la méthylation est encore mal compris. L'ensemble de ces modifications interviennent dans la formation d'une surface d'interaction sur la chromatine pour d'autres protéines responsables de la traduction du "code des histones" en un pattern d'expression transcriptionnelle (Strahl and Allis 2000).

#### ***b) Méthylation de l'ADN***

Chez les eucaryotes, seules les cytosines précédant une guanine peuvent être méthylées. Les îlots CpG sont définis comme étant des régions de plus de 200 paires de bases, dont le pourcentage en G+C est supérieur à 50% (Gardiner-Garden and Frommer 1987). La méthylation des cytosines au niveau des îlots CpG par des ADN méthyltransférases empêche la transcription du gène situé en aval de ces îlots. Des protéines, se liant à l'ADN méthylé, présentent une activité HDAC, suggérant ainsi qu'elles peuvent convertir la chromatine en état inactif au niveau du site d'initiation de la transcription et donc assurer la quiescence des gènes (Nan et al. 1998). La méthylation permettrait par exemple d'inactiver certains gènes lors de la différenciation (Jones 1999).

#### ***c) Remodelage dépendant de l'ATP***

Le remodelage conformationnel des nucléosomes implique la cassure et la reformation des contacts entre les histones et l'ADN sans pour autant modifier la structure des nucléosomes (Emerson 2002). Les complexes de remodelage identifiés chez les eucaryotes, tels que SWI/SNF (Workman and Kingston 1998), RSC de la levure et NURF, CHRAC et ACF de la drosophile contiennent tous une sous-unité ATPase dont l'activité va permettre de libérer l'énergie nécessaire au déplacement des nucléosomes sur l'ADN. La nouvelle position des nucléosomes peut être plus permissive (Schnitzler et al. 1998) ou au contraire inhiber l'accès des facteurs de transcription à leurs sites de liaison (Becker and Horz 2002; Sudarsanam et al. 2000).

#### ***d) Remodelage architectural : les facteurs de la famille HMGB***

Une autre famille de protéines intervient dans le remodelage de la chromatine. Ces protéines participent au contrôle de l'équilibre entre l'état condensé, structuré de l'ADN et une configuration plus relâchée et accessible aux facteurs de transcription à leurs séquences cibles : ce sont les protéines HMG (High Mobility Group). Très abondantes dans les cellules eucaryotes, les protéines HMG sont aussi appelées les protéines chromosomiques non-histone, et sont associées à une chromatine active (Nemeth and Langst 2004). Par opposition à l'histone de liaison H1, elles participent à la mobilité du nucléosome. Elles ont été initialement définies par leurs propriétés physiques, telles que leur vitesse de migration sur gel d'acrylamide, leur extraction possible de la chromatine en milieu acide et leur solubilité (Baxevanis and Landsman 1995). La nomenclature des protéines HMG nucléaires a récemment été révisée pour définir trois superfamilles classées par leur motif fonctionnel de liaison à l'ADN (Bustin 2001):

- les protéines HMGA, caractérisées par le motif "AT-hook" de neuf acides aminés qui se lient aux régions riches en A+T (anciennement HMG-I, HMG-Y et HMG-C).

- les protéines HMGB, caractérisées par un domaine "HMG-box" de 80 acides aminés qui lie l'ADN dans le petit sillon (anciennement HMG-1 et HMG-2),

- les protéines HMGN, caractérisées par un domaine "nucleosomal-binding domain" qui se lie entre les spirales de l'ADN et l'octamère d'histones (anciennement HMG-14 et 17),

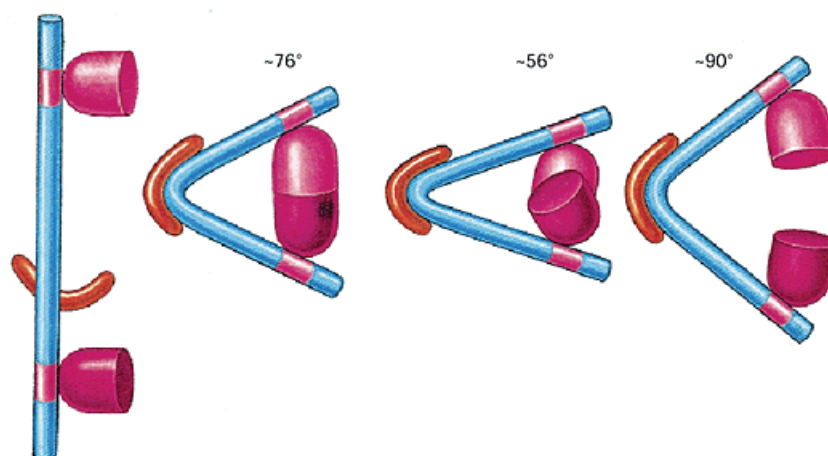
Je vais particulièrement décrire les facteurs HMGB car nous en avons annoté plusieurs chez *P. falciparum* que nous étudions actuellement. La famille HMGB est très ancienne et très conservée au cours de l'évolution. On trouve des membres chez les animaux, les plantes, les levures et les eucaryotes unicellulaires comme *Trypanosoma* (Erondu and Donelson 1992), *Schistosoma* (Gnanasekar et al. 2006) et *Plasmodium* (Kun and Anders 1995).

HMGB, facteur nucléaire. Dans la superfamille HMGB, il existe :

- des facteurs de transcription classiques qui se lient à l'ADN en reconnaissant une séquence linéaire spécifique, tels que les facteurs TCF1 et LEF1 des cellules T, les protéines de typage sexuel MAT $\alpha$  et MAT $\beta$  des levures, le facteur de déterminisme du sexe SRY chez les mammifères et les nombreuses protéines SOX,

- des protéines HMGB qui se lient à l'ADN sans spécificité de séquence mais avec une grande affinité pour une variété de structures particulières de l'ADN autre que la double hélice comme l'ADN cruciforme (Bianchi et al. 1989), l'ADN modifié par le *cis*-platine (Billings et al. 1992), les jonctions entre ADN de forme B et de forme Z (Hamada and Bustin 1985) et les boucles d'ADN maintenues par un hémicaténane (Gaillard and Strauss 2000; Jaouen et al. 2005).

Principalement, les facteurs HMGB ont la capacité de courber l'ADN hélicoïdal générant ainsi une conformation qui facilite l'assemblage des complexes nucléoprotéiques, c'est pourquoi ils sont appelés "facteurs architecturaux" (Bustin 1999; Bustin and Reeves 1996; Thomas and Travers 2001). Ils participent à de multiples activités cellulaires liées à l'ADN telles que transcription, réplication, recombinaison, réparation....



**Figure 9. Assemblage de complexes nucléoprotéiques par l'intermédiaire de la protéine HMGB.**

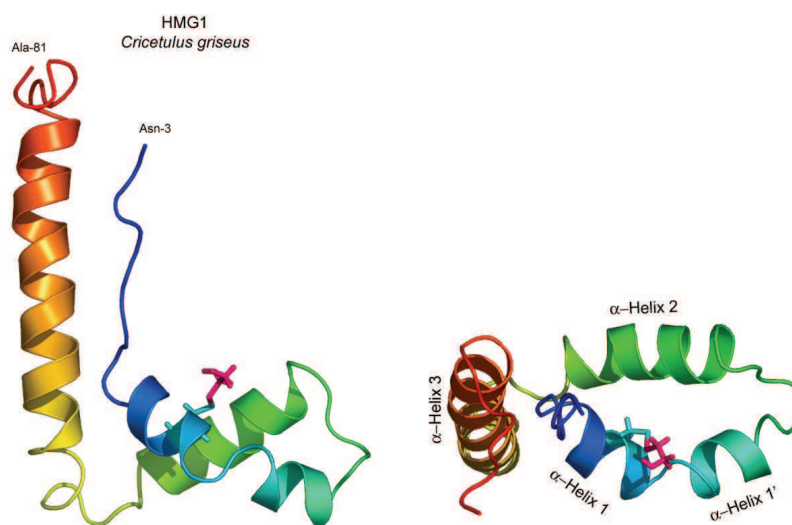
La courbure de l'ADN induite par la protéine HMGB (visualisée en orange) permet l'assemblage de complexes nucléoprotéiques fonctionnels (Wolffe 1999).

Le nombre de domaines "HMG-box" dans les protéines HMG est très variable et n'est pas corrélé à l'appartenance aux facteurs de transcription, si la protéine contient un domaine, ou aux facteurs architecturaux, si la protéine contient plusieurs domaines comme il avait été précédemment proposé (Grosschedl et al. 1994). En effet, les facteurs de transcription UBF de souris et de xénope possèdent cinq domaines tandis que les protéines HMG-D et HMG-Z de la drosophile et les protéines NHP6A et NHP6B de la levure *S. cerevisiae* possèdent un seul domaine. Cependant, la majorité des protéines HMGB, particulièrement chez les vertébrés, possèdent deux domaines "HMG-box" en tandem appelées boîte A et B. Les plantes, les



champignons et les eucaryotes unicellulaires possèdent rarement des protéines HMGB avec des boîtes en tandem (Bianchi and Agresti 2005). En fait, les facteurs de transcription et les facteurs architecturaux peuvent être clairement séparés sur un arbre phylogénétique (Soullier et al. 1999) et leur séparation a lieu avant la divergence entre les levures et le règne animal. De plus, trois déterminants biochimiques caractéristiques des deux groupes de facteurs ont été mis en évidence (Murphy et al. 1999) : la présence dans la séquence protéique aux positions 10 et 32 selon la numérotation de la protéine HMG-D, d'une asparagine et d'un acide aminé hydrophile, respectivement, chez les facteurs de transcription tandis que les facteurs architecturaux présentent une sérine et un acide aminé hydrophobe et la présence d'un cœur hydrophobe.

Le domaine "HMG-box" est très conservé au cours de l'évolution et est composé d'environ 80 acides aminés, repliés en 3 hélices  $\alpha$  agencées en forme de L (Baxevanis and Landsman 1995; Weir et al. 1993).

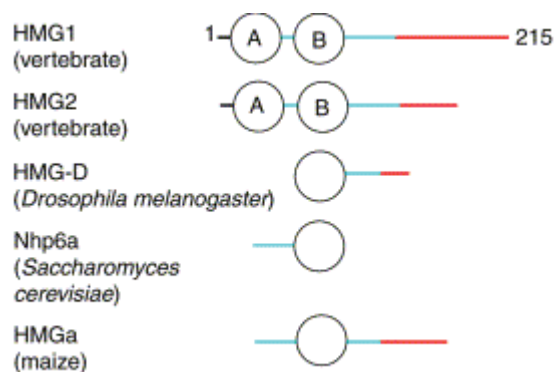


**Figure 10. Structure du domaine "HMG box" du facteur HMG1 de *Cricetulus griseus*, le hamster chinois.**

Obtenue par spectroscopie RMN (Read et al. 1993), la structure du domaine "HMG box" présente 3 hélices  $\alpha$  qui se replient en forme de L, visualisées sous deux angles différents, colorée selon le spectre de la lumière, le bleu désignant la partie N-terminale et le rouge la partie C-terminale.

Chez les vertébrés et chez les plantes, les protéines HMGB présentent un domaine basique aux extrémités N- et C-terminale du domaine "HMG- box" (boîte B chez les vertébrés) ainsi qu'une queue acide en C-terminale de la protéine (Thomas and Travers 2001). Néanmoins, chez les eucaryotes inférieurs, l'extension basique (*D. melanogaster*) ou la

queue acide (*S. cerevisiae*) peuvent manquer. L'extension basique semble jouer un rôle dans l'activité de liaison et de courbure de la protéine (Dai et al. 2005; Teo et al. 1995) conférant à la boîte A la fonction principale d'interaction à l'ADN tandis que la boîte B pourrait être impliquée dans les interactions protéine-protéine, par exemple avec les facteurs de transcription (Webb and Thomas 1999). Le rôle de la queue acide est plus vague, elle pourrait interagir avec les charges positives des histones (Thomsen et al. 2004).



**Figure 11. Représentation schématique de l'alignement multiple des séquences de facteurs architecturaux.**

Les cercles représentent les domaines "HMG box" de chaque protéine, les extensions basiques sont en bleu et les queues acides sont en rouge. Adapté de (Thomas and Travers 2001).

Les protéines HMGB participent ainsi aux interactions physiques entre l'ADN et de multiples facteurs, tels que p53, NF- $\kappa$ B, certaines protéines à homéodomains, les protéines RAG1/2 et des récepteurs aux hormones stéroïdes (Bianchi 2004). Les fonctions nucléaires de la protéine HMGB1 sont essentielles à la vie puisque des souris déficientes pour le gène *hmgb1* meurent rapidement après la naissance (Calogero et al. 1999).

Par leurs interactions continues et extrêmement dynamiques avec les nucléosomes, les protéines HMGB confèrent une flexibilité à la structure de la chromatine et surtout, elles contribuent finement au contrôle de la transcription des gènes en réponse à de rapides changements environnementaux.

HMGB, médiateur extracellulaire. Le rôle nucléaire des facteurs HMGB dans le remodelage de la chromatine n'exclut pas d'autres fonctions cellulaires puisqu'il a été décrit

récemment une facette inattendue de ces protéines en tant que médiateur extracellulaire. En 1999, Tracey et ses collaborateurs rapportent que les protéines HMGB sont sécrétées au cours de l'activation des monocytes et macrophages et agissent en fait comme des cytokines inflammatoires (Wang et al. 1999). Les modifications par acétylation de multiples lysines délocaliseraient les protéines HMGB du noyau vers le cytoplasme pour la sécrétion (Bonaldi et al. 2003). Alors que les macrophages secrètent activement les protéines HMGB hyperacétylées, les cellules en nécrose les libèrent passivement et les cellules en apoptose quant à elles, retiennent HMGB irréversiblement liées à la chromatine condensée. La libération de HMGB conduit au recrutement des cellules de l'inflammation et déclenche la production de cytokines TNF $\alpha$  et IL-1 par les monocytes (Andersson et al. 2000). L'activité cytokinique de HMGB est liée à un domaine de 20 acides aminés présent dans la boîte B de la protéine HMG1 humaine (Li et al. 2003). Les protéines HMGB agissent donc comme médiateurs dans la réponse immunitaire et comme facteurs d'activation des monocytes/macrophages (Dumitriu et al. 2005; Muller et al. 2001) auxquels elles se lieraient au récepteur RAGE (receptor for advanced glycation end products) ainsi qu'à TLR2 et TLR4 (Toll-like receptor) (Kokkola et al. 2005; Park et al. 2004). De plus, les protéines HMGB extracellulaires sont un signal lors de la lésion tissulaire, elles induisent la migration et la prolifération des mésoangioblastes et participeraient ainsi à la régénération des tissus musculaires (Palumbo et al. 2004). La surexpression de HMGB dans certains cancers et dans les états septiques montrent que plus largement, HMGB est impliqué dans les mécanismes de l'inflammation, de développement des tumeurs et des métastases (Lotze and Tracey 2005; Taguchi et al. 2000; Wang et al. 1999).

#### I.4.2 – Les facteurs de transcription

Aux diverses séquences cis-régulatrices décrites précédemment viennent se fixer les facteurs protéiques agissant en *trans* ou facteurs de transcription. La régulation de l'expression des gènes *via* la liaison de facteurs de transcription est le mécanisme principal qui contrôle les processus complexes de développement et de différenciation. Les facteurs de transcription reconnaissent leurs sites spécifiques grâce à de petits domaines de liaison à l'ADN et coopèrent avec d'autres protéines pour stimuler ou réprimer la transcription par l'intermédiaire d'un domaine effecteur (Lemon and Tjian 2000). Les domaines de liaison se replient de façon à présenter une protubérance ou une structure flexible qui entrera en contact avec l'ADN. Ces contacts se font dans le grand sillon de l'ADN, souvent par l'intermédiaire d'une hélice  $\alpha$ , avec des liaisons hydrogène et des interactions de Van der Waals.

Différents groupes de facteurs de transcription ont été identifiés sur la base de la séquence de leurs domaines caractéristiques ainsi que de leur structure tridimensionnelle et leur façon d'interagir avec l'ADN (Harrison 1991; Pabo and Sauer 1992; Tan and Richmond 1998; Warren 2002). Ils ont été classés selon la base de données dédiée aux facteurs de transcription eucaryotes et à leurs éléments de régulation : TRANSFAC (Knuppel et al. 1994; Matys et al. 2003; Wingender et al. 2001).

##### *a) Protéines à domaines basiques*

Ces protéines possèdent un domaine de liaison à l'ADN basique, long de 15 à 30 acides aminés, leur permettant de se dimériser. Dans cette catégorie, se trouvent les protéines portant une agrafe à leucines (leucine zipper ou bZIP), motif largement partagé des levures à l'homme et caractérisé par une répétition de sept leucines (Landschulz et al. 1988). Grâce à leur motif, les protéines à domaine bZIP forment des homodimères ou des hétérodimères et la nature du dimère influe sur l'activité biologique du facteur de transcription. Un autre groupe de protéines dont le domaine basique est suivi d'une région permettant la formation de dimères est constitué des protéines à motif hélice-boucle-hélice (bHLH), motif formé par une hélice  $\alpha$  amphipatique, d'une boucle puis d'une autre hélice  $\alpha$  amphipatique (Murre et al. 1989; Voronova and Baltimore 1990).

Tout comme les protéines à domaine bZIP, les protéines à motif bHLH ont d'importants rôles dans la différenciation et le développement. Leur activité est modulée par la formation d'hétérodimères constitués d'activateurs, de protéines ubiquitaires et de répresseurs (Barinaga 1991).

#### ***b) Protéines à motif en doigt de zinc***

Les motifs en doigt de zinc forment de petits domaines, repliés indépendamment du reste de la structure protéique autour d'un ion zinc qui stabilise la formation d'un feuillet  $\beta$  contre une hélice  $\alpha$  (Tan and Richmond 1998). La liaison à l'ADN de ces motifs passe par l'interaction entre trois acides aminés de l'hélice  $\alpha$  et trois bases de l'ADN. Le domaine en doigt de zinc  $C_2H_2$ , formé de 30 acides aminés repliés sur un ion zinc central chélaté par deux cystéines et deux histidines, est le plus répandu chez les eucaryotes (Klug 1999). Les protéines comportant un motif en doigt de zinc sont impliqués dans la différenciation et la croissance, les proto-oncogènes, les facteurs généraux de la transcription (Berg 1986; Berg 1990; Klug and Rhodes 1987).

#### ***c) Protéines à motif hélice-tour-hélice : les facteurs de la famille Myb***

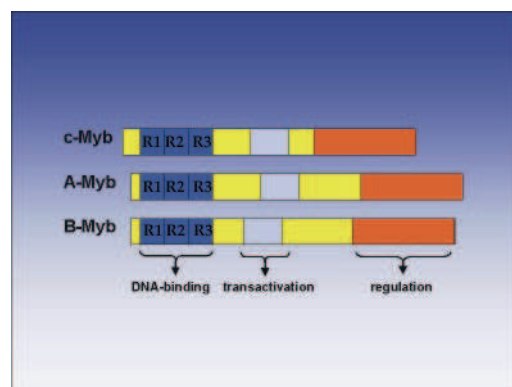
Le motif hélice-tour-hélice (HTH) est un des domaines de liaison à l'ADN le plus répandu au cours de l'évolution, des procaryotes aux eucaryotes (Harrison and Aggarwal 1990). Il est généralement composé de 20 résidus formant deux hélices  $\alpha$  connectées par un tour de 4 résidus (Kohn et al. 1997), la seconde hélice étant celle qui permet la reconnaissance de la séquence d'ADN cible dans le grand sillon (Harrison 1991). Le domaine HTH peut présenter des conformations différentes suivant le nombre de résidus insérés dans le tour. Des résidus hydrophobes conservés dans les hélices ainsi qu'une glycine dans le tour participent à la stabilisation de la structure HTH (Pabo and Sauer 1992). Les protéines à motifs HTH peuvent être classées en plusieurs groupes structuraux dont les protéines à hélices ailées (Gajiwala and Burley 2000) et les protéines présentant un homéodomaine (Gehring et al. 1990; Scott et al. 1989).

Les facteurs de transcription de la famille Myb, qui a fait l'objet d'une étude chez *P. falciparum* lors de la préparation de ma thèse, appartiennent à la classe des protéines HTH. La famille Myb est caractérisée par la présence d'un domaine de liaison à l'ADN appelé le "domaine Myb", présent le plus souvent dans la partie N-terminale de la protéine (Ralston

and Bishop 1983) et qui contient généralement deux ou trois répétitions imparfaites (R1, R2 et R3) d'une cinquantaine d'acides aminés suivant le schéma : trois tryptophanes conservés et régulièrement espacés de 18 ou 19 acides aminés (Peters et al. 1987). Le domaine Myb a été identifié initialement à partir de la protéine virale v-Myb, impliquée dans la transformation cancéreuse des cellules hématopoïétiques (Introna et al. 1994). Depuis l'identification de la protéine c-Myb chez le poulet (Gonda et al. 1982; Klempnauer et al. 1982), de nombreuses protéines Myb eucaryotes ont été découvertes et montrent une grande conservation au cours de l'évolution du domaine Myb (Biedenkapp et al. 1988). Alors que, les répétitions R2 et R3 sont suffisantes pour que la protéine Myb se lie dans le grand sillon de l'ADN de manière séquence-spécifique (Luscher and Eisenman 1990; Ness et al. 1989), la répétition R1 ne semble pas avoir d'interaction spécifique avec l'ADN (Ogata et al. 1994) mais elle augmenterait l'affinité de la protéine pour sa séquence cible (Tanikawa et al. 1993) et la stabilité du complexe protéine-ADN (Dini and Lipsick 1993). Chacune des répétitions se replie en trois hélices  $\alpha$  autour d'un cœur hydrophobe formé par les trois tryptophanes, les deux dernières hélices formant la structure en HTH (Ogata et al. 1994). Cependant, plusieurs protéines Myb possèdent des répétitions imparfaites où les tryptophanes peuvent être remplacés par des résidus phénylalanine ou tyrosine, de la même manière une insertion de quelques acides aminés dans l'une des répétitions peut être observée (Hovring et al. 1994; Morrow et al. 1993). De plus, un autre résidu très conservé et présent dans les répétitions R1 et R2, a une signification fonctionnelle importante dans le domaine de liaison à l'ADN : il s'agit d'une cystéine, dont l'état oxydé ou réduit dépend la conformation de la protéine et par conséquent son interaction à l'ADN (Pinson et al. 2001). Le domaine de liaison à l'ADN des protéines Myb interagit avec une séquence cible sur l'ADN  $^T/cAAC^T/cG$  appelée Myb regulatory element (MRE) (Friedman 1996). La spécificité de séquence semble être relativement bien conservée : la protéine Myb de *Dictyostelium* est capable de lier le même motif que des protéines Myb de vertébrés ou de la drosophile (Stober-Grasser et al. 1992). De plus, la présence de séquences AT riches en aval de la séquence cible consensus participe à l'efficacité de l'interaction (Ganter et al. 1999). Cette donnée est particulièrement intéressante dans le cadre de l'extrême richesse en AT des séquences promotrices chez *P. falciparum*.

Les protéines Myb sont essentielles à la survie de la cellule, elles sont impliquées dans la régulation du cycle cellulaire et de la différenciation. Elles agissent souvent de concert

avec d'autres protéines se liant à l'ADN, dont les comparaisons de séquences ont permis d'identifier d'autres domaines fonctionnels (voir figure 12). Un domaine de transactivation, d'une cinquantaine de résidus, hydrophobe et légèrement acide, est situé en C-terminal du domaine de liaison à l'ADN (Weston and Bishop 1989) et un domaine de régulation négative, situé en C-terminal du domaine de transactivation, se compose de deux sous-domaines indépendants (1 et 2) séparés par une structure en forme de leucine zipper (LZ) et capables d'inhiber l'activation transcriptionnelle de la protéine (Dubendorff et al. 1992). Alors que le domaine de liaison à l'ADN situé dans la partie N-terminale des protéines D-Myb de la drosophile (Katzen et al. 1985), BAS1 de la levure (Tice-Baldwin et al. 1989) ainsi que A-Myb et B-Myb humaines (Nomura et al. 1988) est le mieux conservé, les domaines de transactivation et de régulation négative ne sont pas retrouvés dans toutes les protéines de type Myb. Il est à noter que le domaine de liaison à l'ADN peut se trouver dans la partie C-terminale des protéines comme c'est le cas dans la protéine REB 1 de *Kluyveromyces lactis* (Morrow et al. 1993).



**Figure 12. Représentation schématique des trois domaines fonctionnels des protéines c-Myb de souris, A-Myb et B-Myb humaines.**

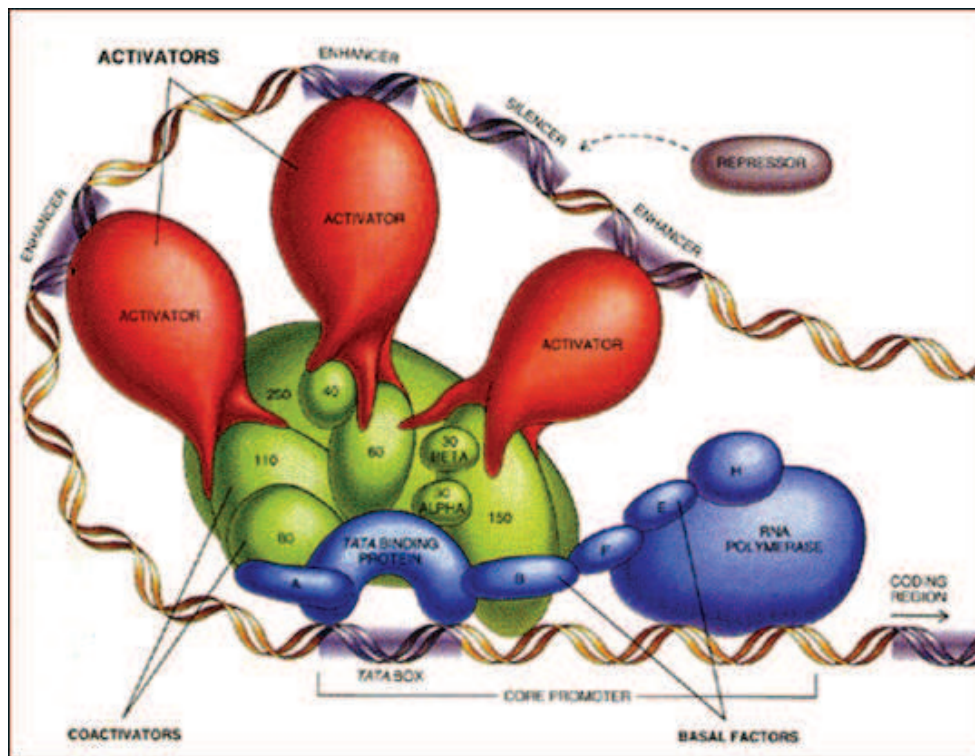
En bleu foncé : le domaine de liaison à l'ADN présentant les 3 répétitions R1, R2 et R3 ; en gris : le domaine de transactivation ; en rouge : le domaine de régulation négative, d'après (Klempnauer).

#### ***d) Protéines à architecture $\beta$***

Il est difficile de trouver une caractéristique commune à tous les facteurs composant cette superclasse tant et si bien que même la caractéristique en architecture  $\beta$  n'est pas partagée par tous. Par exemple, les facteurs de type HMG (présentés dans le paragraphe I.4.1) comprennent seulement des segments en hélices  $\alpha$  mais ressemblent beaucoup aux

membres de cette superclasse par son mode d'interaction à l'ADN qui se traduit par une insertion dans le petit sillon de l'ADN causant une déformation de la double hélice (Gehring et al. 1990).

#### I.4.3 – Schéma de la régulation transcriptionnelle chez les eucaryotes



**Figure 13. Machinerie basale de la transcription, éléments de régulation et facteurs de transcription (Tjian 1995).**

Le recrutement de la machinerie basale de transcription (en bleu) a lieu sur le promoteur basal en association avec les TAF (en vert). L'accessibilité des FT (activateur en rouge ou represson en marron) à leurs séquences cibles (enhancer ou silencer en violet) sur le promoteur distal vont réguler le niveau de transcription. Les boucles formées par l'ADN permettent de rapprocher les facteurs de transcription pour qu'ils puissent former un complexe multiprotéique.

## II– Le génome de *P. falciparum*.

En 1996, un effort concerté de la communauté scientifique internationale a permis de mettre sur pied un projet de séquençage et d'annotation du génome du clone 3D7 de *P. falciparum* (Hoffman et al. 1997). Entre 1998 et 1999, les chromosomes 2 et 3 ont été entièrement séquencés (Bowman et al. 1999; Gardner et al. 1998) et il faudra attendre fin 2002 pour la publication du génome complet, la totalité des séquences sont en libre accès avec néanmoins quelques parties à compléter (Gardner et al. 2002).



## II.1 – Composition et organisation du génome.

Le génome du clone 3D7 de *P. falciparum* est composé de 22.8 mégabases (Mb) réparties en 14 chromosomes dont la taille varie de 0,643 à 3,29 Mb. La composition des nucléotides présente une forte richesse en adénines (A) et en thymines (T), en moyenne de 80.6 % du génome voire plus de 90 % pour les régions intergéniques et les introns et porte *P. falciparum* au premier rang des génomes riches en A+T devant *Dictyostelium discoideum* (Glockner et al. 2002). Cette particularité confère à l'ADN une grande instabilité et rend le génome difficile à manipuler.

**Tableau 2. Données générales sur les génomes de *P. falciparum* et de cinq autres eucaryotes unicellulaires (Abrahamsen et al. 2004).**

Feature	<i>C. parvum</i>	<i>P. falciparum</i>	<i>S. pombe</i>	<i>S. cerevisiae</i>	<i>E. cuniculi</i>
Size (Mbp)	9.1	22.9	12.5	12.5	2.5
(G+C) content (%)	30	19.4	36	38.3	47
No. of genes	3807	5268	4929	5770	1997
Mean gene length (bp) excluding introns	1795	2283	1426	1424	ND
Gene density (bp per gene)	2382	4338	2528	2088	1256
Percent coding	75.3	52.6	57.5	70.5	90
Genes with introns (%)	5	53.9	43	5	ND
Intergenic regions					
(G+C) content %	23.9	13.6	32.4	35.1	45
Mean length (bp)	566	1694	952	515	129
RNAs					
No. of tRNA genes	45	43	174	299	44
No. of 5S rRNA genes	6	3	30	100–200	3
No. of 5.8S, 18S, and 28S rRNA units	5	7	200–400	100–200	22

La dimension des gènes de *P. falciparum* (taille moyenne, 2300 pb sans les introns présents à 54%) souligne une caractéristique importante du génome d'autant qu'aucun élément transposable n'a été identifié dans le génome. Les séquences sont plus largement enrichies en étendues A-T destinées à former des structures non-globulaires. Cette particularité introduit un biais par la présence de successions homopolymériques d'acides aminés (souvent asparagine) faisant apparaître des domaines globulaires à basse entropie ou basse complexité (souvent glutamate et glutamine) qui distingue *Plasmodium* des autres eucaryotes (excepté *D. Discoideum*) (Aravind et al. 2003).

Bien que la taille du génome de *P. falciparum* soit considérablement plus importante que chez les levures *Schizosaccharomyces pombe* et *Saccharomyces cerevisiae*, un nombre

comparable de gènes (5300) a été prédit. A capacité codante équivalente, la différence de taille entre les génomes est reflétée par la taille des protéines parfois jusqu'à 50% plus grande que les protéines orthologues de levure et en contraste avec ce qui est observé chez *Encephalitozoon* (Katinka et al. 2001).

Environ 60 % des protéines prédites, soit 3208 protéines, n'ont pas de fonction assignée par manque d'homologie suffisante avec les protéines d'autres eucaryotes, probablement lié à l'extrême richesse en (A+T). Cette proportion est bien plus importante que pour la plupart des génomes déjà séquencés. Mais, quelques protéines (5%) ont une similitude avec des protéines putatives non annotées d'autres organismes. Un tiers des protéines prédites présentent au moins un domaine transmembranaire et 17.3 % possèdent un signal peptidique putatif.

Le génome code pour 43 ARN de transfert (ARNt) représentant l'ensemble des anticodons. Chez la plupart des eucaryotes séquencés jusqu'à aujourd'hui, les gènes codant les ARNt sont assez redondants, mais le génome de *P. falciparum* ne semble pas comporter de répétitions de gènes codant pour les ARNt. *P. falciparum* partagerait cette organisation avec d'autres protozoaires parasites comme un autre apicomplexe, *Cryptosporidium parvum* et la microsporidie *Encephalitozoon cuniculi* (Tableau 2).

De la même manière que chez *C. parvum* (Abrahamsen et al. 2004) et contrairement à de nombreux eucaryotes, les gènes codant pour les ARN ribosomiaux (ARNr) sont peu nombreux et non répétés en tandem et les unités codant pour les ARNr 5.8 S, 18 S et 28 S sont distribués sur plusieurs chromosomes.

## II.2 – Structure des chromosomes.

*P. falciparum* ressemble aux autres eucaryotes dans l'organisation générale des chromosomes. D'une part, les chromosomes s'organisent autour de centromères composés de petites répétitions en tandem sur une longueur de 2-3 kb et dont le contenu en bases A et T excède 97 %. D'autre part, les régions sub-téломériques contenant les grandes familles multigéniques *var*, *rifin* et *stevor* sont agencées en 5 blocs très conservés et composées de répétitions en tandem, dont l'un d'eux contient des répétitions de 21 pb formant l'élément Rep20. Cet élément est impliqué dans un mécanisme permettant les échanges inter-chromosomiques participant à la variation antigénique chez *P. falciparum* (Figueiredo et al.

2002; Freitas-Junior et al. 2000; O'Donnell et al. 2002). La localisation des familles multigéniques dans les régions sub-téломériques est d'ailleurs partagée par d'autres parasites eucaryotes.

### III – La transcription chez *P. falciparum*.

#### III.1 – Organisation des gènes

Les gènes du parasite sont organisés d'une façon comparable aux gènes eucaryotes ; ils sont monocistroniques, les phases ouvertes de lecture sont flanquées d'une région 5' et 3' non traduite et contiennent des séquences introniques ; les ARNm ont une coiffe en 5' en plus d'être polyadénylés (Horrocks et al. 1998; Lanzer et al. 1993). Enfin, en amont de ces gènes se trouvent des promoteurs dont la structure suit le modèle bi-partite commun aux eucaryotes, comprenant une région promotrice basale capable d'induire un niveau minimal de transcription, contrôlée en amont par des éléments de régulation spécifiques (Crabb and Cowman 1996; Osta et al. 2002).

#### III.2 – Composition en protéines assignées à la régulation génétique

De manière plutôt surprenante compte tenu du cycle complexe de *P. falciparum*, l'analyse du génome révèle un faible pourcentage (1,3 %) de protéines assignées à la régulation de la transcription. Pourtant, il apparaît que la proportion de cette catégorie de protéines augmente avec la complexité des organismes : 4 % pour *S. cerevisiae*, 4,9 % pour *C. elegans*, 6,3 % pour *D. melanogaster* et 6,8 % pour l'homme (Tableau 3).

**Tableau 3. Pourcentage de protéines assignées à la régulation de la transcription pour 8 génomes d'eucaryotes (Coulson et al. 2004).**

Species	Size <sup>a</sup>	Matches	% Genome
<i>P. falciparum</i>	5273	69	1.3
<i>S. cerevisiae</i>	6292	249	4.0
<i>S. pombe</i>	4882	205	4.2
<i>C. elegans</i>	19,099	936	4.9
<i>D. melanogaster</i>	13,707	865	6.3
<i>A. thaliana</i>	27,387	1807	6.6
<i>H. sapiens</i>	29,303	1987	6.8
<i>A. gambiae</i>	15,101	1045	6.9

Size: nombre de protéines dans le génome ; Matches: nombre de protéines assignées à la régulation de la transcription ; % Genome: pourcentage des protéines assignées à la régulation de la transcription par rapport à la totalité des protéines du génome.

Cependant, il faut reconsidérer cette assertion dans l'optique des 60% de phases ouvertes de lecture encore non assignées à une fonction chez *Plasmodium* et dans l'optique de notre travail où nous avons déjà annoté différents facteurs de transcription putatifs avant le séquençage complet du génome. L'importante proportion de séquences orphelines pourrait être expliquée par le fait que les protéines du parasite auraient évolué de telle sorte qu'elles ne peuvent plus être identifiées par de simples recherches de similarité de séquences (Coulson et al. 2004; McConkey et al. 2004). De plus, l'extraordinaire richesse en A+T (80%) dans la composition nucléotidique du parasite, causant un remarquable biais dans l'usage des codons, introduirait d'importants changements dans les séquences d'ADN et d'acides aminés qui affecteraient les procédures de recherche. En effet, les algorithmes d'alignement de séquence qui déterminent la similarité entre deux séquences sont très peu efficaces en dessous de 30% de similarité (Rost 1999). Cela pourrait expliquer, par conséquent, l'apparente pauvreté en facteurs de transcription annoncée dans les analyses du génome et il est légitime de penser qu'un nombre substantiel de cette inhabituelle proportion de protéines non annotées pourrait correspondre à des orthologues "cachés". Par ailleurs, l'analyse des structures secondaires, qui sont souvent plus conservées que les structures primaires permet d'identifier avec des outils appropriés des régions hydrophobes dans les domaines globulaires partagés par différents organismes (Hydrophobic Cluster Analysis). Ce programme d'analyse a conduit ainsi à la prédiction de neuf facteurs de transcription généraux supplémentaires aux dix autres prédits par simples recherches de similarité (Callebaut et al. 2005). L'analyse du contenu protéique du génome par HCA pourrait donner un éclairage nouveau sur la régulation génétique chez *P. falciparum*.

L'ensemble des données que je vais présenter maintenant provient en plus de la littérature, des séquences annotées dans le laboratoire en utilisant l'outil de recherche des domaines de la banque Pfam (<http://www.sanger.ac.uk/Software/Pfam/>) dans PlasmoDB (<http://www.plasmodb.org/>) après vérification avec l'outil MotifScan ([http://myhits.isb-sib.ch/cgi-bin/motif\\_scan](http://myhits.isb-sib.ch/cgi-bin/motif_scan)) ou de recherches grâce à l'outil BLAST de PlasmoDB. L'ensemble des numéros d'accèsion est disponible en annexe I.1.

### III.2.1 – La machinerie générale de la transcription

Le génome de *P. falciparum* code pour l'ensemble des douze sous-unités de l'ARN polymérase II [(Coulson et al. 2004), annexe I.2]. La plus grande sous-unité de l'ARN

polymérase II contient une extension de son domaine C-terminal comme c'est le cas chez un grand nombre d'eucaryotes où elle joue un rôle important dans les processus vitaux tels que l'élongation, l'épissage et la régulation de l'expression des gènes (Carlson 1997; Corden 1990). La présence de cette extension constitue une donnée importante dans la compréhension de la régulation des gènes chez *Plasmodium*.

Le recrutement de la machinerie basale de transcription au niveau du promoteur basal nécessite la présence de plusieurs complexes protéiques (voir paragraphe I.4) dont la TBP et les TAF. Contrairement aux levures, le génome de *P. falciparum* code pour une TBP (Hirtzlin et al. 1994) et une protéine homologue de la TBP (McAndrew et al. 1993); la présence de plusieurs TBP étant une caractéristique des eucaryotes supérieurs. Les séquences de ces deux TBP présentent les caractéristiques fonctionnelles des TBP en terme de liaison à l'ADN et de dimérisation. Certaines protéines composant les complexes TFIIB, TFIIE et TFIIH sont aussi présentes dans le génome.

Bien que les TAF présentent une faible conservation chez les métazoaires (Coulson and Ouzounis 2003), quatre TAF viennent d'être identifiées récemment grâce à la méthode HCA : TAF1, TAF2, TAF7 et TAF10 (Callebaut et al. 2005). De la même manière, la grande sous-unité du facteur TFIIA, deux candidats pour sa petite sous-unité, TFIIE $\beta$  et TFIIF $\beta$  et une sous-unité du facteur TFIIH ont été prédits. Ces données suggèrent qu'il y a plus de facteurs généraux de la transcription dans le génome de *Plasmodium* qu'il avait été annoncé initialement. Enfin, sont présentes également au moins trois protéines, conservées des levures à l'homme, MED7, MED10, MED31 appartenant au complexe Mediator, qui intervient à la phase finale du recrutement de la machinerie de transcription et qui permet de transmettre l'information provenant de la liaison des FT (Bourbon et al. 2004; Linder and Gustafsson 2004).

La découverte de ces protéines suggère que *P. falciparum* possède des représentants de la majorité des facteurs de la machinerie basale, dont la fonction reste à valider, lui permettant d'assurer la synthèse des ARNm ainsi que la régulation transcriptionnelle de ses gènes.

### III.2.2 – Régulateurs de la structure de la chromatine

Il est communément admis que la chromatine est dynamique et influence la régulation génétique (voir paragraphe I.4.1). Chez les parasites apicomplexes, la régulation des gènes doit nécessiter également la participation d'un appareil de remodelage de la chromatine bien développé. Le génome de *Plasmodium* code pour les 4 histones classiques (Longhurst and Holder 1997) composant le nucléosome et il est intéressant de noter que les gènes codant pour H2A et H3 ont un sens de transcription opposé et partagent le même promoteur sur le chromosome 6. Il en est de même pour H2B et H4 sur le chromosome 11. En revanche, l'histone de liaison H1 ne semble pas être présente dans le génome. Même si ces protéines H1 sont impliquées dans la régulation transcriptionnelle de certains gènes, elles ne sont pas responsables de la régulation globale du niveau de transcription (Harvey and Downs 2004). De plus, elles ne semblent pas indispensables à la survie des eucaryotes unicellulaires comme les levures (Ausio 2000). L'apparente absence d'histone de liaison chez *P. falciparum* ne perturberait pas fondamentalement la régulation transcriptionnelle.

La compaction de l'ADN nucléaire en nucléosomes, corrélée à une répression transcriptionnelle nécessitent l'intervention de deux types d'activités enzymatiques définies précédemment permettant une régulation de la structure chromatinienne. Les complexes multiprotéiques intervenant lors de la modification et le remodelage de la chromatine semblent être présents chez *P. falciparum*. Des homologues de la famille GCN5 ont été caractérisés impliquant l'existence de l'acétylation des histones (Fan et al. 2004b). De même, la présence d'un homologue d'ADA2 qui est une sous-unité du complexe HAT GCN5 très conservée au cours de l'évolution a été rapportée (Fan et al. 2004a). L'acétylation des histones est un processus dynamique et réversible régulée également par les histones déacétylases dont *P. falciparum* semble posséder plusieurs représentants conventionnelles (Joshi et al. 1999) et deux membres dans la famille SIRTUIN, dont l'un des membres les plus étudiés est Sir2. Tout comme chez la levure, où Sir2 est nécessaire à l'extinction de la transcription de certains gènes télomériques (Blander and Guarente 2004), chez *P. falciparum* PfSir2 est impliqué dans la répression de certains gènes de la famille *var* situés dans les télomères (Duraisingh et al. 2005{Freitas-Junior, 2005 #197; Freitas-Junior et al. 2005). La méthylation des histones comme l'acétylation est associée à une transcription active et est contrôlée par des histones méthylases qui ont pu être identifiées dans le génome de *P. falciparum* au

nombre de quatre, toutes appartenant à la famille SET. Pour finir, le remodelage de la chromatine aboutit à la modification des interactions histone-ADN grâce à l'intervention de complexes munis d'une sous-unité ATPase et *P. falciparum* dispose au moins de neuf ATPases homologues de SNF2/SWI2, une ATPase de type ISWI et une ATPase homologue de CHD. Il est à noter la présence également de plusieurs gènes codant pour des protéines à domaines PHD ou Bromodomaine connus pour être impliqués dans les processus de remodelage de la chromatine.

Dans la catégorie des facteurs de la classe HMG, seules les protéines de type architectural HMGB ont été trouvées dans le génome de *P. falciparum*. En effet, trois protéines de petite taille contenant un domaine HMG-box et une protéine contenant deux domaines HMG-box ont été annotées. Cette dernière protéine HMGB présente en plus un domaine de liaison à l'ADN caractéristique de la famille Myb.

Il apparaît donc que *P. falciparum* possède une machinerie typique des eucaryotes capable d'engendrer les modifications et le remodelage de la chromatine pour contrôler la régulation transcriptionnelle.

### III.2.3 – Facteurs de transcription spécifiques de séquence

Comme annoncé au début de ce chapitre, l'analyse du génome de *P. falciparum* a révélé une apparente pauvreté en facteurs de transcription (FT) potentiels. En effet, suivant la classification de ces protéines selon la base de données TRANSFAC, plusieurs classes n'ont pas actuellement de représentant dans le génome. Mais l'analyse actuelle du génome permet d'identifier par homologie de séquences des protéines appartenant à une classe de FT au moins parmi les quatre superclasses définies au paragraphe I.4.3 (voir annexe I.1).

La superclasse des protéines à domaines basiques figure dans le génome avec un nombre élevé de protéines, elle est représentée par 26 FT putatifs possédant un domaine bZIP.

La superclasse des protéines à motif en doigt de zinc est sans doute la mieux représentée avec :

- près de 20 protéines à domaines C<sub>2</sub>H<sub>2</sub> où l'ion zinc central est chélaté par deux cystéines et deux histidines,

- au moins trois protéines à domaine en doigt de zinc B-box de type Constans impliqué notamment dans la régulation transcriptionnelle de la floraison chez *Arabidopsis thaliana* (Samach et al. 2000),

- quatre protéines à domaine en doigt de zinc de type MIND, motif associé à la répression transcriptionnelle chez l'homme (Lutterbach et al. 1998),

- douze protéines à domaine en doigt de zinc de type DHHC, motif très conservé au cours de l'évolution des eucaryotes qui participerait à la liaison de l'ADN (Putilina et al. 1999).

La superclasse des Hélice-Tour-Hélice ne figure dans le génome de *P. falciparum* que par la présence d'une classe et une seule famille dans cette classe : les facteurs à domaines tryptophane ou famille Myb. Au moins sept protéines possèdent au minimum les deux régions R2 et R3 nécessaires à la liaison à l'ADN (voir paragraphe I.4.2.c). La plupart de ces protéines ont des domaines putatifs dans lesquels certains tryptophanes sont remplacés par des phénylalanines ou des tyrosines, ce qui a été observé aussi chez l'amibe *D. discoideum* où ces facteurs sont fonctionnels.

Enfin, dans la superclasse des protéines à architecture  $\beta$ , cinq classes sont représentées parmi les onze classes très variées y figurant. Nous ne reparlerons ni des facteurs de la classe des TBP qui sont des facteurs généraux de la transcription (voir paragraphe III.2.1) ni des facteurs de la classe HMGB qui ne sont pas des facteurs de transcription mais des facteurs architecturaux (voir paragraphe III.2.2) :

- une protéine de la classe des boîtes MADS, connues chez les eucaryotes pour lier spécifiquement des séquences d'ADN riches en AT (Mo et al. 2001). Il est intéressant de noter que ces FT interagissent avec les protéines à motif en doigt de zinc de type C<sub>2</sub>H<sub>2</sub>, B-box et avec les protéines à domaines basiques de type bZIP, tous potentiellement présents dans le génome de *P. falciparum*.

- plusieurs protéines de la classe des hétérodimères CCAAT est également présente dans le génome par l'existence des trois gènes codant pour les trois sous-unités du facteur CBF/NF-Y. CBF/NF-Y interviendrait dans la stabilisation de la liaison à l'ADN d'autres facteurs et serait capable d'interagir avec la TBP ainsi qu'à des HAT (Maity and de Crombrugghe 1998).



- deux protéines de la classe des domaines Cold Shock sont présentes dans le génome de *P. falciparum*. Ces protéines sont capables de se lier spécifiquement à l'ADN ou à l'ARN chez les eucaryotes et sont impliquées aussi bien dans la régulation transcriptionnelle que post-transcriptionnelle (Matsumoto and Wolffe 1998; Wistow 1990). De plus, ces facteurs ont gardé leur capacité de réponse à une baisse brutale de la température chez certaines plantes (Karlson and Imai 2003), ce qui pourrait leur donner une signification fonctionnelle chez *P. falciparum* lors de son passage de l'hôte vertébré (37 °C) à l'hôte invertébré (25°C).

Une grande partie des voies de régulation entre ADN et protéines, connues à ce jour, possède des représentants dans le génome de *Plasmodium* mais nécessite encore pour la plupart une validation biologique de leur rôle lors du cycle du parasite. Cette analyse du génome de *P. falciparum* et des protéines potentiellement impliquées dans la régulation génétique révèle une grande diversité dans les mécanismes disponibles participant à la régulation transcriptionnelle.

## **IV – La régulation génétique chez *P. falciparum***

### **IV.1 - L'expression des gènes plasmodiaux**

La capacité du parasite à se développer et à se multiplier dans les très différents environnements que sont ses deux hôtes vertébré et invertébré est associée à une expression des gènes hautement spécifique de chaque stade du développement. En effet, les changements morphologiques résultent de la variation du niveau et de la qualité des ARNm au cours du cycle cellulaire reflétée par des profils d'expression de protéines distincts d'un stade à un autre, comme cela a été décrit il y a maintenant plus de 10 ans pour les gènes *MSA2*, *kahrp* et *SERA* (Barron 1992; Fox and Bzik 1994; Lanzer et al. 1992a).

La génomique fonctionnelle a fondamentalement changé l'approche traditionnelle gène par gène en tirant partie des efforts du séquençage de *Plasmodium*. Néanmoins, trois études à grande échelle avaient été menées, avant la complétion du génome réalisée en Octobre 2002, dans le but de comparer l'abondance relative des transcrits au cours du cycle érythrocytaire. La première description de transcrits spécifiques des stades sexués (gamétocytes) et asexués (trophozoïtes) a été réalisée en utilisant une puce à ADN comprenant 3700 séquences aléatoires à partir d'une banque génomique (Hayward et al.

2000). La technique d'analyse quantitative et qualitative en série (SAGE) de l'abondance des transcrits a ensuite permis d'identifier 4900 gènes exprimés au cours du cycle érythrocytaire, démontrant qu'une grande partie du génome était exprimée à cette phase du cycle (Patankar et al. 2001). Une analyse par puce à ADN plus précise, sur 5 temps du cycle asexué, mais sur un nombre de gènes réduits à 944, a fourni la première image de l'expression temporelle d'une partie des gènes de *P. falciparum* où 60 % des gènes étudiés montraient une régulation significative (Ben Mamoun et al. 2001).

Avec la publication du génome, les bio-puces à ADN pan-génomiques ou presque ont été utilisées avec succès pour étudier les différents stades de développement du *Plasmodium*. Cependant, les analyses du transcriptome requièrent des quantités importantes d'ARN ( $\mu\text{g}$ ) et ont été malgré tout restreintes aux stades cultivables *ex vivo*, limitant les études à grande échelle aux seuls stades sanguins de *P. falciparum* et particulièrement d'isolats adaptés au laboratoire. L'utilisation de milliers d'oligonucléotides longs (70 mers) par puce a fait preuve d'efficacité pour l'étude du génome entier d'abord sur deux stades synchronisés asexués trophozoïtes et schizontes (Bozdech et al. 2003b) puis toutes les heures sur les 48 h du cycle érythrocytaire (Bozdech et al. 2003a). La technologie Affymetrix a permis de valider l'utilisation d'oligonucléotides courts (25 mers) tout d'abord pour l'analyse du transcriptome des gènes du chromosome 2 qui montre que plus d'un tiers des gènes étaient significativement régulés au cours des stades asexués anneaux, trophozoïtes et schizontes (Le Roch et al. 2002). Puis, une étude plus approfondie avec 22000 cibles couvrant la totalité du génome du clone 3D7 a fourni le transcriptome de neuf points du cycle dont le stade sporozoïte, sept points de stades asexués de l'anneau au mérozoïte et un point de stade sexué (Le Roch et al. 2003). Aucun profil général d'expression n'avait été défini chez *P. falciparum* avant ces deux dernières études. Il semble que 80 % des gènes du parasite soient exprimés au cours du cycle érythrocytaire, laissant penser que peu de gènes ont une expression strictement exo-érythrocytaire. Les analyses du transcriptome de *P. falciparum* montre une expression séquentielle des gènes avec un maximum d'expression des ARNm codant pour les protéines de synthèse en premier, suivis des gènes associés aux protéines du métabolisme, puis des gènes associés à l'adhésion/invasion et enfin des gènes codant pour les protéines kinases. Du contrôle fin de l'expression des gènes requis au cours du cycle de

vie de *Plasmodium* résulte en outre un répertoire protéique singulièrement différent à chaque stade du cycle (Florens et al. 2002; Lasonder et al. 2002).

Très récemment, une bio-puce à ADN comprenant 7462 oligonucléotides longs représentant la quasi-totalité du génome a été utilisée pour identifier de nouveaux gènes impliqués dans la gamétytogenèse (Silvestrini et al. 2005). Une analyse comparative a été réalisée, en condition d'induction de la gamétytogenèse, entre le clone parasitaire de référence 3D7 (Walliker et al. 1987) produisant des gamétocytes en culture et son clone dérivé F12, incapable de produire des gamétocytes (Alano et al. 1995). Un groupe de 117 gènes présente un profil d'expression similaire à ceux des gènes *pfs16* et *pfgr27*, marqueurs précoces de la gamétytogenèse (Bruce et al. 1994; Lobo et al. 1999). Deux protéines Pfpeg-3 et Pfpeg-4 ont été identifiées comme spécifiques des gamétocytes de stades I et II, respectivement. Ces résultats seront discutés après la présentation de nos travaux.

Les différentes expériences décrivant le transcriptome du parasite ont montré que l'abondance de la plupart des ARNm varie au cours du cycle cellulaire, suggérant que l'expression des gènes est contrôlée en partie au niveau de la synthèse des ARN.

#### **IV.2 - La régulation transcriptionnelle**

La régulation de l'expression des gènes au niveau transcriptionnel a été décrite pour différents types de transcrits chez *P. falciparum* : i) les gènes des *ARNr* sont exprimés d'une manière spécifique de chaque stade (Waters 1994) ii) un seul des 50 gènes *var* est exprimé au sein d'un parasite alors que tous les autres sont réprimés (Scherf et al. 1998) et iii) l'activation des promoteurs des gènes *pfs16* et *pfs25* marque la transition vers la différenciation sexuée (Dechering et al. 1999). Les profils d'expression des gènes qui accompagnent le cycle de vie sexué et asexué du parasite montrent que le contrôle de la régulation génétique est crucial pour la survie du parasite aussi bien que pour sa virulence.

Chez les eucaryotes, il est établi que le niveau de transcription est contrôlé par les éléments agissant en *cis* que sont les séquences promotrices et les éléments agissant en *trans* que sont les facteurs de transcription; la disponibilité des FT et l'accessibilité à leurs séquences cibles permettant le contrôle de la régulation transcriptionnelle. Il est raisonnable de penser que cela se passe aussi chez *Plasmodium*.

#### IV.2.2 - Les éléments de régulation

L'identification d'éléments de régulation impliqués dans la régulation transcriptionnelle est difficile en raison de la richesse en A+T du génome de *P. falciparum*. En effet, identifier les boîtes TATA des promoteurs dans un océan d'étendues A et T représente un réel défi aussi bien que de cloner ces séquences. La plupart des informations sur la structure des régions promotrices a essentiellement été obtenue grâce à l'étude du gène *gbp130* (Lanzer et al. 1992b) et plus récemment sur des gènes exprimés aux stades sexuels (Alano et al. 1996; Dechering et al. 1999). L'avènement de la technique de transfection a permis de valider l'activité de quelques promoteurs par des études fonctionnelles en utilisant des gènes rapporteurs (Horrocks and Kilbey 1996).

Bien que la structure générale des promoteurs du parasite soit conforme à la structure bi-partite caractéristique des promoteurs eucaryotes (voir paragraphe III.1), quelques différences sont observées comme tout d'abord la distance qui sépare les éléments *cis* caractéristiques de la région basale et le site d'initiation de la transcription, qui est plus longue que chez les autres eucaryotes (Crabb and Cowman 1996; Horrocks and Kilbey 1996; Levitt 1993). Par exemple, la boîte TATA des promoteurs des gènes *kahrp* et *gbp130* se situent à -81 pb et -186 pb du site d'initiation de la transcription, respectivement (Ruvalcaba-Salazar et al. 2005). En général elle se situe entre -25/-30 pb quoique chez la levure, cette distance peut aller jusqu'à -140 pb. De plus, aucun élément de régulation typique des eucaryotes tels que les boîtes CCAAT ou les sites de liaison SP1 n'a été retrouvé comme l'a révélé l'étude du promoteur proximal du gène *gbp130*.

Toutefois, dans la partie distale de ce même promoteur, un élément ayant une homologie avec la séquence activatrice SV40 du virus simien a été identifiée et est capable d'interagir spécifiquement avec des protéines nucléaires (Lanzer et al. 1992b). Mais jusqu'à ce jour, aucune étude menée sur *Plasmodium* ne fait état d'une activité de promoteurs viraux comme SV40 actifs chez les eucaryotes supérieurs et réciproquement il semble que les promoteurs du parasite ne fonctionnent pas dans les cellules de mammifères (Horrocks, 1998). De plus, une étude sur les promoteurs des gènes *var* a révélé l'existence de trois séquences *cis*-activatrices SPE1, SPE2 et CPE, toutes exclusivement associées aux gènes *var* qu'ils soient positionnés au niveau des télomères ou au centre des différents chromosomes et ne présentant aucune similitude avec des sites de liaison de FT eucaryotes répertoriés dans la

base de données TRANSFAC (Voss et al. 2003). Ces données tendent vers une divergence dans l'évolution des éléments de régulation de *Plasmodium* et suggère l'utilisation de différents mécanismes dans la régulation transcriptionnelle du parasite (Aravind et al. 2003).

Néanmoins, quelques études ont permis l'identification, dans les régions distales des promoteurs des gènes *pfs25*, *gbp130* et *CDP diacylglycerol synthase*, d'éléments de régulation non homologues à ceux des eucaryotes, mais responsables de l'activation et de la répression de l'expression du gène rapporteur ; ces éléments de régulation étant capables de lier spécifiquement des facteurs nucléaires (Dechering et al. 1999; Horrocks and Lanzer 1999b; Lanzer et al. 1992a; Osta et al. 2002). Plus récemment, une étude a porté sur une analyse bioinformatique de 18 régions promotrices des gènes *hsp* suivie par une validation biologique par transfection (Militello et al. 2004). Un élément palindromique boîte-G, positionné à -195 pb du site d'initiation de la transcription, présent en 2 exemplaires dans le promoteur du gène *hsp86*, apparaît conservé parmi les espèces de *Plasmodium* (*falciparum*, *yoelii*, *berghei* et *vivax*). Son implication dans la stimulation de la transcription chez *Plasmodium spp* pourrait être liée à sa structure tige-boucle car aucun complexe d'interaction ADN-protéine n'a été obtenu en présence d'extraits nucléaires. Ces expériences montrent l'importance potentielle de facteurs régulant localement la structure de la chromatine dans les mécanismes de régulation des gènes de *Plasmodium*. Une autre région USE (upstream sequence element) située juste avant les 2 boîtes-G du gène *hsp86* participe à la stimulation de la transcription. L'arrangement de ces éléments de régulation est conforme au modèle standard eucaryote contenant un promoteur basal et des séquences activatrices dans la région distale en amont. Pour autant, la difficulté d'identifier des éléments de régulation sur les 5300 gènes prédits reste entière, compte tenu de l'apparente spécificité des ces éléments chez *P. falciparum*.

Par ailleurs, une étude sur le gène *pgs28* exprimé lors du stade sexué de *P. gallinaceum* a révélé la présence d'un élément *cis*-régulateur en aval du site d'initiation de la transcription dont la fonction est indépendante de la boîte TATA ou d'un élément Inr (Chow and Wirth 2003). Cet élément situé à +25 pb du site d'initiation de la transcription interagit spécifiquement avec des facteurs nucléaires et semble important pour réguler l'activité du promoteur en participant à l'effet des éléments activateurs présents en amont. A la lumière de ce résultat, il apparaît intéressant de reconsidérer les séquences en aval du site d'initiation

de la transcription chez *Plasmodium spp* compte tenu de la prépondérance des gènes dépourvus de boîte TATA.

Enfin, un autre élément situé en dehors du promoteur intervient dans le contrôle transcriptionnel du gène *var7b* chez *P. falciparum* (Deitsch et al. 2001). En effet, l'intron de ce gène agit comme un élément de répression de la transcription, il participe à la régulation en coopération avec des éléments présents en amont de la région codante. La présence d'une région très riche en A+T semble suffisante pour réprimer le promoteur du gène *var* associé (Calderwood et al. 2003). L'intervention coopérative de ces différents éléments implique la nécessité d'une configuration de la chromatine appropriée. Cependant, le contrôle de la transcription effectué par ces éléments de régulation analysé avec un vecteur épisomique pourrait être différent du contrôle du gène au sein du chromosome (Voss et al. 2003), comme cela a aussi été montré pour le gène *gfp130* (Horrocks and Lanzer 1999a).

En conclusion, même si peu de travaux ont porté sur la régulation transcriptionnelle des gènes, la régulation génétique chez *Plasmodium* semble évidente au niveau transcriptionnelle grâce à l'identification dans les promoteurs d'éléments capables de promouvoir l'expression du gène rapporteur. Pour autant les promoteurs étudiés contiennent des éléments de régulation différents des autres eucaryotes. Cette différence majeure réside probablement dans l'extrême richesse en A+T des régions intergéniques, qui peut atteindre 90%. Une forte occurrence d'étendues homopolymériques (dA:dT) (> 10 pb) procure à ces séquences promotrices une propriété structurale et fonctionnelle de courbure de l'ADN. Cette flexibilité peut jouer un rôle dans la modulation de l'activité transcriptionnelle des gènes en augmentant l'affinité de l'ADN pour les FT ou l'accessibilité des FT à l'ADN par une distorsion locale du nucléosome (Dechering et al. 1998). Ces étendues (dA:dT) sont connues depuis longtemps pour être des éléments de régulation fonctionnels chez la levure (Struhl 1985). La difficulté de clonage de ces régions liée à leur composition complique l'étude précise de l'influence de ces étendues homopolymériques sur l'activité des promoteurs. Néanmoins, il a été montré qu'une telle région était impliquée dans l'activité transcriptionnelle des promoteurs de l'ADN polymérase  $\delta$  et de la calmoduline de *P. falciparum*, cet effet étant corrélé à une courbure de l'ADN (Polson and Blackman 2005; Porter 2001).

La signification fonctionnelle et structurale des régions homopolymériques des promoteurs s'ajoute à la complexité de la régulation transcriptionnelle chez *P. falciparum*. L'ensemble de ces données tend à montrer qu'au-delà d'un cadre commun aux eucaryotes, il est raisonnable de penser que *P. falciparum* a évolué vers une différenciation des éléments de régulation.

#### IV.2.1 – Les facteurs de transcription

En complément à la recherche des éléments *cis*, des efforts pour élucider la présence de facteurs *trans* restent à faire. Bien que différents auteurs aient montré par analyse informatique une faible représentation des FT chez *Plasmodium* (Callebaut et al. 2005; Coulson et al. 2004), notre annotation du génome de *P. falciparum* ainsi que les données de PlasmoDB ont cependant permis d'identifier un bon nombre de protéines candidates à la fonction de régulation génétique. A ce jour, aucune étude n'a abouti avant notre travail à la caractérisation de facteur de transcription, seul le facteur général de la transcription TBP a été décrit (McAndrew et al. 1993).

En 2004, une technique d'identification de facteurs de transcription à haut débit a été employée avec des extraits nucléaires de *P. falciparum* (Kumar et al. 2004) pour leur capacité à interagir avec des éléments de régulation trouvés chez les eucaryotes. La mise en évidence d'activité de liaison aux 54 séquences *cis*-régulatrices immobilisées sur une membrane permet de déterminer la présence de certaines classes de FT dans l'échantillon. Cependant, ce système a été prévu initialement pour sa compatibilité avec l'homme, la souris et le rat, les séquences cibles choisies étant conservées entre ces espèces. Pour cette raison, il n'est pas étonnant que seuls deux des huit signaux positifs obtenus aient été validés par la traditionnelle expérience de retard sur gel. Ces signaux correspondent à la présence de protéines des familles Myb et MADS, qui toutes deux ont été retrouvées au cours de l'exploration du génome de *P. falciparum*. De plus, une étude récente sur le réseau d'interactions des protéines de *P. falciparum* par la méthode du double hybride chez la levure a mis en évidence des groupes d'interactions entre protéines impliquées dans la transcription et la régulation de la transcription (LaCount et al. 2005). Une histone acétyltransférase de la famille Gcn5 est au centre d'un réseau d'interactions de nombreuses protéines potentielles incluant des protéines de modification de la chromatine comportant des domaines PHD, bromo, BTB/POZ et SET et aussi deux facteurs de transcription putatifs de la famille Myb,

que nous avons annotés lors de notre exploration du génome (PF11\_0241, PFL0815w, voir Annexes I.1). Ces données indiquent que ces complexes formés pour la modification de la chromatine pourraient bien correspondre à des régions spécifiques dans le génome pour réguler la transcription et ont une signification particulière dans le contexte de "pénurie de FT" annoncée précédemment. Il apparaît donc essentiel d'étudier la fonction biologique de ces facteurs et leur implication dans la régulation transcriptionnelle de *P. falciparum*.

