



**HAL**  
open science

# Etude des mécanismes épigénétiques impliqués dans la kystogénèse chez le pathogène humain *Toxoplasma gondii*

Nehmé Saksouk

► **To cite this version:**

Nehmé Saksouk. Etude des mécanismes épigénétiques impliqués dans la kystogénèse chez le pathogène humain *Toxoplasma gondii*. Biochimie [q-bio.BM]. Université Joseph-Fourier - Grenoble I, 2005. Français. NNT: . tel-00011144

**HAL Id: tel-00011144**

**<https://theses.hal.science/tel-00011144>**

Submitted on 20 Dec 2005

**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.

Université Joseph-Fourier – Grenoble I

**THESE**

Pour obtenir le grade de

**DOCTEUR DE L'UNIVERSITE JOSEPH-FOURIER**

Discipline : BIOLOGIE

Présentée et soutenue publiquement par

**NEHME SAKSOUK**

Le 16 Décembre 2005

**Etude des mécanismes épigénétiques  
impliqués dans la kystogénèse  
chez le pathogène humain  
*Toxoplasma gondii.***

**Composition du jury**

Dr. HAKIMI Mohamed-Ali  
Dr. CESBRON-DELAUW Marie-France  
Dr. MENARD Robert  
Dr. LAGRANGE Thierry  
Dr. GAGNON Jean

Directeur de thèse  
Directeur de thèse  
Rapporteur  
Rapporteur  
Président

## REMERCIEMENTS

Je tiens avant tout à remercier les lecteurs de ce manuscrit pour les corrections et les critiques ; Dr. Thierry Lagrange, Robert Ménard et Jean Gagnon pour l'honneur qu'ils m'ont fait en acceptant cette tâche laborieuse qu'est l'évaluation d'un travail de thèse.

Je remercie la directrice Marie-France CESBRON-DELAUW de m'avoir permis d'effectuer ma thèse dans le Laboratoire Adaptation et Pathogénie des Microorganismes.

Je remercie Mohamed Ali-Hakimi, pour m'avoir accueilli dans son équipe ATIP « Epigénétique et Parasites », pour son encadrement et sa disponibilité. D'avoir cru avec tant d'enthousiasme aux différents projets qui ont jalonné mon passage dans son équipe. Merci également pour sa confiance et la grande liberté de travail qu'il m'a accordé. Je remercie aussi Anne-Lise et des gros bisous aux deux petits mignons et jolies enfants Rani et Inaya.

Je remercie le centre National de Recherche Scientifique Libanais (CNRSL) de leur aide financière durant les trois années de thèse effectué dans le laboratoire LAPM.

Je dois un grand "merci" à tous ceux qui, par leurs diverses compétences, leur savoir-faire, leur disponibilité ou simplement leur amitié, m'ont rendu la vie plus facile au sein du laboratoire LAPM. Je tiens en particulier à remercier Karine Musset (qui m'a tout appris sur l'art et la manière de manipuler les petits parasites), Corinne Mercier pour ses conseils et encouragement, Laurence Braun pour son encouragement et nos discussions scientifiques autour d'un café, Laetitia Travier, Claire Gendrin et Céline Sautel pour leur amitié et les belles soirées passées ensemble.

Il est légitime de remercier les membres de ma famille et mes proches pour leur patience et leur réconfort. En plus d'un grand "merci" collectif à tous les membres de ma famille, je tiens à ajouter une pensée particulière à ma petite maman et mon papa (pour toutes ses petites attentions qui remontent le moral), à mes sœurs Hanane et Kady et à mes frères Ali, Mohamed, Ibrahim et Hassan.

Merci aussi à mes indispensables amis : Mohamed Roumani, Abbas Hodroj, Rani Maké, Ali Chokr, Wassim Rammal, Ayman Mourad et Kamal Baydun (pour nos 4 années de belle amitié) et tout le troupeau de l'ARPEJ.

Je finis par une petite pensée à ma belle et bien aimé Julie Monnet Saksouk, ma compagne de galères, avec qui se fut un plaisir de partager des discussions, des moments d'enthousiasme et d'angoisse pendant les derniers moments de cette thèse et avec qui ma vie s'enjolivera de pleins de bonheurs.

En espérant n'avoir oublié personne..... Merci à tous !

# AVANT-PROPOS

Arrivé en France en 2001, j'ai débuté mon Master 2 (anciennement DEA) de Biologie Structurale et Fonctionnelle (BSF) à l'université Joseph Fourier de Grenoble, puis intégré l'équipe du Dr. Roberto GEREMIA au sein du laboratoire LAPM. Mon sujet fut ciblé sur la caractérisation fonctionnelle et structurale des enzymes impliquées dans la biosynthèse de la paroi capsulaire de *Streptococcus Pneumoniae* de type 8. Ces recherches, qui furent aussi l'objet de ma première année de thèse, ont fait l'objet de deux publications :

- **Saksouk N**, Pelosi L, Colin-Morel P, Boumedienne M, Abdian PL, Geremia RA. The capsular polysaccharide biosynthesis of *Streptococcus pneumoniae* serotype 8: functional identification of the glycosyltransferase WciS (Cap8H). *Biochem J.* 2005 Mar 15.
- Pelosi L, Boumedienne M, **Saksouk N**, Geiselman J, Geremia RA. The glucosyl-1-phosphate transferase WchA (Cap8E) primes the capsular polysaccharide repeat unit biosynthesis of *Streptococcus pneumoniae* serotype 8. *Biochem Biophys Res Commun.* 2005 Feb 18;327(3):857-65.

Suite au départ du Dr. Roberto GEREMIA de l'unité en 2004, j'ai pu poursuivre mes études doctorales au sein de l'équipe ATIP/CNRS dirigée par le Dr. Mohamed-Ali HAKIMI.

## RESUME

Le parasite intracellulaire *Toxoplasma gondii* est l'agent pathogène de la toxoplasmose. Cette maladie est gravissime pour le fœtus et pour l'individu immunodéprimé. L'interconversion du parasite de la forme tachyzoïte virulente à la forme bradyzoïte quiescente est au centre de la pathogénèse de cette infection. Ce processus engage une régulation coordonnée des gènes du parasite qui se traduit par une cascade d'évènements moléculaires au niveau de l'ADN. Des études suggèrent un contrôle transcriptionnel de l'interconversion avec l'expression exclusive de certains gènes dans une forme donnée. Cependant, ce parasite et son phylum se distinguent des autres eucaryotes par une quasi-pénurie des facteurs spécifiques de transcription. Nous avons émis l'hypothèse que le niveau d'expression des gènes du *Toxoplasme* est étroitement régulé par la structure physique et la nature chimique de la chromatine. Ce manuscrit illustre l'influence majeure du « code histone » sur la différenciation parasitaire. Nous avons identifié plusieurs enzymes en charge de l'écriture de ce code. L'exemple le plus frappant est la découverte d'une méthyltransférase TgCARM1 qui méthyle l'arginine 17 de l'histone H3, une marque activatrice de la transcription. Nous avons également identifié le premier complexe co-repressor du *Toxoplasme* (TgCRC). TgCRC en opposition avec l'acétylase TgGCN5 régule en partie la balance acétylation/déacétylation, qui en retour influe sur la différenciation parasitaire. L'ensemble de nos résultats converge vers l'idée de l'existence d'un « code histone parasitaire » hautement sophistiqué, qui a co-évolué avec celui de la cellule hôte parasitée.

## ABSTRACT

The intracellular parasite *Toxoplasma gondii* is the pathogen responsible for toxoplasmosis. This opportunistic parasite causes cerebral toxoplasmosis in immunodepressed individuals and the fetus. Interconversion from the virulent tachyzoïte form of the parasite to its quiescent bradyzoïte form is a critical point in the pathogenesis. This process involves a coordinated regulation of the expression of genes in the parasite, which results in a cascade of molecular events at the level of DNA. Studies suggest a transcriptional control of interconversion, including a stage specific expression of genes. However, this parasite and the whole phylum Apicomplexa are characterised by a quasi-shortage of specific transcription factors. We hypothesised that the level of expression of genes in *toxoplasma* is narrowly regulated by physical structure and chemical nature of chromatin. This manuscript illustrates the major influence of the "histone code" on parasite differentiation. We identified enzymes which are responsible for editing this code. The most striking example is the discovery of a methyltransferase TgCARM1 which methylates the arginine-17 residue of the histone H3, which is a marker of transcription activation. We also identified the first co-repressor complex of *Toxoplasma* (TgCRC). TgCRC partly controls the acetylation/deacetylation balance, in opposition to the acetylase TgGCN5, which influences parasitic differentiation. Our results leads mainly to the idea of the existence of a highly sophisticated "parasitic histone code", which has co-evoluted with that of the parasitized host cell.

# SOMMAIRE

## SOMMAIRE

### AVANT PROPOS :

**ARTICLE 1:** The capsular polysaccharide biosynthesis of *Streptococcus pneumoniae* serotype 8: functional identification of the glycosyltransferase WciS (Cap8H). **Biochem J. 2005 Mar 15**.....109

**ARTICLE 2:** The glucosyl-1-phosphate transferase WchA (Cap8E) primes the capsular polysaccharide repeat unit biosynthesis of *Streptococcus pneumoniae* serotype 8. **Biochem Biophys Res Commun. 2005 Feb 18;327(3):857-65**.....119

## I. INTRODUCTION GENERALE

### 1. TOXOPLASMA GONDII, DE LA TOXOPLASMOSE AU MODELE D'ETUDE...1

1.1. LA TOXOPLASMOSE.....1

1.1.1. LA TOXOPLASMOSE ACQUISE DU SUJET IMMUNOCOMPETENT.....1

1.1.2. LA TOXOPLASMOSE CONGENITALE.....2

1.1.3. LA TOXOPLASMOSE DU SUJET IMMUNODEPRIME.....2

1.2. TOXOPLASMA GONDII.....4

1.2.1. LES DIFFERENTES SOUCHES GENOTYPIQUES.....4

1.2.2. SON CYCLE BIOLOGIQUE.....4

1.2.3. L'ULTRASTRUCTURE GENERALE DU TOXOPLASME.....7

1.2.4. LE PROCESSUS D'INTERCONVERSION, AU CENTRE DE LA PATHOGENESE.....7

### 2. CHROMATINE ET REGULATION GENIQUE CHEZ LES EUCARYOTES.....12

2.1. REMODELAGE DE LA CHROMATINE ATP-DEPENDANT.....14

2.2. LES MODIFICATIONS COVALENTES DES QUEUES D'HISTONES.....16

2.2.1. ACETYLATION DES HISTONES ET TRANSCRIPTION.....17

2.2.1.a. LES HISTONES ACETYLTRANSFERASES .....19

2.2.1.b. LES HISTONES DEACETYLASES.....21

2.2.2. METHYLATION DES HISTONES ET TRANSCRIPTION.....23

2.2.2.a. LES HISTONE-LYSINE-METHYLTRANSFERASES (HMT/SET).....24

2.2.2.b. LES HISTONE-ARGININE-METHYLTRANSFERASES (PRMT).....26

2.3. LE CODE HISTONE.....26

## II. RESULTATS

<b>CHAPITRE I :</b> .....	<b>30</b>
AVANT PROPOS.....	30
1. PREMIERE LECTURE DU CODE HISTONE.....	30
2. ANALYSE DES HISTONES DEACETYLASES CHEZ <i>T. GONDII</i> .....	32
3. ANALYSE DES HISTONES ARGININES METHYLTRANSFERASES DU TOXOPLASME.....	32
4. ARTICLE 3: HISTONE MODIFYING COMPLEXES REGULATE GENE EXPRESSION PERTINENT TO THE DIFFERENTIATION OF PROTOZOAN PARASITE TOXOPLASMA GONDII.....	34
5. DONNES SUPPLEMENTAIRES.....	48
<b>CHAPITRE II :</b> .....	<b>53</b>
1. ETUDE MOLECULAIRE DE L'HISTONE DEACETYLASE TgHDAC2.....	53
2. DETERMINATION DE L'INTERACTOME SPECIFIQUE DE TgHDAC2.....	53
3. ETUDE DE LA STRUCTURE MODULAIRE DE LA PROTEINE TgHDAC2.....	59

## III. DISCUSSION.....63

AVANT-PROPOS.....	63
1. L'HYPOTHESE EPIGENETIQUE.....	63
2. ÉVOLUTION DU CODE HISTONE CHEZ <i>T. GONDII</i> .....	64
3. LECTURE DU CODE PAR TgCARM1.....	66
4. BRISER LE CODE & INTERFERER AVEC LA DIFFERENTIATION PARASITAIRE.....	67
5. LA DEACETYLATION AU CŒUR DE LA REPRESSION GENIQUE.....	68
6. EN GUISE DE CONCLUSION.....	70

## IV. MATERIEL ET METHODE

<b>1. MATERIEL.....</b>	<b>72</b>
1.1. VECTEURS DE CLONAGE.....	72
1.2. ENZYMES UTILISEES.....	72
1.3. PREPARATION D'ADN PLASMIDIQUE ET D'ARN TOTAUX DE TACHYZOÏTES.....	72
1.4. LES AMORCES UTILISEES.....	72
1.5. LES PLASMIDES RECOMBINANTS .....	72
1.6. SOUCHES BACTERIENNES.....	76
1.7. LES ANTICORPS .....	76
<b>2. METHODE DE BIOLOGIE MOLECULAIRE.....</b>	<b>77</b>
2.1. CONDITIONS DE CULTURE ET DE STOCKAGE DES BACTERIES.....	77
2.2. PREPARATION DE E.COLI COMPETENTES.....	77
2.3. PCR (POLYMERASE CHAIN REACTION).....	78
2.4. CLONAGE DANS LE VECTEUR TOPO.....	78
2.5. METHODE BLANC-BLEU.....	78
2.6. MINI PREPARATION DE PLASMIDE PAR LYSE ALCALINE DE BACTERIES.....	79
2.7. MIDI PREPARATION DE PLASMIDES.....	79
2.8. SEQUENÇAGE DE L'ADN.....	80
2.9. DIGESTION TOTALE DE L'ADN PLASMIDIQUE.....	80
2.10. SEPARATION DES FRAGMENTS D'ADN SUR GEL.....	80
2.11. PURIFICATION D'UN FRAGMENT D'ADN PAR ELUTION DU GEL D'AGAROSE.....	81
2.12. LIGATION.....	81
2.13. TRANSFORMATION DE LA LIGATION.....	81

### **3. METHODES DE BIOLOGIE CELLULAIRE.....81**

3.1. SUREXPRESSION DE PROTEINES RECOMBINANTES CHEZ E.COLI.....	82
3.2. CULTURE IN VITRO DES PARASITES.....	82
3.3. OBTENTION DE PARASITES RECOMBINANTS STABLES.....	82
3.4. TEST DES CLONES PAR IMMUNOFLUORESCENCE INDIRECTE (IFI).....	83

### **4. METHODES DE BIOCHIMIE.....83**

4.1. ANALYSE D'EXTRAITS PROTEIQUES.....	83
4.2. TRANSFERT PAR WESTERN BLOTS.....	84
4.3. IMMUNOREVELATION DES PROTEINES.....	84
4.4. DOSAGE DES PROTEINES PAR LA METHODE DE BRADFORD.....	84
4.5. PRODUCTION ET PURIFICATION DE LA PROTEINE TAGUEE HISTIDINE ET GST EXPRIMEE DANS LES BACTERIES.....	85
4.5.1. SUREXPRESSION DE LA PROTEINE DANS LES BACTERIES.....	85
4.5.2. PURIFICATION DE PROTEINE RECOMBINANTE AVEC UNE ETIQUETTE 6-HIS SUR COLONNE NICKEL.....	85
4.5.3. PURIFICATION DE PROTEINE RECOMBINANTE AVEC UNE ETIQUETTE GST SUR COLONNE GSH-SEPHAROSE.....	85
4.5.4. PURIFICATION DE LA PROTEINE TAGUEE HA-FLAG EXPRIMEE DANS T.GONDII.....	86
4.6. PURIFICATION DES ANTICORPS SUR COLONNE D'AFFINITE.....	86
4.6.1. LIAISON DES ANTIGENES SUR COLONNE D'AFFINITE.....	86
4.6.2. LIAISON DES ANTICORPS SUR COLONNE D'AFFINITE.....	87
4.7. IMMUNOPRECIPITATION DE PROTEINES PAR DES ANTICORPS IMMUNOPURIFIES.....	87
4.8. EXTRACTION DES CORES HISTONES.....	88
4.9. TEST D'ACTIVITE DES PROTEINES HISTONES-METHYLTRANSFERASE (HMT).....	88
4.10. TEST D'ACTIVITE DES HISTONES-ACETYLLTRANSFERASE (HAT).....	88
4.11. TEST D'ACTIVITE DES HISTONES DEACETYLASE (HDAC).....	88

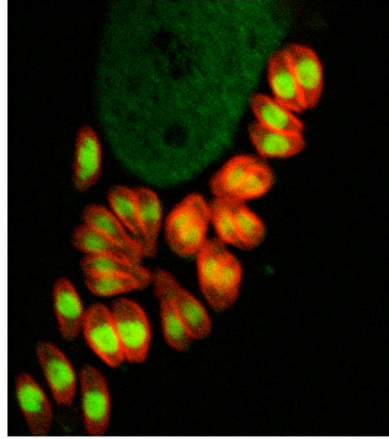
### **V. BIBLIOGRAPHIE.....89**

### **LISTES DES ABREVIATIONS UTILISEES.....105**

### **REMERCIEMENT**



# INTRODUCTION



# 1. TOXOPLASMA GONDII, DE LA TOXOPLASMOSE AU MODELE D'ETUDE

## 1.1 LA TOXOPLASMOSE

*Toxoplasma gondii* est un parasite protozoaire appartenant au phylum des *Apicomplexa*. Ces proches cousins sont *Eimeria*, *Neospora* et *Sarcocystis*, pathogènes des animaux ou encore *Plasmodium*, l'agent responsable de la malaria.

La toxoplasmose est une anthroponose cosmopolite. Des études séro-épidémiologiques chez l'Homme et les animaux ont montré la large distribution géographique et la forte prévalence du toxoplasme. Plus du tiers de la population mondiale aurait été exposé au parasite (Jackson et *al.*, 1989). Cependant le pourcentage de personnes séropositives pour l'infection toxoplasmique varie d'un pays à l'autre, en fonction des groupes ethniques, des habitudes alimentaires et des conditions d'hygiène. En France, 60% environ de la population est séropositive, l'infection résultant principalement de la consommation de kystes contenus dans des viandes crues ou peu cuites (Morlat et *al.*, 1993). La toxoplasmose est également un problème de santé vétérinaire car elle affecte un grand nombre d'hôtes intermédiaires, source potentielle de contamination pour l'Homme. De plus, elle est responsable d'avortements chez les ovins, entraînant de ce fait, des pertes économiques pour les éleveurs. L'avortement n'ayant lieu qu'en cas de primo-infection en cours de gestation, un vaccin destiné à contaminer les brebis en dehors d'une grossesse a été commercialisé. Il repose sur l'inoculation d'une souche atténuée de parasite (la souche S48) incapable de se différencier en bradyzoïtes (Buxton et *al.*, 1995). *T. gondii* est transmis à l'Homme par l'ingestion d'oocystes sporulés présents dans les aliments souillés, ou par la consommation de viande contenant des kystes et insuffisamment cuite.

### 1.1.1 TOXOPLASMOSE ACQUISE DU SUJET IMMUNOCOMPETENT

Maladie généralement bénigne, voire asymptomatique, elle peut parfois se traduire par un syndrome pseudogrippal accompagné ou non d'adénopathies (Ho Yen et *al.*, 1992; Montoya et Liesenfeld, 2004). Des atteintes viscérales sévères sont possibles mais exceptionnelles. La primo-infection évolue ensuite spontanément vers

une phase chronique asymptomatique qui persiste toute la vie en l'absence de déficit immunitaire. Le diagnostic de la toxoplasmose acquise est basé sur la mise en évidence d'une séroconversion toxoplasmique. L'enkystement rapide du parasite dans les muscles et le système nerveux le met à l'abri des atteintes de la réponse immune spécifique qui se développe parallèlement à l'enkystement. La toxoplasmose peut se révéler dangereuse en cas d'immatunité ou de déficience du système immunitaire que l'on retrouve dans la toxoplasmose congénitale ou celle de l'immunodéprimé (SIDA, greffe...).

### 1.1.2 LA TOXOPLASMOSE CONGENITALE

La toxoplasmose congénitale résulte du passage transplacentaire des parasites circulant dans le sang maternel au cours d'une primo-infection chez la femme enceinte. L'immatunité du système immunitaire du fœtus l'empêche alors de réagir contre le parasite. La probabilité de transmission du parasite au fœtus et la gravité de l'atteinte fœtale évoluent inversement au cours de la grossesse (Desmonts et Couvreur, 1974 ; Montoya et Liesenfeld, 2004). La manifestation la plus courante d'une infection précoce du fœtus par le toxoplasme, est une encéphalomyélite aux conséquences graves : avortement ou mortalité néonatale. Les survivants souffrent généralement de retards mentaux progressifs ou d'autres déficiences neurologiques. Si la transmission materno-fœtale du parasite intervient tardivement au cours de la grossesse, les effets sur le fœtus sont moins sévères. Cependant, ils risquent de développer à l'adolescence des symptômes oculaires (rétinchoroïdite, strabisme, cécité), du système nerveux central (déficience psychomotrice, convulsions, retards mentaux) ou des oreilles (surdité). L'incidence de la maladie varie selon les pays, touchant de 1 à 20 naissances sur 10 000. En France, le taux actuel est d'environ 10 cas pour 10 000 (soit environ 700 cas annuels), ce qui place la toxoplasmose congénitale au premier rang des affections congénitales pour lesquelles un dépistage prénatal ou néonatal est effectué.

### 1.1.3 TOXOPLASMOSE DU SUJET IMMUNODEPRIME

Elle touche des sujets présentant une immunosuppression sévère (SIDA, allogreffe de moelle osseuse, greffe d'organe) et résulte soit d'une primo-infection soit

d'une réactivation à partir des kystes présents chez un sujet antérieurement infecté (Ho Yen *et al.*, 1992 ; Mele *et al.*, 2002). La réactivation parasitaire correspond à une conversion des bradyzoïtes en tachyzoïtes, consécutive à l'état d'immunodépression, suivie d'une multiplication et d'une dissémination des tachyzoïtes dans l'organisme du patient. La toxoplasmose cérébrale chez les patients infectés par le virus du SIDA et dont l'immunodéficiences est avancée représente la deuxième infection opportuniste après la pneumocystose pulmonaire. Suivant les enquête, on estime qu'environ 12 à 47% des personnes infectées par le virus du SIDA et séropositives pour *T.gondii* souffriront d'une réactivation du parasite en l'absence d'un traitement prophylactique efficace. Les patients séronégatifs pour le toxoplasme et recevant le coeur d'un donneur séropositif présentent un risque élevé de développer la maladie. De plus, la réactivation des parasites présents dans le greffon est facilitée par le traitement immunosuppresseur mis en place pour éviter les épisodes de rejet. Dans le cas des patients subissant une greffe de moelle osseuse, le risque de réactivation toxoplasmique endogène est élevé chez les sujets séropositifs recevant la moelle d'un donneur séronégatif. En effet, irradié avant la greffe, le receveur porteur latent de *T.gondii*, perd la capacité de répondre contre le parasite. De plus, cette capacité n'est pas restaurée par le greffon qui ne contient pas les cellules immunocompétentes sensibilisées vis à vis du parasite.

Il n'existe pas de vaccin contre la toxoplasmose. Le diagnostic actuel permet seulement de déterminer s'il y a eu infection, mais n'a pas de valeur prédictive en termes de physiopathologie (Ils ne permettent donc pas d'éliminer le parasite de l'organisme), qui persiste sous forme enkystée. Enfin, si des traitements relativement efficaces sont utilisés comme la Pyriméthamine ou la Sulfadiazine, ceux-ci ne sont pas dénués d'effets toxiques (Tenant-Flowers *et al.*, 1991). De plus, leur action se limite à la forme parasitaire proliférative (tachyzoïte) et ne prévient pas l'enkystement parasitaire. Chez les malades immunodéprimés (chimiothérapie du cancer, SIDA, greffe), les traitements ne sont donc prescrits qu'après la réactivation parasitaire, c'est à dire après l'apparition des premiers symptômes d'encéphalite toxoplasmique (désorientation, somnolence, céphalées...). Dès lors, la maladie a déjà provoqué des séquelles irrémédiables. Par ailleurs, ces traitements ne sont pas dénués d'effets secondaires. Aussi, les progrès pour le contrôle de cette maladie passent par une

meilleure compréhension des facteurs de pathogénicité, en particulier des facteurs de développement intracellulaire et des facteurs de résistance de l'hôte.

### 1.2 *TOXOPLASMA GONDII*

#### 1.2.1 LES DIFFERENTES SOUCHES GENOTYPIQUES

A l'heure actuelle, il est admis que la population de *T. gondii* est constituée de trois lignées clonales prédominantes : les souches de génotype I, II et III. Globalement, cette classification corrèle avec d'une part, la virulence des souches dans le modèle murin et d'autre part, l'origine de la souche (l'homme *versus* réservoirs animaux). Les souches de type I, représentées par la souche RH (communément utilisée en laboratoire), sont des souches très virulentes pour la souris (DL 100 < 10 parasites). Les souches de type II et III, moins virulentes, déterminent chez la souris, une infection chronique avec persistance de kystes intracérébraux. Dans les isolats humains, il existe une prépondérance des souches de type II (81 %). La fréquence des souches de type I et III ne serait que de 10% et 7%, respectivement. Les études réalisées dans le cadre du Centre de Ressources Biologiques ToxoBS, consacré aux souches de toxoplasmes (290 isolats avec données cliniques) confirment que le type II est responsable de la quasi-totalité des toxoplasmoses en France métropolitaine, qu'elles soient asymptomatiques ou non. Par contre, les rares génotypes parasitaires différents (type I, recombinants ou atypiques) sont associés à des formes sévères suggérant un rôle de la souche dans les manifestations cliniques. Ce rôle est également évident dans les toxoplasmoses acquises sévères, voire mortelles, qui ont récemment été mises en évidence chez des patients immunocompétents en Guyane (Ajzenberg *et al.*, 2004). L'analyse de ces isolats montre un plus grand polymorphisme génétique qui ne permet pas de les classer dans les trois génotype déjà décrit.

#### 1.2.2 SON CYCLE BIOLOGIQUE

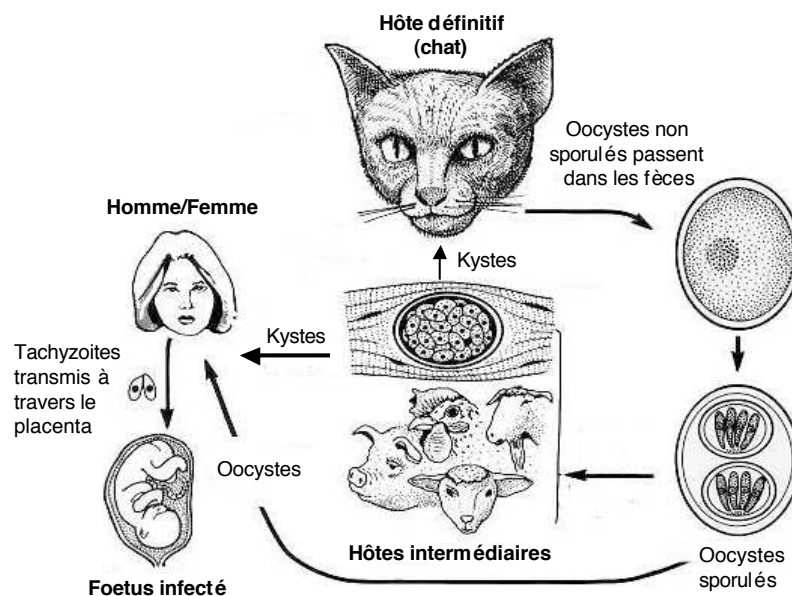
*Toxoplasma gondii* est un parasite intracellulaire obligatoire. Il possède un cycle biologique hétéroxène (Figure A). L'hôte définitif est un féliné (le chat en particulier) où se déroule la reproduction sexuée. La multiplication asexuée prend place chez les

hôtes intermédiaires, les homéothermes, parmi lesquels on retrouve l'Homme. Trois formes de différenciation parasitaire sont infectieuses et peuvent ré-initier le cycle coccidien chez le féliné ou le cycle asexué chez un hôte homéotherme :

(i) **le tachyzoïte** est une forme végétative intracellulaire hautement répliquative, mesurant  $6-8 \mu\text{m} \times 2-3 \mu\text{m}$ . Présent pendant la phase aiguë de l'infection toxoplasmique, il infecte les cellules nucléées de l'hôte et s'y multiplie rapidement au sein d'une vacuole parasitophore qui le protège de l'acidification et empêche la fusion avec les lysosomes.

(ii) **le bradyzoïte** est une forme parasitaire à multiplication lente, ou végétative, issus de la conversion de certains tachyzoïtes. Ils forment les kystes tissulaires ( $20$  à  $200 \mu\text{m}$ ) qui caractérisent la phase chronique de l'infection toxoplasmique.

(iii) **le sporozoïte**, issu de la reproduction sexuée est contenu dans l'ocyste, forme de résistance du parasite dans le milieu extérieur.



**Figure A. Cycle biologique du toxoplasme.** Le parasite possède un seul hôte définitif (félin) dans lequel s'effectue la reproduction sexuée par fécondation des gamétoocytes et de nombreux hôtes intermédiaires homéothermes dans lesquels s'effectue la reproduction asexuée.

## Introduction

L'hôte définitif félin s'infeste le plus souvent en ingérant des kystes présents dans le cerveau ou les muscles de petits rongeurs. Il peut également se contaminer par l'absorption d'oocystes. Les bradyzoïtes ou les sporozoïtes envahissent alors les cellules épithéliales de l'intestin et s'y multiplient activement par schizogonie. Puis, les mérozoïtes produits se différencient au cours de la gamétogenèse, en microgamètes mâles flagellés et macrogamètes femelles immobiles. Issus de la fécondation, les oocystes sont, après rupture de la cellule intestinale infestée, éliminés dans le milieu extérieur avec les fèces du félin. Ils subissent alors une phase de sporulation à l'issue de laquelle ils deviennent infectieux. Ingérés par un chat, ils initient un nouveau cycle sexué. Par contre, chez les hôtes intermédiaires, ils sont à l'origine du cycle asexué.

Le cycle asexué se déroule chez les homéothermes (Figure A). La contamination se fait par ingestion des kystes ou des oocystes. Les bradyzoïtes ou sporozoïtes pénètrent l'épithélium intestinal où ils se transforment en tachyzoïtes. Les tachyzoïtes vont se multiplier activement et entraîner la lyse des cellules parasitées. Libérés, ils peuvent infecter les cellules adjacentes, et notamment celles du système réticulo-histiocytaire, qui ouvre la voie à une dissémination rapide par voie lymphatique et sanguine de l'organisme entier. Cette phase de dissémination et de prolifération intense qui affecte le foie, les poumons, les tissus lymphoïdes, les muscles striés et le cerveau, constitue la phase aiguë de l'infection. Dans la plupart des cas, des défenses immunitaires spécifiques s'établissent pour contrôler l'infection. Les tachyzoïtes sont alors progressivement éliminés. Certains se différencient en bradyzoïtes, forme à multiplication lente, qui s'enkystent au niveau des tissus tels que les muscles et le cerveau. L'enkystement correspond à la phase chronique de l'infection qui persiste de manière latente pendant toute la durée de vie de l'hôte.

Comme beaucoup d'Apicomplexes, *T. gondii* est un parasite intracellulaire obligatoire, pour lequel l'invasion de la cellule-hôte est une étape capitale du cycle biologique, et par là même un composant essentiel de la pathogenèse. Ce phénomène consiste pour le parasite à induire la formation dans la cellule cible d'un nouveau compartiment subcellulaire (la vacuole parasitophore) dans lequel se trouveront réunies les conditions adéquates à son développement. Le parasite va résider dans cette vacuole qui constitue une interface dynamique entre le parasite et la cellule hôte.

Les tachyzoïtes s'y divisent très rapidement par endodyogénie. Après six à huit cycles de multiplication, les parasites vont lyser la cellule hôte et infecter les cellules environnantes. Les parasites qui se répliquent par fission binaire s'organise en rosette, à l'intérieur de la vacuole parasitophore. Les tachyzoïtes sont ensuite disséminés dans l'hôte par les voies sanguine et lymphatique.

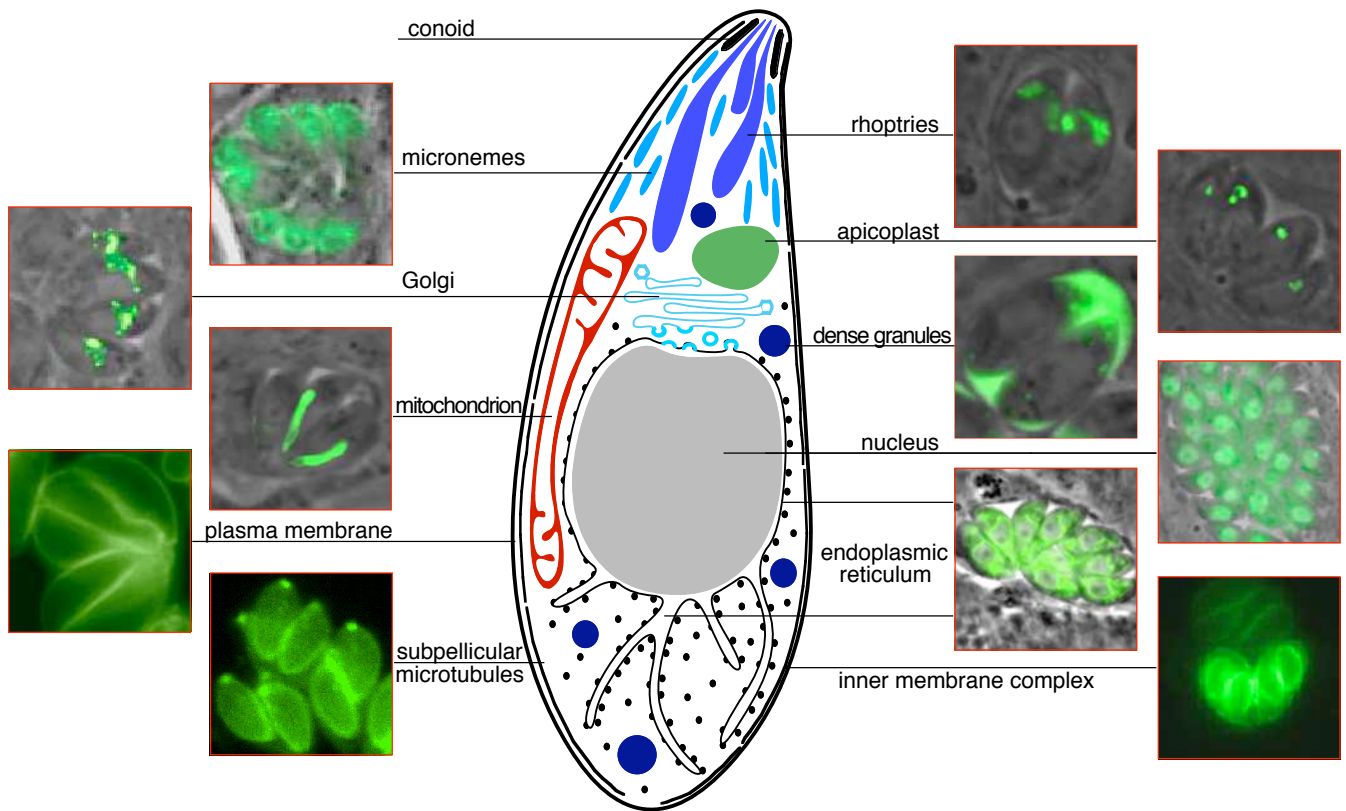
### 1.2.3 L'ULTRASTRUCTURE GENERALE DU TOXOPLASME.

Comme tous les parasites intracellulaires de l'Embranchement des *Apicomplexa*, *T. gondii* possède une organisation cellulaire polarisée contrôlée par un cytosquelette élaboré et caractérisée par la présence d'un complexe apical. Composé du conoïde, des rhoptries et des micronèmes, le complexe apical permet l'entrée active rapide (en 15 à 30 secondes) du parasite dans les cellules dans lesquels il se multiplie à l'intérieur d'une vacuole parasitophore. *T. gondii* possède trois compartiments sécrétoires, les micronèmes, les rhoptries et les granules denses, qui déchargent successivement leur contenu au cours du processus d'invasion cellulaire (Carruthers, 2002). Au niveau moléculaire, on distingue les protéines de surface et les protéines contenues dans les organelles sécrétoires du parasite - les micronèmes, les rhoptries et les granules denses - qui participent directement à la pénétration du parasite et à sa survie dans la cellule cible. La Figure B illustre les différents organites de ce parasite.

### 1.2.4 LE PROCESSUS D'INTERCONVERSION, AU CENTRE DE LA PATHOGENESE

Le processus d'interconversion tachyzoïte-bradyzoïte est un élément central dans la pathogénèse et la longévité de l'infection. Définir les facteurs qui influencent l'interconversion pourrait contribuer significativement au développement de nouvelles thérapies qui préviendraient la réactivation du parasite chez les patients immunodéprimés. Des études ont montré que la morphologie des bradyzoïtes diffère de celle des tachyzoïtes en divers aspects: (i) la modification de la membrane de la vacuole parasitophore (VP) par l'addition de chitine, de glycoprotéines, et probablement de glycolipides pour former la paroi kystique (Boothroyd *et al.*, 1997; Zhang *et al.*, 2001); (ii) l'accumulation de granules d'amylopectine, reflétant un substantiel stockage de glucose; et (iii) la réorganisation sub-cellulaire avec la reloca-





**Figure B - Marquage fluorescent de protéines chimériques dans les organites sub-cellulaires chez *T. gondii*.**

Le dessin central montre un tachyzoïte avec ses structures sub-cellulaires. Autour de celui-ci sont indiqués les organites marqués à la GFP ou à la YFP : le conoïde et les microtubules sont marqués avec la YFP-Tubuline  $\alpha$ , les micronèmes avec MIC3-GFP, le golgi avec MIC3(68-137)-GFP, la mitochondrie avec HSP60-GFP, la membrane plasmique avec P30-GFP-GPI, les rhoptries avec ROP1-GFP, les granules denses avec P30-GFP, le noyau avec PCNA-GFP, le réticulum endoplasmique avec P30-GFP-HDEL et le complexe interne membranaire avec IMC1-YFP. Ce schéma est tiré d'une revue de Joiner & Roos (2002).

-lisation postérieure du noyau et la redistribution des organes sécrétoires importants pour l'invasion et l'entretien de cellule hôte du VP. La mise en place et l'évolution de ces modifications structurales ne sont pas connues avec précision.

En culture tissulaires *in vitro*, les conversions spontanées tachyzoïte-bradyzoïte et réciproquement sont observées. Le taux de conversion est apparemment souche-dépendante. Ainsi, la souche de type II, ME49, a un plus grand taux de conversion spontanée que la souche de type I, RH (Weiss et Kim, 2000). L'inhibition de la réplication des tachyzoïtes (réplication rapide) s'accompagne en culture d'une augmentation du nombre de bradyzoïte (réplication lente). L'application de stress aux conditions de culture (température, pH) se traduit aussi par une augmentation des conversions tachyzoïte-bradyzoïte. La présence d'interféron  $\gamma$  ou encore l'inhibition de la chaîne respiratoire mitochondrial induisent aussi la transition du tachyzoïte vers la forme bradyzoïte. Au cours de cette transition le *Toxoplasme* convertit son métabolisme and son ultrastructure morphologique. Des marqueurs antigéniques sont disponibles pour établir avec précision le *ratio* tachyzoïte/bradyzoïte dans le milieu de culture (Weiss et Kim, 2000). Cependant la différenciation *in vitro* reste asynchrone et les cultures de bradyzoïte sont en général envahies par des tachyzoïtes qui se divisent plus rapidement ; ce qui complique l'analyse des populations de bradyzoïtes.

L'interconversion est associée avec des changements morphologiques du parasite, l'expression d'antigènes de surface spécifiques à une forme parasitaire donnée, ou encore la mise en route de nouvelles voies métaboliques (Ferguson et Hutchison, 1987; Denton H *et al.*, 1996). De nombreux gènes sont exprimés spécifiquement aux stades tachyzoïte (SAG1, SAG2A, SAG2B, LDH1, ENO2, etc) ou bradyzoïte (BSR4, SAG2C, SAG4, BAG1, LDH2, ENO1, etc) (Lyons *et al.*, 2002). Différentes approches (banques d'ADNc soustractives, « differential display ») ont permis d'appuyer l'idée d'une expression différentielle des gènes selon le stade étudié. Le passage d'une forme à l'autre reposerait alors sur un équilibre strict entre activation et répression de grandes familles de gènes (Singh *et al.*, 2002). Chacune de ces voies correspond à la mise en route de programmes génétiques spécifiques et c'est donc au niveau de la transcription, en grande partie au moins, que devrait être contrôlé l'engagement de *Toxoplasma gondii* vers l'une ou l'autre forme. L'interconversion

semble donc mettre en jeu un système complexe de contrôle transcriptionnel dont les mécanismes moléculaires sont encore inconnus.

L'hypothèse d'une régulation transcriptionnelle de l'interconversion nous a conduit à rechercher dans le protéome prédit du toxoplasme des domaines structuraux propres à certains facteurs de transcription eucaryotiques. Le domaine de liaison à l'ADN "helix-turn-helix" (HTH) structure de nombreux facteurs de transcription spécifique (FTs) bactériens et archaebactériens. Les FTs eucaryotiques présentent de nombreuses versions du domaine HTH comme les domaines homeo, Forkhead (Fkh), Bright (ARID), MYB ou encore PSQ (Aravind *et al.*, 2005). En revanche, les eucaryotes ont acquis au cours de l'évolution de nombreux facteurs indépendants structurellement du domaine HTH avec une affinité à l'ADN différente, constituant ainsi des nouvelles familles de FTs. Ces familles incluent les facteurs chélatant le zinc comme les 'doigts de zinc' (zinc Finger) C2H2, les facteurs présentant un repliement hélicale (HMG, bZip et bHLH) ou encore ceux arborant des repliements plus complexes comme des domaines de repliement VP1, AP2, GCM ou TIG (Riechmann *et al.*, 2000). L'analyse génomique comparative des FTs eucaryotiques suggère que les grandes familles ont émergé principalement à travers le processus d'expansion des espèces (Riechmann *et al.*, 2000; Chervitz *et al.*, 1998; Lespinet *et al.*, 2002). Par exemple, le 'zinc-finger' le plus répandu chez les levures de type C6 binuclear Zn-finger est totalement absent chez les plantes et les animaux qui en revanche ont acquis leur propre facteurs comme le domaine VP1 pour les plantes et les récepteurs nucléaires hormone-dépendant chez les animaux.

Des études génomiques similaires chez les Apicomplexa, qui s'embranchent sur l'arbre de la vie en dehors des grands groupes eucaryotiques, montrent une quasi-pénurie de FTs discernable bio-informatiquement à travers leurs protéomes (Templeton *et al.*, 2004; Balaji *et al.*, 2005). L'analyse détaillée des génomes de *Plasmodium falciparum* et de *Cryptosporidium parvum* montre une absence des domaines de liaison à l'ADN conservés chez l'ensemble des autres eucaryotes comme les domaines HTH, bZip, homeo, bHLH ou Fkh. Seuls quelques rares domaines de type Zinc Finger C2H2 et E2F ont pu être identifiés (Templeton *et al.*, 2004). Alors que la proportion de FTs semble augmenter avec la complexité biologique de

l'organisme (3,4% pour *Saccharomyces cerevisiae*, 4,2% pour *Caenorhabditis elegans*, 5,5% pour *Drosophila melanogaster* et 8% pour *Homo sapiens*; Messina *et al.*, 2004) il semble que les parasites Apicomplexa ne suivent pas cette règle (Templeton *et al.*, 2004). Une partie de cette "anomalie" pourrait être expliquée par le style de vie parasitaire, qui n'exige pas de répondre aux contraintes environnementales et aux défis homéostatiques qui se posent, par exemple, à une levure. Mais le style de vie parasitaire n'est pas une réponse suffisante puisque les Apicomplexes harborent un arsenal sophistiqué et comparable à celui des autres eucaryotes de protéines de signalisation comme les kinases ou les GTPases.

Les bases moléculaires de la régulation transcriptionnel de la différenciation chez *T.gondii* reposent probablement sur une autre mécanistique beaucoup moins dépendante de la diversification des facteurs spécifiques de transcription, à l'opposé des autres cellules eucaryotes. Il est maintenant clairement admis que le contrôle transcriptionnel dans la cellule eucaryote est largement sous la dépendance de la structure de la chromatine. Le remodelage de la chromatine a ainsi un impact drastique sur la régulation de la transcription des gènes. Plus la compaction est importante, plus la transcription est réprimée et inversement, la décompaction se traduit par une activité transcriptionnelle intense.

Le terme épigénétique définit les modifications transmissibles et réversibles de l'expression des gènes et ne s'accompagnant pas de changements des séquences nucléotidiques. Ce type de régulation peut cibler l'ADN, l'ARN ou les protéines et agir au niveau du noyau ou du cytoplasme. Le niveau d'activité génique peut changer par «épimutation», ou modification épigénétique, de manière héritable et sans altération génétique concomitante. Une épimutation peut avoir des conséquences phénotypiques importantes. Par exemple, des épimutations sont responsables de la divergence entre un neurone et une cellule musculaire provenant du même œuf et contenant le même génome. Cette différence phénotypique est considérable au regard de celle qui sépare les neurones de deux espèces animales aux génotypes dissemblables, soulignant en cela l'importance des modifications épigénétiques. Ainsi, les épimutations constituent l'un des fondements de la diversité biologique. Des altérations intempestives de ce code épigénétique auront des répercussions sur la

prolifération et la différenciation et peuvent résulter de la surexpression (fonction oncogénique) ou de la perte (fonction de gène suppresseur) d'acteurs clefs de cette régulation. De tels cas ont été décrits dans de nombreux types de cancers. En conséquence, un programme épigénétique aberrant peut tout à fait constituer un événement principal dans l'étiogénèse de certaines formes de cancers.

Notre hypothèse de travail est (i) que la pénurie de facteurs de transcription spécifiques chez le *Toxoplasme* est compensée par une régulation fine de la structure chromatinienne et (ii) que la différenciation parasitaire répond à des signaux épigénétiques complexes. Le passage de la forme proliférative tachyzoïte à la forme enkystée bradyzoïte serait alors la conséquence phénotypique d'épimutations réversibles. Ces épimutations définissent un état transcriptionnel propre à une forme parasitaire donnée ou à un état transitionnel qui autorise le passage d'une forme à l'autre. Pour mieux comprendre l'impact de la chromatine sur la régulation génique, les paragraphes suivants introduisent sommairement les éléments mécanistiques majeurs fondateurs de cette discipline.

## 2. CHROMATINE ET REGULATION GENIQUE CHEZ LES EUCARYOTES

L'ADN des cellules eucaryotes est compacté dans le noyau sous forme de chromatine constituée d'ADN et de protéines histone et non-histone (Agresti et Bianchi, 2003 ; Catez *et al.*, 2003 ; Reeves, 2001). En interphase de mitose, la chromatine occupe tout l'espace nucléaire limité par une membrane. Deux types de chromatine peuvent être distingués: la chromatine condensée, ou hétérochromatine et la chromatine décondensée, ou euchromatine. L'hétérochromatine est inactive pour la transcription, et, est répliquée tardivement. Elle est dite constitutive si elle est condensée dans toutes les cellules de l'organisme, et facultative dans le cas contraire. L'hétérochromatine constitutive est retrouvée essentiellement au niveau des centromères et des télomères et contient surtout de l'ADN satellite. Les gènes structuraux sont localisés dans l'euchromatine et dans l'hétérochromatine facultative. Au cours de la mitose, l'enveloppe nucléaire se disloque, et la chromatine se condense progressivement, rendant visible les chromosomes individuellement.

Le compactage de l'ADN dans le noyau est assuré par différents niveaux de repliement et d'enroulement. Le premier niveau est constitué par l'enroulement de la molécule d'ADN autour des histones (H3, H4, H2A et H2B) pour former les nucléosomes. L'empilement en colonne des nucléosomes, par l'intermédiaire de l'histone H1, conduit à ce qu'on appelle la fibre de 10 nm de diamètre. L'enroulement de cette fibre conduit à la fibre de 30 nm (~6 nucléosomes par tour), dont le modèle le plus couramment présenté est celui du solénoïde. Dans la fibre de 30 nm, le facteur de compaction de l'ADN est de 40 fois. Dans le noyau, le facteur de compaction est d'environ 1000 fois, et il augmente encore d'un facteur 10 dans les chromosomes en mitose. Au-delà de la fibre de 30 nm, la compaction de la chromatine se fait par interaction avec les protéines non histones qui constituent la matrice nucléaire (noyau en interphase) ou le squelette chromosomal (chromosomes à la mitose). Le degré supplémentaire de compaction dans les chromosomes est obtenu par repliement de la fibre chromatidienne.

Les nucléosomes constituent donc une barrière physique et chimique qui limite l'accès à la molécule d'ADN. Mais la structure du nucléosome n'est pas figée : elle est influencée par de nombreux facteurs qui peuvent la modifier. Cette structure dynamique est déterminante pour toutes les fonctions du génome : la transcription, la réplication, la réparation et la recombinaison. Outre son rôle structural, le nucléosome a la capacité de moduler l'accessibilité des séquences d'ADN aux facteurs transcriptionnels, de participer à la signalisation, et même de transmettre de génération en génération des informations dites 'épigénétiques'. Deux grands mécanismes généraux contribuant à la plasticité chromatinienne ont été abondamment décrits dans la littérature ces dernières années. Le premier implique de grands complexes protéiques qui utilisent l'énergie libérée par l'hydrolyse de l'ATP pour altérer physiquement la structure du nucléosome. Le deuxième fait intervenir des modifications chimiques covalentes sur les queues d'histones ou directement sur l'ADN. Pour l'ADN, les séquences cibles sont les séquences CpG, sur lesquelles un groupement méthyl est ajouté au niveau 5' de l'anneau de la cytosine. Ce processus est probablement absent chez *T. gondii*. En ce qui concerne les queues d'histones, elles subissent différentes modifications: l'acétylation, la phosphorylation, la méthylation, l'ADP-ribosylation, la sumoylation, ou encore l'ubiquitination. Ces

modifications ont des effets variables en fonction de la nature du groupement, de l'emplacement de l'acide aminé modifié, et de son environnement proche (même queue d'histone) ou éloigné (queues d'histones différentes). L'ensemble de ces modifications et de leurs variations spatiales et temporelles constitue le « code histone ». La lecture de ce code permet en théorie de déterminer si un gène est réprimé ou activé.

La chromatine n'est pas seulement un obstacle à l'initiation de la transcription, elle est également un frein à l'élongation. Il existe des facteurs qui aident l'ARN Polymérase II à transcrire à travers les nucléosomes. La purification de la polymérase en élongation a en effet permis l'identification du facteur FACT (*FAcilitates Chromatin Transcription*) qui permet à une ARN Pol II arrêtée au niveau d'un nucléosome de reprendre la transcription indépendamment de l'hydrolyse d'ATP et des facteurs TFIIF ou TFIIS (Orphanides *et al.*, 1998). La fonction de FACT implique probablement la fixation directe des dimères H2A/H2B et la dissociation des octamères d'histone (Orphanides *et al.*, 1999). Chez la levure, le complexe *Elongator*, identifié chez la levure (Otero *et al.*, 1999) et chez l'Homme (Hawkes *et al.*, 2002; Kim *et al.*, 2002), semble jouer un rôle similaire. Le complexe DSIF (formé des protéines Spt4 et Spt5) participe également à l'élongation de la transcription *in vivo* (Andrulis *et al.*, 2000; Kaplan *et al.*, 2000), probablement par un mécanisme impliquant le remodelage de la chromatine (Winston et Sudarsanam, 1998). Enfin, il faut noter que la transcription ne requiert pas le relargage complet des nucléosomes puisque la machinerie de transcription semble pouvoir tourner autour du nucléosome déplaçant ainsi la bulle de transcription autour de lui (Studitsky *et al.*, 1997).

### 2.1. Remodelage de la Chromatine ATP-dépendant

Le remodelage physique du nucléosome représente l'ensemble des réarrangements structuraux qui rendent l'ADN nucléosomal potentiellement plus accessible. Les facteurs de remodelage ont tous en commun une sous-unité catalytique ATPase ayant un domaine hélicase. L'activité ATPase permet au complexe de modifier la structure nucléosomique, en partie grâce à l'énergie libérée par l'hydrolyse de l'ATP (Kingston et Narlikar, 1999). Les facteurs de remodelage ayant

des sous-unités ATPase identiques ou très similaires ont été classés, ici, dans une même famille. La sous-unité ATPase, appelée *swi2/snf2* (*mating-type switching/sucrose non-fermenting*) définit la première famille appelée SWI/SNF. L'ATPase ISWI (*imitation of switch*) définit la deuxième famille ainsi nommée ISWI ; enfin la dernière famille est appelée Mi2/NuRD (*nucleosome remodeling histone deacetylase complex*) d'après le nom de sa sous-unité ATPase, Mi2.

Des approches biochimiques ont montré que les différents complexes SWI/SNF ont la capacité de rendre l'ADN plus accessible en créant une structure nucléosomique qui oscille entre la forme canonique d'origine et une nouvelle forme, dite altérée, récemment mise en évidence (Schnitzler *et al.*, 1998). Lors de ce processus, les histones nucléosomiques peuvent être transférées en *trans* sur de l'ADN libre. L'état «altéré» rendrait l'ADN plus accessible à des protéines telles que des facteurs de transcription. Un mécanisme de « glissement », entraînant le déplacement des octamères d'histones en *cis* le long de l'ADN, est ainsi proposé pour la famille ISWI (Lorch *et al.*, 1999; Studitsky *et al.*, 1994). La spécificité des enzymes catalysant l'ATP semble varier selon le substrat. Ainsi par exemple, l'ATPase SWI/SNF ne semble pas faire la différence entre l'ADN nu et les nucléosomes tandis que les ATPases ISWI et Mi-2 montrent une spécificité accrue pour l'ADN enroulé autour des nucléosomes (Johnson *et al.*, 2005).

Certains facteurs de remodelage agissent en concert avec des facteurs qui catalysent des modifications chimiques des queues d'histone, voire de la molécule d'ADN. Ainsi, le facteur de remodelage Mi2/NuRD (*nucleosome remodeling histone deacetylase*) a été identifié lors de la recherche de complexes protéiques contenant des activités histone-désacétylases (HDAC1/2) (Zhang *et al.*, 1998; Knoepfler *et al.*, 1999). Plusieurs complexes ont été isolés qui comprennent tous une sous-unité ATPase désignée selon les cas Mi2 ou CHD (*chromodomain*). Ce complexe participerait à l'établissement d'une structure répressive de la chromatine. Ainsi, chez la drosophile, Mi2 interviendrait à la fois dans le déclenchement et le maintien de l'état réprimé des gènes développementaux *HOX*. De même, chez les mammifères, Mi2 interagit avec Ikaros, un régulateur transcriptionnel essentiel dans la différenciation et la prolifération des cellules lymphoïdes (Kim *et al.*, 1999).



Enfin, le complexe Mi2/NuRD pourrait être ciblé dans les régions méthylées du génome *via* sa sous-unité MBD3 (*methyl-CpGbinding domain*) capable de lier l'ADN méthylé (Wade *et al.*, 1998). La méthylation de l'ADN apparaît comme un marqueur épigénétique relayé par des modifications de la chromatine telles que la désacétylation (Bird et Wolffe, 1999); le complexe Mi2/NuRD pourrait ainsi établir un lien entre méthylation et structure chromatiniennne répressive (Wade *et al.*, 1998).

### 2.2. Les Modifications Covalentes des Queues d'Histones

Afin de pouvoir accéder à un espace suffisant pour que les facteurs d'initiation puissent prendre place au niveau du promoteur et permettre l'ouverture du double brin d'ADN, la chromatine doit être aussi modifiée chimiquement. Ce remodelage fait intervenir une machinerie cellulaire spécialisée comportant des enzymes capables de modifier covalamment les histones par acétylation, phosphorylation ou méthylation, rompant ainsi les liaisons ADN-histones et permettant la translocation des nucléosomes (Berger, 2002).

Les queues N-terminales des histones ont des interactions fortes avec l'ADN et sont le siège de nombreuses modifications covalentes post-traductionnelles. (Peterson et Laniel, 2004). La Figure C identifie les modifications caractérisées à ce jour chez les eucaryotes (Marks *et al.*, 2001). Les modifications majeures sont l'acétylation des résidus lysines (K) ; les mono-, di- et tri-méthylations des résidus lysine (K) ; les mono- et di-méthylations symétriques ou asymétriques des résidus arginines (R) ; les phosphorylations des résidus sérine (S) et thréonine (T) ; la biotinylation, l'ubiquitination et la sumoylation des résidus lysine (K) ; la citrullination des résidus arginines (R) et enfin la mono-ADP-ribosylation des résidus arginines (R) et glutamate (E) (Burzio *et al.*, 1979 ; Kreimeyer *et al.*, 1984 ; Lachner et Jenuwein, 2002 ; Lachner *et al.*, 2003 ; Nowak et Corces 2004 ; Osley, 2004 ; Lo *et al.*, 2001). De nombreux complexes avec une (ou des) activité(s) enzymatique(s) associée(s) ont été caractérisée(s) biochimiquement ou génétiquement chez différents eucaryotes. Les complexes associés à des co-activateurs transcriptionnels et des activités enzymatiques de type acétylase activent la transcription (Kuo et Allis 1998 ; Marmorstein, 2001).

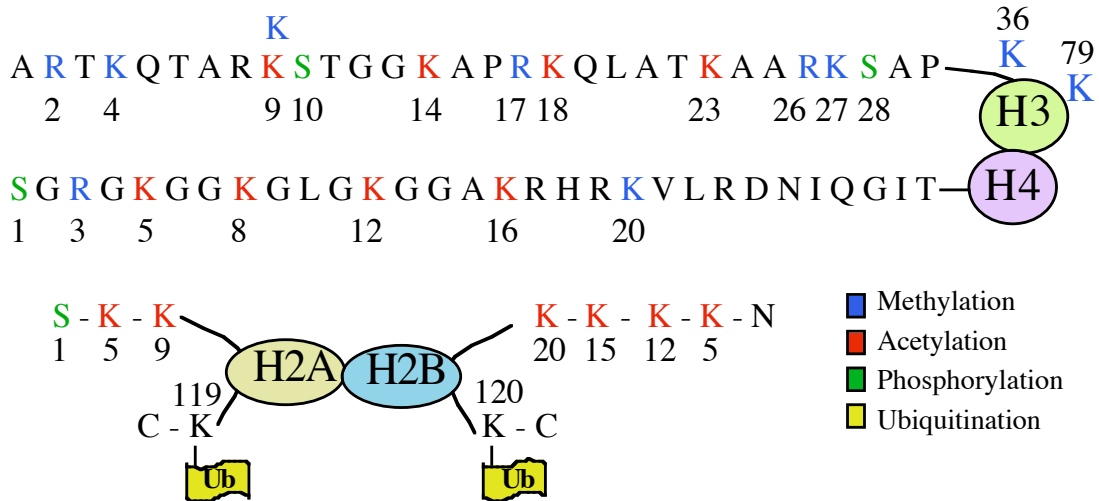


Figure C. Résidus des histones sujets à des modifications post-traductionnelles

Inversement, les complexes associés à des co-répresseurs et présentant des activités antagonistes aux précédents comme la déacétylation, réprime eux la transcription. À l'échelle cellulaire, le niveau d'acétylation délimite ainsi topographiquement l'euchromatine et l'hétérochromatine au moyen d'un gradient d'acétylation/désacétylation (Suka *et al.*, 2002). Cette dualité dans la fonction entre répression et activation se retrouve dans la méthylation et déméthylation des résidus lysine et arginine (Kouzarides, 2002 ; Zhang et Reinberg, 2001).

### 2.2.1 Acétylation des Histones Et Transcription

Les histones H3 et H4 peuvent être acétylés au niveau de résidus lysine situés dans leur partie amino-terminale, qui s'étend en dehors de la structure globulaire de l'octamère d'histone (Kuo *et al.*, 1996). La stabilité des interactions ADN-histones semble être réduite par cette neutralisation des charges basiques portées par les lysines. Cette acétylation conduit donc à une forme moins condensée de la chromatine, où les nucléosomes ne sont néanmoins que partiellement déstructurés (Oliva *et al.*, 1990; van Holde *et al.*, 1992; Lee *et al.*, 1993; Vettese-Dadey *et al.*, 1996). L'état d'acétylation résulte d'un équilibre dynamique entre deux activités antagonistes, régulées par deux familles d'enzymes ; les histone-acétyltransférases (HAT) qui ajoute un groupement acétyl sur le groupement amine en position  $\epsilon$  d'un résidu lysine et les histone-déacétylases (HDAC) qui le retire. Généralement, les

activités d'acétylation sont ciblées vers les régions promotrices des gènes qui doivent être exprimés, rendant ainsi l'ADN accessible aux facteurs de transcription. L'activité transcriptionnelle peut également être contrôlée par le biais des régulateurs qui désacétylent les histones et qui, en conséquence, influencent négativement la liaison des facteurs de transcription à l'ADN (Hendzel *et al.*, 1994; Wade *et al.*, 1997, Kuo et Allis, 1998, Wolffe et Hayes, 1999).

L'hyperacétylation des lysines des parties amino-terminales des histones est engendrée par des activités histones acétyltransférases (HATs) qui permettent le transfert d'un groupe acétyl sur des lysines spécifiques modifiant ainsi la charge présente. Plusieurs mécanismes ont été proposés pour expliquer comment l'acétylation des histones pouvait faciliter la transcription. Tout d'abord, des analyses *in vitro* ont montré que l'acétylation des histones permettait aux facteurs de transcription d'accéder plus facilement à l'ADN. Des expériences réalisées avec des enzymes de restriction le suggèrent également. Une des possibilités est que la réduction de la charge positive sur la partie acétylée de l'histone entraîne une stabilité plus faible d'interaction avec l'ADN. Ensuite, il semble que l'acétylation des histones diminue la compaction des nucléosomes en détruisant les interactions entre les nucléosomes qui se font par l'intermédiaire de la queue des histones. En plus de modifier la structure chromatinienne, l'acétylation des histones pourrait constituer un marquage pour le recrutement d'autres protéines (Narlikar *et al.*, 2002). Les résidus Lys acétylés peuvent en effet être reconnus par un motif d'environ 110 résidus appelé 'bromodomaine' et présent dans p/CAF, p300/CBP et TAFII250 (Dhalluin *et al.*, 1999; Hudson *et al.*, 2000; Jacobson, 2000 et Manning *et al.*, 2001). TAFII250 a la particularité de posséder deux bromodomains en tandem qui interagissent fortement avec l'histone H4 doublement acétylé sur des résidus lysine (Jacobson *et al.*, 2000). Cette double acétylation pourrait donc constituer un signal, reconnu par TAFII250, qui permettrait le recrutement de TFIID et l'assemblage et/ou la stabilisation du complexe de pré-initiation autour du promoteur ainsi marqué (Naar, 2001).

En général, les domaines euchromatiques transcriptionnellement actifs sont hyperacétylés par les HAT (Hebbes *et al.*, 1988), tandis que les domaines inactifs hétérochromatiques sont hypoacétylés par les HDAC (Johnson *et al.*, 1998).

Exceptionnellement, la lysine 12 de l'histone H4, chez la levure, est préférentiellement acétylée au niveau de l'hétéchromatine transcriptionnellement silencieuse (Braunstein *et al.*, 1996) et Rpd3, une histone déacétylase, a pour effet de contrebalancer ce signal (De Rubertis *et al.*, 1996; Vannier *et al.*, 1996). Ces observations contrastent avec la corrélation usuelle de lier l'acétylation des histones à l'activation de la transcription. Une des explications possibles de ce paradoxe est que la forme acétylée de la lysine 12 de l'histone H4 soit reconnue par une protéine qui soit impliquée dans la genèse de l'hétérochromatine.

### 2.2.1.a Les histones acétyltransférases

La plupart des histones acétyltransférases ont initialement été décrites comme des co-activateurs transcriptionnels de natures diverses. En effet, TAFII250 est une sous-unité du complexe TFIID et possède une activité HAT (Mizzen *et al.*, 1996) ; p300/CBP (*E1A-associated 300 KDa protein / CREB Binding Protein*), initialement décrit comme co-activateur transcriptionnel, a ensuite été caractérisé par son activité HAT (Bannister et Kouzarides, 1996; Ogryzko *et al.*, 1996) . La protéine PCAF (p300/CBP associated factor) possède également ce type d'activité et peut interagir avec p300 (Yang *et al.*, 1996) mais aussi avec ACTR (*ACTivator of Retinoid receptor*) et SRC-1 (*Steroid Receptor Coactivator 1*) (Chen *et al.*, 1997; Spencer *et al.*, 1997) - d'autres protéines HATs- impliqués dans la régulation de récepteurs nucléaires répondant aux hormones, pour former un complexe multi-HATs. Enfin Gcn5 a été la première HAT identifiée (Brownell *et al.*, 1996), elle est présente chez tous les eucaryotes (Struhl, 1998). Les HATs sont généralement associées à des complexes qui ont été définis en 2 groupes : les HATs de type A sont nucléaires et responsables de l'acétylation des histones nucléosomales lors de l'activation de la transcription alors que les HATs de type B sont cytoplasmiques et participeraient à l'acétylation des histones nouvellement synthétisées avant l'assemblage de la chromatine pendant la réplication de l'ADN (Kuo et Allis, 1998).

Parmi les HATs de type A, on distingue, les GNAT (*GCN5-related N-AcetylTTransferase*), les MYST (*Moz, Ybf2/Sas3, Sas2 et Tip60*), la famille p300/CBP, les facteurs généraux de la transcription (tel que TAF250), ainsi que les co-activateurs des récepteurs nucléaires hormones dépendantes tels que SRC1 et

ACTR. Cette famille se distingue par la présence d'un motif très conservé de fixation de l'acétyl coA et par l'absence de domaine de liaison à l'ADN. Leur recrutement à des adresses chromosomiques spécifiques et le choix des résidus lysine méthylés sur les histones sont influencés par leur interactome directe au sein des complexes multi-protéiques où elles sont intégrées.

La famille des GNAT englobe les enzymes homologues à GCN5 (*General Control Nonrepressed protein 5*) : GCN5, PCAF, Hat1, ATF-2, Elp3, Nut1 et Hpa2. Ces enzymes ont en commun des domaines caractéristiques: (i) un domaine amino-terminal, variable en longueur suivant les espèces, qui joue un rôle très important dans la reconnaissance des substrats nucléosomiques ; (ii) un domaine catalytique conservé HAT contenant un motif (Arg/Gln-X-X-Gly-X-Gly/Ala) de liaison au substrat (acétyl-CoA) ; (iii) un domaine d'interaction à la protéine co-activatrice Ada2 ; et (4) un bromodomaine carboxy-terminal requis notamment pour permettre les interactions protéines/protéines et probablement le recrutement de ces facteurs au niveau du promoteur. Des études *in vivo* ont montré que le bromodomaine est non-essentiel à l'activité HAT mais intervient dans le remodelage médié par SWI/SNF en stabilisant la liaison de SWI/SNF au promoteur (Hassan *et al.*, 2001 ; Syntichaki *et al.*, 2000).