Au cours de mon travail de thèse, nous nous sommes concentrés sur l'aspect transcriptionnel de la régulation des gènes chez *P. falciparum* lors du développement érythrocytaire. Il n'en reste pas moins qu'une régulation au niveau post-transcriptionnel des gènes est vraisemblable bien qu'elle ait fait l'objet d'un nombre limité d'études.

### **IV.3 - La régulation post-transcriptionnelle**

#### **IV.3.1 – La régulation de l'épissage et de la stabilité des ARNm**

L'épissage des ARNm permet d'obtenir, à partir du même transcrit primaire, les transcrits pouvant être transportés vers le cytoplasme pour y être traduits. Cependant il peut apparaître plusieurs produits d'épissage différant dans la composition de leurs exons, donnant ainsi naissance à des protéines similaires mais dont la fonction peut être différente. C'est un mécanisme très important chez les métazoaires car il permet de générer de la diversité à partir d'un même gène. Les jonctions entre les exons de *P. falciparum* sont démarqués par des signaux communs à la plupart des eucaryotes (Vinkenoog et al. 1995). L'épissage alternatif a été décrit pour quelques gènes de *P. falciparum* (Knapp et al. 1991; Pace et al. 1998; Singh et al. 2004; van Lin et al. 2001), mais seules les isoformes du gène de l'adényl cyclase présenteraient des différences fonctionnelles (Muhia et al. 2003). Le contrôle de l'épissage pourrait être impliqué dans les mécanismes de régulation post-transcriptionnelle du *Plasmodium* en faveur d'une diversité fonctionnelle à partir de son faible nombre de gènes et au regard de la complexité de son cycle.

Lors de son séjour dans le cytoplasme, la vie d'un ARNm et donc le temps durant lequel il pourra servir de matrice pour l'expression d'une protéine sont déterminés par sa



stabilité. Ainsi contrôler la stabilité d'un ARNm constitue un moyen de moduler l'expression d'une protéine. Un tel phénomène prend place chez *Plasmodium* comme l'indique la traduction extrêmement retardée de la protéine Pbs21, qui est produite seulement dans les stades zygote et ookinète alors que son message est transcrit beaucoup plus tôt au stade gametocyte (Shaw et al. 1996). Les régions non traduites (3'UTR) des ARNm possèdent des éléments responsables de la stabilité des transcrits (Wickens et al. 1997) qui interagissent spécifiquement avec des facteurs de la classe RBP (RNA-binding proteins), dont *P. falciparum* possède deux membres appartenant à la famille Puf (Cui et al. 2002). La conservation des protéines PfPuf et leur expression au cours du développement et de la différenciation sexuée chez *Plasmodium spp* suggèrent l'existence d'un mécanisme de régulation de la stabilité des ARN commun aux eucaryotes. De plus, la mise en évidence chez *P. falciparum* de réseaux d'interactions faisant intervenir plusieurs protéines comportant des motifs de reconnaissance de l'ARN (RRM), par la méthode du double hybride chez la levure, indique une éventuelle implication de ces protéines dans la maturation ou l'épissage des ARN (LaCount et al. 2005).

#### IV.3.2 – Les ARN anti-sens

La dernière décennie a montré un intérêt grandissant pour l'ARN en tant que régulateur de l'expression génétique et la présence d'ARN anti-sens naturels a été rapportée chez une grande variété d'eucaryotes (Vanhee-Brossollet and Vaquero 1998). Les ARN anti-sens sont généralement issus du même locus que les transcrits sens, synthétisés à partir du brin d'ADN opposé et chevauchent partiellement leur contrepartie sens avec une complémentarité parfaite par opposition aux microARN plus récemment décrits (Ambros 2001). Initialement décrits chez les procaryotes, ils ont pour rôle de diminuer l'expression des transcrits sens en formant des complexes sens/anti-sens mais ils peuvent aussi coder pour des protéines (Vanhee-Brossollet and Vaquero 1998). Ce mécanisme de régulation "anti-sens" pourrait avoir un sens chez *Plasmodium* au regard de la complexité de son génome et de sa capacité codante limitée.

La première étude faisant état d'une accumulation inverse de transcrit sens et anti-sens chez *P. falciparum* concerne le gène *Pfmsp2* (Kyes et al. 2002). Mais dans ce cas précis, le rôle régulateur de l'ARN anti-sens a été écarté car mis dans un autre contexte chromosomique, *Pfmsp2* sous le contrôle de son promoteur présente le même profil

d'expression (Wickham et al. 2003). Ensuite, des expériences de SAGE réalisées sur le cycle érythrocytaire de *P. falciparum* indiquent une transcription substantielle de 15 % à partir du brin anti-sens (Gunasekera et al. 2004; Patankar et al. 2001). Pourtant, le modèle anti-sens chez *Plasmodium* reste controversé puisque la détection directe d'ARN anti-sens n'a pas été rapportée. Malgré tout, il a été montré par des expériences de "nuclear run-on" que la production d'ARN anti-sens est dépendante de l'activité de l'ARN polymérase II (Militello et al. 2005). Les auteurs proposent un mécanisme de "chromatine ouverte" dans lequel le contexte chromatinien de chaque locus assumerait un rôle dans la synthèse exclusivement simultanée par l'ARN polymérase II d'ARN sens et anti-sens sans distinction. Cette hypothèse est renforcée d'une part, par le fait que la transcription sens et anti-sens a été détectée dans des parasites synchronisés et d'autre part qu'aucun élément promoteur canonique dictant la polarité de la transcription n'ait été décrit chez *P. falciparum* à ce jour. Plus récemment, la répression des gènes *var* a été étudiée en fonction de la production d'ARN anti-sens et il apparaît que l'augmentation des transcrits anti-sens au niveau des loci réprimés n'intervient pas dans l'inhibition des gènes *var* (Ralph et al. 2005). En effet, dans certains cas, l'abondance des transcrits anti-sens coïncident avec l'abondance des transcrits sens suivant le modèle de "chromatine ouverte".

La question de l'implication potentielle de l'ARN anti-sens dans la régulation post-transcriptionnelle de *P. falciparum* reste entière. Même si actuellement un mécanisme de régulation anti-sens naturel ne semble pas exister chez *Plasmodium*, la capacité d'oligonucléotides anti-sens (Noonpakdee et al. 2003; Wanidworanun et al. 1999) ainsi que de double brins d'ARN (Malhotra et al. 2002; McRobert and McConkey 2002) à moduler l'expression des gènes semble néanmoins être une preuve raisonnable de la présence d'une machinerie capable de prendre en charge des duplexes sens/anti-sens.

Enfin, en plus de la régulation apportée au niveau de l'abondance des messagers (transcription, stabilité, transport), il est vraisemblable que la régulation au niveau de la traduction et de la maturation des protéines existe aussi chez *Plasmodium*.

## PRESENTATION DES TRAVAUX DE THESE

Au cours du développement de *Plasmodium falciparum*, la phase érythrocytaire chez l'homme est cruciale car elle est responsable de la maladie et de la transmission du parasite *via* le moustique. Cette étape clé du cycle alterne entre prolifération intense et différenciation en gamétocytes avec des stades asexués et sexués, respectivement. Notre intérêt s'est porté sur cette alternance qui nécessite vraisemblablement un contrôle fin des activités transcriptionnelles de chaque gène impliqué dans ces événements.

Le séquençage complet du génome de *P. falciparum* a donné un élan formidable à la recherche sur la biologie moléculaire du parasite mais un obstacle considérable subsiste : environ 60 % des protéines prédites n'ont pas de fonction connue. Ainsi, l'ensemble des données sur la régulation de l'expression des gènes chez *P. falciparum* se traduit par une grande méconnaissance des processus fins et des acteurs responsables de ceux-ci accompagnée cependant de nombreuses évidences d'une proximité avec les autres eucaryotes. Les travaux que j'ai réalisés pendant ma thèse s'inscrivent dans cette optique de compréhension du rôle des acteurs de la régulation génétique chez *P. falciparum*, pour mieux appréhender les mécanismes impliqués dans le contrôle transcriptionnel de l'expression des gènes au cours du cycle érythrocytaire.

Dans la première partie de ma thèse, nous avons cherché à identifier des gènes impliqués dans la régulation des événements intervenant lors du développement érythrocytaire du parasite notamment lors de la transition de la prolifération asexuée à la différenciation sexuée. Pour cela, nous avons réalisé une analyse comparative du transcriptome entre deux clones de *P. falciparum* : le clone de référence 3D7 produisant des gamétocytes en culture et son clone dérivé F12 qui a perdu la capacité de produire des gamétocytes. Cette étude, initiée avant le séquençage complet du génome de *P. falciparum*, a été réalisée avec une puce à ADN construite au sein de notre laboratoire et représentant 150 gènes impliqués dans la régulation transcriptionnelle, la transduction du signal et le cycle cellulaire.

Dans la deuxième partie, nous avons décidé d'étudier un facteur de transcription spécifique de séquence, PfMyb1 de la famille Myb des protéines à domaines tryptophane qui participe à la régulation de l'expression des gènes impliqués dans le développement et la différenciation. L'analyse bioinformatique de la séquence de PfMyb1 a permis de révéler que ce FT possédait toutes les caractéristiques de la famille Myb. Nous avons ensuite montré que la protéine PfMyb1, présente dans les extraits nucléaires du parasite, était capable de se lier spécifiquement à des séquences cibles MRE identifiées *in silico* dans les promoteurs de deux gènes de *P. falciparum*. Pour évaluer l'implication de PfMyb1 dans la régulation transcriptionnelle de *P. falciparum*, nous avons analysé l'effet de l'extinction partielle de l'expression de son transcrit à son pic d'expression au stade trophozoïte. Nous avons pu identifier certains gènes cibles sous le contrôle de PfMyb1 en utilisant notre puce à ADN thématique.

Enfin, dans la troisième partie, nous nous sommes intéressés à la famille des facteurs nucléaires architecturaux HMGB qui interviennent dans le remodelage de la chromatine. Deux protéines HMGB, PfHMGB1 et PfHMGB2, ont été annotées et leurs séquences possèdent les caractéristiques de ces facteurs spécifiques de structure. Nous avons montré les propriétés architecturales des protéines recombinantes PfHMGB1 et PfHMGB2 par leur capacité à se lier à des structures particulières de l'ADN et à courber l'ADN linéaire. Les différences observées tant au niveau de l'expression de ces deux facteurs dans les stades asexués et sexués, que de la localisation cellulaire et des activités d'interaction avec l'ADN, tendent vers une non-redondance de leur fonction dans la régulation transcriptionnelle au cours du développement érythrocytaire de *P. falciparum*.

## **MATERIELS ET METHODES**

## I – MATERIELS

### I.1 – Clones de *Plasmodium falciparum*

Au cours de ce travail, deux clones de *P. falciparum* ont été utilisés, 3D7 et F12, clone dérivé de 3D7 ayant perdu la capacité à produire des gamétocytes ; ces clones ont été gracieusement fournis par les Dr Walliker et Dr Alano, respectivement.

### I.2 – Souches bactériennes et plasmides

Les bactéries *Escherichia coli* TOP10 (Invitrogen) et SURE (Stratagene), transformées avec le plasmide pCR II-TOPO (Invitrogen), nous ont permis de réaliser le clonage des produits de PCR. La souche d'*E. coli* SG13009 (Qiagen), transformée avec le plasmide pQE30 (Qiagen), a été utilisée pour la production des protéines recombinantes PfMyb1, PfHMGB1 et PfHMGB2. Les cellules transformées avec le vecteur de clonage sont cultivées en milieu LB (Luria-Bertani, Invitrogen) sous pression de sélection antibiotique avec de l'ampicilline (100 µg/ml) et de la kanamycine (50 µg/ml) ; ou en milieu 2YT (Invitrogen) avec de l'ampicilline (100 µg/ml) pour la production de protéine recombinante.

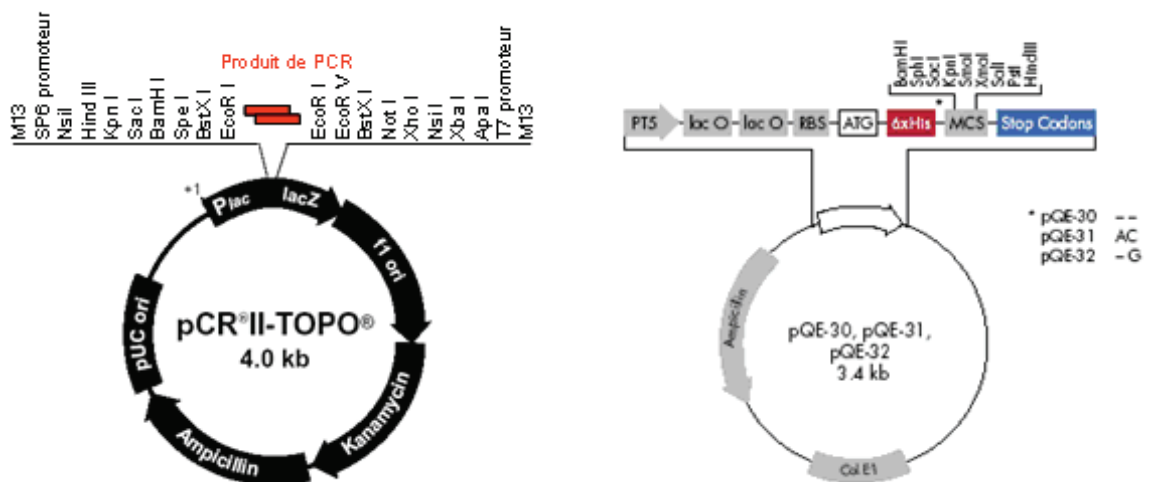


Figure 14. Cartes du vecteur de clonage pCR II-TOPO et du vecteur d'expression pQE30.

### I.3 – Anticorps

Anti-PfMyb1 : Le sérum anti-peptide a été préparé par Eurogentec. Deux lapins ont été immunisés avec deux peptides (RKKY EYKKN SWTKK et KRRYLRLISATNNMEQ),

correspondants aux acides aminés 103-117 (en amont du domaine R1) et 400-414 (à la fin du domaine R3) des domaines tryptophane de la protéine PfMyb1. Un sérum capable de reconnaître la protéine recombinante PfMyb1 a été obtenu après trois inoculations (jour 100) et a donc été sélectionné pour être purifié par immuno-affinité avec un des deux peptides. Toutes les études ont été réalisées avec le sérum anti-PfMyb1 purifié contre le peptide 400-414 (0,46 mg/ml). Les sera préimmun, immun et immuno-purifié ont été utilisés pour les analyses par Western blot et les retards sur gel.

Anti-p50 : Cet anticorps dirigé contre la protéine p50 du facteur de transcription NF-kB a été gracieusement fourni par le Pr A. Israël. Il a servi comme témoin négatif dans les expériences de retard sur gel.

Anti-PfPk5 : Cet anticorps est un don du Dr C. Doerig et a été utilisé dans les expériences d'interférence avec le transcrit *pfMyb1*.

Anti-PfHMGB1 et anti-PfHMGB2 : Les sera anti-HMGB ont été préparés dans le laboratoire par J. F. Franetich. Cinq souris Balb/c ont reçu deux injections sous-cutanée d'environ 50 µg de protéines recombinantes dans l'adjuvant de Freund. Les sera avec une activité anti-HMGB élevée ont été collectés après le premier rappel (jour 30) et sélectionnés pour les analyses immuno-chimiques.

Anti-PfHSP70 : Cet anticorps est un don d'Olivier Silvie et a été utilisé pour normaliser le niveau d'expression des protéines par western blot.

#### **I.4 – Oligonucléotides**

Pour le clonage des phases ouvertes de lecture des gènes *pfmyb1*, *pfhmg1* et *pfhmg2*, les oligonucléotides suivants ont servi d'amorces pour les réactions d'amplification par PCR :

*pfmyb1* sens : 5'-GGTGGATCCATGTTTAATAATAATAACAAATATGAGTGT-3'

*pfmyb1* antisens : 5'-GGTAAGCTTCTTATTGTTCCATGTCATTGTTGCAC-3'

*pfhmg1* sens : 5'-GGTGGATCCATGAAGAATACAGGAAAAGAAG-3'

*pfhmg1* antisens : 5'-ATTGGATCCATGGCTTCAAAATCTCAAAA-3'

*pfhmg2* sens : 5'-GGTCCCGGGGCCAATTAAGCTTTCATTTTC-3'

*pflmgb2* antisens : 5'-ATTGGTACCTTATTCTTGATTTTTCTTTC-3'

Pour l'analyse de l'activité de liaison du facteur de transcription PfMyb1, trois oligonucléotides ont été synthétisés (Invitrogen) comprenant la séquence cible "Myb Regulatory Element" (MRE) conforme à la séquence consensus  $T/cAAC^T/cG$  : deux MRE putatifs de *Plasmodium* présents dans les promoteurs de deux gènes des stades sexué *pfcrk1* (-287 à -266) et asexué *pfmap1* (-1503 à -1482) et un MRE prototype présent dans le promoteur du gène *mim-1* de poulet. L'élément de régulation NF- $\kappa$ B situé dans le promoteur du gène *IL2-R $\alpha$*  a été utilisé pour les compétitions non spécifiques (Galio et al. 1997).

Pour l'analyse de l'activité de liaison à l'ADN cruciforme des protéines PfHMGB, les oligonucléotides partiellement complémentaires 1 à 4 ont été utilisés pour former l'ADN cruciforme complet à quatre brins 4H et les formes incomplètes 3H et 2H :

brin 1 : 5'-CCCTATACCCCTGCATTGAATTCCAGTCTGATAA-3'

brin 2 : 5'-GTAGTCGTGATAGGTGCAGGGGTTATAGG-3'

brin 3 : 5'-AACAGTAGCTCTTATTCGGCTCGCGCCCTATCACGACTA-3'

brin 4 : 5'-TTTATCAGACTGGAATTCAAGCGCGAGCTCGAATAGAGCTACTGT-3'

## II – METHODES

### II.1 – Culture cellulaire

#### II.1.1 - Culture continue

Les parasites ont été cultivés selon une méthode développée par Trager et Jensen (Trager and Jensen 1976) à 37 °C, sous atmosphère contrôlée dans un milieu de culture RPMI 1640 complété avec 0,5 % d'Albumax II (Sérum Albumine Bovine et acides gras). Le milieu de culture est composé de 10,4 g/l de RPMI 1640 avec 0,3 g de glutamine, 35 mM d'Hépès, 25 mM de NaHCO<sub>3</sub>, 50 µg/ml d'hypoxantine, 10 mg/l de gentamycine (Invitrogen) et équilibré à un pH de 6,72. Les parasites sont maintenus à un hématocrite de 5 % en flacon gazé avec un mélange d'oxygène 1 % et de dioxyde de carbone 3 % dans l'azote. Le milieu est renouvelé quotidiennement, la parasitémie est déterminée par numération sur frottis coloré au GIEMSA et lorsqu'elle avoisine 5 à 10 %, selon l'enrichissement en stade cellulaire, la culture est diluée avec des érythrocytes humains frais.



### **II.1.2 – Synchronisation au sorbitol**

L'obtention de parasites "synchrones" ou enrichis en stades cellulaires précis est effectuée par lyse au sorbitol des formes matures (Lambros and Vanderberg 1979). Après centrifugation de la culture pendant 5 min. à 25°C, les parasites sont resuspendus dans une solution de sorbitol à 5 % puis incubés 10 min. à 25°C et lavés deux fois dans du milieu pré-incubé à 37 °C. Trois traitements successifs de sorbitol à 44 h d'intervalle sont effectués pour diminuer au maximum l'écart de temps qui persiste entre les parasites les plus matures des plus jeunes. Cinq stades de développement sont récoltés en fonction du temps et de la morphologie des parasites pour les préparations de transcrits et de protéines : anneaux (10-14 h) ; trophozoïtes jeunes (20-22 h) et matures (26-28 h); schizontes jeunes (32-34 h) et matures (40-42 h post-invasion).

### **II.1.3 - Cryopréservation et décongélation**

Pour la conservation et la préservation des clones, la congélation régulière de parasites est nécessaire. De même que pour limiter la dérive génétique des clones, la culture est établie à partir d'une nouvelle ampoule congelée environ tous les 3 mois.

Après centrifugation d'une culture parasitaire enrichie au stade anneau pendant 5 min. à 25°C, le culot est repris dans 1,5 volume de sérum humain décomplémenté, puis 2,5 volume d'une solution de cryopréservation (28 % glycérol, 3 % sorbitol, 0,65 % NaCl) sont délicatement ajoutés. Les parasites sont aliquotés en 500 µl dans des cryotubes puis directement immergés dans l'azote liquide.

La décongélation des parasites s'effectue directement au bain-marie à 37 °C pendant 1 à 2 min. Un gradient décroissant de sels est alors appliqué : 0,1 volume d'une solution de NaCl 12 % pendant 5 min. à 25°C puis 10 volumes d'une solution de NaCl 1,6 %. Après centrifugation et lavage avec du milieu de culture complet chauffé à 37 °C, les parasites sont mis en culture.

## **II.2 – Extraction des acides nucléiques et clonage**

Le parasite étant intracellulaire, il est nécessaire de procéder à une lyse préalable des globules rouges par la saponine afin d'améliorer les rendements d'extraction. La culture de parasites du clone 3D7 (25 ml) est centrifugée pendant 5 min à 2000 rpm et le culot globulaire

est lavé avec 25 ml de tampon PBS 1X et centrifugé pendant 5 min à 2000 rpm. Les globules rouges sont lysés dans le tampon PBS contenant 0,15 % de saponine pendant 10 min à 25 °C. Après centrifugation pendant 5 min à 2000 rpm à 4 °C, les parasites sont lavés 4 fois avec 1 ml de tampon PBS froid. L'ADN génomique est ensuite préparé avec le kit DNeasy Tissue suivant le protocole "animal tissues" fourni par Qiagen. L'ADN est quantifié par mesure de l'absorbance à une longueur d'onde de 260 nm, sachant qu'une unité d'absorbance correspond à une concentration de 50 µg/ml d'ADN double brin. La pureté des préparations est évaluée en calculant le rapport  $A_{260\text{ nm}}/A_{280\text{ nm}}$  qui doit être compris entre 1,6 et 1,8. Les échantillons sont conservés à 4 °C dans le tampon Tris-HCl 10 mM pH 7,5, EDTA 1mM.

Les phases ouvertes de lecture des gènes *pfmyb1*, *pflmgb1* et *pflmgb2* ont été amplifiées à partir de l'ADN génomique avec les amorces citées au paragraphe I.4 en utilisant les conditions de PCR décrites par (Su et al. 1996) où la température d'élongation est diminuée jusqu'à 60 °C. Les amplicons à la taille attendue sont clonés directement dans le vecteur pCR II-TOPO en suivant le protocole "TA-cloning" d'Invitrogen.

### **II.3 – Extraction et analyse des ARN totaux**

Le protocole utilisé pour l'extraction des ARN totaux de *Plasmodium* a été établi selon la méthode décrite par Kyes et al. (Kyes et al. 2000) en utilisant l'extraction en phase liquide au TRIzol (Invitrogen). Après contrôle du stade cellulaire sur frottis sanguin coloré au Giemsa, les parasites sont centrifugés pendant 5 min à 2000 rpm à 25 °C. Le culot est repris rapidement dans 10 volumes de TRIzol pré-incubé à 37 °C, lorsque la culture est enrichie en stades anneaux ou 20 volumes de TRIzol, lorsque la culture est majoritairement composée de formes matures trophozoïtes ou schizontes, puis incubé à 37 °C pendant 5 min. Les ARN sont extraits par l'ajout de 0,2 volume de chloroforme du volume initial de TRIzol puis agités au vortex pendant 3 min. Après 30 min de centrifugation à 4000 rpm à 4 °C, 0,5 volume d'isopropanol est ajouté à la phase aqueuse précautionneusement prélevée pour la précipitation des ARN pendant 15 h à -20 °C. Après 30 min de centrifugation à 11000 rpm à 4 °C, chaque culot est lavé avec de l'éthanol à 75 % et centrifugé dans les mêmes conditions pour être séchés puis repris dans environ 30 à 50 µl d'eau traitée par 0,1 % de diéthylpyrocarbonate pour être exempte de RNases. Enfin, les échantillons sont traités par 5 unités de DNase I-RNase free/ µg d'ARN pendant 15 min à 37 °C et purifiés par le protocole

"clean-up" du kit Qiagen RNA mini. La qualité et la concentration des ARN peut être déterminée par mesure spectrophotométrique comme pour l'ADN (voir paragraphe II.1) sachant qu'une unité d' $A_{260nm}$  correspond à 40  $\mu\text{g/ml}$  d'ARN et par visualisation des ARN ribosomiques sur gel dénaturant d'agarose 1 %. Cependant, nous avons privilégié l'analyse des profils des ARN totaux par le BioAnalyser 2100 (Agilent Technologies) à l'aide des puces RNA LabChip, qui a pour avantage de ne consommer que quelques centaines de nanogrammes d'ARN contre quelques microgrammes pour un gel classique. Douze échantillons sont déposés sur des "puces" ou unités pré-formés de 16 puits constituant un réseau micro-tubulaire rempli par un gel d'acrylamide linéarisé contenant un intercalant fluorescent de l'ARN détectable à 260 nm. Le temps d'élution mis pour parcourir la distance fixe entre deux électrodes permet de déterminer la taille des fragments. Une représentation de l'intensité de fluorescence en fonction du temps d'élution permet de mettre en évidence les ARN qui apparaissent sous forme de pics. Si le rapport entre l'intensité du pic correspondant à l'ARN 28S et l'intensité du pic correspondant à l'ARN 18S est proche de 1,8, la qualité des ARN est très satisfaisante.

Les transcrits *pfmyb1*, *pfhmgb1* et *pfhmgb2* ont été caractérisés par Northern blot en utilisant les ARN polyA+ purifiés à partir de 35  $\mu\text{g}$  d'ARN totaux d'un mélange des 5 stades de développement de *P. falciparum* (mRNA direct kit de Dynal) pour la détection du transcrit *pfmyb1* ou directement avec 20  $\mu\text{g}$  d'ARN totaux pour les transcrits *pfhmgb*. Après fractionnement sur un gel d'agarose 1% en conditions dénaturantes, les ARNm sont transférés sur une membrane de nitrocellulose 0,2  $\mu$  (Pall) en présence de tampon SSC 20X. L'hybridation est effectuée à 55 °C (*pfmyb1*) ou 65 °C (*pfhmgb*) dans le tampon Ultrahyb (Ambion) avec les ribosondes marquées par l'isotope UTP [ $\alpha$ - $^{32}\text{P}$ ] (Paillard et al. 1990) préparées à partir des vecteurs de clonages pCR-II TOPO respectifs des gènes étudiés en utilisant les ARN polymérase T7 ou SP6 (Promega) en fonction de l'orientation de l'insert.

## II.4 – Expression et purification de protéines recombinantes

La souche d'*E. coli* SG 13009 est transformée avec les constructions pQE30-*pfmyb1*, pQE30-*pfhmgb1* et pQE30-*pfhmgb2* et précultivée toute la nuit à 37 °C dans le milieu 2YT contenant 100  $\mu\text{g/ml}$  d'ampicilline pour une culture ultérieure de 200 ml. L'expression des protéines est induite avec 0,1 mM d'IPTG pendant 3 h à 37 °C. Les cellules sont centrifugées

15 min à 3500 rpm à 4 °C et les culots sont solubilisés dans un tampon de lyse contenant 25 mM Tris-HCl pH 8, 300 mM NaCl, 10 mM imidazole, 10 mM β-mercaptoéthanol, 0,5 % triton X-100, à raison de 1 ml/mg de culot sec, en présence de lysozyme et 1/25 du cocktail d'inhibiteurs de protéases de Roche. La sonication (Bioblock Scientific, Vibracell) est effectuée à 30 W cinq fois pendant 10 s. Les lysats sont ensuite incubés avec 100 µl de billes d'agarose Ni-NTA à 50 % par ml de lysat pendant 2 h à 4 °C sous agitation. Après trois lavages dans le tampon de lyse à 20 mM d'imidazole, les protéines sont éluées à 250 mM d'imidazole dans le tampon Tris-HCl 20 mM pH 8, NaCl 300 mM, excepté pour His-PfHMGB2 qui est éluée à 50 mM. La qualité des préparations est analysée par fractionnement sur gel SDS-PAGE et par western blot en utilisant un anticorps polyclonal de lapin Ig G anti-His (Santa Cruz Biotechnology) à une dilution 1/2000 et un anticorps secondaire couplé à la peroxydase (Sigma) à une dilution 1/10000, la réaction est révélée par chimioluminescence (Supersignal ECL, Perkin Elmer).

## **II.5 – Immunoprécipitation de protéines et western blot en série**

Pour la détection de la protéine PfMyb1 dans les lysats parasitaires, une étape d'immunoprécipitation a dû être réalisée avant l'analyse par western blot. Les extraits protéiques de *P. falciparum* sont préparés comme décrit par Merckx et al. (Merckx et al. 2003). Environ 100 µg de protéines sont incubées avec  $2 \times 10^7$  billes magnétiques activées par un groupement tosyl M-280 (Dynal Biotech) recouvertes d'anticorps anti-PfMyb1 (2,5 µg) ou de serum préimmun en suivant les recommandations du fabricant. Les protéines immunoprécipitées sont fractionnées sur un gel SDS-PAGE 10 % et analysées par western blot après transfert actif sur une membrane PVDF 0,2 µ (Biorad), avec une dilution 1/1000 de l'anticorps anti-PfMyb1 et révélées comme précédemment (paragraphe II.4).

## **II.6 – Localisation de protéines par immunofluorescence**

Les parasites (1 ml à 7-10 % de parasitémie) lavés deux fois dans le milieu de culture RPMI 1640 sont fixés avec 2 % de paraformaldéhyde pendant 20 min à 25 °C et 30 µl de suspension sont déposés sur des lames en verre recouvertes de poly-L-lysine (Sigma). Les cellules lavées cinq fois avec 50 µl de tampon phosphate PBS 1X (Invitrogen) sont traitées avec 75 mM de NH<sub>4</sub>Cl pendant 10 min pour l'extinction de l'autofluorescence,

perméabilisées avec 0,5 % de Triton X100 pendant 3 min puis incubées avec du tampon PBS + 0,5 % BSA pH 7,4 pour saturer les sites non spécifiques pendant 1 h. L'incubation avec une dilution 1/200 des anticorps primaires (anti-PfHMGB, anti-HSP70 et sera preimmun) dans le tampon de blocage est effectuée pendant 1 h à 25 °C, suivie de cinq lavages avec le tampon PBS et d'une incubation avec l'anticorps secondaire anti-IgG de souris couplé au FITC (Sigma) à une dilution de 1/500 pendant 30 min à l'obscurité. Après lavage, l'ADN nucléaire est marqué par le DAPI (1/100) pendant 10 min. Les lames sont recouvertes d'une solution de montage (90% glycerol contenant 0,1 M Tris-HCl, pH 8.0 et 2 mg/ml de *o*-phenylenediamine) avant d'être scellées sous lamelles et observées au microscope à fluorescence Leica DM RD équipés de filtres UV et fluoresceine. Le temps d'exposition est sélectionné en fonction de l'intensité de signal obtenu avec chaque anticorps et les images sont acquises avec le logiciel Lucia 4.7 et superposées avec Adobe Photoshop 7.0.

## **II.7 – Retard sur gel ou EMSA (Electrophoretic Mobility Shift Assay)**

### **II.7.1 – Expérience de liaison aux séquences MRE**

La méthode de retard sur gel a pour principe la mise en présence de la séquence cible d'ADN étudiée radiomarquée, sous forme d'oligonucléotide double brin, avec des extraits protéiques nucléaires. L'apparition d'une bande retardée après migration sur gel de polyacrylamide traduit la formation d'un complexe ADN/protéines et correspond à la mise en évidence de l'activité de liaison d'un facteur de transcription. La spécificité de liaison ADN/protéines est testée par compétition avec des séquences cibles froides en excès ou avec des anticorps.

Les globules rouges infectés par *P. falciparum* aux stades anneaux, trophozoïtes et schizontes (50 ml de culture à 10-12 % de parasitémie) sont lavés dans du tampon PBS froid et lysés par un traitement à la saponine 0,15 % pour récolter les parasites. Les extraits nucléaires de parasite sont préparés suivant le protocole établi par Osta et al. (Osta et al. 2002). Les protéines sont dialysées pendant 1 h à 4 °C à travers un filtre VS 0,025 µ (Millipore) posé sur 25 ml de tampon de dialyse (25 mM HEPES pH 7,5, 50 mM KCl, 1 mM MgCl<sub>2</sub>, 0,5 mM EDTA, 0,2 mM PMSF, 1 mM DTT, 10 % glycérol et 1/25 du cocktail d'inhibiteurs de protéases Complete de Roche) puis congelés dans l'azote liquide et conservés à -80 °C.

Les oligonucléotides complémentaires (10  $\mu$ l à 25  $\mu$ M) correspondant aux séquences MRE des promoteurs *pfcrk1*, *pfmap1* et *mim-1* sont incubés pendant 5 min à 95 °C dans 40  $\mu$ l de tampon d'hybridation (67 mM Tris pH 7,5, 13 mM MgCl<sub>2</sub>, 1,3 mM EDTA, 1,3 mM spermidine et 6,7 mM DTT), l'hybridation a lieu le temps que le bain-marie revienne à 25 °C. Les oligonucléotides double brin (0,5  $\mu$ l à 25  $\mu$ M) sont marqués avec 40  $\mu$ Ci de [ $\gamma$ -<sup>32</sup>P] ATP (< 5000 Ci/mmmole et 10 mCi/ml) en présence de T4 ADN kinase dans le tampon "kinase exchange" (Invitrogen) pendant 30 min à 37 °C et sont purifiés avec le kit "QIAquick nucleotide removal" (Qiagen).

Les extraits nucléaires (1  $\mu$ g) sont pré-incubés avec ou sans compétiteurs spécifiques *pfcrk1*, *pfmap1* et *mim-1* froids en excès molaire 5X, 20X et 200X pendant 30 min à 4 °C dans le tampon de liaison (25 mM HEPES pH 7,5, 50 mM KCl, 1 mM MgCl<sub>2</sub>, 0,5 mM EDTA, 0,2 mM PMSF, 1 mM DTT, 5 % glycérol et 2  $\mu$ g poly dI-dC). La compétition non spécifique est réalisée dans les mêmes conditions avec un excès molaire 200X de l'oligonucléotide NF- $\kappa$ B (voir paragraphe I.4). Le complexe ADN-protéines est analysé après incubation en présence de 0,5  $\mu$ l d'anticorps anti-PfMyb1 ou 0,5  $\mu$ l de serum preimmun ou 0,5  $\mu$ l d'un anticorps non pertinent anti-NF $\kappa$ B. L'EMSA est réalisé avec 50 000 cpm d'oligonucléotides marqués dans un volume final de 10  $\mu$ l pendant 30 min supplémentaire. Les échantillons sont déposés en conditions natives sur un gel de polyacrylamide 6 % et l'électrophorèse est effectuée à 160 V pendant 2 h à 4 °C dans un tampon TBE à faible force ionique (0,25X). Les gels séchés sont autoradiographés à -80 °C.

### II.7.2 – Expérience de liaison à l'ADN cruciforme

L'ADN cruciforme à 4 brins et les structures incomplètes à 2 et 3 brins sont obtenus à partir d'un mélange équimolaire des oligonucléotides appropriés dans un tampon d'hybridation composé de Tris-HCl 10 mM pH 7,2, NaCl 50 mM, MgCl<sub>2</sub> 10 mM. Pour obtenir le cruciforme marqué, le brin 1 (1  $\mu$ M) est radiomarké avec 125  $\mu$ Ci de [ $\gamma$ -<sup>32</sup>P]ATP (< 5000 Ci/mmmole et 10 mCi/ml), 20 unités de T4 ADN kinase (Invitrogen), le tampon kinase 5X dans un volume final de 20  $\mu$ l pendant 1 h à 37 °C. Le mélange avec les autres brins est porté à 100°C pendant 5 min et l'hybridation a lieu alors que la température du Bain-Marie revient à 25 °C. Le mélange réactionnel est déposé sur un gel de polyacrylamide 6,5 % dans un tampon TBE 0,5X. Après 2 h de migration à 120 V, le gel est autoradiographié pendant 15

min. Le cruciforme, mis en évidence après révélation et excisé du gel, est élué dans 1 ml de tampon d'hybridation.

Pour les expériences de retard sur gel, l'ADN cruciforme 4 brins marqué est incubé avec des concentrations croissantes (0 à 25  $\mu$ M) de la protéine PfHMGB étudiée pendant 20 min à 25 °C dans un tampon de liaison composé de Tris-HCl 10 mM pH 8, EDTA 1 mM, NaCl 25 mM, glycérol 2 %, 0,14  $\mu$ g d'ADN de sperme de saumon et 1,4  $\mu$ g de BSA. L'ADN cruciforme à 4 brins ou les structures incomplètes à 2 et 3 brins non marqués sont ajoutés à la réaction en excès molaire de 100X ou 500X pour les tests de compétition et une incubation supplémentaire de 20 min est effectuée. Les échantillons sont déposés sur un gel de polyacrylamide 6,5 % dans un tampon TBE 0,5X. Après 2 h de migration à 120 V, le gel est séché et autoradiographié pendant 5 jours à -80 °C.

### **II.7.3 – Expérience de courbure de l'ADN**

Des fragments d'ADN de 123 pb sont obtenus par digestion du marqueur de taille "123 pb ladder" (Invitrogen) par l'enzyme de restriction Ava I et sont marqués à l'extrémité 5' avec 40  $\mu$ Ci de [ $\gamma$ - $^{32}$ P]ATP comme précédemment. Après purification des fragments à l'aide du kit "Pure PCR Product Purification" (Qiagen), 10 ng d'ADN de 123 pb sont pré-incubés avec des concentrations croissantes (0,25 à 100  $\mu$ M) de la protéine PfHMGB étudiée pendant 20 min à 25°C en présence du tampon T4 ADN ligase 1X (Biolabs) dans un volume final de 10  $\mu$ l. L'ADN se circularise grâce à l'ajout de 40 unités de T4 ADN ligase (Biolabs) pendant 1 h à 25 °C. La réaction est stoppée par une incubation de 10 min à 65 °C. Pour éliminer les fragments linéaires, certains échantillons sont traités avec 30 unités d'exonucléase III (Biolabs) pendant 30 min à 37 °C, les échantillons non traités servent de témoins. Toutes les réactions sont incubées avec 2  $\mu$ g de protéinase K (Invitrogen) dans 1 % de SDS et 2 % de glycérol pendant 30 min à 37 °C avant d'être déposées sur un gel de polyacrylamide 6,5 % dans un tampon TBE 0,5 X. Après 2 h de migration à 120 V, le gel est séché et autoradiographié pendant 5 jours à -80 °C.

## **RESULTATS**



# I – ANALYSE COMPARATIVE DU TRANSCRIPTOME DE DEUX CLONES 3D7 ET F12 DE *P. FALCIPARUM*, LORS DU CYCLE ERYTHROCYTAIRE.

## I.1 - Résumé

La régulation des événements cruciaux pour le développement de *P. falciparum* suppose un contrôle fin de l'expression des gènes. De nombreux facteurs environnementaux semblent influencer la transition de la prolifération asexuée à la différenciation sexuée mais le déclenchement et les mécanismes impliqués dans la gamétocytogénèse restent largement inconnus. Néanmoins, il est admis que l'engagement vers le développement des gamétocytes est établi au stade asexué schizonte. Le clone de *P. falciparum* F12, dérivé de 3D7, incapable de produire des gamétocytes a été caractérisé par une perte de l'expression de transcrits marqueurs de la différenciation sexuée. Nous avons donc entrepris l'étude comparative du transcriptome de ces deux clones au cours du cycle érythrocytaire afin de découvrir des gènes dont le profil d'expression serait différent en fonction de la potentialité à produire des gamétocytes.

L'implication des facteurs de transcription et des protéines de transduction du signal dans l'initiation de la gamétocytogénèse ayant été suggéré, nous avons construit et produit une puce à ADN thématique au sein de notre laboratoire. Celle-ci contient environ 200 fragments de PCR dont 153 sont représentatifs de gènes de *P. falciparum* potentiellement impliqués dans la régulation du cycle cellulaire et la transcription.

Cette étude nous a permis de répertorier d'une part, 106 gènes dont les profils d'expression sont similaires entre les deux clones et de confirmer ainsi la proximité génétique des clones 3D7 et F12. Parmi ces gènes, figurent une quinzaine de FT dont l'expression est modulée au cours du développement, avec cependant des profils d'expression différents (Fig. 3, p. 273 de l'article 1). Certains d'entre eux présentent un profil d'expression conservé avec un autre clone HB3 sans proximité génétique et aurait alors un rôle lors du cycle asexué érythrocytaire et non pas dans l'incapacité à produire des gamétocytes du clone F12. D'autre part, plusieurs gènes ont présenté des profils d'expression différents entre les deux clones 3D7 et F12, notamment des gènes de kinases et de facteurs intervenant possiblement dans la régulation transcriptionnelle (Fig.2, p. 272 de l'article 1). En particulier, un gène codant pour

un FT de la famille des protéines Constans et un gène codant pour une protéine comportant un motif NAP, impliqué dans le remodelage de la chromatine ont retenu notre attention.

**ARTICLE 1:**

**Transcriptome of 3D7 and its gametocyte-less derivative F12 Plasmodium falciparum clones during erythrocytic development using a gene-specific microarray assigned to gene regulation, cell cycle and transcription factors.**

Mathieu Gissot, Philippe Refour, Sylvie Briquet, Charlotte Boschet, Stéphane Coupé, Dominique Mazier et Catherine Vaquero.

Accepté à la publication dans *Gene*.

# Transcriptome of 3D7 and its gametocyte-less derivative F12 *Plasmodium falciparum* clones during erythrocytic development using a gene-specific microarray assigned to gene regulation, cell cycle and transcription factors

Mathieu Gissot, Philippe Refour, Sylvie Briquet, Charlotte Boschet, Stéphane Coupé<sup>1</sup>,  
Dominique Mazier, Catherine Vaquero\*

INSERM U511, CHU Pitié-Salpêtrière, Université Paris 6, 91 boulevard de l'Hôpital, 75013 Paris, France

Received 28 March 2004; received in revised form 21 June 2004; accepted 5 July 2004

Available online 13 September 2004

## Abstract

During the complex life cycle of *Plasmodium falciparum*, through mosquito and human, the erythrocytic cycle is responsible for malarial disease and transmission. The regulation of events that occur during parasite development, such as proliferation and differentiation, implies a fine control of transcriptional activities that in turn governs the expression profiles of sets of genes. Pathways that underline gametocyte commitment are yet poorly understood even though kinases and transcription factors have been assumed to play a crucial role in this event. In order to understand the molecular mechanisms controlling the variation of gene expression profiles that might participate in early gametocytogenesis, the transcriptome of two clones, 3D7 and its gametocyte-less derivative F12, was compared at five time points of the erythrocytic asexual development. We have used a thematic DNA microarray containing 150 PCR fragments, representative of *P. falciparum* genes involved in signal transduction, cell cycle and transcriptional regulation. We identified several genes eliciting different expression profiles among which some implicated in gene regulation or encoding putative transcription factors. The differential expression of transcription factor and kinase transcripts observed in the two clones may enlighten genes that might have a role in impairment of the early gametocytogenesis of the F12 clone.

© 2004 Elsevier B.V. All rights reserved.

**Keywords:** *Plasmodium*; Transcription factor; Microarray; Erythrocytic development; Gametocytogenesis

## 1. Introduction

*Plasmodium* is responsible for 1.5–2.7 million deaths annually, mostly children and pregnant women. Among the

four species of parasites infecting humans, *Plasmodium falciparum* causes the highest morbidity and mortality. Global efforts to eradicate malaria have failed and no effective vaccine is available. Therefore, insight into the parasite biology and mechanisms of gene and cell cycle regulation during the development of *Plasmodium* is urgently required to provide new targets for the control of malaria. The cell cycle of *Plasmodium* between vertebrate host and mosquito passes by a succession of different stages characterized by an active cellular division (as the exoerythrocytic and erythrocytic schizogony) and an arrest of proliferation associated to differentiation (as for gametocytes and sporozoites). Actually, two major events of the *P. falciparum* life cycle occur in the erythrocyte, an asexual multiplication responsible for the clinical manifestations and the initiation of sexual differentiation responsible for dissemination of the disease via the mosquito. Those two

**Abbreviations:** aa, amino acid; CO, Constans; ES, early schizont; ET, early trophozoite; LS, late schizont; LT, late trophozoite; NAP, nucleosome assembly protein; NLS, nuclear localization site; ORF, open reading frame; PCR, polymerase chain reaction; qPCR, real time quantitative PCR; R, ring; RT, reverse transcription; TF, transcription factor; dNTP, deoxyribonucleoside triphosphate; DNase, deoxyribonuclease; RNase, ribonuclease; UTR, untranslated region(s); cDNA, DNA complementary to RNA; IFN, interferon; SDS, sodium dodecyl sulfate; SSC, 0.15 M NaCl/0.015 M Na<sub>3</sub> citrate pH 7.6.

\* Corresponding author. Tel.: +33 1 40778111; fax: +33 1 45838858.

E-mail address: [vaquero@ext.jussieu.fr](mailto:vaquero@ext.jussieu.fr) (C. Vaquero).

<sup>1</sup> Present address: Laboratoire de Parasitologie-Mycologie, Centre Hospitalier Universitaire Lariboisière-Saint-Louis, 75010 Paris, France.

events are under the control of coordinated expression of distinct sets of genes governed by transcriptional regulation even though post-transcriptional events could also be implicated.

Pathways that control asexual cycle regulation and switch to sexual commitment are yet poorly understood although a number of genes described to command cell cycle in other eukaryotes, e.g. kinases and transcription factors, have been predicted in our laboratory or by the sequencing project of *P. falciparum* (Gardner et al., 2002). Two publications have described the whole genome examination of the transcriptome of HB3 (Bozdech et al., 2003) and 3D7 (Le Roch et al., 2003) clones of *P. falciparum* giving the first global idea of the expression profiles of almost all genes during erythrocytic asexual cycle and at gametocyte stage, therefore opening new perspectives for the discovery of gene implication in those pathways. Various genes were reported to be associated with induction of gametocytogenesis, as *pfg27* and *pfs16* (Dyer and Day, 2000b). Moreover, implication of cyclic AMP-dependent pathway (Inselburg, 1983) and trimeric G proteins (Dyer and Day, 2000a) suggests participation of kinases and transcription factors in this phenomenon. Sexual commitment have been reported to be sensitive to many environmental conditions including host immunity, anti-malarial drugs, host hormones, erythrocyte intracellular environment and autocrine growth control (reviewed in Dyer and Day, 2000b). The parasite response to environment that influences the developmental decision, point out the key role of signaling pathways comprising cell cycle regulators, kinases and transcription factors leading to transcriptional control of sexual commitment. Clones of *P. falciparum* lacking or with a reduced ability to produce gametocytes have been reported. Although those alterations in sexual development have been considered to be irreversible, drug-driven induction of gametocytogenesis was observed in clones thought to be impaired in sexual development (Ono and Nakabayashi, 1990) suggesting a reduced sensitivity to environmental changes. Subtelomeric deletion of chromosome 9 also induces almost total loss of gametocyte production (Day et al., 1993). However, total or partial failure of the ability to produce gametocytes has been shown to occur in clones, derived from 3D7, carrying an intact chromosome 9 and no visible changes in karyotype (Alano et al., 1995). F12 clone has been reported to be one of the gametocyte-less 3D7 derivative. In addition, loss and decrease of mRNA expression of the two earliest landmarks of gametocyte differentiation, the cytoplasmic protein *Pfg27* and the membrane protein *Pfs16*, respectively, have been stated (Alano et al., 1995). However, the role of *Pfg27* in sexual differentiation remains unclear. Its expression have been described to be developmentally regulated (Alano et al., 1996), required for early stage of gametocytogenesis and reported to play a role in a part of *Pfs16* expression induction (Lobo et al., 1994, 1999). Moreover, crystal structure of *Pfg27* raised the hypothesis of potential binding

to other proteins such as kinases through SH3 domains and to RNA (Sharma et al., 2003). Nature of the impairment of *pfg27* transcript expression could be due to absence or altered-expression of genes implicated in the control of its expression, since major rearrangement upstream to the *pfg27* coding sequence was not observed in F12 clone (Alano et al., 1996). The absence of expression of *pfg27* transcript and therefore loss of gametocyte production ability remain unclear in F12 clone; whether it has lost the sensitivity to external stimuli or the ability to respond to these stimuli also remains to be elucidated.

In the investigation described here, we constructed a gene-specific microarray encompassing 153 *P. falciparum* PCR products representing genes that might be implicated in regulation of cell cycle and transcription. We hybridized cDNA, obtained from total RNA prepared from five time points of asexual erythrocytic culture of the two clones, 3D7 and F12, and labeled with two distinct radioisotopes. We compared expression profiles of the 153 genes in these two clones grown under normal culture condition. Indeed, we tried to identify genes with differential expression profiles in the F12 clone that does not produce gametocytes and in 3D7 that bear the potentiality to produce gametocytes. This analysis was carried out to underline differentially expressed genes particularly focusing on putative transcription factors. Differential expression of transcription factors and kinases might have a role in the impairment in early stage of gametocytogenesis.

## 2. Materials and methods

### 2.1. *P. falciparum* clones and cultures

*P. falciparum* 3D7 and F12 (Alano et al., 1995) were provided by Dr. D. Walliker and Dr. P. Alano, respectively, and were cultured as described in Trager and Jensen (1976) with slight modifications. To obtain synchronized erythrocytic stages, ring cultures were enriched by three successive treatments of 5% sorbitol at 44-h intervals. Five stages of development, rings (10–14 h), early (20–22 h) and late (26–28 h) trophozoites, and early (32–34 h) and late (40–44 h post-invasion) schizonts, were determined according to time and morphological analysis by GIEMSA staining. At each indicated stage, parasite culture at 8% parasitemia was collected by centrifugation and used for preparation of transcripts for further analyses.

### 2.2. Amplification of PCR products and microarray construction

Microarrays were prepared by printing 170 PCR products 300–1500 bp long produced from sequences cloned in the Topo 2.1 vector (Invitrogen), representing mostly genes implicated in signal transduction and cell cycle of *P. falciparum*. PCR fragments from all clones (Eurogentec)

were purified by isopropanol precipitation and resuspended in spotting buffer (Tris 10 mM pH 7.4, EDTA 1 mM, DMSO 50%) adjusted to an average concentration of 200 ng/ $\mu$ l and produced with specific pairs of primers found in the vector; T7 (5' TAATACGACTCACTATAGGG 3') and M13R (5' GGATAACAATTTACACAGG 3').

Seventeen spots, representing positive and negative controls such as sequences unrelated to *P. falciparum*, were printed. Alien PCR products from Stratagene, with no apparent cross-similarity with *P. falciparum* genomic sequence, were used for normalization of the reverse transcription. 3D7 genomic DNA was also printed in order to normalize the total amount of each cDNA present during the hybridization. All targets were spotted in duplicate with a Qarray arrayer (Genetix) onto poly-L-lysine-coated slides prepared as described previously (Schena et al., 1995) with a spot spacing of 400  $\mu$ m, center to center onto a 1 cm<sup>2</sup> area. DNA was cross-linked by UV irradiation at 3 kJ with Stratalinker model 1800 UV Illuminator (Stratagene). Neutralization of free poly-L-lysine was performed by chemical treatment with succinic anhydride (Sigma). Prior to hybridization, the slides were incubated for 1 min at 95 °C to denature the spotted DNA and dehydrated with a 96% ethanol immersion for 1 min and dried by centrifugation.

### 2.3. Probe labeling, hybridization and data analysis

Parasitized red blood cells were rapidly lysed in TRIZOL (Invitrogen) and total RNA was extracted at different times during erythrocytic development from 3D7 and F12 *P. falciparum* cultures. Concentration was determined by spectrophotometer and integrity of the RNA preparations was verified by ethidium bromide staining on agarose gel and by Agilent 2100 bioanalyser.

Ten micrograms of total RNA extracted at different times of erythrocytic development were labeled during reverse transcription reaction in a sample mixture containing 2  $\mu$ g of 15-mer poly-T primer (Invitrogen), 5 mM of MgCl<sub>2</sub>, 10 mM of DTT, 80 U of RNAout (Invitrogen), 0.2 mM of each dCTP, dGTP and dTTP, 50  $\mu$ Ci/ $\mu$ l of <sup>35</sup>S-dATP (Amersham, catalog number SJ1134) or 40  $\mu$ Ci/ $\mu$ l of <sup>3</sup>H dATP (Amersham, catalog number TRK633). Reaction was performed at 42 °C for 2 h with 250 U of Superscript II reverse transcriptase and subjected to ribonuclease (RNase)-H treatment with 2 U of enzyme for 20 min at 37 °C. Labeled cDNA were purified on QIAquick nucleotide removal columns as indicated by the manufacturer (Qiagen) and eluted in 40  $\mu$ l. Amount of radioactivity contained in 1  $\mu$ l of purified eluate was evaluated using a Beckman Coulter LS 6500 apparatus.

For hybridization, the labeled probe mixture corresponding to  $1.6 \times 10^6$  disintegrations per minute for each cDNA was used per slide. Each labeled cDNA was added to the hybridization buffer (3.5 $\times$  SSC, 0.3% SDS, 0.5  $\mu$ g/ $\mu$ l DNA Salmon Sperm and 0.5  $\mu$ g/ $\mu$ l yeast tRNA), heated to 95 °C for 2 min, cooled down at room temperature and added on

the microarray slide under a cover slip (Easy seal, Hybaid). Hybridization was performed in a cassette chamber (Tele-Chem) submerged in a water bath at 62 °C for 14–16 h. Arrays were washed at room temperature in a 2 $\times$  SSC, 0.1% SDS solution prior to washing in a 2 $\times$  SSC and in a 0.2 $\times$  SSC solution successively, each step for 2 min, and finally dried by centrifugation.

Acquisition of arrays was performed as described previously on a Microimager (Biospace Measure) (Salin et al., 2002), and stopped after 24 h acquisition. Data were collected as an image file and the ARRAYVISION software (Imaging Research) was used for quantification of the hybridization intensities and for normalization. The local background was subtracted from each spot intensity. The intensity of each spot was normalized according to the median value of intensities of spots corresponding to Alien cDNA PCR products on each array in order to assess normalization of the reverse transcription. Furthermore, genomic DNA was used to normalize the amount of cDNA present in the hybridization. For all genes, expression profile was expressed as the log<sub>2</sub> ratio of labeling intensity of experiment over reference. Two independent experiments and one technical replicate were performed.

### 2.4. Genomic DNA labeling

Genomic DNA (30 ng) isolated from F12 and 3D7 clones were <sup>35</sup>S-labeled using a random-primed DNA labeling kit (Amersham). Labeled genomic DNAs were purified using QIAquick columns (Qiagen) and hybridized as described above.

### 2.5. Real time quantitative RT-PCR (qPCR)

Primers were designed using PrimerExpress software (Applied Biosystem). Total RNA (10  $\mu$ g) was treated with 10 U of RNase-free deoxyribonuclease (DNase) I (Qiagen), and cDNA synthesis was performed for 50 min at 42 °C using 1  $\mu$ g of RNA, 50 U of MMLV reverse transcriptase (Superscript II, Invitrogen), 0.5 ng of random hexamers, 5 mM of MgCl<sub>2</sub>, 20 U of RNaseout (Invitrogen), 10 mM of DTT and 0.25 mM of each dNTP. Finally, reaction was subjected to RNase-H treatment with 2 U of enzyme for 20 min at 37 °C. Real time quantitative PCR reaction was performed in a 20  $\mu$ l reaction volume containing 5  $\mu$ l of a dilution of cDNA preparation, 10  $\mu$ l of SYBR green PCR master mix (Sigma) and 0.375  $\mu$ M of gene-specific primers (Invitrogen). Amplification and detection of specific products were performed with the MX4000 light cycler (Stratagene) with the following cycle profile: one cycle at 95 °C for 2 min, 40 cycles with 30 s denaturation at 95 °C and 1 min annealing-elongation at 60 °C. Size of each PCR product and number of bands was verified on a 10% acrylamide gel. Two additional reactions were performed, either without reverse transcriptase or without the RNA sample, to verify the absence of DNA contamination. The

quantity of cDNA for each experimental gene was normalized to the 18S concentration (chr5.rRNA-1–18s-A) in each sample. For each time point of each gene, experiments were performed twice and in triplicates. Relative gene expression was expressed via the  $\log_2$  of the ratio of the expression time point over ring stage using the  $2_T^{-\Delta\Delta C}$  method (Livak and Schmittgen, 2001).

### 3. Results

#### 3.1. *P. falciparum* gene-specific microarray, experimental design and validation of the results

We constructed a microarray encompassing 170 PCR products among which 153 amplified from the *P. falciparum* genomic sequences. This thematic microarray could evolve by adding new genes of interest when needed. All sequences to be amplified were localized within the open reading frame (ORF) close to the 3' end and were sequenced prior to be arrayed on glass slides.

The *Plasmodium* genes were selected for their homology to known proteins previously described in eukaryotes to be involved in genetic regulation from DNA to proteins. Indeed, glass slides were printed with PCR fragments coming from genes encoding putative kinases controlling signal transduction and cell cycle, proteins involved in replication (DNA polymerases, DNA topoisomerases and helicases), proteins involved in general transcription machinery (TATA binding protein, small nuclear ribonucleoprotein, RNA polymerases), transcription factors (TFs) (such as three members of the Myb family, two members of the HMG family and several proteins encompassing zinc finger domains), as well as proteins involved in the stability and the structure of mRNA (ribosomal RNA methylases, helicases), in translation (initiation and elongation factors), and in chromatin remodeling (histone deacetylase, nucleosome assembly protein). Several of these proteins have already been described in the literature, others annotated in databases such as PlasmoDB ([www.plasmodb.org](http://www.plasmodb.org)) or annotated by our group among which putative transcription factors and kinases. For DNA printing, several negative controls were used among which *pflsa1* only expressed during exo-erythrocytic cycle, unrelated genes as human gamma IFN and HIV-1 gp120 envelope protein, empty plasmid vector used to clone each PCR product and mouse or human genomic DNA. Several *Plasmodium* genes reported to be expressed during asexual erythrocytic cycle, among which *hsp70* (PF08\_0054), *H2A* (MAL6P1.249) and *H3* (MAL6P1.248), *kahrp* (PFB0100c), *pcna1* (PF13\_0328) and *gbp130* (PF10\_0159), were spotted as positive controls at different positions on the microarray. Finally, PCR products corresponding to the Alien (Stratagene) synthetic RNAs, added into total RNA samples to normalize the effectiveness of the cDNA probe production, as well as two different *Plasmodium* genomic DNA

concentrations, used to normalize the total amount of each cDNA present in the hybridization, were also printed.

We performed radioactive labeling of the cDNA reverse transcribed from all the total RNA samples, in order to minimize background signal, enhance signal detection of low abundant mRNA species and avoid signal saturation. This experimental approach was also further adapted to very low concentrations of biological material (Refour et al., submitted for publication). A CCD camera determined, in a real time manner, the signals obtained after hybridization of the microarray with the two differentially  $^{35}\text{S}$  or  $^3\text{H}$  labeled cDNA. For each *P. falciparum* clone, two independent technical replicates as well as two different biological RNA preparations from independent culture were used for cDNA synthesis. A representative picture of co-hybridization data obtained with the cDNA labeled with the two radioisotopes is presented in Fig. 1. Negative controls showed no signal as for *pflsa1* (PF10\_0356) (a), human genomic DNA (b), human gamma IFN (c) and HIV-1 gp120 protein (d), in contrast to positive controls, genes expressed during asexual erythrocytic cycle (1–3). In addition, three PCR products along the sequence of PF14\_0316 ORF were spotted to control the PCR product sequence effect on the expression profile (4–6). Moreover, the same PCR product of PF11\_0098 ORF (7–8) was spotted twice to control the localization effect and all gave similar signals. Duplicate spots appeared to give similar signal levels. In addition, self-hybridization between ring stage cDNA labeled with  $^{35}\text{S}$  and ring stage cDNA labeled with  $^3\text{H}$  gave ratio very close to one (data not shown).

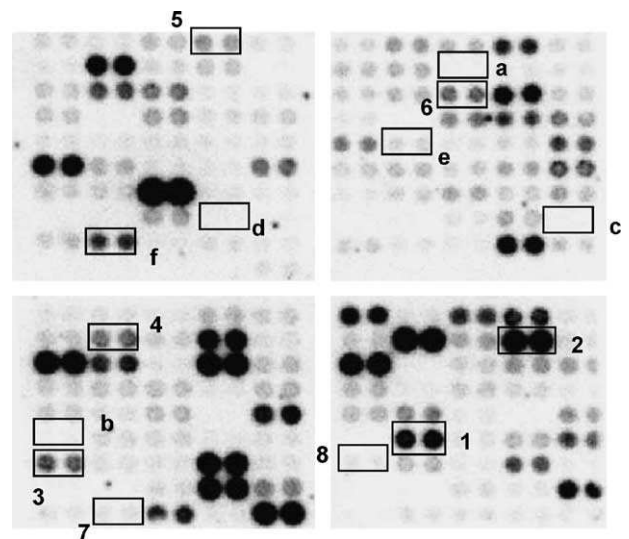


Fig. 1. A representative picture of a co-hybridization of two cDNA labeled in the presence of  $^{35}\text{S}$  dATP and  $^3\text{H}$  dATP. Framed duplicated spots represent (a) *pflsa1* (PF10\_0356), (b) human genomic DNA (20 ng/ $\mu\text{l}$ ), (c) human gamma IFN, (d) HIV gp120 protein, (e) Alien 1 from Stratagene, (f) Alien 2 from Stratagene; (1) *pflhsp70* (PF08\_0054), (2) *histone H2A* (MAL6P1.249), (3) *gbp130* (PF10\_0159); (4, 5, 6) three different PCR products for PF11\_0098; (7, 8) the same PCR fragment for PF14\_0316 spotted twice.

When performing the first time course experiment for the 3D7 clone, we selected five genes as positive controls for which expression has been previously described. We compared throughout development the peaks of expression obtained either by our microarray or qPCR to published data (Table 1). The different expression profiles were equivalent and in good agreement with the published data, therefore validating the quality of our experimental approach.

Finally, prior to the first microarray differential experiments, genomic DNA from F12 clone were labeled and hybridized to the microarray in order to verify that the level of homology between F12 and 3D7 sequences was sufficient to determine and compare the abundance of all messengers of each clone. Actually, all immobilized PCR products amplified from 3D7 clones hybridized to labeled genomic DNA extracted from F12 (data not shown). In addition, we verified by real time quantitative RT-PCR that total RNA prepared from the F12 different stages was devoid of *pfg27* transcript (data not shown).

### 3.2. Comparison between 3D7 and F12 expression profiles

The expression profiles of 3D7 clone and its gametocyte-less derivative F12 clone were monitored during the erythrocytic cycle. Highly synchronized cultures were harvested at five time points during the asexual cycle. Five stages of development, rings (10–14 h), early (20–22 h) and late (26–28 h) trophozoites, and early (32–34 h) and late (40–44 h post-invasion) schizonts, were assessed according to time and morphological analysis. The cDNA obtained from ring total RNA of each *Plasmodium* clone was selected as the reference RNA, labeled with <sup>35</sup>S dATP and co-hybridized with the cDNA labeled with <sup>3</sup>H dATP

prepared from total RNA extracted at each indicated time of development. The five time points are rings (R), early (ET) and late trophozoites (LT), and early (ES) and late schizonts (LS). For all genes, the steady state RNA profile was expressed as the log<sub>2</sub> ratio of labeling intensity of experiment (R, ET, LT, ES, or LS) over reference (R). In addition, signal intensity was normalized twice, first with the Alien cDNA to normalize the RT efficacy and second with the genomic *Plasmodium* DNA in order to normalize the amount of labeled cDNA loaded during the hybridization process.

Among the 153 *Plasmodium* genes spotted on the array, 23 genes including the seven negative controls, reported to be expressed only during exo-erythrocytic cycle, gave no detectable signal. Among the remaining 130 genes, 16 exhibited too low expression intensity and therefore no reproducible results for the three replicates assayed. When the 114 expression profiles obtained for 3D7 clone were compared to the two recently published transcriptome data for HB3 (Bozdech et al., 2003) and 3D7 clones (Le Roch et al., 2003), more than 80% (82,5%) of our expression results (see Supplementary data 1) were similar to at least one of these published data.

We compared the expression profiles of the 114 genes expressed during the erythrocytic cycle of the two clones and found that 106 genes shared similar profiles, consistent with the fact that F12 clone derived from 3D7 (Alano et al., 1995) (see Supplementary data 2). However, eight genes displayed differential expression and those results were confirmed by real time quantitative PCR experiments. These genes were clustered in three groups highlighting the differences between the two clones (Fig. 2). First, two kinases genes (*MAL6P1.146* and *PF13\_0258*) with nearly no modification in expression throughout 3D7 cycle in contrast to what was observed during F12 development with a maximal expression at LT for *MAL6P1.146* and ES for *PF13\_0258* (Fig. 2A). Second, two genes (*PFC0485w* and *PFB0865w*) presented opposite profiles. The *PFC0485w* gene, a putative Ser/Thr kinase, was found to have a maximal expression during ring stage and a minimal expression during schizont stage in 3D7 clone in contrast to the expression profiles of F12 clone. The expression of *PFB0865w*, a putative small nuclear ribonucleoprotein, peaked in 3D7 clone in LT and was always higher than the level observed in rings, whereas mRNA relative amount decreased from ring to schizont in F12 (Fig. 2B).

Third, the maximal peak of expression of genes represented in Fig. 2C was reached for 3D7 at different stages of development forwarded or delayed when compared to the F12 clone. *PF10930c*, a nucleosome assembly protein and *PFC0805w*, a putative DNA-directed RNA polymerase II largest subunit, showed a maximal peak at ET stage in 3D7 clone whereas the maximal peak was observed later in F12 cycle. On the contrary, in the case of *PF07\_0073*, a seryl-tRNA synthetase and *PF14\_0383*, a Constans-like putative transcription factor, the maximal

Table 1  
Comparison of the stage where maximal expression was observed for five control genes analyzed in 3D7 clone by microarray experiments, real time quantitative PCR or published data

Gene	Technical procedure		
	Microarray <sup>a</sup>	qPCR <sup>b</sup>	References
Histone H2A MAL6P1.249	S <sup>c</sup>	S	S (Lobo and Kumar, 1999)
KAHRP PFB0100c	R	R	R (Lanzer et al., 1992a)
GBP130 PF10_0159	LT	LT	LT (Lanzer et al., 1992b)
Pfs 40 <sup>d</sup> PF11_0098	S	S	S (La Greca et al., 1997)
PCNA1 PF13_0328	ES	ES	ES (Horrocks et al., 1996)

<sup>a</sup> The stage of maximal expression was determined from the microarray data for the 3D7 clone.

<sup>b</sup> By real time quantitative RT-PCR experiments.

<sup>c</sup> S stands for schizonts, R for rings, LT for late trophozoite and ES for early schizonts.

<sup>d</sup> Results are given for the three different PCR products as stated in Materials and methods.



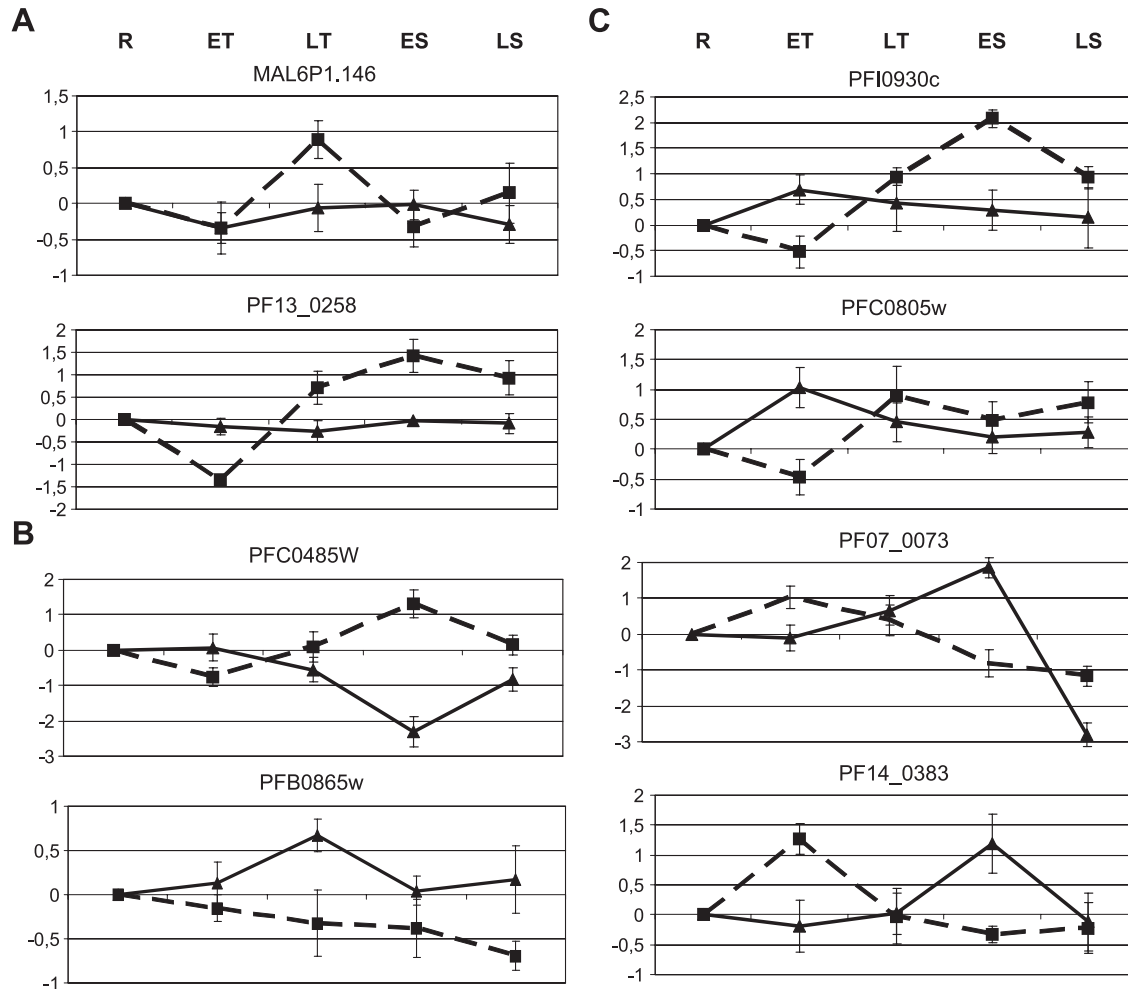


Fig. 2. Graphic representation of the expression profile of eight genes showing differential expression between the two clones 3D7 and F12. For each gene, the steady state RNA profile was expressed as  $\log_2$  ratio of labeling intensity of experiment (any time point) over reference (ring). 3D7 expression profile is represented by a continuous line with black triangles and F12 expression profile by a dash line with black squares. R stands for rings, ET for early trophozoites, LT for late trophozoites ES for early schizonts and LS for late schizonts, stated at the top of the figure. For each clone, the average of three experiments are shown as well as standard deviation. The genes have been clustered in three groups according to their expression profiles: (A) two kinases MAL6P1.146 and PF13\_0258, (B) PFB0865w, a putative small nuclear ribonucleoprotein, and PFC0485w, a putative serine/threonine kinase, and, (C) PFC0805w, a putative DNA-directed RNA polymerase II largest subunit, PFI0930c, a nucleosome assembly protein, PF14\_0383, a Constans-like putative transcription factor, and PF07\_0073, a seryl-tRNA synthetase.

peak was reached in the ES in 3D7 and in ET for the F12 clone (Fig. 2C).

### 3.3. Comparison of the transcriptional machinery proteins in 3D7 and F12 clones

We particularly focused our attention on the putative transcription factors and proteins of the transcriptional machinery, some previously annotated by our group, among which several are under molecular and functional investigation.

All 15 transcription factors were developmentally regulated during erythrocytic asexual cycle of the two clones with essentially similar patterns of expression. In Fig. 3, we selected five putative TFs presenting different expression patterns throughout erythrocytic development however similar in 3D7 and F12 clones by microarray experiment

(Fig. 3, panel A) and confirmed the data by real time quantitative PCR experiment (Fig. 3, panel B). *Pfhmg2* (PFL0290w) was expressed at a similar level throughout the asexual cycle with slight variations when compared to ring stage. In contrast, the relative amount of *pfphD2* declined strongly more than four times from R to ES re-increasing thereafter. For the last three transcripts, maximal expression was observed in ET for *pfmyb3* (PF10\_0143), in ES for *pfphDB* (PFL1905w) and LS for *pfkrox* (MAL13P1.76). All the messengers coding for these TF were developmentally regulated peaking at different times during the erythrocytic asexual cycle.

However, as already stated in Fig. 2, several messengers encoding various proteins of the transcription machinery exhibited differential expression profiles when comparing the 3D7 to F12 clone. Those transcripts encode a putative nuclear ribonucleoprotein (PFB0865w), a DNA-directed

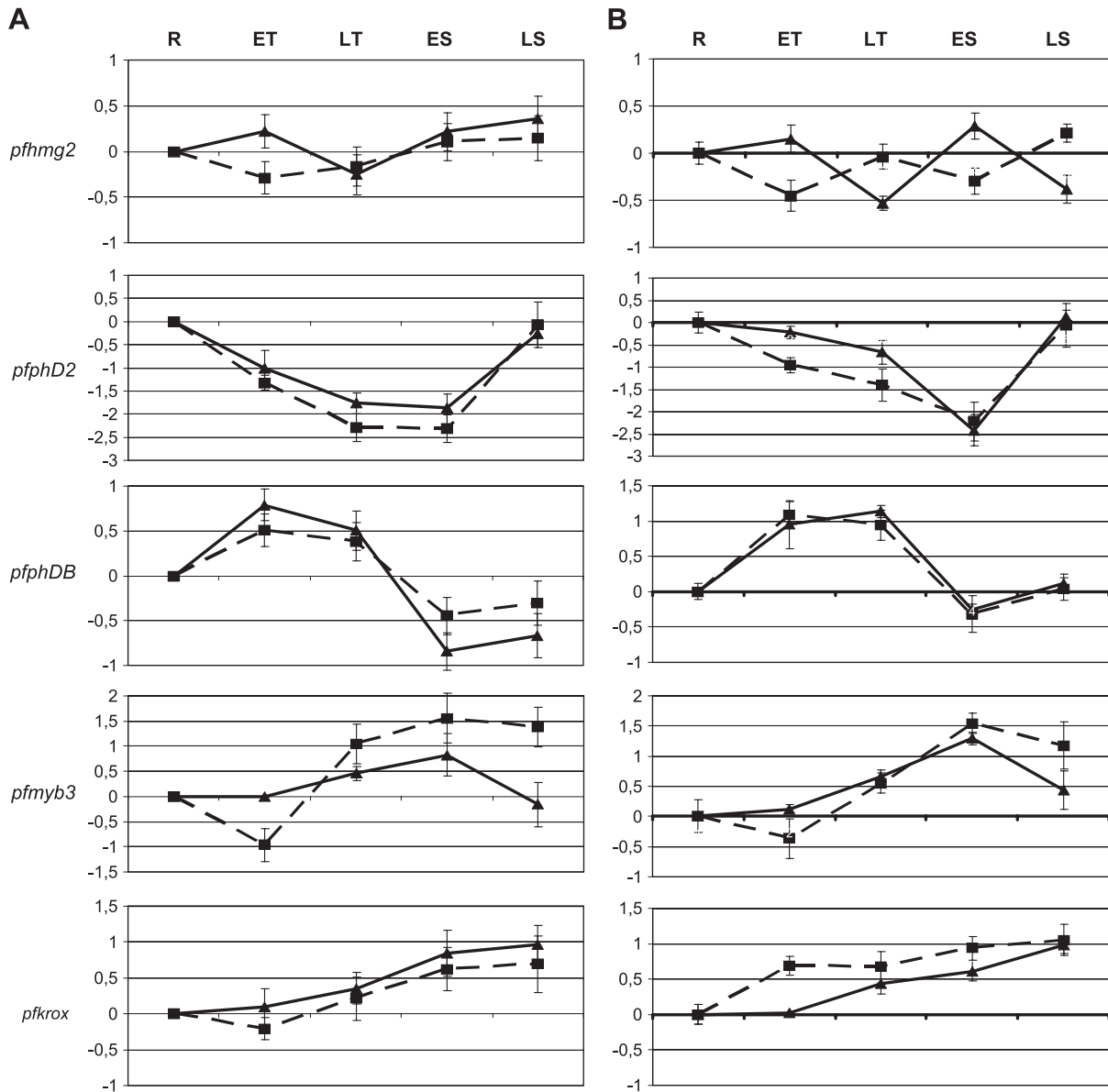


Fig. 3. Graphic representation of the expression profile of five putative transcription factors in the two clones 3D7 and F12 monitored via microarray and qPCR. For each gene, expression profile was expressed and represented as mentioned in Fig 2. (A) Expression profile obtained from microarray experiments. (B) Expression profile derived from real time quantitative PCR experiments.

RNA polymerase II (PFC0805w), a nucleosome assembly protein (NAP) (PFI0930c) and a Constans-like transcription factor (PF14\_0383). These messengers and proteins should be analysed to determine if their differential expression could participate in the transcriptional regulation of varied sets of genes during the asexual parasite cycle.

#### 4. Discussion

In order to determine differential transcript expression during erythrocytic development of *P. falciparum* clones 3D7 and its gametocyte-less derivative F12, we have used a gene-specific microarray printed with 153 PCR products

representing genes potentially implicated in cell cycle and transcriptional regulation. F12 clone has been isolated as a gametocyte non-producer from a long-term culture of the 3D7 clone. In contrast to 3D7 clone, F12 is impaired for early stages of gametocytogenesis without apparent difference in chromosomes number or size (Alano et al., 1995). Due to the genomic proximity of these two clones, one might expect that differential gene expression during the erythrocytic asexual cycle would participate in the early stage of gametocytogenesis including sensitivity to environment stimuli known to induce differentiation into gametocyte. Therefore, we designed a gene-specific microarray to study and compare temporal expression of roughly 150 messengers, during erythrocytic asexual development, of

highly synchronized cultures to identify some genes that might play a role in the parasite differentiation.

Before performing the time course transcriptome experiments we verified the clone properties, the quality of synchronization procedure and the feasibility of using F12 derived cDNA with a microarray encompassing 3D7 derived PCR product. First, we monitored expression of *pfgr27* transcript by real time quantitative RT-PCR and verified that no expression of the transcript could be detected in F12 in contrast to 3D7. Second, we carefully synchronized the parasite cultures, by three successive treatments with sorbitol at 44-h intervals. Subsequently, the five stages of erythrocytic asexual development were harvested with a minimum of 90% enrichment at each stage. Finally, we verified that all the PCR fragments obtained after amplification of the 3D7 genomic sequence and printed on the array could cross-hybridize with the genetic material prepared from F12 clone.

The radioactive labeled cDNAs were prepared from total RNA extracted at the indicated times of asexual life cycle. Formerly, the quality of the RNAs was controlled by Agilent analyzer and by fractionation on agarose gel. We performed radioactive-labeling of the cDNA probes, in order to minimize the background signal, avoid signal saturation and enhance signal detection especially for transcription factor mRNA that are believed to have a low abundance level. At the level of the microarray experimental protocol, we controlled the novel radioactive labeling and hybridization procedures carried out with the two probes (Salin et al., 2002; Refour et al., submitted for publication). Indeed, this was necessary to certify the reliability and reproducibility of this novel experimental approach.

Fig. 1 presents a representative picture corresponding to a co-hybridization with two cDNA labeled with  $^{35}\text{S}$  dATP and  $^3\text{H}$  dATP. Printing and hybridization were verified by the (i) appropriate responses given by the positive and negative controls, (ii) quality of the duplicates, and (iii) similarity of the signals given by the same PCR product (PF11\_0098) spotted twice at different places in the array as well as that given by three different PCR products along the PF14\_0316 ORF. In addition, each raw signal was normalized twice, first with the Alien cDNAs and second with the genomic *Plasmodium* DNA. Finally, some of the microarray data were verified using qPCR and compared with the results reported in the literature (Table 1).

The comparison of the steady state profiles of mRNA expression between the two clones, 3D7 and F12, led us to list 106 genes with essentially similar profiles and only eight genes with differential expression (Fig. 2). Additionally, we focused our attention on all the transcripts probably implicated in the transcriptional machinery and in particular those coding for transcription factors. Indeed, among the transcripts of the 15 putative transcription factors, most of them gave similar profiles throughout erythrocytic development of 3D7 and F12 clones, as shown for five of them verified by quantitative real time PCR (Fig. 3A and B). This

is not surprising since F12 was derived from 3D7. In addition, the conserved transcriptional profiles of several of these putative transcription regulators during erythrocytic cycle in 3D7, F12 and HB3 clones (Bozdech et al., 2003), favour their significant role for asexual growth and not for the impairment of sexual differentiation observed in the F12 clone.

All genes and in particular those encoding kinases and proteins of the transcriptional machinery displaying different expression profiles, were compared to results reported in the Scripps/GNF Malaria Array data (Le Roch et al., 2003) and the De Risi *P. falciparum* HB3 time course microarray data (Bozdech et al., 2003). Most of our data obtained with the 3D7 clone were in accordance with these two published experiments. However, some differences were observed that might be due to the clones features or/and to the cDNA preparation and labeling protocol (fluorescence and use of RNA amplification).

Even though most of the transcript profiles were similar between 3D7 and F12 clones, differential profiles were observed for genes potentially coding for kinases. First, two kinases, *MAL6P1.146* (haem-regulated inhibitory (HRI) kinase) and *PF13\_0258* (a putative serine/threonine protein kinase) were clustered since expression was quite similar throughout 3D7 cycle in contrast to a maximal peak of expression occurring in the F12 clone, at LT for *MAL6P1.146* and ES for *PF13\_0258* (Fig. 2A). Since *PF13\_025* expression peaked in the schizont stage for 3D7 (Scripps/GNF Malaria Array data) and HB3 clones as well as for F12 clone, this gene does not appear to be essential for either asexual or sexual growth.

*MAL6P1.146*, also known as PfPK4, has been annotated as a protozoan eIF-2 $\alpha$ -related protein kinase due to its homology to HRI kinase, and reported to participate to the inhibition of protein synthesis, via the phosphorylation of the small subunit of the eIF-2 initiation factor (Clemens, 2001). Indeed, activity of the recombinant PfPK4 has been described to be inhibited by haemin (Mohrle et al., 1997). In erythrocytes infected by the K1 isolate, the level of the *MAL6P1.146* protein appeared quite similar during the parasite cycle even though its cellular distribution varied (Mohrle et al., 1997). Since PfPK4 accumulates in rhoptries, it probably plays a role in invasion and early development of the ring stage. In addition, erythrocytes contain small concentrations of free haemin and this level is increased in aged erythrocytes and pathological cells of sickle cell anaemia (Shaklai et al., 1985). The inhibition of PfPK4 activity by haemin might play a role in sensing internal erythrocyte environment. Since differential expression takes place between 3D7 and F12 clones and since enhanced gametocyte production occurs in young erythrocytes and in blood from patients with sickle cell anaemia (Trager and Gill, 1992), PfPK4 might be implicated in commitment of the young ring to gametocytogenesis by sensing internal erythrocyte environment. Over-expression of this gene in trophozoite might be at least in part

responsible for a decreased sensitivity to erythrocyte internal environment leading to its inability to respond to the commitment stimulus.

Second, *PFC0485w* encoding a putative Ser/Thr kinase, displayed opposite expression profiles (Fig. 2B) with a maximal expression for F12 in early schizonts, and a minimal for expression 3D7 clone. Interestingly, expression has also been shown to be increased in 3D7 gametocytes (Le Roch et al., 2003), and therefore, might be involved in gametocytogenesis providing an explanation for the gametocyte-less phenotype of F12 clone.

Finally, *PF07\_0073*, a seryl-tRNA ligase, expression has been shown to be maximal in the early schizonts in 3D7 as also observed in HB3 in contrast to F12 where expression was maximal at early trophozoite stage (Fig. 2C) as seen for 3D7 (Scripps/GNF Malaria Array data, sorbitol synchronization). The expression profile of this gene may not be crucial for sexual development since it is different in the gametocyte producing clones.

In addition to kinases, several transcripts coding for proteins involved in transcriptional machinery showed differential profiles between the two clones even though most of the TF transcripts investigated presented similar expression during erythrocytic development.

The relative level of *PFB0865w* transcript, encoding a putative small nuclear ribonucleoprotein probably involved in RNA splicing, was higher than ring in 3D7 clone, in good agreement with the Scripps/GNF Malaria Array data, and continuously lower than ring in F12, as stated for the HB3 clone. Since HB3 clone is a gametocyte producer, this gene may not play a crucial role in controlling gametocyte production.

The maximal peak of expression of *PFC0805w* (a putative DNA-directed RNA polymerase II largest subunit) was reached at different stages of development, delayed in F12 when compared to 3D7. The delay observed in F12 clone may have very limited effects considering the crucial role of RNA polymerase II in transcription.

Finally, two putative transcriptional regulators *PF10930c* and *PF14\_0383*, showed distinct expression profiles in the two clones. Expression of *PF10930c* transcript encoding a protein also known as PfB7, bearing a NAP motif, is not greatly altered throughout 3D7 parasite while it is markedly enhanced in F12 early schizonts. Proteins bearing a NAP motif have been shown to act as histone chaperones, shuttling both core and linker histones from their site of synthesis in the cytoplasm to the nucleus (Akey and Luger, 2003). These proteins may be involved in regulating gene expression by interaction either with chromatin remodelling factors therefore promoting transcriptional silencing (Singer et al., 1998) or with architectural factors such as HMGB proteins interacting with nucleosomes and promoting their sliding (Agregi and Bianchi, 2003). Interestingly, as we have annotated various HMGB proteins and since they are expressed during erythrocytic cycle (to be published), PfB7 might be

responsible for the modulation of the functional activity of these proteins. We have identified orthologs protein of *PF10930c* in the four *Plasmodium* species accessible in PlasmoDB that showed a very high level of similarity (about 95% of the amino acid content), as stated before (Birago et al., 1996), enlightening the crucial role of this protein (data not shown). Moreover, evidence of two differently-sized mRNAs, encoding this protein, have been obtained in *P. falciparum* (Pace et al., 1998). In *P. berghei*, two different lengths of the 5' untranslated region (UTR) have been proposed to explain these two forms of mRNAs, one accumulating during gametocyte differentiation and the other specific for asexual growth. Since the two transcripts share an identical coding sequence, the PCR product used in this study detects either asexual or sexual specific transcript. Different species of *PfB7* mRNA have been hypothesized to be the results of two different promoters specific for asexual or sexual growth. Therefore, *PF10930c* regulation has been demonstrated to be linked to gametocyte differentiation in *P. berghei* (Pace et al., 1998) and one can assume that it is also the case in *P. falciparum*. Differential expression in 3D7 and F12 clone may therefore be a consequence of differential expression of transcription factors and kinases implicated in the gametocyte differentiation. Indeed, chromatin remodelling may play a role in the impairment of F12 to produce gametocytes.

In the case of *PF14\_0383* transcript coding for a Constans-like transcription factor, the maximal peak was obtained in the early schizont stage in 3D7 when it was observed during early trophozoite stage in F12 clone (Fig. 2C). The *PF14\_0383* protein comprises two B-box motifs as seen in Constans (CO) proteins (Robson et al., 2001) and two nuclear localization signals (NLS) in tandem. Indeed, the B-box zinc finger motifs are present in a large number of proteins involved in transcription regulation, usually associated with RING finger and coil-coil motifs to form a tripartite motif, and if mutated, associated with human disease and cancer (Torok and Etkin, 2001). In plants, CO belongs to a family of transcription factors characterized by two zinc-finger B-boxes in tandem in their N-termini, believed to mediate protein-protein interaction (Robson et al., 2001; Samach et al., 2000). CO is involved in transcriptional regulation of flowering in *Arabidopsis thaliana* and promotes the expression of specific sets of genes (Samach et al., 2000). Again, we identified protein orthologs in the four *Plasmodium* species accessible in PlasmoDB (data not shown) showing conserved region located in the two B-boxes and in two other locations of the coding region including the two putative NLS. The first B-box matches the exact consensus (C x2 C x8 C x7 C x2 C x4 H x8 H) derived from CO B-box (Robson et al., 2001) where x represents any amino acid (aa). In addition, two other amino acids conserved in CO motif were shown, if mutated, to result in a flowering phenotype mutant of *A. thaliana* (Koornneef et al., 1991), and were conserved in the *PF14\_0383* first B-box motif. The second B-box appeared

quite similar to the canonical B-box consensus [C x2 H x(4–9) C x2 C x4 C x2 H (C)] (Torok and Etkin, 2001). Bioinformatics analyses of the two B-boxes appeared to reveal a genuine CO-like transcription factors. Although its function as a transcription factor, has not yet been demonstrated, it possesses highly conserved motifs that have been involved in transcription regulation. Differential expression of this gene between the two clones may promote differential expression of sets of genes implicated in gametocytogenesis.

This study is the first attempt to analyse, by using a gene-specific microarray, the expression profile of messengers encoding proteins implicated in the cell cycle and transcription machinery, among which potential transcription factors not involved in the basal transcription but in the modulation of transcription level. Our results led us to identify, during the erythrocytic development of the 3D7 and F12 clones, transcription factors and kinases differentially expressed. They might in turn lead to differential expression of different sets of genes directly or indirectly implicated in the gametocyte differentiation. Nonetheless, true functional involvement in gametocytogenesis remains to be demonstrated. More work is needed and already engaged in order to decipher the potential involvement of each differentially expressed gene in F12 clone gametocyte-less phenotype.

### Acknowledgements

We are indebted to Dr. Walliker and Dr. Alano for the kind gift of the *P. falciparum* clones. We are grateful for the microarray facilities provided by L. Galio and A. Doukani at P3S (Pitié-Salpêtrière). M.G., P.R. and C.B. were financially supported by the Ministère de l'Éducation Nationale, de la Recherche et de la Technologie and the PAL+ program. This work was supported by INSERM, France, to C.V. and D.M.

### Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.gene.2004.07.004](https://doi.org/10.1016/j.gene.2004.07.004).

### References

- Agresti, A., Bianchi, M.E., 2003. HMGB proteins and gene expression. *Curr. Opin. Genet. Dev.* 13, 170–178.
- Akey, C.W., Luger, K., 2003. Histone chaperones and nucleosome assembly. *Curr. Opin. Struck. Biol.* 13, 6–14.
- Alano, P., Roca, L., Smith, D., Read, D., Carter, R., Day, K., 1995. *Plasmodium falciparum*: parasites defective in early stages of gametocytogenesis. *Exp. Parasitol.* 81, 227–235.
- Alano, P., Silvestrini, F., Roca, L., 1996. Structure and polymorphism of the upstream region of the pfg27/25 gene, transcriptionally regulated in gametocytogenesis of *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* 79, 207–217.
- Birago, C., Pace, T., Barca, S., Picci, L., Ponzi, M., 1996. A chromatin-associated protein is encoded in a genomic region highly conserved in the *Plasmodium* genus. *Mol. Biochem. Parasitol.* 80, 193–202.
- Bozdech, Z., Llinas, M., Pulliam, B.L., Wong, E.D., Zhu, J., De Risi, J.L., 2003. The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. *PLoS Biol.* 1, 5.
- Clemens, M.J., 2001. Initiation factor eIF2 alpha phosphorylation in stress responses and apoptosis. *Prog. Mol. Subcell. Biol.* 27, 57–89.
- Day, K.P., Karamalis, F., Thompson, J., Barnes, D.A., Peterson, C., Brown, H., Brown, G.V., Kemp, D.J., 1993. Genes necessary for expression of a virulence determinant and for transmission of *Plasmodium falciparum* are located on a 0.3-megabase region of chromosome 9. *Proc. Natl. Acad. Sci. U. S. A.* 90, 8292–8296.
- Dyer, M., Day, K., 2000a. Expression of *Plasmodium falciparum* trimeric G proteins and their involvement in switching to sexual development. *Mol. Biochem. Parasitol.* 110, 437–448.
- Dyer, M., Day, K.P., 2000b. Commitment to gametocytogenesis in *Plasmodium falciparum*. *Parasitol. Today* 16, 102–107.
- Gardner, M.J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R.W., Carlton, J.M., Pain, A., Nelson, K.E., Bowman, S., Paulsen, I.T., James, K., Eisen, J.A., Rutherford, K., Salzberg, S.L., Craig, A., Kyes, S., Chan, M.S., Nene, V., Shallom, S.J., Suh, B., Peterson, J., Angiuoli, S., Pertea, M., Allen, J., Selengut, J., Haft, D., Mather, M.W., Vaidya, A.B., Martin, D.M., Fairlamb, A.H., Fraunholz, M.J., Roos, D.S., Ralph, S.A., McFadden, G.I., Cummings, L.M., Subramanian, G.M., Mungall, C., Venter, J.C., Carucci, D.J., Hoffman, S.L., Newbold, C., Davis, R.W., Fraser, C.M., Barrell, B., 2002. Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* 419, 498–511.
- Horrocks, P., Jackson, M., Cheesman, S., White, J.H., Kilbey, B.J., 1996. Stage specific expression of proliferating cell nuclear antigen and DNA polymerase delta from *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* 79, 177–182.
- Inselburg, J., 1983. Stage-specific inhibitory effect of cyclic AMP on asexual maturation and gametocyte formation of *Plasmodium falciparum*. *J. Parasitol.* 69, 592–597.
- Koornneef, M., Hanhart, C.J., van der Veen, J.H., 1991. A genetic and physiological analysis of late flowering mutants in *Arabidopsis thaliana*. *Mol. Gen. Genet.* 229, 57–66.
- La Greca, N., Hibbs, A.R., Riffkin, C., Foley, M., Tilley, L., 1997. Identification of an endoplasmic reticulum-resident calcium-binding protein with multiple EF-hand motifs in asexual stages of *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* 89, 283–293.
- Lanzer, M., de Bruin, D., Ravetch, J.V., 1992a. A sequence element associated with the *Plasmodium falciparum* KAHRP gene is the site of developmentally regulated protein–DNA interactions. *Nucleic Acids Res.* 20, 3051–3056.
- Lanzer, M., de Bruin, D., Ravetch, J.V., 1992b. Transcription mapping of a 100 kb locus of *Plasmodium falciparum* identifies an intergenic region in which transcription terminates and reinitiates. *EMBO J.* 11, 1949–1955.
- Le Roch, K.G., Zhou, Y., Blair, P.L., Grainger, M., Moch, J.K., Haynes, J.D., De La Vega, P., Holder, A.A., Batalov, S., Carucci, D.J., Winzeler, E.A., 2003. Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science*.
- Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2-(Delta Delta C(T)) Method. *Methods* 25, 402–408.
- Lobo, C.A., Kumar, N., 1999. Differential transcription of histone genes in asexual and sexual stages of *Plasmodium falciparum*. *Int. J. Parasitol.* 29, 1447–1449.
- Lobo, C.A., Konings, R.N., Kumar, N., 1994. Expression of early gametocyte-stage antigens Pfg27 and Pfs16 in synchronized game-

- toocytes and non-gametocyte producing clones of *Plasmodium falciparum*. Mol. Biochem. Parasitol. 68, 151–154.
- Lobo, C.A., Fujioka, H., Aikawa, M., Kumar, N., 1999. Disruption of the Pfg27 locus by homologous recombination leads to loss of the sexual phenotype in *P. falciparum*. Mol. Cell. 3, 793–798.
- Mohrle, J.J., Zhao, Y., Wemli, B., Franklin, R.M., Kappes, B., 1997. Molecular cloning, characterization and localization of PfPK4, an eIF-2alpha kinase-related enzyme from the malarial parasite *Plasmodium falciparum*. Biochem. J. 328 (Pt. 2), 677–687.
- Ono, T., Nakabayashi, T., 1990. Gametocytogenesis induction by ammonium compounds in cultured *Plasmodium falciparum*. Int. J. Parasitol. 20, 615–618.
- Pace, T., Birago, C., Janse, C.J., Picci, L., Ponzi, M., 1998. Developmental regulation of a *Plasmodium* gene involves the generation of stage-specific 5' untranslated sequences. Mol. Biochem. Parasitol. 97, 45–53.
- Refour, P., Siau, A., Gissot, M., Briquet, S., Boschet, C., Mazier, D., Vaquero, C., 2004. High-accuracy of nanogram total RNA amounts for gene profiling without amplification (submitted for publication).
- Robson, F., Costa, M.M., Hepworth, S.R., Vizir, I., Pineiro, M., Reeves, P.H., Putterill, J., Coupland, G., 2001. Functional importance of conserved domains in the flowering-time gene CONSTANS demonstrated by analysis of mutant alleles and transgenic plants. Plant J. 28, 619–631.
- Salin, H., Vujasinovic, T., Mazurie, A., Maitrejean, S., Menini, C., Mallet, J., Dumas, S., 2002. A novel sensitive microarray approach for differential screening using probes labelled with two different radioelements. Nucleic Acids Res. 30, e17.
- Samach, A., Onouchi, H., Gold, S.E., Ditta, G.S., Schwarz-Sommer, Z., Yanofsky, M.F., Coupland, G., 2000. Distinct roles of CONSTANS target genes in reproductive development of *Arabidopsis*. Science 288, 1613–1616.
- Schena, M., Shalon, D., Davis, R.W., Brown, P.O., 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 270, 467–470.
- Shaklai, N., Shviro, Y., Rabizadeh, E., Kirschner-Zilber, I., 1985. Accumulation and drainage of hemin in the red cell membrane. Biochim. Biophys. Acta 821, 355–366.
- Sharma, A., Sharma, I., Kogkasuriyachai, D., Kumar, N., 2003. Structure of a gametocyte protein essential for sexual development in *Plasmodium falciparum*. Nat. Struct. Biol. 10, 197–203.
- Singer, M.S., Kahana, A., Wolf, A.J., Meisinger, L.L., Peterson, S.E., Goggin, C., Mahowald, M., Gottschling, D.E., 1998. Identification of high-copy disruptors of telomeric silencing in *Saccharomyces cerevisiae*. Genetics 150, 613–632.
- Torok, M., Etkin, L.D., 2001. Two B or not two B? Overview of the rapidly expanding B-box family of proteins. Differentiation 67, 63–71.
- Trager, W., Gill, G.S., 1992. Enhanced gametocyte formation in young erythrocytes by *Plasmodium falciparum* in vitro. J. Protozool. 39, 429–432.
- Trager, W., Jensen, J.B., 1976. Human malaria parasites in continuous culture. Science 193, 673–675.

## II – CARACTERISATION FONCTIONNELLE ET ROLE DE PFMYB1, UN FACTEUR DE TRANSCRIPTION SPECIFIQUE DE SEQUENCE, DANS LA REGULATION DES GENES AU COURS DU CYCLE ERYTHROCYTAIRE DE DEUX CLONES DE *P. FALCIPARUM*, 3D7 ET F12.

### II.1 – Caractérisation fonctionnelle de PfMyb1

#### II.1.1 Résumé

Comme chez les eucaryotes, la transcription chez *Plasmodium* semble être régulée par les interactions entre les éléments *cis*- et *trans*-régulateurs. Mais le nombre d'acteurs de la régulation de la transcription décrits chez *P. falciparum* est faible, tant au niveau des facteurs de transcription qu'au niveau des éléments de régulation identifiés dans les promoteurs des gènes. La caractérisation de **FT spécifiques de séquence ADN** est essentielle à l'étude de la régulation de l'expression des gènes. Avant que le génome de *P. falciparum* ne soit complètement séquencé, plusieurs séquences de FT avaient déjà été identifiées dans notre laboratoire. Les FT sont définis selon leur domaine de liaison à l'ADN. Nous avons identifié dans le génome de *P. falciparum* par homologie de séquences avec les FT de la famille Myb, connue chez les eucaryotes pour être impliquée dans le développement et la différenciation, une phase ouverte de lecture que nous avons appelée : PfMyb1. La séquence de la protéine PfMyb1 partage les caractéristiques communes à la classe des FT à domaines tryptophane (Fig.1, p. 160 de l'Article 2).

Cependant, ces protéines prédites par analyse bioinformatique nécessitent une validation biologique de leur fonction pour une annotation définitive. Nous avons alors entrepris la caractérisation fonctionnelle de la protéine PfMyb1 par des expériences biologiques. La fonction d'un FT spécifique de séquence est caractérisée par sa capacité à se lier à des motifs de l'ADN de manière spécifique. Les protéines de la famille Myb sont connues pour interagir spécifiquement avec une séquence d'ADN consensus appelée Myb Regulatory Element (MRE). Nous avons pu identifier des motifs de régulation MRE caractéristiques des protéines Myb eucaryotes dans des promoteurs de gènes plasmodiaux. Par des expériences de retard sur gel, nous avons montré qu'une ou plusieurs protéines contenues dans les extraits nucléaires de *P. falciparum* sont capables de se lier spécifiquement à plusieurs séquences d'oligonucléotides contenant un MRE (Fig. 2, p. 161 de l'Article 2). Les

éléments de régulation MRE putatifs retrouvés dans deux promoteurs de *P. falciparum* ont permis d'obtenir une interaction comparable à celle obtenue avec une séquence prototype MRE présente dans le promoteur du gène *mim-1* de poulet connue pour interagir avec des protéines de la famille Myb. Cette interaction est spécifique puisqu'elle est chassée par un anticorps spécifique anti-PfMyb1 et qu'elle n'est pas perturbée par un anticorps non pertinent anti-NF- $\kappa$ B. La protéine PfMyb1 possède donc les qualités essentielles d'un FT spécifique de séquence de la famille Myb.

Enfin, nous avons analysé la cinétique d'interaction de la protéine au cours du développement érythrocytaire des deux clones de *P. falciparum*, 3D7 et F12 déficient pour la gamétocytogénèse afin d'évaluer l'implication éventuelle de PfMyb1 dans l'initiation de la différenciation sexuée. Ces deux cinétiques d'interaction sont différentes et marquées par l'absence de complexes ADN/protéines entre les extraits nucléaires d'anneaux et de schizontes des parasites F12 et le MRE. Ces résultats sont en contraste avec les cinétiques d'expression semblables du transcrite *pfmyb1* et de sa protéine dans les deux clones. L'absence d'un partenaire de PfMyb1 ou encore l'absence d'une modification post-traductionnelle pourrait perturber l'activité de liaison de PfMyb1 aux stades anneaux et schizontes du clone F12. Ces résultats indiquent une possible implication du FT PfMyb1 dans l'incapacité de ce clone à produire des gamétocytes.



**ARTICLE 2:**

**Characterization of PfMyb1 transcription factor during erythrocytic development of 3D7 and F12 *Plasmodium falciparum* clones.**

Charlotte Boschet\*, Mathieu Gissot\*, Sylvie Briquet, Zuhail Hamid, Clotilde Claudel-Renard et Catherine Vaquero.

\* : les deux auteurs ont contribué également à ce travail.

Accepté à la publication dans *Molecular and Biochemical Parasitology*.

Short communication

# Characterization of PfMyb1 transcription factor during erythrocytic development of 3D7 and F12 *Plasmodium falciparum* clones<sup>☆</sup>

Charlotte Boschet<sup>a,1</sup>, Mathieu Gissot<sup>a,1</sup>, Sylvie Briquet<sup>a</sup>, Zuhail Hamid<sup>a,2</sup>,  
Clotilde Claudel-Renard<sup>b</sup>, Catherine Vaquero<sup>a,\*</sup>

<sup>a</sup> INSERM U511, CHU Pitié-Salpêtrière, 91 boulevard de l'Hôpital, 75013 Paris, France

<sup>b</sup> Department of Biochemistry, University of Pretoria, Pretoria, South Africa

Received 8 May 2004; received in revised form 19 July 2004; accepted 20 July 2004

**Keywords:** *Plasmodium*; Transcription factor; Expression; Erythrocyte; Development

Two major events of the *Plasmodium falciparum* life cycle occur in the erythrocyte: an asexual multiplication responsible for clinical manifestations and initiation of sexual differentiation responsible for dissemination of the disease via mosquito. These developmental pathways require the coordinated and modulated expression of different sets of genes, involving transcriptional controls, even though post-transcriptional regulation could also be implicated. These gene clusters are therefore likely to share regulatory elements or modules within their promoters [1]. Regulation of transcription is also accomplished by the availability of transcription factors, which is presumably modulated throughout parasite development.

In *Plasmodium* very little is known at the level of *cis*- and *trans*-regulatory elements, and to our knowledge only a few regulatory elements and no clearly defined transcription factors have been reported in the literature. Nevertheless, *Plasmodium* transcriptional machinery appeared to share com-

mon features with that of eukaryotes [2]. Moreover, regulatory sequences within the promoters were reported to contain motifs in common with the binding sites of eukaryotic transcription factors [2–5] and our own computer analyses confirmed this statement, including DNA motifs potentially interacting with Myb factors (Boschet and Vaquero, to be published elsewhere).

Herein, we present the first description of a *Plasmodium* transcription factor: a Myb-related protein of the tryptophan cluster family. The Myb proteins are highly conserved in eukaryotes and generally, their characteristic DNA-binding domain (DBD) contains three tandem repeats (R1, R2 and R3) of approximately 50 residues with 3 regularly spaced (18 or 19 amino acids) tryptophan residues [6,7]. They have been shown to bind DNA in a sequence-specific manner and to participate in the regulation of the expression of genes implicated in growth control and differentiation [7]. To identify Myb proteins in the *P. falciparum* genome, more than 200 non-redundant eukaryotic Myb proteins were aligned with MultAlin and ClustalW [8,9] to generate a consensus sequence corresponding to the characteristic DNA-binding domain. This consensus was used as query for the *Plasmodium* database and allowed the annotation of a 414 amino acid-long ORF: PfMyb1. Since PFSCAN [10] identified only one Myb domain (R2 in Fig. 1A), the complete sequence of PfMyb1 was aligned with the DBD of three proteins, DdMybH, DdMyb2 and DdMyb3 [11–13] of the slime mold *Dictyostelium discoideum*, which has an A + T-rich genome (78% overall) like *Plasmodium* (85%). This resulted in the

**Abbreviations:** aa, amino-acid; DBD, DNA-binding domain; EMSA, electrophoretic mobility shift assay; MRE, Myb regulatory element; NE, nuclear extract; ORF, open reading frame; PCR, polymerase chain reaction; RT, reverse transcription

<sup>☆</sup> Note: EMBL accession number AJ291747.

\* Corresponding author. Tel.: +33 1 40778111; fax: +33 1 45838858.

E-mail address: [vaquero@ext.jussieu.fr](mailto:vaquero@ext.jussieu.fr) (C. Vaquero).

<sup>1</sup> The authors wish it to be known that, in their opinion, Charlotte Boschet and Mathieu Gissot should be regarded as joint first authors.

<sup>2</sup> Permanent address: INMO, University of Gezira, Wad Madani, Box 20, Sudan.

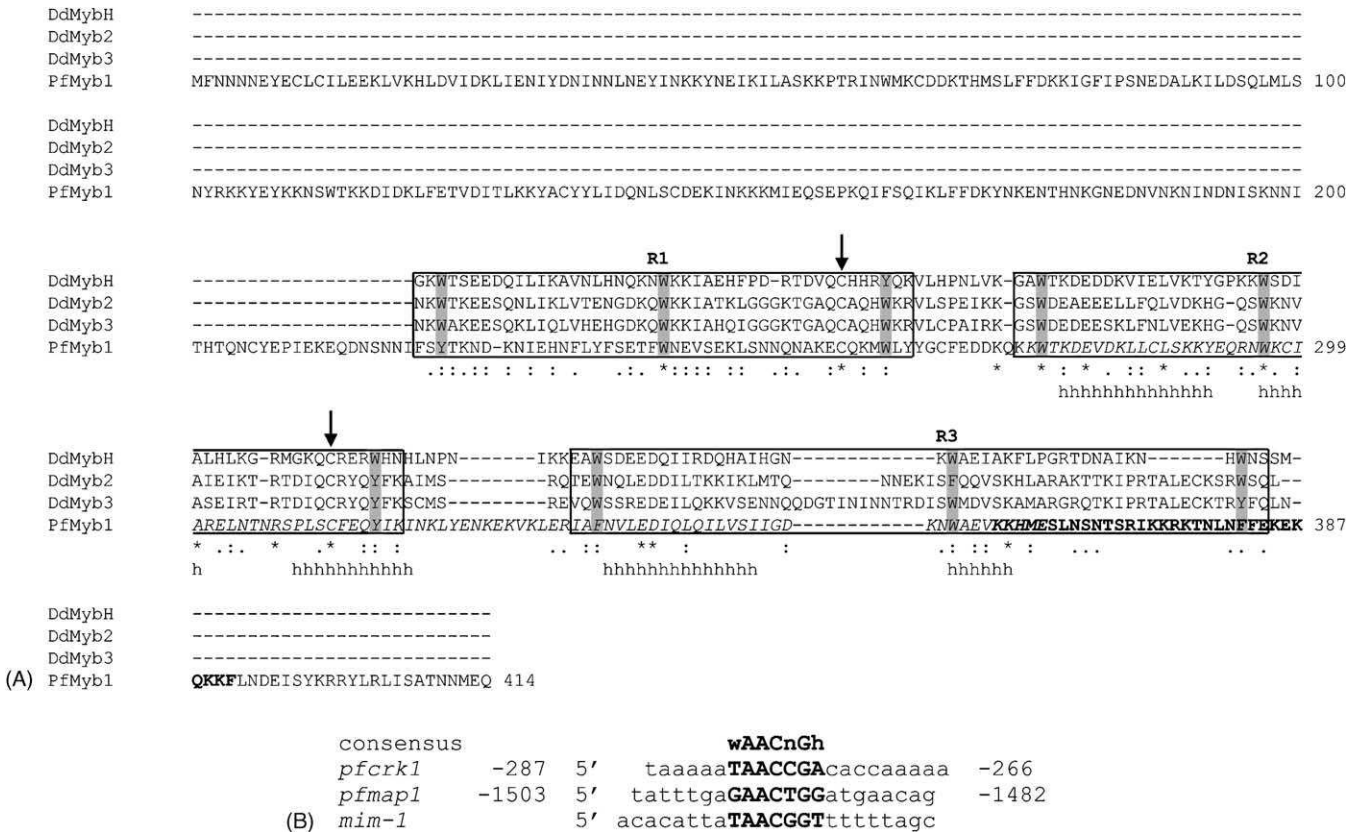


Fig. 1. The DNA-binding domain of the PfMyb1 sequence and three different Myb regulatory elements. (A) The protein consensus of the characteristic DNA-binding domain of the Myb family was obtained with an alignment of eukaryotic Myb selected throughout evolution and allowed the annotation PfMyb1 within the *Plasmodium* chromosome 13. The complete sequence of PfMyb1 is shown as well as an optimal alignment between its putative DNA-binding domain and the three DBD of *Dictyostelium discoideum* (DdMybH, DdMyb2 and DdMyb3; SwissProt accession numbers: P34127, O15816, Q9GUB3, respectively). Identities are indicated by asterisks, strong similarities by two points and weak similarities by one point. The residues W, Y or F that characterize the Myb domains are highlighted in gray and R1, R2 and R3 are boxed. The highly conserved cysteine residues are marked by an arrow. The sequence from Lys 275 to Glu 364 threaded by Meta-Server and TITO [21,35] is indicated in italics and aa predicted to take part in  $\alpha$ -helices are marked with an h. A putative nuclear localization sequence, encompassing the unusually long C-terminal part of the third repeat, is shown in bold letters. (B) Putative MRE were localized among others in *Plasmodium pferck1* and *pfmap1* gene promoters using MatInspector [36] according to three matrix families (one from the plant and two from the vertebrate subsections). Core and matrix similarity parameters were left by default. These two potential MRE were aligned with a prototype MRE found in the chicken *mim-1* gene promoter reported to be functional [26] and with the consensus sequence of the Myb binding domain. Letters in upper-case represent the MRE and in lower-case the flanking sequences. The positions are numbered from the first ATG of the open reading frames. W: A or T; H: A, C or T; N: A, C, G or T.

identification of three Myb domains (Fig. 1A), located in the C-terminus of the protein as in DdMyb2 and DdMyb3 of *D. discoideum*, and these two factors were reported to be functional [12,13] whereas in most of the Myb proteins reported so far, the DBD is located in the N-terminus. However, PfMyb1 contains imperfect repeats with a tyrosine or a phenylalanine in place of tryptophan and with an aa insertion in R3 as observed in *D. discoideum* (Fig. 1A) as well as in *Saccharomyces cerevisiae* Bas1 [14,15] and *Kluyveromyces lactis* Reb1 [16]. Moreover, a critical cysteine residue, as highly conserved as the tryptophan residues, [17] was also found in R1 and R2 at position 258 and 312, respectively, and was reported to play a role in redox regulation [18–20]. Furthermore, 3D-structure was determined with Meta-Server [21] and it appeared that the sequence from Lys 275 to Glu 364 (Fig. 1A), encompassing the R2 domain and the first part of

the R3 domain, could be threaded by homology with the human Myb proto-oncogene protein, whose structure has been determined by X-ray diffraction [22]. Indeed, three  $\alpha$ -helices were predicted in R2 and only two in R3 in contrast to the R1 repeat. Nevertheless, in Myb proteins, R2 and R3 constitute a minimal DBD sufficient for sequence-specific DNA binding, while R1 does not appear to engage in specific interactions with DNA [23]. In summary, the diverse computational analyses of PfMyb1 provide evidence for a genuine Myb protein, which is strikingly conserved in all *Plasmodium* species listed in PlasmoDB.

After annotation of the putative transcription factor PfMyb1, the characterization of its cognate RNA transcript was carried out by Northern blotting experiment. A messenger of 2.6 kb was evidenced (data not shown) in *P. falciparum* infected erythrocytes. Our aim was to analyze the differences

occurring at the level of transcription factors that might control expression of transcripts and/or proteins involved in sexual differentiation, a process crucial for the dissemination of the disease via the mosquito. Using a semi-quantitative RT-PCR, the temporal expression of *pfmyb1* transcript (data not shown) was analyzed at different times of erythrocytic development of highly synchronized parasites (rings, early and late trophozoites, as well as early and late schizonts). The profile of expression was determined in two clones of *P. falciparum*, 3D7 and the gametocyte-less F12 derived from 3D7. The major difference in the mRNA profiles resides in a lower abundance of the *pfmyb1* transcript in the ring stage of F12 when compared to 3D7, followed by a sharp increase in early F12 trophozoites. In contrast, the mRNA level was

quite similar in both clones in the early and late schizonts after a maximal expression in trophozoites. The 3D7 expression profile is in good agreement with the published transcriptome data of HB3 [24] and 3D7 clones [25].

We then analyzed the presence of the transcription factor in the nuclear extracts of the parasite. The *pfmyb1* ORF has been cloned and the His-tagged recombinant protein was expressed, purified and used as a positive control in Western blots. The PfMyb1 protein was shown to be present in the parasite trophozoite nuclear extracts by a serial immunoprecipitation and Western blotting experiments (Fig. 2A). Briefly, nuclear extracts were immunoprecipitated, using tosylactivated magnetic Dynabeads coated with either anti-PfMyb1 serum (lane 2) or preimmune (lane 3). Immunoprecipitated

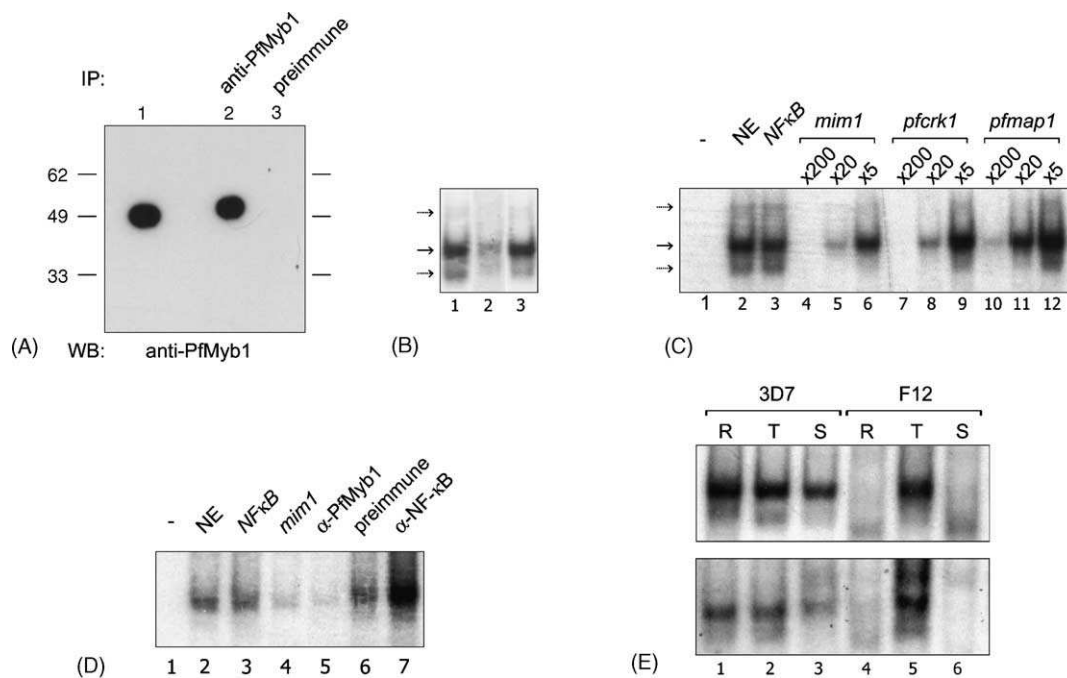


Fig. 2. Functional study of PfMyb1 in *P. falciparum* nuclear extracts. (A) Serial immunoprecipitation and Western blotting experiments. *Plasmodium falciparum* protein extracts were prepared as described in [37]. A polyclonal antipeptide serum was prepared by Eurogentec (Belgium). Two rabbits were immunized with two peptides (RKKYEYKKNNSWTKK and KRRYLRLISATNMEQ), corresponding to amino acids (aa) 103–117 (upstream to the R1 domain) and 400–414 (encompassing the very last C-terminus aa of R3) of the PfMyb1 amino acid sequence, respectively. One hundred micrograms proteins of *P. falciparum* trophozoites were subjected to immunoprecipitation using  $2 \times 10^7$  magnetic Dynabeads tosylactivated M-280 (DynaL Biotech) coated with 2.5  $\mu$ g of anti-PfMyb1 serum (lane 2) or preimmune (lane 3) following manufacturer recommendations. Lane 1 stands for PfMyb1 recombinant protein. Immunoprecipitated proteins were run on a 10% SDS-PAGE and subjected, after transfer onto nitrocellulose membrane, to Western blotting using the anti-PfMyb1 serum and revealed by a peroxidase-conjugated anti-rabbit antibody (Sigma). (B–E) EMSA. Nuclear extracts (NE) from 50 mL of red blood cells at 10%–12% parasitaemia infected with rings (1–18 h), trophozoites (20–28 h) and schizonts (32–42 h post-invasion) were prepared as described by Osta et al. [37]. EMSA experiments were conducted with nuclear extracts of 3D7 and F12 clones for 30 min at 4 °C in binding buffer (25 mM HEPES pH 7.5, 50 mM KCl, 1 mM MgCl<sub>2</sub>, 0.5 mM EDTA, 0.2 mM PMSF, 1 mM DTT, 5% glycerol and 2  $\mu$ g of poly dI-dC). Addition of 50,000 cpm of labeled probes is followed in a final volume of 10  $\mu$ l for an additional 30 min. Samples were loaded on pre-run 6% non-denaturing polyacrylamide gel and electrophoresed at 160 V for 2 h at 4 °C in a low-ionic strength TBE buffer (0.25 $\times$ ). (B) Double stranded oligonucleotides (100 ng) of the prototype and putative *Plasmodium* MRE (Fig. 1B) were end-labeled with 40  $\mu$ Ci of [ $\gamma$ -<sup>32</sup>P] ATP and purified with QIAquick nucleotide removal kit (Qiagen). One micrograms of 3D7 trophozoite nuclear extract was incubated with labeled chicken *mim-1* prototype, *pfmap1* and *pferk1* putative *Plasmodium* MRE (lanes 1–3). A bold arrow shows the major complex and dotted arrows indicate two minor complexes. (C) Competitions were carried out as in (B) only with the labeled chicken *mim-1* MRE prototype in presence of a 200-fold molar excess of unlabeled non-specific *NF- $\kappa$ B* (lane 3) and a 200-, 20- and 5-fold excess of specific *mim-1* (lanes 4–6), *pferk1* (lanes 7–9) and *pfmap1* (lanes 10–12) oligonucleotides. Controls without and with NE alone are presented in lanes 1 and 2, respectively. (D) Interaction between 3D7 nuclear extract and *mim-1* was assayed in the presence of either 0.5  $\mu$ l of anti-PfMyb1 polyclonal antibody (0.6  $\mu$ g/ $\mu$ l, lane 5), or 0.5  $\mu$ l of the corresponding preimmune serum (lane 6), or 0.5  $\mu$ l of an irrelevant anti-NF- $\kappa$ B antibody (lane 7). Lanes 1–4 are as in (C). (E) Labeled *mim-1* oligonucleotide was incubated with 1  $\mu$ g of nuclear extracts of 3D7 and F12 rings (R, lanes 1 and 4), trophozoites (T, lanes 2 and 5) and schizonts (S, lanes 3 and 6). Two independent experiments are presented.

proteins were subjected to SDS-PAGE and Western blotting with the anti-PfMyb1 serum. A single band, whose molecular mass (50 kDa) is slightly higher than the recombinant PfMyb1 (lane 1) was observed (lane 2) in the parasite nuclear extracts. In contrast, when using preimmune coated beads (lane 3) no protein was revealed by the anti-PfMyb1 serum.

In order to activate or repress transcription of particular sets of genes, the transcription factors have to bind *cis*-regulatory DNA elements in a sequence-specific manner. A Myb-related protein was evidenced in the 3D7 trophozoite nuclear extracts (NE) since specific interactions were detected with diverse Myb-regulatory elements (Fig. 1B), such as a prototype MRE (*mim-1* gene promoter) [26] and two putative MRE naturally occurring in the promoters of *Plasmodium* *pfmap1* and *pfcrk1* genes. These genes were originally reported to be expressed preferentially during erythrocytic asexual and sexual stages, respectively [27,28]. The EMSA profiles (Fig. 2B) appeared quite similar for the three MRE assayed, all giving rise to one major and two minor complexes, even though the magnitude of interaction was different, with its highest level being observed for *mim-1* and *pfcrk1* MRE, in line with the nucleotide sequence of the elements (Fig. 1B). These retarded bands could result from homo-dimerization of the PfMyb1 protein, as well as from interaction with other proteins, which have been shown to interact with the Myb proteins, as the bZip transcription factor C/EBP $\beta$  [22] and the poly(ADP-ribose) polymerase (PARP) [29]. The DNA/protein interaction was specific, as shown by competition experiments with unlabeled oligonucleotides corresponding to each MRE (Fig. 2C). The interaction was not competed out by a 200-fold excess of an irrelevant unlabeled NF- $\kappa$ B element present in the *IL-2R $\alpha$*  promoter [30] (lane 3), in contrast to a 200-fold excess of unlabeled relevant *mim-1*, *pfcrk1* and *pfmap1* (lanes 4, 7 and 10, respectively). In Fig. 2D, the specificity of the interaction was verified by using the anti-PfMyb1 antibody (lane 5), which competed out interaction with the labeled probe as efficiently as the *mim-1* competitor (lane 4), in contrast to the pre-immune serum (lane 6) and an irrelevant anti-NF- $\kappa$ B antibody [31] (lane 7). Indeed, anti-PfMyb1 antibody, raised against the C-terminus of the DBD, impaired complex formation in place of generating a super-shifted band as already reported for other transcription factors [30,32]. For all these reasons we think that PfMyb1 is a genuine transcription factor. However, it was not possible, under the experimental conditions assayed so far, to demonstrate any interaction between these MRE and the recombinant protein, whether complete or incomplete, encompassing the three Myb domains. This might be due to inappropriate folding, processing, and/or phosphorylation of the recombinant protein, as has already been reported for other recombinant *P. falciparum* proteins expressed in *E. coli* [33].

Finally, we compared the temporal expression of PfMyb1 protein during 3D7 and F12 erythrocytic development with nuclear extracts prepared from the three main forms: rings, trophozoites and schizonts. EMSA experiments using the

three nuclear extracts and labeled *mim-1* oligonucleotide are shown in Fig. 2E and correspond to two representative and independent biological experiments. The major complex was present throughout development of the 3D7 clone, decreasing only slightly up to the schizont stage (lanes 1–3). In contrast, for the F12 clone the major complex, while quite apparent in trophozoites (lane 5), was almost undetectable in rings and schizonts (lanes 4 and 6). A discrepancy between transcript and EMSA profiles was quite apparent, especially in the F12 clone. The messenger was observed in F12 rings, despite absence of Myb/MRE interaction. Moreover, in early and late schizonts of both clones, the *pfmyb1* transcript was present to equal extents, while the DNA/protein interaction was detectable in 3D7 and undetectable in F12 nuclear extracts. The absence of interaction was not due to the degradation or lower concentration of the nuclear extracts used in the EMSA, since the integrity and protein content of all nuclear preparations were verified by SDS-PAGE analysis and Bradford assay (data not shown). Decreased mRNA translation [34] or absence of expression of a co-factor responsible for the formation of EMSA complexes, among other explanations, could account for the absence of retarded band. We are not, however, as yet able to address these questions. Nevertheless, the lack of PfMyb1 or partners in rings and schizonts might participate to the impairment of sexual differentiation towards gametocytes observed in F12 clone.

In summary, by computational annotation we identified, within the genomic DNA of *P. falciparum*, an open reading frame sharing common features with the Myb transcription factors. We assumed that this factor could regulate expression of particular genes since appropriate binding sites were identified within their promoters. In nuclear extracts, a Myb DNA binding activity was observed with a prototype and two putative *Plasmodium* Myb regulatory elements. This interaction was demonstrated to be specific, since it was inhibited by specific competitors and anti-PfMyb1 antibody in bandshift assays. Finally, during erythrocytic development, the bandshift profiles were markedly different in the 3D7 and the gametocyte-less F12 clones, in contrast to the transcript profile. These experiments provide the first evidence of the presence, in *Plasmodium*, of a transcription factor belonging to the highly conserved Myb family that might be implicated in the regulation of gene expression.

## Acknowledgements

We thank Pietro Alano for his gift of F12 clone, Leila Gannoun (UMR-CNRS 5539) for the nuclear extract protocol and Alain Israël for the anti-NF- $\kappa$ B antibody. We are grateful to Jennifer Richardson for critical reading of the manuscript, and to Professors M.M. Magzoub (Minister of Higher Education and Scientific Research, Sudan) and Dominique Mazier, as well as Philippe Refour (INSERM U511), for continuous encouragement and support. C.B. and M.G. were financially supported by the Ministère de l'Éducation Nationale, France,

and Z.H. by the WHO Mediterranean Office and the Ministry of Higher Education and Scientific Research, Sudan. The project was supported by INSERM funds to C.V.