Il est maintenant bien établi que ces activités HATs sont regroupées sous la forme de complexe. L'environnement de la protéine portant l'activité HAT semble important puisque pour PCAF par exemple, lorsqu'elle est associée à son complexe, contenant 20 polypeptides, le niveau d'acétylation est plus important (Ogryzko *et al.*, 1998). Chez la levure, l'enzyme Gcn5 est associé au complexes SAGA (*Spt-Ada-GCN5-Acetyltransferase complex*), SAGA<sub>alt</sub> (*alternate form of SAGA*), SLIK (*SAGA LIKe*), SALSA (*SAGA A Ltered Spt8 Absent*), HAT-A2 et ADA (*Alteration/Deficiency in Activation*) (Grant *et al.*, 1997 ; Grant *et al.*, 1998 ; Pray-Grant *et al.*, 2002 ; Sterner *et al.*, 2002). Les spécificités d'action sont déterminées par les partenaires avec lesquels la protéine Gcn5 de levure interagit. Ainsi, Gcn5 seule acétyle essentiellement *in vitro* le résidu lysine-14 de l'histone H3 alors qu'insérée dans le complexe SAGA, l'enzyme acétyle les résidus K14, K9, K18 et K23 (Grant *et al.*, 1999).

Dans le phylum des apicomplexes, quelques membres de la famille GNAT ont été identifiés : PfGCN5 chez *Plasmodium falciparum* (Fan *et al.*, 2004) et TgGCN5a (Sullivan *et al.*, 2000 ; Bhatti *et al.*, 2005 ; Saksouk *et al.*, 2005) et TgGCN5b (Sullivan WJ, communication personnelle) chez *Toxoplasma gondii*.

### 2.2.1.b Les histones déacétylases

L'importance fonctionnelle de l'acétylation dépend complètement de la précision et de l'efficacité d'une part de l'acétylation mais aussi de sa réaction opposée : la désacétylation réalisée par les désacétylases d'histones (HDAC). Il est généralement admis que l'hypoacétylation des histones est reliée à la répression de la transcription mais quelques données génétiques suggèrent un rôle également dans l'activation transcriptionnelle (Wang *et al.*, 2002). De plus, tout récemment il a été montré que les HDAC pouvaient directement activer la transcription en désacétylant certains promoteurs (De Nadal *et al.*, 2004).

Les HDAC peuvent être divisées en trois classes : classe I, II et III. Les classes I et II sont sensibles à l'inhibiteur de HDAC nommé trichostatin A (TSA) alors que la classe III ne l'est pas. Les classes I et II sont définies par rapport à l'homologie existant entre deux HDAC de la levure *S. cerevisiae* : Rpd3 (**R**educed **P**otassium **D**ependency **3**) pour la classe I et Hda1 pour la classe II. La classe II est exprimée de façon spécifique à certains tissus et certaines de ces HDAC se retrouvent dans le cytoplasme ainsi que dans le noyau. D'autres possèdent un rôle dans la différenciation cellulaire ainsi que dans le développement (Verdin *et al.*, 2003). La classe III est définie par l'homologie de séquence avec Sir2 (**S**ilent **I**nformation **R**egulator), une histone désacétylase de la levure *S. cerevisiae* dépendante du NAD<sup>+</sup>. La protéine Sir2 est intégrée dans le complexe SIR impliqué dans la formation de régions chromatiniennes silencieuses transcriptionnellement et hypoacétylées sur les histones H3 et H4. Les homologues de Sir2 se retrouvent dans tous les organismes, de la levure aux humains, et dans tous les cas ces enzymes ont été impliquées dans la répression transcriptionnelle (Blander et Guarente, 2004). Le complexe SIR possède un rôle très particulier dans l'établissement de la frontière entre l'hétérochromatine et l'euchromatine.

La classe I est nucléaire et possède de nombreuses activités (Gray et Ekstrom, 2001) associées à une grande variété de complexes qui présentent parfois une taille globulaire supérieur au million de Dalton (Grunstein, 1997; Zhang et Reinberg 2001). Par exemple, Rpd3, est associée à un complexe avec le co-répresseur global de la transcription Sin3 (Kadosh et Struhl, 1997; Zhang *et al.*, 1997) mais aussi dans le complexe de remodelage de la chromatine dépendant de l'ATP, NuRD (Zhang *et al.*, 1998). Tous ces complexes peuvent interagir avec un grand nombre de protéines impliquées dans la régulation de la transcription (Kurdistani et Grunstein, 2003).

Chez l'Homme les complexes NRD (NuRD/Mi2), BHC et SIN3 possèdent un noyau protéique commun composé des enzymes HDAC1 et HDAC2 (Knoepfler et Eisenman ; Ahringer, 2000, 1999 ; Vermaak *et al.*, 1999 ; Chen *et al.*, 2001 ; You *et al.*, 2001 ; Humphrey *et al.*, 2001 ; Hakimi *et al.*, 2002). NRD et SIN3 contiennent également protéines RbAp46 et RbAp48 (*Retinoblastoma Associated proteins*), qui se lient au *core* histone *via* leur domaine WD40. NRD se distingue par la présence en son sein des protéines MBD3 (*Methyl-CpG-Binding-Domain protein 3*) liant l'ADN méthylé, la protéine MTA2 (*Metastasis Associated protein 2*) surexprimée dans les cellules tumorales métastatiques et la protéine Mi2 $\alpha/\beta$  (CHD3/4) qui est un facteur de remodelage ATP dépendant (Zhang *et al.*, 1999 ; Zhang *et al.*, 1998). BHC contient en outre BHC110/LSD1, une histone déméthylase spécifique des résidus lysine (You *et al.*, 2001 ; Humphrey *et al.*, 2001 ; Hakimi *et al.*, 2002 ; Shi *et al.*, 2004, 2005 ; Metzger *et al.*, 2005; Lee *et al.*, 2005).

Le degré de complexité augmente *in vivo* avec la possibilité pour un même facteur de transcription de recruter différents complexes répresseurs. Ainsi, le facteur de transcription Ikaros, impliqué dans le développement des cellules lymphocytaires et l'activation des cellules immunitaires, interagit avec SIN3 et NRD (Knoepfler et Eisenman, 1999). Autre exemple, le facteur de transcription REST recrute SIN3 et BHC à proximité des gènes neuronaux des cellules non neuronales, (Andres *et al.*, 1999 ; Roopra *et al.*, 2000 ; Hakimi *et al.*, 2002).

Les enzymes HDAC contribuent fortement aux processus de prolifération et de différenciation des cellules eucaryotes; où l'activité de déacétylation joue un rôle central dans la répression spécifique des gènes. Des études *in vitro* et *in vivo* ont montré qu'un grand nombre d'inhibiteurs d'enzymes HDAC (HDACi) induisent l'arrêt de croissance, la différenciation et l'apoptose des cellules tumorales, par l'altération sélective de la transcription d'environ 2% des gènes exprimés (Marks *et al.*, 2001; Van Lint *et al.*, 1996). Des inhibiteurs spécifiques comme CI-994, Pyroxamide ou encore SAHA sont d'ores et déjà testés en essais cliniques pour faire régresser certaines tumeurs. De plus certains de ces inhibiteurs ont un pouvoir anti-protozoaire.

L'exemple le plus connu est le métabolite fongique, Apicidin qui provoque l'hyperacétylation des histones par inhibition des enzymes HDAC (Colletti *et al.*, 2001 ; Meinke *et al.*, 2000 ; Kim *et al.*, 2000). Isolé par l'entreprise Merck à partir du champignon *Fusarium pallidoroseum*, l'Apicidin exhibe une activité anti-protozoaire à large spectre, au sein du phylum des Apicomplexes (Darkin-Rattray *et al.*, 1996). Les mécanismes moléculaires en jeu (balance acétylation/déacétylation) sont peu voire pas documentés dans le phylum des apicomplexes. Une seule étude rapporte une caractérisation succincte de l'enzyme PfHDAC1 chez *Plasmodium falciparum* (Joshi *et al.*, 1999). Une autre étude récente montre que la transcription mutuellement exclusive des gènes *var* chez *Plasmodium falciparum* est liée à la transformation dynamique de la chromatine, via la déacétylation médiée par la protéine parasitaire PfSir2, une enzyme de classe III insensible aux inhibiteurs cités ci-dessus (Freitas-Junior *et al.*, 2005). La modulation de la balance acétylation/déacétylation apparaît donc comme un nouveau moyen d'action contre l'invasion, voire la différenciation parasitaire. Reste à en définir les limites moléculaires en recherchant les acteurs impliqués et leur mécanisme d'action. Dans la partie « résultats » nous présentons les premières données sur ce sujet chez *T. gondii*.

### 2.2.2 Méthylation des Histones Et Transcription

Les méthyltransférases d'histones (HMT) sont des enzymes capables de réaliser le transfert d'un groupement méthyle à partir d'une molécule donneuse, le SAM (S-adenosyl-méthionine ou AdoMet), sur la chaîne latérale d'une lysine ou d'une



arginine. Les arginines peuvent être mono- et di-méthylées de façon symétrique ou asymétrique alors que les lysines peuvent être mono-, di- ou tri-méthylées. Les histones méthyltransférases (HMT) sont classées selon l'acide aminé modifié. Les PRMT (Protein Arginine Methyltransferase) peuvent méthyler les arginines tandis que les protéines à domaine SET (Su(var)3-9, Enhancer of Zeste, Trithorax), peuvent méthyler les lysines (Zhang et Reinberg, 2001). Les lysines et les arginines ciblées sont en majorité situées sur les extrémités N-terminales des histones H3 et H4 (Figure C). Dot1, une HMT d'une nouvelle classe, est cependant responsable de la méthylation d'une lysine située à l'intérieur du domaine globulaire de l'histone H3 nucléosomale (Khan et Hampsey, 2002; Lacoste *et al.*, 2002).

La méthylation des histones a été largement reliée à la répression transcriptionnelle ainsi qu'à la régulation épigénétique mais il est évident qu'elle est aussi reliée à l'activation transcriptionnelle. Cette dualité de rôle est flagrante pour la méthylation de K4 et de K9 sur H3; dans le premier cas la marque est associée à l'euchromatine alors que la deuxième est associée à l'hétérochromatine, ces deux marques étant mutuellement exclusives (Noma *et al.*, 2001; Zhang et Reinberg, 2001; Jenuwein et Allis, 2001).

Il a aussi été montré que le nombre de groupements méthyls présents sur une même lysine peut mener à différents types de réponses, augmentant ainsi la complexité du code des histones (Santos-Rosa *et al.*, 2002). La méthylation des arginines semble avoir un effet positif sur la transcription (Zhang et Reinberg, 2001). Les preuves que la méthylation est une modification réversible ont été publiées que très récemment, avec la découverte de BHC110/LSD1, une histone déméthylase spécifique des résidus lysine (Shi *et al.*, 2004, 2005 ; Metzger *et al.*, 2005; Lee *et al.*, 2005) et PAD4 qui intervient sur les résidus arginine en convertissant le monométhylarginine en citrulline (Wang *et al.*, 2004).

### 2.2.2.a Les Histone-Lysine-Méthyltransférases (HMT/SET)

Les HMT à domaine SET (Su(var)3-9, Enhancer of Zeste, Trithorax) possèdent dans la plupart des cas une activité histone méthyltransférase (Jenuwein, 2001;

Pijnappel *et al.*, 2001). Le domaine SET caractérise une superfamille de protéines capables de transférer des résidus méthyl- en position  $\epsilon$ -NH<sub>2</sub> de certaines lysines N-terminales des histones, en altérant ainsi la conformation de la chromatine. En fonction de la famille et du résidu méthylé, l'action sera inhibitrice ou activatrice vis à vis de la transcription. Le domaine SET se trouve plutôt vers l'extrémité C-terminale et est entouré par des séquences pre-SET et post-SET présentant un certain nombre de résidus cystéine invariants, en particulier les 9 résidus du domaine pre-SET qui s'organisent en un triangle équilatéral autour de trois ions Zn. Le domaine SET forme un pli  $\beta$  uniquement observé dans ces méthyl-transférases.

La première HMT de cette famille ayant été caractérisée était homologue à une protéine de drosophile, Su(var)3-9. Celle-ci méthyle spécifiquement la lysine 9 de H3, cette marque étant par la suite reconnue par HP1, une protéine très importante pour la formation et le maintien de l'hétérochromatine (Lachner *et al.*, 2001). Sur la lysine 9 de H3, acétylation et méthylation semblent être mutuellement exclusives. Ainsi, chez la drosophile, l'histone H3 méthyltransférase Suv39h est physiquement et fonctionnellement associée à une histone deacétylase, ce qui suggère un *mécanisme moléculaire unique*, permettant de convertir directement une lysine 9 acétylée en une lysine 9 méthylée. La méthylation de H3-K9 crée en outre un site de haute affinité pour la protéine de l'hétérochromatine HP1. Su(var)3-9 est également reliée à la régulation de la transcription, la mémoire génomique (genomic imprinting) et la régulation épigénétique (Lachner and Jenuwein, 2002). Enfin, chez le champignon *Neurospora crassa* ainsi que chez *Arabidopsis thaliana*, il a été montré que la méthylation de H3-K9 pouvait entraîner la méthylation de l'AND (Jackson *et al.*, 2002; Tamaru et Selker, 2001).

Chez la levure *S. cerevisiae*, il existe 7 protéines de cette famille et elles sont généralement impliquées dans la régulation transcriptionnelle (Pijnappel *et al.*, 2001; Santos-Rosa *et al.*, 2002; Xiao *et al.*, 2003). Set1 fait partie d'un complexe multiprotéique nommé COMPASS contenant 8 protéines (Krogan *et al.*, 2002; Miller *et al.*, 2001; Nagy *et al.*, 2002; Roguev *et al.*, 2001). Set1 est capable de mono- di- ou tri-méthyliser spécifiquement la lysine 4 de l'histone H3. En effet, la lysine 4 de l'histone 3 (H3-K4) est, chez *Saccharomyces cerevisiae*, diméthylée dans les régions

euchromatiques (exprimées) comme hétérochromatiques (réprimées), et est triméthylée dans les seules régions euchromatiques (Santos-Rosa *et al.*, 2002).

Parmi les protéines à domaine SET voici quelques unes dont la fonction est partiellement caractérisée chez les humains. G9a, méthyle principalement la lysine 9 mais également la lysine 27 de H3. Cette enzyme participe à la régulation transcriptionnelle et au développement mais, plus important encore, elle est la HMT majeure capable de modifier la lysine 9 de H3 dans les régions euchromatiniennes (Tachibana *et al.*, 2002). PR-SET7 méthyle la lysine 20 de H4 et est présente dans les domaines non actifs transcriptionnellement de l'euchromatine (Nishioka *et al.*, 2002).

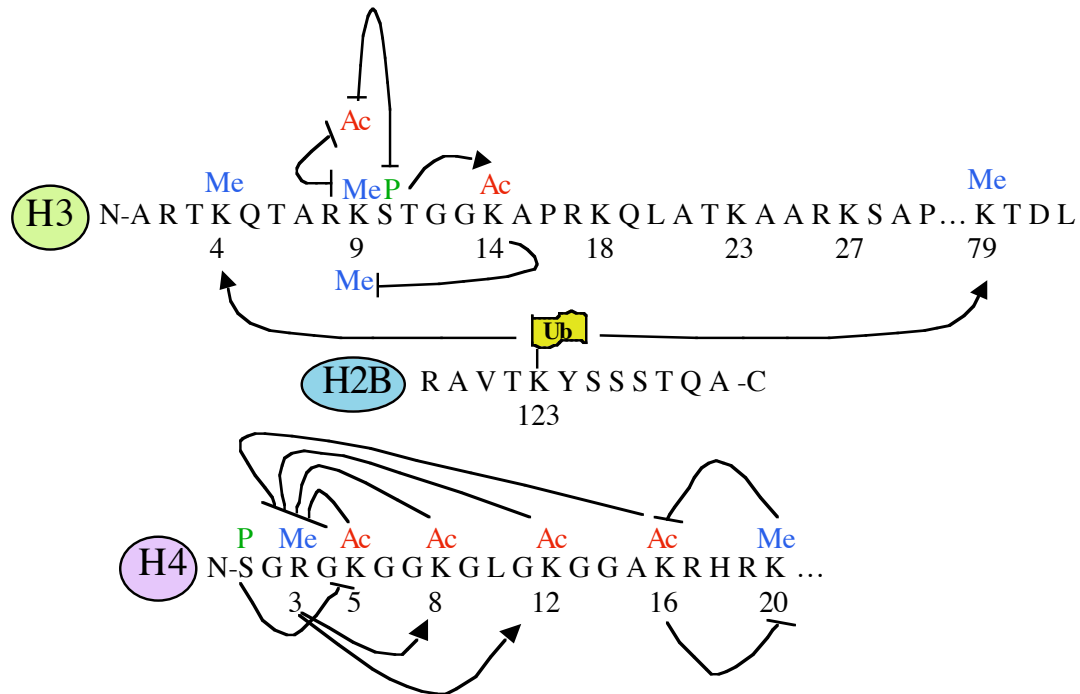
### 2.2.2.b Les Histone-Arginine-Méthyltransférases (PRMT)

Sept PRMT (PRMT1-7) ont été identifiées chez l'humain mais seules PRMT1 et PRMT4 (également appelée CARM1, pour *co-activator-associated arginine methyltransferase 1*) présentent une activité histone méthyltransférase *in vivo*. PRMT1 (Rmt1 chez la levure *S. cerevisiae*) méthyle l'arginine 3 de H4 et est très conservée au cours de l'évolution. Elle possède un rôle dans l'activation transcriptionnelle (Wang *et al.*, 2001). PRMT4/CARM1 méthyle les arginines 2, 17 et 26 de H3 ainsi que d'autres protéines non reliées aux histones comme CBP, sa méthylation inhibant son recrutement sur les différents promoteurs qu'il contrôle (Xu *et al.*, 2001; Chen *et al.*, 1999; Schurter *et al.*, 2001). Dans le cadre de l'activation de gènes par p53, il y a recrutement de PRMT1 suivi un peu plus tard de CARM1 laissant supposer que ces deux protéines possèdent un rôle prédominant dans le contrôle du cycle cellulaire ainsi que dans l'apoptose (An *et al.*, 2004).

## 2.3 LE CODE HISTONE

Il existe une interdépendance entre les différentes modifications covalentes qui ont lieu sur des résidus spécifiques et adjacents des histones. Une combinaison de modifications donnée entraînerait alors l'activation ou la répression de l'expression de gènes. Le cas de l'histone H3 est en effet éloquent. Ses résidus potentiellement modifiés sont les Lys4, 9, 14, 18, 23 et 27, la Ser10 et les Arg2, 17 et 26 (Figure D).

L'acétylation de la Lys14 (qui caractérise un état transcriptionnellement inactif) précède la méthylation de la Lys9, mais est elle-même précédée par (et dépendante de) la phosphorylation de la Ser10.



**Figure D. Interrelations régulatrices entre les différentes modifications des histones. A)** La phosphorylation de la sérine 10 de l'histone H3 stimule l'acétylation de la lysine 14. Ces deux modifications répriment alors la méthylation de la lysine 9. Cette méthylation réprime quant à elle la phosphorylation de la sérine 10. Chez la levure *S. cerevisiae*, la mono-ubiquitination de la lysine 123 de l'histone H2B est prérequisse à la méthylation des lysines 4 et 79 de l'histone H3. **B)** La méthylation de l'arginine 3 de l'histone H4 stimule l'acétylation des lysines 8 et 12 de cette même histone. Par contre l'acétylation des résidus 5, 8, 12 et 16 inhibe la méthylation de l'arginine 3. La phosphorylation de la sérine 1 de H4 inhibe l'acétylation de la lysine 5, une inhibition bloquée par la méthylation de l'arginine 3. L'acétylation de la lysine 16 semble inhibée par la méthylation de la lysine 20 et inversement.

Au contraire, la méthylation de la Lys9 est inhibée par la phosphorylation préalable de la Ser10 adjacente. Ceci suggère que la reconnaissance du résidu Ser10 phosphorylé par l'HAT puisse guider et stabiliser son interaction avec la queue de l'histone, menant ainsi à l'acétylation d'un résidu Lys spécifique. Enfin très récemment il a été montré que, chez la levure, Rad6, une enzyme permettant le transfert de l'ubiquitine, servait de médiateur à la méthylation de la lysine 4 de l'histone H3 en permettant l'ubiquitination de la lysine 123 de l'histone H2B. L'utilisation de mutants a permis de vérifier que ce schéma était unidirectionnel et de montrer l'importance de ces modifications dans la répression de la transcription (Sun and Allis, 2002). Un autre exemple est la méthylation de l'Arg3 de l'histone H4 qui précède et favorise

l'acétylation des Lys8 et Lys12 mais qui est inhibée par l'acétylation préalable de l'histone H4.

Ces connexions définissent un code, appelé 'code histone', qui tient compte à la fois de la nature de ces modifications et de leur chronologie (Jenuwein et Allis, 2001). Ce code émerge comme faisant partie intégrante de la voie de régulation de la transcription qui assure le remodelage de la chromatine autour d'un promoteur spécifique en réponse à un besoin.

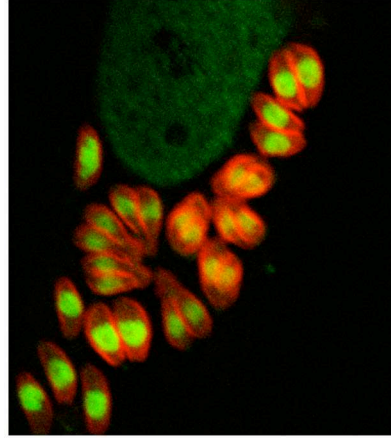
Il faut enfin noter que l'ensemble de ces modifications décrites ci-dessus ne sont qu'une des étapes préparant la chromatine à accueillir le complexe de pré-initiation au niveau des promoteurs. Des études restent à faire pour déterminer une ou des séquence(s) de modifications spécifiques aboutissant à l'activation ou la répression de la transcription et/ou si ces modifications sont spécifiques du gène et/ou du promoteur. En outre, de même que pour les facteurs de remodelage de la chromatine du type complexe SWI/SNF et HAT, il semble important d'arriver à déterminer par quel type de facteur et comment les enzymes de modifications sont recrutées au niveau du promoteur et quelles sont leurs spécificités.

Il est maintenant clairement admis que la chromatine est un élément essentiel dans la régulation de la transcription et ne peut-être considérée comme le simple support porteur de l'information génétique. Sa structure est étroitement régulée et modulée au cours de la transcription. L'ensemble des modifications post-traductionnelles existantes est utilisé pour permettre la régulation structurale mais aussi le recrutement de nouveaux facteurs qui vont agir en synergie pour conduire la cellule à exprimer/réprimer certains gènes. Il ne reste qu'un pas à franchir pour appliquer ce mode de régulation à la différenciation parasitaire, en particulier l'interconversion, chez *T.gondii*.

Dans ce manuscrit de thèse, l'accent a été mis sur l'analyse des modifications des histones de *T.gondii* pour évaluer leur impact sur la répression ou l'activation des gènes impliqués dans l'interconversion. L'hypothèse retenue est alors que la conversion d'une forme à l'autre reposerait alors sur l'équilibre entre les grandes

marques épigénétiques. Dans un premier temps, la cinétique et les inter-dépendances des modifications (acétylation et méthylation), éléments fondateurs du code histone, ont été analysées durant le passage d'une forme à l'autre. Le deuxième grand volet du manuscrit concerne les enzymes, qui catalysent les modifications et qui participent potentiellement en cela au contrôle de l'interconversion. L'étude de deux grandes familles d'enzyme a été privilégiée: les arginine-méthyltransférases (PRMT) et les histones déacétylases (HDAC) et succinctement les acétylases de type GCN5. Nous soulignons dans ce manuscrit le «dialogue moléculaire» entre TgHDAC3 et TgGCN5 qu'ainsi qu'entre TgCARM1 et TgGCN5 et les conséquences de leur enzymologie sur la différenciation. La partie «résultats» est divisée en deux parties: la première est constituée par le papier Saksouk *et al.*, (2005) publié dans *Molecular and Cellular Biology*; la seconde apporte les premières pistes de travail sur une nouvelle histone déacétylase du Toxoplasme, TgHDAC2.

# RESULTATS



# CHAPITRE I



### Avant propos

Le chapitre 1 contient un résumé succinct des résultats publiés dans *Molecular and Cellular Biology* (Saksouk *et al.*, 2005) suivi de l'article et des données supplémentaires.

### 1. Première lecture du code histone

Au départ du projet il existait peu d'informations sur les histones et leurs modifications post-traductionnelles chez *T. gondii*. Une carte des principales modifications post-traductionnelles des histones a pu être dressée par l'analyse biochimique (chromatographie et spectrométrie de masse) et immunologique des histones de *T.gondii* et de leurs variants modifiés. Les difficultés à isoler suffisamment de kystes de cerveaux de souris ne nous ont pas permis à ce jour d'obtenir par spectrométrie de masse une carte différentielle (tachyzoïte *versus* bradyzoïte) exhaustive des isoformes modifiés.

Ce type d'approche n'est finalement pas suffisamment résolutif pour décrire les cinétiques (mise en place et *turn-over*), les interdépendances inter-et intra-histone ou encore les localisations génomiques d'une modification donnée. La méthode d'immunoprécipitation de la chromatine (ChIP, Orlando *et al.*, 2000) a montré ces dernières années qu'il était possible de décrire une modification donnée dans son contexte génomique, en déterminant *in vivo* les adresses chromosomiques de protéines ou de modifications post-traductionnelles d'intérêts.

Nous avons développé et publié récemment la première application chez *T.gondii* de la technique de Chromatine Immunoprecipitation couplée à une marche à haute résolution sur chromosome par PCR (Scanning ChIP). En résumé, nous avons analysé la dynamique de l'acétylation et de la méthylation des histones H3 et H4 au cours du passage tachyzoïte-bradyzoïte (Saksouk *et al.*, 2005, Figure 1).

Nos premiers résultats révèlent que les marques épigénétiques de l'activation ou de la répression transcriptionnelle connues chez la levure ou l'Homme sont conservées dans le Toxoplasme. Par exemple la méthylation de l'arginine 17 sur l'histone H3, marque de l'activation, est présente exclusivement à proximité de gènes ou de locus tachyzoïte-spécifiques dans un contexte cellulaire tachyzoïte et inversement pour la forme bradyzoïte (Saksouk *et al.*, 2005, Figure 6A). Nous montrons également que le phénomène de “*long-range acetylation*” (Vogelauer *et al.*, 2000) contrôle chez *T.gondii* la transcription stade-dépendant de gènes organisés en tandem (Saksouk *et al.*, 2005, Figure 1D).

Nos résultats préliminaires de SChIP soulignent l'influence du « code histone » sur la différenciation parasitaire. Déséquilibrer la balance épigénétique en influant chimiquement ou génétiquement sur les enzymes en charge de l'écriture et de la lecture du code peut aider à mieux comprendre, voire à moduler le processus d'interconversion.

Les enzymes de modification des queues d'histone sont souvent intégrées dans des grands complexes protéiques avec une taille globulaire souvent supérieur au million de Dalton. La détermination de leur activité enzymatique native nécessite souvent la purification biochimique homogène des complexes qui les contiennent. Nous avons concentré nos efforts à l'analyse des deux modifications majeures que sont l'acétylation et la méthylation des histones et des enzymes qui les catalysent: les arginine-méthyltransférases (PRMT), les acétylases (HAT) et les histones déacétylases (HDAC).

Par une approche bioinformatique, nous avons systématiquement cherché dans la banque de données TOXODB (<http://www.toxodb.org>) les gènes codant pour les enzymes putatives catalysant la déacétylation et la méthylation des histones.

Sur le plan biochimique notre équipe a développé une plate-forme à haut-débit d'analyse systématique des interactomes associés aux enzymes étudiées avec (i) l'obtention de toxoplasmes transgéniques exprimant ectopiquement des versions étiquetées des enzymes d'intérêt, (ii) le développement de protocole de chromatographie

conventionnelle (FPLC) et d'affinité pour la purification des complexes associés à ces enzymes et enfin iii) le micro-séquençage des protéines en MS/MS (Coll. Jérôme Garin). Des études enzymologiques avec la détermination du substrat histone et des résidus modifiés complète la procédure.

### **2. Analyse des histones déacétylases chez *T. gondii***

À ce jour, nous avons identifié et cloné cinq enzymes HDAC chez le parasite (TgHDAC1-5). Nous avons obtenu des toxoplasmes transgéniques stables produisant de manière ectopique les protéines TgHDAC2, TgHDAC3, TgHDAC4 (variants 1 et 2) et TgHDAC5 en fusion avec une étiquette HAFlag. L'interactome des enzymes TgHDAC4 (v1 et v2) et TgHDAC5 est encore à l'étude; les résultats de ce travail encore sommaire ne seront pas présentés dans ce manuscrit. Les résultats sur l'histone déacétylase TgHDAC2 sont présentés dans le Chapitre 2.

Concernant les résultats publiés nous avons déterminé l'interactome spécifique de TgHDAC3, une enzyme exclusivement nucléaire. Son complexe associé TgCRC (Co-Repressor Complex) présente une activité enzymatique de déacétylation des histones qui est sensible aux principaux inhibiteurs HDACi (Saksouk *et al.*, 2005, Figures 2 et 3). Le complexe TgCRC contient une protéine homologue à TBL1, une composante majeure des complexes humains N-CoR (nuclear receptor corepressor) and SMRT (silencing mediator of retinoid and thyroid hormone receptors). De plus, nous avons montré *in vivo* par CHIP que l'enzyme est localisée à des *loci* hypoacétylé (Saksouk *et al.*, 2005, Figure 4). En collaboration avec l'équipe du Dr. Sullivan W, on a aussi pu montrer que l'acétylase TgGCN5 antagonise le rôle de TgHDAC3 sur plusieurs *loci* (Saksouk *et al.*, 2005, Figure 4).

### **3. Analyse des histones arginines méthyltransférases du Toxoplasme**

Nous avons identifié deux arginine-méthyltransférases chez *T.gondii*, TgCARM1 et TgPRMT1, qui catalysent *in vitro*, respectivement, la méthylation des arginine

H3(R17) et H4(R3) (Saksouk *et al.*, 2005, Figure 5). Par ChIP, nous avons montré *in vivo* que la protéine TgCARM1 et la modification H3(R17)<sup>methyl</sup> sont co-localisées exclusivement à proximité de gène tachyzoïte-spécifique, qui sont actifs dans un contexte tachyzoïte. Et on observe dans une population de bradyzoïte la co-localisation de l'enzyme et de sa modification à un locus bradyzoïte (Saksouk *et al.*, 2005, Figure 6). Une mutagénèse chimique a été réalisée avec AMI-1, une petite molécule qui inhibe spécifiquement l'activité arginine méthyltransférase des cellules humaines (Chen *et al.*, 2004). *In vitro* AMI-1 inhibe la méthylation de l'histone H3 par TgCARM1 sur le résidu 17 mais agit faiblement sur la méthylation de l'histone H4 médiée par TgPRMT1 (Saksouk *et al.*, 2005, Figure 7). *In vivo* AMI-1 induit une diminution du taux d'occupation à certains *loci* de la modification H3(R17)<sup>methyl</sup> et provoque un enkystement de la souche Prugniaud en culture cellulaire (Saksouk *et al.*, 2005, Figure 7). Nous pensons que AMI-1 en diminuant l'activité de TgCARM1 influence négativement sur la transcription d'un facteur X avec pour conséquence la conversion des tachyzoïtes en bradyzoïtes.

D'autres résultats sur la relation structure/fonction de la protéine TgCARM1 ont été obtenus mais non publiés. Ainsi, l'enzyme FLAG-TgCARM1, purifiée à partir de toxoplasmes transgéniques, se présente sous forme d'un hexamère sans partenaires protéiques apparemment associé. En collaboration avec le Dr. Hassan Berhali (EMBL Station, Grenoble), nous tentons actuellement de résoudre la structure de TgCARM1 par biocristallographie aux rayons X (ESFR *synchrotron* de Grenoble). La difficulté à obtenir des cristaux est due à la multimérisation de la protéine TgCARM1 sous la forme d'hexamère et en quantité plus faible sous forme de dimère (Saksouk *et al.*, 2005, Figure 5E). D'après la littérature, la cristallisation des protéines Hmt1, PRMT1 et PRMT3 ont montré que ses protéines n'étaient actives que sous forme d'homodimère (Zhang *et al.*, 2000 ; Zhang et Cheng 2003). En solution Hmt1 et PRMT1 forme un hexamère (Weiss *et al.*, 2000 ; Tang *et al.*, 1998) tandis que PRMT5 forme un homo-oligomère (Rho *et al.*, 2001).

# Histone-Modifying Complexes Regulate Gene Expression Pertinent to the Differentiation of the Protozoan Parasite *Toxoplasma gondii*†

Nehmé Saksouk,<sup>1</sup> Micah M. Bhatti,<sup>2</sup> Sylvie Kieffer,<sup>3</sup> Aaron T. Smith,<sup>2</sup> Karine Musset,<sup>1</sup>  
Jérôme Garin,<sup>3</sup> William J. Sullivan, Jr.,<sup>2</sup> Marie-France Cesbron-Delauw,<sup>1</sup>  
and Mohamed-Ali Hakimi<sup>1\*</sup>

UMR5163/CNRS-UJF, Jean-Roget Institute,<sup>1</sup> and INSERM ERIT, M 0201, CEA,<sup>3</sup> Grenoble, France, and Department of Pharmacology and Toxicology, Indiana University School of Medicine, Indianapolis, Indiana<sup>2</sup>

Received 11 June 2005/Returned for modification 25 August 2005/Accepted 14 September 2005

**Pathogenic apicomplexan parasites like *Toxoplasma* and *Plasmodium* (malaria) have complex life cycles consisting of multiple stages. The ability to differentiate from one stage to another requires dramatic transcriptional changes, yet there is a paucity of transcription factors in these protozoa. In contrast, we show here that *Toxoplasma* possesses extensive chromatin remodeling machinery that modulates gene expression relevant to differentiation. We find that, as in other eukaryotes, histone acetylation and arginine methylation are marks of gene activation in *Toxoplasma*. We have identified mediators of these histone modifications, as well as a histone deacetylase (HDAC), and correlate their presence at target promoters in a stage-specific manner. We purified the first HDAC complex from apicomplexans, which contains novel components in addition to others previously reported in eukaryotes. A *Toxoplasma* orthologue of the arginine methyltransferase CARM1 appears to work in concert with the acetylase TgGCN5, which exhibits an unusual bias for H3 [K18] *in vitro*. Inhibition of TgCARM1 induces differentiation, showing that the parasite life cycle can be manipulated by interfering with epigenetic machinery. This may lead to new approaches for therapy against protozoal diseases and highlights *Toxoplasma* as an informative model to study the evolution of epigenetics in eukaryotic cells.**

The phylum *Apicomplexa* is composed of a number of protozoan parasites of immense clinical and economic importance, including *Plasmodium*, the causative agent of malaria, and *Toxoplasma gondii*. *Toxoplasma* has long been a major medical and veterinary problem, capable of causing abortion or congenital birth defects in both humans and livestock. The advent of AIDS has drawn even more attention to *Toxoplasma* as a serious opportunistic pathogen.