## References

- [1] Arnone MI, Davidson EH. The hardwiring of development: organization and function of genomic regulatory systems. *Development* 1997;124:1851–64.
- [2] Horrocks P, Decherig K, Lanzer M. Control of gene expression in *Plasmodium falciparum*. *Mol Biochem Parasitol* 1998;95:171–81.
- [3] Lanzer M, de Bruin D, Ravetch JV. Transcription mapping of a 100 kb locus of *Plasmodium falciparum* identifies an intergenic region in which transcription terminates and reinitiates. *EMBO J* 1992;11:1949–55.
- [4] Lanzer M, Wertheimer SP, de Bruin D, Ravetch JV. *Plasmodium*: control of gene expression in malaria parasites. *Exp Parasitol* 1993;77:121–8.
- [5] Horrocks P, Kilbey BJ. Physical and functional mapping of the transcriptional start sites of *Plasmodium falciparum* proliferating cell nuclear antigen. *Mol Biochem Parasitol* 1996;82:207–15.
- [6] Kanei-Ishii C, Nomura T, Ogata K, et al. Structure and function of the proteins encoded by the *myb* gene family. *Curr Top Microbiol Immunol* 1996;211:89–98.
- [7] Lipsick JS. One billion years of Myb. *Oncogene* 1996;13:223–35.
- [8] Corpet F. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res* 1988;16:10881–90.
- [9] Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 1994;22:4673–80.
- [10] PFSCAN, <http://hits.isb-sib.ch/cgi-bin/PFSCAN>.
- [11] Stober-Grasser U, Brydolf B, Bin X, Grasser F, Firtel RA, Lipsick JS. The Myb DNA-binding domain is highly conserved in *Dictyostelium discoideum*. *Oncogene* 1992;7:589–96.
- [12] Otsuka H, Van Haastert PJ. A novel Myb homolog initiates *Dictyostelium* development by induction of adenylyl cyclase expression. *Genes Dev* 1998;12:1738–48.
- [13] Guo K, Anjard C, Harwood A, Kim HJ, Newell PC, Gross JD. A myb-related protein required for culmination in *Dictyostelium*. *Development* 1999;126:2813–22.
- [14] Tice-Baldwin K, Fink GR, Arndt KT. BAS1 has a Myb motif and activates HIS4 transcription only in combination with BAS2. *Science* 1989;246:931–5.
- [15] Hovring I, Bostad A, Ording E, Myrset AH, Gabrielsen OS. DNA-binding domain and recognition sequence of the yeast BAS1 protein, a divergent member of the Myb family of transcription factors. *J Biol Chem* 1994;269:17663–9.
- [16] Morrow BE, Ju Q, Warner JR. A bipartite DNA-binding domain in yeast Reb1p. *Mol Cell Biol* 1993;13:1173–82.
- [17] Pinson B, Brendeford EM, Gabrielsen OS, Daignan-Fornier B. Highly conserved features of DNA binding between two divergent members of the myb family of transcription factors. *Nucleic Acids Res* 2001;29:527–35.
- [18] Myrset AH, Bostad A, Jamin N, Lirsac PN, Toma F, Gabrielsen OS. DNA and redox state induced conformational changes in the DNA-binding domain of the Myb oncoprotein. *EMBO J* 1993;12:4625–33.
- [19] Guehmann S, Vorbrueggen G, Kalkbrenner F, Moelling K. Reduction of a conserved Cys is essential for Myb DNA-binding. *Nucleic Acids Res* 1992;20:2279–86.
- [20] Grasser FA, LaMontagne K, Whittaker L, Stohr S, Lipsick JS. A highly conserved cysteine in the v-Myb DNA-binding domain is essential for transformation and transcriptional trans-activation. *Oncogene* 1992;7:1005–9.
- [21] Douguet D, Labesse G. Easier threading through web-based comparisons and cross-validations. *Bioinformatics* 2001;17:752–3.
- [22] Tahirov TH, Sato K, Ichikawa-Iwata E, et al. Mechanism of c-Myb-C/EBP beta cooperation from separated sites on a promoter. *Cell* 2002;108:57–70.
- [23] Ogata K, Morikawa S, Nakamura H, et al. Solution structure of a specific DNA complex of the Myb DNA-binding domain with cooperative recognition helices. *Cell* 1994;79:639–48.
- [24] Bozdech Z, Llinas M, Pulliam BL, Wong ED, Zhu J, DeRisi JL. The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. *PLoS Biol* 2003;1:5.
- [25] Le Roch KG, Zhou Y, Blair PL, et al. Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science* 2003;301:1503–8.
- [26] Ness SA, Marknell A, Graf T. The v-myb oncogene product binds to and activates the promyelocyte-specific *mim-1* gene. *Cell* 1989;59:1115–25.
- [27] Doerig C, Horrocks P, Coyle J, et al. Pferk-1, a developmentally regulated cdc2-related protein kinase of *Plasmodium falciparum*. *Mol Biochem Parasitol* 1995;70:167–74.
- [28] Doerig CM, Parzy D, Langsley G, Horrocks P, Carter R, Doerig CD. A MAP kinase homologue from the human malaria parasite, *Plasmodium falciparum*. *Gene* 1996;177:1–6.
- [29] Cervellera MN, Sala A. Poly(ADP-ribose) polymerase is a B-MYB coactivator. *J Biol Chem* 2000;275:10692–6.
- [30] Galio L, Briquet S, Cot S, Guillet JG, Vaquero C. Analysis of interactions between huGATA-3 transcription factor and three GATA regulatory elements of HIV-1 long terminal repeat, by surface plasmon resonance. *Anal Biochem* 1997;253:70–7.
- [31] Blank V, Kourilsky P, Israel A. Cytoplasmic retention, DNA binding and processing of the NF-kappa B p50 precursor are controlled by a small region in its C-terminus. *EMBO J* 1991;10:4159–67.
- [32] Markle D, Das S, Ward SV, Samuel CE. Functional analysis of the KCS-like element of the interferon-inducible RNA-specific adenosine deaminase ADAR1 promoter. *Gene* 2003;304:143–9.
- [33] Pandey KC, Singh S, Pattnaik P, et al. Bacterially expressed and refolded receptor binding domain of *Plasmodium falciparum* EBA-175 elicits invasion inhibitory antibodies. *Mol Biochem Parasitol* 2002;123:23–33.
- [34] Decherig KJ, Thompson J, Dodemont HJ, Eling W, Konings RN. Developmentally regulated expression of pfs16, a marker for sexual differentiation of the human malaria parasite *Plasmodium falciparum*. *Mol Biochem Parasitol* 1997;89:235–44.
- [35] Labesse G, Mornon J. Incremental threading optimization (TITO) to help alignment and modelling of remote homologues. *Bioinformatics* 1998;14:206–11.
- [36] Quandt K, Frech K, Karas H, Wingender E, Werner T. MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res* 1995;23:4878–84.
- [37] Osta M, Gannoun-Zaki L, Bonnefoy S, Roy C, Vial HJ. A 24 bp cis-acting element essential for the transcriptional activity of *Plasmodium falciparum* CDP-diacylglycerol synthase gene promoter. *Mol Biochem Parasitol* 2002;121:87–98.

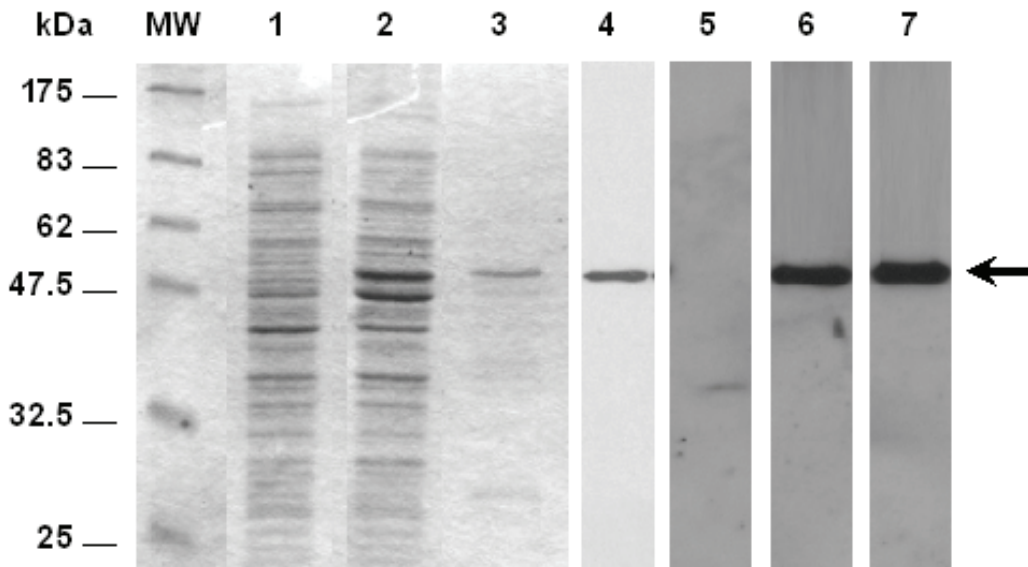
## II.1.2 – Résultats complémentaires

### a) Expression de la protéine PfMyb1 recombinante

La découverte d'une phase ouverte de lecture ayant des similarités avec la famille des FT à domaines tryptophane, dans les séquences accumulées peu à peu lors du séquençage du génome de *P. falciparum*, nous a permis de cloner ce gène à partir de l'ADN génomique du clone 3D7 de *P. falciparum*. La séquence complète de la phase ouverte de lecture a été clonée dans le vecteur PGEM-T easy et séquencée. Deux clones indépendants ont donné une séquence identique et celle-ci a été soumise à la banque de données EMBL (Décembre 2000 numéro d'accèsion AJ291747).

La séquence complète de la phase ouverte de lecture a ensuite été insérée dans le vecteur PQe30 afin de produire la protéine PfMyb1 recombinante additionnée de six histidines dans la partie N-terminale. Après transformation des bactéries, la protéine PfMyb1 recombinante a été exprimée, purifiée et analysée en utilisant les lysats par électrophorèse SDS-PAGE et Western blot. Même si la protéine semblait largement insoluble, la fraction insoluble a pu être purifiée.

La figure 14 présente une expérience typique de purification par affinité au nickel. Les lignes 1 et 2 montrent les lysats d'*E. coli* non induits et induits pour la production de la protéine PfMyb1 recombinante, et la ligne 3 les protéines obtenues après élution de la colonne, après une électrophorèse SDS-PAGE et coloration au bleu de Coomassie. Dans la partie droite de la figure 14, la protéine purifiée a été révélée par Western Blot utilisant un anticorps anti-histidine (ligne 4), le sérum anti-PfMyb1 (ligne 6), l'anticorps immuno-purifié anti-PfMyb1 (ligne 7) ou le sérum pré-immun (ligne 5).



**Figure 15. Analyse de la protéine PfMyb1 recombinante.**

La protéine a été exprimée dans *E. coli* et les lysats non induits et induits ont été fractionnés sur un gel SDS-PAGE 12 % coloré au bleu de Coomassie (lignes 1 et 2). La protéine PfMyb1 recombinante obtenue après élution de la colonne de nickel est présentée en ligne 3. Les lignes 4 à 7 représentent une analyse de la protéine PfMyb1 purifiée en Western blot utilisant un anticorps dirigé contre le tag histidine (ligne 4), un sérum immun anti-PfMyb1 et un anticorps immuno-purifié anti-PfMyb1 (lignes 6 et 7, respectivement), et le sérum pré-immun (ligne 5). Le marqueur de protéines est situé sur la partie gauche et une flèche indique la position de la protéine PfMyb1.

Ces expériences nous ont permis de confirmer la spécificité de l'anticorps qui a été utilisé dans la Fig. 2, p. 161 de l'Article 2.

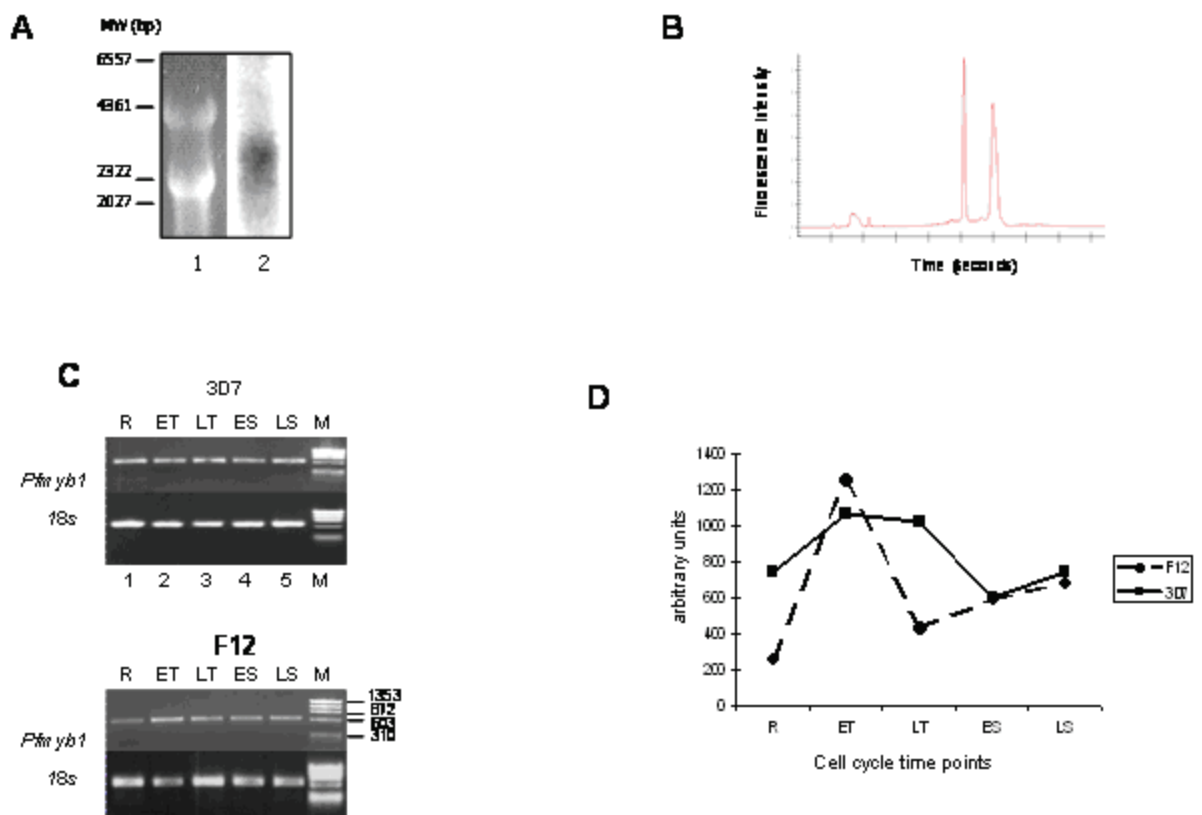
### **b) Etude de la cinétique d'expression du transcrit *pfmyb1***

Le transcrit du gène *pfmyb1* a été caractérisé par Northern Blot à partir d'une préparation d'ARN polyA<sup>+</sup> isolée à partir d'un mélange d'ARN totaux de différents stades du cycle érythrocytaire du parasite. Seul ce protocole, engageant une grande quantité d'ARN, nous a permis d'identifier un transcrit d'une taille de 2,6 kb, suggérant que le transcrit est peu représenté dans le parasite. De plus, la taille du messenger est un exemple supplémentaire de parties transcrites mais non traduites exceptionnellement longues chez *P. falciparum* (Figure 15A).

La cinétique d'expression temporelle du transcrit a été analysée tout au long du développement érythrocytaire par RT-PCR semi-quantitative (anneaux, jeunes et vieux trophozoïtes, de même que jeunes et vieux schizontes) et comparée entre deux clones de *P. falciparum*: 3D7 et F12. Les stades de développement du parasite ont été collectés, après trois synchronisations successives au sorbitol, en fonction du temps et d'une analyse



morphologique de frottis colorés au GIEMSA. L'intégrité des différentes préparations d'ARN a été vérifiée par coloration au bromure d'éthidium (Figure 15A) et par le bioanalyseur d'Agilent (Figure 15B). La spécificité des produits de RT-PCR a été confirmée par hybridation avec un oligonucléotide interne à la séquence (données non présentées). Pour chaque expérience, deux RT-PCR témoins en absence d'ARN et de transcriptase inverse ont été réalisées (données non présentées). De plus, des expériences de RT-PCR réalisées avec des amorces spécifiques du gène MAL13P1.76, situées autour d'un intron, nous ont permis d'exclure une contamination par de l'ADN génomique dans les préparations d'ARN total traitées à la DNase I (données non présentées).



**Figure 16. Analyse de l'expression temporelle des ARNm de *pfmyb1*.**

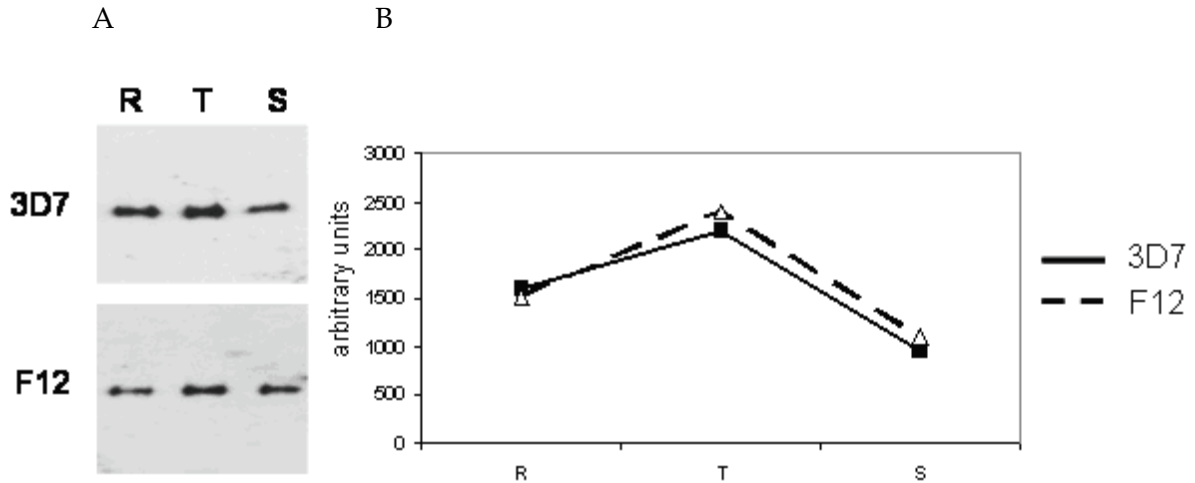
L'ARN poly A<sup>+</sup> a été purifié à partir d'un mélange de 35 µg d'ARN total issu des différents stades de développement du clone 3D7, dont la coloration au bromure d'éthidium est présentée en ligne 1 (A) et dont le profil Agilent montre un ratio *28S/18S* de 1,86 (B). Le Northern Blot hybridé avec une ribosonde de *pfmyb1* est présenté en ligne 2 (A). Les amplicons de *pfmyb1* et du gène de l'ARN ribosomal *18S* (indiqués sur la gauche) ont été obtenus après extraction d'ARN total à différents temps du cycle érythrocytaire : anneaux (R), jeunes (ET) et vieux trophozoïtes (LT), jeunes (ES) et vieux schizontes (LS). Le marqueur de taille est présenté sur la droite (C). Représentation graphique des intensités des fragments amplifiés après normalisation avec l'ARN ribosomal *18S*. Les figures 2C et 2D sont représentatives d'une expérience sur trois réalisées (D).

Les profils de RT-PCR ont été déterminés après vérification des conditions utilisées, afin que ces expériences soient semi-quantitatives (températures d'hybridation des amorces, concentration des différents composants et nombre de cycles pour l'amplification des deux séquences *pfmyb1* et *18S*). La cinétique d'expression du transcrit *pfmyb1* a été suivie en normalisant le fragment de 634 pb provenant de l'ARNm de *pfmyb1* (Figure 15C) par l'intensité des l'amplicons du gène de l'ARN ribosomal *18S* (Figure 15D). Ces expériences montrent que l'abondance du messenger de *pfmyb1* est régulée au cours du développement érythrocytaire. Le profil d'expression du transcrit *pfmyb1* est similaire entre les clones 3D7 et F12 avec un maximum au stade jeune trophozoïte et diminue de manière équivalente lors du passage au stade schizonte. Cependant, le niveau d'expression au stade anneau est nettement plus fort dans le clone 3D7.

### **c) Etude de la cinétique d'expression de la protéine PfMyb1**

Le profil d'expression de la protéine PfMyb1 a été déterminé après immunoprécipitation et Western Blot en série avec le sérum anti-PfMyb1 et à partir d'extraits nucléaires provenant des clones 3D7 et F12 (Figure 16A). Les extraits nucléaires ont été préparés à partir de cultures synchronisées et pour trois points du développement au cours du cycle érythrocytaire : anneaux, trophozoïtes et schizontes. Plusieurs immunoprécipitations, avec le sérum pré-immun ou avec des billes magnétiques seules, suivies d'un Western blot avec le sérum anti-PfMyb1, n'ont pas donné de signal (données non présentées).

Une quantification de l'intensité donnée par la bande représentant la protéine PfHsp70 nous a permis de nous assurer que, pour chaque point de la cinétique, une quantité équivalente de protéines avait été déposée pour le SDS-PAGE (données non présentées). La cinétique d'expression de la protéine PfMyb1 semble très similaire pour les deux clones 3D7 et F12 (Figure 16B). Le niveau d'expression est relativement plus élevé au stade anneau qu'au stade schizonte. Le pic d'expression de la protéine est observé au stade trophozoïte. Dans cette première expérience, la comparaison entre les cinétiques d'expression du transcrit et de la protéine révèle une certaine proximité entre les deux clones, sauf pour le stade anneau du clone F12 où le niveau du transcrit apparaît inférieur à celui de 3D7.



**Figure 17. Analyse de l'expression temporelle de la protéine PfMyb1.**

(A) Immuno-précipitation et Western Blot en série de la protéine PfMyb1. Les extraits nucléaires de stades anneau (R), trophozoïte (T) et schizonte (S) ont été soumis à une immuno-précipitation utilisant des billes magnétiques recouvertes de sérum anti-PfMyb1, suivie d'un Western Blot avec le sérum anti-PfMyb1. (B) Représentation graphique après normalisation de l'intensité donnée par les bandes révélées par un Western Blot utilisant un anticorps anti-PfHsp70. Les figures 16A et 16B sont représentatives d'une expérience sur deux réalisées.

Ces résultats sont à mettre en perspective avec ceux obtenus lors des cinétiques d'interaction avec la séquence d'ADN contenant un MRE, par la technique du retardement sur gel, et pour les clones 3D7 et F12 (Fig. 2, p. 161 de l'Article 2). Tout d'abord, il semble que la cinétique d'expression de la protéine PfMyb1, pour le clone 3D7, soit très proche des résultats obtenus avec la cinétique d'interaction. Cette donnée vient renforcer l'hypothèse que PfMyb1 joue un rôle central dans la détermination de l'interaction avec un MRE pour le clone 3D7. Cependant, les résultats du clone F12 montre une différence entre le niveau d'expression de la protéine analysée par Western blot du lysat total et la cinétique d'interaction obtenue par retardement sur gel. Ces résultats suggèrent que, non pas le facteur PfMyb1, mais un partenaire serait absent dans l'extrait nucléaire du clone F12 aux stades anneaux et schizontes. L'absence du partenaire ou encore l'absence d'une modification post-traductionnelle pouvant perturber la fonction du facteur, pourrait expliquer l'absence de bandes retardées avec les extraits nucléaires d'anneaux et de schizontes.

## II.2 – Rôle de PfMyb1 lors du cycle érythrocytaire et début de caractérisation de son réseau de régulation

### II.2.1 Résumé

Après avoir validé la fonction de la protéine PfMyb1, par son activité de liaison spécifique à l'ADN en interagissant avec les séquences MRE, nous nous sommes intéressés à son rôle dans la régulation de l'expression des gènes au cours du cycle érythrocytaire. Pour cela, nous avons utilisé des ARNdb, spécifiques du gène *pfmyb1*, pour induire une baisse de l'expression du facteur PfMyb1 et ainsi tenter d'identifier des gènes cibles de ce FT.

Le traitement des cultures de parasites par l'ARNdb *pfmyb1* entraîne une inhibition de la croissance de 40 % par rapport à une culture témoin analysée 48 h plus tard par trois techniques différentes. Ceci est dû à la mort des parasites lors du passage du stade trophozoïte au stade schizonte. Cette baisse de la parasitémie n'a pas lieu lors du traitement des parasites par un ARNdb spécifique d'un gène du VIH-1 (Fig. 1, p. 31 de l'Article 3).

L'expression du gène *pfmyb1*, au niveau de son transcrite comme au niveau de sa protéine, est fortement diminuée dans les cultures traitées par l'ARNdb *pfmyb1* en comparation de la culture témoin, au moment de son pic d'expression au stade trophozoïte (Fig. 2 et 3, p. 32-33 de l'Article 3).

Pour analyser les conséquences de l'inactivation partielle de l'expression de PfMyb1, nous avons examiné la modulation de l'expression des gènes avec la bio-puce thématique. Les résultats obtenus ont permis d'identifier huit gènes dont l'expression est altérée dans les cultures traitées par l'ARNdb *pfmyb1* par rapport aux cultures témoin non traitée ou traitée par un ARNdb *gp120* non pertinent (Tableau 1, p. 35 de l'Article 3). La confirmation de ces modifications d'expression nous a permis de désigner ces gènes comme étant directement ou indirectement régulés par PfMyb1. Afin de mieux comprendre les modalités d'action de ce FT, nous avons analysé les interactions de PfMyb1 avec les promoteurs par immunoprécipitation de la chromatine au stade trophozoïte. Ces expériences ont montré que le FT PfMyb1 participait directement à la régulation négative du gène de la kinase PfPK5 ou positive des gènes PCNA, PGK, H3, H2A et PP1-like (Fig. 7, p. 36 de l'Article 3). A l'instar des protéines Myb eucaryotes, PfMyb1 intervient directement dans la régulation de

l'expression de gènes clés impliqués dans le cycle cellulaire et joue un rôle essentiel dans le développement érythrocytaire de *P. falciparum*.

**ARTICLE 3:**

**PfMyb1, a *Plasmodium falciparum* transcription factor, is required for intra-erythrocytic growth and controls key genes for cell cycle regulation.**

Mathieu Gissot, Sylvie Briquet, Philippe Refour, Charlotte Boschet et Catherine Vaquero.

Accepté à la publication dans *Journal of Molecular Biology*.

# PfMyb1, a *Plasmodium falciparum* Transcription Factor, is Required for Intra-erythrocytic Growth and Controls Key Genes for Cell Cycle Regulation

Mathieu Gissot, Sylvie Briquet, Philippe Refour, Charlotte Boschet and Catherine Vaquero\*

INSERM U511, CHU  
Pitié-Salpêtrière, 91 boulevard  
de l'Hôpital, 75013 Paris  
France

During the complex life cycle of *Plasmodium falciparum*, divided between mosquito and human hosts, the regulation of morphologic changes implies a fine control of transcriptional regulation. Transcriptional control, however, and in particular its molecular actors, transcription factors and regulatory motifs, are as yet poorly described in *Plasmodium*. In order to decipher the molecular mechanisms implicated in transcriptional regulation, a transcription factor belonging to the tryptophan cluster family was studied. In a previous work, the PfMyb1 protein, contained in nuclear extracts, was shown to have DNA binding activity and to interact specifically with *myb* regulatory elements. We used long *pfmyb1* double-stranded RNA (dsRNA) to interfere with the cognate messenger expression. Parasite cultures treated with *pfmyb1* dsRNA exhibited a 40% growth inhibition when compared with either untreated cultures or cultures treated with unrelated dsRNA, and parasite mortality occurred during trophozoite to schizont transition. In addition, the *pfmyb1* transcript and protein decreased by as much as 80% in treated trophozoite cultures at the time of their maximum expression. The global effect of this partial loss of transcript and protein was investigated using a thematic DNA microarray encompassing genes involved in signal transduction, cell cycle and transcriptional regulation. SAM software enabled us to identify several genes that were differentially expressed and probably directly or indirectly under the control of PfMyb1. Using chromatin immunoprecipitation, we demonstrated that PfMyb1 binds, within the parasite nuclei, to several promoters and therefore participates directly in the transcriptional regulation of the corresponding genes. This study provides the first evidence of a regulation network involving a *Plasmodium* transcription factor.

© 2004 Elsevier Ltd. All rights reserved.

**Keywords:** *Plasmodium*; gene regulation; transcription; transcription factor; PfMyb1

\*Corresponding author

Present address: P. Refour, Department of Immunology and Infectious Diseases, Harvard School of Public Health, 665 Huntington Ave., Boston, MA 02115, USA.

Abbreviations used: asRNA, anti-sense RNA; cDNA, DNA complementary to RNA; ChIP, chromatin immunoprecipitation; dsRNA, double-stranded RNA; DNase, deoxyribonuclease; dNTP, deoxyribonucleoside triphosphate; MRE, Myb regulatory element; PCR, polymerase chain reaction; qPCR, real-time quantitative PCR; RNase, ribonuclease; RT, reverse transcription; sRNA, sense RNA.

E-mail address of the corresponding author:  
vaquero@ext.jussieu.fr

## Introduction

*Plasmodium* is responsible for 1.5–2.7 million deaths annually,<sup>1</sup> mostly among children and pregnant women. Of the four species of *Plasmodium* infecting humans, *Plasmodium falciparum* causes the highest morbidity and mortality. As global efforts to eradicate malaria have failed, there is an urgent need to decipher the biology of *Plasmodium* and in particular the mechanisms of gene regulation that govern its developmental cycle, so as to propose novel strategies to fight malaria.<sup>2</sup>

The cell cycle of *Plasmodium* between vertebrate

and mosquito hosts implies a high degree of adaptation and strict control of the cellular machinery to correspond precisely to the state of differentiation in the host. These different developmental stages require coordinated modulation of expression of distinct sets of genes, which could be achieved by transcriptional and/or post-transcriptional controls. In *Plasmodium*, as in all eukaryotes, gene expression is governed at the level of transcription by the interaction of elements within promoters acting *in cis* (DNA regulatory elements) and elements acting *in trans* (transcription factors) whose availability is modulated during cellular development. The fine-tuning of transcriptional regulation resides in the interplay between the *cis* and *trans* elements throughout *P. falciparum* development. Little, however, is currently known about these two actors in *Plasmodium*.

The study of *Plasmodium* chromosomes 2 and 3,<sup>3,4</sup> and recently of the complete 3D7 clone sequence,<sup>5</sup> has revealed a genomic organization composed of rather short intergenic sequences, encompassing the basal and distal promoter sequences,<sup>6</sup> between successive open reading frames. The messengers thus far investigated, while few in number, appear to resemble eukaryotic transcripts. They are capped and bear canonic poly-adenylation signals.<sup>7</sup> Moreover, the nucleosome organization,<sup>8</sup> the structure of the promoters and the general transcription factors<sup>9-13</sup> are quite similar to those of other eukaryotes. The regulatory sequences within the promoters appear to contain motifs similar to several binding sites encountered in eukaryotes.<sup>6,14,15</sup> Nevertheless, the majority of the regulatory elements reported so far have not been linked to any known transcription factor<sup>16-18</sup> or found to bind to unknown nuclear factors.<sup>19</sup>

In order to decipher the molecular mechanisms implicated in transcriptional regulation during the erythrocytic development, we have identified, among diverse families of transcription factors, a Myb-related protein belonging to the tryptophan cluster family. In eukaryotes, the Myb proteins are highly conserved throughout evolution and generally contain three repeats of approximately 50 residues with three regularly spaced tryptophan residues.<sup>20,21</sup> Indeed, a large number of Myb-related proteins have been shown to bind DNA in a sequence-specific manner and to participate in the regulation of the expression of genes implicated in growth control and differentiation.<sup>21</sup> PfMyb1 was the first transcription factor not belonging to general transcription factors to be cloned and analyzed in *P. falciparum*.<sup>22</sup> This study showed that *pfmyb1* mRNA reached its highest level at the trophozoite stage, at least as regards the 3D7 clone. In *in vitro* experiments, the PfMyb1 protein present in nuclear extracts was able to bind specifically to a prototype *myb*-regulatory element (MRE) from the chicken *mim-1* gene promoter<sup>23</sup> and two putative MRE annotated within the *Plasmodium* promoters,<sup>22</sup> suggesting that PfMyb1 can be considered to be a Myb transcription factor.

We took advantage of long *pfmyb1* double-stranded RNA (dsRNA) to interfere with expression of the cognate messenger and to address the role of this transcription factor during the erythrocytic cycle. In particular, we investigated the consequences of decrease in *pfmyb1* transcript and protein on parasite growth and transcriptional regulation of *P. falciparum* genes.

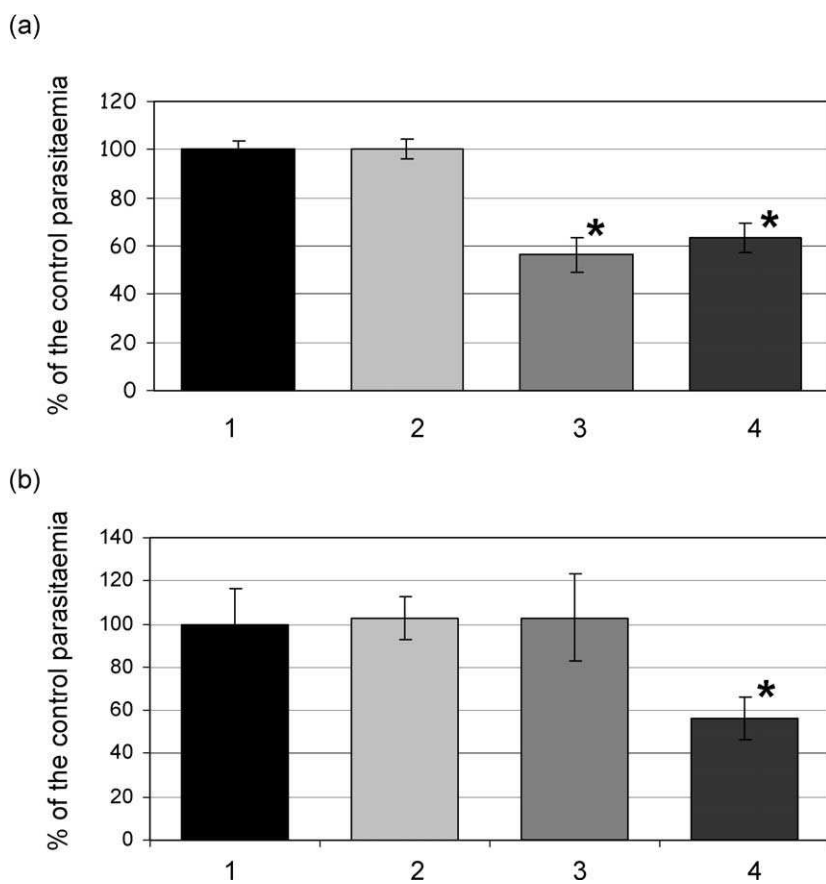
## Results

### *pfmyb1* dsRNA alters parasite growth and trophozoite to schizont transition

We investigated the global effect of double-stranded RNA corresponding to part of the *pfmyb1* sequence on *P. falciparum* growth when added to highly synchronized parasite cultures during a 48 hour assay. Parasites were considered to be synchronized after three successive sorbitol treatments leading to an enrichment in rings of at least 95%. In young ring cultures, *pfmyb1* dsRNA or dsRNA of an unrelated gene (HIV-1 *gp120* AN: K02013) was added (25 µg per ml of culture) and parasitaemia was evaluated using three different methods (Figure 1). As shown in Figure 1(a), addition of unrelated dsRNA to culture media (line 2) did not affect asexual growth when compared with control cultures (line 1), i.e. parasites grown in media supplemented with identical volumes of annealing buffer. In contrast, a statistically significant decrease in parasitaemia (approximately 40%) was observed in cultures treated with *pfmyb1* dsRNA using manual counting (Figure 1(a), line 3) or a [<sup>3</sup>H]hypoxanthine assay (Figure 1(a), line 4). The result obtained by manual counting (line 3) corresponds to the average of nine experiments performed in triplicate, while the result obtained by [<sup>3</sup>H]hypoxanthine assay (line 4) corresponds to two experiments carried out in triplicate. Finally, using a flow-cytometry assay, a similar degree of growth inhibition (40%) was observed after treatment with *pfmyb1* dsRNA as compared with controls treated with single-stranded sense or anti-sense *pfmyb1* RNA (data not shown). Thus three different experimental approaches used for evaluating the inhibition of growth upon addition of dsRNA specific for *pfmyb1* gave essentially the same results (Figure 1).

We investigated the effect of *pfmyb1* dsRNA on trophozoite to schizont transition (Figure 1(b)). During the erythrocytic cycle, smears were performed every four hours and the number of parasites in trophozoite (Figure 1(b), lines 1 and 2) and in late schizont stage (Figure 1(b), lines 3 and 4) was evaluated. The parasitaemia observed in the trophozoite cultures was quite similar in the control (Figure 1(b), line 1) and *pfmyb1* dsRNA-treated cultures (Figure 1(b), line 2), in contrast to that of late schizonts showing a significant decrease in *pfmyb1* dsRNA-treated parasites (Figure 1(b), line 4 versus line 3). Moreover, the growth inhibition





**Figure 1.** Parasite growth and death evaluation during trophozoite to schizont transition. (a) Evaluation of parasite growth 48 hours after addition of dsRNA to the medium. Parasitaemia was either counted on GIEMSA-stained blood smears (1–3) or monitored by  $^3\text{H}$ -Hypoxanthine uptake (1 and 4). Parasitaemia was evaluated in untreated culture (1), unrelated dsRNA (2) and *pfmyb1* dsRNA treated culture (3 and 4). Asterisks indicate statistically significant values ( $p$ -value  $< 0.0001$ ). Parasitaemia was evaluated in nine independent experiments on Giemsa-stained blood smears and two independent experiments as regards [ $^3\text{H}$ ]hypoxanthine uptake assay. (b) Parasitaemia of highly synchronized cultures was determined by counting at late trophozoite (1 and 2) and at late schizont stage (3 and 4) in untreated culture (1 and 3) and *pfmyb1* dsRNA-treated culture (2 and 4). Asterisks are as in (a). The experiment was repeated four times in triplicate. The results are expressed as a percentage of the control cultures.

(Figure 1(a)) and the decrease in schizont number (Figure 1(b)) were of an essentially similar magnitude.

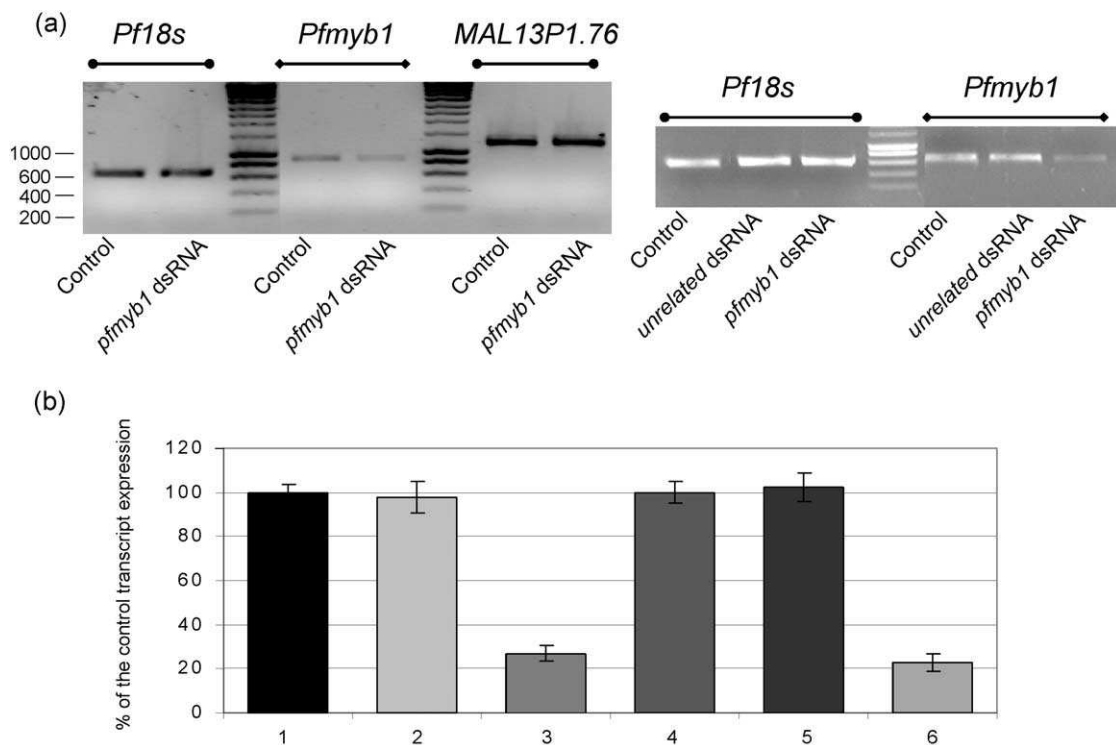
#### Decreased expression of *pfmyb1* messenger is observed in *pfmyb1* dsRNA-treated cultures

The expression of the *pfmyb1* transcript was analysed at the trophozoite stage, at the peak of expression in the erythrocytic cycle, in cultures treated with *pfmyb1* or *gp120* dsRNA, or left untreated, using a semi-quantitative (Figure 2(a) and (b)) and real-time quantitative reverse transcription (RT)-PCR (Figure 2(b)). For each set of experiments, PCR control reactions were performed in the absence of RNA or RT (data not presented), as well as on the MAL13P1.76 gene with two primers encompassing an intron (Figure 2(a) left panel). These data demonstrated the absence of DNA contamination in the RNA preparations. First, by semi-quantitative RT-PCR (Figure 2(a), right and left panel), we showed that the expression of the *pfmyb1* transcript was decreased in *pfmyb1* dsRNA-treated parasites when compared with control parasites, while expression of the MAL13P1.76 transcript was unaltered. After normalization against ribosomal 18 S RNA, a 75% decrease in the *pfmyb1* transcript was observed in *pfmyb1* dsRNA-treated cultures (Figure 2(b), lines 3 and 6) when compared with control cultures (lines 1 and 4) or cultures treated with unrelated dsRNA (lines 2 and 5),

using semi-quantitative (lines 1–3) or real-time quantitative RT-PCR (RT-qPCR) (lines 4–6). RT-qPCR was performed using a multiplex assay amplifying *pfmyb1* and the 18S reference in the same experiment. For the experiments presented herein, different couples of primers (see Materials and Methods) were used, all yielding similar results.

#### Decreased expression of PfMyb1 protein is observed in *pfmyb1* dsRNA-treated culture

In order to follow the effect of *pfmyb1* dsRNA treatment on PfMyb1 protein expression, serial immuno-precipitation and Western blotting experiment were performed. Protein extracts from *P. falciparum* trophozoites were immuno-precipitated using magnetic beads coated with anti-PfMyb1 serum. Immuno-precipitated proteins were subjected to SDS-PAGE followed by Western blotting with the anti-PfMyb1 serum. No bands were detected after immuno-precipitation of lysates with either preimmune serum or uncoated beads and Western blotting with the anti-PfMyb1 serum (data not shown). The abundance of the PfMyb1 protein was markedly decreased in *pfmyb1* dsRNA-treated parasites (Figure 3(a), lower panel) as confirmed by normalization with PfHsp70 (Figure 3(a), upper panel). Indeed, a reduction of approximately 80% in the PfMyb1 protein was observed in protein extracts from trophozoite cultures treated with *pfmyb1* dsRNA when



**Figure 2.** Analysis and quantification of *pfmyb1* mRNA expression. (a) Amplicons of *pfmyb1*, 18 S ribosomal RNA and MAL13P1.76 (indicated on the top) were obtained after semi-quantitative RT-PCR on DNase-treated total RNA prepared from either untreated culture (control), irrelevant *gp120* dsRNA (unrelated dsRNA) or *pfmyb1* dsRNA (*pfmyb1* dsRNA) treated culture. Two independent experiments are presented. Markers are indicated on the left side of the Figure. (b) Graphical representation of the *pfmyb1* amplified fragment intensities after normalization with 18 S ribosomal RNA. Semi-quantitative (1–3) and real time quantitative RT-PCR (4–6) were performed on samples from either untreated culture (1 and 4), or cultures treated with irrelevant (2 and 5) or *pfmyb1* dsRNA (3 and 6). The results are expressed as a percentage of the control transcript expression in three independent experiments.

compared to untreated cultures (Figure 3(b), lane 2 versus lane 1), while no such reduction was observed in controls treated with unrelated dsRNA (lane 3).

### Genes are differentially expressed in *pfmyb1* dsRNA-treated culture

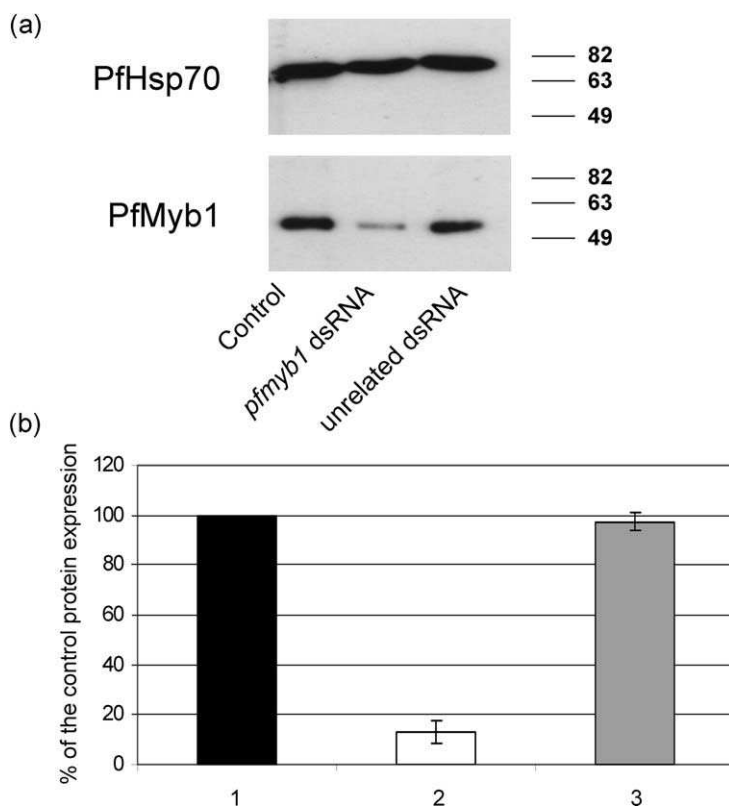
We then decided to investigate the effect of the partial loss of the PfMyb1 transcript and protein on the global expression profile, by comparing untreated culture with cultures treated with either *pfmyb1* dsRNA or irrelevant dsRNA at a time when the expression of the *pfmyb1* transcript was highest, i.e. the trophozoite stage.

With the use of a thematic microarray fully described by Gissot *et al.*<sup>24</sup> and comprising 180 PCR products printed in duplicate, among which 153 amplified from the *P. falciparum* genomic sequences, we monitored the steady state of the corresponding transcripts. The genes were selected for their homology to known proteins previously described in eukaryotes to be involved in genetic regulation from DNA to proteins. The microarray quality controls were performed as detailed by Gissot *et al.*<sup>24</sup> Briefly, for each growth condition, at least two different total RNA preparations from

independent cultures were used for synthesis of cDNA radiolabelled with <sup>3</sup>H or <sup>35</sup>S radioisotope, and for subsequent hybridization of two sets of arrays. Each raw signal was normalized twice, first with the Alien cDNAs in order to normalize for the efficiency of each reverse transcription and second with 3D7 genomic DNA to normalize for the amount of each cDNA used for the hybridization. Statistically significant differences in gene expression were monitored using the Statistical Analysis for Microarrays (SAM) program.<sup>25</sup> Among the 153 *P. falciparum* genes printed on the microarray, 139 genes gave signals at least twice as high as background (see Supplementary Data 1 and 2).

When untreated controls were compared with HIV-1 *gp120* dsRNA-treated culture, no differences in the relative expression level of any gene were detected using the SAM program (see Supplementary Data 2). In contrast, upon comparison of *pfmyb1* dsRNA treated and untreated parasites, 11 genes were identified as statistically different and the 128 remaining genes as similar. Among the 11 genes, one was up-regulated and ten down-regulated, as shown in the SAM output plot presented in Figure 4.

The differential expression of these 11 genes was



**Figure 3.** Analysis and quantification of PfMyb1 protein expression. (a) Serial immunoprecipitation and Western blotting of *P. falciparum* PfMyb1. The protein extracts of trophozoite parasites were first subjected to immuno-precipitation using beads coated with an anti-PfMyb1 serum followed by Western blot analysis with the same anti-PfMyb1 serum (lower panel). Concomitantly a Western blot analysis with an anti-PfHSP70 serum was performed (upper panel). The protein markers are indicated on the right of the Figure. Samples from either untreated culture (control), or cultures treated with irrelevant (unrelated dsRNA) or *pfmyb1* dsRNA (*pfmyb1* dsRNA) were studied. (b) Graphical representation of the PfMyb1 protein intensities after normalization with the intensities of the PfHSP70 protein: Lane 1, untreated culture; lane 3, irrelevant dsRNA-treated culture; lane 2, *pfmyb1* dsRNA (*pfmyb1* dsRNA) treated culture. One representative experiment among three is presented in (a). The results are expressed as for Figure 2.

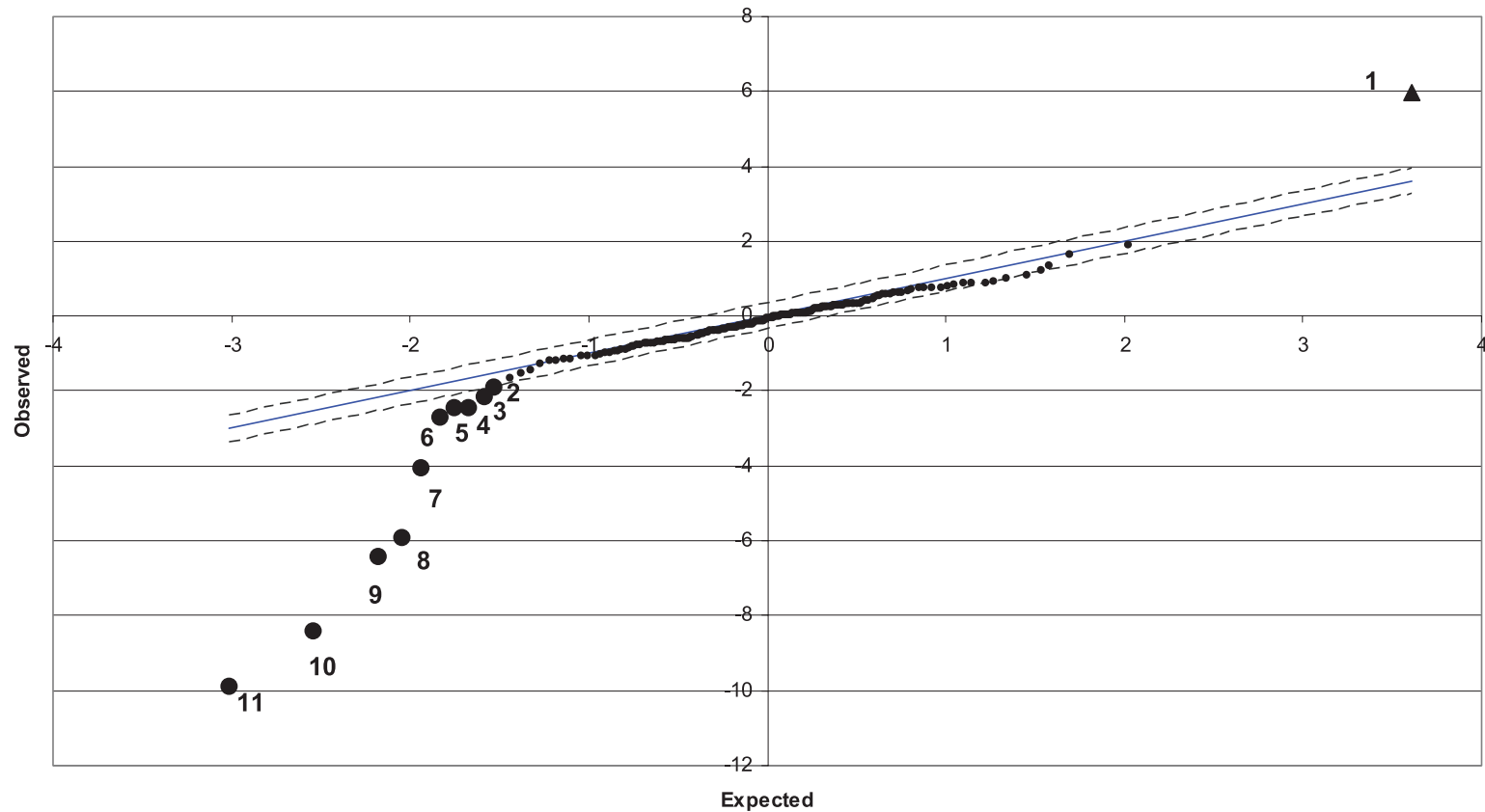
verified by real-time quantitative RT-PCR using the 18 S gene as reference. Using two independent biological samples, three (numbered 2, 3 and 4 in Figure 4) of the ten down-regulated genes did not give sufficiently consistent results to be taken into account. The results of the RT-qPCR (filled bars) and microarray (open bars) analyses, for the genes whose expression was modulated by *pfmyb1* dsRNA (seven down and one up-regulated gene, also listed in Table 1) and diverse additional genes as controls of the experiments, are compared in Figure 5.

Indeed, several genes whose expression profile was similar in control and *pfmyb1* dsRNA-treated parasites were selected and their expression compared by DNA microarray or RT-qPCR experimental approaches. PfHSP70 (PF08\_0054) and MAL13P1.76 were shown to be equally expressed in treated and untreated cultures when analysed by either experimental approach. This was also the case for PF10\_0233, PFD0465w and MAL8P1.154 genes, which showed slight but detectable homology with the *pfmyb1* sequence used to generate *pfmyb1* dsRNA by using the BLASTN tool on the PlasmoDB web site, as well as for *pfmyb2* (PF10\_0327) and *pfmyb3* (PF10\_0143), belonging to the tryptophan cluster family of transcription factors. Moreover, the RT-qPCR also confirmed the results obtained for the seven down and one up-regulated gene, attesting to the reliability of the microarray data (Figure 5 and Table 1).

Finally, the over-expression of the *pfpk5* transcript in *pfmyb1* dsRNA-treated samples was also established at the level of protein expression (Figure 6). Western blotting experiments on whole trophozoite protein extracts showed an increased representation of PfPfk5 protein in *pfmyb1* dsRNA-treated samples as compared with untreated samples, in good agreement with the transcript data.

### PfMyb1 is associated with the promoter of genes differentially expressed using ChIP experiments

In order to determine whether, within the parasite nuclei, the PfMyb1 protein could interact with promoters of genes listed in Table 1, and therefore could be involved in the regulation of the targeted genes, chromatin immuno-precipitation (ChIP) experiments were carried out. PfMyb1 protein present in nuclear extracts has been shown to interact with Myb-related elements (MRE).<sup>22</sup> A computational analysis was performed on the promoters of the genes listed in Table 1 to identify these MRE, and revealed that among the eight promoters only that of PGK (PFI1105w) did not contain any putative MRE. The PfMyb1-DNA complexes were cross-linked within the parasite nuclei and subsequently immuno-precipitated by using the anti-PfMyb1-coated beads. No cross-linked PfMyb1-DNA complexes were immuno-precipitated using either beads alone or beads coated with preimmune serum (data not shown).



**Figure 4.** Graphical output of SAM software. Genes are plotted where the *x*-axis represents the expected distribution of each PCR product intensity and the *y*-axis observed distribution of each PCR product intensity. Genes identified by SAM as up-regulated in *pfmyb1* dsRNA-treated samples or as down-regulated in *pfmyb1* dsRNA-treated samples are represented by triangles or circles, respectively, and are as follows: (1) PfPk5 (MAL13P1.279); (2) PF14\_0366; (3) PFC0385c; (4) PFE0420c; (5) PfPGK (PFI1105w); (6) PfPk2 (PFL1885c); (7) PfTBP (PFL2345c); (8) PfPCNA1 (PFL1285c); (9) PP1-like (PF14\_0224); (10) PfHistone H3 (MAL6P1.248); and (11) PfHistone H2A (MAL6P1.249).

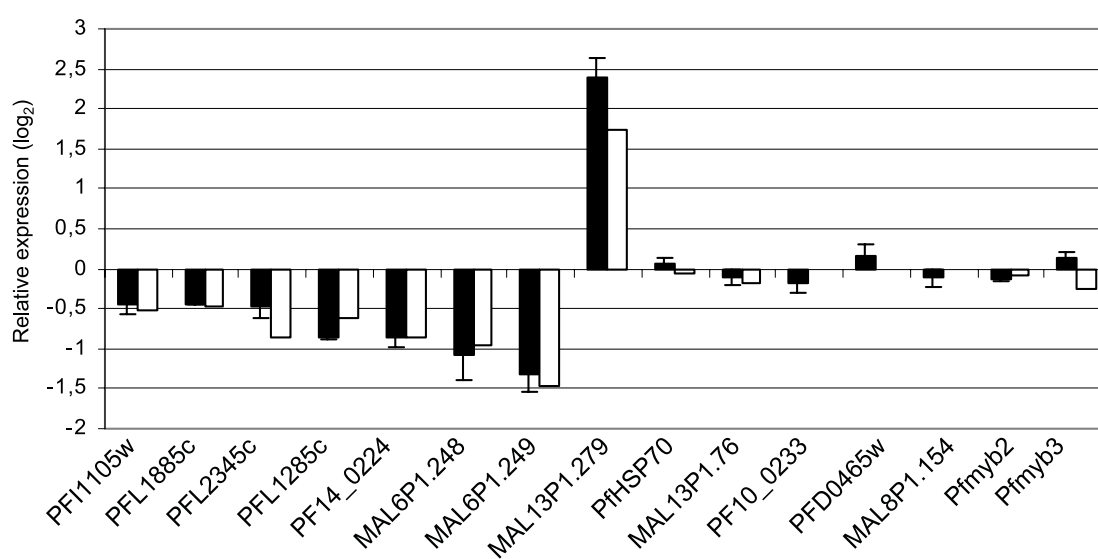
**Table 1.** List of SAM-identified genes verified by qPCR

Gene name	AN <sup>a</sup>	Putative function	Number <sup>b</sup>	Microarray results <sup>c</sup>
PGK	PFI1105w	Phosphoglycerate kinase	5	▼
PfPfk2	PFL1885c	Calcium dependent kinase	6	▼
PfTbp	PFL2345c	TATA binding homologue	7	▼
PCNA	PFL1285c	Proliferating cell nuclear antigen	8	▼
PP1-like	PF14_0224	Phosphatase	9	▼
Histone H3	MAL6P1.248	Histone	10	▼
Histone H2A	MAL6P1.249	Histone	11	▼
PfPfk5	MAL13P1.279	Cyclin dependent kinase	1	▲

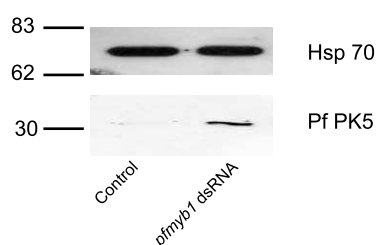
<sup>a</sup> Accession number according to PlasmoDB annotation.

<sup>b</sup> Genes are numbered as in Figure 4.

<sup>c</sup> Gene exhibiting an altered expression microarray experiments; ▲ stands for up-regulated and ▼ stands for down-regulated.



**Figure 5.** Comparison of microarray and RT-qPCR data. Microarray data are indicated by open bars. RT-qPCR data are indicated by filled bars. Relative transcriptional level of each gene (accession number indicated on bottom) is expressed as a log<sub>2</sub> ratio of *pfmyb1* dsRNA-treated samples over untreated samples.

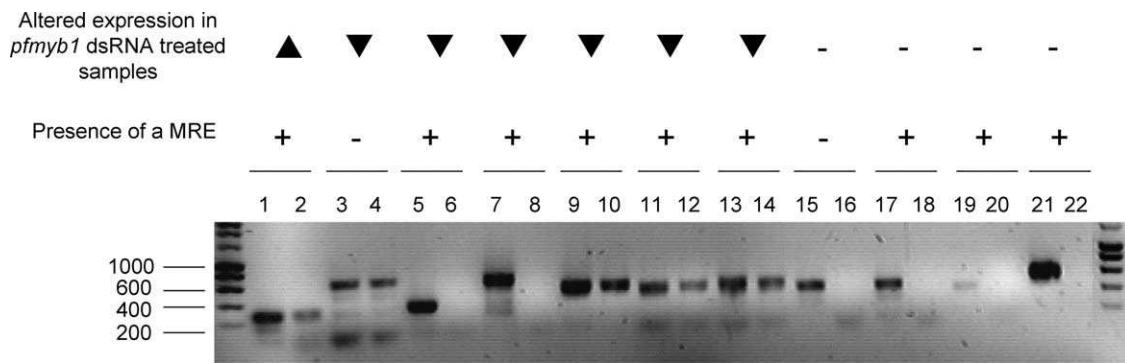


**Figure 6.** PfPfk5 expression in *pfmyb1* dsRNA-treated samples. Trophozoite protein extracts were subjected to SDS-10% PAGE followed by Western blotting with either anti-PfHSP70 serum (upper panel) or anti-PfPfk5 serum (lower panel) as indicated on the right side of the Figure. The protein markers are indicated on the left side of the Figure. Samples of untreated culture (control) and *pfmyb1* dsRNA-treated culture (*pfmyb1* dsRNA) were investigated. The Figure presents one representative experiment among three.

After purification of the DNA fragments, PCR analysis was carried out for the eight previously selected promoters. In Figure 7, odd numbers correspond to PCR performed on input DNA, and even numbers to immuno-precipitated DNA.

Promoters of three genes containing MRE (PF08\_0054, PF10\_0159, PF11\_0096), but whose expression was not altered in parasites cultured in the presence of *pfmyb1* dsRNA in microarray experiments, were amplified in input (Figure 7 lanes 17, 19 and 21) but not ChIP samples (Figure 7 lanes 18, 20 and 22), even though expression of their cognate transcript peaked at the trophozoite stage.<sup>26</sup> In addition, the promoter of the PF07\_0073 gene, devoid of MRE, yielded identical results (Figure 7 lanes 15–16).

In contrast, among the eight genes whose expression was altered in the presence of *pfmyb1* dsRNA, six (MAL13P1\_279, PFI1105w, PFL1285c, MAL6P1.248, MAL6P1.249 and PF14\_0224) gave rise to amplification products with both anti-PfMyb1 immuno-precipitated (Figure 7 lanes 2, 4, 10, 12 and 14) and input DNA (Figure 7 lanes 1, 3, 9,



**Figure 7.** Chromatin immuno-precipitation (ChIP) using an anti-PfMyb1 serum. After DNA/PfMyb1 cross-linking within the parasite nuclei and immuno-precipitation of DNA using anti-PfMyb1 antibody, PCR was carried out with input DNA (odd numbers) and immuno-precipitated DNA (even numbers) using primers amplifying promoters of the following genes: MAL13P1\_279 (lane 1 and 2), PFI1105w (lane 3 and 4), PFL1885c (lane 5 and 6), PFL2345c (lane 7 and 8), PFL1285c (lane 9 and 10), MAL6P1.248 and MAL6P1.249 (lane 11 and 12), PF14\_084 (lane 13 and 14), PF07\_073 (lane 15 and 16), PF11\_0096 (lane 17 and 18), PF08\_0054 (lane 19 and 20) and PF10\_0159 (lane 21 and 22). Amplicons were fractionated on 1.2% agarose gel. Genes identified in previous experiments as displaying up-regulated (▲), down-regulated (▼) or unaltered expression (-) are indicated. The presence (+) or absence (-) of Myb-related elements (MRE) in the promoter of the gene is indicated. **Figure 7** portrays one representative experiment among three.

11 and 13). These six genes corresponded to five promoters, since the genes encoding histones H3 and H2A (MAL6P1.248 and MAL6P1.249) share a bidirectional promoter. Finally, the promoters of two (PFL1885c and PFL2345c) of these eight genes gave no amplification products using ChIP (**Figure 7** lanes 6 and 8), in contrast to input samples (**Figure 7** lanes 5 and 7).

## Discussion

In light of the crucial role of transcriptional regulation during the complex life cycle of *P. falciparum*, we decided to investigate PfMyb1, a protein belonging to the tryptophan cluster family. Indeed, previous experiments have shown that this protein interacts specifically with several MRE.<sup>22</sup> These data, together with computational study of the PfMyb1 amino acid sequence,<sup>22</sup> have led us to consider this protein to be a genuine transcription factor.

In order to investigate the role of PfMyb1 in transcriptional regulation of *P. falciparum*, we decided to interfere, by means of long *pfmyb1* dsRNA, with its transcript and therefore its protein expression at the time of its highest expression, i.e. during the trophozoite stage.<sup>22,26,27</sup> As shown in **Figure 1(a)**, a 40% growth inhibition was observed when parasite cultures were treated with *pfmyb1* dsRNA as compared with either untreated parasite cultures or cultures treated with unrelated long dsRNA. The observed inhibition thus appeared to depend specifically on *pfmyb1* dsRNA. This observation was made by using three different methods and confirmed by repeated experiments that consistently yielded similar results. The unrelated dsRNA was chosen for its relatively high AT-content (60%) and its length (570 nt), approximating those of the *pfmyb1* dsRNA

(70% AT and 634 nt, respectively). Moreover, using the BLASTN tool of PlasmoDB, we did not find any detectable homology between this unrelated sequence and the entire *P. falciparum* genome. We further showed that parasite death occurred during the trophozoite to schizont transition, notably after the *pfmyb1* expression peak (**Figure 1(b)**).

Since addition of *pfmyb1* dsRNA modified parasite growth and development, we investigated the level of the *pfmyb1* transcript at the trophozoite stage. By two different experimental approaches based on RT-PCR (**Figure 2**) and performed with different sets of primers, we showed that *pfmyb1* dsRNA treatment diminished the *pfmyb1* transcript in the parasite by 75%. Moreover, the decrease in the *pfmyb1* transcript was correlated to a decrease in the corresponding protein, as determined by serial immuno-precipitation and Western blotting of the trophozoite protein extracts, using anti-PfMyb1 antibodies<sup>22</sup> and after normalization for protein loading. In conclusion, the decreased availability of *pfmyb1* transcript and protein, together with the growth inhibition occurring during the trophozoite to schizont transition, argue in favour of a major function for this transcription factor during the parasite erythrocytic cycle. This assumption is strengthened by the high level of conservation of *pfmyb1* gene within the diverse *Plasmodium* species listed in PlasmoDB.

In *Plasmodium*, the mechanisms by which gene expression is modified by dsRNA are still not clearly defined. The characterization of such mechanisms, whether involving classical RNA interference as suggested by previous work,<sup>28,29</sup> or an anti-sense effect as shown by Noonpakdee *et al.*,<sup>30</sup> is beyond the scope of this study. Herein, we took advantage of long dsRNA to reduce the cognate messenger and encoded protein, so as to study the consequences of this depletion on gene expression. Interestingly, many antisense transcripts have been

shown to arise during the erythrocytic cycle<sup>31,32</sup> and have been suggested to be potential regulatory elements in gene transcription.<sup>33</sup> These studies might imply that the machinery required to process these antisense transcripts and the resulting dsRNA is present in *P. falciparum*.

In order to identify at least some of the transcripts whose expression is governed by the transcription factor, we compared the transcriptome of cultures treated with *pfmyb1* dsRNA and control parasite cultures by means of our thematic DNA microarray. The slides were hybridized with cDNA prepared from total RNA of control and *pfmyb1*-treated cultures radiolabelled as reported.<sup>24</sup> Total RNA was extracted at the trophozoite stage prior to parasite death and when maximum expression of the *pfmyb1* transcript was reached (Figure 1). When unrelated dsRNA-treated samples were compared with untreated controls, no significant differential expression was identified by SAM (see Supplementary Data 2). The decrease in *pfmyb1* transcript and protein mediated by *pfmyb1* dsRNA altered the relative expression of a set of genes in four independent biological experiments normalized and filtered through SAM software (Figure 4).

In particular, among the 153 *P. falciparum* genes printed on the microarray, the level of expression of 128 transcripts was similar in control and in *pfmyb1* dsRNA-treated parasites (see Supplementary Data 1), suggesting that treatment with *pfmyb1* dsRNA did not markedly alter synchronization and development of the parasites. In this regard, upon examination of Giemsa-stained blood smears no substantial shift was detected between the two growth conditions.