*Toxoplasma* is incurable because of its ability to differentiate from the rapidly replicating tachyzoite stage into a latent cyst form (bradyzoite stage) that is impervious to immunity and current drugs. This developmental process, called stage conversion, is triggered by the host immune response. Impairment of the immune response in immunocompromised patients initiates the reconversion of bradyzoite cysts into cytolytic tachyzoites. Stage conversion is associated with dramatic morphological and physiological changes, including alteration of the parasite surface, changes in parasite metabolism, and induction of genes associated with stress response (12, 17, 34, 57). Differentiation is clearly regulated in part at the transcriptional level (37, 40, 49), but the control of gene expression in apicomplexan parasites is profoundly understudied.

Analysis of apicomplexan genome databases for *Plasmodium* and *Cryptosporidium* reveals a relatively small number of transcription factors (54). Review of the *Toxoplasma* genome database reinforces the paucity of conventional transcription

factors in these parasites (M. A. Hakimi, unpublished data). The proportion of the transcriptome encoding transcription factors appears to increase as the complexity of the organism increases (3.4% for *Saccharomyces cerevisiae*, 4.2% for *Caenorhabditis elegans*, 5.5% for *Drosophila melanogaster*, 8% for humans) (42). Thus, it has been proposed that the regulation of gene expression in apicomplexan parasites may require the concerted action of a well-developed chromatin-remodeling apparatus and a small set of specific DNA-binding proteins (54).

It is now well accepted that chromatin is dynamic and influences gene expression regulation. Chromatin can be fashioned to be transcriptionally repressive or permissive, dependent in part on posttranslational modifications made to histones, a phenomenon referred to as the histone code (29). In this study, we established chromatin immunoprecipitation (ChIP) techniques for *Toxoplasma* to investigate histone modifications in the context of stage conversion. For the first time, we show that acetylation and methylation of parasite histones correlate with stage-specific gene expression. In tachyzoites, a novel histone deacetylase (HDAC) corepressor complex (TgCRC) operates at bradyzoite-specific promoters while a TgGCN5 histone acetyltransferase (HAT) (53) operates at tachyzoite-specific promoters. Surprisingly, TgGCN5 exhibits an unexpected substrate preference and exclusively acetylates histone H3 lysine 18 (H3 [K18]). In humans, this modification is mediated by the metazoan HAT CBP/p300 and facilitates recruitment of an arginine methylase, CARM1, to activate transcription (11). Consequently, we characterized a novel CARM1 complex in *Toxoplasma* that targets H3 [R17] coincident with TgGCN5-mediated stage-specific gene activation. Collectively, the histone modifications associated with differentiation, along with

\* Corresponding author. Mailing address: ATIP-UMR5163-CNRS, Jean-Roget Institute, Domaine de la Merci, 38700 Grenoble, France. Phone: (33) 4 76 63 74 69. Fax: (33) 4 76 63 74 97. E-mail: Mohamed-Ali.Hakimi@ujf-grenoble.fr.

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





the mediators of those activities reported here, represent a substantial increase in our understanding of the control of gene expression in parasitic protozoa.

#### MATERIALS AND METHODS

**Parasite methods.** The Prugniaud and RH strains of *Toxoplasma* were grown in human foreskin fibroblasts, transfected, and cloned by limiting dilution as described previously (34). The RH strain lacking HXGPRT was used in selection experiments (14). In vitro bradyzoite induction procedures were performed by the high-pH (8.1) method (58). To determine the ratio of parasites expressing stage-specific antigens within the entire population, the differentiation process was monitored daily by double-immunofluorescence assay with both tachyzoite-specific  $\alpha$ -SAG1 and cyst-specific  $\alpha$ -CC2 (a kind gift from W. Bohne and U. Gross) antibodies. The ratio of parasites reacting with the antibodies (see Fig. 7B) was determined by counting 50 random fields under a fluorescence microscope.

**IFA and Western blot analysis.** Immunofluorescence assays (IFA) were performed as described previously (17). Immunoblotting with alkaline phosphatase was performed as previously described (26).

**Antibodies.** Primary antibodies for IFA included a rat monoclonal antibody against the hemagglutinin (HA) epitope tag used at 1:500 ( $\alpha$ -HA; Roche Diagnostic), a monoclonal antibody against BSR4 used at 1:300 (p36, Tg4A12; a kind gift from J. F. Dubremetz), a monoclonal antibody against cyst wall antigen used at 1:500 ( $\alpha$ -CC2; a kind gift from W. Bohne and U. Gross), and an anti-dimethyl [R17] H3 polyclonal antibody (ab8284; Abcam). An anti-dimethyl [R3] H3 polyclonal antibody (07-213; Upstate) was used for Western blot analysis. A ChIP grade anti-dimethyl [R17] H3 polyclonal antibody (ab8284; Abcam) was used for Western blot analysis and ChIP assays. Additional antibodies used in ChIP assays were polyclonal anti-acetyl [K9-K14] H3 (06-599; Upstate), anti-acetyl [K5-K8-K12-K16] H4 (06-866; Upstate), anti-acetyl [K18] H3 (ab1191; Abcam),  $\alpha$ -dimethyl [K4] H3 (07-030; Upstate),  $\alpha$ -trimethyl [K4] H3 (ab8580; Abcam), and  $\alpha$ -histone H3 (ab12079; Abcam). A rabbit antiserum was raised against recombinant TgCARM1 according to the protocol of the manufacturer (Eurogentec). A 1:3,000 dilution of polyclonal anti-TgCARM1 antibody ( $\alpha$ -49) was used to perform Western blot analysis, and 3  $\mu$ l of immunopurified  $\alpha$ -49 was used in each ChIP assay.

**ChIP.** ChIP was performed as described previously (26). Briefly, freshly lysed tachyzoites or in vitro bradyzoites were treated for 20 min with 1% formaldehyde at 37°C. The lysate was sonicated under conditions yielding fragments ranging from 400 bp to 800 bp. Samples were purified through PCR Purification Kit columns (QIAGEN) and used as a template in PCRs to detect specific targets. Primer pairs (melting temperature, 55 to 65°C) amplifying 200- to 450-bp fragments were created by using sequences found at ToxoDB.org (see Table S1 in the supplemental material) (33). For each primer pair, the optimal magnesium concentration (1 to 2.5 mM MgCl<sub>2</sub>) was determined. For individual primer pairs, the annealing temperature and number of cycles were adjusted until no signal was detected for the mock-immunoprecipitated DNA (IgG [immunoglobulin G]), but the amplification on the genomic template was not altered (input). Signals obtained with the antibody-immunoprecipitated DNA under these conditions were considered significant. The amplified DNA was separated on 2% agarose and stained with ethidium bromide.

**Affinity chromatographic purification of HA-FLAG-CARM1.** To purify HA-FLAG-TgCARM1, whole-cell extract (WCE) from transgenic parasite clone A3

(5  $\times$  10<sup>10</sup> extracellular tachyzoites) was incubated with anti-FLAG M2 affinity gel (Sigma). After extensive washing with buffer A (20 mM Tris-HCl [pH 7.9], 0.5 M KCl, 10% glycerol, 1 mM EDTA, 5 mM dithiothreitol [DTT], 0.5% NP-40), the affinity column was eluted with buffer A containing FLAG peptide (500  $\mu$ g/ml; Sigma) according to the manufacturer's instructions. HA-FLAG-TgCARM1-containing fractions were fractionated on a Superose 6 HR 10/30 (Pharmacia) equilibrated in 0.5 M KCl in buffer A containing 0.1% NP-40, 1  $\mu$ g/ml aprotinin, leupeptin, and 0.1 mM pepstatin phenylmethylsulfonyl fluoride (PMSF).

#### Chromatographic purification of TgCRC, a TgHDAC3-containing complex.

The TgHDAC3 complex was purified from 2 g of WCE from transgenic extracellular tachyzoites expressing TgHDAC3-HA-FLAG ectopically. WCE was loaded onto a 250-ml column of phosphocellulose (P11; Whatman) and fractionated stepwise by the indicated KCl concentrations in buffer A (20 mM Tris HCl [pH 7.9], 0.2 mM EDTA, 10 mM  $\beta$ -mercaptoethanol, 10% glycerol, 0.2 mM PMSF). The P11 0.3 M KCl fraction (750 mg) was loaded onto a 45-ml DEAE-Sephacel column (Pharmacia) and eluted with 0.35 M KCl. The 0.35 M KCl elution (420 mg) was incubated with 1 ml of anti-FLAG M2 affinity gel (Sigma) for 2 h at 4°C. Beads were washed with 50 ml of BC500 buffer (20 mM Tris [pH 8], 0.5 M KCl, 10% glycerol, 1 mM EDTA, 1 mM DTT, 0.1% NP-40, 0.5 mM PMSF, 1  $\mu$ g/ml each aprotinin, leupeptin, and pepstatin) and one time with 10 ml BC100 buffer (20 mM Tris [pH 8], 0.1 M KCl, 10% glycerol, 1 mM EDTA, 1 mM DTT, 0.1% NP-40, 1  $\mu$ g/ml each aprotinin, leupeptin, and pepstatin). Bound peptides were eluted stepwise with 350  $\mu$ g/ml of FLAG peptide (Sigma) diluted in BC100 buffer. Eluates enriched in TgHDAC3 were fractionated on a Superose 6 HR 10/30 (Pharmacia) and equilibrated in 0.5 M KCl in buffer A containing 0.1% NP-40 and 1  $\mu$ g/ml each aprotinin, leupeptin, and pepstatin. Our current purification protocol yields approximately 10 to 30  $\mu$ g of TgCRC complex from 120 mg of the transgenic parasite clone expressing TgHDAC3-HA-FLAG. This preparation is sufficiently clean that individual peptide bands can be excised and sequenced.

**Mass spectrometry peptide sequencing.** Protein bands were excised from colloidal blue-stained gels (Invitrogen), oxidized with 7% H<sub>2</sub>O<sub>2</sub>, and subjected to in-gel tryptic digestion. Peptides were extracted with a 5% (vol/vol) formic acid solution and acetonitrile and injected into a CapLC (Waters) nanoLC system that was directly coupled to a QTOF Ultima mass spectrometer (Waters) (18). Tandem mass spectrometry data were acquired and processed automatically by using MassLynx 3.5 software (Waters). Tandem mass spectra were searched against a compiled *T. gondii* database by using the MASCOT program (Matrix Sciences, London, United Kingdom) available at <http://www.osc.edu/hpc/software/apps/mascot.shtml>.

**Histone methyltransferase assays.** Recombinant proteins rTgCARM1 (pNS2) and rTgPRMT1 (pMAH2) were purified from bacteria as described previously (26). Protein samples (0.25  $\mu$ g of rTgCARM1, 0.2  $\mu$ g of HA-FLAG-TgCARM1 elution, and 0.25  $\mu$ g of rTgPRMT1) were incubated with recombinant histones or free core histones at 30°C for 60 min in phosphate-buffered saline reaction buffer containing 1 mM DTT and 0.2  $\mu$ g/ $\mu$ l of *S*-adenosyl-L-methionine (Sigma) as previously described (48). Assays were performed with 3  $\mu$ g of free core histones purified from chicken erythrocytes (13-107; Upstate) or from tachyzoites, 1.5  $\mu$ g of recombinant human histone H3 expressed in *Escherichia coli* (14-494; Upstate), or 1  $\mu$ g of purified glutathione *S*-transferase (GST)-histone TgH4 and 1  $\mu$ g of mutated GST-histone TgH4 (R3K). The reactions were stopped by the addition of sodium dodecyl sulfate (SDS) sample buffer. Reaction products were then separated on 15% SDS-polyacrylamide gels and transferred onto PVDF membrane. Immunoblot assays were performed by using

FIG. 1. Histone acetylation is a mark of gene activation during stage conversion. (A) Tachyzoite (T)-to-bradyzoite (B) conversion of *Toxoplasma*. The Prugniaud strain was induced in vitro with an alkaline pH. The kinetics of stage conversion was monitored by immunofluorescence assay with bradyzoite-specific antibodies ( $\alpha$ -CC2 and  $\alpha$ -BRS4) and tachyzoite-specific antibody  $\alpha$ -SAG1. Note that in vitro conversion to bradyzoites is not 100% efficient, so our converted parasites are labeled B<sup>T</sup> to denote parasite populations greatly enriched for bradyzoites (70 to 80%). (B) Stage-specific and constitutively expressed genes examined by ChIP analysis. Genes: *BAG*, bradyzoite antigen; *BSR*, bradyzoite-specific recombinant; *LDH*, lactate dehydrogenase; *SAG*, surface antigen; *SRS*, SAG-related sequences; *DEOC*, deoxyribose-phosphate aldolase; *GRA*, dense granule protein; *eIF4A*, eukaryotic translation initiation factor 4A; *DHFR*, dihydrofolate reductase-thymidylate synthase. (C) ChIP analysis of housekeeping genes (*eIF4A* and *DHFR*), tachyzoite-specific genes (*SAG1* and *SAG2A*), and bradyzoite-specific genes (*DEOC*, *LDH2*, and *BAG1*). Tachyzoites (T) and in vitro-generated bradyzoites (B<sup>T</sup>) were processed for ChIP analysis with  $\alpha$ -acetyl [K9-K14] histone H3 and  $\alpha$ -acetyl [K5-K8-K12-K16] histone H4 and  $\alpha$ -histone H3 antibodies. A chromatin sample with an irrelevant antibody was used as a negative control (IgG) and genomic DNA as a positive control for PCR (input). The input PCR signal was set at 100%, and the numerical value of the ChIP signal represents the percentage of the input. The amount of immunoprecipitated DNA determined by semiquantitative PCR was normalized to the respective input DNA for each sample (shown in arbitrary units). Each bar is an average of three independent experiments. (D) Long-range ChIP analysis of the *BSR4* locus. (E) ChIP analysis of the *GRA* gene family (*GRA1* to *GRA9*).

rabbit antisera as follows: polyclonal antibody anti-dimethyl [R17] H3 (ab8284; Abcam) at 1:1,000 and polyclonal antibody anti-dimethyl [R3] H4 (07-213; Upstate) at 1:1,000. Colorimetric detection with alkaline phosphatase-coupled anti-rabbit IgG (Promega) was performed. Methylation reactions with AMI-1 or AMA-2 were performed as previously described (9).

**HAT and HDAC assays.** The HAT assay procedure used was described previously (16). HDAC activity was measured with the colorimetric HDAC activity assay kit (Calbiochem-Novabiochem, Darmstadt, Germany).

## RESULTS

**Histone acetylation accompanies stage conversion.** Histone acetylation of the promoter precedes the activation of many genes and is thought to establish a chromatin environment suitable for assembly of the transcriptional complex. We established a ChIP (55) assay for *Toxoplasma* to monitor the status of histone modifications in the promoters of both stage-specific and constitutively expressed genes for tachyzoites and in vitro-generated bradyzoites (Fig. 1A). We designed a series of primer pairs proximal to the putative transcription initiation site of each target gene (Fig. 1B; see Table S1 in the supplemental material). As expected, acetylation of histones H3 and H4 upstream of constitutively expressed housekeeping genes is observed in both tachyzoites and bradyzoites (Fig. 1C). In marked contrast, bradyzoite-specific genes are hypoacetylated in tachyzoite populations and become acetylated under differentiation conditions (Fig. 1C and D). We also noted that tachyzoite-specific genes are hyperacetylated in tachyzoites and acetylation at these promoters is diminished under bradyzoite growth conditions (Fig. 1C). ChIP assays with histone H3 antibody show that the decreased signal from the specific antibodies is due to decreased acetylation, rather than decreased H3 occupancy or accessibility to antibodies. Some residual acetylation of tachyzoite-specific promoters is still evident because in vitro stage conversion is not 100% efficient.

Recent evidence indicates that some *Toxoplasma* genes belonging to gene families are located in close proximity to each other (31). This is the case for the majority of the *SRS* genes (*SAG*-related sequences) found throughout the genome in distinct, tandemly arrayed multigene clusters (31). Since transcription of these clusters is coordinated during interconversion, we determined if they exhibit long-range histone acetylation patterns (19). Using ChIP, we scanned a 15-kb stretch of the bradyzoite-specific *BRS4* locus for acetylated histone H4 in both tachyzoite and in vitro-induced bradyzoite stages (Fig. 1D). Coincident with bradyzoite conversion, histone acetylation occurs at each gene within the locus. Whereas no significant acetylation is observed for certain regions, some sites displayed great peaks of acetylation (Fig. 1D). These findings demonstrate that long-range histone acetylation patterns spanning many kilobases can be established in *Toxoplasma*. By contrast, genes from the dense granule family (*GRA*), which are found in various chromosomal locations, exhibited disparate patterns of acetylation between life cycle stages (Fig. 1E).

**Purification of a novel corepressor complex from *Toxoplasma*.** Histone acetylation activities like those described above are mediated by HATs and HDACs. A GCN5 family HAT has been described in *Toxoplasma* (27, 53), but no one has characterized an HDAC enzyme in this parasite. Bioinformatic analysis of the *Toxoplasma* genome (<http://ToxoDB.org>)

(33) reveals a family of six putative HDACs in *Toxoplasma*, which can be grouped into three classes (I, II, and III) based upon similarity to human homologues (data not shown; 13, 23). We generated stable transgenic parasites (RH strain) expressing recombinant *Toxoplasma* HDAC homologues dually tagged with HA and FLAG, referred to as TgHDAC1 to -5 and TgSIR2. Whereas some class I- and class II-related enzymes were predominantly cytoplasmic (data not shown), TgHDAC3 (class I) was exclusively localized in the parasite nucleus, being largely excluded from the nucleolus, as observed for acetylated H4 (Fig. 2A). Its nuclear localization prompted us to examine the TgHDAC3 complex in greater detail by utilizing the epitope tag for chromatography purification (Fig. 2B). Analysis by silver staining of Superose 6 gel filtration fractions reveals that TgHDAC3 is embedded in a remarkably robust (it remains stable in 0.5 M NaCl and 0.1% NP-40) and high-molecular-mass complex (>1 MDa, fractions 20 to 22, Fig. 2C). We also observed a peak of immunoreactivity of the monomeric form (fractions 30 to 32; Fig. 2C). Colloidal blue staining of trichloroacetic acid-precipitated fraction 20 displays multiple bands, whose intensity suggests a stoichiometric complex (named TgCRC for *T. gondii* corepressor complex). Peptide sequences obtained from excised bands were compared to predicted proteins in the *Toxoplasma* sequence database (Fig. 2D; see Table S2 in the supplemental material).

We identified the TgCRC-62 and -100 bands as TgHDAC3 (accession no. DQ004745) and TgTBL1 (*T. gondii* transducin beta-like protein 1; accession no. DQ004746), respectively. Interestingly, TgTBL1 contains one LisH and six WD40 domains most similar to its human counterpart (Fig. 2E). TBL1 was originally identified as deleted in patients with sensorineural deafness associated with an X-linked syndrome of ocular albinism (3). HDAC3 is also found associated with TBL1 in human N-CoR (nuclear receptor corepressor)- and SMRT (silencing mediator of retinoid and thyroid hormone receptors)-containing complexes (24, 28, 36, 63, 64). However, no peptide sequences corresponding to parasite homologues of N-CoR or SMRT proteins were purified with TgHDAC3, consistent with their absence from the *Toxoplasma* genome database. Nonetheless, TgCRC exhibits other features of HDAC3-containing complexes besides the TBL1 association. For instance, the actin-binding protein IR10 was reported as a specific component of the N-CoR complex, suggesting a functional connection with actin (63). Interestingly, we identified TgCRC-45 as actin (accession no. AAC13766). TgHDAC3 also forms a stable and stoichiometric association with two HSP70-like proteins, TgCRC-75 (TgHSP70a; accession no. AAF23321), and TgCRC-80 (TgHSP70b; accession no. AAC72001), as previously reported in human cells (25, 63). We also established TgCRC-64, -66, and -68 bands as subunits of a *Toxoplasma* TCP-1 ring complex, TRiC (see Table S2 in the supplemental material) (50). TRiC stably associates with human HDAC3 (25) and the yeast HDAC HOS2 (46).

TgCRC-120 was identified as a novel parasite HSP90-like protein (accession no. DQ004747) distinct from the one previously described as being secreted into the parasitophorous vacuole (accession no. AY292370) (1). Whereas in human cells the HDAC3-Hsp90 association was described as transient (25), we observed a stable and robust interaction between TgHDAC3 and TgCRC-120 (Fig. 2D). However, treatment of



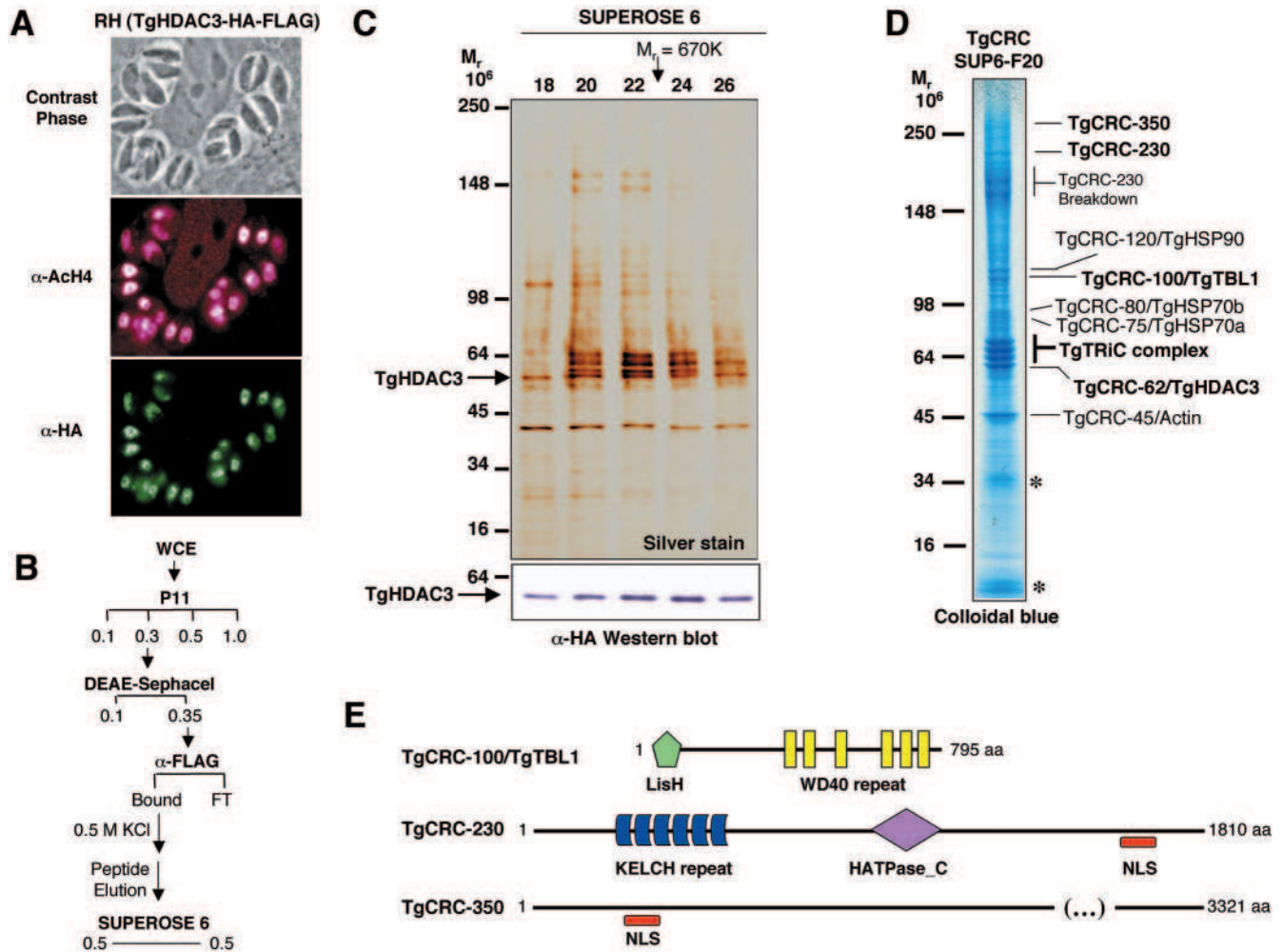


FIG. 2. Chromatographic purification of TgCRC, a TgHDAC3-containing complex. (A) Recombinant protein in transgenic parasites expressing TgHDAC3-HA-FLAG was detected by IFA with HA antibody and compared to nuclear localization of acetylated histone H4 ( $\alpha$ -AcH4). (B) Purification scheme of the TgHDAC3-HA-FLAG complex. WCE from tachyzoites was fractionated by chromatography as described in Materials and Methods. The 0.35 M KCl elution of DEAE-Sephacel was purified with an anti-FLAG M2 affinity column. The bound proteins were further analyzed by Superose 6 gel filtration. The horizontal and diagonal lines indicate stepwise and gradient elutions, respectively. Molar concentrations are given. FT, flowthrough. (C) Silver staining and Western blotting ( $\alpha$ -HA) of Superose 6 fractions (15  $\mu$ l). Molecular weight markers are indicated on the left. (D) Superose 6 fraction 20 was trichloroacetic acid precipitated, separated by SDS-polyacrylamide gel electrophoresis (4% to 12%), and visualized by colloidal blue staining. Molecular weight markers are indicated on the left. The proteins identified by mass spectrometry sequencing are indicated. Polypeptides marked with asterisks are contaminants. (E) Structural features of TgCRC-100/TgTBL1, TgCRC-230, and TgCRC-350. aa, amino acids.

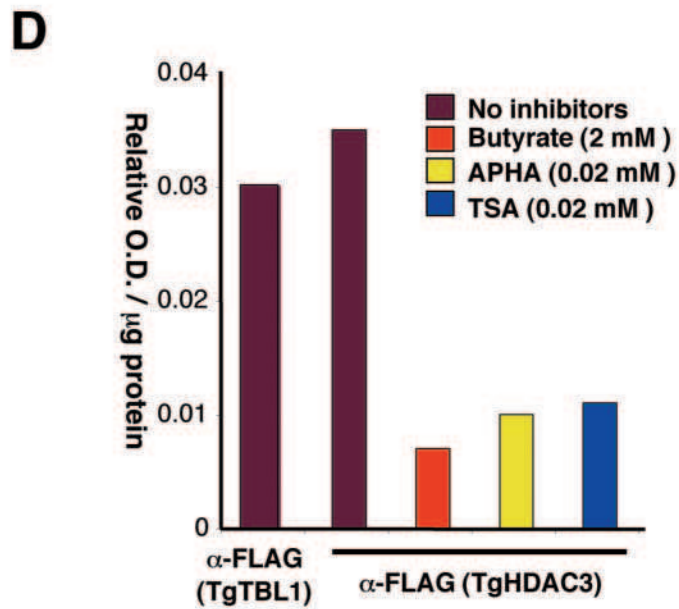
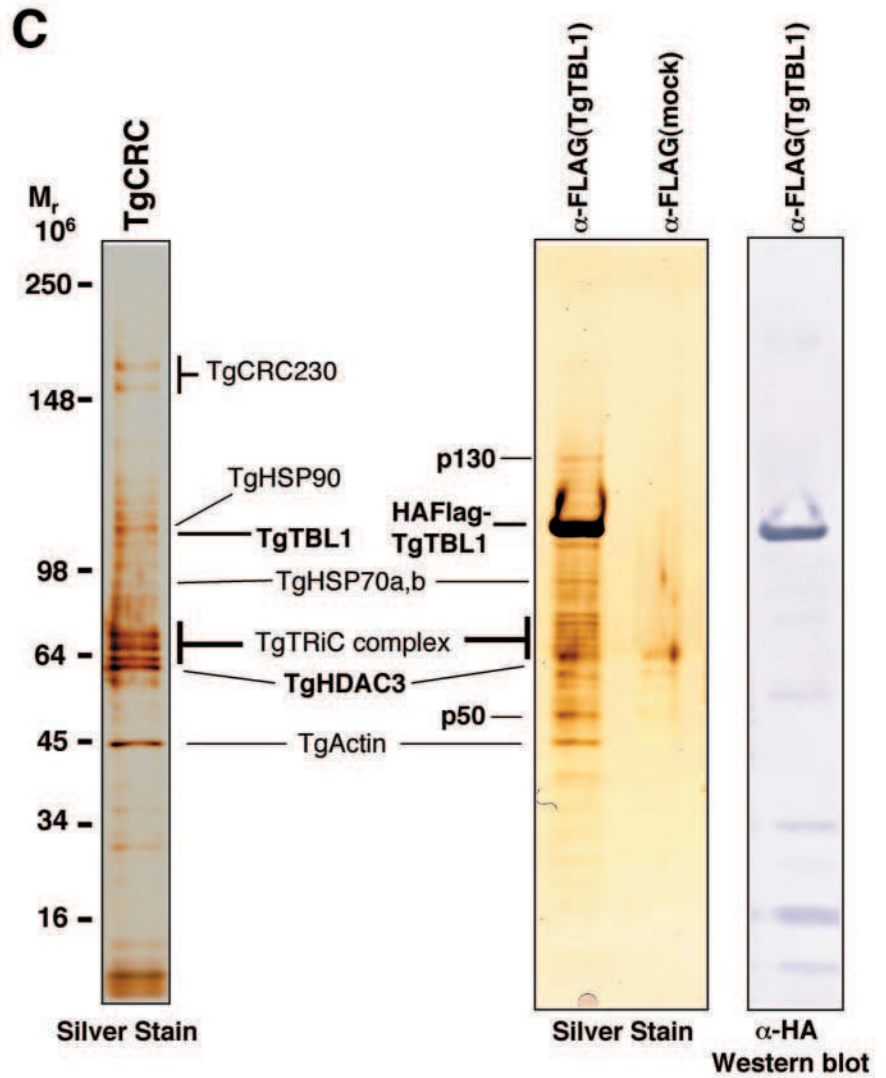
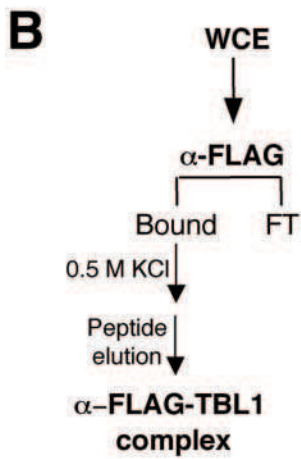
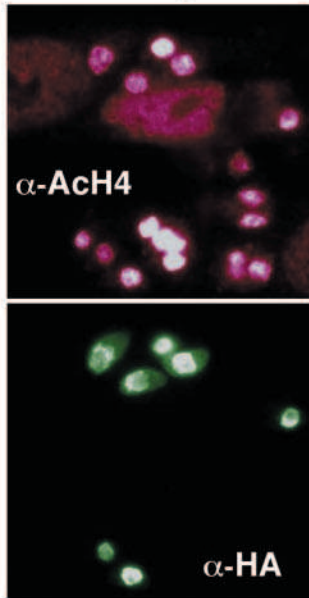
transgenic parasites with the Hsp90 inhibitor geldanamycin does not prevent nuclear localization of TgHDAC3-HA-FLAG, as shown in human cells (data not shown; 25).

Convincing data suggest that human HDAC3 interacts with SMRT only after priming by cellular chaperones, including the TCP-1 ring complex, HSP70, and HSP90 (25). It has also been shown that HDAC3 requires SMRT for HDAC enzyme activity (25, 28, 63, 64). While no clear N-CoR- or SMRT-related homologues exist in *Toxoplasma*, two new polypeptides were discovered that correspond to functionally uncharacterized genes (TgCRC-350 and TgCRC-230, Fig. 2E). Analysis of a number of preparations indicates that TgCRC-350 is a substoichiometric component of this complex whereas TgCRC-230 is sensitive to breakdown and abundant, with more than 30 tryptic peptides detected (Fig. 2D). Both TgCRC-230 and

TgCRC-350 display a bipartite nuclear targeting sequence, as predicted by PSORT (Fig. 2E). SMART analysis reveals that TgCRC-230 possesses six KELCH repeats and a HATPase domain (Fig. 2E). The roles of the TgCRC-350 and TgCRC-230 proteins remain to be elucidated, but it is tempting to speculate that they could be activating factors for TgHDAC3 that can potentiate transcriptional repression in a manner similar to that reported for SMRT or N-CoR (24, 28, 36, 63, 64).