In contrast, 11 genes were identified by SAM as differentially expressed in control and in *pfmyb1* dsRNA-treated parasites, of which one was up-regulated and ten down-regulated. All of the microarray data leading to the identification of the genes of interest as well as several control genes were verified by quantitative real-time RT-PCR. Three genes identified by SAM were eliminated upon performing independent biological experiments. It is noteworthy that these genes were very close to the limit defined by SAM (2, 3 and 4 of Figure 4). The remaining eight genes, listed in Table 1, gave similar expression ratios when microarray data were compared with real-time RT-PCR data. In addition, the expression ratio of seven other gene controls (Figure 5, right part) was verified, confirming the reliability of the microarray experiments. The level of expression of *pfmyb2* (PF10\_0327) and *pfmyb3* (PF10\_0143), two genes belonging to the tryptophan cluster family of transcription factors, was not modified by *pfmyb1* dsRNA. This was also the case for three genes showing slight but detectable homology to the *pfmyb1* DNA sequence used to generate dsRNA. Taken together, these data indicate that the activity of *pfmyb1* dsRNA is sequence-dependent.

Among the eight genes cited as differentially expressed in *pfmyb1* dsRNA-treated samples, the

upregulation of PfPk5 (MAL13P1\_279) expression observed at the transcript level was also detected at the protein level by Western blotting (Figure 6). It is worthy of note that PfPk5 has been annotated as a cdc2-related cyclin-dependent kinase bearing 60% identity to the human cdc2 gene.<sup>34</sup> In all eukaryotes, cyclin-dependent kinases are key regulators of cell cycle control and progression.<sup>35</sup> PfPk5 has been implicated in regulation of the nuclear division cycle<sup>36</sup> and is active during the erythrocytic cycle, and shares common features with the cdc-related kinase activation process.<sup>37</sup> Taken together, these data suggest a crucial role for PfPk5 in regulating cell cycle progression of *P. falciparum*. A clear link between the classical cell cycle (G1, S, G2 and M phases) and the *Plasmodium* erythrocytic cell cycle has not been established.<sup>38</sup> Nevertheless, DNA synthesis appears to occur during the trophozoite stage and asynchronous nuclear division during the trophozoite to schizont transition.<sup>39</sup> These two critical phenomena, however, may overlap during trophozoite to schizont transition. Indeed, PfPk5 may play a crucial role during the trophozoite to schizont transition and its over-expression in *pfmyb1* dsRNA-treated culture may explain the growth inhibition and parasite death at that particular time in the erythrocytic cycle.

Furthermore, in *pfmyb1* dsRNA-treated samples, a decreased expression in the gene PFL1285c, a homologue of proliferating cell nuclear antigen (PCNA), was observed. In eukaryotes, PCNA has been shown to be a key player in DNA replication, DNA repair and cell cycle control.<sup>40</sup> In *Plasmodium*, PCNA1 expression was maximum at the trophozoite stage at the time of DNA replication<sup>41</sup> and its altered expression could therefore have played a role in the death of *pfmyb1* dsRNA-treated parasites. Alteration of the glucose metabolism *via* the partial inhibition of expression of the phosphoglycerate kinase (PGK, PFI1105w) might also have participated in the alteration of the parasite life cycle. PGK has been found to be a key enzyme of the glucose pathway and its activity has been detected during the erythrocytic cycle.<sup>42</sup> Of course, all of the observed variations in gene expression might contribute in a cooperative manner to the alteration of parasite development and therefore participate in promoting parasite death. Finally, two histone (H3 and H2A) genes have been shown to be under-expressed in *pfmyb1* dsRNA-treated samples, in contrast to the two other histones (H4 and H2B). The two histone (H3 and H2A) genes appeared to be under the control of a bidirectional promoter. This observation provided another argument in favour of an activity of PfMyb1 as a transcription factor.

In order to determine, among the eight genes identified by microarray, those whose expression might be directly controlled by PfMyb1, we performed a chromatin immuno-precipitation assay with a specific anti-PfMyb1 antibody, followed by a PCR identification of gene promoters. For six genes (MAL13P1\_279, PFI1105w, PFL1285c,

MAL6P1.248, MAL6P1.249 and PF14\_0224), an interaction was observed *ex vivo* between the promoter and the transcription factor PfMyb1, suggesting that they might be directly regulated by PfMyb1. By contrast, two genes (PFL1885c and PFL2345c, see Figure 7, lanes 5–8) were not amplified in ChIP samples. These genes had in fact shown small variation in the level of gene expression in *pfmyb1* dsRNA samples (spots 5 and 6 of Figure 4). The absence of amplification might be attributable to a weak interaction between their potential MRE and PfMyb1, below the threshold of detection of ChIP methodology, or to poor accessibility of the MRE within the chromatin. Alternatively, these genes might be regulated by a transcription factor itself under PfMyb1 control. Similar results were obtained with three control genes (PF08\_0054, PF10\_0159, PF11\_0096) displaying no significant differential expression in microarray experiments and sharing the same peak of expression during the trophozoite stage. Identical results were obtained as regards the PF07\_0073 promoter that lacks MRE (Figure 7).

Finally, the PFI1105w promoter, despite the apparent absence of MRE, gave rise to a positive signal in ChIP experiments (Figure 7, lanes 3 and 4). This result has been shown to arise with a Myb protein from *Saccharomyces cerevisiae* incapable of binding conventional MRE.<sup>43</sup> *P. falciparum* might have evolved new PfMyb1/DNA binding specificities that might depend on chromatin context and/or availability of the protein partner. Indeed, identification of binding sites for PfMyb1 in the PGK promoter is a new challenge for understanding PfMyb1-dependent transcriptional regulation. Nevertheless, this experiment clearly demonstrated the ability of the PfMyb1 protein to bind to the promoter of a set of genes of critical importance in the *P. falciparum* cell cycle and therefore to regulate their expression in a direct manner. Moreover, these data suggest that PfMyb1 may participate directly in either repression (for expression of PfPk5) or activation (for the five other genes) of transcriptional gene expression. This is a common feature for Myb proteins, among which is the c-Myb factor reported to mediate either up or down-regulation of different genes.<sup>20</sup> Hence, depending on post-translational modifications of the transcription factors,<sup>44,45</sup> the protein partners present on each given promoter and the chromatin context, the activity of transcription factors could be modulated either for activation or repression of transcription.

In conclusion, we describe the first attempt to unravel the PfMyb1 regulation network. Although more work is needed to understand the multiple determinants of PfMyb1-mediated regulation, these experiments provide important clues for future analysis of factors determining regulation of gene expression by PfMyb1. The data reported herein suggest that PfMyb1 is an essential transcription factor for the erythrocytic cycle of *P. falciparum* and,

moreover, that PfMyb1 directly regulates key genes involved in cell cycle regulation and progression.

## Materials and Methods

### *Plasmodium falciparum* culture

The 3D7 clone of *P. falciparum* was provided by Dr D. Walliker and was cultured as described<sup>46</sup> with slight modifications. To obtain synchronized erythrocytic stages, ring cultures were enriched by three successive treatments with 5% (w/v) sorbitol at 44 hour intervals.<sup>47</sup> Synchronization of culture was verified by morphological analysis of Giemsa-stained blood smears at different time-points.

### dsRNA preparation

A polymerase chain reaction (PCR) fragment (634 nt) representing the *pfmyb1* gene (PF13\_0088) was amplified from 3D7 genomic DNA using forward (5'GGATGAAATGTGATGATAAAAAC) and backward primers (5'TACTTCATCTTTTGTCCATTTT). The PCR product was cloned into TOPO PCR II vector (Invitrogen). A PCR fragment (570 nt) representing the portion of the *env* gene encoding gp120 (GeneBank accession number K02013) was amplified from HIV-1 genomic DNA using forward (5'ATGCCCCAGACTGTGAGTTG) and backward primers (5'CTACTGTAATTCAACACAAC) and cloned into pBluescript SK- vector. The clones were sequenced by dideoxy sequencing reactions. Plasmids were used as a template to generate sense RNA (sRNA) and antisense RNA (asRNA) using T7, SP6 and T3 RiboMAX Express Large Scale RNA Production System (Promega). The dsRNA was prepared as follows: equal amounts of sRNA and asRNA were separately denatured at 95 °C for one minute, mixed and then heated to 95 °C for an additional minute. Annealing was performed by slow cooling over several hours and thereafter dsRNA samples were analyzed on native 1% (w/v) agarose gel.

### dsRNA assay

For each experiment using dsRNA, 25 µg of dsRNA was added per millilitre of highly synchronized young ring culture (four to six hours post invasion). Parasitaemia was set as described in the next section for manual counting, [<sup>3</sup>H]hypoxanthine uptake assay and flow cytometry in 24-well plates and incubated for 48 hours. A 20 ml culture of 5% parasitaemia of highly synchronized young ring stages was set for RNA and protein extraction. This culture was incubated in 75 cm<sup>3</sup> culture flask for 18–20 hours as described above *P. falciparum* culture. Control cultures were carried out as described above and an equivalent volume of dsRNA annealing buffer was added to the medium.

### Parasite growth assay

[<sup>3</sup>H]hypoxanthine uptake was evaluated as described<sup>48</sup> with slight modifications. In brief, 200 µl of ring stage parasitized erythrocytes (parasitaemia, 0.5%; hematocrit, 1.8%) were dispensed in 96-well plates preloaded with 25 µg of dsRNA per millilitre in triplicate and with serial dilutions of chloroquine in positive control wells. After 24 hours, [<sup>3</sup>H]hypoxanthine was added to each well and plates were then incubated for an



additional 24 hours. Parasites were harvested, and incorporation of radioactivity was determined by liquid scintillation counting. Experiments were repeated twice. Data are reported as mean and standard deviation (SD) of percentage parasite growth inhibition in triplicate experiments relative to untreated controls after correcting for background [<sup>3</sup>H]hypoxanthine incorporation in uninfected erythrocytes. Alternatively, growth inhibition was determined by microscopic examination of Giemsa-stained thin blood smears. Smears from untreated cultures were always used as controls. Parasitaemia was measured by counting 3000 red blood cells and expressed as percentage of total parasitized erythrocytes. Twenty-five micrograms of dsRNA was added to a ring stage culture (1.5% parasitaemia, 5% hematocrit) and parasitaemia was evaluated 48 hours later. Experiments were repeated nine times in triplicate. Finally, flow cytometry was performed on 48 hour-treated culture (1.5% rings, 5% hematocrit) as described.<sup>49</sup> Experiments were repeated three times in triplicate. Statistical analysis was performed using Student's *t*-test.

### RT-PCR and RT-qPCR

RT-PCR: total RNA (1 µg) was treated with two units of RNase-free DNase I (Qiagen) and purified using RNeasy columns with the Qiagen clean-up procedure to avoid DNA contamination. Synthesis of cDNA was performed for 50 minutes at 42 °C using 1 µg of RNA (in 40 µl final volume), 50 units of MMLV reverse transcriptase (Superscript II Invitrogen), 0.5 ng of random hexamers, 5 mM MgCl<sub>2</sub>, 20 units of RNaseout (Invitrogen), 10 mM DTT and 0.25 mM each dNTP. Samples were then subjected to RNase-H treatment with two units of enzyme for 20 minutes at 37 °C. PCR was carried out with 2 µl of each cDNA sample (or the negative control) in the presence of 1 µM primers; 18 S forward (5'GACTCAACACGGGGAACTCACTAGTT) and backward primers (5'ACAATTCATCATATCTTTAATCG GTA) or *pfmyb1* forward (5'GGATGAAATGTGATGATAAAAC) and backward primers (5'TACTTCATCTTTTGTCCATTTT or 5'TCTCCTTTTGTAAGATATTT CATATTCA) or MAL13P1.76 forward (5'ATGCAA AATCCTCAAAT) and backward (5'TTATGTATCGT TAATTAAC) primers, 200 µM dNTP, and two units of Taq polymerase enzyme. After denaturation at 95 °C for two minutes, 30 cycles with annealing at 55 °C, elongation at 60 °C and denaturation at 95 °C, each step for one minute, were performed to amplify the *pfmyb1* and MAL13P1.76 fragments. Twenty cycles of the same amplification profile were used to amplify the ribosomal 18 S subunit. We verified that, under these conditions and number of cycles, the magnitude of the signals remained proportional to RNA concentrations, making the RT-PCR a semi-quantitative procedure. Fragments of expected lengths (657 bp, 1204 bp and 634 bp or 1015 bp for 18 S, MAL13P1.76 and *pfmyb1*, respectively) were observed by 1% agarose gel electrophoresis and relative amounts of mRNA were determined by densitometry (Densylab, Microvision Instruments, Evry, France).

Real-time quantitative PCR reaction was performed as described.<sup>24</sup> Specific primers are listed in Supplementary Data 1.

Multiplex real-time quantitative PCR was performed with LUX-primers (Invitrogen). Reactions were performed in a 25 µl volume containing 5 µl of 25-fold diluted cDNA preparations, 12.5 µl of Platinum PCR mix (Invitrogen) supplemented with 3 mM MgCl<sub>2</sub>, and 1875 nM *pfmyb1* primers and 94 nM 18 S primers.

Amplification and detection of specific products were performed with the MX4000 light cycler (Stratagene) with the following cycle profile: one cycle at 50 °C for two minutes, one cycle at 95 °C for two minutes, 40 cycles with 15 seconds denaturation at 95 °C and 30 seconds annealing-elongation at 60 °C. The 18 S primers were CACCAT TTTCTTGATTCTTGGATGG labelled with JOE, and GATCTCGTTCGTTATCGGAAT. *pfmyb1* primers were CACATCTGCAAAGGAATGTCAAAGATG labelled with FAM, and CACGACAAAGGACTTCTATTGGT.

The size of each PCR product and number of bands were verified on a 10% (w/v) acrylamide gel. Two additional reactions were performed, either without reverse transcriptase or without the RNA sample, to verify the absence of DNA contamination. The quantity of cDNA for each experimental gene was normalized to the 18 S concentration (chr5.rRNA-1-18s-A) in each sample. For each sample and each gene, experiments were performed twice and in triplicate. Relative gene expression was expressed as the log<sub>2</sub> of the ratio of expression at each time-point over that of the ring stage using the 2<sup>-ΔΔCT</sup> method.

### Serial immuno-precipitation and Western blotting

Protein extracts of *P. falciparum* trophozoites were prepared as described.<sup>16</sup> Protein (100 µg) was subjected to immuno-precipitation using 2 × 10<sup>7</sup> magnetic tosylactivated Dynabeads (M-280, Dynal Biotech) coated with 2.5 µg of either anti-PfMyb1 or preimmune serum, according to the manufacturer's recommendations. Immuno-precipitated proteins were resolved by SDS-10% PAGE and subjected, after transfer onto nitrocellulose membrane, to Western blotting using the anti-PfMyb1. The membrane was first incubated overnight at 4 °C with a 1:1000 dilution of preimmune or anti-PfMyb1 serum and then incubated with a 1:10,000 dilution of peroxidase-conjugated anti-rabbit antibody (Biolabs). Concomitantly, 5 µg of the supernatant proteins was resolved by SDS-10% PAGE and blotted onto nitrocellulose membrane. The membrane was first incubated overnight at 4 °C with a 1:10,000 dilution of anti-PfHsp70 serum and then with a 1:10,000 dilution of peroxidase-conjugated anti-mouse antibody (Biolabs).

Western blotting experiments on PfPfk5 protein were performed on 60 µg of protein extracts prepared as described.<sup>50</sup> Proteins were resolved by SDS-12% PAGE and subjected, after transfer onto nitrocellulose membrane, to Western blotting using the anti-PfPfk5 serum, a kind gift from C. Doerig. The membrane was first incubated overnight at 4 °C with a 1:1000 dilution of anti-PfPfk5 serum and then with a 1:10,000 dilution of peroxidase-conjugated anti-rabbit antibody (Biolabs). PfHSP70 protein was detected as described above.

All membranes were developed using a chemiluminescent substrate, according to the manufacturer's instructions (ECL, Amersham-Pharmacia Biotech). Relative amounts of proteins were determined by densitometry (Densylab, Microvision Instruments, Evry, France).

### Microarray construction, probe labeling, hybridization and data analysis

Microarrays were prepared as described by Gissot *et al.*<sup>24</sup> Briefly, 17 spots representing positive and negative controls such as sequences unrelated to *P. falciparum* were printed. Alien PCR products from Stratagene, with no apparent similarity to the *P. falciparum* genomic sequence,

were used for normalization of the reverse transcription. 3D7 genomic DNA was also printed in order to normalize the total amount of each cDNA present during the hybridization. All targets were spotted in duplicate with a Qarray arrayer (Genetix) onto poly-L-lysine-coated slides prepared as described<sup>51</sup> with a spot spacing of 400 µm, centre to centre onto a 1 cm<sup>2</sup> area.

Probe labeling and hybridization were performed as described.<sup>24</sup> Data were collected as an image file, gridded, and converted into an Excel file using β-Vision software (Biospace mesure). The global background was subtracted for each signal. In addition, the signal intensity of each spot was normalized according to the average value of intensities of spots corresponding to exogenous control PCR products (Alien, Stratagene) on each array in order to control for variation in the efficiency of reverse transcription. Furthermore, *P. falciparum* genomic DNA was used to normalize the amount of cDNA present in the hybridization. Four independent biological experiments were performed. Statistical discrimination of genes expressed differentially in the two stages was performed on the complete normalized set of data using the freely available SAM software†.<sup>25</sup>

### Chromatin immuno-precipitation

Potential MRE were localized in promoter sequences by using MatInspector<sup>52</sup> according to three matrix families (one from the plant and two from the vertebrate subsections). This protocol is adapted from Hecht *et al.*<sup>53</sup> A 100 ml sample of trophozoite culture was subjected to protein-DNA crosslink using 1% (v/v) formaldehyde (Sigma) at 37 °C for 15 minutes. Nuclei were prepared as described<sup>16</sup> and resuspended in RIPA buffer (30 mM Tris-HCl (pH 8), 150 mM NaCl, 20 mM MgCl<sub>2</sub>, 1 mM EDTA, 1 mM DTT, 0.5% (v/v) Triton X-100, 1% (v/v) Nonidet-P40, 1 mM PMSE, 1X Complete TM mixture tablet from Roche Molecular Biochemicals). Samples were sonicated (BioBlock Scientific, VibraCell 72405) three times at 30 W for ten seconds, allowing 30 seconds cooling between sonifications, and centrifuged for five minutes at 14,400 g. The supernatant, containing soluble chromatin, was incubated overnight at 4 °C with 2 × 10<sup>7</sup> magnetic tosylactivated Dynabeads (M-280, Dynal Biotech) coated with 2.5 µg of either preimmune or anti-PfMyb1 serum. Beads were washed twice with 1 ml of RIPA buffer supplemented with 0.01% (w/v) SDS. Immuno-precipitated chromatin was eluted using TE (10 mM Tris (pH 8), 1 mM EDTA) supplemented with 1% SDS. Chromatin was reverse cross-linked during six hours at 65 °C and subjected to proteinase K (20 µg) treatment for two hours at 37 °C. DNA was extracted with phenol/chloroform/isoamyl alcohol and precipitated with ethanol and sodium acetate. DNA from input and ChIP samples were resuspended in 200 µl and 20 µl of TE, respectively.

The amount of template used for PCR was 2 µl of a 1:100 dilution of input DNA and 2 µl of undiluted ChIP DNA. PCR was performed in the presence of 1 µM specific primers (see Supplementary Data 2), 200 µM dNTP, and two units of Taq polymerase. After denaturation at 95 °C for two minutes, 40 cycles were performed with annealing at 50 °C, elongation at 60 °C and denaturation at 95 °C, each step for one minute. PCR products were analyzed on 1.2% agarose gel.

† <http://www-stat.stanford.edu/~tibs/SAM/stat.SAM/>

### Acknowledgements

We thank Dr Christian Doerig for the kind gift of the anti-PfPfk5 serum and Dr Jennifer Richardson for critical reading of the manuscript. We are grateful to Professor Dominique Mazier (INSERM U511) for continuous encouragement and support. M.G., P.R. and C.B. were financially supported by the Ministère de l'Éducation Nationale, France. M.G. held a fellowship from Fondation pour la Recherche Médicale (FRM). The project was supported by INSERM funds to C.V. We are grateful to the Pitié-Salpêtrière genomic core facility (P3S) supported by INSERM, the région Ile-de-France, the Fondation de Recherche HMR-Aventis, and by University Paris VI.

### Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2004.11.045

### References

1. World Health Organization (2000). *WHO Expert Committee on Malaria—Twentieth Report*. WHO, Rome.
2. Demidov, V. V. (2003). Infection stage clues to new antimalarial medicines. *Drug Discov. Today*, **8**, 913–915.
3. Gardner, M. J., Tettelin, H., Carucci, D. J., Cummings, L. M., Aravind, L., Koonin, E. V. *et al.* (1998). Chromosome 2 sequence of the human malaria parasite *Plasmodium falciparum*. *Science*, **282**, 1126–1132.
4. Bowman, S., Lawson, D., Basham, D., Brown, D., Chillingworth, T., Churcher, C. M. *et al.* (1999). The complete nucleotide sequence of chromosome 3 of *Plasmodium falciparum*. *Nature*, **400**, 532–538.
5. Gardner, M. J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R. W. *et al.* (2002). Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature*, **419**, 498–511.
6. Horrocks, P., Dechering, K. & Lanzer, M. (1998). Control of gene expression in *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **95**, 171–181.
7. Levitt, A. (1993). RNA processing in malarial parasites. *Parasitol. Today*, **9**, 465–468.
8. Cary, C., Lamont, D., Dalton, J. P. & Doerig, C. (1994). *Plasmodium falciparum* chromatin: nucleosomal organisation and histone-like proteins. *Parasitol Res.* **80**, 255–258.
9. Li, W. B., Bzik, D. J., Gu, H. M., Tanaka, M., Fox, B. A. & Inselburg, J. (1989). An enlarged largest subunit of *Plasmodium falciparum* RNA polymerase II defines conserved and variable RNA polymerase domains. *Nucl. Acids Res.* **17**, 9621–9636.
10. Li, W. B., Bzik, D. J., Tanaka, M., Gu, H. M., Fox, B. A. & Inselburg, J. (1991). Characterization of the gene encoding the largest subunit of *Plasmodium falciparum* RNA polymerase III. *Mol. Biochem. Parasitol.* **46**, 229–239.
11. Fox, B. A., Li, W. B., Tanaka, M., Inselburg, J. & Bzik, D. J. (1993). Molecular characterization of the largest subunit of *Plasmodium falciparum* RNA polymerase I. *Mol. Biochem. Parasitol.* **61**, 37–48.

12. McAndrew, M. B., Read, M., Sims, P. F. & Hyde, J. E. (1993). Characterisation of the gene encoding an unusually divergent TATA-binding protein (TBP) from the extremely A + T-rich human malaria parasite *Plasmodium falciparum*. *Gene*, **124**, 165–171.
13. Hirtzlin, J., Farber, P. M. & Franklin, R. M. (1994). Isolation of a novel *Plasmodium falciparum* gene encoding a protein homologous to the Tat-binding protein family. *Eur. J. Biochem.* **226**, 673–680.
14. Horrocks, P. & Kilbey, B. J. (1996). Physical and functional mapping of the transcriptional start sites of *Plasmodium falciparum* proliferating cell nuclear antigen. *Mol. Biochem. Parasitol.* **82**, 207–215.
15. Lanzer, M., de Bruin, D. & Ravetch, J. V. (1992). Transcription mapping of a 100 kb locus of *Plasmodium falciparum* identifies an intergenic region in which transcription terminates and reinitiates. *EMBO J.* **11**, 1949–1955.
16. Osta, M., Gannoun-Zaki, L., Bonnefoy, S., Roy, C. & Vial, H. J. (2002). A 24 bp *cis*-acting element essential for the transcriptional activity of *Plasmodium falciparum* CDPdiacylglycerol synthase gene promoter. *Mol. Biochem. Parasitol.* **121**, 87–98.
17. Miller, L. H., Baruch, D. I., Marsh, K. & Doumbo, O. K. (2002). The pathogenic basis of malaria. *Nature*, **415**, 673–679.
18. Calderwood, M. S., Gannoun-Zaki, L., Wellem, T. E. & Deitsch, K. W. (2003). *Plasmodium falciparum* var genes are regulated by two regions with separate promoters, one upstream of the coding region and a second within the intron. *J. Biol. Chem.* **278**, 34125–34132.
19. Dechering, K. J., Kaan, A. M., Mbacham, W., Wirth, D. F., Eling, W., Konings, R. N. & Stunnenberg, H. G. (1999). Isolation and functional characterization of two distinct sexual-stage-specific promoters of the human malaria parasite *Plasmodium falciparum*. *Mol. Cell. Biol.* **19**, 967–978.
20. Kanei-Ishii, C., Nomura, T., Ogata, K., Sarai, A., Yasukawa, T., Tashiro, S. *et al.* (1996). Structure and function of the proteins encoded by the *myb* gene family. *Curr. Top Microbiol. Immunol.* **211**, 89–98.
21. Lipsick, J. S. (1996). One billion years of Myb. *Oncogene*, **13**, 223–235.
22. Boschet, C., Gissot, M., Briquet, S., Hamid, Z., Claudel-Renard, C. & Vaquero, C. (2004). Characterization of PfMyb1 transcription factor during erythrocytic development of 3D7 and F12 *Plasmodium falciparum* clones. *Mol. Biochem. Parasitol.* **138**, 159–163.
23. Ness, S. A., Marknell, A. & Graf, T. (1989). The *v-myb* oncogene product binds to and activates the promyelocyte-specific *mim-1* gene. *Cell*, **59**, 1115–1125.
24. Gissot, M., Refour, P., Briquet, S., Boschet, C., Coupe, S., Mazier, D. & Vaquero, C. (2004). Transcriptome of 3D7 and its gametocyte-less derivative F12 *Plasmodium falciparum* clones during erythrocytic development using a gene-specific microarray assigned to gene regulation, cell cycle and transcription factors. *Gene*, **341**, 267–277.
25. Tusher, V. G., Tibshirani, R. & Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl Acad. Sci. USA*, **98**, 5116–5121.
26. Bozdech, Z., Llinas, M., Pulliam, B. L., Wong, E. D., Zhu, J. & DeRisi, J. L. (2003). The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. *PLoS Biol.* **1**, 5.
27. le Roch, K. G., Johnson, J. R., Florens, L., Zhou, Y., Santrosyan, A., Grainger, M. *et al.* (2004). Global analysis of transcript and protein levels across the *Plasmodium falciparum* life cycle. *Genome Res.* **14**, 2308–2318.
28. Malhotra, P., Dasaradhi, P. V., Kumar, A., Mohammed, A., Agrawal, N., Bhatnagar, R. K. & Chauhan, V. S. (2002). Double-stranded RNA-mediated gene silencing of cysteine proteases (falcipain-1 and -2) of *Plasmodium falciparum*. *Mol. Microbiol.* **45**, 1245–1254.
29. McRobert, L. & McConkey, G. A. (2002). RNA interference (RNAi) inhibits growth of *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **119**, 273–278.
30. Noonpakdee, W., Pothikasikorn, J., Nimitsantiwong, W. & Wilairat, P. (2003). Inhibition of *Plasmodium falciparum* proliferation *in vitro* by antisense oligodeoxynucleotides against malarial topoisomerase II. *Biochem. Biophys. Res. Commun.* **302**, 659–664.
31. Gunasekera, A. M., Patankar, S., Schug, J., Eisen, G., Kissinger, J., Roos, D. & Wirth, D. F. (2004). Widespread distribution of antisense transcripts in the *Plasmodium falciparum* genome. *Mol. Biochem. Parasitol.* **136**, 35–42.
32. Patankar, S., Munasinghe, A., Shoaibi, A., Cummings, L. M. & Wirth, D. F. (2001). Serial analysis of gene expression in *Plasmodium falciparum* reveals the global expression profile of erythrocytic stages and the presence of anti-sense transcripts in the malarial parasite. *Mol. Biol. Cell.* **12**, 3114–3125.
33. Kyes, S., Christodoulou, Z., Pinches, R. & Newbold, C. (2002). Stage-specific merozoite surface protein 2 antisense transcripts in *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **123**, 79–83.
34. Ross-Macdonald, P. B., Graeser, R., Kappes, B., Franklin, R. & Williamson, D. H. (1994). Isolation and expression of a gene specifying a *cdc2*-like protein kinase from the human malaria parasite *Plasmodium falciparum*. *Eur. J. Biochem.* **220**, 693–701.
35. Morgan, D. O. (1997). Cyclin-dependent kinases: engines, clocks, and microprocessors. *Annu. Rev. Cell. Dev. Biol.* **13**, 261–291.
36. Graeser, R., Wernli, B., Franklin, R. M. & Kappes, B. (1996). *Plasmodium falciparum* protein kinase 5 and the malarial nuclear division cycles. *Mol. Biochem. Parasitol.* **82**, 37–49.
37. Graeser, R., Franklin, R. M. & Kappes, B. (1996). Mechanisms of activation of the *cdc2*-related kinase PfPK5 from *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **79**, 125–127.
38. Arnot, D. E. & Gull, K. (1998). The *Plasmodium* cell-cycle: facts and questions. *Ann. Trop. Med. Parasitol.* **92**, 361–365.
39. Read, M., Sherwin, T., Holloway, S. P., Gull, K. & Hyde, J. E. (1993). Microtubular organization visualized by immunofluorescence microscopy during erythrocytic schizogony in *Plasmodium falciparum* and investigation of post-translational modifications of parasite tubulin. *Parasitology*, **106**, 223–232.
40. Kelman, Z. (1997). PCNA: structure, functions and interactions. *Oncogene*, **14**, 629–640.
41. Kilbey, B. J., Fraser, I., McAleese, S., Goman, M. & Ridley, R. G. (1993). Molecular characterisation and stage-specific expression of proliferating cell nuclear antigen (PCNA) from the malarial parasite, *Plasmodium falciparum*. *Nucl. Acids Res.* **21**, 239–243.
42. Grall, M., Srivastava, I. K., Schmidt, M., Garcia, A. M., Mael, J. & Perrin, L. H. (1992). *Plasmodium falciparum*: identification and purification of the phosphoglycerate kinase of the malaria parasite. *Expt. Parasitol.* **75**, 10–18.

43. Hovring, I., Bostad, A., Ording, E., Myrset, A. H. & Gabrielsen, O. S. (1994). DNA-binding domain and recognition sequence of the yeast BAS1 protein, a divergent member of the Myb family of transcription factors. *J. Biol. Chem.* **269**, 17663–17669.
44. Bannister, A. J. & Miska, E. A. (2000). Regulation of gene expression by transcription factor acetylation. *Cell. Mol. Life Sci.* **57**, 1184–1192.
45. Whitmarsh, A. J. & Davis, R. J. (2000). Regulation of transcription factor function by phosphorylation. *Cell. Mol. Life Sci.* **57**, 1172–1183.
46. Trager, W. & Jensen, J. B. (1976). Human malaria parasites in continuous culture. *Science*, **193**, 673–675.
47. Lambros, C. & Vanderberg, J. P. (1979). Synchronization of *Plasmodium falciparum* erythrocytic stages in culture. *J. Parasitol.* **65**, 418–420.
48. Desjardins, R. E., Canfield, C. J., Haynes, J. D. & Chulay, J. D. (1979). Quantitative assessment of antimalarial activity *in vitro* by a semiautomated microdilution technique. *Antimicrob. Agents Chemother.* **16**, 710–718.
49. Makler, M. T., Lee, L. G. & Recktenwald, D. (1987). Thiazole orange: a new dye for Plasmodium species analysis. *Cytometry*, **8**, 568–570.
50. Merckx, A., le Roch, K., Nivez, M. P., Dorin, D., Alano, P., Guiterrez, G. J. *et al.* (2003). Identification and initial characterization of three novel cyclin-related proteins of the human malaria parasite *Plasmodium falciparum*. *J. Biol. Chem.* **278**, 39839–39850.
51. Schena, M., Shalon, D., Davis, R. W. & Brown, P. O. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*, **270**, 467–470.
52. Quandt, K., Frech, K., Karas, H., Wingender, E. & Werner, T. (1995). MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucl. Acids Res.* **23**, 4878–4884.
53. Hecht, A., Strahl-Bolsinger, S. & Grunstein, M. (1999). Mapping DNA interaction sites of chromosomal proteins. Crosslinking studies in yeast. *Methods Mol. Biol.* **119**, 469–479.

*Edited by J. Karn*

*(Received 17 September 2004; received in revised form 18 November 2004; accepted 18 November 2004)*

### III – CARACTERISATION FONCTIONNELLE DE DEUX FACTEURS NUCLEAIRES ARCHITECTURAUX DE *P. FALCIPARUM* PFHMGB1 ET PFHMGB2.

#### III.1 – Résumé

Chez les eucaryotes, le remodelage de la chromatine est une étape cruciale dans la régulation de la transcription. Le contrôle de l'équilibre entre l'état condensé de l'ADN et une configuration plus relâchée est essentiel à l'accessibilité des facteurs de transcription à leurs séquences cibles, dont nous avons montré l'importance chez *P. falciparum*, précédemment dans ce travail de thèse. Nous nous sommes donc intéressés dans cette troisième partie à l'existence des protéines responsables de cet aspect de la régulation de la transcription chez *P. falciparum*. Nous avons identifié quatre phases ouvertes de lecture codant pour des protéines HMGB. Nous avons entrepris l'étude de deux de ces protéines, PfHMGB1 et PfHMGB2 possédant un domaine de liaison à l'ADN "HMG-box" caractéristique de la famille des protéines HMGB. Les analyses bioinformatiques concernant la phylogénie et la structure second-tertiaire de ces séquences (Données supplémentaires III.1, 2 et 3, p. 173-176) soutiennent l'idée que ces protéines sont des **facteurs HMGB spécifiques de structure ADN**.

Nous avons alors tenté de valider cette hypothèse expérimentalement. La fonction d'un facteur nucléaire architectural se caractérise par sa capacité à lier des structures particulières de l'ADN et à induire la courbure de l'ADN linéaire. Par des expériences de retard sur gel, nous avons montré d'une part, que les protéines recombinantes étaient capables d'interagir spécifiquement avec des ADN cruciformes. Tandis que l'interaction de la protéine PfHMGB1 est strictement dépendante d'une structure complète à 4 brins, l'interaction de la protéine PfHMGB2 avec des structures incomplètes est possible (Fig. 2, p. 676 de l'Article 4). D'autre part, pour mettre en évidence la capacité de ces protéines à courber l'ADN, nous avons utilisé de petits fragments d'ADN de 123 pb dont la rigidité naturelle empêche la circularisation à moins d'un changement conformationnel. Nous avons observé que les protéines PfHMGB1 et PfHMGB2 pouvaient induire la flexibilité de ces fragments et permettre l'obtention d'ADN circulaires en présence d'une ADN ligase, avec une meilleure efficacité pour PfHMGB1 (Fig. 3, p. 677 de l'Article 4).

A l'image de leurs protéines orthologues eucaryotes, les protéines PfHMGB1 et PfHMGB2 possèdent donc les propriétés architecturales essentielles à l'activité des facteurs de la famille HMGB, avec toutefois des différences de spécificité et d'efficacité.

De plus, l'étude de la localisation par immunofluorescence s'ajoute à la détection de ces protéines dans les extraits nucléaires du parasite (Fig.4b, p. 677 de l'Article 4) pour confirmer la fonction nucléaire de ces facteurs. En effet, PfHMGB1 et PfHMGB2 sont localisées dans le noyau des parasites aux stades sexués et asexués, PfHMGB2 étant aussi observée dans le cytoplasme des gamétocytes (Fig. 5, p. 679 de l'Article 4). Ces facteurs sont exprimés tout au long du développement érythrocytaire du clone parasitaire 3D7 à la différence notable que PfMGB1 est préférentiellement exprimée pendant les stades asexués tandis que PfHMGB2 est exprimée majoritairement dans les stades sexués (Fig. 4c, p. 677 de l'Article 4).

A la lumière de toutes les différences observées pour les deux protéines, leur rôle non redondant dans la régulation transcriptionnelle apparaît comme une évidence au cours du développement érythrocytaire et notamment lors de la différenciation sexuée.

**ARTICLE 4 :**

**The High Mobility Group Box (HMGB) nuclear factors of  
*Plasmodium falciparum*.**

Sylvie Briquet, Charlotte Boschet, Mathieu Gissot, Emilie Tissandié, Elisa Sevilla, Jean-François Franetich, Isabelle Thiery, Zuhail Hamid, Catherine Bourgouin and Catherine Vaquero.

Accepté à la publication dans *Eukaryotic cell*.

## High-Mobility-Group Box Nuclear Factors of *Plasmodium falciparum*†

Sylvie Briquet,<sup>1\*</sup> Charlotte Boschet,<sup>1</sup> Mathieu Gissot,<sup>1‡</sup> Emilie Tissandié,<sup>1</sup> Elisa Sevilla,<sup>1</sup>  
Jean-François Franetich,<sup>1</sup> Isabelle Thiery,<sup>2</sup> Zuhail Hamid,<sup>1§</sup>  
Catherine Bourguoin,<sup>2</sup> and Catherine Vaquero<sup>1\*</sup>

INSERM, U511, Université Pierre et Marie Curie, Paris VI, Centre Hospitalo-Universitaire de la Pitié-Salpêtrière, Paris, France,<sup>1</sup> and  
Biologie et Génétique du Paludisme, CEPIA (Centre de Production et d'Infection des Anophèles), Institut Pasteur, Paris, France<sup>2</sup>

Received 30 November 2005/Accepted 31 January 2006

In eukaryotes, the high-mobility-group (HMG) nuclear factors are highly conserved throughout evolution and are divided into three families, including HMGB, characterized by an HMG box domain. Some HMGB factors are DNA structure specific and preferentially interact with distorted DNA sequences, trigger DNA bending, and hence facilitate the binding of nucleoprotein complexes that in turn activate or repress transcription. In *Plasmodium falciparum*, two HMGB factors were predicted: PfHMGB1 and PfHMGB2. They are small proteins, under 100 amino acids long, encompassing a characteristic HMG box domain closely related to box B of metazoan factors, which comprises two HMG box domains, A and B, in tandem. Computational analyses supported the conclusion that the *Plasmodium* proteins were genuine architectural HMGB factors, and *in vitro* analyses performed with both recombinant proteins established that they were able to interact with distorted DNA structures and bend linear DNA with different affinities. These proteins were detected in both asexual- and gametocyte-stage cells in Western blotting experiments and mainly in the parasite nuclei. PfHMGB1 is preferentially expressed in asexual erythrocytic stages and PfHMGB2 in gametocytes, in good correlation with transcript levels of expression. Finally, immunofluorescence studies revealed differential subcellular localizations: both factors were observed in the nucleus of asexual- and sexual-stage cells, and PfHMGB2 was also detected in the cytoplasm of gametocytes. In conclusion, in light of differences in their levels of expression, subcellular localizations, and capacities for binding and bending DNA, these factors are likely to play nonredundant roles in transcriptional regulation of *Plasmodium* development in erythrocytes.

Malaria is the most important parasitic disease in the world, and of the 300 to 500 million cases each year, approximately 2 million people die. Among the four species of malaria parasites infecting humans, *Plasmodium falciparum* causes the highest morbidity and mortality. Global efforts to eradicate malaria have failed, and there is presently no effective vaccine available. Therefore, a greater understanding of parasite biology throughout development is urgently needed in order for novel therapeutic strategies to control malaria to be proposed.

During the erythrocytic life cycle, intense multiplication of parasites takes place, as well as gametocyte differentiation associated with cell cycle arrest. These different developmental pathways require the coordinated and modulated expression of diverse sets of genes, involving transcriptional, epigenetic, and posttranscriptional regulation. Currently, it is commonly accepted that general mechanisms involved in gene regulation in eukaryotes also operate in *P. falciparum* (25, 32, 33). Nevertheless, elucidation of the molecular mechanisms involved in transcriptional regulation in *Plasmodium* is still challenging.

Even if very little is known about the *cis*- and *trans*-regulatory elements of the parasite, *Plasmodium* genes exhibit the bipartite structure of eukaryotic promoters, i.e., a basal promoter regulated by upstream regulatory elements (25) that present some homology with the binding sites of eukaryotic transcription factors (TF). The recent completion of the genome sequence of *P. falciparum* revealed a high proportion of orphan proteins (60% of the open reading frames [ORFs] have no match with any of the annotated sequences listed in the data banks [18]). These data might contribute to the low numbers of recognizable, orthologous TF (11). However, it is reasonable to assume that in *Plasmodium* the interplay between regulatory elements and TF, whose availability (49) presumably modulated throughout parasite development, governs also the level of RNA synthesis.

In eukaryotes, in addition to general TF also annotated in *Plasmodium* (10, 23, 37, 38, 43, 44), the factors involved in transcriptional regulation can be divided into factors interacting either with specific DNA sequences (42) or with DNA structures. The latter include the nonhistone proteins of the high-mobility-group (HMG) superfamily (7, 58, 62), which is divided into three families of proteins in line with their characteristic functional motifs (8): HMGA, which interacts with the AT hook; HMGN, which interacts with the nucleosomes; and HMGB, which encompasses one or several copies of the HMG box DNA binding domain (for a review, see reference 7). HMG proteins are present in all metazoan phyla, plants, and yeast and have also been reported in unicellular parasites, including trypanosomes (15, 45), schistosomes (21), and *Plasmodium* (29). They are quite abundant proteins, one molecule

\* Corresponding author. Mailing address: INSERM, U511, Université Pierre et Marie Curie, Paris VI, Centre Hospitalo-Universitaire de la Pitié-Salpêtrière, 91 boulevard de l'Hôpital, 75013 Paris, France. Phone: 33 (0) 1 40 77 81 14. Fax: 33 (0) 1 45 83 88 58. E-mail for Sylvie Briquet: briquet@ext.jussieu.fr. E-mail for Catherine Vaquero: vaquero@ext.jussieu.fr.

† Supplemental material for this article may be found at <http://ec.asm.org/>.

‡ Present address: Albert Einstein College of Medicine, New York, N.Y.

§ Present address: INMO, University of Gezira, Sudan.



for 10 to 15 nucleosomes in vertebrates. It is assumed that the wrapping of DNA by histones and nonhistone proteins, including the HMG proteins, controls the access of the TF to their target sites on nucleosomes (31).

HMGB factors are highly conserved throughout evolution, and their HMG box domain is composed of around 80 amino acids (aa) folded in three  $\alpha$ -helices arranged in an L shape (3, 66). In vertebrates, the HMGB proteins generally present two boxes, A and B, and also basic N- and C-terminal extensions and a rather long C-terminal acidic tail (58). Despite their low sequence homology, both boxes (A and B) present a well-conserved L-shaped structure, even though their DNA binding and bending capacities may display some differences (28, 69). In lower eukaryotes, either the basic extension (*Drosophila melanogaster*) or the negatively charged tail (*Saccharomyces cerevisiae*) is missing, in contrast to plant HMGs (60), which possess both extensions, albeit of different lengths. The basic domains appear to play a role in the stabilization of HMGB-mediated DNA bending. In contrast, the role of the acidic tail remains elusive and may be shaped to interact with the positive charges of histones (for a review, see reference 62). Two subfamilies of HMGB, with either DNA sequence specificity (SOX, SRY, TCF, MATA) or structure specificity (HMGB per se), have been identified. The latter preferentially interacts with distorted DNA sequences and triggers DNA bending, hence altering the positioning of nucleosomes on the DNA fiber, thereby controlling the level of transcription (31). Finally, a linker histone H1 (27), via its interaction with the DNA linker between two nucleosomes, increases the compactness of the chromatin (53), impairing interactions between DNA and TF and therefore repressing gene transcription (24). In contrast, the HMGB proteins appear to be associated with active chromatin (48), increasing nucleosome sliding and target site accessibility and thereby enhancing transcription.

Lately, these proteins, historically known as nuclear proteins, have been reported to be released from mammalian cells and to act as mediators of the immune response and as potent macrophage-activating factors (for reviews, see references 14, 41, and 46).

In *Plasmodium*, very few TF have been annotated (PlasmoDB and our group [19]) and characterized (6, 20). In *Plasmodium falciparum*, four potential HMG factors have been annotated, including one previously reported for the FCQ27 (29) and FCC1/HN parasite clones. Two of these factors, PfHMGB1 and PfHMGB2, were investigated during the erythrocytic cycle to evaluate their molecular implications in transcription regulation.

#### MATERIALS AND METHODS

***Plasmodium falciparum* culture.** The 3D7 clone of *P. falciparum* was provided by D. Walliker and was grown in human erythrocytes, as described by Trager and Jensen (61), except that the culture medium contained 0.5% Albumax instead of human serum. Gametocytes were produced from the gametocyte-producing isolate NF54 (provided by W. Eling) as aforementioned, with 10% AB human serum, using the automated culturing system developed by Ponnudurai et al. (52). Parasitemia was monitored on Giemsa-stained blood smears. Thirteen to 15 days after induction of gametocytogenesis, gametocytes were monitored for maturity by observation under a microscope (1  $\times$  100 immersion) of microgamete exflagellation, with the addition of 5  $\mu$ l of gametocyte culture in a 10- $\mu$ l drop of human serum.

**Antibodies.** Primary antibodies used to characterize the PfHMGB1 and PfHMGB2 proteins were obtained after immunization of BALB/c mice with either recombinant protein (50  $\mu$ g, twice at 2-week intervals). Sera with high levels of

anti-PfHMGB1 and anti-PfHMGB2 activity were collected after the first booster inoculation (day 30) and selected for immunochemical analyses.

**Annotation.** Before the sequence of the *P. falciparum* genome was completed in 2002, PfHMGB1 was annotated within chromosome 12 by homology to a consensus of the HMG box domains of around 50 disparate eukaryotic sequences. Then, the HMG box domain of PfHMGB1, used as a query against the *Plasmodium* database, allowed the annotation of PfHMGB2 and PfHMGB3 within chromosomes 8 and 12, respectively, while that of PfHMGB2 allowed the annotation of PfHMGB4 within chromosome 13. The presence of relevant HMG box domains in these proteins was checked with MotifScan (16).

**Molecular cloning.** The *Pfhmgb1* and *Pfhmgb2* ORF sequences were amplified from genomic DNA of the *P. falciparum* 3D7 clone, with forward (5'-GGTGG ATCCATGAAGAATACAGGAAAAGAAG-3' and 5'-GGTCCCGGGCCCA ATTTAAGCTTTCATTTTC-3') and backward (5'-ATTGGATCCATGGCTTC AAAATCTCAAAA-3' and 5'-ATTGGTACCTTATCTTGATTTTCTTTC-3') primers, respectively, using PCR conditions with an elongation temperature of 60°C, as described previously (57). Fragments of 294 bp and 300 bp, corresponding to the complete *Pfhmgb1* and *Pfhmgb2* ORF sequences, were cloned directly into the pGEM-T Easy vector (Promega) and PCR II-Topo (Invitrogen), respectively, and then sequenced with the ABI Prism kit (Perkin Elmer).

**Northern blot analysis.** Total RNA was purified from isolated parasites with TRIzol (Invitrogen), and the integrity of the RNA preparation was monitored by ethidium bromide staining on an agarose gel and analysis with an Agilent bioanalyzer. The *Pfhmgb1* and *Pfhmgb2* transcripts were characterized from 20  $\mu$ g of 3D7 total RNA by Northern blotting, according to the Ultrahyb protocol of Ambion, with a 65°C hybridization temperature. Antisense  $\alpha$ -<sup>32</sup>P riboprobes were prepared as previously described (50) from the pGEMT-*pfhmgb1* and pCRII Topo-*pfhmgb2* vectors by using T7 and SP6 RNA polymerase (Promega), respectively.

**Expression and purification of recombinant proteins.** The pQE30 vector (QIAGEN) was used to express the recombinant proteins (rePfHMGB1 and rePfHMGB2) in *Escherichia coli* as His<sub>6</sub>-tagged proteins with BamHI-XmaI- and BamHI-KpnI-digested inserts from the above-mentioned vectors, respectively. The bacterial strain SG 13009, harboring the *Pfhmgb1* or *Pfhmgb2* expression construct, was grown at 37°C in 200 ml of 2YT medium containing 100  $\mu$ g/ml ampicillin as well as 50  $\mu$ g/ml kanamycin. Expression was induced with 0.1 mM IPTG (isopropyl- $\beta$ -D-thiogalactopyranoside) for 3 h at 37°C, and collected cells were solubilized in sonication buffer S (25 mM Tris-HCl [pH 8], 300 mM NaCl, 10 mM imidazole, 10 mM  $\beta$ -mercaptoethanol, 0.5% Triton X-100), 1 ml per mg of dry pellet, in the presence of lysozyme and a 1/25 final dilution of a protease inhibitor cocktail tablet (Roche). Purification of His-PfHMGB proteins was performed essentially as previously described with Ni-nitrilotriacetic acid agarose beads (QIAGEN) (34). After three washes with buffer S supplemented with 20 mM imidazole, bound proteins were eluted with either 250 mM or 50 mM imidazole in 20 mM Tris-HCl (pH 8)-300 mM NaCl for rePfHMGB1 or rePfHMGB2, respectively.

**EMSA with synthetic four-way DNA junctions.** The partially complementary oligonucleotides 1 to 4, previously described (5, 67), were used to create four-way DNA junctions (4H), as well as 3H and 2H. In addition, the radiolabeled leg 1 containing cruciform 4H was fractionated and eluted from a 5% polyacrylamide gel in 0.5 $\times$  Tris-borate-EDTA (TBE) buffer. Electrophoretic mobility shift assays (EMSA) were performed by incubating labeled 4H with increasing amounts (0 to 25  $\mu$ M) of rePfHMGB1 or rePfHMGB2 in a 10- $\mu$ l final volume for 20 min at room temperature. This dose-response experiment was carried out to determine the amounts of protein necessary to create a major 4H-PfHMGB complex. In the competition assay, 100- and 500-fold molar excesses of cold complete 4H or incomplete 3H and 2H were added to the reaction for an additional 20 min. Samples were run on a 6.5% polyacrylamide gel in 0.5 $\times$  TBE buffer at 120 V. The vacuum-dried gels were autoradiographed with intensifying screens at -80°C overnight.

**Ligase-mediated circularization assay.** The circularization assay was based on existing protocols (9, 67). Linear DNA fragments of 123 bp were  $\gamma$ -<sup>32</sup>P 5' end labeled and preincubated with increasing concentrations (0.25 to 100  $\mu$ M) of rePfHMGB1 or rePfHMGB2 for 20 min at room temperature in 1 $\times$  DNA ligase buffer (New England Biolabs [NEB]) in a final volume of 10  $\mu$ l. DNA circularization was generated with T4 DNA ligase (NEB), and linear DNA was subsequently digested by exonuclease III (NEB). Samples were treated with proteinase K (Invitrogen) and run on a 6.5% polyacrylamide gel in 0.5 $\times$  TBE buffer at 120 V. The gels were vacuum dried and autoradiographed.

**Preparation of parasite NE, CE, and total extracts and Western blot analysis.** Nuclear extracts (NE), cytoplasmic extracts (CE), and total lysates were prepared from 50 ml of red blood cells at 10 to 12% parasitemia infected with 3D7 asexual-stage cells, as described by Osta et al. (49). Proteins (12  $\mu$ g) were run on

12% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and subjected to Western blotting experiments (see Fig. 4b) after transfer onto polyvinylidene difluoride membranes (Bio-Rad). For sexual-stage analysis, 2 ml of NF54 culture was harvested on day 13 after induction of gametocytogenesis, with 13.8% parasitemia including 55% mature gametocytes (stage V). After erythrocytic lysis in phosphate-buffered saline–0.15% saponin, gametocytes were collected and resuspended in 40  $\mu$ l of Laemmli buffer. Half of this sample was fractionated for detection of either protein, PfHMGB1 or PfHMGB2. A similar amount of asexual-stage culture was run in parallel (see Fig. 4c). The blots were probed with a 1:2,000 dilution of anti-PfHMGB1 or anti-PfHMGB2 serum, followed by incubation with a peroxidase-conjugated anti-mouse immunoglobulin G (IgG) antibody (Sigma), and revealed by chemiluminescence (Perkin Elmer) or the Supersignal West Femto kit (Pierce). Negative controls were performed with preimmune sera and positive controls with the recombinant proteins. Finally, an anti-HSP70 serum (55) was used as a positive control for CE (see Fig. 4b) and for normalization of protein loading between asexual/sexual stages (QuantiScan Biosoft 2.1).

**Immunofluorescence assay.** For localization of PfHMGB1 and PfHMGB2 in *P. falciparum*, asexual and sexual stages obtained as described previously were washed twice in RPMI, fixed in 2% paraformaldehyde for 20 min at room temperature, and laid on poly-L-lysine-coated multiwell glass slides. Endogenous fluorescence was quenched with 75 mM  $\text{NH}_4\text{Cl}$  for 10 min. The parasites were permeabilized with 0.5% Triton X-100 for 3 min, blocked with 0.5% bovine serum albumin in phosphate-buffered saline for 1 h, and then incubated with a 1:200 dilution of primary antibodies (anti-PfHMGB1, anti-PfHMGB2, anti-HSP70) for 1 h at room temperature, followed by incubation with a fluorescein isothiocyanate-conjugated anti-mouse IgG antibody (Sigma) and DAPI (4',6'-diamidino-2-phenylindole) for 1 h before examination by fluorescence microscopy using a Leica DM RD fluorescent microscope equipped with UV and fluorescein filters. The time of exposure was selected in relation to the intensity of the immunofluorescence obtained with each antibody. Images were acquired with Lucia 4.7 and merged with Adobe Photoshop.

**Nucleotide sequence accession numbers.** The nucleotide sequences of the genes encoding the following proteins have been assigned the corresponding PlasmoDB accession numbers: PfHMGB1, PFL0145c; PfHMGB2, MAL8P1.72; PfHMGB3, PFL0290w; PfHMGB4, MAL13P1.290.

## RESULTS

**Several predicted factors in *Plasmodium* belong to the HMGB family.** Four putative HMG proteins in *P. falciparum* were predicted by sequence homology, as described in Materials and Methods. The proteins appeared to belong to the HMGB family and were named PfHMGB1 to PfHMGB4. The first two, PfHMGB1 and PfHMGB2, are small proteins, under 100 aa in length, while PfHMGB4 has been predicted to encode a 160-aa-long protein. All three encompass only one HMG box domain. In contrast, PfHMGB3 is a larger protein (2,284 aa), with two HMG box domains and several additional putative functional motifs, including one Myb domain (2). The two small proteins, PfHMGB1 and PfHMGB2, are presented herein. It is worthy of note that in contrast to most eukaryotic HMGBs, displaying N- and C-terminal extensions of diverse lengths (58) and reported to bear functional roles, the two *Plasmodium* factors have only a short basic extension upstream of the HMG box domain and no acidic C-terminal tail.

Soullier et al. (56) revealed by phylogenetic analysis that the HMGB factors can be separated into two clearly defined subgroups: (i) the SOX/SRY/MATA/TCF family, whose members are able to bind specific linear DNA sequences, and (ii) the HMGB/UBF family, whose members interact with high affinity to distorted DNA structures. After the addition of *Plasmodium* and *Babesia bovis* sequences, the same analysis was performed several times with different random number seeds, and we were able to assign PfHMGB1 and PfHMGB2 to the subgroup of HMGB proteins characterized by DNA

structure specificity (see supplemental material S1) (17, 59). This finding was strengthened by the alignment shown in Fig. 1, performed with several eukaryotic HMG box domains and two sets of complete HMGB1 and HMGB2 sequences issued from diverse *Plasmodium* species. The HMGB of *Plasmodium* possessed two of the three determinants reported to determine the structural DNA specificity (47), that is, the presence in positions 10 and 32 (according to the residue numbering of *Drosophila* HMG-D) of a serine and a hydrophobic residue, respectively. Therefore, we assigned the *Plasmodium* factors to the architectural HMGB family. Finally, a phylogenetic tree was issued from the HMG box domains of PfHMGB1 and PfHMGB2 and those of various metazoan proteins containing two boxes, A and B, in tandem (see supplemental material S2). This analysis, performed several times with different random number seeds, revealed that the HMG box sequences of both *Plasmodium* proteins are more similar to box B than to box A of proteins with two HMG box domains.

The three-dimensional structure of the two *Plasmodium* factors was modeled by homology using as template the structure of box B of the Chinese hamster HMG1 protein (PDB file 1hsn [54]). Four  $\alpha$ -helices, called 1, 1', 2, and 3 (underlined residues in Fig. 1 corresponding to the three  $\alpha$ -helices, I, II, and III, stated at the top of Fig. 1 for *D. melanogaster* HMG-D), were predicted to fold in an L shape (see supplemental material S3) (12, 30) in the PfHMGB1 sequence from His 19 to Tyr 90 and in the PfHMGB2 sequence from Ala 23 to Gln 98.

Therefore, all of these computational analyses agreed in suggesting that the two PfHMGB proteins were genuine factors of the HMGB family and may therefore behave as potential architectural factors.

**PfHMGB1 and PfHMGB2 interact with 4H.** Several sets of in vitro assays were used to validate the computational identification. After the genes were cloned, expression of rePfHMGB1 and rePfHMGB2 was carried out in *Escherichia coli* and His proteins were purified as described in Materials and Methods.

First, increasing amounts of both recombinant rePfHMGB1 and rePfHMGB2 (up to 25  $\mu$ M) were incubated with radiolabeled complete 4H to analyze the formation of 4H-rePfHMGB1 and 4H-rePfHMGB2 complexes by EMSA. When 3  $\mu$ M rePfHMGB1 (Fig. 2a) or 0.6  $\mu$ M rePfHMGB2 (Fig. 2b) was added to the reaction, the 4H labeled cruciform became incorporated into a major 4H-PfHMGB1 or 4H-PfHMGB2 retarded band, respectively. These amounts of recombinant factors were used for the subsequent EMSA experiments.

Second, the binding specificities of the 4H-PfHMGB1 and 4H-PfHMGB2 complexes were analyzed by competition experiments, that is, by adding a 100- or 500-fold molar excess of either complete (4H) or incomplete (3H and 2H) cold DNA junctions, after DNA-protein complex formation. A 500-fold excess of the integral 4H structure abolished the interaction of either protein, whereas a 100-fold molar excess shifted the 4H-PfHMGB2 interaction completely but the 4H-PfHMGB1 interaction only weakly (Fig. 2c and d). In addition, the competition with cold 3H and 2H was ineffective upon analysis of PfHMGB1, in contrast to the substantial competition observed in the presence of PfHMGB2. Thus, binding of PfHMGB1 to 4H required an intact crossover-containing structure. Alto-

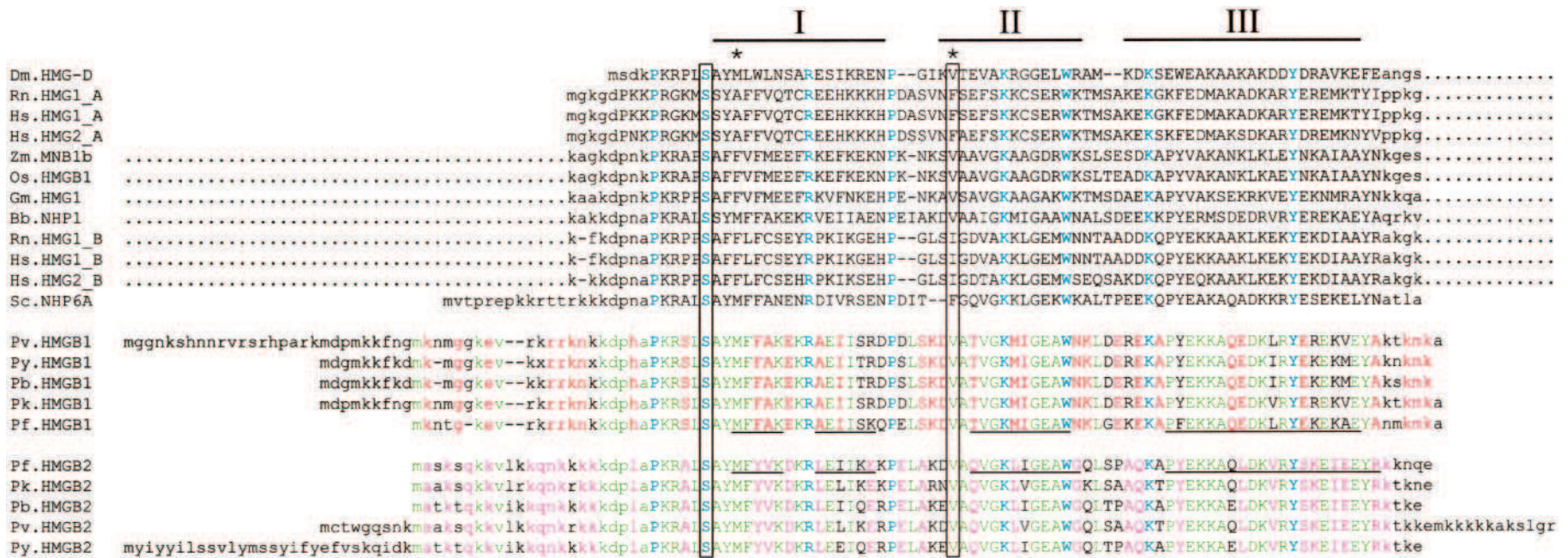


FIG. 1. Multiple alignment of HMG box domains of several eukaryotic HMG proteins and complete sequences of *Saccharomyces* NHP6A and HMGB1 and HMGB2 of diverse *Plasmodium* species. Dashes and dots represent gaps and missing sequences, respectively. Uppercase letters represent HMG box domains identified by the program MotifScan (16). Identities shared by all HMGB sequences presented herein are indicated in bold blue print, by all *Plasmodium* sequences in green, by all *Plasmodium* HMGB1 sequences in red, and by all *Plasmodium* HMGB2 sequences in pink. The boxed amino acids are two crucial determinants that differ between the sequence-specific and the structure-specific HMG box domains: here, a serine and a hydrophobic residue, as found in all non-sequence-specific HMG proteins, whereas all sequence-specific HMG proteins present an asparagine and a hydrophilic residue at these positions. I, II, and III represent the three  $\alpha$ -helices of the *D. melanogaster* HMG-D structure (PDB file 1qrv [47]), and asterisks indicate the two residues of *D. melanogaster* HMG-D that intercalate in the DNA minor groove. Underlined residues in PfHMGB1 and PfHMGB2 represent the four  $\alpha$ -helices modeled by homology with box B of the Chinese hamster HMG1 protein, used as a template (PDB file 1hsn [54]). Abbreviations are as follows: Bb, *Babesia bovis*; Dm, *Drosophila melanogaster*; Gm, *Glycine max*; Os, *Oryza sativa*; Pb, *Plasmodium berghei*; Pf, *Plasmodium falciparum*; Pk, *Plasmodium knowlesi*; Pv, *Plasmodium vivax*; Py, *Plasmodium yoelii*; Rn, *Rattus norvegicus*; Sc, *Saccharomyces cerevisiae*; Zm, *Zea mays*.

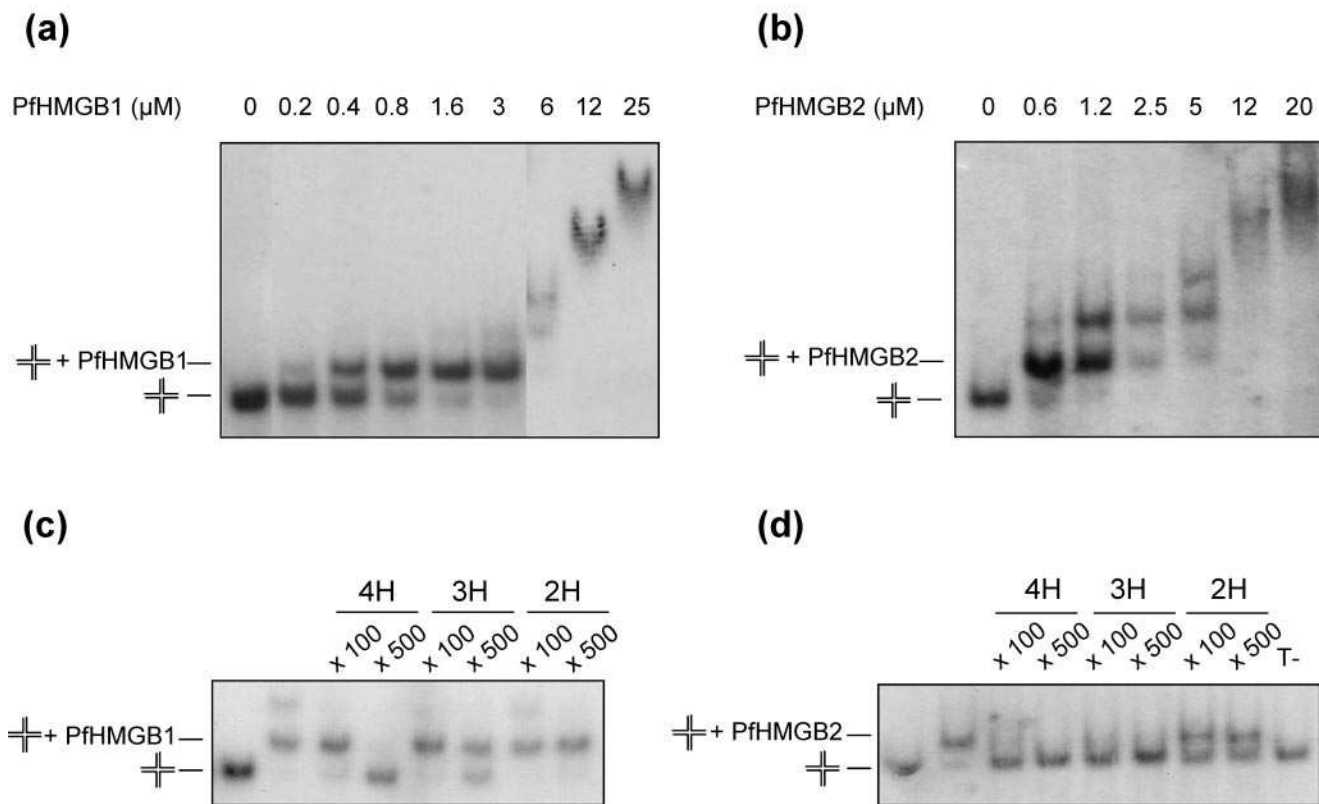


FIG. 2. EMSA interaction between cruciform DNA and either rePfHMGB1 (a) or rePfHMGB2 (b). Increasing concentrations (0 to 25  $\mu\text{M}$ ) of rePfHMGB proteins were incubated with radiolabeled 4H. Competition EMSA experiments were performed between the rePfHMGB1-4H (c) or rePfHMGB2-4H (d) complex and various DNA competitors. After incubation of the rePfHMGB proteins with labeled 4H, either cold incomplete junctions (3H and 2H) or complete 4H were added to the reaction at 100- and 500-fold molar excesses, as indicated above each lane.  $\oplus$ , cruciform DNA. The first lanes correspond to cruciform DNA migration without protein interaction (a to d), and the second lanes correspond to a cruciform DNA-rePfHMGB complex retarded band without competition (c and d).

gether, these results indicated that the two recombinant proteins were able to bind distorted DNA in vitro, even though the nature of their interaction with 4H was quite different, being more efficient and specific for PfHMGB1 than for PfHMGB2.

**PfHMGB1 and PfHMGB2 induce DNA bending.** We compared the efficiencies of increasing concentrations of rePfHMGB1 and rePfHMGB2 to bend and in turn promote T4 DNA ligase-mediated circularization of a labeled synthetic linear DNA fragment. Indeed, when ligase was added to the labeled fragment of around 125 bp, several bands appeared, including a circular DNA form resistant to exonuclease III (Fig. 3a and b). In the presence of exonuclease III alone, which digests only linear DNA molecules, a marked decrease in all labeled bands was observed, showing that in the absence of PfHMGB proteins, only small amounts of minicircles, if any, were produced. In contrast, in the presence of ligase and increasing amounts of PfHMGB proteins, the quantity of minicircles was quite increased, suggesting that both proteins were capable of enhancing DNA flexibility and hence DNA circularization. The capacity for DNA bending is thus an intrinsic property of PfHMGB1 and PfHMGB2. Nevertheless, rePfHMGB1 once again showed greater efficacy (Fig. 3a), since it started to promote circularization at 0.25  $\mu\text{M}$ , at a concentration 10-fold lower than that of rePfHMGB2 (3  $\mu\text{M}$ ), and the maximum

signal was reached with 1  $\mu\text{M}$  compared with 50  $\mu\text{M}$  rePfHMGB2. Moreover, the signal observed with 50 to 100  $\mu\text{M}$  PfHMGB2 (Fig. 3b) was far weaker than that observed with 2  $\mu\text{M}$  PfHMGB1 (Fig. 3a).

**Pfhmgb1 and Pfhmgb2 transcripts are expressed during the *P. falciparum* erythrocytic cycle.** In order to determine the presence of *Pfhmgb1* and *Pfhmgb2* RNA and to characterize them molecularly, we performed a Northern blotting analysis of total RNA prepared from infected red blood cells (Fig. 4a). The lengths of *Pfhmgb1* and *Pfhmgb2* mRNA were estimated at 1.3 and 1.1 kb, respectively. The integrity and quality of total RNA extracts of 3D7 were verified after ethidium bromide staining of the gel. As already described for many other *Plasmodium* messengers, the 5' and 3' untranslated regions of both transcripts are quite long, as the coding regions comprise fewer than 300 nucleotides.

**PfHMGB1 and PfHMGB2 proteins are present in the *P. falciparum* nucleus.** Localization of PfHMGB1 and PfHMGB2 was analyzed by Western blotting analysis using CE and NE of *P. falciparum* asexual stages and specific antisera raised against each recombinant protein. The NE and CE prepared from 3D7 parasites (12  $\mu\text{g}$ ), as well as the recombinant proteins rePfHMGB1 and rePfHMGB2 (50 ng), were fractionated by SDS-PAGE, and after transfer, the membranes were developed

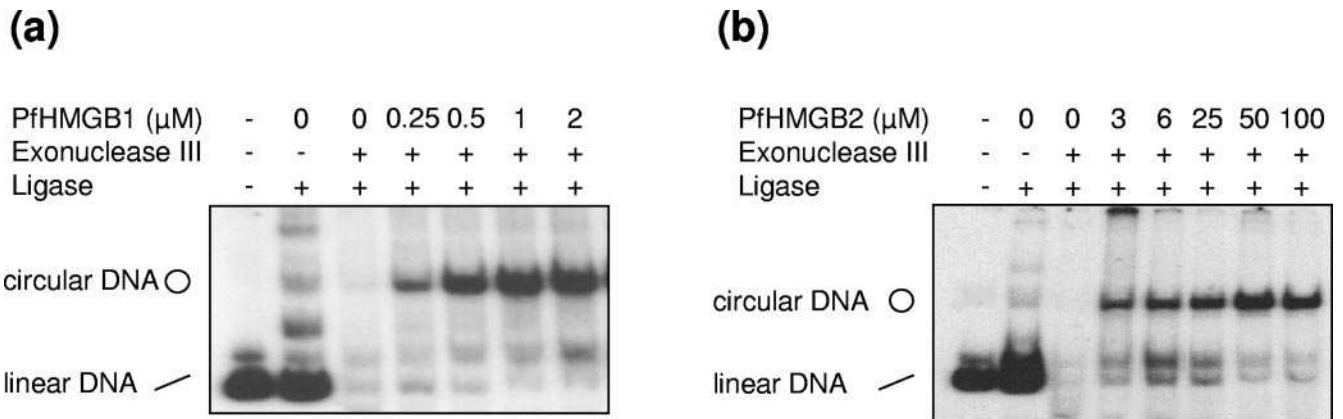


FIG. 3. DNA bending and ligase-mediated circularization assay with either rePfHMGB1 (a) or rePfHMGB2 (b). The  $\gamma$ - $^{32}\text{P}$  5' end-labeled 123-bp DNA fragment was preincubated with increasing amounts of rePfHMGB1 (0 to 2  $\mu\text{M}$ ) or rePfHMGB2 (0 to 100  $\mu\text{M}$ ), followed by ligation with T4 DNA ligase. The ligation products were subjected to electrophoresis after exonuclease III treatment. T4 DNA ligase was added to all samples except that loaded in the first lane. All samples were treated with exonuclease III except samples of the first two lanes. The migration positions of 123-bp linear and 123-bp minicircular DNAs are indicated at left.

with the specific antisera (Fig. 4b). The specificities of the two antibodies were verified. No cross-reaction was observed, as His-PfHMGB1 and His-PfHMGB2 were recognized only by their respective antisera (Fig. 4b, lanes 4 to 5 and lanes 8 to 9 for His-PfHMGB1 and His-PfHMGB2, respectively). The control experiments performed with the two preimmune sera gave no signal (data not shown). Both PfHMGB1 and PfHMGB2

were clearly detected in the NE (lanes 3 and 7), whereas same protein loading of CE does not give any detectable signal (lanes 2 and 6). The quality of the CE preparations was controlled by detection of the HSP protein (70 kDa) (lane 1). The apparent molecular mass of both PfHMGB proteins in the nuclear extracts, around 12 kDa, was in good agreement with the theoretical molecular masses of 11.3 kDa and 11.5 kDa for

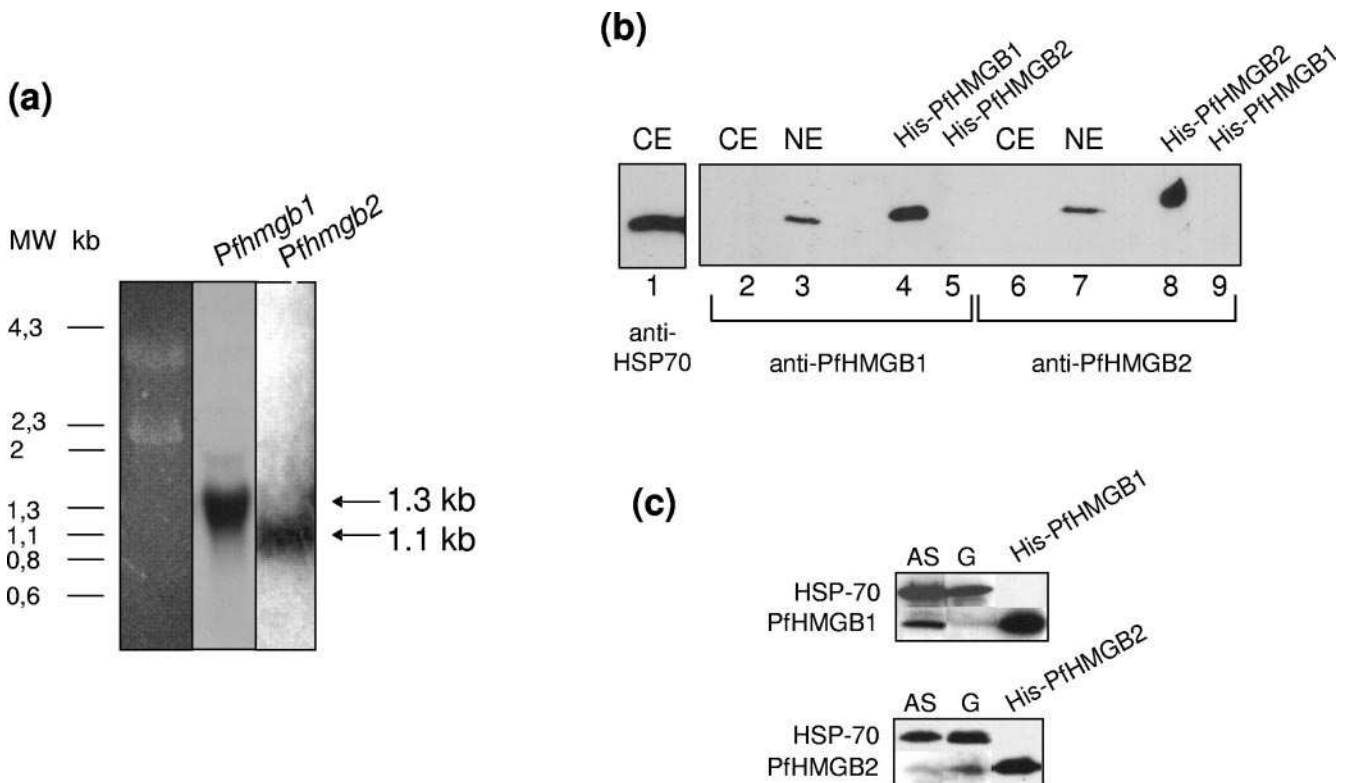


FIG. 4. Characterization of *Pfhmgb1* and *Pfhmgb2* transcripts by Northern blotting (a) and of their corresponding proteins, His-PfHMGB1 and His-PfHMGB2, by Western blotting (b). By SDS-PAGE, the apparent molecular mass of PfHMGB factors was approximately 12 kDa, whereas that of HSP70 was approximately 70 kDa. (c) Expression of PfHMGB1 and PfHMGB2 in total lysates prepared from asexual- versus gametocyte-enriched cultures. HSP70 protein expression was used for normalization of sample loading. AS, asexual stages; G, gametocytes.

PfHMGB1 and PfHMGB2, respectively (lanes 3 and 7), slightly smaller than the recombinant His proteins (12.7 kDa and 12.9 kDa, respectively).

Localization of the factors within the nucleus was confirmed by immunofluorescence analysis of unsynchronized parasite cultures containing all asexual stages of *P. falciparum*. Control experiments carried out in the presence of either preimmune sera or the secondary antibody alone showed no signal (data not shown), in contrast to the labeling obtained with both anti-PfHMGB sera coupled with nucleus-specific DAPI staining (Fig. 5). The two proteins appeared to be present mainly in the parasite nuclei, as shown by the merge of DAPI (lanes a and b), and, in contrast to the HSP70 signal, also readily detectable in the cytoplasm, as indicated by the red fluorescence observed in the superposition insert (lane c).

**PfHMGB1 and PfHMGB2 factors are expressed differentially in the asexual and gametocyte stages.** We compared the expression levels of both factors within total lysates from mixed asexual versus gametocyte cultures by means of Western blotting experiments (Fig. 4c). Each factor was evaluated as detailed in Materials and Methods via HSP70 expression. The level of PfHMGB1 expression was clearly lower in gametocytes than in mixed asexual stages, in contrast to PfHMGB2, whose expression was higher in gametocytes. Indeed, after densitometric quantification of this representative experiment via HSP70 protein normalization (QuantiScan Biosoft 2.1), the normalized values of protein expression were 71%/29% for PfHMGB1 and 27%/73% for PfHMGB2, respectively, when asexual and gametocyte cultures were compared.

We also compared the localizations of both factors in asexual (red immunofluorescence) and gametocyte (green immunofluorescence) stages. As already mentioned, the two PfHMGB factors (Fig. 5, lanes a and b) appeared to be located mainly in the nucleus of the asexual stages (rings, trophozoites, and schizonts), whereas the HSP70 protein (lane c) was also found in the parasite cytoplasm. Surprisingly, in addition to its nuclear localization, PfHMGB2 could also be readily detected within the cytoplasm of different stages (IV and V) of gametocytes (lanes e), as also observed for the HSP70 protein (lane f), whereas PfHMGB1 was associated mainly with the nucleus of gametocytes, as in asexual parasites (Fig. 5, lanes d and a).

## DISCUSSION

The two *Plasmodium* HMG proteins belong to the HMGB subfamily and comprise only one HMG box domain, like proteins of plants and several proteins of yeast and *Drosophila*. The parasite HMG box domains are quite similar to those of all eukaryotic organisms and show characteristic residues that are important for DNA binding and bending, as shown for the *Drosophila* HMG-D protein by Murphy and colleagues (47). The presence of Ser-10 and Val-32 (boxed residues in Fig. 1, according to the numbering of HMG-D) allowed us to assign the *Plasmodium* factors to the architectural HMGB family, as also shown by the phylogenetic analysis performed with 159 HMG box domains (see supplemental material S1). Ser-10 forms water-mediated hydrogen bonds with DNA. The hydrophobic residues Val-32 and Met-13 (Fig. 1, residues with asterisks) partially intercalate between two base pairs, introducing two successive kinks into the bound DNA that enhance the

more uniform bend associated with the widening of the minor groove. In addition to the HMG box domain, the parasite proteins exhibit a basic extension N terminal to the HMG box domain and apparently no acidic C-terminal tail, as also observed for the yeast NHP6A protein. The HMG box domains bind the DNA minor groove, and in the NHP6A and HMG-D proteins, the basic extension binds in the compressed major groove on the face of the helix opposite the widened minor groove (1, 13), so as to stabilize the HMG box domain-induced bending (39) and consequently facilitate circularization (22, 68).

Since the two *Plasmodium* HMGB factors exhibited only one HMG box domain, we asked whether the *Plasmodium* box was more similar to box A or box B of the metazoan HMGB that encompasses two HMG box domains in tandem. All analyses converge to the same conclusion: the *Plasmodium* HMG box domain more closely resembles box B. When the phylogenetic analysis (see supplemental material S2) was performed with the HMG box domains of PfHMGB1 and PfHMGB2 and of various proteins containing box A and box B, the *Plasmodium* factors clearly clustered with all B boxes. For the human HMGB1, it was reported that structure-specific binding to the four-way DNA junction was mediated by the A domain (65) and that box B, flanked by the basic region, displayed a marked DNA recognition activity (69). Hence, the short N-terminal basic domain of the two *Plasmodium* nuclear factors might govern their interaction with distorted DNA and subsequent DNA bending.

In addition, box B of the HMGB of vertebrates was reported to behave as a potent proinflammatory cytokine (64). The tumor necrosis factor (TNF)-stimulating activity was mapped to the KDPNAPKRPPSAFFLFCSEY sequence, corresponding to the first 20 aa of box B in human HMGB1 factor, according to the numbering of Li et al. (36). In *Plasmodium*, a domain sharing 75% and 70% identical or strongly similar residues with the TNF-stimulating domain of the human factor was found at the N-terminal position of the HMG box domains of PfHMGB1 and PfHMGB2, respectively. Presently, experiments are under way to analyze whether the two *Plasmodium* factors exhibit TNF-stimulating function.

An automatic three-dimensional structural prediction was performed, and for both *Plasmodium* factors, four  $\alpha$ -helices, called 1, 1', 2, and 3, were predicted (underlined residues of PfHMGB1 and PfHMGB2 of Fig. 1), folding in an L shape in the HMG box domain. HMG box domains have actually exhibited only three  $\alpha$ -helices (66), but even if four  $\alpha$ -helices were predicted in the two parasite proteins, their positions would be in good agreement with those of *Drosophila melanogaster* HMG-D (PDB file 1qrv [47]), which are indicated (I, II, and III) at the top of Fig. 1. In addition, the sample of the reference template used for nuclear magnetic resonance contained a molecule of  $\beta$ -mercaptoethanol attached to the single cysteine of the protein. The molecule of  $\beta$ -mercaptoethanol, which reduces the affinity of the HMG box domain for the four-way DNA junction, also disrupted the usual first helix (see supplemental material S3). For that reason, helices 1 and 1' could be regarded as a single, kinked helix.

These computational analyses were concordant in suggesting that the *Plasmodium* proteins were genuine architectural HMGB factors close to box B in being able to bind and bend DNA. In vitro analyses performed with both recombinant pro-

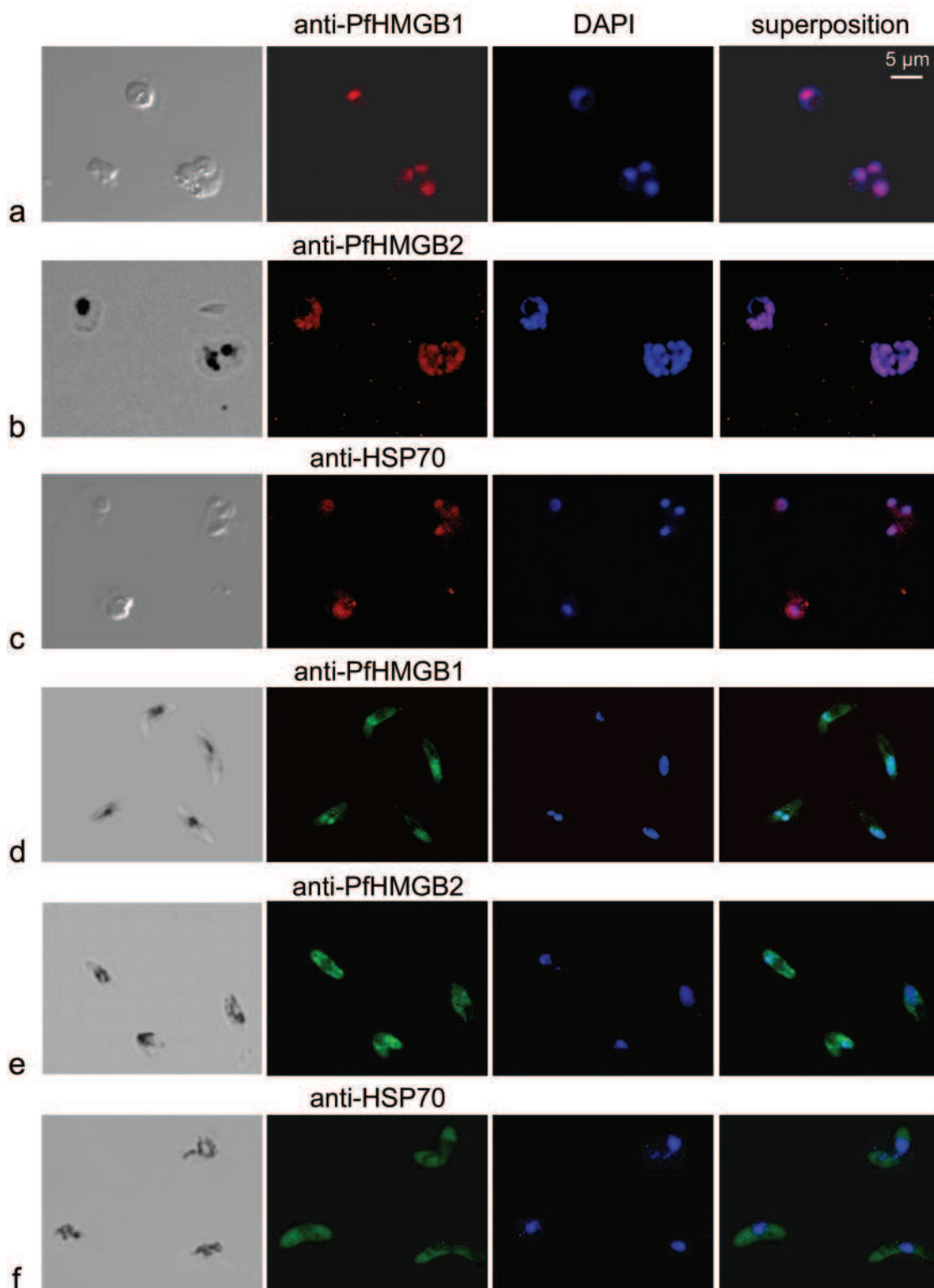


FIG. 5. Immunofluorescence localization of PfHMGB1 and PfHMGB2 in asexual (a, b, and c) and sexual (d, e, and f) stages of *Plasmodium* erythrocytic development. Paraformaldehyde-fixed parasites were labeled with mouse anti-PfHMGB1 and anti-PfHMGB2 antibodies (1:200) and FITC-conjugated anti-mouse IgG (1:100); DNA was stained with DAPI (1:100). Merged fluorescent signals are shown in the “superposition” column. Cells were visualized by phase-contrast (a and c) or transmission (b, d, e, and f) microscopy. Panels: a, trophozoites; b, trophozoite and schizont; c, trophozoites; d to f, gametocytes. Anti-PfHMGB and anti-HSP70 fluorescence is red for panels a to c and green for panels d to f.

teins established that they were indeed able to interact with distorted DNA structures (Fig. 2) and bend linear DNA (Fig. 3), leading to the validation of the computational data (Fig. 1 and supplemental materials S1, S2, and S3). In contrast, EMSA performed with labeled linear DNA binding sites reported to interact specifically with members of the HMGB subfamily comprising the usual TF SOX and SRY (63) gave no detectable retarded complexes (data not shown).

Therefore, it can reasonably be assumed that these architectural HMGB factors might play a role in the remodeling of chromatin. In eukaryotes, one proposed mechanism is that the HMGB nuclear factors might change the nucleosome structure and relax the wrapped DNA so as to enhance the accessibility of the remodeling complexes to chromatin and facilitate interaction of TF with their binding sites (for a review, see reference 62). It has also been observed that the interplay between these factors and the linker histone H1 modulates the balance between alternative conformations of the chromatin, histone H1 enhancing chromatin compaction, in contrast to HMGB. In *Plasmodium*, even though the gene for the H1 linker histone has not yet been annotated, along with 60% of the 5,300 predicted genes, a putative histone H1-like protein might be present and counteract the function of *Plasmodium* HMGB. Indeed, the H1 histones are evolutionarily conserved in metazoans but substantially divergent in protists (51). Some protists appeared to have only a lysine-rich basic protein, whose composition is similar to some of the histone H1-like proteins from eubacteria, animals, and plants (26).

Le Roch et al. reported differential expression of both transcripts. *PfHmgb1* is preferentially expressed during the erythrocytic asexual stages, in contrast to *PfHmgb2*, in which preferential expression occurs in gametocytes (35). Figure 4c shows that expression of the two corresponding proteins is closely related to the level of transcripts and is differentially expressed in mixed asexual and gametocyte stages. The Western blot of cytoplasmic and nuclear extracts prepared from asexual stages (Fig. 4b) and immunofluorescence of asexual and gametocyte stages (Fig. 5a, b, d, and e) revealed that the two factors are localized mainly in the nucleus. Furthermore, PfHMGB2 was clearly detected in the cytoplasm of gametocytes (Fig. 5d and e).

In addition to the differences in the levels of expression and localization within asexual and sexual parasites, these two factors exhibited different affinities when interacting and bending DNA (Fig. 3), PfHMGB2 being less efficient, at least when examined in vitro. All of these results argue in favor of little if any redundancy between the two proteins and in favor of a role in gametocytogenesis.

In summary, a combination of computational and molecular analyses is needed to increase our knowledge of transcriptional regulation of *Plasmodium* genes involved in crucial steps of asexual and sexual erythrocytic development. This report describes the characterization of two *Plasmodium* HMGB factors that appear to exhibit substantial similarity to architectural factors as regards their biological functions, at least when analyzed in vitro for the capacity to interact with distorted DNA and to bend DNA, even though their capacities to do so appeared to be quite different. As in eukaryotes, HMGB factors in *Plasmodium*, since they are highly conserved through evolution, are probably involved in chromatin remodeling. How-

ever, even though these proteins were observed in asexual and gametocyte stages, their levels of expression were clearly different, with PfHMGB1 likely implicated in proliferation and PfHMGB2 implicated in differentiation of *Plasmodium*. Much more work will be needed to understand the functions of these two proteins, both as nuclear factors and as cytokines. Invalidation of either the factors or the interaction between the PfHMGB and DNA via gene silencing strategies (20) and antagonists, etc., will increase our knowledge of transcriptional regulation as well as our control of the erythrocytic development of the parasite (proliferation and differentiation). It might provide exciting new therapeutic possibilities. In this regard, it is worthy of note that this type of approach, via disruption of DNA/factor interaction, is currently being evaluated for human cancer therapy with a special focus on HMG proteins (4, 40).

#### ACKNOWLEDGMENTS

We thank Stephan Soullier for kind help in the phylogenetic studies, Olivier Silvie for the gift of anti-HSP70 serum, and Justine Masson (Inserm U677, Paris) for help with the immunofluorescence facilities. We are grateful to Jennifer Richardson for critical reading of the manuscript.

M.G. and C.B. were financially supported by the Ministère de l'Éducation Nationale, France. The project was supported by Inserm and UPMC funds to C.V.

#### REFERENCES

- Allain, F. H., Y. M. Yen, J. E. Masse, P. Schultze, T. Dieckmann, R. C. Johnson, and J. Feigon. 1999. Solution structure of the HMG protein NHP6A and its interaction with DNA reveals the structural determinants for non-sequence-specific binding. *EMBO J.* **18**:2563–2579.
- Aravind, L., L. M. Iyer, T. E. Wellem, and L. H. Miller. 2003. *Plasmodium* biology: genomic gleanings. *Cell* **115**:771–785.
- Baxevanis, A. D., and D. Landsman. 1995. The HMG-1 box protein family: classification and functional relationships. *Nucleic Acids Res.* **23**:1604–1613.
- Beckerbauer, L., J. J. Tepe, J. Cullison, R. Reeves, and R. M. Williams. 2000. FR900482 class of anti-tumor drugs cross-links oncoprotein HMG I/Y to DNA in vivo. *Chem. Biol.* **7**:805–812.
- Bianchi, M. E., M. Beltrame, and G. Paonessa. 1989. Specific recognition of cruciform DNA by nuclear protein HMG1. *Science* **243**:1056–1059.
- Boschet, C., M. Gissot, S. Briquet, Z. Hamid, C. Claudel-Renard, and C. Vaquero. 2004. Characterization of PfMyb1 transcription factor during erythrocytic development of 3D7 and F12 *Plasmodium falciparum* clones. *Mol. Biochem. Parasitol.* **138**:159–163.
- Bustin, M. 1999. Regulation of DNA-dependent activities by the functional motifs of the high-mobility-group chromosomal proteins. *Mol. Cell. Biol.* **19**:5237–5246.
- Bustin, M. 2001. Revised nomenclature for high mobility group (HMG) chromosomal proteins. *Trends Biochem. Sci.* **26**:152–153.
- Bustin, M., D. A. Lehn, and D. Landsman. 1990. Structural features of the HMG chromosomal proteins and their genes. *Biochim. Biophys. Acta* **1049**:231–243.
- Callebaut, I., K. Prat, E. Meurice, J. P. Mornon, and S. Tomavo. 2005. Prediction of the general transcription factors associated with RNA polymerase II in *Plasmodium falciparum*: conserved features and differences relative to other eukaryotes. *BMC Genomics* **6**:100.
- Coulson, R. M., N. Hall, and C. A. Ouzounis. 2004. Comparative genomics of transcriptional control in the human malaria parasite *Plasmodium falciparum*. *Genome Res.* **14**:1548–1554.
- Douquet, D., and G. Labesse. 2001. Easier threading through web-based comparisons and cross-validations. *Bioinformatics* **17**:752–753.
- Dow, L. K., D. N. Jones, S. A. Wolfe, G. L. Verdine, and M. E. Churchill. 2000. Structural studies of the high mobility group globular domain and basic tail of HMG-D bound to disulfide cross-linked DNA. *Biochemistry* **39**:9725–9736.
- Dumitriu, I. E., P. Baruah, A. A. Manfredi, M. E. Bianchi, and P. Rovere-Querini. 2005. HMGB1: guiding immunity from within. *Trends Immunol.* **26**:381–387.
- Erondu, N. E., and J. E. Donelson. 1992. Differential expression of two mRNAs from a single gene encoding an HMG1-like DNA binding protein of African trypanosomes. *Mol. Biochem. Parasitol.* **51**:111–118.
- Falquet, L., M. Pagni, P. Bucher, N. Hulo, C. J. Sigrist, K. Hofmann, and A.



- Bairoch. 2002. The PROSITE database, its status in 2002. *Nucleic Acids Res.* **30**:235–238.
17. Felsenstein, J. 1989. PHYLIP—Phylogeny Inference Package (version 3.63). *Cladistics* **5**:164–166.
  18. Gardner, M. J., N. Hall, E. Fung, O. White, M. Berriman, R. W. Hyman, J. M. Carlton, A. Pain, K. E. Nelson, S. Bowman, I. T. Paulsen, K. James, J. A. Eisen, K. Rutherford, S. L. Salzberg, A. Craig, S. Kyes, M. S. Chan, V. Nene, S. J. Shallom, B. Suh, J. Peterson, S. Angiuoli, M. Pertea, J. Allen, J. Selengut, D. Haft, M. W. Mather, A. B. Vaidya, D. M. Martin, A. H. Fairlamb, M. J. Fraunholz, D. S. Roos, S. A. Ralph, G. I. McFadden, L. M. Cummings, G. M. Subramanian, C. Mungall, J. C. Venter, D. J. Carucci, S. L. Hoffman, C. Newbold, R. W. Davis, C. M. Fraser, and B. Barrell. 2002. Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* **419**:498–511.
  19. Gissot, M. 2005. Etude de la régulation transcriptionnelle des gènes lors du cycle érythrocytaire de *Plasmodium falciparum*. Ph.D. thesis. Université Pierre et Marie Curie—Paris VI, Paris, France.
  20. Gissot, M., S. Briquet, P. Refour, C. Boschet, and C. Vaquero. 2005. PfMyb1, a *Plasmodium falciparum* transcription factor, is required for intra-erythrocytic growth and controls key genes for cell cycle regulation. *J. Mol. Biol.* **346**:29–42.
  21. Gnanasekar, M., R. Velusamy, Y. X. He, and K. Ramaswamy. 2006. Cloning and characterization of a high mobility group box 1 (HMGB1) homologue protein from *Schistosoma mansoni*. *Mol. Biochem. Parasitol.* **145**:137–146.
  22. Grasser, K. D., S. H. Teo, K. B. Lee, R. W. Broadhurst, C. Rees, C. H. Hardman, and J. O. Thomas. 1998. DNA-binding properties of the tandem HMG boxes of high-mobility-group protein 1 (HMGI). *Eur. J. Biochem.* **253**:787–795.
  23. Hirtzlin, J., P. M. Farber, and R. M. Franklin. 1994. Isolation of a novel *Plasmodium falciparum* gene encoding a protein homologous to the Tatabinding protein family. *Eur. J. Biochem.* **226**:673–680.
  24. Horn, P. J., and C. L. Peterson. 2002. Molecular biology. Chromatin higher order folding—wrapping up transcription. *Science* **297**:1824–1827.
  25. Horrocks, P., K. Dechering, and M. Lanzer. 1998. Control of gene expression in *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **95**:171–181.
  26. Kasinsky, H. E., J. D. Lewis, J. B. Dacks, and J. Ausio. 2001. Origin of H1 linker histones. *FASEB J.* **15**:34–42.
  27. Khochbin, S. 2001. Histone H1 diversity: bridging regulatory signals to linker histone function. *Gene* **271**:1–12.
  28. Knapp, S., S. Muller, G. Digilio, T. Bonaldi, M. E. Bianchi, and G. Musco. 2004. The long acidic tail of high mobility group box 1 (HMGB1) protein forms an extended and flexible structure that interacts with specific residues within and between the HMG boxes. *Biochemistry* **43**:11992–11997.
  29. Kun, J. F., and R. F. Anders. 1995. A *Plasmodium falciparum* gene encoding a high mobility group protein box. *Mol. Biochem. Parasitol.* **71**:249–253.
  30. Labesse, G., and J. Mornon. 1998. Incremental threading optimization (TITO) to help alignment and modelling of remote homologues. *Bioinformatics* **14**:206–211.
  31. Langst, G., and P. B. Becker. 2004. Nucleosome remodeling: one mechanism, many phenomena? *Biochim. Biophys. Acta* **1677**:58–63.
  32. Lanzer, M., D. de Bruin, and J. V. Ravetch. 1992. Transcription mapping of a 100 kb locus of *Plasmodium falciparum* identifies an intergenic region in which transcription terminates and reinitiates. *EMBO J.* **11**:1949–1955.
  33. Lanzer, M., S. P. Wertheimer, D. de Bruin, and J. V. Ravetch. 1993. *Plasmodium*: control of gene expression in malaria parasites. *Exp. Parasitol.* **77**:121–128.
  34. Le Roch, K., C. Sestier, D. Dorin, N. Waters, B. Kappes, D. Chakrabarti, L. Meijer, and C. Doerig. 2000. Activation of a *Plasmodium falciparum* cdc2-related kinase by heterologous p25 and cyclin H. Functional characterization of a P. falciparum cyclin homologue. *J. Biol. Chem.* **275**:8952–8958.
  35. Le Roch, K. G., Y. Zhou, P. L. Blair, M. Grainger, J. K. Moch, J. D. Haynes, P. De La Vega, A. A. Holder, S. Batalov, D. J. Carucci, and E. A. Winzler. 2003. Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science* **301**:1503–1508.
  36. Li, J., R. Kokkola, S. Tabibzadeh, R. Yang, M. Ochani, X. Qiang, H. E. Harris, C. J. Czura, H. Wang, L. Ulloa, H. S. Warren, L. L. Moldawer, M. P. Fink, U. Andersson, K. J. Tracey, and H. Yang. 2003. Structural basis for the proinflammatory cytokine activity of high mobility group box 1. *Mol. Med.* **9**:37–45.
  37. Li, W. B., D. J. Bzik, H. M. Gu, M. Tanaka, B. A. Fox, and J. Inselburg. 1989. An enlarged largest subunit of *Plasmodium falciparum* RNA polymerase II defines conserved and variable RNA polymerase domains. *Nucleic Acids Res.* **17**:9621–9636.
  38. Li, W. B., D. J. Bzik, M. Tanaka, H. M. Gu, B. A. Fox, and J. Inselburg. 1991. Characterization of the gene encoding the largest subunit of *Plasmodium falciparum* RNA polymerase III. *Mol. Biochem. Parasitol.* **46**:229–239.
  39. Lnenicek-Allen, M., C. M. Read, and C. Crane-Robinson. 1996. The DNA bend angle and binding affinity of an HMG box increased by the presence of short terminal arms. *Nucleic Acids Res.* **24**:1047–1051.
  40. Lotze, M. T., and R. A. DeMarco. 2003. Dealing with death: HMGB1 as a novel target for cancer therapy. *Curr. Opin. Investig. Drugs* **4**:1405–1409.
  41. Lotze, M. T., and K. J. Tracey. 2005. High-mobility group box 1 protein (HMGB1): nuclear weapon in the immune arsenal. *Nat. Rev. Immunol.* **5**:331–342.
  42. Martinez, E. 2002. Multi-protein complexes in eukaryotic gene transcription. *Plant Mol. Biol.* **50**:925–947.
  43. McAndrew, M. B., M. Read, P. F. Sims, and J. E. Hyde. 1993. Characterisation of the gene encoding an unusually divergent TATA-binding protein (TBP) from the extremely A+T-rich human malaria parasite *Plasmodium falciparum*. *Gene* **124**:165–171.
  44. Meissner, M., and D. Soldati. 2005. The transcription machinery and the molecular toolbox to control gene expression in *Toxoplasma gondii* and other protozoan parasites. *Microbes Infect.* **7**:1376–1384.
  45. Morales, M., E. Onate, M. Imschenetzky, and N. Galanti. 1992. HMG-like chromosomal proteins in *Trypanosoma cruzi*. *J. Cell. Biochem.* **50**:279–284.
  46. Muller, S., P. Scaffidi, B. Degryse, T. Bonaldi, L. Ronfani, A. Agresti, M. Beltrame, and M. E. Bianchi. 2001. New EMBO members' review. The double life of HMGB1 chromatin protein: architectural factor and extracellular signal. *EMBO J.* **20**:4337–4340.
  47. Murphy, F. V. T., R. M. Sweet, and M. E. Churchill. 1999. The structure of a chromosomal high mobility group protein-DNA complex reveals sequence-neutral mechanisms important for non-sequence-specific DNA recognition. *EMBO J.* **18**:6610–6618.
  48. Nemeth, A., and G. Langst. 2004. Chromatin higher order structure: opening up chromatin for transcription. *Brief. Funct. Genomics Proteomics* **2**:334–343.
  49. Osta, M., L. Gannoun-Zaki, S. Bonnefoy, C. Roy, and H. J. Vial. 2002. A 24 bp cis-acting element essential for the transcriptional activity of *Plasmodium falciparum* CDP-diacylglycerol synthase gene promoter. *Mol. Biochem. Parasitol.* **121**:87–98.
  50. Paillard, F., G. Sterkers, and C. Vaquero. 1990. Transcriptional and post-transcriptional regulation of TcR, CD4 and CD8 gene expression during activation of normal human T lymphocytes. *EMBO J.* **9**:1867–1872.
  51. Parseghian, M. H., and B. A. Hamkalo. 2001. A compendium of the histone H1 family of somatic subtypes: an elusive cast of characters and their characteristics. *Biochem. Cell Biol.* **79**:289–304.
  52. Ponnudurai, T., J. H. Meuwissen, A. D. Leeuwenberg, J. P. Verhave, and A. H. Lensen. 1982. The production of mature gametocytes of *Plasmodium falciparum* in continuous cultures of different isolates infective to mosquitoes. *Trans. R. Soc. Trop. Med. Hyg.* **76**:242–250.
  53. Ragab, A., and A. Travers. 2003. HMG-D and histone H1 alter the local accessibility of nucleosomal DNA. *Nucleic Acids Res.* **31**:7083–7089.
  54. Read, C. M., P. D. Cary, C. Crane-Robinson, P. C. Driscoll, and D. G. Norman. 1993. Solution structure of a DNA-binding domain from HMGI. *Nucleic Acids Res.* **21**:3427–3436.
  55. Silvie, O., J. F. Franetich, S. Charrin, M. S. Mueller, A. Siau, M. Bodescot, E. Rubinstein, L. Hannoun, Y. Charoenvit, C. H. Kocken, A. W. Thomas, G. J. Van Gemert, R. W. Sauerwein, M. J. Blackman, R. F. Anders, G. Pluschke, and D. Mazier. 2004. A role for apical membrane antigen 1 during invasion of hepatocytes by *Plasmodium falciparum* sporozoites. *J. Biol. Chem.* **279**:9490–9496.
  56. Soullier, S., P. Jay, F. Poulat, J. M. Vanacker, P. Berta, and V. Laudet. 1999. Diversification pattern of the HMG and SOX family members during evolution. *J. Mol. Evol.* **48**:517–527.
  57. Su, X. Z., Y. Wu, C. D. Sifri, and T. E. Wellems. 1996. Reduced extension temperatures required for PCR amplification of extremely A+T-rich DNA. *Nucleic Acids Res.* **24**:1574–1575.
  58. Thomas, J. O., and A. A. Travers. 2001. HMGI and 2, and related 'architectural' DNA-binding proteins. *Trends Biochem. Sci.* **26**:167–174.
  59. Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**:4673–4680.
  60. Thomsen, M. S., L. Franssen, D. Launholt, P. Fojan, and K. D. Grasser. 2004. Interactions of the basic N-terminal and the acidic C-terminal domains of the maize chromosomal HMGB1 protein. *Biochemistry* **43**:8029–8037.
  61. Trager, W., and J. B. Jensen. 1976. Human malaria parasites in continuous culture. *Science* **193**:673–675.
  62. Travers, A. A. 2003. Priming the nucleosome: a role for HMGB proteins? *EMBO Rep.* **4**:131–136.
  63. van de Wetering, M., M. Oosterwegel, K. van Norren, and H. Clevers. 1993. Sox-4, an Sry-like HMG box protein, is a transcriptional activator in lymphocytes. *EMBO J.* **12**:3847–3854.
  64. Wang, H., O. Bloom, M. Zhang, J. M. Vishnubhakat, M. Ombrellino, J. Che, A. Frazier, H. Yang, S. Ivanova, L. Borovikova, K. R. Manogue, E. Faist, E. Abraham, J. Andersson, U. Andersson, P. E. Molina, N. N. Abumrad, A. Sama, and K. J. Tracey. 1999. HMGB-1 as a late mediator of endotoxin lethality in mice. *Science* **285**:248–251.
  65. Webb, M., and J. O. Thomas. 1999. Structure-specific binding of the two tandem HMG boxes of HMGI to four-way junction DNA is mediated by the A domain. *J. Mol. Biol.* **294**:373–387.
  66. Weir, H. M., P. J. Kraulis, C. S. Hill, A. R. Raine, E. D. Laue, and J. O. Thomas. 1993. Structure of the HMG box motif in the B-domain of HMGI. *EMBO J.* **12**:1311–1319.

67. **Wu, Q., W. Zhang, K. H. Pwee, and P. P. Kumar.** 2003. Rice HMGB1 protein recognizes DNA structures and bends DNA efficiently. *Arch. Biochem. Biophys.* **411**:105–111.
68. **Yen, Y. M., B. Wong, and R. C. Johnson.** 1998. Determinants of DNA binding and bending by the *Saccharomyces cerevisiae* high mobility group protein NHP6A that are important for its biological activities. Role of the unique N terminus and putative intercalating methionine. *J. Biol. Chem.* **273**:4424–4435.
69. **Yoshioka, K., K. Saito, T. Tanabe, A. Yamamoto, Y. Ando, Y. Nakamura, H. Shirakawa, and M. Yoshida.** 1999. Differences in DNA recognition and conformational change activity between boxes A and B in HMG2 protein. *Biochemistry* **38**:589–595.

## DISCUSSION

L'absence de vaccin efficace contre le paludisme ainsi que l'apparition grandissante des résistances de *P. falciparum* face aux anti-paludiques conduisent à l'étude de nouvelles voies pour combattre la maladie. Il est donc nécessaire d'étudier les mécanismes moléculaires qui dirigent le cycle cellulaire du parasite. Dans cette optique, la compréhension du cycle érythrocytaire chez l'homme semble essentielle car il est responsable de la multiplication intense du parasite à l'origine de la maladie et de ses complications sévères, le neuropaludisme. De plus, le développement érythrocytaire est le point de départ de la dissémination du parasite qui se différencie en gamétocytes.

Comme nous l'avons vu dans l'Introduction, *P. falciparum* subit de nombreuses modifications tant morphologiques que métaboliques au cours de son cycle complexe de développement. Au cours de la phase érythrocytaire, chaque étape de prolifération et de différenciation peut être caractérisée par l'expression préférentielle de certains gènes (Bozdech et al. 2003a; Le Roch et al. 2003). La capacité de cet eucaryote unicellulaire à s'adapter aux différents environnements auxquels il est confronté en prenant des formes si variées dénote un contrôle très subtil de l'expression des gènes. Bien que *P. falciparum* soit un organisme encore assez mal connu du point de vue génétique (60 % des phases ouvertes de lecture n'ont pas encore d'annotation), plusieurs études ont montré des caractéristiques communes aux autres eucaryotes comme l'organisation de la chromatine en nucléosomes, la structure des ARNm et la structure bi-partite des promoteurs. L'identification dans le génome de plus de 150 phases ouvertes de lecture [Annexes I.1-2, p. 166-169 et (Coulson et al. 2004)] codant pour des acteurs de la régulation transcriptionnelle des gènes suggère que ce contrôle joue vraisemblablement un rôle déterminant dans le cycle de vie de *P. falciparum*. La diversité des voies disponibles à *P. falciparum* pour assurer la régulation transcriptionnelle de ses gènes semble finalement aller à l'encontre des conclusions sur l'apparente pauvreté en FT (Aravind et al. 2003; Coulson et al. 2004). L'analyse des structures secondaires des séquences de *P. falciparum* par HCA (Hydrophobic Cluster Analysis) pour la prédiction de nouvelles protéines orthologues semble une voie à suivre pour participer à l'annotation des protéines de *P. falciparum* (Baussand et al. 2006).

Au cours de ce travail de thèse, nous avons abordé divers aspects de la régulation génétique par l'étude de certains acteurs de la régulation transcriptionnelle dans les événements clés du cycle érythrocytaire : la prolifération asexuée et la différenciation sexuée.

Lorsque le projet sur l'analyse du transcriptome de *P. falciparum* a commencé au laboratoire, la séquence du génome du parasite n'était que très partielle. Aucune base de données ordonnée n'était alors disponible et le choix des gènes à étudier a donc été fait avant l'achèvement de la séquence génomique de *P. falciparum* (Octobre 2002). Il a fallu définir et recenser les gènes prédits pour leur rôle dans la régulation du cycle cellulaire. Nous avons aussi commencé à annoter des séquences pouvant coder pour des facteurs de transcription et des kinases potentiels. Le choix de familles de gènes potentiellement impliqués dans la régulation transcriptionnelle, traductionnelle, dans la transduction du signal et le cycle cellulaire, a été motivé par la pertinence de ces voies de régulation de l'expression génétique au sens large dans le contrôle du développement et de la différenciation. Une bio-puce thématique a donc été élaborée dans le but de comparer les profils d'expression d'environ 150 gènes de deux clones de *P. falciparum* ayant un potentiel de différenciation en gamétocytes différent. En effet, le clone F12 bien qu'issu du clone 3D7, est déficient pour la production des gamétocytes contrairement au clone parental (Alano et al. 1995). Les profils d'expression de ces deux clones ont été comparés à cinq temps précis du cycle érythrocytaire post-invasion, le stade anneau (10-14 h), le stade trophozoïte jeune (20-22 h), le stade trophozoïte âgé (26-28 h), le stade schizonte jeune (32-34 h) et le stade schizonte âgé (40-44 h). Les parasites ont été cultivés dans des conditions de culture n'induisant pas la production de gamétocytes chez le clone 3D7. La technique des puces à ADN a été couplée à l'utilisation de la radioactivité comme marqueur différentiel des ADNc car la limite de détection théorique de ce marqueur permettait une plus grande fiabilité des données pour des gènes dont l'expression risquait d'être faible comme pour les facteurs de transcription que nous avons annotés.

Nous avons ainsi comparé tous les profils d'expression obtenus pour le clone 3D7 avec les deux publications décrivant ces profils au cours du cycle complet du parasite pour les clones 3D7 et HB3 (Bozdech et al. 2003a; Le Roch et al. 2003). Cette comparaison indique

que 84 % des gènes présents sur la bio-puce thématique ont le même profil que celui décrit dans au moins une des deux publications. En effet, tous les gènes que nous avons étudiés ne sont pas toujours présents sur les puces du laboratoire de De Risi (Operon) ou de Winzeler (Affymetrix). De plus, l'expression temporelle de certains gènes spécifiques de stades cellulaires a permis de contrôler la qualité de la procédure de synchronisation et les propriétés des clones notamment par l'absence d'expression du transcrite *pfg27*, marqueur de la gamétoctogénèse, pour le clone F12 contrairement au clone 3D7.

Enfin, l'objectif principal de nos expériences était de vérifier le pouvoir discriminant de notre approche entre deux clones ayant un même fond génétique. Les données concernant le clone F12 déficient pour la gamétoctogénèse suggère que le phénotype observé pourrait être dû à l'absence ou à l'altération de l'expression de gènes impliqués dans le contrôle de l'expression de gènes gouvernant la différenciation en gamétoctes. Existe-t-il alors une prédisposition potentielle à la gamétoctogénèse sans induction de celle-ci ? L'intérêt s'est porté sur plusieurs facteurs de transcription pouvant avoir un rôle direct ou indirect dans la différenciation du parasite. Malgré tout, les données sur le profil d'expression des transcrits ne présument pas totalement du profil d'expression des protéines correspondantes. En effet, les régulations post-transcriptionnelles et post-traductionnelles peuvent modifier la présence et l'activité des protéines. Les données issues du transcriptome constituent un premier élément de réponse à notre interrogation et devront donc être suivis par l'analyse moléculaire des protéines.

La comparaison des transcriptomes entre les deux clones 3D7 et F12 nous a permis de répertorier d'une part, 106 gènes qui présentent des profils similaires confirmant le fond isogénique de ces clones. Parmi ces gènes, une quinzaine de FT pourrait avoir un rôle important dans le développement asexué du parasite puisqu'ils conservent le même profil d'expression dans le clone HB3 (Bozdech et al. 2003a). La plupart de ces FT ont des profils d'expression similaires entre les clones 3D7 et F12, même si selon les transcrits les profils sont différents : stable pour le gène *pfhmg2* ou bien avec des pics d'expression distincts pour les gènes *pfphD2*, *pfphDB*, *pfmyb3* et *pfkrox* (Fig. 3, p. 273 de l'Article 1). Enfin, parmi tous les gènes analysés, huit ont des profils d'expression différents entre les deux clones 3D7 et F12, dont deux protéines ayant une fonction putative dans la régulation transcriptionnelle et qui pourraient jouer un rôle intéressant.

- *pf7* (PFI0930c) : on observe un niveau stable de ce transcrit lors du cycle du clone 3D7 tandis qu'une forte augmentation de son expression marque le passage à la forme schizonte du clone F12. Ce gène code pour une protéine possédant un motif NAP (Nucleosome Assembly Protein). Chez les eucaryotes, ces protéines chaperonnes des histones ont un rôle fondamental dans l'organisation et le remodelage de la chromatine (Mello and Almouzni 2001). Elles participeraient à la régulation de l'expression des gènes notamment en interagissant avec les facteurs architecturaux HMGB (Agresti and Bianchi 2003). Ces dernières ayant des représentants dans le génome de *P. falciparum* (voir notre étude de ces facteurs nucléaires), cette fonction pourrait être conservée dans le parasite. D'autres activités ont été attribuées aux protéines NAP comme l'activité déacétylase des histones (Seo et al. 2001), l'activité inhibitrice des phosphatases (Saito et al. 1999) et la capacité d'interaction avec les cyclines (Kellogg et al. 1995) leur conférant un rôle mutuel dans la régulation de la transcription et du cycle cellulaire. Des protéines cyclines ont d'ailleurs aussi été caractérisées chez *P. falciparum* (Merckx et al. 2003).

- *constans* (PF14\_0383) : ce transcrit présente un pic d'expression précoce dès le stade jeune trophozoïte dans le clone F12 alors que ce pic n'est atteint qu'au stade jeune schizonte dans le clone 3D7. Ce gène code pour un FT possédant deux motifs en doigt de zinc de type B-box. Ces motifs présents dans la protéine CONSTANS de la plante *A. thaliana* sont responsables de la régulation transcriptionnelle de la floraison vraisemblablement *via* des interactions protéine-protéine (Robson et al. 2001). Les acides aminés déterminants pour cette fonction (Koornneef et al. 1991) sont conservés dans la séquence de la protéine chez *P. falciparum*. L'expression différentielle de ce gène dans les clones 3D7 et F12 pourrait être à l'origine de la modulation de l'expression d'un ensemble de gènes impliqués dans la gamétocytogénèse. Des expériences d'inhibition ou de surexpression de cette protéine chez *P. falciparum* pourrait nous renseigner sur son rôle dans ce processus.

Cette étude a été la première tentative d'analyse comparative entre deux clones 3D7 et F12 par bio-puce, des profils d'expression de gènes intervenant dans la régulation de la transcription et le cycle cellulaire, principalement des facteurs de transcription et des kinases. Ce travail constitue aussi les premiers pas vers l'identification de gènes directement ou indirectement impliqués dans la prédisposition au processus de différenciation en gamétocytes.

Très récemment, d'autres études comparatives entre clones de *P. falciparum* ont été publiées. Les transcriptomes des clones 3D7 et F12 (Silvestrini et al. 2005) ou de clones 3D7 producteurs ou non de gamétocytes présentant des formes tronquées de la protéine de surface du gamétocyte Pfs230 (Eksi et al. 2005) ont été comparés cette fois en condition d'induction de la gamétocytogénèse sur la quasi-totalité du génome en utilisant des oligonucléotides de 70 mers. Afin d'identifier de nouveaux gènes spécifiques des stades précoces de la gamétocytogénèse, les profils d'expression similaires à ceux des marqueurs précoces de la gamétocytogénèse *pfs16* et *pfg27* ont été analysés. Dans l'étude de F. Silvestrini, figurent notamment des gènes codant pour des enzymes de modification (kinases, phosphatases, transférases) et trois gènes annotés comme facteurs de transcription généraux (PF11\_0477, MAL7P1.86, PFB0290c). Deux nouvelles protéines Pfpeg-3 et Pfpeg-4 ont été identifiées comme spécifiques des gamétocytes de stades I et II, respectivement. Les travaux d'Eksi et al. se sont focalisés sur une famille de gènes subtélomériques surexprimés lors de la transition du stade asexué vers le stade sexué, comme *pfg14.744* et *pfg14.748*. Il est important de noter que ces études n'ont révélé aucun FT impliqué dans la régulation du niveau de transcription des gènes. Vraisemblablement les gènes de ces FT manqueraient dans cette bio-puce élaborée par l'équipe de De Risi (voir Données supplémentaires II p. 170-171) comme c'est le cas des facteurs PfKrox et PfphDB qui sont régulés différemment au cours du développement ou encore Pfb7 dont nous avons retenu un profil différent entre les clones 3D7 et F12 avec notre bio-puce thématique (Fig. 2, p. 272 de l'Article 1).

En parallèle à notre étude avec la bio-puce thématique, nous avons cherché à poursuivre notre analyse comparative entre les transcriptomes des clones 3D7 et F12 avec une bio-puce pan-génomique de *P. falciparum*. Ces expériences en condition de non induction de la gamétocytogénèse, comme dans notre étude de l'article 1, sont réalisées en collaboration avec E. Bischof et P. David grâce à la participation du Consortium de la Plateforme Post-génomique *Plasmodium Pasteur* (P4). L'analyse des données expérimentales est en cours. Nous espérons obtenir un groupe de gènes exprimés différemment entre les clones 3D7 et F12 en analysant plus particulièrement les gènes impliqués potentiellement dans la régulation transcriptionnelle de la gamétocytogénèse. De plus, l'analyse des promoteurs des gènes groupés suivant leurs profils d'expression pourrait conduire à la découverte de motifs conservés responsables de la régulation transcriptionnelle (en

collaboration avec J. Van Helden). Cette approche a été suivie en utilisant le transcriptome des gamétocytes du clone 3D7 et de l'isolat NF54. L'analyse des régions promotrices d'un groupe de 246 gènes, dont la majorité n'est pas annotée, a révélé la présence notamment d'une séquence palindromique TGTANNTACA sur 65 de ces gènes (Young et al. 2005). L'assomption que ce motif pourrait être considéré comme une séquence *cis*-régulatrice fonctionnelle est renforcé par le fait que la délétion de la région contenant ce motif dans le promoteur *pfs25* conduit à une diminution significative de l'activité promotrice de ce gène (Dechering et al. 1999).

Jusqu'à ce jour, peu de facteurs de transcription à l'exclusion des facteurs généraux de la transcription, ont été annotés chez *P. falciparum* à l'image du faible pourcentage (1.3 %) de protéines assignées à la régulation de la transcription annoncé par R. M. Coulson (Coulson et al. 2004), contre 4 % chez la levure *S. cerevisiae*, dont le génome possède un nombre de gènes équivalent. Notre exploration du génome nous a permis d'annoter plusieurs membres de deux familles de facteurs nucléaires et d'analyser leur fonction biologique. Les premiers appartenant à la Famille Myb sont spécifiques de séquence d'ADN et interviennent dans la régulation positive ou négative du niveau de la transcription. Les seconds appartenant à la famille HMGB sont spécifiques de structure d'ADN et participent au remodelage de la chromatine.

Nous avons alors entrepris l'étude d'un acteur de la régulation transcriptionnelle, PfMyb1, un facteur de transcription qui se lie spécifiquement à des séquences d'ADN, annoté dans notre équipe et défini par ces domaines tryptophane comme appartenant à la famille des protéines Myb de la classe Hélice-Tour-Hélice. La caractérisation moléculaire de ce FT a permis de montrer que le gène *pfmyb1* est exprimé tout au long du cycle érythrocytaire des clones 3D7 et F12, avec un pic d'expression au stade trophozoïte, comme pour le clone HB3 (Bozdech et al. 2003a). La cinétique d'expression de la protéine apparaît similaire à celle de son transcrit pour les deux clones étudiés dans des expériences complémentaires (p. 102). Aussi, la conservation de la séquence de PfMyb1 parmi les différentes espèces de *Plasmodium* s'ajoute aux données précédentes pour souligner l'importance de cette protéine au cours du cycle érythrocytaire.



Pour activer ou réprimer la transcription d'un ensemble de gènes, les FT doivent se lier à des séquences d'ADN *cis*-régulatrices d'une manière spécifique, *i. e.* MRE pour Myb Regulatory Element. D'une part, nous avons pu montrer que PfMyb1, présente dans les extraits nucléaires du parasite, était capable de se lier à la séquence MRE prototype du gène *mim1* de poulet. D'autre part, deux séquences MRE putatives retrouvées dans les promoteurs des gènes *pfmap1* et *pfcrk1* de *P. falciparum* se sont avérées efficaces pour permettre la formation de complexes comparables à ceux obtenus avec la séquence MRE prototype. La cinétique d'expression de la protéine PfMyb1 pour le clone 3D7 étant similaire au profil des bandes retardées obtenu par retard sur gel, cela nous permet de présumer de la participation de PfMyb1 à ces complexes ADN-protéine. L'activité spécifique liée à l'anticorps anti-PfMyb1 a permis de confirmer cette hypothèse. Nous avons donc montré par ces expériences que la protéine PfMyb1 possède tous les attributs d'un vrai FT de la famille Myb.

Dans la littérature, de nombreux partenaires pour les protéines Myb ont été décrits, tel que les FT à domaine bZip (Tahirov et al. 2002), qui figurent en grand nombre dans le génome de *P. falciparum* (voir Introduction p. 51 et annexes I.1 p. 166). Ces partenaires éventuels ont d'autant plus d'importance que la cinétique d'interaction avec les extraits nucléaires des clones 3D7 et F12 est différente. L'expression de PfMyb1 suivant le même profil dans les deux clones, justifie la disponibilité de la protéine au cours du développement érythrocytaire. En revanche, la disparition de l'interaction avec le MRE dans les stades anneau et schizonte du clone F12 suggère l'absence d'un ou plusieurs partenaires déterminant pour la formation de complexes MRE-protéine (Fig. 2, p. 161 de l'Article 2). Ces partenaires seraient présents dans le clone 3D7 et absents dans le clone F12. Cette différence pourrait avoir un lien avec le phénotype du clone F12. Dans ce cadre, la recherche de partenaires PfMyb1 prend d'autant plus d'importance qu'ils pourraient être impliqués dans le processus de gamétocytogénèse. Des techniques d'immunoprécipitation associées à la spectrométrie de masse sont employées actuellement au laboratoire pour tenter d'identifier des partenaires de PfMyb1. Par ailleurs, l'effet de modifications post-traductionnelles ne peut être exclu, ces dernières pouvant réguler la liaison de PfMyb1 à l'ADN.

**En conclusion**, cette étude a permis de mettre en évidence, pour la première fois chez *P. falciparum*, l'existence d'un facteur de transcription spécifique de séquence présent au

cours du cycle érythrocytaire et possédant tous les attributs de la famille des facteurs Myb très conservée chez les eucaryotes.

Afin de poursuivre l'étude du rôle de PfMyb1 dans le cycle érythrocytaire de *P. falciparum*, nous avons tenté d'interférer avec l'expression de son transcrite et ainsi de sa protéine en utilisant de longs ARN double brin (ARNdb) spécifique d'une partie du gène *pfmyb1*. L'addition d'ARNdb *pfmyb1* au milieu de culture entraîne une inhibition de croissance de 40 % vraisemblablement due à la mort des parasites lors de la transition du stade trophozoïte au stade schizonte. Cette inhibition de croissance de 40 % (Fig. 1, p. 31 de l'Article 3) s'accompagne d'une baisse considérable du niveau de transcrite (75 %) et de protéine PfMyb1 (80 %) dans la culture traitée par l'ARNdb *pfmyb1* (Fig. 2-3, p. 32-33 de l'Article 3). Ces données abondent dans le sens d'un rôle crucial de PfMyb1 pour la survie de *P. falciparum* lors du cycle érythrocytaire.

Afin d'analyser les conséquences de l'inhibition partielle de l'expression du transcrite *pfmyb1* et de sa protéine, nous avons utilisé la bio-puce à ADN thématique. Cette étude a permis d'identifier huit gènes dont l'expression est altérée dans un contexte de sous-expression de la protéine PfMyb1 (Tableau 1, p. 35 de l'Article 3). Tout d'abord, pour s'assurer de la spécificité d'action des ARNdb, l'expression de plusieurs gènes a été comparée entre les cultures témoins et les cultures traitées. Les gènes *pfmyb2* et *pfmyb3*, en particulier, présentant un niveau d'expression similaire dans les deux conditions, il semble que l'action de ces ARNdb soit spécifique de séquence. De plus, 128 autres gènes, dont l'expression est similaire dans les deux conditions, sont autant de marqueurs de synchronisation de la culture et d'arguments pour penser que l'altération de l'expression des huit gènes dans les cultures traitées est bien la conséquence de la baisse de disponibilité de la protéine PfMyb1 et non pas le produit d'un décalage de croissance entre les deux cultures.

Chez *Plasmodium*, le mécanisme par lequel ces ARNdb agissent est encore largement inconnu et l'existence même d'une voie d'action par ARN interférence *per se* est mise en doute liée à l'incapacité de trouver, dans le génome de *P. falciparum*, les enzymes clés de ce mécanisme (Aravind et al. 2003). Pourtant, ces ARNdb ont été utilisés avec succès pour interférer avec l'expression des gènes chez *P. falciparum* (Li et al. 2004; Malhotra et al. 2002),

ce qui laisse penser qu'une machinerie capable de prendre en charge les ARNdb pourrait exister, mais dont les activités n'auraient pas encore été mises à jour. De plus, il est possible que les acteurs de ce processus ne soient pas encore identifiés par manque d'homologie suffisante avec ces protéines chez les autres eucaryotes. L'inhibition des transcrits pourrait être due à un "mécanisme antisens" d'autant plus que la découverte de transcrits anti-sens chez *P. falciparum* (Gunasekera et al. 2004; Patankar et al. 2001) a relancé le débat sur l'éventuelle participation de ces ARN-antisens à la régulation transcriptionnelle (Militello et al. 2005; Ralph et al. 2005) (voir Introduction p.62). Il n'en reste pas moins que de nombreux mécanismes peuvent induire un effet anti-sens (Vanhee-Brossollet and Vaquero 1998). Les oligonucléotides anti-sens ont d'ailleurs montré leur efficacité sur le parasite (Ikemoto et al. 1999). De plus, l'introduction d'un vecteur exprimant une partie du transcrit anti-sens du gène *clag9* entraîne la baisse de l'expression du transcrit et de la protéine ayant pour conséquence l'apparition d'un phénotype de cytoadhérence réduite (Gardiner et al. 2000). L'état actuel des connaissances sur "l'effet anti-sens" chez *P. falciparum* ne permet pas la compréhension des mécanismes moléculaires, néanmoins nous avons choisi cette approche expérimentale d'inactivation de gène par les ARNdb. Cet outil s'est avéré fonctionnel et utile pour révéler la modulation de huit gènes dans un contexte d'inactivation du transcrit *pfmyb1* au cours du cycle érythrocytaire.

L'expression de ces huit gènes serait donc sous la dépendance directe ou indirecte du facteur de transcription PfMyb1. Parmi ces gènes, la surexpression du messager *pfpk5* et la sous-expression du transcrit *pcna1* pourraient contribuer d'une manière coopérative à l'altération du développement des parasites et finalement participer à induire la mort lors de leur passage au stade schizonte. En effet, les protéines produites par ces deux gènes seraient responsables de la régulation du cycle cellulaire. La protéine PfPk5, dont la surexpression a été vérifiée par Western blot (Fig. 6, p. 35 de l'Article 3), paraît impliquée dans la régulation du cycle cellulaire et de la synthèse d'ADN chez *P. falciparum* (Graeser et al. 1996), tout comme la protéine PCNA chez les eucaryotes (Kelman 1997).