To rigorously demonstrate that these proteins are components of the TgCRC, we employed a HA-FLAG-tagged version of TgTBL1 as bait to reisolate it. This protein ectopically expressed in tachyzoites localizes exclusively in the nucleus and is also excluded from the nucleolus, as observed previously for TgHDAC3 (Fig. 3A). No stable transgenic parasites could be obtained, suggesting that an additional copy of TgTBL1 is not

**A** RH (HAFLAG-TgTBL1)  
transient expression



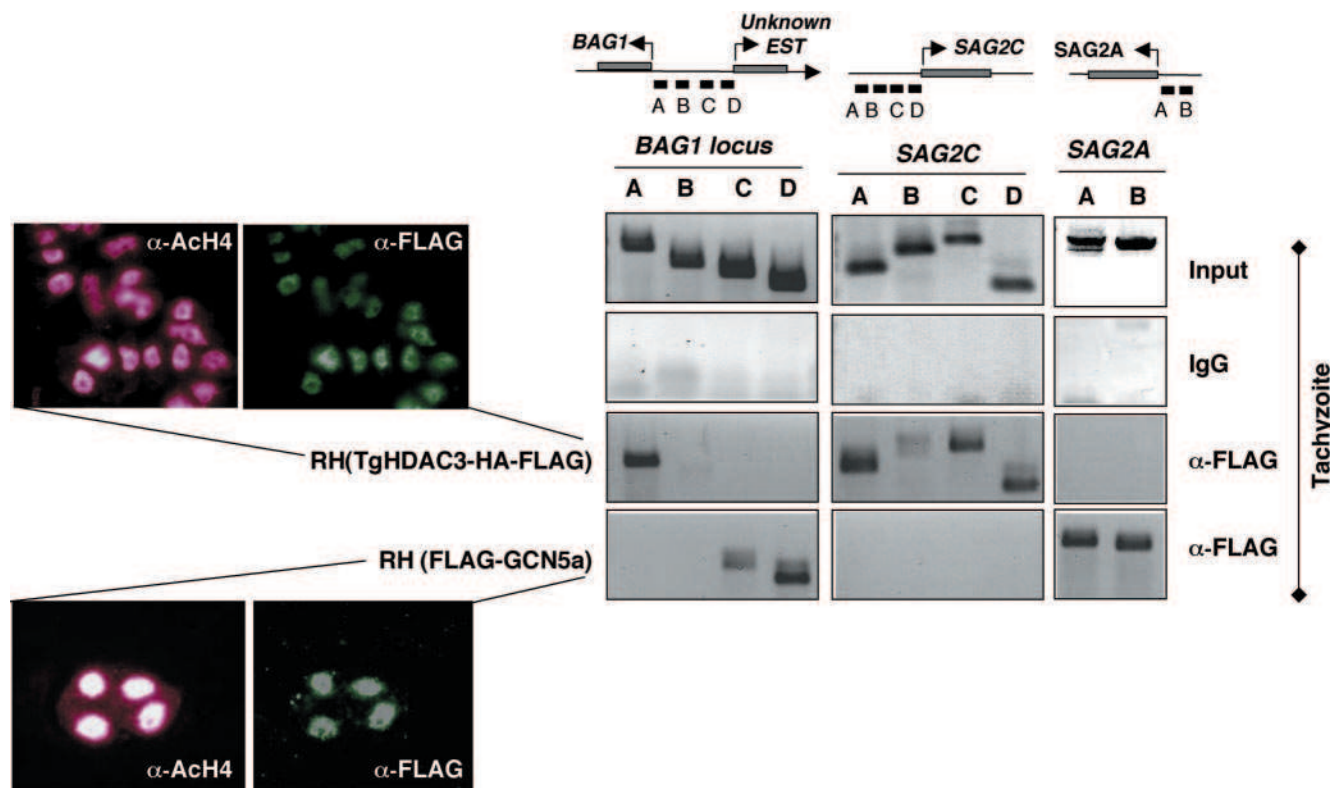


FIG. 4. ChIP analysis of TgHDAC3 and TgGCN5 binding sites. Tachyzoites from stable transgenic parasite clones expressing either FLAG-TgGCN5 or TgHDAC3-HA-FLAG were processed for ChIP with  $\alpha$ -FLAG antibody. Stage-specific promoters that were tested included *SAG2A* and an unknown EST (tachyzoite specific), as well as *BAG1* and *SAG2C* (bradyzoite specific). A chromatin sample with an irrelevant antibody was used as a negative control (IgG) and genomic DNA as a positive control for PCR (input).

tolerated. However, silver staining analysis of the affinity eluates from transiently transfected tachyzoites confirms that many of the polypeptides are members of the TgCRC. The differences include the lack of the TgCRC-230 subunit and the presence of two novel uncharacterized subunits named p50 and p130 (Fig. 3B and C).

HDAC activities of the TgCRC purified by virtue of HA-FLAG-TBL1 or TgHDAC3-HA-FLAG were determined and found to be susceptible to known HDAC inhibitors (Fig. 3D). The purified complexes also displayed HDAC activity on core histones (data not shown).

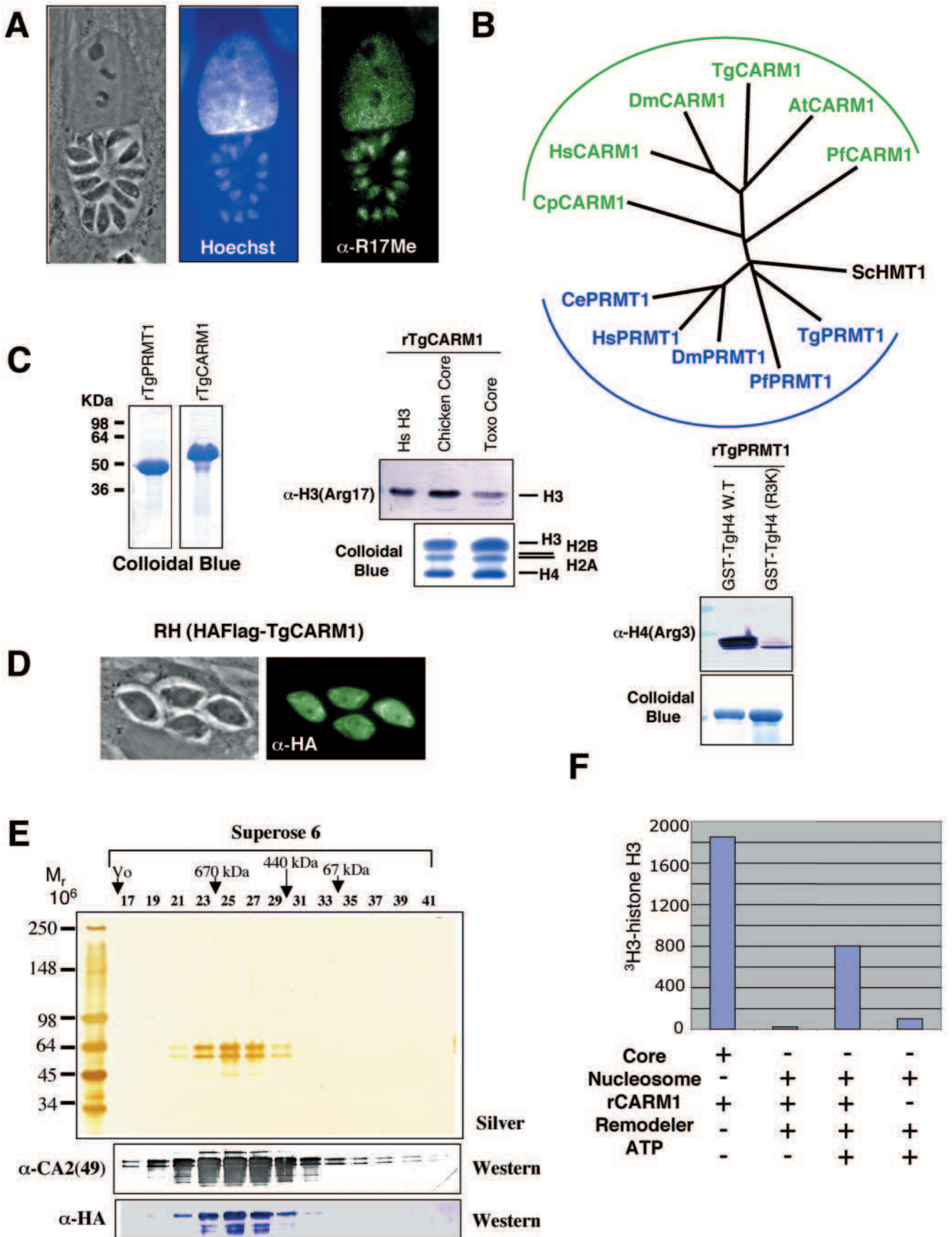
**Differential association of TgGCN5 and TgHDAC3 at stage-specific promoters.** We hypothesized that during the tachyzoite stage HATs would be present at tachyzoite-specific promoters while HDACs would be at bradyzoite-specific promoters. We performed ChIP assays on stable transgenic parasite clones expressing either TgGCN5 or TgHDAC3 as a FLAG-tagged

fusion protein (Fig. 4) (5). ChIP assays using anti-FLAG show that TgHDAC3 is recruited upstream of the bradyzoite-specific *SAG2C* and *BAG1* genes in tachyzoites (Fig. 4). Importantly, as shown by ChIP assays using antibodies to acetylated H3 and H4, hypoacetylation is observed at the same loci (data not shown). In contrast, TgGCN5 is found in association with tachyzoite-specific promoters (*SAG2A* and an unknown expressed sequence tag [EST], Fig. 4). Therefore, the differential association of TgGCN5 and TgHDAC3 at stage-specific promoters contributes to gene regulation via acetylation of histone tails. Other acetylases should be involved, since hyperacetylation of H3 [K9-K14] and H4 [K5-K8-K12-K16] is detected at these loci (data not shown).

**Arginine methylation of *Toxoplasma* histones.** In mammalian cells, methylation of histone arginine residues by CARM1 (coactivator-associated arginine methyltransferase 1) and PRMT1 (protein arginine methyltransferase 1) occurs at promoters during

FIG. 3. Repurification of TgCRC with TgTBL1 as bait and determination of enzymatic activity. (A) Transiently expressed recombinant protein HA-FLAG-TgTBL1 was detected by IFA with an HA antibody and compared to nuclear localization of acetylated histone H4 ( $\alpha$ -AcH4). (B) Purification scheme of TgTBL1-associated proteins. FT, flowthrough. (C) Silver staining of affinity-purified TgTBL1-containing complex compared to TgCRC (Superose 6, fraction 20). The proteins analyzed are indicated on the right. Western blot analysis using HA antibody detects tagged TgTBL1. WCE from untransfected wild-type RH tachyzoites was processed similarly as a control (mock). Molecular weight markers are indicated on the left. (D) HDAC activity assays of FLAG affinity-purified TgTBL1 or TgHDAC3. TSA, trichostatin A; APAH, aroyl-pyrrole-hydroxyamides.





transcriptional activation in response to hormone induction (4, 38, 51, 56) or in association with p53 (2). No evidence exists that arginine methylation affects gene expression in protozoa. We detected methylation of H3 [R17] within tachyzoite nuclei by using immunofluorescent staining (Fig. 5A). Bioinformatic analysis of the *Toxoplasma* genome sequence identifies five putative arginine methyltransferases with a conserved catalytic core. Since CARM1 and PRMT1 are the methyltransferases predominantly involved in gene regulation in human cells, we focused on the putative orthologues *TgCARM1* (accession no. AY820755) and *TgPRMT1* (accession no. AY820756) (Fig. 5B).

In vitro histone methylation assays show that recombinant TgCARM1 methylates H3 [R17] while TgPRMT1 methylates H4 [R3], consistent with their respective human counterparts (Fig. 5C). Moreover, mutation of arginine 3 of H4 abolishes the enzymatic activity of TgPRMT1 (Fig. 5C). Ectopically expressed HA-FLAG-TgCARM1 is able to localize to the parasite nucleus, consistent with a role in transcriptional regulation (Fig. 5D). Affinity-purified HA-FLAG-TgCARM1 from the transgenic parasite clone reveals a dimer of tagged and endogenous TgCARM1 capable of methylating histone H3 [R17] in free core histones (Fig. 5E). The apparent molecular mass is 600 kDa, suggesting that TgCARM1 exists as a hexamer, as reported previously for HMT1 and PRMT3 (59, 65).

Human CARM1 acquires the ability to covalently modify nucleosomal histones when complexed with SWI/SNF ATPases (60). We examined this in *Toxoplasma* since SWI/SNF proteins have been identified (52) by mixing recombinant TgCARM1 with parasite extract enriched in ATP-dependent nucleosome disruption activity. In the presence of ATP, this mixture was capable of methylating nucleosomal histones (Fig. 5F), suggesting that TgCARM1 likely acts in concert with SWI/SNF chromatin remodelers in *Toxoplasma* in vivo.

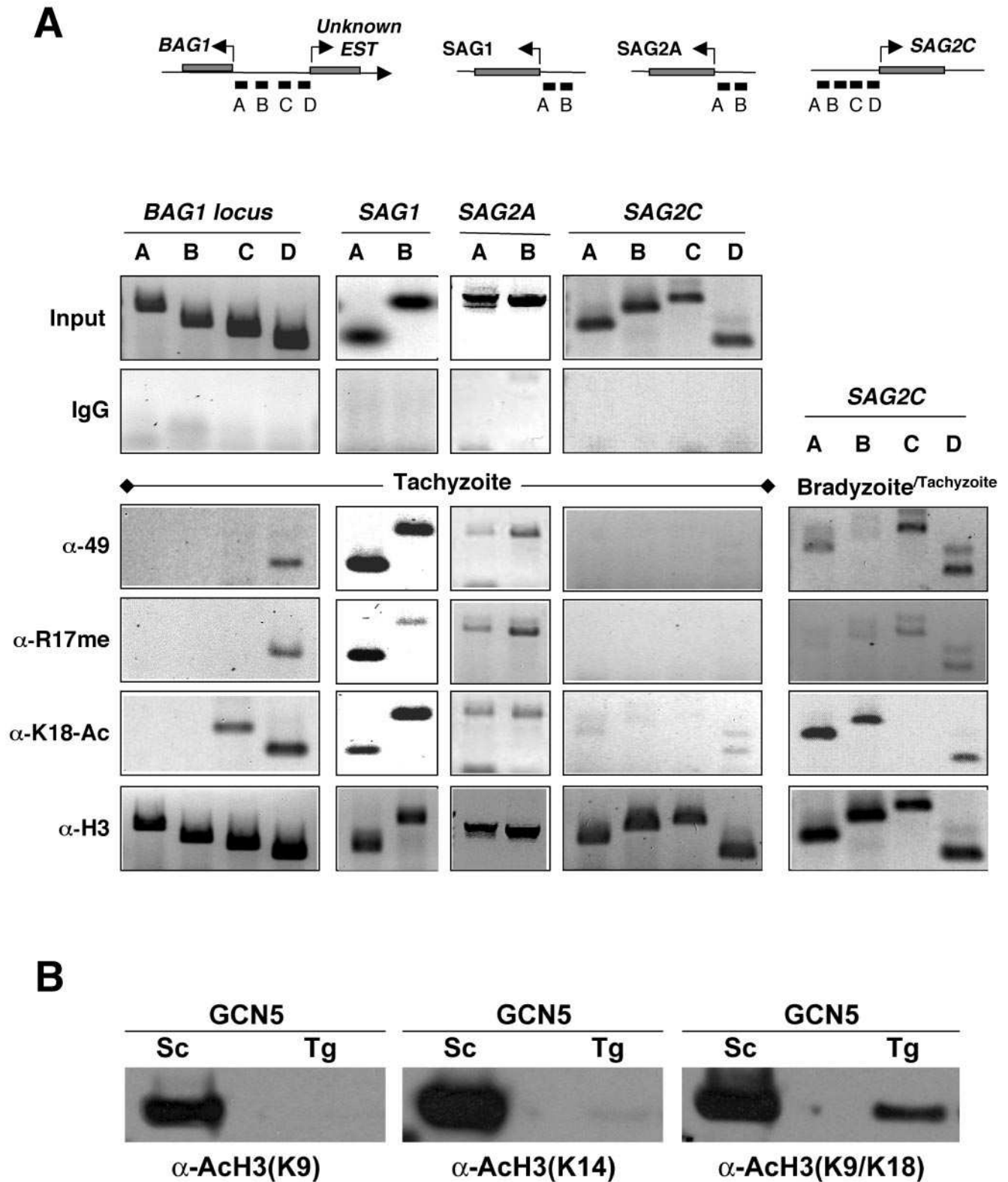
**Concurrent histone methylation and acetylation on stage-specific promoters.** We generated a polyclonal antibody against TgCARM1 ( $\alpha$ -49) for use in ChIP experiments to identify target genes of TgCARM1. ChIP analysis of tachyzoites reveals that TgCARM1 is associated with the promoters of tachyzoite-specific genes *SAG1* and *SAG2A* and the unknown EST proximal to the *BAG1* locus (Fig. 6A). After induction of stage conversion, TgCARM1 was found in association with the

bradyzoite-specific *SAG2C* promoter (Fig. 6A). Importantly, methylated arginine 17 and acetylated lysine 18 of histone H3 were consistently present at the same loci (Fig. 6A). Cross talk between these histone modifications was previously reported in human cells, where CBP/p300 acetylates H3 [K18] to facilitate CARM1 recruitment and subsequent H3 [R17] methylation (11). Apicomplexans do not possess a CBP/p300 homologue, but purified FLAG-TgGCN5 exhibits an unexpected bias to acetylate only H3 [K18] (Fig. 6B). This is intriguing because the preferred substrate for GCN5 in other species is H3 [K14]. These studies suggest that TgGCN5 has evolved an unusual substrate preference and may contribute to combinatorial histone modifications to affect gene expression in *Toxoplasma*. However, it should be noted that yeast GCN5 exhibits an increased ability to acetylate multiple H3 lysines when partnered with additional proteins found in native GCN5 complexes (22).

We also observed accumulation of multiple methyl groups at H3 [K4] ( $\alpha$ -K4-2ME and -3ME [data not shown]). Dimethylation occurs at both inactive and active euchromatic *Toxoplasma* genes, whereas trimethylation is present exclusively at active genes, as reported in yeast (data not shown; 47). Trimethylation is significantly enriched in the neighborhood of tachyzoite-specific genes during that life cycle stage and enhanced for the bradyzoite gene *SAG2C* following differentiation. Therefore, the presence of a trimethylated lysine 4 defines an active state of gene expression in *Toxoplasma*, much like arginine 17 methylation and K18 acetylation.

**Inhibiting TgCARM1 induces cyst formation.** Multiple attempts to generate a knockout of the *TgCARM1* gene failed, suggesting that the gene is essential in tachyzoites (data not shown). This is consistent with the observation that mice null for CARM1 die perinatally and are small (61). We next attempted to express a dominant-negative form of TgCARM1 by mutating the conserved catalytic residues to abolish enzyme activity (VMD $\Delta$ AAA) (7). Surprisingly, HA-FLAG-tagged TgCARM1<sup>VMD $\Delta$ AAA</sup> is exclusively cytoplasmic (data not shown). Unfortunately, no stable parasites could be obtained expressing this mutant protein. Even when transiently expressed, HA-FLAG-TgCARM1<sup>VMD $\Delta$ AAA</sup> significantly reduces parasite division (data not shown).

FIG. 5. Arginine methylation of *Toxoplasma* histones. (A) Immunofluorescence colocalization of methylated [R17] H3 and Hoechst 33258 within the nuclei of intracellular tachyzoites. (B) Evolutionary relationships between arginine methyltransferases of *T. gondii* TgCARM1 and TgPRMT1, *Plasmodium falciparum* PfCARM1 (accession no. CAD51260) and PfPRMT1 (accession no. AAN36855), *Cryptosporidium parvum* CpCARM1 (genomic annotation), *Arabidopsis thaliana* AtCARM1 (accession no. AAF26997), *D. melanogaster* DmCARM1 (CARMER, accession no. AAF54471) and DmPRMT1 (accession no. AAM11369), *C. elegans* CePRMT1 (accession no. T26447), *S. cerevisiae* SCHMT1 (accession no. NP\_009590), and *Homo sapiens* HsCARM1 (accession no. AAH46240) and HsPRMT1 (accession no. AAF62895). The unrooted phylogenetic tree was inferred from the core domain alignment. (C) Recombinant histidine-tagged proteins (rTgCARM1 and rTgPRMT1) were expressed in *E. coli* and purified by Ni-nitrilotriacetic acid affinity chromatography. Methylation of H3 [R17] is mediated in vitro by rTgCARM1 (0.25  $\mu$ g). The assays with recombinant H3 (1.5  $\mu$ g) and free core histones (3  $\mu$ g) were performed by immunoblotting with  $\alpha$ -methyl [R17] H3 antibody. Methylation of H4 [R3] is mediated in vitro by rTgPRMT1 (2.5  $\mu$ g). The assays with wild-type GST-histone TgH4 (1.5  $\mu$ g), mutant GST-histone TgH4 (R3K, 1.5  $\mu$ g), and free core histones (3  $\mu$ g) were performed by immunoblotting with  $\alpha$ -methyl [R3] H4 antibody. (D) Immunofluorescence assay with  $\alpha$ -HA of transgenic parasite clone A3, engineered to ectopically express HA-FLAG-TgCARM1. (E) WCE of  $5 \times 10^{10}$  tachyzoites expressing HA-FLAG-TgCARM1 was fractionated with an anti-FLAG M2 affinity column. Bound proteins were further analyzed by chromatography on a Superose 6 column. Silver staining analysis of the anti-FLAG affinity eluates corresponding to fractions 17 to 41 was done. Western blotting was performed with the antibodies indicated on the left. W.T., wild type. (F) Methylation of nucleosomal histones (3  $\mu$ g) by rTgCARM1 (0.25  $\mu$ g) supplemented with 5  $\mu$ g of *Toxoplasma* protein extract enriched in ATP-dependent nucleosome disruption activity. The concentration of ATP was 1.0 mM where added. The reaction mixtures were resolved by SDS-polyacrylamide gel electrophoresis, and [<sup>3</sup>H]histone H3 was excised for counting in scintillation buffer.





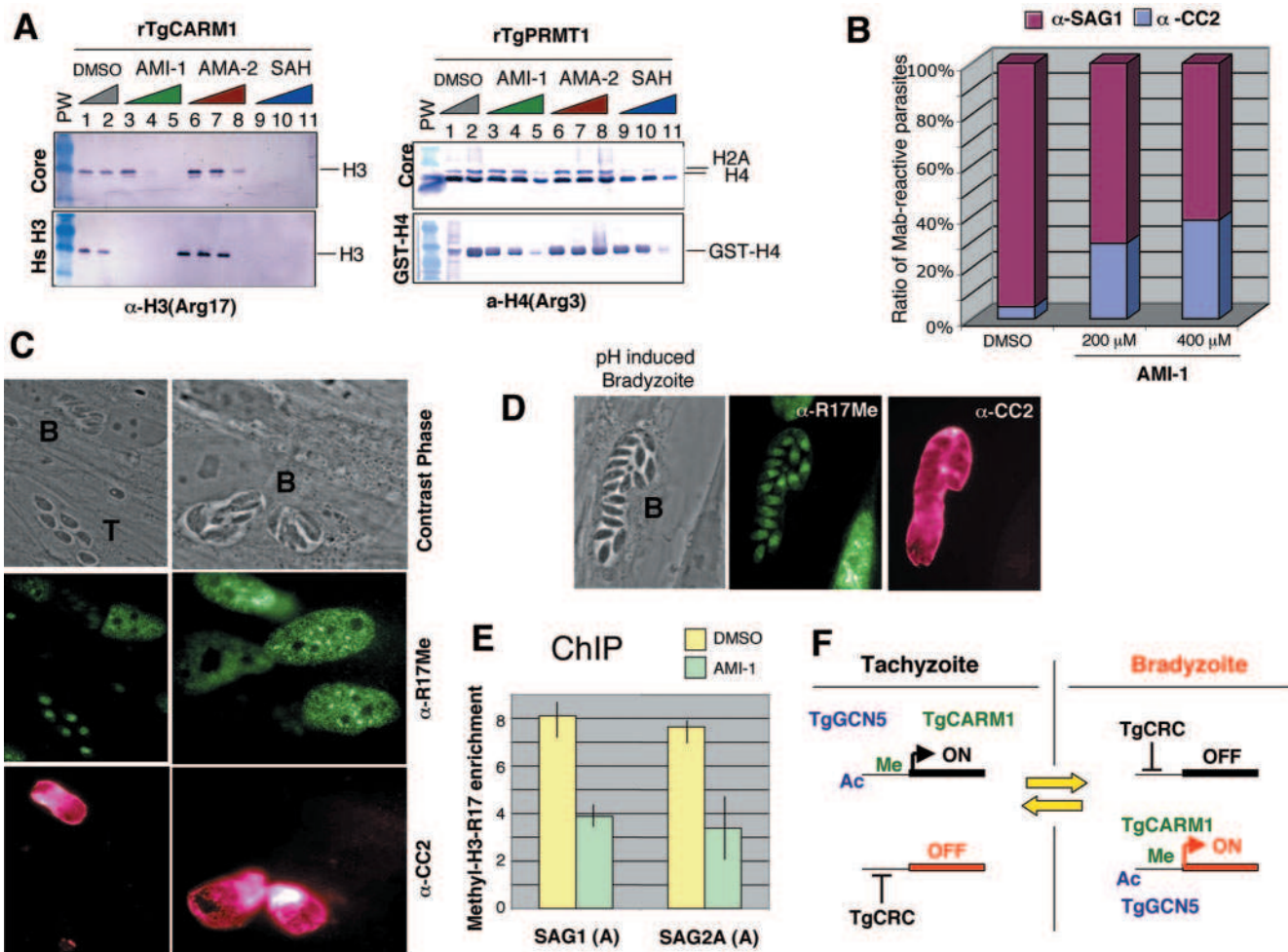


FIG. 7. AMI-1 inhibits TgCARM1 activity in vitro and induces cyst formation in vivo. (A) In vitro inhibition by AMI-1 and activation by AMA-2 of methylation reactions mediated by rTgCARM1 (0.25 μg) and rTgPRMT1 (0.2 μg). Methylation reactions were performed with histone H3 (1.5 μg), wild-type GST-histone TgH4 (1.5 μg), and free core histone substrates (3 μg) in the presence of dimethyl sulfoxide (DMSO), AMI-1 (50, 200, and 600 μM), AMA-2 (50, 200, and 600 μM), and S-adenosylhomocysteine (SAH; 0.04, 0.16, and 0.32 μM). PW, DNA molecular weight. (B) Estimation of AMI-1-dependent in vitro conversion. The ratio of parasites that were reactive either with tachyzoite-specific α-SAG1 antibody or bradyzoite-specific α-CC2 antibody was determined by double immunofluorescence as described in Materials and Methods. Mab, monoclonal antibody. (C) Phase-contrast microscopy and immunofluorescence assays with AMI-1-treated tachyzoites with α-methyl [R17] H3 and α-CC2 antibodies. The decrease of methylation at arginine 17 localizes within the induced bradyzoites as assessed by α-CC2 immunofluorescence. (D) Immunofluorescence of in vitro-generated bradyzoites induced by an alkaline pH with α-methyl [R17] H3 and α-CC2 antibodies. (E) ChIP experiments with extracellular tachyzoites treated with AMI-1 or dimethyl sulfoxide (control) display the binding of methylated H3 [R17] on TgCARM1 target *SAG1* and *SAG2A* genes (Fig. 6A). The *n*-fold enrichment was calculated by dividing the α-methyl [R17] H3 intensity signal by the input control signal. The data from five independent experiments were pooled in order to graph the average enrichment. The primers used in the PCR analysis were designed to span the region showing the highest peak of enrichment for each promoter (region A, Fig. 6A). (F) Model of epigenetic control of stage-specific gene expression in *T. gondii*.

We then explored pharmacological options to study TgCARM1 in greater detail. AMI-1 is a small molecule that specifically inhibits arginine *N*-methyltransferase activity (9). AMI-1 does not compete for the AdoMet binding site and inhibits arginine but not lysine methyltransferases in vitro (9). AMA-2 is an activator of Hmt1p (9). We investigated the effect of AMI-1 and AMA-2 in methylation reactions performed with recombinant TgCARM1 and TgPRMT1 (Fig. 7A). AMA-2 stimulates the methylation activity of TgCARM1 up to 10-fold, whereas it has no effect on TgPRMT1; AMI-1 efficiently inhibits TgCARM1-mediated methylation of H3 (Fig. 7A). In contrast, AMI-1 exhibits a minimal effect on TgPRMT1 activ-

ity whereas S-adenosylhomocysteine totally abrogates both TgCARM1 and TgPRMT1 activities at 0.04 μM and 0.40 μM, respectively (Fig. 7A).

Specific inhibition by AMI-1 of TgCARM1-mediated methylation prompted us to test its effect on *Toxoplasma*. Treatment of intracellular parasites had no observable effect on cell division or cyst formation (data not shown). However, pretreatment of extracellular tachyzoites with AMI-1 prior to infection induced cyst formation following two to three rounds of intracellular division (Fig. 7B). Interestingly, AMI-1-induced stage conversion correlates with a significant decrease in global H3 [R17] methylation (Fig. 7C). We also used ChIP assays to

confirm that AMI-1 treatment triggers an up to 50% decrease in methylated H3 [R17] in promoters of the tachyzoite-specific genes *SAG1* and *SAG24* (Fig. 7E). These observations imply that methylated H3 [R17] might be a hallmark of the tachyzoite stage; however, in vitro-generated bradyzoites also exhibit methylation of H3 [R17] (Fig. 7D), showing clearly that TgCARM1 is active in both stages. One explanation for the effect of AMI-1 is that by inhibiting TgCARM1 activity, unknown “differentiation checkpoint genes” may have been turned off. Down-regulation of these genes would therefore initiate differentiation of tachyzoites to bradyzoites. Alternatively, arginine methyltransferases have been shown to target nonhistone proteins, and inhibiting those modifications may be responsible for the effects of AMI-1.

## DISCUSSION

Regulation of gene expression is vastly understudied in pathogenic protozoan parasites, where it is likely to be of critical importance given their complex multiple life cycle stages. Recent completion of several apicomplexan genomes reveals a glaring absence of conventional transcription factors; however, it is becoming apparent that an abundance of chromatin remodeling machinery is present. Insight into the epigenetics and how transcription is regulated in these ancient eukaryotes has not only important evolutionary implications but also the potential to reveal novel ways to therapeutically intervene to treat protozoal diseases. Indeed, antigenic variation in *Trypanosoma* and *Plasmodium* involves epigenetics, and disruption of the SIR2 HDAC in the latter leads to a loss of control over *var* gene expression (15, 20, 45). Further, the antiprotozoal agent apicidin targets a parasite HDAC (10). In this study, we were unable to disrupt the *TgCARM1* loci, implying that the gene product is essential for parasite welfare. Hence, we sought to explore epigenetics as a means of controlling stage-specific gene expression in the medically relevant apicomplexan *Toxoplasma*. This is the first study to demonstrate the use of histone modifications for life cycle stage conversion and elucidates novel parasite-specific HDAC and CARM1 complexes mediating these activities in the context of a previously described GCN5 HAT.

**Histone modifications and life cycle stage conversion.** Having made strong correlations between histone acetylation status and stage-specific gene expression in *Toxoplasma* (Fig. 1), we have made our next objective to identify the chromatin remodeling complexes responsible for mediating these activities. A GCN5 HAT has been described previously (53), and here we characterize TgCRC, a multisubunit TgHDAC3-containing complex found in association with transcriptionally inactive promoters in *Toxoplasma*. While *Toxoplasma* lacks unambiguous homologues to N-CoR and SMRT, the TgCRC is a likely evolutionary precursor to the complexes containing NcoR/SMRT in higher eukaryotes. While some components, such as TBL1, HSP70, and the TCP-1 ring complex, are conserved between TgCRC and N-CoR/SMRT complexes, novel parasite-specific proteins (TgCRC-230 and -350) highlight the uniqueness of the *Toxoplasma* machinery, whose physiological and phylogenetic properties warrant further investigation. Of particular note is that TgCRC-350 is a substoichiometric component that hence may not always be associated with the

complex, suggesting that it may function as the DNA-binding component. Consistent with a crucial role for TgCRC in apicomplexans in general, HDAC3 (30) and TgTBL1 (accession no. Q8IEK5) orthologs are present in *Plasmodium*, as are possible orthologs to TgCRC-230 (accession no. AAN37262) and -350 (accession no. CAD49240).

We also show here for the first time that arginine methylation, linked to transcriptional activation in other organisms, is a key histone modification in *Toxoplasma* that also correlates with parasite differentiation. This is in agreement with other observations showing that arginine methylation functions in differentiation (8). We noted that TgCARM1 acquires the ability to methylate nucleosomes in the presence of ATP and parasite fractions enriched for SWI/SNF activities (Fig. 5F). It is tempting to speculate that it interacts with TgSRCAP, a SWI/SNF ATPase noted to increase in expression during bradyzoite differentiation (52) and whose human homologue does interact with CARM1 (43). Finally, it would be of interest to study if TgCARM1 methylates nonhistone substrates as a means of triggering differentiation, as noted in thymocytes (32).

**Evolution of histone modifications.** It was surprising to discover that *Toxoplasma* possesses a sophisticated capacity to modify histones, rivaling the system observed in higher eukaryotic cells. For instance, a more extensive repertoire of arginine methylation machinery is present in apicomplexans compared to yeast and *C. elegans*, which each possess only one PRMT-related enzyme (Fig. 5B). This observation is consistent with the idea that the lack of many known transcription factors in apicomplexans may be compensated for at an epigenetic level. Even more surprising is the parasite’s employment of H3 [K18] acetylation and H3 [R17] methylation to activate transcription in a stage-specific manner. This combination has particular relevance as it has only been reported to occur in humans, whereby CBP/p300 recruits CARM1 by virtue of H3 [K18] acetylation (11). While there is no CBP/p300 in apicomplexans, we made the remarkable discovery that recombinant TgGCN5 exquisitely targets H3 [K18] in in vitro HAT assays. As observed in yeast, TgGCN5 may display greater H3 acetylation capacity in vivo (22). Nevertheless, this concurrent acetylation-methylation observed in *Toxoplasma* may represent an example of convergent evolution.

**Model for epigenetic involvement in parasite differentiation.** Our model thus far for the epigenetic regulation of stage conversion in *Toxoplasma* is shown in Fig. 7F. As noted, key signatures of *Toxoplasma* gene activation are acetylation of H3 [K18] and methylation of H3 [R17], which we have shown to be mediated by TgGCN5 and TgCARM1, respectively. We have also shown that *Toxoplasma* can silence genes via deacetylases, such as the TgHDAC3-containing TgCRC. There are undoubtedly other histone modifications employed by *Toxoplasma* to modulate gene expression, and our establishment of a working ChIP protocol for this organism will be invaluable in teasing them apart. The parasite-specific subunits contained within these chromatin-remodeling complexes are some of the most compelling components worthy of further investigation. These or other identified remodeling components (44) are likely to be master regulators or members of the stress-signaling pathway triggering bradyzoite differentiation. Either will be critical clues illuminating the molecular mechanisms driving



parasite differentiation. Finally, our report demonstrates that besides its importance in pathogenesis, stage conversion in *Toxoplasma* can serve as a model system to study epigenetic mechanisms of eukaryotic cell differentiation.

#### ACKNOWLEDGMENTS

Thanks go to J. F. Dubremetz, U. W. Bohne, and U. Gross for antibodies and D. S. Roos for the  $\lambda$ ZAPII library. We thank J. Gagnon for reviewing the manuscript. Preliminary genomic and/or cDNA sequence data were accessed via <http://ToxoDB.org> (33).

W. J. Sullivan is supported by an R01 grant from the National Institutes of Health (GM065051). N. Saksouk is supported by a fellowship from CNRSL (Lebanon). M. A. Hakimi is supported by grants from CNRS (ATIP program), the Fondation pour la Recherche Médicale (FRM), Emergence (Rhône-Alpes), and INSERM. Genomic data were provided by the Institute for Genomic Research (supported by NIH grant AI05093) and by the Sanger Center (Wellcome Trust). EST sequences were generated by Washington University (NIH grant 1R01AI045806-01A1).

#### REFERENCES

- Ahn, H. J., S. Kim, and H. W. Nam. 2003. Molecular cloning of the 82-kDa heat shock protein (HSP90) of *Toxoplasma gondii* associated with the entry into and growth in host cells. *Biochem. Biophys. Res. Commun.* **311**:654–659.
- An, W., J. Kim, and R. G. Roeder. 2004. Ordered cooperative functions of PRMT1, p300, and CARM1 in transcriptional activation by p53. *Cell* **117**:735–748.
- Bassi, M. T., R. S. Ramesar, B. Caciotti, I. M. Winship, A. De Grandi, M. Riboni, P. L. Townes, P. Beighton, A. Ballabio, and G. Borsani. 1999. X-linked late-onset sensorineural deafness caused by a deletion involving OA1 and a novel gene containing WD-40 repeats. *Am. J. Hum. Genet.* **64**:1604–1616.
- Bauer, U. M., S. Daujat, S. J. Nielsen, K. Nightingale, and T. Kouzarides. 2002. Methylation at arginine 17 of histone H3 is linked to gene activation. *EMBO Rep.* **3**:39–44.
- Bhatti, M. M., and W. J. Sullivan, Jr. 2005. Histone acetylase GCN5 enters the nucleus via importin- $\alpha$  in protozoan parasite *Toxoplasma gondii*. *J. Biol. Chem.* **280**:5902–5908.
- Bohne, W., U. Gross, D. J. Ferguson, and J. Heesemann. 1995. Cloning and characterization of a bradyzoite-specifically expressed gene (*hsp30/bag1*) of *Toxoplasma gondii*, related to genes encoding small heat-shock proteins of plants. *Mol. Microbiol.* **16**:1221–1230.
- Chen, D., H. Ma, H. Hong, S. S. Koh, S. M. Huang, B. T. Schurter, D. W. Aswad, and M. R. Stallcup. 1999. Regulation of transcription by a protein methyltransferase. *Science* **284**:2174–2177.
- Chen, S. L., K. A. Löffler, D. Chen, M. R. Stallcup, and G. E. Muscat. 2002. The coactivator-associated arginine methyltransferase is necessary for muscle differentiation: CARM1 coactivates myocyte enhancer factor-2. *J. Biol. Chem.* **277**:4324–4333.
- Cheng, D., N. Yadav, R. W. King, M. S. Swanson, E. J. Weinstein, and M. T. Bedford. 2004. Small molecule regulators of protein arginine methyltransferases. *J. Biol. Chem.* **279**:23892–23899.
- Darkin-Rattray, S. J., A. M. Gurnett, R. W. Myers, P. M. Dulski, T. M. Crumley, J. J. Allocco, C. Cannova, P. T. Meinke, S. L. Colletti, M. A. Bednarek, S. B. Singh, M. A. Goetz, A. W. Dombrowski, J. D. Polishook, and D. M. Schmatz. 1996. Apicidin: a novel antiprotozoal agent that inhibits parasite histone deacetylase. *Proc. Natl. Acad. Sci. USA* **93**:13143–13147.
- Daujat, S., U. M. Bauer, V. Shah, B. Turner, S. Berger, and T. Kouzarides. 2002. Crosstalk between CARM1 methylation and CBP acetylation on histone H3. *Curr. Biol.* **12**:2090–2097.
- Denton, H., C. W. Roberts, J. Alexander, K. W. Thong, and G. H. Coombs. 1996. Enzymes of energy metabolism in the bradyzoites and tachyzoites of *Toxoplasma gondii*. *FEMS Microbiol. Lett.* **137**:103–108.
- de Ruijter, A. J., A. H. van Gennip, H. N. Caron, S. Kemp, and A. B. van Kuilenburg. 2003. Histone deacetylases (HDACs): characterization of the classical HDAC family. *Biochem. J.* **370**:737–749.
- Donald, R. G., D. Carter, B. Ullman, and D. S. Roos. 1996. Insertional tagging, cloning, and expression of the *Toxoplasma gondii* hypoxanthine-xanthine-guanine phosphoribosyltransferase gene. Use as a selectable marker for stable transformation. *J. Biol. Chem.* **271**:14010–14019.
- 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.
- Eberharter, A., S. John, P. A. Grant, R. T. Utley, and J. L. Workman. 1998. Identification and analysis of yeast nucleosomal histone acetyltransferase complexes. *Methods* **15**:315–321.
- Ferguson, D. J. 2004. Use of molecular and ultrastructural markers to evaluate stage conversion of *Toxoplasma gondii* in both the intermediate and definitive host. *Int. J. Parasitol.* **34**:347–360.
- Ferro, M., D. Salvi, S. Brugiere, S. Miras, S. Kowalski, M. Louwagie, J. Garin, J. Joyard, and N. Rolland. 2003. Proteomics of the chloroplast envelope membranes from *Arabidopsis thaliana*. *Mol. Cell. Proteomics* **2**:325–345.
- Forsberg, E. C., and E. H. Bresnick. 2001. Histone acetylation beyond promoters: long-range acetylation patterns in the chromatin world. *Bioessays* **23**:820–830.
- 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.
- Gastens, M. H., and H. G. Fischer. 2002. *Toxoplasma gondii* eukaryotic translation initiation factor 4A associated with tachyzoite virulence is down-regulated in the bradyzoite stage. *Int. J. Parasitol.* **32**:1225–1234.
- Grant, P. A., A. Eberharter, S. John, R. G. Cook, B. M. Turner, and J. L. Workman. 1999. Expanded lysine acetylation specificity of Gcn5 in native complexes. *J. Biol. Chem.* **274**:5895–5900.
- Gray, S. G., and T. J. Ekstrom. 2001. The human histone deacetylase family. *Exp. Cell Res.* **262**:75–83.
- Guenther, M. G., W. S. Lane, W. Fischle, E. Verdine, M. A. Lazar, and R. Shiekhattar. 2000. A core SMRT corepressor complex containing HDAC3 and TBL1, a WD40-repeat protein linked to deafness. *Genes Dev.* **14**:1048–1057.
- Guenther, M. G., J. Yu, G. D. Kao, T. J. Yen, and M. A. Lazar. 2002. Assembly of the SMRT-histone deacetylase 3 repression complex requires the TCP-1 ring complex. *Genes Dev.* **16**:3130–3135.
- Hakimi, M. A., D. A. Bochar, J. A. Schmiesing, Y. Dong, O. G. Barak, D. W. Speicher, K. Yokomori, and R. Shiekhattar. 2002. A chromatin remodelling complex that loads cohesin onto human chromosomes. *Nature* **418**:994–998.
- Hettmann, C., and D. Soldati. 1999. Cloning and analysis of a *Toxoplasma gondii* histone acetyltransferase: a novel chromatin remodelling factor in apicomplexan parasites. *Nucleic Acids Res.* **27**:4344–4352.
- Ishizuka, T., and M. A. Lazar. 2003. The N-CoR/histone deacetylase 3 complex is required for repression by thyroid hormone receptor. *Mol. Cell. Biol.* **23**:5122–5131.
- Jenuwein, T., and C. D. Allis. 2001. Translating the histone code. *Science* **293**:1074–1080.
- Joshi, M. B., D. T. Lin, P. H. Chiang, N. D. Goldman, H. Fujioka, M. Aikawa, and C. Syin. 1999. Molecular cloning and nuclear localization of a histone deacetylase homologue in *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **99**:11–19.
- Jung, C., C. Y. Lee, and M. E. Grigg. 2004. The SRS superfamily of *Toxoplasma* surface proteins. *Int. J. Parasitol.* **34**:285–296.
- Kim, J., J. Lee, N. Yadav, Q. Wu, C. Carter, S. Richard, E. Richie, and M. T. Bedford. 2004. Loss of CARM1 results in hypomethylation of thymocyte cyclic AMP-regulated phosphoprotein and deregulated early T cell development. *J. Biol. Chem.* **279**:25339–25344.
- Kissinger, J. C., B. Gajria, L. Li, I. T. Paulsen, and D. S. Roos. 2003. ToxoDB: accessing the *Toxoplasma gondii* genome. *Nucleic Acids Res.* **31**:234–236.
- Knoll, L. J., and J. C. Boothroyd. 1998. Isolation of developmentally regulated genes from *Toxoplasma gondii* by a gene trap with the positive and negative selectable marker hypoxanthine-xanthine-guanine phosphoribosyltransferase. *Mol. Cell. Biol.* **18**:807–814.
- Lekutis, C., D. J. Ferguson, and J. C. Boothroyd. 2000. *Toxoplasma gondii*: identification of a developmentally regulated family of genes related to SAG2. *Exp. Parasitol.* **96**:89–96.
- Li, J., J. Wang, Z. Nawaz, J. M. Liu, J. Qin, and J. Wong. 2000. Both corepressor proteins SMRT and N-CoR exist in large protein complexes containing HDAC3. *EMBO J.* **19**:4342–4350.
- Lyons, R. E., R. McLeod, and C. W. Roberts. 2002. *Toxoplasma gondii* tachyzoite-bradyzoite interconversion. *Trends Parasitol.* **18**:198–201.
- Ma, H., C. T. Baumann, H. Li, B. D. Strahl, R. Rice, M. A. Jelinek, D. W. Aswad, C. D. Allis, G. L. Hager, and M. R. Stallcup. 2001. Hormone-dependent, CARM1-directed, arginine-specific methylation of histone H3 on a steroid-regulated promoter. *Curr. Biol.* **11**:1981–1985.
- Manger, I. D., A. B. Hehl, and J. C. Boothroyd. 1998. The surface of *Toxoplasma* tachyzoites is dominated by a family of glycosylphosphatidylinositol-anchored antigens related to SAG1. *Infect. Immun.* **66**:2237–2244.
- Matrajt, M., R. G. Donald, U. Singh, and D. S. Roos. 2002. Identification and characterization of differentiation mutants in the protozoan parasite *Toxoplasma gondii*. *Mol. Microbiol.* **44**:735–747.
- Matrajt, M., C. D. Platt, A. D. Sagar, A. Lindsay, C. Moulton, and D. S. Roos. 2004. Transcript initiation, polyadenylation, and functional promoter mapping for the dihydrofolate reductase-thymidylate synthase gene of *Toxoplasma gondii*. *Mol. Biochem. Parasitol.* **137**:229–238.
- Messina, D. N., J. Glasscock, W. Gish, and M. Lovett. 2004. An ORFeome-

- based analysis of human transcription factor genes and the construction of a microarray to interrogate their expression. *Genome Res.* **14**:2041–2047.
43. **Monroy, M. A., N. M. Schott, L. Cox, J. D. Chen, M. Ruh, and J. C. Chrivia.** 2003. SNF2-related CBP activator protein (SRCAP) functions as a coactivator of steroid receptor-mediated transcription through synergistic interactions with CARM-1 and GRIP-1. *Mol. Endocrinol.* **17**:2519–2528.
  44. **Nallani, K. C., and W. J. Sullivan, Jr.** 2005. Identification of proteins interacting with Toxoplasma SRCAP by yeast two-hybrid screening. *Parasitol. Res.* **95**:236–242.
  45. **Navarro, M., G. A. Cross, and E. Wirtz.** 1999. Trypanosoma brucei variant surface glycoprotein regulation involves coupled activation/inactivation and chromatin remodeling of expression sites. *EMBO J.* **18**:2265–2272.
  46. **Pijnappel, W. W., D. Schaft, A. Roguev, A. Shevchenko, H. Tekotte, M. Wilm, G. Rigaut, B. Seraphin, R. Aasland, and A. F. Stewart.** 2001. The *S. cerevisiae* SET3 complex includes two histone deacetylases, Hos2 and Hst1, and is a meiotic-specific repressor of the sporulation gene program. *Genes Dev.* **15**:2991–3004.
  47. **Santos-Rosa, H., R. Schneider, A. J. Bannister, J. Sherriff, B. E. Bernstein, N. C. Emre, S. L. Schreiber, J. Mellor, and T. Kouzarides.** 2002. Active genes are tri-methylated at K4 of histone H3. *Nature* **419**:407–411.
  48. **Schneider, R., and A. J. Bannister.** 2002. Protein N-methyltransferase assays in the study of gene transcription. *Methods* **26**:226–232.
  49. **Singh, U., J. L. Brewer, and J. C. Boothroyd.** 2002. Genetic analysis of tachyzoite to bradyzoite differentiation mutants in *Toxoplasma gondii* reveals a hierarchy of gene induction. *Mol. Microbiol.* **44**:721–733.
  50. **Spies, C., A. S. Meyer, S. Reissmann, and J. Frydman.** 2004. Mechanism of the eukaryotic chaperonin: protein folding in the chamber of secrets. *Trends Cell. Biol.* **14**:598–604.
  51. **Strahl, B. D., S. D. Briggs, C. J. Brame, J. A. Caldwell, S. S. Koh, H. Ma, R. G. Cook, J. Shabanowitz, D. F. Hunt, M. R. Stallcup, and C. D. Allis.** 2001. Methylation of histone H4 at arginine 3 occurs in vivo and is mediated by the nuclear receptor coactivator PRMT1. *Curr. Biol.* **11**:996–1000.
  52. **Sullivan, W. J., Jr., M. A. Monroy, W. Bohne, K. C. Nallani, J. Chrivia, P. Yaciuk, C. K. Smith II, and S. F. Queener.** 2003. Molecular cloning and characterization of an SRCAP chromatin remodeling homologue in *Toxoplasma gondii*. *Parasitol. Res.* **90**:1–8.
  53. **Sullivan, W. J., Jr., and C. K. Smith II.** 2000. Cloning and characterization of a novel histone acetyltransferase homologue from the protozoan parasite *Toxoplasma gondii* reveals a distinct GCN5 family member. *Gene* **242**:193–200.
  54. **Templeton, T. J., L. M. Iyer, V. Anantharaman, S. Enomoto, J. E. Abrahamte, G. M. Subramanian, S. L. Hoffman, M. S. Abrahamsen, and L. Aravind.** 2004. Comparative analysis of apicomplexa and genomic diversity in eukaryotes. *Genome Res.* **14**:1686–1695.
  55. **Vogelauer, M., J. Wu, N. Suka, and M. Grunsteins.** 2000. Global histone acetylation and deacetylation in yeast. *Nature* **408**:495–498.
  56. **Wang, H., Z. Q. Huang, L. Xia, Q. Feng, H. Erdjument-Bromage, B. D. Strahl, S. D. Briggs, C. D. Allis, J. Wong, P. Tempst, and Y. Zhang.** 2001. Methylation of histone H4 at arginine 3 facilitating transcriptional activation by nuclear hormone receptor. *Science* **293**:853–857.
  57. **Weiss, L. M., and K. Kim.** 2000. The development and biology of bradyzoites of *Toxoplasma gondii*. *Front. Biosci.* **5**:D391–D405.
  58. **Weiss, L. M., D. Laplace, P. M. Takvorian, H. B. Tanowitz, A. Cali, and M. Wittner.** 1995. A cell culture system for study of the development of *Toxoplasma gondii* bradyzoites. *J. Eukaryot. Microbiol.* **42**:150–157.
  59. **Weiss, V. H., A. E. McBride, M. A. Soriano, D. J. Filman, P. A. Silver, and J. M. Hogle.** 2000. The structure and oligomerization of the yeast arginine methyltransferase, Hmt1. *Nat. Struct. Biol.* **7**:1165–1171.
  60. **Xu, W., H. Cho, S. Kadam, E. M. Banayo, S. Anderson III, J. R. Yates, B. M. Emerson, and R. M. Evans.** 2004. A methylation-mediator complex in hormone signaling. *Genes Dev.* **18**:144–156.
  61. **Yadav, N., J. Lee, J. Kim, J. Shen, M. C. Hu, C. M. Aldaz, and M. T. Bedford.** 2003. Specific protein methylation defects and gene expression perturbations in coactivator-associated arginine methyltransferase 1-deficient mice. *Proc. Natl. Acad. Sci. USA* **100**:6464–6468.
  62. **Yang, S., and S. F. Parmley.** 1997. *Toxoplasma gondii* expresses two distinct lactate dehydrogenase homologous genes during its life cycle in intermediate hosts. *Gene* **184**:1–12.
  63. **Yoon, H. G., D. W. Chan, Z. Q. Huang, J. Li, J. D. Fondell, J. Qin, and J. Wong.** 2003. Purification and functional characterization of the human N-CoR complex: the roles of HDAC3, TBL1 and TBLR1. *EMBO J.* **22**:1336–1346.
  64. **Zhang, J., M. Kalkum, B. T. Chait, and R. G. Roeder.** 2002. The N-CoR-HDAC3 nuclear receptor corepressor complex inhibits the JNK pathway through the integral subunit GPS2. *Mol. Cell* **9**:611–623.
  65. **Zhang, X., L. Zhou, and X. Cheng.** 2000. Crystal structure of the conserved core of protein arginine methyltransferase PRMT3. *EMBO J.* **19**:3509–3519.

## Supplemental Data

### Experimental procedures

#### Construction of *T. gondii* and *E. coli* expression vectors

The *T. gondii* RH (EP) cDNA expression library in  $\lambda$ ZAPII was obtained from NIH AIDS reagents (kindly provided by D.S. Roos). The library was used to amplify by PCR both TgPRMT1 and TgCARM1 cDNAs. The coding sequence of TgCARM1 and TgPRMT1 was amplified by PCR to introduce HindIII and XhoI sites at the start and the stop codons, respectively. Primers forward (5'-ACCCAAGCTTGACGCGTCCGTCGTCGAAACG-3') and reverse (5'-ACCGCTCGAGTTACCTCCCTGAGAACACTCG-3') for TgCARM1, and forward (5'-CCCAAGCTTAGTTCCAGCAGCCCCAACTC-3') and reverse (5'-CCGCTCGAGCTATCGGAGTCTGTAGAACTGG-3') for TgPRMT1, were used for PCR amplification. PCR products were cloned into the bacterial expression vectors pET28b+ (Novagen) to create pNS2 (His-TgCARM1) and pMAH2 (His-TgPRMT1). The coding sequence of TgCARM1 was also amplified by PCR (Forward primer: 5'-GGATCCGACGCGTCCGTCGTCGAAACGTAC-3'; Reverse primer: 5'-ACCGCTC GAGTTACCTCCCTGAGAACACTCG-3') to introduce BamHI and XhoI sites, respectively. The PCR product allowed cloning of TgCARM1 in frame with an N-terminal HA-FLAG tag into the *T. gondii* expression vector pMAH12. The vector pMAH12 used in this work carries the promoter region and the 3'-flanking sequence of the *SAG1* gene. A HA-FLAG dual tag was also introduced to allow N- and C-terminal tag fusions. Tachyzoite RNAs were used to amplify by RT-PCR both TgHDAC3 and TgTBL1 cDNAs. The coding sequences were amplified by PCR to introduce BamHI and PstI sites for TgHDAC3 and BglII and HindIII sites for TgTBL1. Primers forward (5'-GGATCCGCGCTCAGTGCGCTGCGCAAGC-3') and reverse (5'-CTGCAGGATCGG AACTTTTGGTCTCTG-3') for TgHDAC3, and forward (5'-AGATCTCGG ATCCACTTGTCTTCGGACGAAGTG-3') and reverse (5'-AAGCTTCGTTACGCA GAGTCTTCCAGCTTCGC-3') for TgTBL1, were used for PCR amplification. The PCR product allowed cloning of TgHDAC3 and TgTBL1 in frame with either a C- or N-terminal HA-FLAG tag into the *T. gondii* expression vector pMAH14.

### **Recombinant proteins rTgCARM1 and rTgPRMT1**

The recombinant proteins were expressed in *E. coli* strain BL21 (Novagen). Bacteria were transformed with the corresponding vectors and grown in Luria-Bertoni medium at 37 °C to an OD<sub>600</sub> of ~0.6. The cells were induced with 0.5 mM IPTG and grown for an additional period of 3 hours at 37 °C. Bacteria were harvested by centrifugation and lysed by mild sonication at 4 °C in 20 mM HEPES (pH 7.9), 500 mM KCl, 0.1% (v/v) Nonidet P-40, 10% (v/v) glycerol, 5 mM imidazole, and 0.1 mM PMSF. The lysates were cleared by centrifugation and were loaded onto a 0.5-ml column of nitriloacetic-acid-agarose containing immobilized Ni<sup>2+</sup> ions (Ni<sup>2+</sup>-NTA-agarose, Qiagen). The column was washed with 10 column volumes of lysis buffer supplemented with 20 mM imidazole and eluted with 5-ml linear gradient of imidazole (20-300 mM) in lysis buffer. Fractions containing the recombinant proteins were pooled and dialyzed against PBS reaction buffer with 10% glycerol and 0.1 mM PMSF.

### **Chromatographic enrichment of nucleosome remodeling activities.**

The parasite enriched chromatin-remodeling activities were assessed using the restriction enzyme accessibility assay, which relies on the ability of DNA incorporated into chromatin to be refractory to restriction endonuclease digestion (Barak *et al.* 2003). Chromatin-remodeling activities were enriched as following. Whole cell extract (2 grams) from wild-type RH strain was loaded on a 250 ml column of phosphocellulose (P11, Whatman) and fractionated stepwise by the indicated KCl concentration in buffer A (20 mM Tris-HCl, pH 7.9, 0.2 mM EDTA, 10 mM βME, 10% glycerol, 0.2 mM PMSF). 80 mg of the P11 0.5 M KCl fraction was loaded on a 1 ml HiTrap DEAE FF column (Pharmacia) and eluted with 0.5 M KCl in buffer A. The 0.5 M KCl elution (40 mg) was dialyzed to 100 mM KCl in Buffer A, and loaded on Hiprep 16/10 Heparin FF (Pharmacia). The column was resolved using a linear 20 column volume gradient of 100 to 500 mM KCl in Buffer A. Chromatin remodeling-enriched fractions were used for the remodeling assay.

## **Reference**

Barak O, Lazzaro MA, Lane WS, Speicher DW, Picketts DJ, Shiekhattar R. Isolation of human NURF: a regulator of Engrailed gene expression. EMBO J. 2003 Nov 17;22(22):6089-100.

**Table 1 - ChIP Primers**

<b>Gene</b>	<b>Forward</b>	<b>Reverse</b>
GRA1	5-GTACTCGTACGAATACCAATCGCTG-3	5-GCTTTACCGATTGCCCAATGCGAAGTG-3
GRA2	5-GGTTAATTACCTACGCACCACGAAGG-3	5-GACCCTTTTCTGTCATCAGTTTCGAGGAC-3
GRA3	5-GAGAGGTTTCATTTCAAACATCCTAATGGC-3	5-CGTAGGTGCAGGTGTCCCAGTGTCTGTC-3
GRA4	5-CAAGAGCGCGACTAAATTTTACGTAGGC-3	5-CGAAACTCCGGGTATGCGTTGGTTGTG-3
GRA5	5-GTCGCATGACTCCCTCAGGTGGTTAG-3	5-CAAAATAAAATGACTGACGGTAGGATC-3
GRA6	5-GCTAAATTTTCGAAAGAGGATTGTTGACG-3	5-CGACGACGCAAATATTCTACTAATGGG-3
GRA7	5-CTTAAAAGCAGTTGGGTAGGTCGAAGC-3	5-GTCAATGGTAAACGAAATCAGGGATCTC-3
GRA8	5-CTTGTTTTTCAGCAGTGCCGTGTTTGTG-3	5-GGATATTTTCTGGTGGAGATCGAATCAC-3
GRA9	5-CATTTCTGTGCTTCTTGGGGGTTTATG-3	5-GACGACTCTCTTTGATGGCGACGGTTC-3
eIF4A	5-CACGCACTCCTCACTCGTCTGTAAG-3	5-CTTCGGAACCCTCGAGGAAATGATGAG-3
DHFR	5-CATGCGCGAACTACGTGGACATTAAG-3	5-CATTTTCGCGCACGGCAGTCAGATAAC-3
DEOC-A	5-GGACTTTACCCAGAGTAGAGATACC-3	5-CTTGTGTACGGTGTCTCTTTCAGC-3
DEOC-B	5-GCTGAAAGAGAACACCCGTAACAAAG-3	5-GAGAACGGTGTGCAAATGGAGCTAC-3
LDH2-A	5-GTACTGCCCGGAGCATGCGTCC-3	5-GTAAGTATGTTGCTACGAGAAGC
LDH2-B	5-GTAAGCCTGTTTAGCTGCCTTTCG-3	5-GGACGCATGCTCCGGGCAGTAC
SAG2C-A	5-CAGTAGCTGGTCAAATCGTGTAGAAG-3	5-GAAGTAGGCGCGCCGGTCTTTCAGC-3
SAG2C-B	5-GCTGAAGGACCGGCGCGCCTAGTTC-3	5-GACGATATACGGCTTTCTGCGATAG-3
SAG2C-C	5-CTATCGCAGAAAAGCCGTATATCGTC-3	5-CAAAGTTATCACTTGTGTGCTGATC-3
SAG2C-D	5-GATACAGCAACAAGTGATAAGTTTG-3	5-GTTCGCGCACACTACCAGTAGAAAAG-3
SAG1-A	5-GGATGCGGGACACGCCTTCCTG-3	5-CAGCTTCTACAGGTGCACATTG-3
SAG1-B	5-CAGTGCTCACAGTGCCGCAACG-3	5-CAGCTTCTACAGGTGCACATTG-3
SAG2A-A	5-CGTAAAGCGCGCAGATGACAG-3	5-GTTCCGACAAGACAAAGATTG-3
SAG2A-B	5-GAGAGGTACTTCTGCAACGAG-3	5-CTGTCATCTGCGCGCTTTACG-3
BAG1 locus - A	5-CTACTGATCAGTGCCTCTACGAGATG-3	5-GACGCTCTCACAGACGACGC
BAG1 locus - B	5-CTCACTCGCCAGTCGAAAGCC-3	5-CATCTCGTAGAGTGCCTGATCAGTAG
BAG1 locus - C	5-GCGATAGCGTAGCAAGAAATG-3	5-GGCTTTTCGACTGGCGAGTGAG
BAG1 locus - D	5-CAAGGAGGAGCTGCTCCGTAG-3	5-CATTTCTTGTACGCTATCGC
BSR4-A	5-CTTGCCTGTACTTTTCGAGCACATGG-3	5-GATTTGCATCGGCGCTTATCAGTAG-3
BSR4-B	5-CTACTGATAAGCGCCGATGCAAATC-3	5-GTGGAGGCAAGCAAATCTCACGACAG-3
BSR4-C	5-GCATATTGGTATTCGTACGTCAATTTG-3	5-GAACGCGAAGCGCTCTCATTCATTTAG-3
BSR4-D	5-CTAAATGAATGAGAGCGTTCCGCTT-3	5-CAAAGTAGTACTCAGGTGTCATCTCG-3
BSR4-E	5-CGAGATGACACCTGAGTACTAGTTTG-3	5-CTGTGCGACAACGAGAGATTTCTGC-3
BSR4-F	5-GCAGAAATCTCTCGTTGTGCGACAG-3	5-GGCGAATTCAGAGCGTGTGCTAATG-3
BSR4-G	5-GATGTGAAAATGTGCGTCATAAAAC-3	5-GCTATAGCTGTAGTCTGATCGTTCG-3
BSR4-H	5-GCGACGATCAGACTACAGCTATAGC-3	5-GGTGAAAATAAGATATTGGGACGAG-3
BSR4-I	5-CTCGTCCCAATATCTTATTTTACC-3	5-CTCTTTAATGAATTCAGTCTACAGCC-3
BSR4-J	5-CTTATAGTTTCTGCTATTGGCTGTG-3	5-GGAAACCGTCTCAGTGAACAAGGC-3
BSR4-K	5-CATACTTTTCTGTGTTTCTTGC-3	5-CTCTGCAAGTATGGGCTCAGGTAG-3
BSR4-L	5-CTACCTGAGCCATACTTGCAGAGAG-3	5-GGTTCTAGTCTACAGTCACAAACGC-3
BSR4-M	5-GATGCGTGCCTCAGCGTTGGAG-3	5-CCGAAGTGTGAGTACGCAACGCTTGG-3
BSR4-N	5-CACGACCATGCTTCACTGTAACGCTG-3	5-GTTTTCAAAGGAGAGTCGTGAGAG-3

**Supplemental Table 2 – The TgCRC subunits**

<b>Bands</b>	<b>Name</b>	<b>TOXODB annotation<sup>1</sup></b>	<b>Genbank</b>
TgCRC-45	<b>Actin</b>		AAC13766
TgCRC-62	<b>TgHDAC3</b>		DQ004745
	<b><u>TRiC complex</u></b>		
TgCRC-64	TCP-1-theta	TgTigrScan_4156	
	TCP-1-delta	TgGlmHMM_4680	
	TCP-1-eta	TgTigrScan_0599	
TgCRC-66	TCP-1-alpha	TgTigrScan_6329	
	TCP-1-epsilon	TgTigrScan_3211	
	TCP-1-gamma	TgTigrScan_7268	
TgCRC-68	TCP-1-beta	TgTigrScan_2113	
	TCP-1-zeta	TgTigrScan_4832	
TgCRC-75	<b>TgHSP70a</b>		AAF23321
TgCRC-80	<b>TgHSP70b</b>		AAC72001
TgCRC-100	<b>TgTBL1</b>		DQ004746
TgCRC-120	<b>TgHSP90b</b>		DQ004747
<b>TgCRC-230</b>	<b>TgCRC-230</b>	TgTigrScan_2931 + TgTigrScan_2930	
<b>TgCRC-350</b>	<b>TgCRC-350</b>	TgTigrScan_6883	

1 <http://www.toxodb.org/ToxODB.shtml> (reference 16)

# CHAPITRE II



### 1. Etude moléculaire de l'histone déacétylase TgHDAC2

Le gène TgHDAC2 (7081 pbs, Figure 8A) code pour un ARN messager de 1929 paires de base avec un cadre de lecture ouvert, codant pour une protéine de 68 KDa. L'ADNc a été amplifié par RT-PCR à partir d'ARN de tachyzoïtes de la souche RH puis sous-cloné dans le vecteur parasitaire pMAH14 pour former le plasmide recombinant pNS13. Le vecteur pMAH14 permet d'exprimer des protéines chimériques chez le toxoplasme sous le contrôle du promoteur fort *GRA1*, et offre la possibilité de faire une fusion traductionnelle avec la double étiquette HA-FLAG.

Notre étude phylogénétique prédit que TgHDAC2 est une histone déacétylase de classe I. Cependant la sur-expression ectopique de TgHDAC2-HAFlag dans des toxoplasmes transgéniques stables montre une localisation cytoplasmique dans les tachyzoïtes (Figure 9A), alors que la classe I est connu pour être nucléaire (de Ruijter *et al.*, 2003). Nous avons précédemment montré que TgHDAC3 qui appartient aussi à la classe I est localisée exclusivement dans le compartiment nucléaire (Saksouk *et al.*, 2005; Figure 2A).