Afin d'identifier les gènes dont l'expression serait directement régulée par PfMyb1, nous avons réalisé des expériences d'immuno-précipitation de la chromatine. Parmi les huit gènes, dont l'expression est altérée en présence d'ARNdb *pfmyb1* dans la culture, six de leur promoteur semblent interagir avec la protéine PfMyb1. Cette donnée indique clairement que

PfMyb1 joue bien le rôle d'un FT en se liant à l'ADN dans le noyau du parasite et participe à la régulation transcriptionnelle de certains gènes comme *pfpk5* et *pcna1*.

Les promoteurs de deux gènes, dont l'expression est altérée dans un contexte de baisse de disponibilité de PfMyb1, n'ont pu être mis en évidence dans ces expériences. Dans ces deux cas, l'interaction PfMyb1-promoteur serait éventuellement trop faible pour permettre la réaction d'immunoprécipitation de la chromatine.

Les cinq promoteurs interagissant avec PfMyb1 possèdent tous un MRE à l'exception du promoteur *pgk*. En effet, l'analyse bioinformatique de la séquence de ce promoteur n'a pas pu nous permettre d'identifier des séquences correspondant au consensus du MRE en dépit de la présence de PfMyb1 à son promoteur. Cette donnée implique que PfMyb1 posséderait d'autres moyens de se lier à l'ADN que par le biais d'un MRE. Cette caractéristique a déjà été décrite pour une protéine de la famille Myb, Bas1p de *S. cerevisiae* (Hovring et al. 1994), mais aussi pour d'autres protéines contenant des domaines de liaison à l'ADN de type Myb (Ohi et al. 1998). Cette flexibilité quant à la liaison à un promoteur ne possédant pas de site de liaison spécifique pourrait être aussi une caractéristique de PfMyb1.

**En conclusion**, le fait que PfMyb1 régule directement le gène *pfpk5* et les cinq autres gènes implique qu'il possède la capacité d'induire soit l'activation, soit la répression de la transcription. Cette étude nous a permis de caractériser un facteur de transcription spécifique de séquence chez *P. falciparum*. Nous avons aussi pu comprendre l'importance de son rôle lors du cycle érythrocytaire, et notamment dans la régulation positive ou négative de gènes clés pour le cycle cellulaire. Ces résultats soulignent aussi que la régulation des gènes chez *P. falciparum* possède des spécificités, à l'image de la liaison de PfMyb1 à l'ADN dont les déterminants restent encore flous. Ce travail révèle la première ébauche à partir de données expérimentales d'un réseau de régulation chez *P. falciparum*. Un travail identique pourra être mené en utilisant des puces à ADN représentant tout le génome de *P. falciparum*, permettant ainsi d'accumuler plus de données sur les gènes régulés par PfMyb1. De plus, des expériences d'immunoprécipitation de la chromatine au niveau de tout le génome (ChIP on chip) nous renseigneraient sur tous les sites de liaisons occupés par PfMyb1 et nous permettraient de mieux comprendre les déterminants de la liaison de PfMyb1 à l'ADN.

Au fil de notre recherche sur les acteurs de la régulation transcriptionnelle, notre intérêt s'est aussi porté sur les facteurs qui se lient à l'ADN sans spécificité de séquence mais avec une spécificité de structure. Nous avons identifié quatre phases ouvertes de lecture dont deux codant pour de petites protéines contenant un seul domaine de liaison à l'ADN ou motif fonctionnel "HMG-box", que nous avons appelé PfHMGB1 et PfHMGB2. Chez les eucaryotes et en particulier chez les vertébrés, les facteurs de type HMG ont été décrits pour leur rôle dans la régulation de l'expression des gènes (Bustin 1999). Le profil d'expression tout à fait séduisant de ces deux gènes au cours du cycle érythrocytaire du clone 3D7 a immédiatement retenu notre attention (Le Roch et al. 2003). En effet, le transcrit *pflmgb1* est préférentiellement exprimé dans les stades asexués du parasite tandis que le transcrit *pflmgb2* présente une expression abondante dans le stade sexué. Ces données intéressantes du transcriptome en plus de l'identité importante entre les deux protéines et de leur conservation parmi les différentes espèces plasmodiales (Fig. 1, p. 675 de l'Article 4) ont suscité un grand intérêt dans notre équipe quant à l'implication de tels facteurs dans le contrôle des événements clés du cycle érythrocytaire de *P. falciparum*.

L'analyse phylogénétique basée sur le travail de S. Soullier, permettant de séparer les facteurs de transcription et les facteurs architecturaux de la superfamille HMGB (Soullier et al. 1999), a permis de regrouper les protéines PfHMGB1 et PfHMGB2 de *P. falciparum* avec les protéines architecturales HMGB de vertébrés, de plantes, de la drosophile, de la levure *S. cerevisiae* et d'une manière très proche avec la protéine NHP1 d'un autre hématozoaire *Babesia bovis* (Données supplémentaires III.1 p. 173). De plus, la présence des acides aminés Ser-10 et Val-32, déterminants caractéristiques du groupe des facteurs architecturaux (Murphy et al. 1999), renforce les résultats obtenus avec la phylogénie et nous permet d'assigner les protéines plasmodiales à la famille HMGB de type architectural. Contrairement aux protéines HMGB des vertébrés qui possèdent deux domaines en tandem A et B, les protéines PfHMGB1 et PfHMGB2 ne possèdent qu'un seul domaine "HMG-box" comme les plantes, la levure et la drosophile. Cet unique domaine "HMG-box" apparaît plus proche des boîtes B sur un arbre phylogénétique généré uniquement avec les boîtes A et B des protéines de vertébrés et les boîtes des protéines parasitaires (Données supplémentaires III.2, p. 175). Il est à noter aussi la présence d'une petite région basique en N-terminal et l'absence de région basique et de queue acide en C-terminal du domaine "HMG-box" des

protéines plasmodiales, comme c'est aussi le cas des protéines HMGB de levure. Enfin, une méthode "tout-automatique" de prédiction de structure (@TOME) a été utilisée pour modéliser les facteurs PfHMGB1 et PfHMGB2 en utilisant comme structure support la boîte B du facteur HMG1 de *Cricetulus griseus*, le hamster chinois, obtenue par spectroscopie RMN (Read et al. 1993). Celle-ci a montré que les domaines "HMG-box" parasitaires comportent trois hélices  $\alpha$  (Données supplémentaires III.3, p. 176), dont l'une est rompu par la présence d'une molécule de  $\beta$ -mercaptoéthanol utilisé lors de l'étude structurale, qui se replie en forme de L conforme à la structure de tous les domaines "HMG-box" conservés au cours de l'évolution (Weir et al. 1993).

Toutes les analyses *in silico* soutiennent l'idée que les protéines PfHMGB1 et PfHMGB2 sont de vrais facteurs architecturaux de la famille HMGB. Pour jouer ce rôle architectural, les facteurs HMGB doivent être capables de se lier à des structures d'ADN d'une manière spécifique et aussi d'induire une courbure de l'ADN facilitant l'assemblage de complexes nucléoprotéiques importants. Par des expériences biologiques de retard sur gel, nous avons pu montrer que les protéines recombinantes PfHMGB1 et PfHMGB2 sont capables (i) d'interagir *in vitro* avec l'ADN cruciforme (Fig. 2, p. 676 de l'Article 4), cette interaction étant plus spécifique pour PfHMGB1, et (ii) d'induire la courbure de l'ADN linéaire, encore une fois, PfHMGB1 étant plus efficace que PfHMGB2 (Fig. 3, p. 677 de l'Article 4). Chez l'homme, il a été montré que la capacité de liaison à l'ADN cruciforme est associée à la boîte A (Webb and Thomas 1999), mais que la boîte B flanquée d'une région basique présente une activité de reconnaissance de l'ADN plus marquée (Yoshioka et al. 1999). Ainsi, chez *P. falciparum*, la petite région basique des facteurs HMGB pourrait diriger l'activité d'interaction avec l'ADN.

Il est raisonnable de penser que chez *P. falciparum*, les facteurs HMGB participent au remodelage de la chromatine par leurs propriétés architecturales. Les protéines HMGB plasmodiales pourraient avoir un rôle essentiel dans le glissement du nucléosome sur l'ADN et dans la modification de l'accessibilité des facteurs de transcription aux éléments de régulation comme chez les autres eucaryotes (Travers 2003). De cette manière, elles seraient capables de réguler la transcription. Des approches moléculaires telles que l'inactivation des gènes *pfhmgb1* et *pfhmgb2* par interférence avec des ARNdb et l'immuno-précipitation de la chromatine comme cela a été effectué pour le FT PfMyb1, devrait nous permettre de

déterminer la fonction des facteurs PfHMGB dans le parasite et en particulier d'identifier leurs gènes cibles.

L'implication vraisemblable des protéines PfHMGB dans la régulation des gènes de *P. falciparum* a d'autant plus d'importance que les différences d'expression de ces facteurs en fonction du stade de développement suggèrent une intervention spécifique de PfHMGB1 dans la prolifération asexuée et de PfHMGB2 dans la différenciation sexuée (Fig. 4c, p. 677 de l'Article 4). En effet, la cinétique d'expression des protéines PfHMGB1 et PfHMGB2 au cours du cycle érythrocytaire du clone 3D7 est similaire au profil d'expression des transcrits obtenu par le laboratoire de E. Winzeler (Le Roch et al. 2003). La localisation de PfHMGB1 et PfHMGB2 dans le noyau des différents stades de développement confirme d'une part leur fonction nucléaire (Fig. 5, p. 679 de l'Article 4). D'autre part, la présence de PfHMGB2 dans le cytoplasme des gamétocytes renforce l'hypothèse du rôle non-redondant de ces protéines dans la cellule tout comme les différences observées vis-à-vis de leur interaction avec l'ADN. L'inactivation spécifique de l'un ou l'autre des deux gènes *pfhmgb1* et *pfhmgb2* pourrait nous renseigner sur la redondance de ces facteurs dans le parasite. Il est intéressant de noter que les protéines humaines Hmgb1 et Hmgb2 ont aussi une expression différente suivant le contexte cellulaire et le stade de développement et ne sont pas interchangeables en dépit de leur extrême similarité (+ 80 % d'identité) (Ronfani et al. 2001). En effet, Hmgb1 est une protéine ubiquitaire et les souris déficientes *hmgb1*<sup>-/-</sup> ne sont pas viables (Calogero et al. 1999) tandis que HMGB2 est majoritairement exprimée dans les organes lymphoïdes et les testicules et les souris *hmgb2*<sup>-/-</sup> sont viables mais présentent une sévère baisse de la fertilité (Ronfani et al. 2001).

Chez les eucaryotes, plusieurs études ont rapporté l'interaction des facteurs HMGB avec diverses protéines partenaires impliquées dans la transcription comme la TBP (Ge and Roeder 1994), la protéine SET d'assemblage du nucléosome (NAP) (Fan et al. 2002), la protéine HOX (Zappavigna et al. 1996), les protéines de la famille Rel comme NF- $\kappa$ B et sa sous-unité p50 (Brickman et al. 1999) et la protéine p53 (Jayaraman et al. 1998). Parmi ces protéines qui n'ont pas encore toutes d'homologues chez *P. falciparum*, une protéine à domaine NAP, connue pour interagir avec la machinerie de remodelage de la chromatine (Mello and Almouzni 2001), Pfb7 a été identifiée par homologie avec Pbb7 de *P. berghei* (Birago et al. 1996). Le gène *pfb7* est de surcroît l'un de ceux qui présentait un profil

d'expression intéressant dans notre étude comparative du transcriptome des clones 3D7 et F12, stable au cours du développement du clone 3D7 et fortement augmenté en fin de cycle du clone F12. Il est donc important d'étudier l'interaction entre Pfb7 et les facteurs PfHMGB1 et PfHMGB2 pour analyser son implication dans l'expression des gènes de ces deux clones de *P. falciparum*, producteur ou non de gamétocytes. Dans cette optique, le gène *pfb7* a été cloné, la protéine Pfb7 et des anticorps ont été produits. L'identification de partenaires protéiques, éventuellement encore non annotés, des facteurs PfHMGB a été initiée par des expériences d'immuno-précipitation couplée à la spectrométrie de masse (en collaboration avec O. Azlor et C. Marinach-Patrice).

Même si la fonction nucléaire des protéines HMGB est essentielle chez les eucaryotes, la littérature porte un intérêt grandissant à la double vie de ces protéines quant à leur rôle pro-inflammatoire (Muller et al. 2001). En effet, les protéines HMGB libérées activement ou passivement dans le milieu extracellulaire sont de puissants activateurs de monocytes et déclenchent la production de cytokines comme le TNF $\alpha$  (Wang et al. 1999; Yang et al. 2005) en faisant intervenir un autre domaine fonctionnel de la protéine appelé "domaine de stimulation du TNF $\alpha$ " (Li et al. 2003). D'une manière intéressante, un domaine présentant 75 % d'identité avec le "domaine de stimulation du TNF $\alpha$ " identifiée dans la boîte B de la protéine humaine Hmgb1, est présent dans l'unique domaine des protéines plasmodiales. Ceci suggère que les facteurs PfHMGB pourraient être aussi une cytokine cruciale impliquée dans la réponse inflammatoire à l'infection par *P. falciparum*.

La physiopathologie du neuropaludisme est encore mal comprise mais elle semble résulter de la séquestration des globules rouges infectés et des réponses pro-inflammatoires (Engwerda et al. 2005). Des niveaux élevés de cytokines, en particulier de TNF $\alpha$ , ont été corrélés à la sévérité de la maladie (Grau et al. 1989; Kwiatkowski et al. 1990). Plus récemment, une étude portant sur 16 enfants, dont 10 sont morts par neuropaludisme, a rapportée que le taux sérique de la protéine HMGB humaine pouvait être associé à la gravité de l'infection à *P. falciparum* (Alleva et al. 2005). Des expériences préliminaires montrent que les deux protéines PfHMGB1 et PfHMGB2 sont libérées par le parasite et seraient capables d'activer des monocytes primaires humains induisant la production de cytokines (en collaboration avec V. Maréchal). Ces premiers résultats sur l'activité cytokinique potentielle des facteurs PfHMGB ont d'autant plus d'importance que ces protéines pourraient jouer un



rôle clé dans le développement de la pathologie cérébrale, le neuropaludisme. L'utilisation d'un modèle murin de neuropaludisme pourrait nous donner un élément de réponse quant à la fonction de ces facteurs *in vivo*. Il serait intéressant de déterminer si la neutralisation de la protéine parasitaire PbHMGB après immunisation ou injection d'anticorps pourrait entraver le développement de la pathologie cérébrale et influencer ainsi sur la survie de souris infectée par *P. berghei*.

**En conclusion**, des études couplées de bioinformatique et de biologie moléculaire ont permis de montrer l'existence d'authentiques facteurs Myb et HMGB impliqués dans la régulation transcriptionnelle des événements clés du cycle érythrocytaire de *P. falciparum*. L'importance des facteurs de transcription dans la régulation du cycle cellulaire, la multiplication et la différenciation leur a donné un potentiel thérapeutique certain et ouvre de nouvelles perspectives dans la lutte contre le paludisme.

## REFERENCES BIBLIOGRAPHIQUES

- Abrahamsen, M.S., T.J. Templeton, S. Enomoto, J.E. Abrahante, G. Zhu, C.A. Lancto, M. Deng, C. Liu, G. Widmer, S. Tzipori, G.A. Buck, P. Xu, A.T. Bankier, P.H. Dear, B.A. Konfortov, H.F. Spriggs, L. Iyer, V. Anantharaman, L. Aravind, and V. Kapur. 2004. Complete genome sequence of the apicomplexan, *Cryptosporidium parvum*. *Science* **304**: 441-445.
- Agresti, A. and M.E. Bianchi. 2003. HMGB proteins and gene expression. *Curr Opin Genet Dev* **13**: 170-178.
- Alano, P., L. Roca, D. Smith, D. Read, R. Carter, and K. Day. 1995. *Plasmodium falciparum*: parasites defective in early stages of gametocytogenesis. *Exp Parasitol* **81**: 227-235.
- Alano, P., F. Silvestrini, and L. Roca. 1996. Structure and polymorphism of the upstream region of the pfg27/25 gene, transcriptionally regulated in gametocytogenesis of *Plasmodium falciparum*. *Mol Biochem Parasitol* **79**: 207-217.
- Alleva, L.M., H. Yang, K.J. Tracey, and I.A. Clark. 2005. High mobility group box 1 (HMGB1) protein: possible amplification signal in the pathogenesis of falciparum malaria. *Trans R Soc Trop Med Hyg* **99**: 171-174.
- Ambros, V. 2001. microRNAs: tiny regulators with great potential. *Cell* **107**: 823-826.
- Andersson, U., H. Wang, K. Palmblad, A.C. Aveberger, O. Bloom, H. Erlandsson-Harris, A. Janson, R. Kokkola, M. Zhang, H. Yang, and K.J. Tracey. 2000. High mobility group 1 protein (HMG-1) stimulates proinflammatory cytokine synthesis in human monocytes. *J Exp Med* **192**: 565-570.
- Aravind, L., L.M. Iyer, T.E. Wellems, and L.H. Miller. 2003. Plasmodium biology: genomic gleanings. *Cell* **115**: 771-785.
- Ausio, J. 2000. Are linker histones (histone H1) dispensable for survival? *Bioessays* **22**: 873-877.
- Baldari, M., A. Tamburro, G. Sabatinelli, R. Romi, C. Severini, G. Cuccagna, G. Fiorilli, M.P. Allegri, C. Buriani, and M. Toti. 1998. Malaria in Maremma, Italy. *Lancet* **351**: 1246-1247.
- Bannister, L. and G. Mitchell. 2003. The ins, outs and roundabouts of malaria. *Trends Parasitol* **19**: 209-213.
- Bannister, L.H., J.M. Hopkins, R.E. Fowler, S. Krishna, and G.H. Mitchell. 2000. A brief illustrated guide to the ultrastructure of *Plasmodium falciparum* asexual blood stages. *Parasitol Today* **16**: 427-433.
- Barinaga, M. 1991. Dimers direct development. *Science* **251**: 1176-1177.
- Barron, C. 1992. Transcription of the gene for the merozoite surface antigen MSA2 of the human malaria parasite *Plasmodium falciparum* during the asexual cycle. *FEBS Lett* **300**: 77-81.
- Baussand, J., C. Deremble, and C. A. 2006. Periodic distribution of hydrophobic amino acids allows to define fundamental building blocks to align distantly related proteins. *to be submitted*.
- Baxevanis, A.D. and D. Landsman. 1995. The HMG-1 box protein family: classification and functional relationships. *Nucleic Acids Res* **23**: 1604-1613.
- Becker, P.B. and W. Horz. 2002. ATP-dependent nucleosome remodeling. *Annu Rev Biochem* **71**: 247-273.

- Ben Mamoun, C., I.Y. Gluzman, C. Hott, S.K. MacMillan, A.S. Amarakone, D.L. Anderson, J.M. Carlton, J.B. Dame, D. Chakrabarti, R.K. Martin, B.H. Brownstein, and D.E. Goldberg. 2001. Co-ordinated programme of gene expression during asexual intraerythrocytic development of the human malaria parasite *Plasmodium falciparum* revealed by microarray analysis. *Mol Microbiol* **39**: 26-36.
- Bentley, D. 2002. The mRNA assembly line: transcription and processing machines in the same factory. *Curr Opin Cell Biol* **14**: 336-342.
- Berg, J.M. 1986. Potential metal-binding domains in nucleic acid binding proteins. *Science* **232**: 485-487.
- Berg, J.M. 1990. Zinc finger domains: hypotheses and current knowledge. *Annu Rev Biophys Chem* **19**: 405-421.
- Bianchi, M.E. 2004. Significant (re)location: how to use chromatin and/or abundant proteins as messages of life and death. *Trends Cell Biol* **14**: 287-293.
- Bianchi, M.E. and A. Agresti. 2005. HMG proteins: dynamic players in gene regulation and differentiation. *Curr Opin Genet Dev* **15**: 496-506.
- Bianchi, M.E., M. Beltrame, and G. Paonessa. 1989. Specific recognition of cruciform DNA by nuclear protein HMG1. *Science* **243**: 1056-1059.
- Biedenkapp, H., U. Borgmeyer, A.E. Sippel, and K.H. Klempnauer. 1988. Viral myb oncogene encodes a sequence-specific DNA-binding activity. *Nature* **335**: 835-837.
- Billings, P.C., R.J. Davis, B.N. Engelsberg, K.A. Skov, and E.N. Hughes. 1992. Characterization of high mobility group protein binding to cisplatin-damaged DNA. *Biochem Biophys Res Commun* **188**: 1286-1294.
- Billker, O., V. Lindo, M. Panico, A.E. Etienne, T. Paxton, A. Dell, M. Rogers, R.E. Sinden, and H.R. Morris. 1998. Identification of xanthurenic acid as the putative inducer of malaria development in the mosquito. *Nature* **392**: 289-292.
- Birago, C., T. Pace, S. Barca, L. Picci, and M. Ponzi. 1996. A chromatin-associated protein is encoded in a genomic region highly conserved in the *Plasmodium* genus. *Mol Biochem Parasitol* **80**: 193-202.
- Blander, G. and L. Guarente. 2004. The Sir2 family of protein deacetylases. *Annu Rev Biochem* **73**: 417-435.
- Bonaldi, T., F. Talamo, P. Scaffidi, D. Ferrera, A. Porto, A. Bachi, A. Rubartelli, A. Agresti, and M.E. Bianchi. 2003. Monocytic cells hyperacetylate chromatin protein HMGB1 to redirect it towards secretion. *Embo J* **22**: 5551-5560.
- Bourbon, H.M., A. Aguilera, A.Z. Ansari, F.J. Asturias, A.J. Berk, S. Bjorklund, T.K. Blackwell, T. Borggrefe, M. Carey, M. Carlson, J.W. Conaway, R.C. Conaway, S.W. Emmons, J.D. Fondell, L.P. Freedman, T. Fukasawa, C.M. Gustafsson, M. Han, X. He, P.K. Herman, A.G. Hinnebusch, S. Holmberg, F.C. Holstege, J.A. Jaehning, Y.J. Kim, L. Kuras, A. Leutz, J.T. Lis, M. Meisterernest, A.M. Naar, K. Nasmyth, J.D. Parvin, M. Ptashne, D. Reinberg, H. Ronne, I. Sadowski, H. Sakurai, M. Sipiczki, P.W. Sternberg, D.J. Stillman, R. Strich, K. Struhl, J.Q. Svejstrup, S. Tuck, F. Winston, R.G. Roeder, and R.D. Kornberg. 2004. A unified nomenclature for protein subunits of mediator complexes linking transcriptional regulators to RNA polymerase II. *Mol Cell* **14**: 553-557.
- Bowman, S., D. Lawson, D. Basham, D. Brown, T. Chillingworth, C.M. Churcher, A. Craig, R.M. Davies, K. Devlin, T. Feltwell, S. Gentles, R. Gwilliam, N. Hamlin, D. Harris, S. Holroyd, T. Hornsby, P. Horrocks, K. Jagels, B. Jassal, S. Kyes, J. McLean, S. Moule, K. Mungall, L. Murphy, K. Oliver, M.A. Quail, M.A. Rajandream, S. Rutter, J. Skelton, R. Squares, S. Squares, J.E. Sulston, S. Whitehead, J.R. Woodward, C. Newbold, and B.G.

- Barrell. 1999. The complete nucleotide sequence of chromosome 3 of *Plasmodium falciparum*. *Nature* **400**: 532-538.
- Bozdech, Z., M. Llinas, B.L. Pulliam, E.D. Wong, J. Zhu, and J.L. DeRisi. 2003a. The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. *PLoS Biol* **1**: E5.
- Bozdech, Z., J. Zhu, M.P. Joachimiak, F.E. Cohen, B. Pulliam, and J.L. DeRisi. 2003b. Expression profiling of the schizont and trophozoite stages of *Plasmodium falciparum* with a long-oligonucleotide microarray. *Genome Biol* **4**: R9.
- Brickman, J.M., M. Adam, and M. Ptashne. 1999. Interactions between an HMG-1 protein and members of the Rel family. *Proc Natl Acad Sci U S A* **96**: 10679-10683.
- Bruce, M.C., P. Alano, S. Duthie, and R. Carter. 1990. Commitment of the malaria parasite *Plasmodium falciparum* to sexual and asexual development. *Parasitology* **100 Pt 2**: 191-200.
- Bruce, M.C., R.N. Carter, K. Nakamura, M. Aikawa, and R. Carter. 1994. Cellular location and temporal expression of the *Plasmodium falciparum* sexual stage antigen Pfs16. *Mol Biochem Parasitol* **65**: 11-22.
- Buratowski, S. 1994. The basics of basal transcription by RNA polymerase II. *Cell* **77**: 1-3.
- Buratowski, S., S. Hahn, L. Guarente, and P.A. Sharp. 1989. Five intermediate complexes in transcription initiation by RNA polymerase II. *Cell* **56**: 549-561.
- Bustin, M. 1999. Regulation of DNA-dependent activities by the functional motifs of the high-mobility-group chromosomal proteins. *Mol Cell Biol* **19**: 5237-5246.
- Bustin, M. 2001. Revised nomenclature for high mobility group (HMG) chromosomal proteins. *Trends Biochem Sci* **26**: 152-153.
- Bustin, M. and R. Reeves. 1996. High-mobility-group chromosomal proteins: architectural components that facilitate chromatin function. *Prog Nucleic Acid Res Mol Biol* **54**: 35-100.
- Calderwood, M.S., L. Gannoun-Zaki, T.E. Wellems, and K.W. Deitsch. 2003. *Plasmodium falciparum* var genes are regulated by two regions with separate promoters, one upstream of the coding region and a second within the intron. *J Biol Chem* **278**: 34125-34132.
- Callebaut, I., K. Prat, E. Meurice, J.P. Mornon, and S. Tomavo. 2005. Prediction of the general transcription factors associated with RNA polymerase II in *Plasmodium falciparum*: conserved features and differences relative to other eukaryotes. *BMC Genomics* **6**: 100.
- Calogero, S., F. Grassi, A. Aguzzi, T. Voigtlander, P. Ferrier, S. Ferrari, and M.E. Bianchi. 1999. The lack of chromosomal protein Hmg1 does not disrupt cell growth but causes lethal hypoglycaemia in newborn mice. *Nat Genet* **22**: 276-280.
- Carlson, M. 1997. Genetics of transcriptional regulation in yeast: connections to the RNA polymerase II CTD. *Annu Rev Cell Dev Biol* **13**: 1-23.
- Chalkley, G.E. and C.P. Verrijzer. 1999. DNA binding site selection by RNA polymerase II TAFs: a TAF(II)250-TAF(II)150 complex recognizes the initiator. *Embo J* **18**: 4835-4845.
- Chitnis, C.E. and M.J. Blackman. 2000. Host cell invasion by malaria parasites. *Parasitol Today* **16**: 411-415.
- Chow, C.S. and D.F. Wirth. 2003. Linker scanning mutagenesis of the *Plasmodium gallinaceum* sexual stage specific gene pgs28 reveals a novel downstream cis-control element. *Mol Biochem Parasitol* **129**: 199-208.
- Coatney, G., W. Collins, M. Warren, and P. Contacos. 1971. The Primate Malariae. *Bethesda: U.S. Department of Health, Education and Welfare*.

- Coin, F. and J.M. Egly. 1998. Ten years of TFIIH. *Cold Spring Harb Symp Quant Biol* **63**: 105-110.
- Conaway, R.C. and J.W. Conaway. 1993. General initiation factors for RNA polymerase II. *Annu Rev Biochem* **62**: 161-190.
- Cooper, G.M. 2000. The cell : a molecular approach. *Amer. Soc. Microbiol., Washington and Sinauer Assoc., Sunderland, MA* **2d ed.**
- Corden, J.L. 1990. Tails of RNA polymerase II. *Trends Biochem Sci* **15**: 383-387.
- Cosma, M.P. 2002. Ordered recruitment: gene-specific mechanism of transcription activation. *Mol Cell* **10**: 227-236.
- Coulson, R.M., N. Hall, and C.A. Ouzounis. 2004. Comparative genomics of transcriptional control in the human malaria parasite *Plasmodium falciparum*. *Genome Res* **14**: 1548-1554.
- Coulson, R.M. and C.A. Ouzounis. 2003. The phylogenetic diversity of eukaryotic transcription. *Nucleic Acids Res* **31**: 653-660.
- Crabb, B.S. and A.F. Cowman. 1996. Characterization of promoters and stable transfection by homologous and nonhomologous recombination in *Plasmodium falciparum*. *Proc Natl Acad Sci U S A* **93**: 7289-7294.
- Cui, L., Q. Fan, and J. Li. 2002. The malaria parasite *Plasmodium falciparum* encodes members of the Puf RNA-binding protein family with conserved RNA binding activity. *Nucleic Acids Res* **30**: 4607-4617.
- Dai, Y., B. Wong, Y.M. Yen, M.A. Oettinger, J. Kwon, and R.C. Johnson. 2005. Determinants of HMGB proteins required to promote RAG1/2-recombination signal sequence complex assembly and catalysis during V(D)J recombination. *Mol Cell Biol* **25**: 4413-4425.
- Davison, B.L., J.M. Egly, E.R. Mulvihill, and P. Chambon. 1983. Formation of stable preinitiation complexes between eukaryotic class B transcription factors and promoter sequences. *Nature* **301**: 680-686.
- Dechering, K.J., K. Cuelenaere, R.N. Konings, and J.A. Leunissen. 1998. Distinct frequency-distributions of homopolymeric DNA tracts in different genomes. *Nucleic Acids Res* **26**: 4056-4062.
- Dechering, K.J., A.M. Kaan, W. Mbacham, D.F. Wirth, W. Eling, R.N. Konings, and H.G. Stunnenberg. 1999. Isolation and functional characterization of two distinct sexual-stage-specific promoters of the human malaria parasite *Plasmodium falciparum*. *Mol Cell Biol* **19**: 967-978.
- Deitsch, K.W., M.S. Calderwood, and T.E. Wellems. 2001. Malaria. Cooperative silencing elements in var genes. *Nature* **412**: 875-876.
- Desowitz, R.S. 1991. The malaria capers. In *More Tales of Parasites and People, Research and Reality* (ed. W.W.N. Company), New York.
- Dini, P.W. and J.S. Lipsick. 1993. Oncogenic truncation of the first repeat of c-Myb decreases DNA binding in vitro and in vivo. *Mol Cell Biol* **13**: 7334-7348.
- Dobson, M.J. 1999. The malariology centenary. *Parassitologia* **41**: 21-32.
- Douguet, D. and G. Labesse. 2001. Easier threading through web-based comparisons and cross-validations. *Bioinformatics* **17**: 752-753.
- Dubendorff, J.W., L.J. Whittaker, J.T. Eltman, and J.S. Lipsick. 1992. Carboxy-terminal elements of c-Myb negatively regulate transcriptional activation in cis and in trans. *Genes Dev* **6**: 2524-2535.
- Dumitriu, I.E., P. Baruah, A.A. Manfredi, M.E. Bianchi, and P. Rovere-Querini. 2005. HMGB1: guiding immunity from within. *Trends Immunol* **26**: 381-387.

- Duraisingh, M.T., T.S. Voss, A.J. Marty, M.F. Duffy, R.T. Good, J.K. Thompson, L.H. Freitas-Junior, A. Scherf, B.S. Crabb, and A.F. Cowman. 2005. Heterochromatin silencing and locus repositioning linked to regulation of virulence genes in *Plasmodium falciparum*. *Cell* **121**: 13-24.
- Dyer, M. and K.P. Day. 2000. Commitment to gametocytogenesis in *Plasmodium falciparum*. *Parasitol Today* **16**: 102-107.
- Eksi, S., Y. Haile, T. Furuya, L. Ma, X. Su, and K.C. Williamson. 2005. Identification of a subtelomeric gene family expressed during the asexual-sexual stage transition in *Plasmodium falciparum*. *Mol Biochem Parasitol* **143**: 90-99.
- Emerson, B.M. 2002. Specificity of gene regulation. *Cell* **109**: 267-270.
- Engwerda, C., E. Belnoue, A.C. Gruner, and L. Renia. 2005. Experimental models of cerebral malaria. *Curr Top Microbiol Immunol* **297**: 103-143.
- Erondu, N.E. and J.E. Donelson. 1992. Differential expression of two mRNAs from a single gene encoding an HMG1-like DNA binding protein of African trypanosomes. *Mol Biochem Parasitol* **51**: 111-118.
- Escalante, A.A. and F.J. Ayala. 1994. Phylogeny of the malarial genus *Plasmodium*, derived from rRNA gene sequences. *Proc Natl Acad Sci U S A* **91**: 11373-11377.
- Fan, Q., L. An, and L. Cui. 2004a. PfADA2, a *Plasmodium falciparum* homologue of the transcriptional coactivator ADA2 and its in vivo association with the histone acetyltransferase PfGCN5. *Gene* **336**: 251-261.
- Fan, Q., L. An, and L. Cui. 2004b. *Plasmodium falciparum* histone acetyltransferase, a yeast GCN5 homologue involved in chromatin remodeling. *Eukaryot Cell* **3**: 264-276.
- Fan, Z., P.J. Beresford, D. Zhang, and J. Lieberman. 2002. HMG2 interacts with the nucleosome assembly protein SET and is a target of the cytotoxic T-lymphocyte protease granzyme A. *Mol Cell Biol* **22**: 2810-2820.
- Fantini, B. 1999. The concept of specificity and the Italian contribution to the discovery of the malaria transmission cycle. *Parassitologia* **41**: 39-47.
- Felsenstein, J. 1989. PHYLIP, Phylogeny Inference Package (version 3.63). *Cladistics* **5**: 164-166.
- Figueiredo, L.M., L.H. Freitas-Junior, E. Bottius, J.C. Olivo-Marin, and A. Scherf. 2002. A central role for *Plasmodium falciparum* subtelomeric regions in spatial positioning and telomere length regulation. *Embo J* **21**: 815-824.
- Fire, A., M. Samuels, and P.A. Sharp. 1984. Interactions between RNA polymerase II, factors, and template leading to accurate transcription. *J Biol Chem* **259**: 2509-2516.
- Fischer, L., M. Sterneck, M. Claus, A. Costard-Jackle, B. Fleischer, H. Herbst, X. Rogiers, and C.E. Broelsch. 1999. Transmission of malaria tertiana by multi-organ donation. *Clin Transplant* **13**: 491-495.
- Florens, L., M.P. Washburn, J.D. Raine, R.M. Anthony, M. Grainger, J.D. Haynes, J.K. Moch, N. Muster, J.B. Sacci, D.L. Tabb, A.A. Witney, D. Wolters, Y. Wu, M.J. Gardner, A.A. Holder, R.E. Sinden, J.R. Yates, and D.J. Carucci. 2002. A proteomic view of the *Plasmodium falciparum* life cycle. *Nature* **419**: 520-526.
- Flores, O., H. Lu, M. Killeen, J. Greenblatt, Z.F. Burton, and D. Reinberg. 1991. The small subunit of transcription factor IIF recruits RNA polymerase II into the preinitiation complex. *Proc Natl Acad Sci U S A* **88**: 9999-10003.
- Fox, B.A. and D.J. Bzik. 1994. Analysis of stage-specific transcripts of the *Plasmodium falciparum* serine repeat antigen (SERA) gene and transcription from the SERA locus. *Mol Biochem Parasitol* **68**: 133-144.

- Freitas-Junior, L.H., E. Bottius, L.A. Pirrit, K.W. Deitsch, C. Scheidig, F. Guinet, U. Nehrass, T.E. Wellems, and A. Scherf. 2000. Frequent ectopic recombination of virulence factor genes in telomeric chromosome clusters of *P. falciparum*. *Nature* **407**: 1018-1022.
- Freitas-Junior, L.H., R. Hernandez-Rivas, S.A. Ralph, D. Montiel-Condado, O.K. Ruvalcaba-Salazar, A.P. Rojas-Meza, L. Mancio-Silva, R.J. Leal-Silvestre, A.M. Gontijo, S. Shorte, and A. Scherf. 2005. Telomeric heterochromatin propagation and histone acetylation control mutually exclusive expression of antigenic variation genes in malaria parasites. *Cell* **121**: 25-36.
- Frevert, U. 2004. Sneaking in through the back entrance: the biology of malaria liver stages. *Trends Parasitol* **20**: 417-424.
- Friedman, A.D. 1996. Regulation of immature myeloid cell differentiation by PEBP2/CBF, Myb, C/EBP and Ets family members. *Curr Top Microbiol Immunol* **211**: 149-157.
- Gaillard, C. and F. Strauss. 2000. High affinity binding of proteins HMG1 and HMG2 to semicatenated DNA loops. *BMC Mol Biol* **1**: 1.
- Gajiwala, K.S. and S.K. Burley. 2000. Winged helix proteins. *Curr Opin Struct Biol* **10**: 110-116.
- Galio, L., S. Briquet, S. Cot, J.G. Guillet, and C. Vaquero. 1997. Analysis of interactions between huGATA-3 transcription factor and three GATA regulatory elements of HIV-1 long terminal repeat, by surface plasmon resonance. *Anal Biochem* **253**: 70-77.
- Ganter, B., S.T. Chao, and J.S. Lipsick. 1999. Transcriptional activation by the myb proteins requires a specific local promoter structure. *FEBS Lett* **460**: 401-410.
- Garcia, G.E., R.A. Wirtz, J.R. Barr, A. Woolfitt, and R. Rosenberg. 1998. Xanthurenic acid induces gametogenesis in *Plasmodium*, the malaria parasite. *J Biol Chem* **273**: 12003-12005.
- Gardiner, D.L., D.C. Holt, E.A. Thomas, D.J. Kemp, and K.R. Trenholme. 2000. Inhibition of *Plasmodium falciparum* clag9 gene function by antisense RNA. *Mol Biochem Parasitol* **110**: 33-41.
- Gardiner-Garden, M. and M. Frommer. 1987. CpG islands in vertebrate genomes. *J Mol Biol* **196**: 261-282.
- Gardner, M.J., N. Hall, E. Fung, O. White, M. Berriman, R.W. Hyman, J.M. Carlton, A. Pain, K.E. Nelson, S. Bowman, I.T. Paulsen, K. James, J.A. Eisen, K. Rutherford, S.L. Salzberg, A. Craig, S. Kyes, M.S. Chan, V. Nene, S.J. Shallom, B. Suh, J. Peterson, S. Angiuoli, M. Perte, J. Allen, J. Selengut, D. Haft, M.W. Mather, A.B. Vaidya, D.M. Martin, A.H. Fairlamb, M.J. Fraunholz, D.S. Roos, S.A. Ralph, G.I. McFadden, L.M. Cummings, G.M. Subramanian, C. Mungall, J.C. Venter, D.J. Carucci, S.L. Hoffman, C. Newbold, R.W. Davis, C.M. Fraser, and B. Barrell. 2002. Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* **419**: 498-511.
- Gardner, M.J., H. Tettelin, D.J. Carucci, L.M. Cummings, L. Aravind, E.V. Koonin, S. Shallom, T. Mason, K. Yu, C. Fujii, J. Pederson, K. Shen, J. Jing, C. Aston, Z. Lai, D.C. Schwartz, M. Perte, S. Salzberg, L. Zhou, G.G. Sutton, R. Clayton, O. White, H.O. Smith, C.M. Fraser, M.D. Adams, J.C. Venter, and S.L. Hoffman. 1998. Chromosome 2 sequence of the human malaria parasite *Plasmodium falciparum*. *Science* **282**: 1126-1132.
- Ge, H. and R.G. Roeder. 1994. The high mobility group protein HMG1 can reversibly inhibit class II gene transcription by interaction with the TATA-binding protein. *J Biol Chem* **269**: 17136-17140.
- Gehring, W.J., M. Muller, M. Affolter, A. Percival-Smith, M. Billeter, Y.Q. Qian, G. Otting, and K. Wuthrich. 1990. The structure of the homeodomain and its functional implications. *Trends Genet* **6**: 323-329.

- Geiger, J.H., S. Hahn, S. Lee, and P.B. Sigler. 1996. Crystal structure of the yeast TFIIA/TBP/DNA complex. *Science* **272**: 830-836.
- Giglioli, G. 1951. Nation-wide malaria eradication projects in the Americas. III. Eradication of *Anopheles darlingi* from the inhabited areas of British Guiana by DDT residual spraying. *J Natl Malar Soc* **10**: 142-161.
- Glockner, G., L. Eichinger, K. Szafranski, J.A. Pachebat, A.T. Bankier, P.H. Dear, R. Lehmann, C. Baumgart, G. Parra, J.F. Abril, R. Guigo, K. Kumpf, B. Tunggal, E. Cox, M.A. Quail, M. Platzer, A. Rosenthal, and A.A. Noegel. 2002. Sequence and analysis of chromosome 2 of *Dictyostelium discoideum*. *Nature* **418**: 79-85.
- Gnanasekar, M., R. Velusamy, Y.X. He, and K. Ramaswamy. 2006. Cloning and characterization of a high mobility group box 1 (HMGB1) homologue protein from *Schistosoma mansoni*. *Mol Biochem Parasitol* **145**: 137-146.
- Gonda, T.J., D.K. Sheiness, and J.M. Bishop. 1982. Transcripts from the cellular homologs of retroviral oncogenes: distribution among chicken tissues. *Mol Cell Biol* **2**: 617-624.
- Graeser, R., B. Wernli, R.M. Franklin, and B. Kappes. 1996. Plasmodium falciparum protein kinase 5 and the malarial nuclear division cycles. *Mol Biochem Parasitol* **82**: 37-49.
- Grau, G.E., T.E. Taylor, M.E. Molyneux, J.J. Wirima, P. Vassalli, M. Hommel, and P.H. Lambert. 1989. Tumor necrosis factor and disease severity in children with falciparum malaria. *N Engl J Med* **320**: 1586-1591.
- Grosschedl, R., K. Giese, and J. Pagel. 1994. HMG domain proteins: architectural elements in the assembly of nucleoprotein structures. *Trends Genet* **10**: 94-100.
- Grunstein, M. 1997. Histone acetylation in chromatin structure and transcription. *Nature* **389**: 349-352.
- Gunasekera, A.M., S. Patankar, J. Schug, G. Eisen, J. Kissinger, D. Roos, and D.F. Wirth. 2004. Widespread distribution of antisense transcripts in the Plasmodium falciparum genome. *Mol Biochem Parasitol* **136**: 35-42.
- Hamada, H. and M. Bustin. 1985. Hierarchy of binding sites for chromosomal proteins HMG 1 and 2 in supercoiled deoxyribonucleic acid. *Biochemistry* **24**: 1428-1433.
- Harrison, S.C. 1991. A structural taxonomy of DNA-binding domains. *Nature* **353**: 715-719.
- Harrison, S.C. and A.K. Aggarwal. 1990. DNA recognition by proteins with the helix-turn-helix motif. *Annu Rev Biochem* **59**: 933-969.
- Harvey, A.C. and J.A. Downs. 2004. What functions do linker histones provide? *Mol Microbiol* **53**: 771-775.
- Hayward, R.E., J.L. Derisi, S. Alfadhli, D.C. Kaslow, P.O. Brown, and P.K. Rathod. 2000. Shotgun DNA microarrays and stage-specific gene expression in Plasmodium falciparum malaria. *Mol Microbiol* **35**: 6-14.
- He, H. and N. Lehming. 2003. Global effects of histone modifications. *Brief Funct Genomic Proteomic* **2**: 234-243.
- Hebbes, T.R., A.L. Clayton, A.W. Thorne, and C. Crane-Robinson. 1994. Core histone hyperacetylation co-maps with generalized DNase I sensitivity in the chicken beta-globin chromosomal domain. *Embo J* **13**: 1823-1830.
- Hebbes, T.R., A.W. Thorne, and C. Crane-Robinson. 1988. A direct link between core histone acetylation and transcriptionally active chromatin. *Embo J* **7**: 1395-1402.
- Hirtzlin, J., P.M. Farber, and R.M. Franklin. 1994. Isolation of a novel Plasmodium falciparum gene encoding a protein homologous to the Tat-binding protein family. *Eur J Biochem* **226**: 673-680.



- Hoffman, S.L., W.H. Bancroft, M. Gottlieb, S.L. James, E.C. Burroughs, J.R. Stephenson, and M.J. Morgan. 1997. Funding for malaria genome sequencing. *Nature* **387**: 647.
- Holt, R.A. G.M. Subramanian A. Halpern G.G. Sutton R. Charlab D.R. Nusskern P. Wincker A.G. Clark J.M. Ribeiro R. Wides S.L. Salzberg B. Loftus M. Yandell W.H. Majoros D.B. Rusch Z. Lai C.L. Kraft J.F. Abril V. Anthouard P. Arensburger P.W. Atkinson H. Baden V. de Berardinis D. Baldwin V. Benes J. Biedler C. Blass R. Bolanos D. Boscus M. Barnstead S. Cai A. Center K. Chaturverdi G.K. Christophides M.A. Chrystal M. Clamp A. Cravchik V. Curwen A. Dana A. Delcher I. Dew C.A. Evans M. Flanigan A. Grundschober-Freimoser L. Friedli Z. Gu P. Guan R. Guigo M.E. Hillenmeyer S.L. Hladun J.R. Hogan Y.S. Hong J. Hoover O. Jaillon Z. Ke C. Kodira E. Kokoza A. Koutsos I. Letunic A. Levitsky Y. Liang J.J. Lin N.F. Lobo J.R. Lopez J.A. Malek T.C. McIntosh S. Meister J. Miller C. Mobarry E. Mongin S.D. Murphy D.A. O'Brochta C. Pfannkoch R. Qi M.A. Regier K. Remington H. Shao M.V. Sharakhova C.D. Sitter J. Shetty T.J. Smith R. Strong J. Sun D. Thomasova L.Q. Ton P. Topalis Z. Tu M.F. Unger B. Walenz A. Wang J. Wang M. Wang X. Wang K.J. Woodford J.R. Wortman M. Wu A. Yao E.M. Zdobnov H. Zhang Q. Zhao S. Zhao S.C. Zhu I. Zhimulev M. Coluzzi A. della Torre C.W. Roth C. Louis F. Kalush R.J. Mural E.W. Myers M.D. Adams H.O. Smith S. Broder M.J. Gardner C.M. Fraser E. Birney P. Bork P.T. Brey J.C. Venter J. Weissenbach F.C. Kafatos F.H. Collins and S.L. Hoffman. 2002. The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science* **298**: 129-149.
- Horrocks, P., K. Dechering, and M. Lanzer. 1998. Control of gene expression in *Plasmodium falciparum*. *Mol Biochem Parasitol* **95**: 171-181.
- Horrocks, P. and B.J. Kilbey. 1996. Physical and functional mapping of the transcriptional start sites of *Plasmodium falciparum* proliferating cell nuclear antigen. *Mol Biochem Parasitol* **82**: 207-215.
- Horrocks, P. and M. Lanzer. 1999a. Differences in nucleosome organization over episomally located plasmids coincides with aberrant promoter activity in *P. falciparum*. *Parasitol Int* **48**: 55-61.
- Horrocks, P. and M. Lanzer. 1999b. Mutational analysis identifies a five base pair cis-acting sequence essential for GBP130 promoter activity in *Plasmodium falciparum*. *Mol Biochem Parasitol* **99**: 77-87.
- Hovring, I., A. Bostad, E. Ording, A.H. Myrset, and O.S. Gabrielsen. 1994. DNA-binding domain and recognition sequence of the yeast BAS1 protein, a divergent member of the Myb family of transcription factors. *J Biol Chem* **269**: 17663-17669.
- Ikemoto, N., H.S. Kim, M. Kanazaki, Y. Ueno, S. Shuto, A. Matsuda, and Y. Wataya. 1999. Antisense oligodeoxynucleotides: useful tool for search and assessment of new targets for anti-malarial drugs. *Nucleic Acids Symp Ser*: 89-90.
- Introna, M., M. Luchetti, M. Castellano, M. Arsura, and J. Golay. 1994. The myb oncogene family of transcription factors: potent regulators of hematopoietic cell proliferation and differentiation. *Semin Cancer Biol* **5**: 113-124.
- Janse, C.J., P.F. van der Klooster, H.J. van der Kaay, M. van der Ploeg, and J.P. Overdulve. 1986. DNA synthesis in *Plasmodium berghei* during asexual and sexual development. *Mol Biochem Parasitol* **20**: 173-182.
- Jaouen, S., L. de Koning, C. Gaillard, E. Muselikova-Polanska, M. Stros, and F. Strauss. 2005. Determinants of specific binding of HMGB1 protein to hemicatenated DNA loops. *J Mol Biol* **353**: 822-837.

- Jayaraman, L., N.C. Moorthy, K.G. Murthy, J.L. Manley, M. Bustin, and C. Prives. 1998. High mobility group protein-1 (HMG-1) is a unique activator of p53. *Genes Dev* **12**: 462-472.
- Jones, P.A. 1999. The DNA methylation paradox. *Trends Genet* **15**: 34-37.
- Joshi, M.B., D.T. Lin, P.H. Chiang, N.D. Goldman, H. Fujioka, M. Aikawa, and C. Syn. 1999. Molecular cloning and nuclear localization of a histone deacetylase homologue in *Plasmodium falciparum*. *Mol Biochem Parasitol* **99**: 11-19.
- Kaiser, K., K. Matuschewski, N. Camargo, J. Ross, and S.H. Kappe. 2004. Differential transcriptome profiling identifies *Plasmodium* genes encoding pre-erythrocytic stage-specific proteins. *Mol Microbiol* **51**: 1221-1232.
- Karlson, D. and R. Imai. 2003. Conservation of the cold shock domain protein family in plants. *Plant Physiol* **131**: 12-15.
- Katinka, M.D., S. Duprat, E. Cornillot, G. Metenier, F. Thomarat, G. Prensier, V. Barbe, E. Peyretilade, P. Brottier, P. Wincker, F. Delbac, H. El Alaoui, P. Peyret, W. Saurin, M. Gouy, J. Weissenbach, and C.P. Vivares. 2001. Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. *Nature* **414**: 450-453.
- Katzen, A.L., T.B. Kornberg, and J.M. Bishop. 1985. Isolation of the proto-oncogene c-myc from *D. melanogaster*. *Cell* **41**: 449-456.
- Kellogg, D.R., A. Kikuchi, T. Fujii-Nakata, C.W. Turck, and A.W. Murray. 1995. Members of the NAP/SET family of proteins interact specifically with B-type cyclins. *J Cell Biol* **130**: 661-673.
- Kelman, Z. 1997. PCNA: structure, functions and interactions. *Oncogene* **14**: 629-640.
- Kim, J.L., D.B. Nikolov, and S.K. Burley. 1993. Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature* **365**: 520-527.
- Klempnauer, K.H. <http://www.uni-muenster.de/Chemie.bc/fields/ak-klempnauer/khk-topics.html>.
- Klempnauer, K.H., T.J. Gonda, and J.M. Bishop. 1982. Nucleotide sequence of the retroviral leukemia gene v-myc and its cellular progenitor c-myc: the architecture of a transduced oncogene. *Cell* **31**: 453-463.
- Klug, A. 1999. Zinc finger peptides for the regulation of gene expression. *J Mol Biol* **293**: 215-218.
- Klug, A. and D. Rhodes. 1987. Zinc fingers: a novel protein fold for nucleic acid recognition. *Cold Spring Harb Symp Quant Biol* **52**: 473-482.
- Knapp, B., U. Nau, E. Hundt, and H.A. Kupper. 1991. Demonstration of alternative splicing of a pre-mRNA expressed in the blood stage form of *Plasmodium falciparum*. *J Biol Chem* **266**: 7148-7154.
- Knuppel, R., P. Dietze, W. Lehnberg, K. Frech, and E. Wingender. 1994. TRANSFAC retrieval program: a network model database of eukaryotic transcription regulating sequences and proteins. *J Comput Biol* **1**: 191-198.
- Kohn, W.D., C.T. Mant, and R.S. Hodges. 1997. Alpha-helical protein assembly motifs. *J Biol Chem* **272**: 2583-2586.
- Kokkola, R., A. Andersson, G. Mullins, T. Ostberg, C.J. Treutiger, B. Arnold, P. Nawroth, U. Andersson, R.A. Harris, and H.E. Harris. 2005. RAGE is the major receptor for the proinflammatory activity of HMGB1 in rodent macrophages. *Scand J Immunol* **61**: 1-9.
- Kongkasuriyachai, D., H. Fujioka, and N. Kumar. 2004. Functional analysis of *Plasmodium falciparum* parasitophorous vacuole membrane protein (Pfs16) during gametocytogenesis and gametogenesis by targeted gene disruption. *Mol Biochem Parasitol* **133**: 275-285.

- Koornneef, M., C.J. Hanhart, and J.H. van der Veen. 1991. A genetic and physiological analysis of late flowering mutants in *Arabidopsis thaliana*. *Mol Gen Genet* **229**: 57-66.
- Kornberg, R.D. and Y. Lorch. 1999. Chromatin-modifying and -remodeling complexes. *Curr Opin Genet Dev* **9**: 148-151.
- Koshland, D. and A. Strunnikov. 1996. Mitotic chromosome condensation. *Annu Rev Cell Dev Biol* **12**: 305-333.
- Kumar, N., G. Cha, F. Pineda, J. Maciel, D. Haddad, M. Bhattacharyya, and E. Nagayasu. 2004. Molecular complexity of sexual development and gene regulation in *Plasmodium falciparum*. *Int J Parasitol* **34**: 1451-1458.
- Kun, J.F. and R.F. Anders. 1995. A *Plasmodium falciparum* gene encoding a high mobility group protein box. *Mol Biochem Parasitol* **71**: 249-253.
- Kwiatkowski, D., A.V. Hill, I. Sambou, P. Twumasi, J. Castracane, K.R. Manogue, A. Cerami, D.R. Brewster, and B.M. Greenwood. 1990. TNF concentration in fatal cerebral, non-fatal cerebral, and uncomplicated *Plasmodium falciparum* malaria. *Lancet* **336**: 1201-1204.
- Kyes, S., Z. Christodoulou, R. Pinches, and C. Newbold. 2002. Stage-specific merozoite surface protein 2 antisense transcripts in *Plasmodium falciparum*. *Mol Biochem Parasitol* **123**: 79-83.
- Kyes, S., R. Pinches, and C. Newbold. 2000. A simple RNA analysis method shows var and rif multigene family expression patterns in *Plasmodium falciparum*. *Mol Biochem Parasitol* **105**: 311-315.
- LaCount, D.J., M. Vignali, R. Chettier, A. Phansalkar, R. Bell, J.R. Hesselberth, L.W. Schoenfeld, I. Ota, S. Sahasrabudhe, C. Kurschner, S. Fields, and R.E. Hughes. 2005. A protein interaction network of the malaria parasite *Plasmodium falciparum*. *Nature* **438**: 103-107.
- Lambros, C. and J.P. Vanderberg. 1979. Synchronization of *Plasmodium falciparum* erythrocytic stages in culture. *J Parasitol* **65**: 418-420.
- Landschulz, W.H., P.F. Johnson, and S.L. McKnight. 1988. The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. *Science* **240**: 1759-1764.
- Lanzer, M., D. de Bruin, and J.V. Ravetch. 1992a. A sequence element associated with the *Plasmodium falciparum* KAHRP gene is the site of developmentally regulated protein-DNA interactions. *Nucleic Acids Res* **20**: 3051-3056.
- Lanzer, M., D. de Bruin, and J.V. Ravetch. 1992b. Transcription mapping of a 100 kb locus of *Plasmodium falciparum* identifies an intergenic region in which transcription terminates and reinitiates. *Embo J* **11**: 1949-1955.
- Lanzer, M., S.P. Wertheimer, D. de Bruin, and J.V. Ravetch. 1993. *Plasmodium*: control of gene expression in malaria parasites. *Exp Parasitol* **77**: 121-128.
- Lasonder, E., Y. Ishihama, J.S. Andersen, A.M. Vermunt, A. Pain, R.W. Sauerwein, W.M. Eling, N. Hall, A.P. Waters, H.G. Stunnenberg, and M. Mann. 2002. Analysis of the *Plasmodium falciparum* proteome by high-accuracy mass spectrometry. *Nature* **419**: 537-542.
- Le Roch, K.G., Y. Zhou, S. Batalov, and E.A. Winzeler. 2002. Monitoring the chromosome 2 intraerythrocytic transcriptome of *Plasmodium falciparum* using oligonucleotide arrays. *Am J Trop Med Hyg* **67**: 233-243.
- Le Roch, K.G., Y. Zhou, P.L. Blair, M. Grainger, J.K. Moch, J.D. Haynes, P. De La Vega, A.A. Holder, S. Batalov, D.J. Carucci, and E.A. Winzeler. 2003. Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science* **301**: 1503-1508.

- Lee, T.I. and R.A. Young. 2000. Transcription of eukaryotic protein-coding genes. *Annu Rev Genet* **34**: 77-137.
- Leech, J.H., J.W. Barnwell, L.H. Miller, and R.J. Howard. 1984. Identification of a strain-specific malarial antigen exposed on the surface of Plasmodium falciparum-infected erythrocytes. *J Exp Med* **159**: 1567-1575.
- Legube, G. and D. Trouche. 2003. Regulating histone acetyltransferases and deacetylases. *EMBO Rep* **4**: 944-947.
- Lemon, B. and R. Tjian. 2000. Orchestrated response: a symphony of transcription factors for gene control. *Genes Dev* **14**: 2551-2569.
- Levitt, A. 1993. RNA processing in malarial parasites. *Parasitol Today* **9**: 465-468.
- Lewis, B.A. and D. Reinberg. 2003. The mediator coactivator complex: functional and physical roles in transcriptional regulation. *J Cell Sci* **116**: 3667-3675.
- Li, H., Z. Han, Y. Lu, Y. Lin, L. Zhang, Y. Wu, and H. Wang. 2004. Isolation and functional characterization of a dynamin-like gene from Plasmodium falciparum. *Biochem Biophys Res Commun* **320**: 664-671.
- Li, J., R. Kokkola, S. Tabibzadeh, R. Yang, M. Ochani, X. Qiang, H.E. Harris, C.J. Czura, H. Wang, L. Ulloa, H.S. Warren, L.L. Moldawer, M.P. Fink, U. Andersson, K.J. Tracey, and H. Yang. 2003. Structural basis for the proinflammatory cytokine activity of high mobility group box 1. *Mol Med* **9**: 37-45.
- Linder, T. and C.M. Gustafsson. 2004. The Sohl/MED31 protein is an ancient component of Schizosaccharomyces pombe and Saccharomyces cerevisiae Mediator. *J Biol Chem* **279**: 49455-49459.
- Lobo, C.A., H. Fujioka, M. Aikawa, and N. Kumar. 1999. Disruption of the Pfg27 locus by homologous recombination leads to loss of the sexual phenotype in *P. falciparum*. *Mol Cell* **3**: 793-798.
- Longhurst, H.J. and A.A. Holder. 1997. The histones of Plasmodium falciparum: identification, purification and a possible role in the pathology of malaria. *Parasitology* **114 ( Pt 5)**: 413-419.
- Lotze, M.T. and K.J. Tracey. 2005. High-mobility group box 1 protein (HMGB1): nuclear weapon in the immune arsenal. *Nat Rev Immunol* **5**: 331-342.
- Luger, K. and T.J. Richmond. 1998. The histone tails of the nucleosome. *Curr Opin Genet Dev* **8**: 140-146.
- Luscher, B. and R.N. Eisenman. 1990. New light on Myc and Myb. Part I. Myc. *Genes Dev* **4**: 2025-2035.
- Lutterbach, B., D. Sun, J. Schuetz, and S.W. Hiebert. 1998. The MYND motif is required for repression of basal transcription from the multidrug resistance 1 promoter by the t(8;21) fusion protein. *Mol Cell Biol* **18**: 3604-3611.
- MacPherson, G.G., M.J. Warrell, N.J. White, S. Looareesuwan, and D.A. Warrell. 1985. Human cerebral malaria. A quantitative ultrastructural analysis of parasitized erythrocyte sequestration. *Am J Pathol* **119**: 385-401.
- Maity, S.N. and B. de Crombrughe. 1998. Role of the CCAAT-binding protein CBF/NF-Y in transcription. *Trends Biochem Sci* **23**: 174-178.
- Malhotra, P., P.V. Dasaradhi, A. Kumar, A. Mohammed, N. Agrawal, R.K. Bhatnagar, and V.S. Chauhan. 2002. Double-stranded RNA-mediated gene silencing of cysteine proteases (falcipain-1 and -2) of Plasmodium falciparum. *Mol Microbiol* **45**: 1245-1254.
- Matsumoto, K. and A.P. Wolffe. 1998. Gene regulation by Y-box proteins: coupling control of transcription and translation. *Trends Cell Biol* **8**: 318-323.

- Matys, V., E. Fricke, R. Geffers, E. Gossling, M. Haubrock, R. Hehl, K. Hornischer, D. Karas, A.E. Kel, O.V. Kel-Margoulis, D.U. Kloos, S. Land, B. Lewicki-Potapov, H. Michael, R. Munch, I. Reuter, S. Rotert, H. Saxel, M. Scheer, S. Thiele, and E. Wingender. 2003. TRANSFAC: transcriptional regulation, from patterns to profiles. *Nucleic Acids Res* **31**: 374-378.
- McAndrew, M.B., M. Read, P.F. Sims, and J.E. Hyde. 1993. Characterisation of the gene encoding an unusually divergent TATA-binding protein (TBP) from the extremely A+T-rich human malaria parasite *Plasmodium falciparum*. *Gene* **124**: 165-171.
- McConkey, G.A., J.W. Pinney, D.R. Westhead, K. Plueckhahn, T.B. Fitzpatrick, P. Macheroux, and B. Kappes. 2004. Annotating the *Plasmodium* genome and the enigma of the shikimate pathway. *Trends Parasitol* **20**: 60-65.
- McKnight, S.L. 1982. Functional relationships between transcriptional control signals of the thymidine kinase gene of herpes simplex virus. *Cell* **31**: 355-365.
- McRobert, L. and G.A. McConkey. 2002. RNA interference (RNAi) inhibits growth of *Plasmodium falciparum*. *Mol Biochem Parasitol* **119**: 273-278.
- Mello, J.A. and G. Almouzni. 2001. The ins and outs of nucleosome assembly. *Curr Opin Genet Dev* **11**: 136-141.
- Merckx, A., K. Le Roch, M.P. Nivez, D. Dorin, P. Alano, G.J. Gutierrez, A.R. Nebreda, D. Goldring, C. Whittle, S. Patterson, D. Chakrabarti, and C. Doerig. 2003. Identification and initial characterization of three novel cyclin-related proteins of the human malaria parasite *Plasmodium falciparum*. *J Biol Chem* **278**: 39839-39850.
- Militello, K.T., M. Dodge, L. Bethke, and D.F. Wirth. 2004. Identification of regulatory elements in the *Plasmodium falciparum* genome. *Mol Biochem Parasitol* **134**: 75-88.
- Militello, K.T., V. Patel, A.D. Chessler, J.K. Fisher, J.M. Kasper, A. Gunasekera, and D.F. Wirth. 2005. RNA polymerase II synthesizes antisense RNA in *Plasmodium falciparum*. *Rna* **11**: 365-370.
- Millet, P., G.H. Campbell, A.J. Sulzer, K.K. Grady, J. Pohl, M. Aikawa, and W.E. Collins. 1993. Immunogenicity of the *Plasmodium falciparum* asexual blood-stage synthetic peptide vaccine SPf66. *Am J Trop Med Hyg* **48**: 424-431.
- Mo, Y., W. Ho, K. Johnston, and R. Marmorstein. 2001. Crystal structure of a ternary SAP-1/SRF/c-fos SRE DNA complex. *J Mol Biol* **314**: 495-506.
- Morrow, B.E., Q. Ju, and J.R. Warner. 1993. A bipartite DNA-binding domain in yeast Reb1p. *Mol Cell Biol* **13**: 1173-1182.
- Muhia, D.K., C.A. Swales, U. Eckstein-Ludwig, S. Saran, S.D. Polley, J.M. Kelly, P. Schaap, S. Krishna, and D.A. Baker. 2003. Multiple splice variants encode a novel adenylyl cyclase of possible plastid origin expressed in the sexual stage of the malaria parasite *Plasmodium falciparum*. *J Biol Chem* **278**: 22014-22022.
- Muller, S., P. Scaffidi, B. Degryse, T. Bonaldi, L. Ronfani, A. Agresti, M. Beltrame, and M.E. Bianchi. 2001. New EMBO members' review: the double life of HMGB1 chromatin protein: architectural factor and extracellular signal. *Embo J* **20**: 4337-4340.
- Murphy, F.V.t., R.M. Sweet, and M.E. Churchill. 1999. The structure of a chromosomal high mobility group protein-DNA complex reveals sequence-neutral mechanisms important for non-sequence-specific DNA recognition. *Embo J* **18**: 6610-6618.
- Murre, C., P.S. McCaw, and D. Baltimore. 1989. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. *Cell* **56**: 777-783.