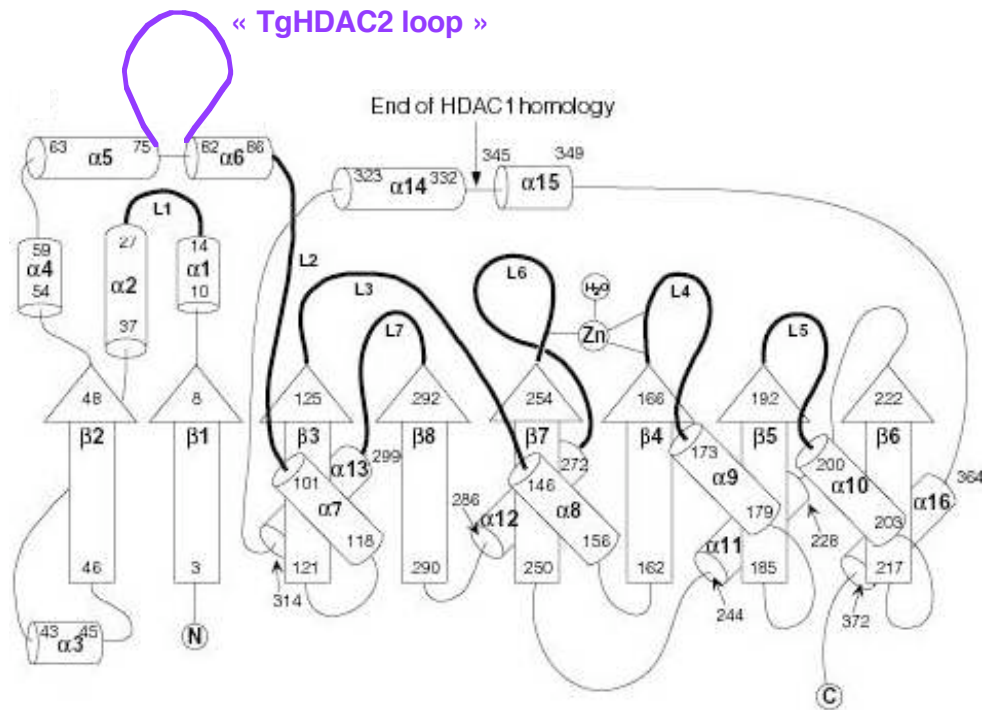
### 2. Détermination de l'interactome spécifique de TgHDAC2

Les HDAC sont souvent intégrées dans de grands complexes protéiques avec une taille globulaire parfois supérieure au million de Dalton (Burke et Baniahmad, 2000). La caractérisation des activités enzymatiques natives impose parfois la purification biochimique des complexes qui les contiennent. L'interactome de TgHDAC2 a été déterminé par chromatographie d'affinité et d'exclusion de taille, couplée à de la spectrométrie de masse (*Coll. Jérôme Garin*). L'étiquette Flag a permis d'immunopurifier la protéine chimérique par chromatographie d'affinité  $\alpha$ -Flag selon le schéma de purification décrit en Figure 9B. Les fractions éluées à l'aide du peptide compétiteur FLAG, testées positives par Western Blot ( $\alpha$ -HA), ont été réunies et analysées par chromatographie d'exclusion de taille (Superdex 200, FPLC Akta, Figure 9C).



3640 3660 3680 3700 3720 3740 3760 3780 3800 3820 3840 3860 3880 3900 3920 3940 3960 3980 4000 4020 4040 4060 4080 4100 4120 4140 4160 4180 4200 4220 4240 4260 4280 4300 4320 4340 4360 4380 4400 4420 4440 4460 4480 4500 4520 4540 4560 4580 4600 4620 4640 4660 4680 4700 4720 4740 4760 4780 4800 4820 4840 4860 4880 4900 4920 4940 4960 4980 5000 5020 5040 5060 5080 5100 5120 5140 5160 5180 5200 5220 5240 5260 5280 5300 5320 5340 5360 5380 5400 5420 5440 5460 5480 5500 5520 5540 5560 5580 5600 5620 5640 5660 5680 5700 5720 5740 5760 5780 5800 5820 5840 5860 5880 5900 5920 5940 5960 5980 6000 6020 6040 6060 6080 6100 6120 6140 6160 6180 6200 6220 6240 6260 6280 6300 6320 6340 6360 6380 6400 6420 6440 6460 6480 6500 6520 6540 6560 6580 6600 6620 6640 6660 6680 6700 6720 6740 6760 6780 6800 6820 6840 6860 6880 6900 6920 6940 6960 6980 7000 7020 7040 7060 7080 7100 7120 7140 7160 7180 7200 7220 7240 7260 7280 7300 7320 7340 7360 7380 7400 7420 7440 7460 7480 7500 7520 7540 7560 7580 7600 7620 7640 7660 7680 7700 7720 7740 7760 7780 7800 7820 7840 7860 7880 7900 7920 7940 7960 7980 8000 8020 8040 8060 8080 8100 8120 8140 8160 8180 8200 8220 8240 8260 8280 8300 8320 8340 8360 8380 8400 8420 8440 8460 8480 8500 8520 8540 8560 8580 8600 8620 8640 8660 8680 8700 8720 8740 8760 8780 8800 8820 8840 8860 8880 8900 8920 8940 8960 8980 9000 9020 9040 9060 9080 9100 9120 9140 9160 9180 9200 9220 9240 9260 9280 9300 9320 9340 9360 9380 9400 9420 9440 9460 9480 9500 9520 9540 9560 9580 9600 9620 9640 9660 9680 9700 9720 9740 9760 9780 9800 9820 9840 9860 9880 9900 9920 9940 9960 9980 10000

B)



Source figure: Finnin MS et al., 1999

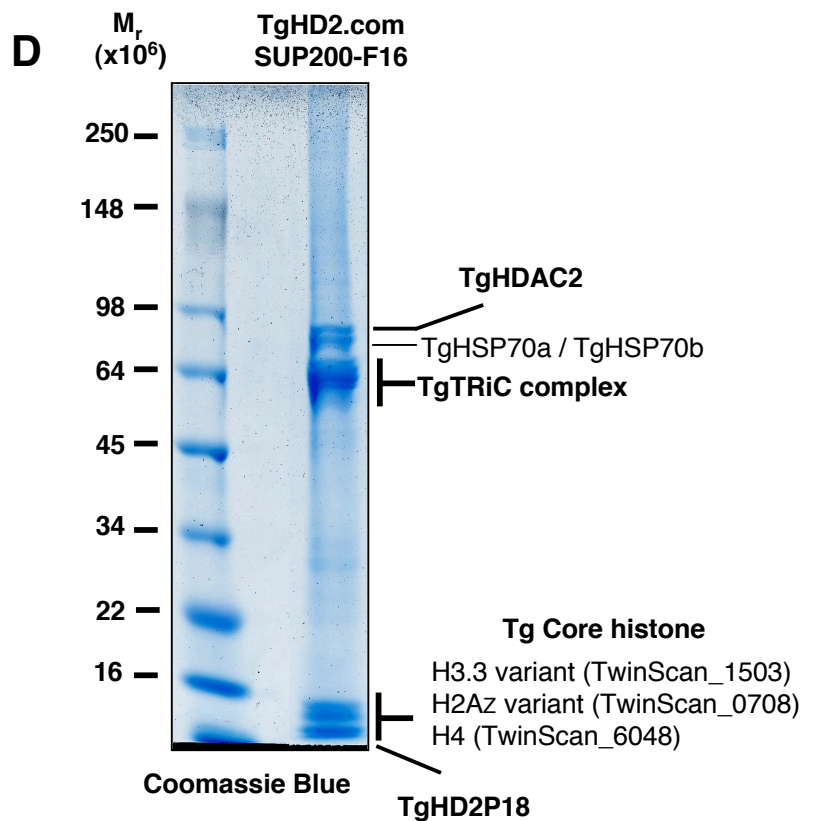
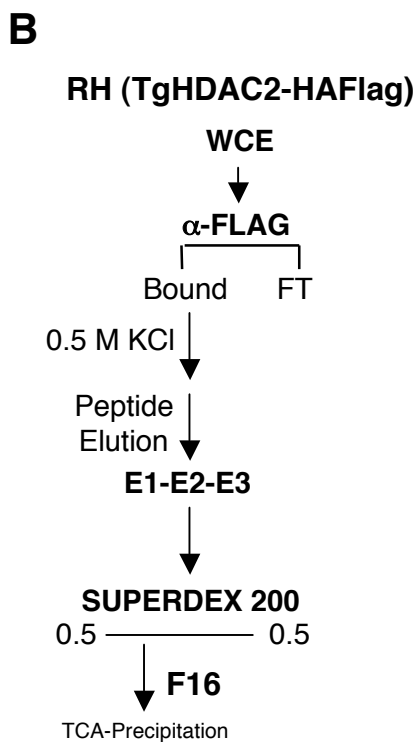
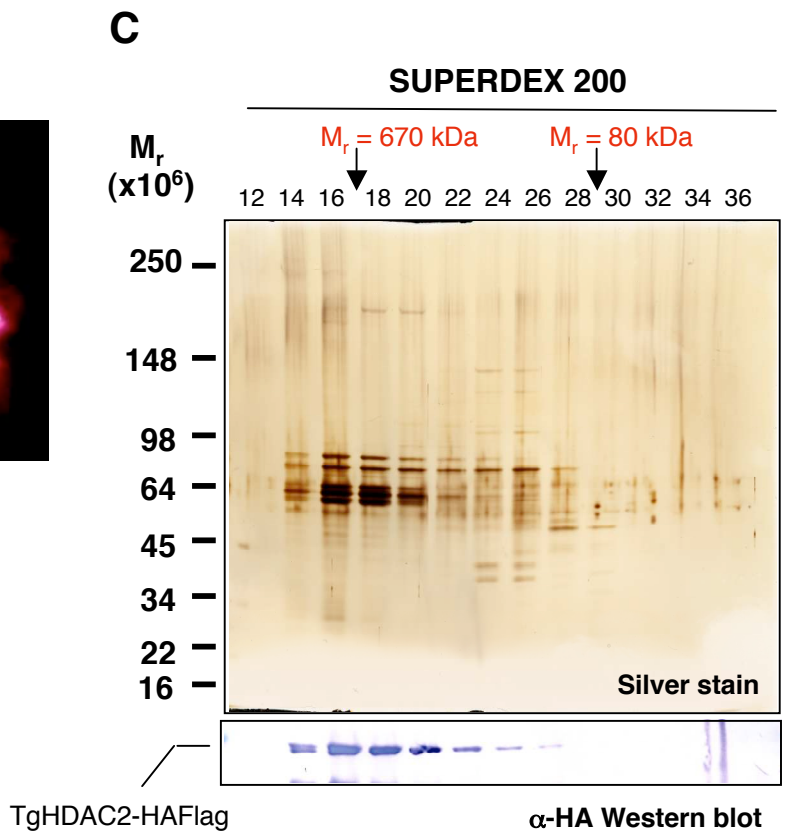
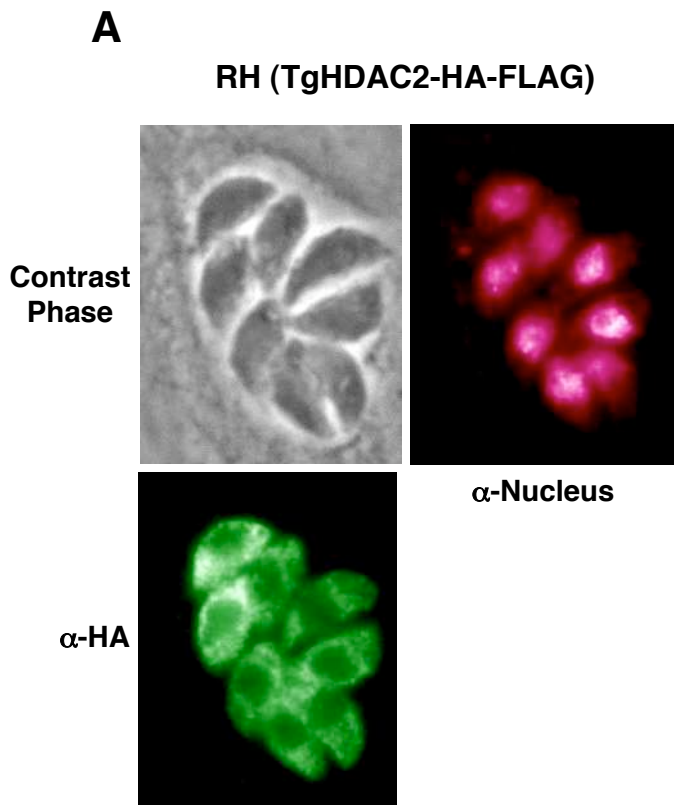
**Figure 8 – A) Organisation génomique du gène *TgHDAC2*.** La position des introns et exons a été déterminée par comparaison de la séquence génomique à l'ARNm mature qui lui correspond, à partir du logiciel *DNA strider*. L'ADN génomique est le produit du séquençage systématique de la souche ME49 (type II; [www.TOXODB.org](http://www.TOXODB.org)) alors que le cDNA provient lui du produit de RT-PCR amplifié à partir d'ARN-messagers de la souche RH (type I). Les différences en nucléotides observées sont dues à la variation génique entre les souches de type I et II (estimée généralement à 2%). En Jaune, sont soulignés les codons d'initiation et de terminaison de la traduction; en rouge, le domaine correspondant à la séquence codante de la boucle insérée. B) **Organisation structurale de la protéine *TgHDAC2*.** Schéma de la localisation potentielle de la boucle dans la structure tridimensionnelle de l'histone déacétylase HDLP de la bactérie hyperthermophile *Aquifex aeolicus*.

L'étude par coloration au nitrate d'argent et par western-blot des fractions issues de la filtration sur gel révèle que la protéine HA-Flag-TgSet13 (68-kDa) est présente dans les fractions 14 à 20 ; ce qui indique qu'elle appartient à un complexe d'une taille globulaire d'environ 700-kDa (Figure 9C). Les différentes protéines de ce complexe ont été identifiées par spectrométrie de masse de type MS-MS en collaboration avec Jérôme Garin (INSERM-CEA) (Figure 9D).

L'analyse des peptides indique la présence de plusieurs protéines du choc thermique HSP (**H**eat **S**hock **P**rotein), TgHSP70a (AAF23321) and TgHSP70b (AAC72001) (Figure 9D). La sur-expression ectopique de protéines dans le toxoplasme induit généralement la co-purification de chaperones de type HSP70 et HSP90 en charge du repliement de la protéine chimérique en excès dans la cellule. On note cependant ici l'absence de TgCRC-120 (HSP90-like protein, DQ004747). Par contre, le TRiC complexe est aussi co-purifié avec TgHDAC2 comme pour TgHDAC3 (Saksouk *et al.*, 2005, Figure 2D) et alors que nous n'observons pas de co-purification du TRiC complexe avec TgHDAC4 ou encore TgHDAC5 (résultats non montrés ici). Le complexe protéique TRiC (aussi nommé TCP-1 ou CCT) est constitué de huit sous-unités protéiques (TCP-1-thêta, delta, êta, alpha, epsilon, gamma, bêta et zêta, Saksouk *et al.*, 2005, supplemental data) qui ont un arrangement similaire au complexe bactérien GroEL (Dunn *et al.*, 2001 ; Hartl et Hayer-Hartl, 2002). Ce complexe utilise l'énergie libérée par la dégradation de l'ATP pour procéder au repliement des histones déacétylases (Melki *et al.*, 1997).

Le plus surprenant dans les résultats de spectrométrie de masse est le séquençage de peptides correspondants à un core histone « variant ». Le noyau du nucléosome est constitué par un premier tétramère H3/H4 qui est d'abord déposé sur l'ADN puis complété par deux hétérodimères H2A-H2B autour duquel 146 paires de bases s'enroulent. *T. gondii* possède les quatre histones constitutifs du nucléosome, TgH3, TgH4, TgH2A et TgH2B et certains de leur variants. Le plus surprenant ici est que nous co-purifions avec TgHDAC2 un core histone contenant deux variants, les histones TgH3.3 et TgH2AZ (Figure 9D).

**Figure 9 - Détermination de l'interactome spécifique de TgHDAC2.** **A)** Immunofluorescence du clone transgénique stable A4 exprimant ectopiquement la protéine chimérique TgHDAC2-HAFlag. La protéine détectée par l'anticorps  $\alpha$ -HA est semblée être exclue du compartiment nucléaire. L'expérience est contrôlée avec un anticorps qui détecte spécifiquement une protéine nucléaire du toxoplasme. L'expérience est réalisée avec des tachyzoïtes et la souche transgénique est issue d'un RH, une souche de type I. **B)** Schéma de purification de la protéine chimérique TgHDAC2-HAFlag. Brièvement un extrait total de protéines issu de la souche RH (TgHDAC2-HAFlag) est enrichi pour la protéine TgHDAC2 par chromatographie d'affinité ( $\alpha$ -FLAG) suivi d'une chromatographie d'exclusion de taille (Superdex 200). **C)** Les protéines issues de la purification sont révélées par coloration au nitrate d'argent et western blot avec un anticorps  $\alpha$ -HA. La protéine TgHDAC2-HAFlag chimérique co-migre avec d'autres protéines dans un complexe appelé TgHD2.com d'une taille globulaire d'environ 700 kDa. **D)** La fraction 16 de la superdex 200 est précipitée au TCA et analysée par coloration au bleu de coomassie. Les bandes sont excisées du gel et soumises à un micro-séquençage peptidique de type MS/MS.



Le micro-séquençage peptidique révèle aussi la présence d'une protéine hypothétique de 18-kDa que nous avons nommé TgHD2P18 (Figure 9D). L'ARNm a été amplifié par RT-PCR et son ADNc correspondant cloné dans pMAH14. La protéine semble être localisée dans la zone périnucléaire (résultats non montrés). Cette étude est poursuivie dans le laboratoire afin de déterminer la relevance de cette interaction avec TgHDAC2.

### 3. Etude de la structure modulaire de la protéine TgHDAC2

Une autre caractéristique majeure de TgHDAC2 nous a interpellé. Cette enzyme présente une boucle unique de 207 acide-aminés (Figure 8B et 10A). En effet, aucune enzyme HDAC de classe I publiée à ce jour ne présente ce type de boucle.

Nous avons d'abord pensé à un épissage alternatif possible qui retire à façon la boucle. Cependant, nous ne sommes pas arrivés à identifier l'ADNc correspondant à l'ARNm alternativement épissés ni dans les tachyzoïtes, ni dans les bradyzoïtes. Sans résoudre l'hypothèse d'une forme « épissé » de TgHDAC2, nous avons construit une protéine chimérique qui élimine la boucle (Figure 10A). L'objectif fut de déterminer si cette boucle influençait sa localisation cellulaire, son activité enzymatique, voire son interactome.

L'ADNc chimérique TgHDAC2 $\Delta$ loop a été sous-cloné dans pMAH14 pour former le vecteur d'expression parasitaire pNS90. L'expression ectopique de TgHDAC2 $\Delta$ loop-HAFlag dans des toxoplasmes transgéniques ne modifie pas la localisation cytoplasmique observée avec TgHDAC2-HAFlag (Figure 10B). La purification conjointe de TgHDAC2 et sa version tronquée TgHDAC2 $\Delta$ loop montre une symétrie parfaite dans l'interactome avec la conservation des HSP70 et du complexe TRiC (Figure 10C). Notons ici que les protéines basiques histone se colorent mal au nitrate d'argent mais leur présence a été confirmée par coloration au bleu de coomassie après précipitation au TCA (résultats non montrés).



**Figure 10 – Characterisation de la « boucle » peptidique.** **A)** Alignement de TgHDAC2, TgHDAC2 $\Delta$ loop avec l'homologue humain HDAC2 (HsHDAC2). Les comparaisons de séquence ont été effectuées en utilisant le logiciel online Clustalw (<http://www.ebi.ac.uk/clustalw/>). **B)** Analyse par immunofluorescence de la localisation des protéines chimériques TgHDAC2-HAFlag et TgHDAC2 $\Delta$ loop-HAFlag dans des tachyzoites. Les toxoplasmes sont issus de clones transgéniques exprimant ectopiquement ces protéines : « FL » pour (RH::TgHDAC2 full-length-HAFlag) et «  $\Delta$  » pour (RH::TgHDAC2 $\Delta$ loop-HAFlag). Les protéines sont détectées par l'anticorps  $\alpha$ -HA. L'expérience est contrôlée avec un anticorps qui détecte spécifiquement une protéine nucléaire du toxoplasme. **C)** Les extraits totaux « FL » et de «  $\Delta$  » sont enrichis pour les protéines TgHDAC2-HAFlag et TgHDAC2 $\Delta$ loop-HAFlag par chromatographie d'affinité ( $\alpha$ -FLAG). Les protéines issues de la purification sont révélées par coloration au nitrate d'argent et western blot avec un anticorps  $\alpha$ -HA. **D)** Les activités enzymatiques sont évaluées pour les produits d'éluion  $\alpha$ -FLAG de « FL » et de «  $\Delta$  ». Les mesures sont indiquées unité relative par  $\mu$ g de protéines.

**A**

```

-----MGRRKSVLYFYDENIGNFYHGPSHPMKPHRVMTDLIQAIDLLQSMDIVKPLKP 55 TgHDAC2
-----MGRRKSVLYFYDENIGNFYHGPSHPMKPHRVMTDLIQAIDLLQSMDIVKPLKP 55 TgHDAC2Δloop
MAYSOGGKVCYIYDGDIGNYYGGQHPMKPHRIRMTNLLMLNYGLYRMEIYRPHKA 60 HsHDAC2
*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*
AVEDELTRFHANDYVDFLRASAQSRLEASLSFRPFLFFPSPPAARASASRPSPPPQAEAGAA 115 TgHDAC2
AVEDELTRFHANDYVDFLRASAQSRLE----- 81 TgHDAC2Δloop
TAEMTKYHSDIYKFLRSIRPDNMS----- 86 HsHDAC2
*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*

AVSAGASGVSVVQSGSRSVRSTLHPLRSESRSVSPATSTFDRSRVPPEGVGADALSSTNA 175 TgHDAC2
----- 175 TgHDAC2Δloop
----- 175 HsHDAC2

RGRTPPKSLGVLDLPGPGGSRAATQKPPFLQASSGDGAHGLGLQCPSAAEETGTGGQAPA 235 TgHDAC2
----- 235 TgHDAC2Δloop
----- 235 HsHDAC2

WAPPGVSPSLRRGEDANGLLSAAFDVEEVAAPVPLSSSLADQMARFNVGEDCPVFDGLW 295 TgHDAC2
-----LADQMARFNVGEDCPVFDGLW 102 TgHDAC2Δloop
-----EYSKQMHIFNVGEDCPAFDGLF 108 HsHDAC2
*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*

EYCQYISGGSIEGARRVVRGEYDWAINVAGLHHGKKHEASGFCYINDCVLAALEFLRFK 355 TgHDAC2
EYCQYISGGSIEGARRVVRGEYDWAINVAGLHHGKKHEASGFCYINDCVLAALEFLRFK 162 TgHDAC2Δloop
EFCQLSTGGSVAGAVKLNQQDTDMAVNWAGLHHAKKYEASGFCYVNDIVLAILLELLKYH 168 HsHDAC2
*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*

HRVLYVDVDDIHHGDGVVEEAFYTSRVLCCSFHKYGDYFPGTGALDDVIGDEGLGYSVNV 415 TgHDAC2
HRVLYVDVDDIHHGDGVVEEAFYTSRVLCCSFHKYGDYFPGTGALDDVIGDEGLGYSVNV 222 TgHDAC2Δloop
QRVLYIDIDIIHHGDGVVEEAFYTDVRMTVTSFHKYGEYFPGTGLRDIGAGKGYAVNFP 228 HsHDAC2
*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*

LSEGVDDAMYTRLFQVMDIMESYRPEAKVLQCGAGSLSGDRLCFNLSIDGHSEAVRY 475 TgHDAC2
LSEGVDDAMYTRLFQVMDIMESYRPEAVLQCGADSLSGDRLCFNLSIDGHSEAVRY 282 TgHDAC2Δloop
MCDGIDDESIGQIFKPIISKVMEMYPQSAVVLQCGADSLSGDRLCFNLTVKGHAKCVEV 288 HsHDAC2
*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*

FCSKDVDPALFLGGGGYTLRNVPRCWAETGVLGSLNPNPEESEYVGYGPDSTLHVR 535 TgHDAC2
FCSKDVDPALFLGGGGYTLRNVPRCWAETGVLGSLNPNPEESEYVGYGPDSTLHVR 342 TgHDAC2Δloop
VTFNPLMLLGGGGYTLRNVARCWYETVAVALDCEIPNELPNY-DYFEYFGDFKLIHS 347 HsHDAC2
*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*

TSMNENQEDSVQHVIERISQTLKDHVRPVGQAISAN----- 573 TgHDAC2
TSMNENQEDSVQHVIERISQTLKDHVRPVGQAISAN----- 380 TgHDAC2Δloop
PSMNTNQTFEYMEKIKQRLFENLRMLPHAPGVQMQAIPEDAVHEDSGDEGEDPDKRIS 407 HsHDAC2
*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*

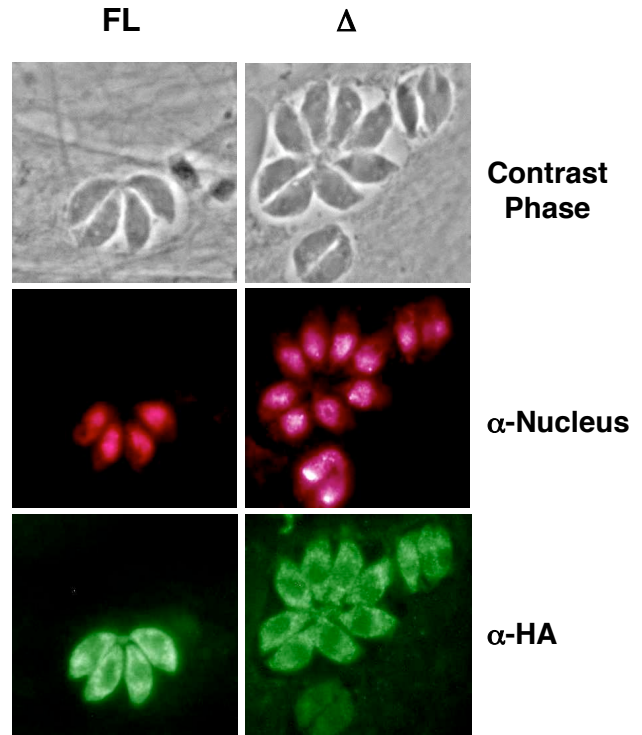
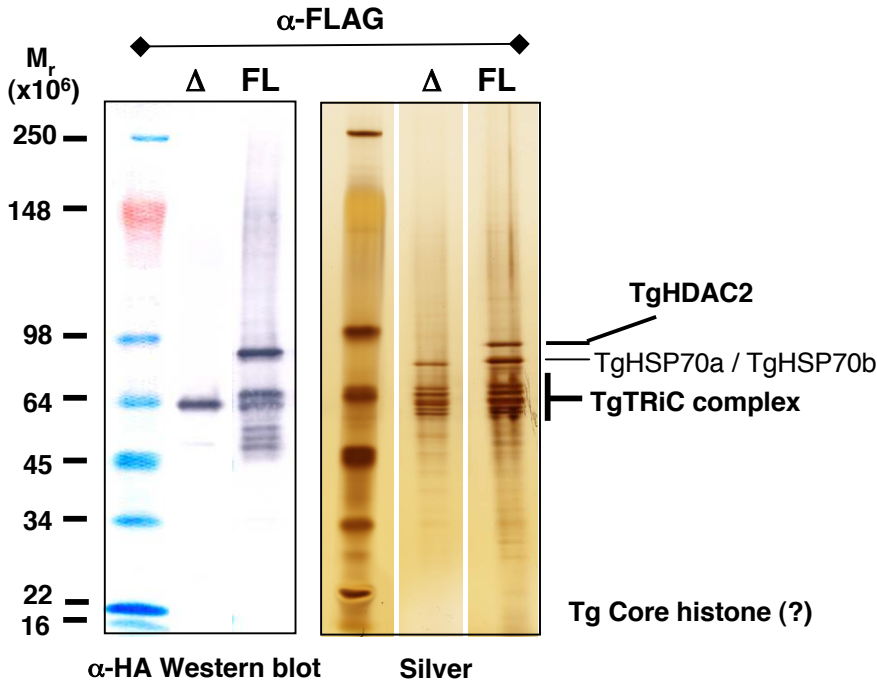
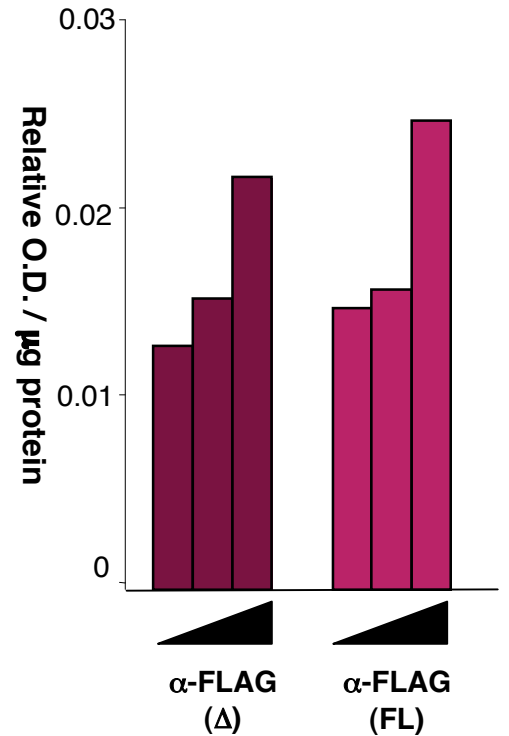
-----YFLDSDVMKSDQ-----IRALLADE-----EEEKA 599 TgHDAC2
-----YFLDSDVMKSDQ-----IRALLADE-----EEEKA 406 TgHDAC2Δloop
IRASDKRIACDEEFSDSEDEGGRRNVADHKKGAKKARIEEDKKETEDKKTQVKEEDKS 467 HsHDAC2
*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*

MDSGDAEDADGAGRFEIISNRNRG 624 TgHDAC2
MDSGDAEDADGAGRFEIISNRNRG 431 TgHDAC2Δloop
KDNSGERTDTKGT-KSEQLSNP--- 488 HsHDAC2
*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*

```

**B**

**FL > RH::TgHDAC2 full-length-HAFlag**  
**Δ > RH::TgHDAC2Δloop-HAFlag**

**C****D**

L'analyse par western blot (anticorps  $\alpha$ -HA) des produits de la purification  $\alpha$ -Flag montre une « instabilité » de la protéine TgHDAC2 qui a tendance à se dégrader en isoformes plus petits. Cette « instabilité » n'est pas observée pour TgHDAC2 $\Delta$ loop (Figure 10C). L'analyse de la composition en acide aminé de la boucle montre un enrichissement des résidus P (13%), D/E (8%), S/T (22%). Or il est décrit dans la littérature que la séquence PEST (Proline, Glutamate, Sérine, Thréonine) est une signature qui dirige les protéines qui les contiennent à la dégradation contrôlée par les Calpain. Les séquences PEST sont retrouvées dans 10% des protéines cellulaires des organismes analysés à ce jour. Elles sont retrouvées le plus souvent dans des protéines dont la fonction est hautement régulée, qui ont de ce fait une demi-vie très courte (0,5 à 2 heures) par rapport aux autres protéines (>24 heures). Le retrait de la séquence PEST de ces protéines augmente en conséquence leur demi-vie (Rechsteiner *et al.*, 1996 ; Mitchell *et al.*, 2003).

La préparation de nos extraits totaux de protéines de Toxoplasme et nos purifications biochimiques sont réalisées en présence de nombreux inhibiteurs de protéases connus pour être normalement efficaces contre les Calpains et les autres protéases. Si l'« instabilité » observée est due à la présence d'un domaine PEST dans la boucle, cela signifie que la dégradation en différents isoformes a eu lieu intracellulairement avant la lyse des tachyzoïtes.

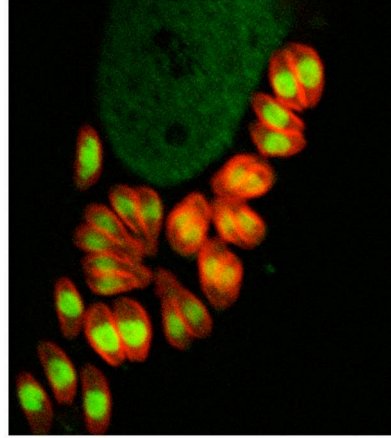
La deuxième hypothèse que nous avons retenue est que cette boucle constitue un domaine *Intein* impliqué dans l'épissure protéinique. Un intein est un segment d'une protéine qui peut s'exciser et rejoindre les parties restantes (les exteins) avec un lien de peptide. Inteins se sont également appelés les "introns de protéine" (Petrokovski *et al.*, 2001). L'épissure Intein-négociée de protéine se produit après que le mRNA ait été traduit en protéine. Cette protéine de précurseur contient trois segments - un N-extein suivi de l'intein suivi d'un C-extein. Après l'épissure a eu lieu, le résultat s'appelle également un extein. Le premier intein a été découvert en 1987. Depuis lors, des inteins ont été trouvés dans chacun des trois domaines de la vie (eukaryotes, bactéries, et

archaea). Dans l'hypothèse où la boucle de TgHDAC2 formerait un domaine intein on peut aisément expliquer l'instabilité de la protéine entière par rapport à la protéine « épissée » (TgHDAC2 $\Delta$ loop, Figure 10C). Cependant, la plupart des inteins rapportés contiennent également un domaine de ribonucléase qui joue un rôle dans la propagation d'intein. Cela n'est pas observé pour TgHDAC2. De plus les domaines Intein ont motifs conservés (Petrokovski *et al.*, 2001), partiellement retrouvés dans la boucle de TgHDAC2.

D'autres expériences seront nécessaires pour déterminer si la boucle de TgHDAC2 est une séquence PEST ou un domaine cryptique Intein. Une fois posée la question de la nature de cette boucle il faudra ensuite en rechercher sa fonction biologique.

La résolution de la structure de l'histone déacétylase HDLP de la bactérie hyperthermophile *Aquifex aeolicus* a permis d'établir un modèle structural pour les enzymes de classe I et leur interaction avec les inhibiteurs spécifiques tels que le TSA et le SAHA (Finnin *et al.*, 1999). Cette structure tridimensionnelle révèle la présence d'une cavité catalytique dans laquelle se loge le substrat acétyl-CoA ou les inhibiteurs TSA ou SAHA (Finnin *et al.*, 1999). La boucle de TgHDAC2 brise la structure de la cavité en insérant ses 200 acide aminés (Figure 11). Les activités enzymatiques de TgHDAC2 et de TgHDAC2 $\Delta$ loop ne laissent pas apparaître de différences majeures (Figure 10D). Mais attention, ces activités ont été déterminées à partir de l'analyse du clivage d'un peptide qui imite la liaison acétyle sur une histone en libérant un fluorochrome. Il se peut que le substrat « naturel » de TgHDAC2 soit affecté dans un sens ou dans un autre par l'intégration de 200 acide aminé dans sa cavité réactionnelle. Ce substrat dit naturel est probablement ce core histone « variant » (H2AZ/H3.3), co-purifié avec l'enzyme TgHDAC2 et TgHDAC2 $\Delta$ loop. Des études enzymologiques avec le core histone variant acétylé comme substrat permettront de déterminer si la boucle influence la reconnaissance du substrat. Et vu que la cavité réactionnelle peut aussi accueillir les inhibiteurs, il sera intéressant de déterminer si l'enzyme perd ou gagne en sensibilité au contact de ces inhibiteurs.