- Nakajima, N., M. Horikoshi, and R.G. Roeder. 1988. Factors involved in specific transcription by mammalian RNA polymerase II: purification, genetic specificity, and TATA box-promoter interactions of TFIID. *Mol Cell Biol* **8**: 4028-4040.
- Nan, X., H.H. Ng, C.A. Johnson, C.D. Laherty, B.M. Turner, R.N. Eisenman, and A. Bird. 1998. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* **393**: 386-389.
- Narum, D.L. and A.W. Thomas. 1994. Differential localization of full-length and processed forms of PF83/AMA-1 an apical membrane antigen of Plasmodium falciparum merozoites. *Mol Biochem Parasitol* **67**: 59-68.
- Nemeth, A. and G. Langst. 2004. Chromatin higher order structure: opening up chromatin for transcription. *Brief Funct Genomic Proteomic* **2**: 334-343.
- Ness, S.A., A. Marknell, and T. Graf. 1989. The v-myb oncogene product binds to and activates the promyelocyte-specific mim-1 gene. *Cell* **59**: 1115-1125.
- Nijhout, M.M. and R. Carter. 1978. Gamete development in malaria parasites: bicarbonate-dependent stimulation by pH in vitro. *Parasitology* **76**: 39-53.
- Nikolov, D.B., H. Chen, E.D. Halay, A.A. Usheva, K. Hisatake, D.K. Lee, R.G. Roeder, and S.K. Burley. 1995. Crystal structure of a TFIIB-TBP-TATA-element ternary complex. *Nature* **377**: 119-128.
- Nomura, N., M. Takahashi, M. Matsui, S. Ishii, T. Date, S. Sasamoto, and R. Ishizaki. 1988. Isolation of human cDNA clones of myb-related genes, A-myb and B-myb. *Nucleic Acids Res* **16**: 11075-11089.
- Noonpakdee, W., J. Pothikasikorn, W. Nimitsantiwong, and P. Wilairat. 2003. Inhibition of Plasmodium falciparum proliferation in vitro by antisense oligodeoxynucleotides against malarial topoisomerase II. *Biochem Biophys Res Commun* **302**: 659-664.
- Nosny, P. 1980. [Alphonse Laveran and the discovery of the malarial parasite]. *Bull Acad Natl Med* **164**: 80-87.
- O'Donnell, R.A., L.H. Freitas-Junior, P.R. Preiser, D.H. Williamson, M. Duraisingh, T.F. McElwain, A. Scherf, A.F. Cowman, and B.S. Crabb. 2002. A genetic screen for improved plasmid segregation reveals a role for Rep20 in the interaction of Plasmodium falciparum chromosomes. *Embo J* **21**: 1231-1239.
- Oelgeschlager, T., C.M. Chiang, and R.G. Roeder. 1996. Topology and reorganization of a human TFIID-promoter complex. *Nature* **382**: 735-738.
- Ogata, K., S. Morikawa, H. Nakamura, A. Sekikawa, T. Inoue, H. Kanai, A. Sarai, S. Ishii, and Y. Nishimura. 1994. Solution structure of a specific DNA complex of the Myb DNA-binding domain with cooperative recognition helices. *Cell* **79**: 639-648.
- Ohi, R., A. Feoktistova, S. McCann, V. Valentine, A.T. Look, J.S. Lipsick, and K.L. Gould. 1998. Myb-related Schizosaccharomyces pombe cdc5p is structurally and functionally conserved in eukaryotes. *Mol Cell Biol* **18**: 4097-4108.
- Osta, M., L. Gannoun-Zaki, S. Bonnefoy, C. Roy, and H.J. Vial. 2002. A 24 bp cis-acting element essential for the transcriptional activity of Plasmodium falciparum CDP-diacylglycerol synthase gene promoter. *Mol Biochem Parasitol* **121**: 87-98.
- Pabo, C.O. and R.T. Sauer. 1992. Transcription factors: structural families and principles of DNA recognition. *Annu Rev Biochem* **61**: 1053-1095.
- Pace, T., C. Birago, C.J. Janse, L. Picci, and M. Ponzi. 1998. Developmental regulation of a Plasmodium gene involves the generation of stage-specific 5' untranslated sequences. *Mol Biochem Parasitol* **97**: 45-53.

- Paillard, F., G. Sterkers, and C. Vaquero. 1990. Transcriptional and post-transcriptional regulation of TcR, CD4 and CD8 gene expression during activation of normal human T lymphocytes. *Embo J* **9**: 1867-1872.
- Palumbo, R., M. Sampaolesi, F. De Marchis, R. Tonlorenzi, S. Colombetti, A. Mondino, G. Cossu, and M.E. Bianchi. 2004. Extracellular HMGB1, a signal of tissue damage, induces mesoangioblast migration and proliferation. *J Cell Biol* **164**: 441-449.
- Park, J.S., D. Svetkauskaite, Q. He, J.Y. Kim, D. Strassheim, A. Ishizaka, and E. Abraham. 2004. Involvement of toll-like receptors 2 and 4 in cellular activation by high mobility group box 1 protein. *J Biol Chem* **279**: 7370-7377.
- Patankar, S., A. Munasinghe, A. Shoaibi, L.M. Cummings, and D.F. Wirth. 2001. Serial analysis of gene expression in *Plasmodium falciparum* reveals the global expression profile of erythrocytic stages and the presence of anti-sense transcripts in the malarial parasite. *Mol Biol Cell* **12**: 3114-3125.
- Payne, D. 1987. Spread of chloroquine resistance in *Plasmodium falciparum*. *Parasitol Today* **3**: 241-246.
- Pe Than, M. and S. Tin. 1986. The efficacy of artemether (qinghaosu) in *Plasmodium falciparum* and *P. vivax* in Burma. *Southeast Asian J Trop Med Public Health* **17**: 19-22.
- Peters, C.W., A.E. Sippel, M. Vingron, and K.H. Klempnauer. 1987. Drosophila and vertebrate myb proteins share two conserved regions, one of which functions as a DNA-binding domain. *Embo J* **6**: 3085-3090.
- Pinson, B., E.M. Brendeford, O.S. Gabrielsen, and B. Daignan-Fornier. 2001. Highly conserved features of DNA binding between two divergent members of the Myb family of transcription factors. *Nucleic Acids Res* **29**: 527-535.
- Polson, H.E. and M.J. Blackman. 2005. A role for poly(dA)poly(dT) tracts in directing activity of the *Plasmodium falciparum* calmodulin gene promoter. *Mol Biochem Parasitol* **141**: 179-189.
- Porter, M.E. 2001. The DNA polymerase delta promoter from *Plasmodium falciparum* contains an unusually long 5' untranslated region and intrinsic DNA curvature. *Mol Biochem Parasitol* **114**: 249-255.
- Ptashne, M.a.G., A. 2002. Genes and signals. *New York : Cold Spring Harbor*.
- Pugh, B.F. and R. Tjian. 1991. Transcription from a TATA-less promoter requires a multisubunit TFIID complex. *Genes Dev* **5**: 1935-1945.
- Putilina, T., P. Wong, and S. Gentleman. 1999. The DHHC domain: a new highly conserved cysteine-rich motif. *Mol Cell Biochem* **195**: 219-226.
- Ralph, S.A., E. Bischoff, D. Mattei, O. Sismeiro, M.A. Dillies, G. Guigon, J.Y. Coppee, P.H. David, and A. Scherf. 2005. Transcriptome analysis of antigenic variation in *Plasmodium falciparum*--var silencing is not dependent on antisense RNA. *Genome Biol* **6**: R93.
- Ralston, R. and J.M. Bishop. 1983. The protein products of the myc and myb oncogenes and adenovirus E1a are structurally related. *Nature* **306**: 803-806.
- Read, C.M., P.D. Cary, C. Crane-Robinson, P.C. Driscoll, and D.G. Norman. 1993. Solution structure of a DNA-binding domain from HMG1. *Nucleic Acids Res* **21**: 3427-3436.
- Reichenow, E. 1920. *Zentralbl. Bakteriologie*. **85**: 207-216.
- Reinberg, D. and R.G. Roeder. 1987. Factors involved in specific transcription by mammalian RNA polymerase II. Purification and functional analysis of initiation factors IIB and IIE. *J Biol Chem* **262**: 3310-3321.

- Rich, S.M. and F.J. Ayala. 2003. Progress in malaria research: the case for phylogenetics. *Adv Parasitol* **54**: 255-280.
- Robson, F., M.M. Costa, S.R. Hepworth, I. Vizir, M. Pineiro, P.H. Reeves, J. Putterill, and G. Coupland. 2001. Functional importance of conserved domains in the flowering-time gene CONSTANS demonstrated by analysis of mutant alleles and transgenic plants. *Plant J* **28**: 619-631.
- Ronfani, L., M. Ferraguti, L. Croci, C.E. Ovitt, H.R. Scholer, G.G. Consalez, and M.E. Bianchi. 2001. Reduced fertility and spermatogenesis defects in mice lacking chromosomal protein Hmgb2. *Development* **128**: 1265-1273.
- Rost, B. 1999. Twilight zone of protein sequence alignments. *Protein Eng* **12**: 85-94.
- Ruvalcaba-Salazar, O.K., M. del Carmen Ramirez-Estudillo, D. Montiel-Condado, F. Recillas-Targa, M. Vargas, and R. Hernandez-Rivas. 2005. Recombinant and native Plasmodium falciparum TATA-binding-protein binds to a specific TATA box element in promoter regions. *Mol Biochem Parasitol* **140**: 183-196.
- Saito, S., M. Miyaji-Yamaguchi, T. Shimoyama, and K. Nagata. 1999. Functional domains of template-activating factor-I as a protein phosphatase 2A inhibitor. *Biochem Biophys Res Commun* **259**: 471-475.
- Samach, A., H. Onouchi, S.E. Gold, G.S. Ditta, Z. Schwarz-Sommer, M.F. Yanofsky, and G. Coupland. 2000. Distinct roles of CONSTANS target genes in reproductive development of Arabidopsis. *Science* **288**: 1613-1616.
- Scherf, A., R. Hernandez-Rivas, P. Buffet, E. Bottius, C. Benatar, B. Pouvelle, J. Gysin, and M. Lanzer. 1998. Antigenic variation in malaria: in situ switching, relaxed and mutually exclusive transcription of var genes during intra-erythrocytic development in Plasmodium falciparum. *Embo J* **17**: 5418-5426.
- Schnitzler, G.R., S. Sif, and R.E. Kingston. 1998. A model for chromatin remodeling by the SWI/SNF family. *Cold Spring Harb Symp Quant Biol* **63**: 535-543.
- Scott, M.P., J.W. Tamkun, and G.W. Hartzell, 3rd. 1989. The structure and function of the homeodomain. *Biochim Biophys Acta* **989**: 25-48.
- Seo, S.B., P. McNamara, S. Heo, A. Turner, W.S. Lane, and D. Chakravarti. 2001. Regulation of histone acetylation and transcription by INHAT, a human cellular complex containing the set oncoprotein. *Cell* **104**: 119-130.
- Shaw, M.K., J. Thompson, and R.E. Sinden. 1996. Localization of ribosomal RNA and Pbs21-mRNA in the sexual stages of Plasmodium berghei using electron microscope in situ hybridization. *Eur J Cell Biol* **71**: 270-276.
- Silamut, K., N.H. Phu, C. Whitty, G.D. Turner, K. Louwrier, N.T. Mai, J.A. Simpson, T.T. Hien, and N.J. White. 1999. A quantitative analysis of the microvascular sequestration of malaria parasites in the human brain. *Am J Pathol* **155**: 395-410.
- Silvestrini, F., P. Alano, and J.L. Williams. 2000. Commitment to the production of male and female gametocytes in the human malaria parasite Plasmodium falciparum. *Parasitology* **121 Pt 5**: 465-471.
- Silvestrini, F., Z. Bozdech, A. Lanfrancotti, E. Di Giulio, E. Bultrini, L. Picci, J.L. Derisi, E. Pizzi, and P. Alano. 2005. Genome-wide identification of genes upregulated at the onset of gametocytogenesis in Plasmodium falciparum. *Mol Biochem Parasitol* **143**: 100-110.
- Sinden, R.E. 1983. Sexual development of malarial parasites. *Adv Parasitol* **22**: 153-216.
- Sinden, R.E. and R.H. Hartley. 1985. Identification of the meiotic division of malarial parasites. *J Protozool* **32**: 742-744.



- Sinden, R.E., R.H. Hartley, and L. Winger. 1985. The development of *Plasmodium* ookinetes *in vitro*: an ultrastructural study including a description of meiotic division. *Parasitology* **91 ( Pt 2)**: 227-244.
- Singh, N., P. Preiser, L. Renia, B. Balu, J. Barnwell, P. Blair, W. Jarra, T. Voza, I. Landau, and J.H. Adams. 2004. Conservation and developmental control of alternative splicing in *maeb1* among malaria parasites. *J Mol Biol* **343**: 589-599.
- Smale, S.T. and J.T. Kadonaga. 2003. The RNA polymerase II core promoter. *Annu Rev Biochem* **72**: 449-479.
- Snow, R.W., C.A. Guerra, A.M. Noor, H.Y. Myint, and S.I. Hay. 2005. The global distribution of clinical episodes of *Plasmodium falciparum* malaria. *Nature* **434**: 214-217.
- Sopta, M., R.W. Carthew, and J. Greenblatt. 1985. Isolation of three proteins that bind to mammalian RNA polymerase II. *J Biol Chem* **260**: 10353-10360.
- Soullier, S., P. Jay, F. Poulat, J.M. Vanacker, P. Berta, and V. Laudet. 1999. Diversification pattern of the HMG and SOX family members during evolution. *J Mol Evol* **48**: 517-527.
- Stober-Grasser, U., B. Brydolf, X. Bin, F. Grasser, R.A. Firtel, and J.S. Lipsick. 1992. The Myb DNA-binding domain is highly conserved in *Dictyostelium discoideum*. *Oncogene* **7**: 589-596.
- Strahl, B.D. and C.D. Allis. 2000. The language of covalent histone modifications. *Nature* **403**: 41-45.
- Struhl, K. 1985. Naturally occurring poly(dA-dT) sequences are upstream promoter elements for constitutive transcription in yeast. *Proc Natl Acad Sci U S A* **82**: 8419-8423.
- Struhl, K. 1995. Yeast transcriptional regulatory mechanisms. *Annu Rev Genet* **29**: 651-674.
- Su, X.Z., Y. Wu, C.D. Sifri, and T.E. Wellems. 1996. Reduced extension temperatures required for PCR amplification of extremely A+T-rich DNA. *Nucleic Acids Res* **24**: 1574-1575.
- Sudarsanam, P., V.R. Iyer, P.O. Brown, and F. Winston. 2000. Whole-genome expression analysis of *snf/swi* mutants of *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* **97**: 3364-3369.
- Taddei, A., Ray-Gallet, D. and Almouzni, G. 2000. Assemblage et remodelage: le nucléosome sous influence. *Med Sci* **16**: 603-610.
- Taguchi, A., D.C. Blood, G. del Toro, A. Canet, D.C. Lee, W. Qu, N. Tanji, Y. Lu, E. Lalla, C. Fu, M.A. Hofmann, T. Kislinger, M. Ingram, A. Lu, H. Tanaka, O. Hori, S. Ogawa, D.M. Stern, and A.M. Schmidt. 2000. Blockade of RAGE-amphotericin signalling suppresses tumour growth and metastases. *Nature* **405**: 354-360.
- Tahirov, T.H., K. Sato, E. Ichikawa-Iwata, M. Sasaki, T. Inoue-Bungo, M. Shiina, K. Kimura, S. Takata, A. Fujikawa, H. Morii, T. Kumasaka, M. Yamamoto, S. Ishii, and K. Ogata. 2002. Mechanism of c-Myb-C/EBP beta cooperation from separated sites on a promoter. *Cell* **108**: 57-70.
- Takagi, Y., H. Komori, W.H. Chang, A. Hudmon, H. Erdjument-Bromage, P. Tempst, and R.D. Kornberg. 2003. Revised subunit structure of yeast transcription factor IIIH (TFIIH) and reconciliation with human TFIIH. *J Biol Chem* **278**: 43897-43900.
- Talisuna, A.O., P. Bloland, and U. D'Alessandro. 2004. History, dynamics, and public health importance of malaria parasite resistance. *Clin Microbiol Rev* **17**: 235-254.
- Tan, S., Y. Hunziker, D.F. Sargent, and T.J. Richmond. 1996. Crystal structure of a yeast TFIIA/TBP/DNA complex. *Nature* **381**: 127-151.
- Tan, S. and T.J. Richmond. 1998. Eukaryotic transcription factors. *Curr Opin Struct Biol* **8**: 41-48.

- Tanese, N., B.F. Pugh, and R. Tjian. 1991. Coactivators for a proline-rich activator purified from the multisubunit human TFIID complex. *Genes Dev* **5**: 2212-2224.
- Tanikawa, J., T. Yasukawa, M. Enari, K. Ogata, Y. Nishimura, S. Ishii, and A. Sarai. 1993. Recognition of specific DNA sequences by the c-myc protooncogene product: role of three repeat units in the DNA-binding domain. *Proc Natl Acad Sci U S A* **90**: 9320-9324.
- Teo, S.H., K.D. Grasser, and J.O. Thomas. 1995. Differences in the DNA-binding properties of the HMG-box domains of HMG1 and the sex-determining factor SRY. *Eur J Biochem* **230**: 943-950.
- Thomas, J.O. and A.A. Travers. 2001. HMG1 and 2, and related 'architectural' DNA-binding proteins. *Trends Biochem Sci* **26**: 167-174.
- Thompson, J.D., D.G. Higgins, and T.J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**: 4673-4680.
- Thomsen, M.S., L. Franssen, D. Launholt, P. Fojan, and K.D. Grasser. 2004. Interactions of the basic N-terminal and the acidic C-terminal domains of the maize chromosomal HMGB1 protein. *Biochemistry* **43**: 8029-8037.
- Tice-Baldwin, K., G.R. Fink, and K.T. Arndt. 1989. BAS1 has a Myb motif and activates HIS4 transcription only in combination with BAS2. *Science* **246**: 931-935.
- Tjian, R. 1995. Molecular machines that control genes. *Sci Am* **272**: 54-61.
- Trager, W. and J.B. Jensen. 1976. Human malaria parasites in continuous culture. *Science* **193**: 673-675.
- Travers, A.A. 2003. Priming the nucleosome: a role for HMGB proteins? *EMBO Rep* **4**: 131-136.
- van Lin, L.H., T. Pace, C.J. Janse, C. Birago, J. Ramesar, L. Picci, M. Ponzi, and A.P. Waters. 2001. Interspecies conservation of gene order and intron-exon structure in a genomic locus of high gene density and complexity in Plasmodium. *Nucleic Acids Res* **29**: 2059-2068.
- Vanhee-Brossollet, C. and C. Vaquero. 1998. Do natural antisense transcripts make sense in eukaryotes? *Gene* **211**: 1-9.
- Vinkenoog, R., B. Veldhuisen, M.A. Speranca, H.A. del Portillo, C. Janse, and A.P. Waters. 1995. Comparison of introns in a cdc2-homologous gene within a number of Plasmodium species. *Mol Biochem Parasitol* **71**: 233-241.
- Vivier, E.a.D., I. 1989. Phylum Apicomplexa. *Handbook of Protoctista* **Chapter 30**: 549-573.
- Voronova, A. and D. Baltimore. 1990. Mutations that disrupt DNA binding and dimer formation in the E47 helix-loop-helix protein map to distinct domains. *Proc Natl Acad Sci U S A* **87**: 4722-4726.
- Voss, T.S., M. Kaestli, D. Vogel, S. Bopp, and H.P. Beck. 2003. Identification of nuclear proteins that interact differentially with Plasmodium falciparum var gene promoters. *Mol Microbiol* **48**: 1593-1607.
- Walliker, D., I.A. Quakyi, T.E. Wellems, T.F. McCutchan, A. Szarfman, W.T. London, L.M. Corcoran, T.R. Burkot, and R. Carter. 1987. Genetic analysis of the human malaria parasite Plasmodium falciparum. *Science* **236**: 1661-1666.
- Wang, H., O. Bloom, M. Zhang, J.M. Vishnubhakat, M. Ombrellino, J. Che, A. Frazier, H. Yang, S. Ivanova, L. Borovikova, K.R. Manogue, E. Faist, E. Abraham, J. Andersson, U. Andersson, P.E. Molina, N.N. Abumrad, A. Sama, and K.J. Tracey. 1999. HMG-1 as a late mediator of endotoxin lethality in mice. *Science* **285**: 248-251.

- Wang, W., M. Carey, and J.D. Gralla. 1992. Polymerase II promoter activation: closed complex formation and ATP-driven start site opening. *Science* **255**: 450-453.
- Wanidworanun, C., R.L. Nagel, and H.L. Shear. 1999. Antisense oligonucleotides targeting malarial aldolase inhibit the asexual erythrocytic stages of *Plasmodium falciparum*. *Mol Biochem Parasitol* **102**: 91-101.
- Warren, A.J. 2002. Eukaryotic transcription factors. *Curr Opin Struct Biol* **12**: 107-114.
- Waters, A.P. 1994. The ribosomal RNA genes of *Plasmodium*. *Adv Parasitol* **34**: 33-79.
- Webb, M. and J.O. Thomas. 1999. Structure-specific binding of the two tandem HMG boxes of HMG1 to four-way junction DNA is mediated by the A domain. *J Mol Biol* **294**: 373-387.
- Weideman, C.A., R.C. Netter, L.R. Benjamin, J.J. McAllister, L.A. Schmiedekamp, R.A. Coleman, and B.F. Pugh. 1997. Dynamic interplay of TFIIA, TBP and TATA DNA. *J Mol Biol* **271**: 61-75.
- Weir, H.M., P.J. Kraulis, C.S. Hill, A.R. Raine, E.D. Laue, and J.O. Thomas. 1993. Structure of the HMG box motif in the B-domain of HMG1. *Embo J* **12**: 1311-1319.
- Wernsdorfer, W.H. 1991. The development and spread of drug-resistant malaria. *Parasitol Today* **7**: 297-303.
- Weston, K. and J.M. Bishop. 1989. Transcriptional activation by the v-myb oncogene and its cellular progenitor, c-myb. *Cell* **58**: 85-93.
- WHO. 2000. WHO Expert Committee on Malaria - Twentieth Report. World Health Organization, Geneva.
- Wickens, M., P. Anderson, and R.J. Jackson. 1997. Life and death in the cytoplasm: messages from the 3' end. *Curr Opin Genet Dev* **7**: 220-232.
- Wickham, M.E., J.K. Thompson, and A.F. Cowman. 2003. Characterisation of the merozoite surface protein-2 promoter using stable and transient transfection in *Plasmodium falciparum*. *Mol Biochem Parasitol* **129**: 147-156.
- Wingender, E., X. Chen, E. Fricke, R. Geffers, R. Hehl, I. Liebich, M. Krull, V. Matys, H. Michael, R. Ohnhauser, M. Pruss, F. Schacherer, S. Thiele, and S. Urbach. 2001. The TRANSFAC system on gene expression regulation. *Nucleic Acids Res* **29**: 281-283.
- Wistow, G. 1990. Cold shock and DNA binding. *Nature* **344**: 823-824.
- Wolffe, A.P. 1999. Architectural regulations and Hmg1. *Nat Genet* **22**: 215-217.
- Workman, J.L. and R.E. Kingston. 1998. Alteration of nucleosome structure as a mechanism of transcriptional regulation. *Annu Rev Biochem* **67**: 545-579.
- Yang, H., H. Wang, C.J. Czura, and K.J. Tracey. 2005. The cytokine activity of HMGB1. *J Leukoc Biol* **78**: 1-8.
- Yoshioka, K., K. Saito, T. Tanabe, A. Yamamoto, Y. Ando, Y. Nakamura, H. Shirakawa, and M. Yoshida. 1999. Differences in DNA recognition and conformational change activity between boxes A and B in HMG2 protein. *Biochemistry* **38**: 589-595.
- Young, J.A., Q.L. Fivelman, P.L. Blair, P. de la Vega, K.G. Le Roch, Y. Zhou, D.J. Carucci, D.A. Baker, and E.A. Winzeler. 2005. The *Plasmodium falciparum* sexual development transcriptome: a microarray analysis using ontology-based pattern identification. *Mol Biochem Parasitol* **143**: 67-79.
- Zappavigna, V., L. Falciola, M. Helmer-Citterich, F. Mavilio, and M.E. Bianchi. 1996. HMG1 interacts with HOX proteins and enhances their DNA binding and transcriptional activation. *Embo J* **15**: 4981-4991.
- Zawel, L. and D. Reinberg. 1993. Initiation of transcription by RNA polymerase II: a multi-step process. *Prog Nucleic Acid Res Mol Biol* **44**: 67-108.

Zheng, X.M., V. Moncollin, J.M. Egly, and P. Chambon. 1987. A general transcription factor forms a stable complex with RNA polymerase B (II). *Cell* **50**: 361-368.

# ANNEXES

## I - NUMEROS D'ACCESSION DES PROTEINES LIEES A LA REGULATION DE L'EXPRESSION DES GENES CHEZ *P. FALCIPARUM*

### I.1 – Protéines identifiées par nos soins

<b>Numéro d'accession</b>	<b>Superclasse ou fonction</b>	<b>Domaine caractéristique</b>
<a href="#">MAL13P1.61</a>	Architecture bêta	cold shock
<a href="#">PFA0470c</a>	Architecture bêta	cold shock protein, putative
<a href="#">PFL0145c</a>	Architecture bêta	HMG
<a href="#">MAL8P1.72</a>	Architecture bêta	HMG
<a href="#">PFL0290w</a>	Architecture bêta	HMG
<a href="#">MAL13P1.290</a>	Architecture bêta	HMG
<a href="#">PF10_0024</a>	Architecture bêta	MADS box
<a href="#">PFC0345w</a>	Domaines basiques	B-box
<a href="#">PFE0895c</a>	Domaines basiques	B-box
<a href="#">MAL13P1.37</a>	Domaines basiques	B-box
<a href="#">PFA0550w</a>	Domaines basiques	b-Zip
<a href="#">PFB0615c</a>	Domaines basiques	b-Zip
<a href="#">PFB0690w</a>	Domaines basiques	b-Zip
<a href="#">PFC0235w</a>	Domaines basiques	b-Zip
<a href="#">PFC0365w</a>	Domaines basiques	b-Zip
<a href="#">PFC0760c</a>	Domaines basiques	b-Zip
<a href="#">PFC0960c</a>	Domaines basiques	b-Zip
<a href="#">PFD0330w</a>	Domaines basiques	b-Zip
<a href="#">PFF0140c</a>	Domaines basiques	b-Zip
<a href="#">PFF0375c</a>	Domaines basiques	b-Zip
<a href="#">PFF1285w</a>	Domaines basiques	b-Zip
<a href="#">PF07_0038</a>	Domaines basiques	b-Zip
<a href="#">MAL7P1.87</a>	Domaines basiques	b-Zip
<a href="#">PFI0175w</a>	Domaines basiques	b-Zip
<a href="#">PF11_0268</a>	Domaines basiques	b-Zip
<a href="#">PF11_0342</a>	Domaines basiques	b-Zip
<a href="#">PF11_0392</a>	Domaines basiques	b-Zip
<a href="#">PF11_0433</a>	Domaines basiques	b-Zip
<a href="#">PFL1930w</a>	Domaines basiques	b-Zip
<a href="#">PFL2185w</a>	Domaines basiques	b-Zip
<a href="#">PF13_0047</a>	Domaines basiques	b-Zip
<a href="#">MAL13P1.72</a>	Domaines basiques	b-Zip
<a href="#">MAL13P1.97</a>	Domaines basiques	b-Zip
<a href="#">MAL13P1.140</a>	Domaines basiques	b-Zip
<a href="#">MAL13P1.188</a>	Domaines basiques	b-Zip
<a href="#">MAL13P1.295</a>	Domaines basiques	b-Zip
<a href="#">PFC0690c</a>	Zinc finger	C2H2
<a href="#">PFD0375w</a>	Zinc finger	C2H2
<a href="#">PFD0595w</a>	Zinc finger	C2H2
<a href="#">PFD0765w</a>	Zinc finger	C2H2
<a href="#">PFI0590c</a>	Zinc finger	C2H2
<a href="#">PFI1255w</a>	Zinc finger	C2H2

<a href="#">PF10_0091</a>	Zinc finger	C2H2
<a href="#">PFL0455c</a>	Zinc finger	C2H2
<a href="#">PFL2075c</a>	Zinc finger	C2H2
<a href="#">PF13_0278</a>	Zinc finger	C2H2
<a href="#">PF14_0479</a>	Zinc finger	C2H2
<a href="#">PF14_0559</a>	Zinc finger	C2H2
<a href="#">PF14_0612</a>	Zinc finger	C2H2
<a href="#">PF14_0643</a>	Zinc finger	C2H2
<a href="#">PF14_0657</a>	Zinc finger	C2H2
<a href="#">PF14_0707</a>	Zinc finger	C2H2
<a href="#">PF14_0383</a>	Zinc finger	Constans
<a href="#">PFB0140w</a>	Zinc finger	DHHC-type zinc finger protein
<a href="#">PFB0725c</a>	Zinc finger	DHHC-type zinc finger protein
<a href="#">PFE1415w</a>	Zinc finger	DHHC-type zinc finger protein
<a href="#">PFF0485c</a>	Zinc finger	DHHC-type zinc finger protein
<a href="#">MAL7P1.68</a>	Zinc finger	DHHC-type zinc finger protein
<a href="#">PFI0665w</a>	Zinc finger	DHHC-type zinc finger protein
<a href="#">PFI1580c</a>	Zinc finger	DHHC-type zinc finger protein
<a href="#">PF10_0273</a>	Zinc finger	DHHC-type zinc finger protein
<a href="#">PF11_0167</a>	Zinc finger	DHHC-type zinc finger protein
<a href="#">PF11_0217</a>	Zinc finger	DHHC-type zinc finger protein
<a href="#">MAL13P1.117</a>	Zinc finger	DHHC-type zinc finger protein
<a href="#">MAL13P1.126</a>	Zinc finger	DHHC-type zinc finger protein
<a href="#">PFL0465c</a>	Zinc finger	krox
<a href="#">PFF0105w</a>	Zinc finger	MYND finger domain protein
<a href="#">PFF0350w</a>	Zinc finger	MYND finger protein
<a href="#">PF07_0124</a>	Zinc finger	MYND finger protein
<a href="#">PF13_0293</a>	Zinc finger	MYND finger protein
<a href="#">MAL6P1.142</a>	Hélice-Tour-Hélice	Myb
<a href="#">MAL6P1.277</a>	Hélice-Tour-Hélice	Myb
<a href="#">PF10_0327</a>	Hélice-Tour-Hélice	PfMyb2
<a href="#">PF11_0241</a>	Hélice-Tour-Hélice	Myb
<a href="#">PF13_0088</a>	Hélice-Tour-Hélice	PfMyb1
<a href="#">PFL0815w</a>	Hélice-Tour-Hélice	Myb
<a href="#">PFL1215c</a>	Hélice-Tour-Hélice	Myb
<a href="#">PFL0290w</a>	Facteur général	TFIIIB'
<a href="#">PF14_0333</a>	Mediator	MED10 Mediator component
<a href="#">PF14_0718</a>	Mediator	MED31 Mediator component
<a href="#">PF08_0037</a>	Mediator	MED7 Mediator component
<a href="#">PF11_0192</a>	Modification de la chromatine	moz acetyl transferase
<a href="#">PF11_0053</a>	Modification de la chromatine	PfSNF2L
<a href="#">PFF1440w</a>	Modification de la chromatine	putative SET-domain protein
<a href="#">PFD0190w</a>	Modification de la chromatine	SET
<a href="#">PFL0690c</a>	Modification de la chromatine	SET
<a href="#">PF13_0293</a>	Modification de la chromatine	SET
<a href="#">PF14_0489</a>	Modification de la chromatine	sir2
<a href="#">PF13_0152</a>	Modification de la chromatine	protein sir2 homologue, putative
<a href="#">PFB0730w</a>	Remodelage de la chromatine	brahma snf2/iswi
<a href="#">PFF1185w</a>	Remodelage de la chromatine	iswi protein homologue
<a href="#">PFF1140c</a>	Remodelage de la chromatine	PhD
<a href="#">PF10_0079</a>	Remodelage de la chromatine	PhD
<a href="#">PF11_0429</a>	Remodelage de la chromatine	PhD
<a href="#">PFL0575w</a>	Remodelage de la chromatine	PhD
<a href="#">PFL1010c</a>	Remodelage de la chromatine	PhD

<a href="#">PFL1210w</a>	Remodelage de la chromatine	PhD
<a href="#">PFL1905w</a>	Remodelage de la chromatine	PhD
<a href="#">MAL13P1.302</a>	Remodelage de la chromatine	PhD
<a href="#">PF14_0315</a>	Remodelage de la chromatine	PhD
<a href="#">PFF0225w</a>	Remodelage de la chromatine	SNF2
<a href="#">PF08_0048</a>	Remodelage de la chromatine	SNF2
<a href="#">MAL8P1.65</a>	Remodelage de la chromatine	SNF2
<a href="#">MAL8P1.73</a>	Remodelage de la chromatine	SNF2
<a href="#">PF10_0232</a>	Remodelage de la chromatine	SNF2
<a href="#">MAL13P1.216</a>	Remodelage de la chromatine	SNF2
<a href="#">PF13_0308</a>	Remodelage de la chromatine	SNF2

## I.2 – Protéines identifiées par Coulson et collaborateurs (Coulson et al. 2004)

PFC0805w	RNA POLYMERASE II::Rpb-1
PFB0715w	RNA POLYMERASE II::Rpb-2
PF11.130c	RNA POLYMERASE II::Rpb-3
PFB0245c	RNA POLYMERASE II::Rpb-4
PF13.0341	RNA POLYMERASE I/II/III::Rpb-5 [ABC27]
PFC0155c	RNA POLYMERASE I/II/III::Rpb-6 [ABC23]
PF10.0269	RNA POLYMERASE II::Rpb-7
PFL0665c	RNA POLYMERASE I/II/III::Rpb-8 [ABC14.5]
PFA0505c	RNA POLYMERASE II::Rpb-9
PF07.0027	RNA POLYMERASE I/II/III::Rpb-10 [ABC10 beta]
PF13.0023	RNA POLYMERASE II::Rpb-11
MAL13P1.213	RNA POLYMERASE I/II/III::Rpb-12 [ABC10 alpha]
PFE0465c	RNA POLYMERASE I::A190
PF11.0358	RNA POLYMERASE I::A135
PFD0360w	RNA POLYMERASE I::A12.2
PF11.0445	RNA POLYMERASE I/III::AC40
PF14.0150	RNA POLYMERASE I/III::AC19
PF13.0150	RNA POLYMERASE III::C160
PFL0330c	RNA POLYMERASE III::C128
PF14.0207	RNA POLYMERASE III::C34
PF11.0058	RNA POLYMERASE III::C25
PFB0290c	RNA POLYMERASE III::C11
PF11.0204	MITOCHONDRIAL RNA POLYMERASE
PF14.0695	APICOPLAST RNA POLYMERASE: alpha subunit
PF13.0040	APICOPLAST RNA POLYMERASE: alpha subunit
PF10.0367	APICOPLAST RNA POLYMERASE: beta, beta' & beta'' subunits
PF11.0315	APICOPLAST RNA POLYMERASE: beta, beta' & beta'' subunits
PFE0635c	APICOPLAST RNA POLYMERASE: beta, beta' & beta'' subunits
PF10330c	APICOPLAST RNA POLYMERASE: beta, beta' & beta'' subunits
PFA0525w	TFIIB: TFIIB
PFE0415w	TFIIB: TFIIB-like
PFE0305w	TFIID: TBP
PF14.0267	TFIID: TBP-like factor
MAL7P1.06	TFIIE: TFIIE-alpha
PF10.0369	TFIIN: XPB
PF1050w	TFIIN: XPD
PF12125c	TFIIN: p52
MAL13P1.76	TFIIN: p4
PF13.0279	TFIIN: p34
PFE0610c	TFIIN: MAT1 (PF00097)
PF13.0022	TFIIN: cycEa K
PF14.0469	TFIIB: 70 kDa subunit
PFL0520c	TFIIB: 63 kDa subunit
PF10.0293	Transcription-elongation: SPT4
MAL6P1.111	Transcription-elongation: SPT5
PF14.0059	Transcription-elongation: SPT6
PF07.0057	Transcription elongation factor TFIIS
PFE0870w	Transcription elongation factor CDC68
PFA0355w	CCR4-NOT complex: CCR4
PFE0980c	CCR4-NOT complex: CCR4
PFC0850c	CCR4-NOT complex: CCR4
MAL8P1.104	CCR4-NOT complex: CAF1
PF11.0049	CCR4-NOT complex: NOT1
PF14.0170	CCR4-NOT complex: NOT1
PF11.0297	CCR4-NOT complex: NOT2
PF10.0062	CCR4-NOT complex: NOT3/NOT5
PFL1705w	CCR4-NOT complex: NOT4
PF11.0477	CCAAT BOX-BINDING COMPLEX::CBF-A/NF-Y-B (PF00808)
PF13.0043	CCAAT BOX-BINDING COMPLEX::CBF-B/NF-Y-A
PF14.0374	CCAAT BOX-BINDING COMPLEX::CBF-C/NF-Y-C (PF00808)
PF08.0034	Histone acetyltransferase GCN5 (PF00439)
PF10.0143	Transcriptional adaptor ADA2 (PF00249, PF00549)
PF11.0053	Nucleosome remodelling: SNF2L
MAL6P1.183	Nucleosome remodelling: ISWI (PF00628)
PF11260c	Histone deacetylase RPD3
PF10.0078	Histone deacetylase
PF14.0690	Histone deacetylase
MAL7P1.37	SIN3-REPRESSING COMPLEX: p18
MAL6P1.269	Transcription co-activator ALY
PF11.0293	Transcription co-activator MBF1 (PF01381)
PFC0175w	Transcription co-activator TRIAD2
PF14.0489	Sir2-like protein
PF07.0083	TBP-binding protein
PF14.0393	Single-strand recognition protein
MAL6P1.142	DNA methyltransferase I-associated protein 1 DNMAP1
PFB0875c	Chromatin-binding protein (SKI/SNW family) SkiP
PF14.0241	General transcription factor BTF3
PFL2390c	Pleiotropic regulatory protein THO2
PF14.0718	RNA polymerase II-associated factor SOH1
PF14.0649	Transcription factor C1
MAL8P1.131	Transcription factor GAS41 (PF03366)
PF07.0123	Transcriptional regulator SPO8
PF13.0293	SET-domain transcriptional regulator
PFE0090w	Histone transcription regulator
PF11470c	Histone transcription regulator
PF14.0492	Calcium-dependent transcriptional repressor DREAM
MAL8P1.111	Transcriptional repressor RPH1
Minichromosome maintenance proteins : PF07.0023, PF13.0095, PF13.0291, PF14.0177, PFD0790c, PFE1345c, PFL0560c, PFL0580w	



## II - DONNEES SUPPLEMENTAIRES DE L'ARTICLE 1

### Comparison of 3D7 clone transcriptome data obtained from our experiments to Scripps/GNF Malaria Array data using 3D7 clone and De Risi *Plasmodium falciparum* HB3 time course microarray data.

Genes that fit (+) or do not fit (-) the data are listed. Genes whose expression patterns were not available are marked "none". "ND" means that the expression profile was discordant between sorbitol and temperature synchronised culture in Scripps/GNF. Malaria Array data. Average of three experiments for each clone is shown.

gene name	R	ET	LT	ES	LS	accession number	Le Roch <i>et al.</i>	Bozdeck <i>et al.</i>
CHR13 41763	0	-0,47	0,93	0,09	-0,69	MAL13P1.196	+	+
ZnSnf2	0	-0,9	0,68	-2,12	-0,5	MAL13P1.216	+	+
CHR13 40872	0	-0,46	-0,7	-0,97	0,97	MAL13P1.278	+	+
PHd2	0	-1,01	-1,77	-1,87	-0,27	MAL13P1.76	+	+
CHR13 40983	0	0,27	0,36	-0,06	0,24	MAL13P1.78	+	+
PF CDPK2	0	-0,13	0,48	0,14	-0,1	MAL6P1.108	+	-
PF PK2	0	0,95	0,16	-0,44	0,04	MAL6P1.108	+	-
PF PK4	0	-0,34	-0,07	-0,02	-0,3	MAL6P1.146	+	+
CHRBLOB 36958	0	0,05	-1,9	-2,81	1,11	MAL6P1.191	+	+
cdc48	0	0,79	0,41	1,56	1,31	MAL6P1.232	+	+
HISTONE H3	0	0,48	0,34	0,83	1,44	MAL6P1.248	none	none
HISTONE H2A	0	-0,38	0,28	1,07	1,86	MAL6P1.249	none	+
CHRBLOB 68156	0	0,32	1,16	0,59	-0,39	MAL6P1.271	+	+
PF FEST	0	0,14	0,83	0,24	0,29	MAL7P1.91	+	+
PF BO520W	0	1,25	0,35	0,21	0,14	PF BO520W	-	-
PF BO665W	0	0,24	0,21	0,56	1,29	PF BO665W	+	+
PF BO815W	0	-0,24	-0,66	0,8	1,75	PF BO815W	+	+
PF STARP	0	-0,16	-0,29	-1,2	-2,65	PF07_0006	+	+
PF CDC ATP	0	-0,65	0,46	-0,35	-1,02	PF07_0047	+	+
C2143	0	-0,12	0,67	1,86	-2,81	PF07_0073	+	+
CHRBLOB BU60D08	0	-0,12	-0,84	-1,22	-0,19	PF08_0044	nd	+
PF MYB3	0	0,01	-0,26	0,83	-0,16	PF10_0143	+	+
GBP130	0	0,08	1,33	0,45	-0,12	PF10_0159	+	+
PF PUGBP130	0	-0,61	0,98	-0,13	-0,92	PF10_0159	+	+
ADP-RIBO FAC1	0	-0,14	0,8	-0,16	0,1	PF10_0203	-	-
PFMYB2	0	-0,52	-1,21	-0,73	-0,53	PF10_0327	+	+
HISTONE H4	0	0,04	0,63	1,43	2,21	PF11_0061	+	-
HISTONE H2B	0	1,03	1,09	2,08	2,44	PF11_0062	+	-
CHR11 S189	0	0,33	0,38	0,37	0,23	PF11_0096	+	+
PF ERB	0	-0,77	0,35	0,93	0,85	PF11_0098	+	+
PF ERC	0	0,67	0,63	1,09	0,78	PF11_0098	+	+
PF S40	0	0,36	0,35	1,1	0,82	PF11_0098	+	+
PF MAP2	0	-1,27	-1,87	-1,94	-1,77	PF11_0147	nd	+
PF RAN-1	0	-0,15	0,09	-0,15	-0,09	PF11_0183	+	+
CHR11 S51	0	-0,37	0,7	1,08	1,2	PF11_0220	+	+
PF CK1	0	-0,37	-1,59	0,4	0,42	PF11_0377	+	+
CHR11 S127	0	-0,96	-1,5	-0,13	1,45	PF11_0464	+	+
CYCLIN ANIA	0	-0,65	-0,24	-1,86	-1,01	PF13_0022	none	-
PFMYB1	0	-0,27	-0,69	-1,09	-0,55	PF13_0088	-	-
RNA POLBP	0	-0,17	-2,13	-3,3	-0,01	PF13_0123	none	none
RNA POLIII	0	-0,08	-0,46	-0,45	-0,35	PF13_0150	+	+
PF PK6	0	-0,34	1,43	2,08	-1,02	PF13_0206	+	none
CHR13 32969	0	0,59	-0,99	2,11	1,11	PF13_0211	+	+

CHR13 21780	0	-0,16	-0,26	-0,03	-0,1	PF13_0258	-	+
PF PCNA	0	-0,6	1,06	2,22	2,04	PF13_0328	+	+
PF RNR2	0	1,03	1,22	2,03	1,48	PF14_0053	+	+
PF ACT11	0	-0,31	-0,39	-0,67	-0,91	PF14_0124	+	+
PP1-LIKE	0	-0,53	0,37	-0,19	-0,32	PF14_0224	-	-
D11CY3	0	-0,52	-0,96	-1	0,28	PF14_0316	+	+
PF TOPO 2	0	0,13	-0,22	-0,72	-0,01	PF14_0316	+	+
CALMODULIN	0	0,78	1,07	1,42	1,58	PF14_0323	+	+
CHR14 132	0	0,12	-0,38	-1,25	-0,89	PF14_0346	-	-
PF RNR1	0	0,58	0,14	1,08	1,31	PF14_0352	+	+
DNA PRIM	0	-0,3	0,42	0,37	-0,27	PF14_0366	+	+
Constans	0	-0,2	0,02	1,2	-0,12	PF14_0383	nd	+
CHR9 9716	0	0,3	0,47	0,52	0,35	PF14_0392	+	+
CHR14 128 B	0	0,32	-0,25	-1	-0,55	PF14_0423	+	+
CALCINEURIN	0	-0,11	0,34	-0,03	-0,51	PF14_0492	-	-
PF CYC-1	0	0,06	-0,22	-1,04	0,28	PF14_0605	-	+
KHARP	0	0,37	-1,34	-1,35	-2	PFB0100c	+	+
RPB4-LIKE	0	0,47	-0,26	-1	0,29	PFB0245c	+	+
PF RAD2	0	-0,57	0,28	-1,12	-2,01	PFB0265c	+	-
TFIIS	0	0,23	-0,23	-1,53	-1,55	PFB0290c	+	+
YOU2	0	1,32	-0,08	0,04	0,22	PFB0620w	+	+
PF MBP	0	-0,59	-0,1	1,29	0,02	PFB0725c	+	none
DNA HELICASE	0	-1,17	-0,97	-0,24	-2,12	PFB0730w	+	-
PF CPK	0	0,05	-0,5	-0,21	2,61	PFB0815w	+	+
SPOU	0	0,06	0,96	-0,29	-0,23	PFB0855c	+	+
RNA HELICASE	0	1,14	-1,38	-1,58	-1,41	PFB0860c	+	+
PF SNRNP	0	0,14	0,67	0,05	0,18	PFB0865w	+	-
RSUA	0	0,13	1,14	0,51	-0,79	PFB0890c	+	+
PFC0060C	0	-0,55	-1,68	-0,95	0,02	PFC0060c	+	+
RNA POL1	0	0,37	0,34	-1,88	-1,35	PFC0155c	+	+
PFC0195W	0	0,01	0,67	-0,13	-0,62	PFC0195w	+	+
EIF-4A	0	-0,04	0,69	-0,51	-0,61	PFC0445w	nd	none
PFC0485W	0	0,07	-0,57	-2,31	-0,82	PFC0485w	nd	+
PFC0510W	0	1,67	1,86	3,41	2,05	PFC0510w	+	+
PFC0525C	0	-0,17	-0,62	0,68	-0,51	PFC0525c	+	+
PFC0635C	0	0,1	-0,68	-0,53	0,39	PFC0635c	+	+
PFC0740C	0	-0,48	-0,54	-0,83	-0,04	PFC0740c	+	+
PFC0755C	0	-0,6	0,01	1,34	1,22	PFC0755C	+	+
PFC0805W	0	1,03	0,46	0,21	0,29	PFC0805w	+	+
PFC0825C	0	0,62	0,91	0,29	0,43	PFC0825C	nd	-
PFC0915W	0	-0,2	0,88	0,75	-0,84	PFC0915w	+	+
PFC0920W	0	-0,41	0,92	0,17	-0,51	PFC0920w	+	+
HOX 1	0	-0,07	-0,32	0,4	0,46	PFD0340c	+	+
PF MT C	0	-0,68	-2,12	-1,5	0,38	PFD0370w	-	-
PF POLA	0	-0,83	0,71	1,7	2,56	PFD0590c	+	+
PF CRK2	0	-0,31	-0,31	-0,6	-1,1	PFD0740w	+	+
SAR1	0	-0,71	0,42	0,52	0,27	PFD0810w	+	+
PF CRK1	0	-0,48	-0,85	0,74	-1,06	PFD0865c	+	+
HOX 2	0	0,36	-0,46	0,24	0,97	PFD1160w	nd	-
PF ATBP	0	-0,85	-0,85	-3,74	-1,31	PFE0305w	nd	+
CHR5 21208	0	-0,4	0,61	0,78	-0,14	PFE0420c	+	+
PF TOPO 1	0	-0,5	-0,82	0,79	1,04	PFE0520c	-	-
PFCYC-3 (TRYP-2)	0	-0,01	-1,22	-3,53	-1,33	PFE0920c	nd	-
EF-1B	0	0,27	0,68	-0,11	-0,5	PFI0645w	+	+

PF B7	0	0,69	0,44	0,49	0,15	PFI0930c	-	none
3-PHOSPHO	0	0,84	1,85	1	-1,31	PFI1105w	+	+
PF PHOS	0	0,2	1,46	0,03	-0,91	PFI1245c	+	+
CHR9 IX78E09,P1C	0	0,83	-0,96	-0,42	-1,47	PFI1280c	+	+
PF PKA	0	0,82	0,41	2,08	2,84	PFI1685w	+	+
CHR11 N3190	0	0,47	0,45	1,03	0,64	PFL0080c	nd	+
PF HMG1	0	-0,24	0,15	0,5	0,4	PFL0145c	+	+
PF HMG2	0	-0,3	-0,17	0,11	0,15	PFL0290w	+	+
KROX	0	0,09	0,36	0,85	0,98	PFL0465c	+	none
NF-LIK1	0	-0,58	-0,18	-2,09	-1	PFL1030w	-	-
CHR12 06	0	-0,43	-1,47	-2,12	-1,17	PFL1285c	-	-
PFCYC-2	0	-0,56	-0,88	0,23	0,1	PFL1330c	+	+
PF NEK-1	0	0,03	0,9	0,9	0,47	PFL1370w	+	+
phdB	0	0,79	0,51	-0,85	-0,67	PFL1905w	+	none
FSTH	0	0,18	1,13	1,69	1,48	PFL1925w	+	+
PF TBP	0	0,14	0,97	0,98	0,84	PFL2345c	+	+

### III – DONNEES SUPPLEMENTAIRES DE L'ARTICLE 4

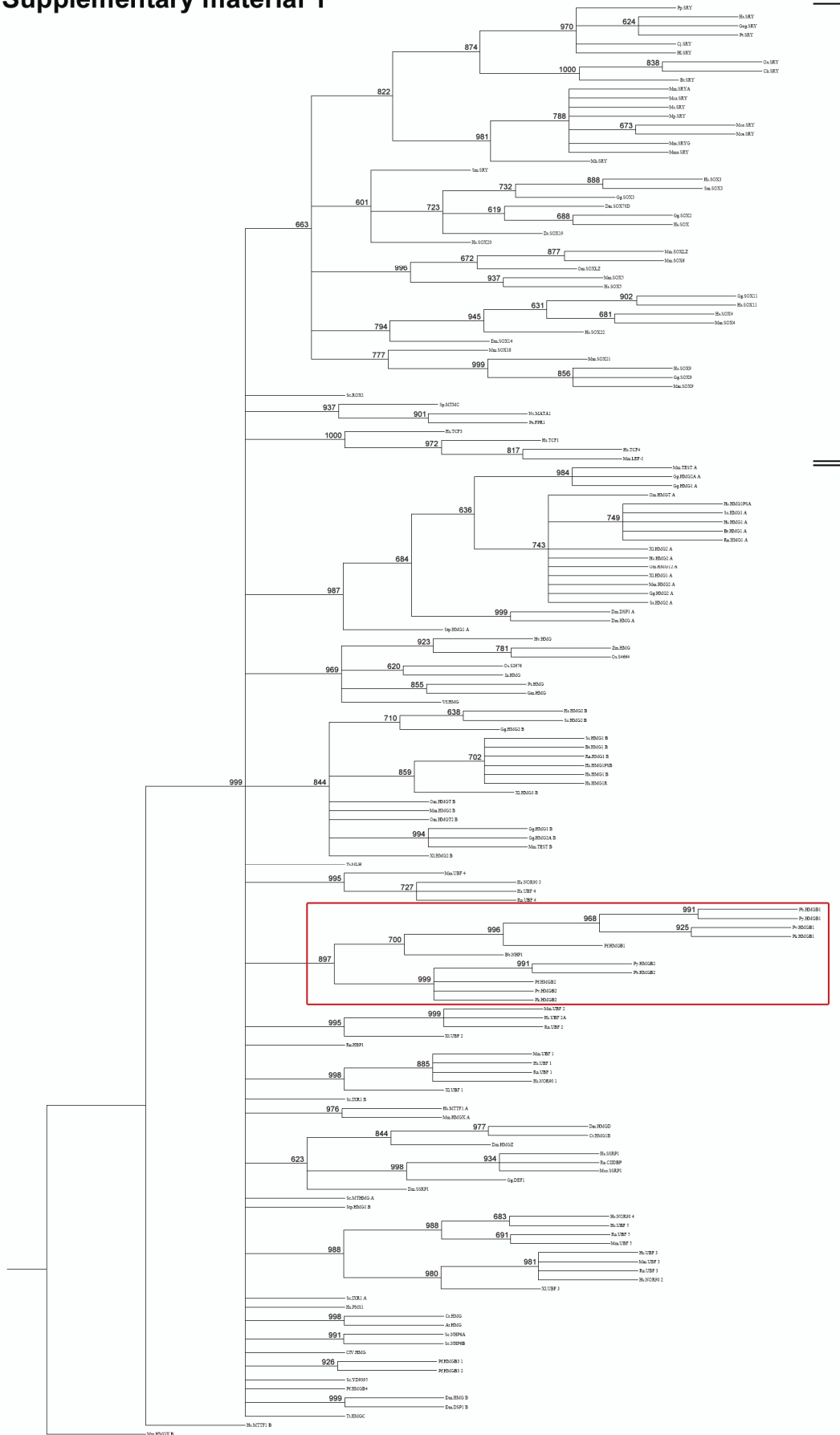
#### III.1 - Donnée supplémentaire 1. Arbre phylogénétique de 159 domaines "HMG-box".

Les séquences des domaines 159 "HMG-box" ont été alignés avec le programme ClustalW (Thompson et al. 1994) et l'arbre phylogénétique a été construit avec le programme PHYLIP (Felsenstein 1989). Il se décompose en deux catégories : les facteurs de transcription SOX/SRY/MATA/TCF et les facteurs architecturaux HMGB/UBF. Sont encadrés en rouge les domaines "HMG-box" des protéines PfHMGB1 et PfHMGB2, ceux de leurs orthologues chez *P. yoelii*, *P. berghei*, *P. vivax* et *P. knowlesi* ainsi que celui de la protéine NHP1 de *Babesia bovis*, un parasite apicomplexe dont le génome est riche en A+T comme celui de *Plasmodium*.

Pour agrandir ce document :

Site <http://ec.asm.org/cgi/content/full/5/4/672/DC1>

# Supplementary material 1



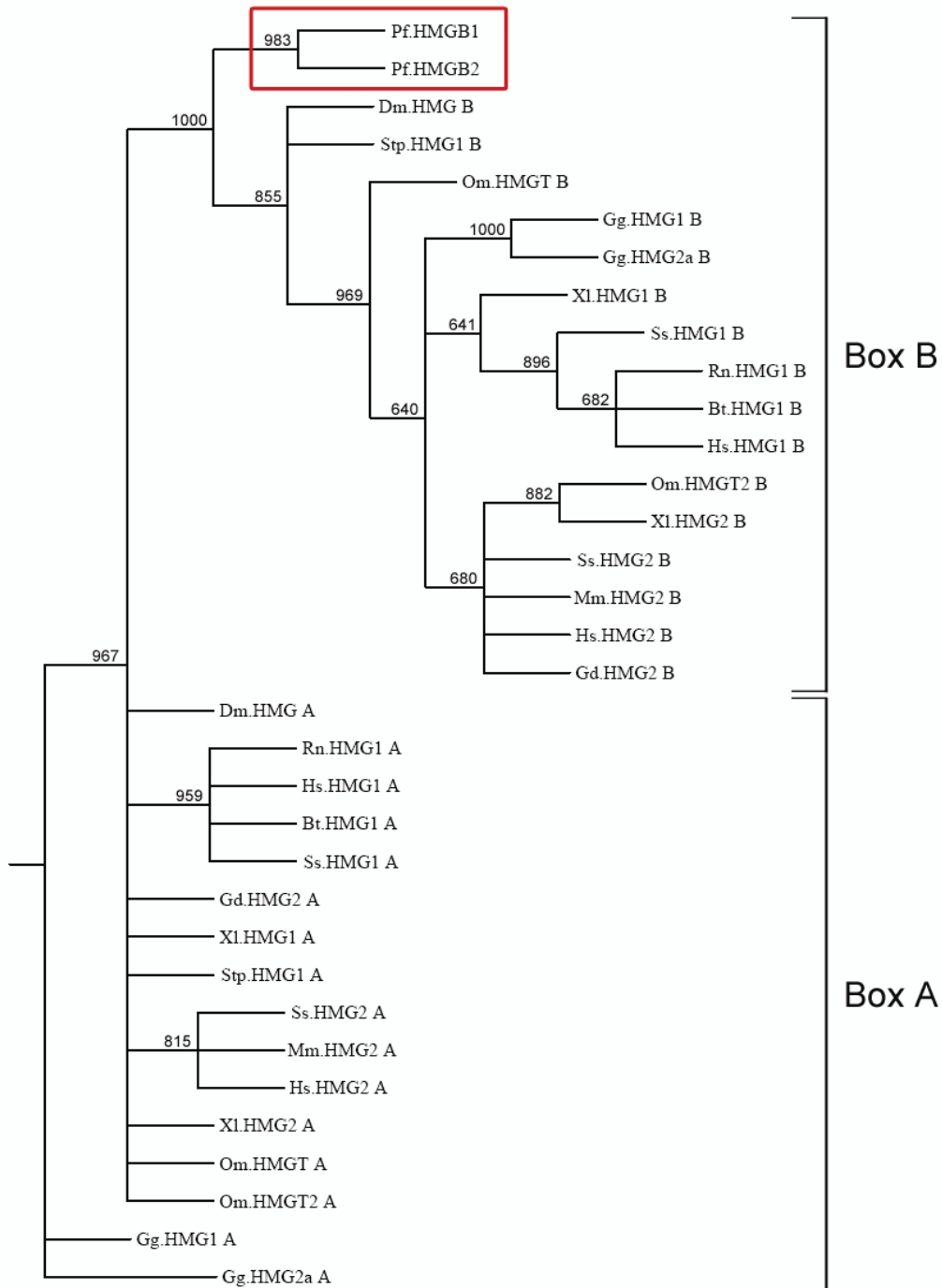
SOX/SRY/MATA/TCF

HMGB/UBF

### III.1 - Donnée supplémentaire 2. Arbre phylogénétique de 34 boîtes A et boîtes B.

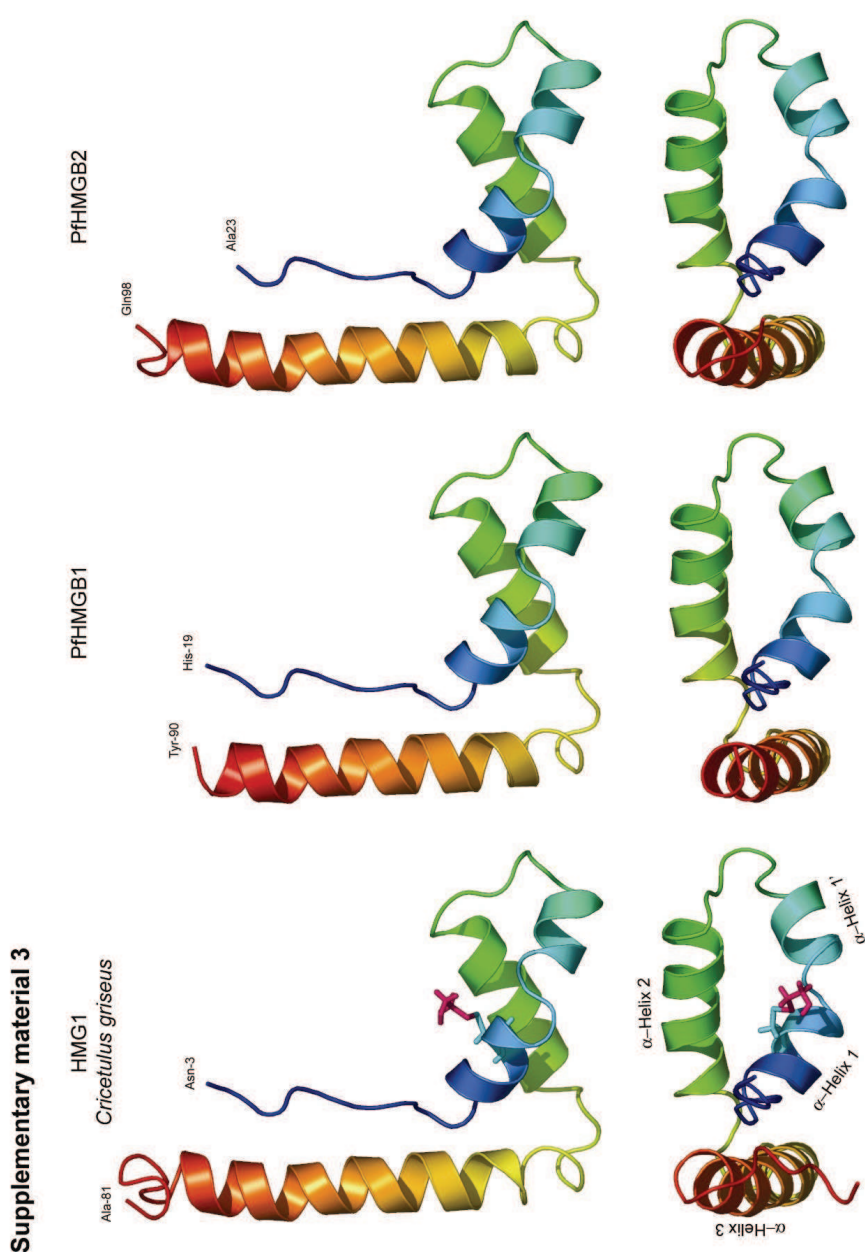
L'arbre construit comme précédemment est composé des boîtes A et B des protéines HMGB de divers métazoaires et des "HMG-box" des protéines parasitaires. Il se divise en deux groupes : un premier avec les boîtes B et les "HMG-box" de PfHMGB1 et PfHMGB2, un second regroupant toutes les boîtes A.

#### Supplementary material 2



### III.3 - Donnée supplémentaire 3. Modèles de structures des protéines PfHMGB1 et PfHMGB2 obtenus par homologie avec la protéine HMG1 du hamster chinois *Cricetulus griseus*.

La méthode @utomatic Threading Optimisation Modelling & Evaluation a été utilisée pour la prédiction de structure (Douguet and Labesse 2001). Chaque structure, modèles de PfHMGB1 et PfHMGB2 et support de HMG1 de *C. griseus* (Read et al. 1993), est visualisée sous deux angles différents. Les trois structures sont colorées selon le spectre de la lumière, le bleu désignant la partie N-terminale et le rouge la partie C-terminale de la protéine. En rose figure la molécule de  $\beta$ ME, initialement fixée à la cystéine (en cyan) située dans la première hélice de la structure support du hamster chinois.



## LISTE DES PUBLICATIONS

Galio, L., Dias Pereira, S. & Vaquero, C. (1996). Fine screening of regulatory transcription factors. *BIAjournal* **3**, 30-31.

Galio, L., Briquet, S., Cot, S., Guillet, J.G. & Vaquero, C. (1997). Analysis of interactions between huGATA-3 transcription factor and three GATA regulatory elements of HIV-1 Long Terminal Repeat, by Surface Plasmon Resonance. *Analytical Biochemistry* **253**, 70-77.

Galio, L., Briquet, S. & Vaquero, C. (1999). Real-time study of interactions between a composite DNA regulatory region (HIV-1 LTR NRE) and several transcription factors of nuclear extracts. *BBRC* **264**, 6-13

Briquet, S., Richardson, J., Vanhée-Brossollet, C. & Vaquero, C (2001). Natural antisense transcripts are detected in different cell lines and tissues of cats infected with Feline Immunodeficiency Virus. *Gene* **267**, 157-164.

Briquet, S. & Vaquero, C (2002). Immunolocalization studies of an antisense protein in HIV-1 infected cells and viral particles. *Virology* **292**, 177-184.

Boschet, C. J.\*, Gissot, M.\*, Briquet, S., Hamid, Z., Claudel-Renard, C. & Vaquero, C (2004). Characterization of PfMyb1 transcription factor during erythrocytic development of 3D7 and F12 *Plasmodium falciparum* clones. *Mol Biochem Parasitol* **138** 159-163.

Gissot, M., Refour, P., Briquet, S., Boschet, C., Coupé, S., Mazier, D. & Vaquero C (2004). Transcriptome of 3D7 and its gametocyte-less derivative F12 *Plasmodium falciparum* clones during erythrocytic development using a gene specific microarray assigned to gene regulation, cell cycle and transcription factors. *Gene* **341** 267-277.

Gissot, M., Briquet, S., Refour, P., Boschet, C. & Vaquero C (2005). PfMyb1, a *Plasmodium falciparum* transcription factor, is required for intra-erythrocytic growth and controls key genes for cell cycle regulation. *J. Mol. Biol.* **346** 29-42.

Briquet, S., Boschet, C., Gissot, M., Tissandié, E., Sevilla, E., Franetich, J. F., Thiery, I., Hamid, Z., Bourgouin, C. & Vaquero, C (2006). High-Mobility-Group Box nuclear factors of *Plasmodium falciparum*. *Eukaryotic cell* **5** 672-682.