# DISCUSSION



### Avant-Propos

Le décryptage du « code épigénétique universel » n'en est qu'à ses prémices et son implication dans la régulation génique chez *Toxoplasma gondii* et son phylum est peu, voire pas, étudiée. Ce manuscrit apporte les premières lectures d'un code parasitaire et son impact sur la différenciation parasitaire chez l'organisme modèle *T. gondii*.

### 1. L'hypothèse épigénétique

Le séquençage systématique des génomes des parasites *Apicomplexa* et l'analyse prédictive de leur contenu débouche aujourd'hui sur des observations très inattendues. Ces parasites ont conservé une machinerie transcriptionnelle de type eucaryotique au centre duquel on retrouve bien les ARN polymérases I, II et III. Cependant, il existe des différences considérables entre les machineries nucléaires de transcription parasitaires et celles des autres eucaryotes. Ainsi les facteurs homologues à TFIIA, TFIIE-beta, TFIIF et la plupart des facteurs associés à TBP (TAFs) ne sont pas trouvées dans les bases de données des génomes parasitaires (Hakimi, communication personnel ; Meissner *et al.*, 2005). En particulier l'absence de TAFs conservés est une explication plausible à l'incapacité du transactivator viral VP16 d'activer la transcription chez ces parasites (Meissner *et al.*, 2005). Les Apicomplexes pourraient donc avoir maintenu des machines de transcription plus simplistes des autres cellules eucaryotiques par la perte progressive de certains gènes au cours de l'évolution. Par exemple, le facteur général TFIIA-alpha a pu être identifié chez *C. parvum* (accession number EAK87567), tandis qu'aucun homologue n'est identifiable chez ses cousins *Plasmodium* ou *T. gondii*. Alternativement, quelques gènes homologues pourraient ne pas être identifiables par nos algorithmes bioinformatiques à cause d'une conservation faible de leurs domaines protéiques. La réalité est encore plus criante pour les facteurs de transcription dits spécifiques qui se distinguent par leur rareté (Aravind *et al.*, 2003 ; Templeton *et al.*, 2004). Un paradoxe difficile à interpréter quand on considère le style de vie « complexe » de ces parasites, comparé à celui de la levure *S. cerevisiae* par exemple qui possède un arsenal de 200 facteurs de transcription spécifiques (Messina *et al.*, 2004).

Nous avons supposé au début de ce travail de thèse que la régulation génique des *Apicomplexa* et en particulier celle de *T. gondii* répondait à une mécanique moléculaire moins dépendante des facteurs de transcription spécifiques, voire de certains facteurs généraux ;

finalement une mécanique finement adaptée au cycle de vie complexe de ces parasites. Les transitions d'une forme à l'autre, véritables événements de différenciation, conditionnent la pathogénécité et la survie de ces parasites dans leurs hôtes. Ainsi chez *T. gondii*, la capacité d'une souche à opérer une conversion efficace tachyzoïte-bradyzoïte et échapper ainsi à la réponse immunitaire détermine potentiellement sa virulence chez l'Homme.

L'hypothèse clé formulée dans ce manuscrit est que la différenciation parasitaire chez *T. gondii* est sous le contrôle quasi-exclusif des facteurs de remodelage physique et chimique de la chromatine. L'étude de cette barrière nucléosomale et de sa dynamique au cours de l'interconversion nous a permis d'apporter les premières briques fondatrices du « code histone parasitaire » et son impact sur la différenciation de ce parasite.

### 2. Évolution du code histone chez *T. gondii*

Nous avons montré que l'acétylation et la méthylation, deux grandes marques épigénétiques, étaient conservées chez *T. gondii*. Nous observons par SchIP que ces modifications ont des effets variables en fonction de la nature du groupement, de l'emplacement de l'acide aminé modifié, et de son environnement proche (même queue d'histone) ou éloigné (queues d'histones différentes). Un gène réprimé porte alors les marques de la répression (déacétylation par exemple) et inversement un gène actif les marques de l'activation (acétylation). À l'échelle cellulaire, le niveau d'acétylation délimite bien topographiquement l'euchromatine active de celle inactive au moyen d'un gradient d'acétylation/désacétylation et ce, à proximité de gènes exprimés de manière différentielle dans les deux stades différenciés du Toxoplasme (Saksouk *et al.*, 2005, Figure 1).

Le plus étonnant dans nos résultats de SchIP est littéralement « l'explosion » de l'acétylation sur 15 Kbs autour du *locus* bradyzoïte *BSR4* et ce, après un événement de conversion *in vitro* tachyzoïte-bradyzoïte (Saksouk *et al.*, 2005, Figure 1D). De nombreux domaines chromatinien sont régulés par des LCRs (locus control regions) qui créent des *loci* autonomes avec une expression tissu-spécifique (Forsberg *et al.*, 2001). Le membre fondateur de cette classe d'éléments régulateurs est la LCR du gène de la  $\beta$ -globine. Des études ont montrées que ce LCR augmentait l'expression du gène de la  $\beta$ -globine chez l'Homme et la souris. La distribution des histones acétylés à ce *locus* déterminée par ChIP a démontré pour la première fois l'existence d'une mécanique d'acétylation de large région chromatinienne qui

telle une vague hyperacétyle les résidus de chaque mononucléosome, de proche en proche jusqu'à trouver une 'barrière' limitant la propagation à la zone adjacente hypoacétylé. Il semblerait donc que *T. gondii* ait gardé cette capacité à stimuler un tel phénomène.

La démonstration d'une activité R17(me) sur l'histone H3 chez *T. gondii* a aussi bouleversé notre vision évolutive. En effet, cette modification présente chez le Toxoplasme est absente chez la levure et *C.elegans*. Elle était jusqu'à maintenant considérée comme une marque épigénétique propre au répertoire des eucaryotes supérieures. Cette observation est conforme à l'idée que l'absence de facteurs de transcription spécifiques dans les apicomplexes peut être compensé à un niveau épigénétique. Nous démontrons également que le « dialogue » acétylation/méthylation intra-histone H3 des marques R17(me) et K18(ac) a aussi lieu chez le Toxoplasme (Saksouk *et al.*, 2005, Figure 6A). Cette combinaison a été décrite uniquement chez l'Homme où CBP/p300 recrute sa méthylase CARM1 via l'acétylation du résidu H3(K18) (Daujat *et al.*, 2002). L'absence d'homologue à CBP/P300 a été palliée chez *T. gondii* par la duplication du gène *GCN5* en deux orthologues, *TgGCN5a* et *TgGNC5b*. L'équipe du Dr. Sullivan a pu ainsi montrer que l'acétylase *TgGCN5a* acétyle exclusivement le résidu H3(K18) alors que son orthologue est spécifique des résidus multiples sur la queue de l'histone H3 (Sullivan *et al.*, 2000 ; Saksouk *et al.*, 2005; Bhatti *et al.*, Manuscrit). Ainsi les duplications des gènes PRMT et GCN5 représentent un exemple d'évolution convergente entre le parasite et sa cellule hôte avec un objectif commun affiché, l'adaptation à un environnement commun.

D'autres modifications confirment la complexité du « code histone parasitaire ». Par SchIP, notre équipe a pu montrer que le nombre de groupements méthyles sur le résidu H3 (K4) détermine l'état transcriptionnel chez le Toxoplasme. Il a été montré chez les autres eucaryotes que le nombre de groupements méthyles présents sur une même lysine (la mono-, di- et tri-méthylations) peut mener à différents types de réponses, augmentant ainsi la complexité du code histones. Une spécificité fonctionnelle de ces degrés de modification a été mis en évidence par la génération d'anticorps permettant la distinction entre les formes di- et tri-méthylées de la lysine : chez la levure, tandis que la tri-méthylation de la K4-H3 n'est présente que sur les gènes actifs, la di-méthylation (également présente exclusivement dans les régions euchromatiques) est présente indifféremment sur les gènes actifs et inactifs (Santos-Rosa *et al.*, 2002). Chez *T. gondii*, la marque H3 (K4)<sup>dimethyl</sup> est une marque de gènes



inactif et actif, alors que la triméthylation de ce résidu marque uniquement les gènes actifs, et ce en fonction du stade parasitaire étudié (résultats non montrés).

### **3. Lecture du code par TgCARM1**

L'arginine-méthyltransférase CARM1 est au centre de plusieurs régulations transcriptionnelles au sein de la cellule humaine (Figure 12). L'une des cibles de CARM1 sont les histones H3 (arginine 17) et H4, qui sont aussi cibles de l'activité acétylase des coactivateurs CBP/p300 : en se fixant aux régions des promoteurs via les coactivateurs, CARM1 augmente la méthylation des histones et rend ces régions plus accessibles pour la transcription. Elle est associée aux coactivateurs de transcription de la classe p160 impliqués dans l'activation des gènes (Stallcup *et al.*, 2000 ). CARM1 interagit aussi avec les coactivateurs de transcription CBP/p300. CARM1 méthyle aussi des substrats non-histone comme l'arginine 2142 (domaine de fixation de GRIP1) de p300. Une autre cible de la méthylation par CARM1 est un coactivator qui interagit avec CBP. La méthylation de CBP par CARM1 empêche CBP d'agir en tant que coactivator auprès de CREB (*Cyclic AMP Response Element Binding protein*) et rend ainsi disponible le pool faible de CBP dans la cellule pour les récepteurs nucléaires sensibles aux hormones stéroïdes (Chen *et al.*, 2000 ; Koh *et al.*, 2001). Notons ici que d'autres modifications post-traductionnelles interviennent dans ce schéma comme la phosphorylation.

Dans ce manuscrit nous avons étudié uniquement l'activité nucléosomale de l'homologue parasitaire TgCARM1, qui méthyle spécifiquement le résidu R17 de l'histone H3, à l'image de son homologue humain (Saksouk *et al.*, 2005, Figure 5C ; Bauer *et al.*, 2002). Nous avons montré que l'enzyme seule ne peut catalyser la réaction à ce résidu que sur un substrat « core histone ». La même activité sur un substrat « mono- et/ou poly-nucléosomes » n'a pu être obtenue qu'après incubation des nucléosomes avec un extrait enrichi en activité de remodelage ATP-dépendant (Saksouk *et al.*, 2005, Figure 5F). Chez l'Homme, CARM1 interagit avec la sous-unité BRG1 du complexe de remodelage ATP-dépendant SWI/SNF pour formé le complexe NUMAC (*NUclear Methylation Activator Complex*) (Xu *et al.*, 2004). Les auteurs ont ainsi montré que le complexe NUMAC était capable de méthyle les les nucléosome, signe que l'activité de remodelage ouvre l'accès de la queue d'histone.

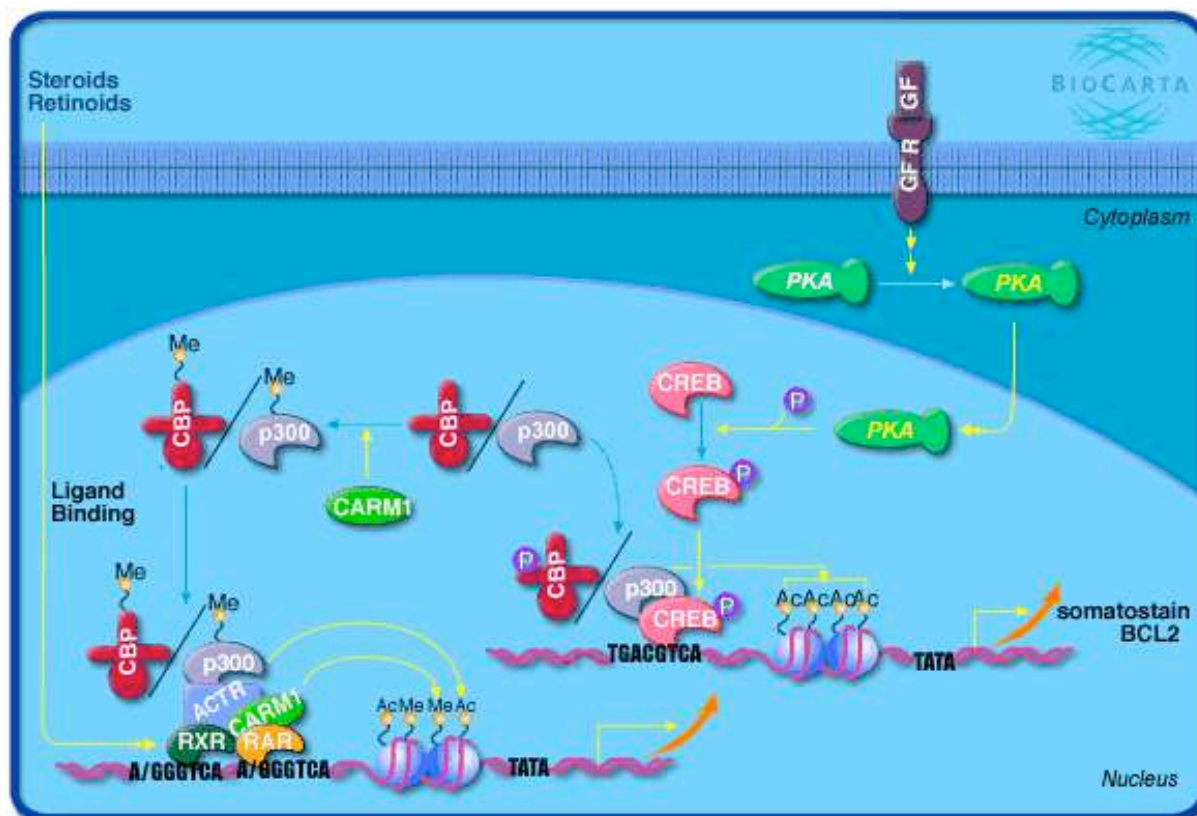


FIGURE 12 (Source figure : Biocarta)

La prochaine étape sera d'identifier plus spécifiquement chez le *Toxoplasme* la (ou les) enzyme(s) qui interagisse(nt) fonctionnellement avec TgCARM1 sur le nucléosome. Une autre question se posera alors : celle de l'identification potentielle des substrats non histone de cette enzyme chez le *Toxoplasme*.

#### 4. Briser le code & Interférer avec la différenciation parasitaire

Une délétion du gène CARM1 est létal au stade embryonique chez la souris. En 2004, des inhibiteurs spécifiques des arginine méthyltransférases ont été découverts (Cheng *et al.*, 2004). Nous avons saisi l'opportunité de les tester dans notre système d'étude avec l'objectif de réaliser un mutant chimique de TgCARM1. Nos résultats d'inhibition *in vitro* avec la molécule AMI-1 diffèrent légèrement de ceux obtenus par l'équipe du Dr. Bedford publié dans Cheng, D *et al.* (2004). En effet, nous observons une spécificité accrue de l'inhibiteur pour TgCARM1 comparée à TgPRMT1, et ce pour de faible quantité ajoutée (Saksouk *et al.*, 2005, Figure 7A).

La molécule AMI-1 a été décrite comme perméable aux cellules humaines sans effet cytotoxique apparent (Cheng *et al.*, 2004). Ce fut donc pour nous l'occasion de la tester *in vivo* sur des cellules humaines infectées de toxoplasme. Les premières analyses n'ont montré aucun changement significatif sur l'infection ou la réplication des parasites traités intracellulairement. Nous avons alors cherché à déterminer l'effet de la drogue sur la conversion. Nous avons observé une induction de la conversion par AMI-1 chez la souche de type II prugniaud. La conversion fut faible lorsque le traitement était opéré avec des tachyzoïtes intracellulaires. Par contre, un traitement en extracellulaire suivi d'une infection, provoquait chez environ 40% de la population de tachyzoïte un arrêt de division et une conversion progressive vers la forme bradyzoïte (Saksouk *et al.*, 2005, Figure 7). Le plus surprenant est que cette conversion s'accompagne d'une diminution globale (détectée par immunofluorescence, Figure 7C) et spécifique (détectée par SchIP, Figure 7E) de la méthylation au résidu H3 (R17). L'hypothèse que la marque R17<sup>dimethyl</sup> soit une modification exclusive du tachyzoïte a été vite écartée car une induction au pH de la conversion ne diminuait pas le taux de méthylation globale du toxoplasme (détectée par immunofluorescence, Figure 7D). La co-localisation au locus bradyzoïte *SAG2C* de l'enzyme TgCARM1 et de sa modification R17<sup>dimethyl</sup> après un événement de conversion tachyzoïte-bradyzoïte (Figure 6A) est venue confirmer l'implication de cette enzyme dans la régulation épigénétique du processus d'interconversion.

Il est maintenant clairement établi que la dérégulation ou le dysfonctionnement des phénomènes épigénétiques sur des histones (ou d'autres substrats) et la modification de la structure chromatinienne sont des étapes critiques pouvant participer aux processus de diverses maladies comme les cancers. La modulation de l'activité des différentes enzymes (HMT, HAT et HDAC) contrôlant le degré de modifications apparaît donc comme un nouveau moyen d'action contre l'invasion parasitaire et/ou contre les phénomènes d'interconversion chez *T. gondii*.

### 5. La déacétylation au cœur de la répression génique

Nous avons identifié le premier complexe de déacétylation chez les Apicomplexes (Saksouk *et al.*, 2005, Figure 2 et 3). TgCRC (pour *T. gondii* Co-Repressor Complex) est un complexe qui partage plusieurs sous-unités communes avec les complexes humains N-CoR et SMRT comme le co-repressor TgTBL1. Ce complexe a une activité de déacétylation

spécifique et son enzyme centrale TgHDAC3 co-localise à des *loci* hypo-acétylés (Saksouk *et al.*, 2005, Figure 3 et 4).

L'identification de plusieurs autres déacétylases chez le toxoplasme est un signe important de la diversification fonctionnelle et de l'importance de ces enzymes dans l'écriture du code. L'étude de TgHDAC2 résume bien la complexité introduite par le parasite pour réguler cette fonction, probablement essentielle. TgHDAC2 se distingue en effet de toutes ses homologues de classe I, à ce jour publiés, par une insertion peptidique de 200 acide aminés au sein d'une région potentiellement importante pour son activité catalytique.

D'autres particularismes ont pu être observés dans cette famille enzymatique parasitaire. Par exemple, le gène *TgHDAC4* code pour deux ARNm alternativement épissés v1 et v2. Les deux variants présentent en plus du domaine catalytique HDAC un motif NES (*Nuclear Export Signal*) qui permet d'exporter les protéines TgHDAC4v1 et TgHDAC4v2 vers le cytoplasme. Cependant, aucune séquence d'adressage nucléaire NLS (*Nuclear Localization Signal*) n'a pu être déterminée par PSORT (<http://psort.nibb.ac.jp>). L'expression ectopique dans des tachyzoïtes du variant TgHDAC4v1 confirme l'exclusion nucléaire alors que l'expression de TgHDAC4v2 montre une localisation cytoplasmique et nucléaire (résultats non montrés). Le variant 2 diffère du variant 1 par la présence unique en C-terminal de la protéine d'une insersion (due à un épissage alternatif) de 15 acide aminés (PMTSLTLYGGTLFFF). Son implication dans la localisation nucléaire du variant 2 reste à démontrer.

Ces éléments suggèrent que la complexité du code épigénétique chez *T. gondii* s'opère au niveau i) de la duplication évolutive créant ainsi des familles géniques pour une fonction donnée (par exemple 5 gènes TgHDAC), mais aussi ii) des modifications des domaines structuraux des enzymes (exemple de TgHDAC2) qui créent encore plus de diversité fonctionnelle. Une couche de complexité supplémentaire est ajoutée lorsque l'on découvre que ces enzymes sont associés à des complexes multi-protéiques géants, comme TgCRC.

Souvent les enzymes HDAC sont inactives sous leur forme recombinante exprimée dans les bactéries. Il a été montré chez l'Homme que HDAC3 néo-synthétisée est prise en charge par la protéine HSP90 afin d'être présentée au complexe de protéines chaperonines Hsc70-TRiC. Ce complexe utilise l'énergie libérée par la dégradation de l'ATP pour procéder au

repliement de l'HDAC3 (Melki *et al.*, 1997). La protéine SMRT, en se liant à l'HDAC3, conduit alors à la dissociation du complexe TRiC. Le complexe HDAC3/SMRT possède alors une activité HDAC stable (Guenther *et al.*, 2002). Chez *T. gondii*, nous retrouvons les chaperones au sein du complexe TgCRC et nous pensons que TgCRC-230 et TgCRC-350, deux protéines nucléaires, spécifiques au phylum des Apicomplexes, jouent un rôle similaire à celui de SMRT chez l'Homme.

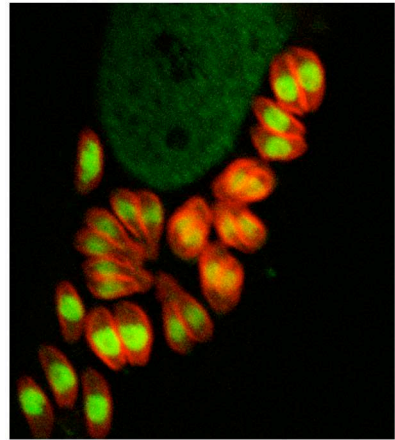
Le co-represser TgTBL1 qui contient un domaine LISH et six domaines WD40, joue probablement lui aussi un rôle très important au sein du complexe TgCRC dans la régulation transcriptionnelle. Ainsi, chez l'Homme, les protéines TBL1 et TBLR1 vont interagir d'une part par leur domaine LISH avec les protéines SMRT/N-CoR et d'autre part par leur domaine WD40 avec les histones H2B et H4 hypoacétylées, pour stabiliser le complexe co-répresseur HDAC3 au nucléosome (Yoon *et al.*, 2003). Il a récemment aussi été démontré que le domaine WD40 serait un module de liaison aux lysines méthylées (Wysocka *et al.*, 2005). Retour donc ici au dialogue entre la méthylation et l'acétylation/déacétylation, pierre angulaire du code histone.

### 6. En guise de conclusion...

Nous avons émis l'hypothèse en introduction de ce manuscrit que la régulation transcriptionnelle de la différenciation chez *T. gondii* est moins dépendante de la diversification des facteurs spécifiques de transcription, observées chez les eucaryotes « primitifs » (levure) et supérieurs. En revanche, nous avons montré que *T. gondii*, et par extension son phylum, avait développé une machinerie épigénétique sophistiquée, plus proche de celle de l'Homme que celle de la levure ; l'exemple le plus frappant étant la co-évolution d'un homologue à CARM1. Le cycle biologique de ces parasites intracellulaires apporte une explication sérieuse à cette évolution particulière des systèmes de régulation génique. L'explication la plus simple est que face à un environnement protégé, exemple de la vacuole parasitophore pour le Toxoplasme, la réponse du parasite aux *stimuli* bénéficie d'un temps plus long que celui d'une levure dans un milieu naturel constamment bouleversé. Ce délai dans la réponse a permis au parasite de réduire, voire de ne pas augmenter, son arsenal en facteur de transcription spécifiques et généraux. Cela dit, l'interaction « cellule hôte-parasite » implique aussi une régulation génique concertée, voire quasi-symbiotique. Nous pensons que la sophistication du code épigénétique, résultat d'une co-évolution des systèmes d'écriture et de

lecture entre la cellule hôte et son parasite prend en charge ce « dialogue cellulaire ». Si notre hypothèse de départ a maintenant des bases solides, il faudra décrypter ce code histone parasite. Puis dans un deuxième temps il faudra découvrir les épimutations qui peuvent déséquilibrer le système d'écriture et/ou de lecture pour aboutir au développement de candidat-médicaments à même de briser le cycle biologique parasite.

# MATERIELS ET METHODES



## 1. MATERIEL

### 1.1. Vecteurs de clonage :

- Clonage de produits PCR : TOPO TA cloning (*Invitrogen*).
- Production de protéines recombinantes chez *E. coli* : pET28a+ et pET41a+ (*Novagen*).
- Expression ectopique de protéines chez *T. gondii* : Le vecteur pMAH14 a été utilisé ; il est issu d'une matrice pKS dans laquelle ont été ajouté un promoteur fort GRA1, un terminateur SAG1 et la double étiquette HA-Flag permettant de réaliser des fusions traductionnelles.

### 1.2. Enzymes utilisées :

- Reverse transcription : reverse transcriptase Superscript III (*Invitrogen*).
- Amplification par PCR : TAQ DNA polymerase Dynazyme EXT (*Finnzyme*).
- Ligation de fragments d'ADN dans un vecteur d'intérêt : T4 DNA ligase (*Biolabs*).
- Digestion de fragments d'ADN : Enzymes de restriction (*Fermentas & Biolabs*).

### 1.3. Préparation d'ADN plasmidique et d'ARN totaux de tachyzoïtes :

Les mini-préparations d'ADN plasmidique sont effectuées selon Sambrook and Russell. Les plasmides d'intérêt sont amplifiés par Maxiprep avec le kit Nucleobond (*Machery-Nagel*). Le séquençage des molécules d'ADN est réalisé par *Genome express*. Les ARN totaux de tachyzoïtes sont extraits avec le kit Quiagen RNeasy Minikit.

### 1.4. Les amorces utilisés :

Les amorces utilisées pour les constructions des plasmides recombinants sont répertoriées dans le tableau 1.

### 1.5. Les plasmides recombinants :

les vecteurs et insert de clonages réalisés durant la thèse sont répertoriés dans le tableau 2.



Vecteur	Nom des amorces	Séquences des amorces
<b>pNS2</b>	6646-PCR5-HdIII :	F 5' ACCCAAGCTTGACGCGTCCGTCGTCGAAACG 3'
	6646-PCR3-XhoI :	R 5' ACCGCTCGAGTTACCTCCCT GAGAACACTCG 3'
<b>pNS4 et pNS7</b>	CARMI-PCR5-BamHI :	F 5' GGATCCGACGCGTCCGTCGTCGAAACGTAC 3'
	6646-PCR3-XhoI :	R 5' ACCGCTCGAG TTACCTCCCT GAGAACACTCG 3'
<b>pMAH2</b>	TgPRMT1-PCR5-SSS-HindIII	F 5' CCCAAGCTTAGTTCCAGCAGCCCCAACTC 3'
	TgPRMT1-PCR3-RLR-XhoI	R 5' CCGCTCGAGCTATCGGAGTCTGTAGAACTGG 3'
<b>pNS14 à pNS18</b>	TgTBL1-PCR5-BglII	F 5' GAGATCTCGGATCCACTTGTCTTCGGACGAAGTG 3'
	TgTBL1-PCR3-HindIII	F 5' CAAGCTTCGTTACGCAGAGTCTTCCAGCTTCGC 3'
<b>pNS9 à pNS12</b>	TgGCN5-PCR5-BamHI	F 5' GGATCCGGCGCTCCAACAGGTCTGGGGCTCG 3'
	TgGCN5-PCR3-HindIII	F 5' AAGCTTCGTTACGAAACTCCCAGAGCCTCG 3'
<b>pNS25, 29, 32 et 33</b>	PCR5-CARM1(VMD-AAA)	F 5' GAAGGAAAAACTGCCGCCCGTCGGAGCCGG 3'
	PCR3-CARM1(VMD-AAA)	F 5' CCGGCTCCGACGGCGGCGGAGTTTTCCTTC 3'
<b>pMAH21</b>	TgHDAC3-PCR5-BamHI	F 5' CGGGATCCGCGCTCAGTGCCTGCGCAAGC 3'
	TgHDAC3-PCR3-PstI	F 5' CAGAGACCAAAGGTTCCGATCCTGCAG 3'
<b>pNS19, 26 et 29</b>	TgHDAC3-PCR5-BamHI	F 5' GGATCCGCGAAGAAGCTGAATCACCAC 3'
	TgHDAC3-PCR3-HidIII	F 5' AAGCTTGATCGGAACCTTTTGGTCTCTG 3'
<b>pNS14 à 18</b>	TgTBL1-PCR5-BglII	F 5' GAGATCTCGGATCCACTTGTCTTCGGACGAAGTG 3'
	TgTBL1-PCR3-HindIII	F 5' CAAGCTTCGTCCACGCAGAGTCTTCCAGCTTCGC 3'
<b>pNS43 à 46</b>	TgTBL1-PCR5-BglII	F 5' GAGATCTCGGATCCACTTGTCTTCGGACGAAGTG 3'
	TgLISH-PCR-HindIII	F 5' AAGCTTCAACGCGACTTATTCTTGGCATG 3'
<b>pNS34, 35, 47, 55 et 104</b>	TgWD40-PCR-BglII	F 5' AGATCTTCTCCTTCTGTGGAGATCGTTGGC 3'
	TgTBL1-PCR3-HindIII	F 5' CAAGCTTCGTCCACGCAGAGTCTTCCAGCTTCGC 3'
<b>pNS36, 38 et 54</b>	TgKELCH-PCR5-BamHI	F 5' GGATCCGATGAGGTCGTGAAGGCGTTGCGG 3'
	TgKELCH-PCR3-XbaI	F 5' TCTAGAGAAATGCCACATGTCATCGAGGAG 3'
<b>pNS13</b>	TgHDAC2-PCR5-BamHI	F 5' CGGGATCCCTGGGGAGGAGGAAGAGCGTTC 3'
	TgHDAC2-PCR3-EcoRI	F 5' CGGAATTCACCACGATTCTGTTCGAGATC 3'
<b>pNS88 et 90</b>	HD2-PCR5-BamHI	F 5' GGATCCACTAGTAACGGCCGCCAG 3'
	HD2-fusion-PCR5	F 5' CAATCGCGTCTCGAACTGGCGGATCAAATG 3'
	HD2-Fusion6PCR3	F 5' CATTGATCCGCCAGTTCGAGACGCGATTG 3'

## Matériel et méthode

	HD2-PCR3-HindIII	F 5' GATA <u>AAGCTT</u> CTATCAAGATCTCTTGTCG 3'
<b>pNS39 à 42, 49 à 52, 102 et 103</b>	TgHDAC4-PCR5-BamHI	F 5' <u>GGATCC</u> GACAGCACACGAGCGACTTTGGGC 3'
	TgHDAC4-PCR3-EcoRI	F 5' GTGTGCTG <u>GAAATC</u> CGCCCTTGAC 3'
<b>pNS70, 71, 74 et 75</b>	TgHDAC5-ATN-BglII	F 5' GA <u>AGATCT</u> GCCACCAACAACATTCGGGGCGTC 3'
	TgHDAC5-GTD-PCR3-EcoRI	F 5' ACGGAATTCATCGGTTCCGCGGCACAAAGCTCG 3'
	TgHDAC5-THV-BglII	F 5' <u>AGATCT</u> ACTCATGTCCGAAATATCGTGAATG 3'
<b>pNS31</b>	TgSIR2-PCR5-XmaI	F 5' <u>CCCGGG</u> CACTCAGTTTATGGGAGTGAAGTAG 3'
	TgSIR2-PCR3-EcoRI	F 5' <u>GAAATC</u> GTTGCCTTGCCTTACCTCAACAAC 3'

**Tableau 1 :** Liste des amorces utilisées pour les clonages (cf. : tableau 2). Les sites de restrictions sont soulignés sur les séquences des amorces.

Plasmide	Vecteur	Insert	Taille en pb	Poids moléculaire	Enzymes de restriction pour le clonage
<b>pNS2</b>	pET28b+	TgCARM1	1491 pb	54.61 KDa	HindIII + XhoI
<b>pNS4</b>	pMAH12	TgCARM1	1509 pb	55.57 KDa	BglII/BamHI +XhoI
<b>pNS7</b>	pET41a+	TgCARM1	2220 pb	82.45 KDa	BglII/BamHI +XhoI
<b>pNS25</b>	pTOPO	TgCARM1 (VMD_AAA)	1384 pb		BamHI & XhoI
<b>pNS29</b>	pMAH12	TgCARM1 (VMD_AAA)	1509 pb	55,4 KDa	BglII/BamHI+XhoI
<b>pNS32</b>	pET28a+	TgCARM1 (VMD_AAA)	1467 pb	53,6 KDa	BamHI+XhoI
<b>pNS33</b>	pET41a+	TgCARM1 (VMD_AAA)	2220 pb	82,3 KDa	BamHI+XhoI
<b>pMAH2</b>	pET28b+	TgPRMT1	1182 pb	44 KDa	HindIII + XhoI
<b>pNS14</b>	pTOPO	TgTBL1	2399 pb		BglII & HindIII
<b>pNS15</b>	pMAH14	TgTBL1	2472 pb	88,163 KDa	BglII + HindIII
<b>pNS16</b>	pET28a+	TgTBL1	2481 pb	88,111 KDa	BglII + HindIII
<b>pNS17</b>	pET41a+	TgTBL1	3234 pb	117 KDa	BglII + HindIII
<b>pNS18</b>	pET28a+	TgTBL1	1749 pb	61,6 KDa	pNS16+Not1
<b>pNS35</b>	pTOPO	TgWD40 de TBLI	1414 pb		BglII & HindIII

## Matériel et méthode

pNS34	pMAH14	TgWD40 de TBLI	1488 pb	53,37 Kda	BgIII + HindIII
pNS47	pET28a+	TgWD40 de TBLI	1494 pb	53,32 KDa	BamHI&BgIII+HindIII
pNS55	pET41a+	TgWD40 de TBLI	2250 pb	82 KDa	BamHI&BgIII+HindIII
pNS104	pNS85	TgWD40 de TBLI	1494 pb	53,56 KDa	BgIII + HindIII
pNS43	pTOPO	TgLISH de TBLI	1414 pb		BgIII & HindIII
pNS44	pMAH14	TgLISH de TBLI	942 pb	33,54 KDa	BgIII + HindIII
pNS45	pET28a+	TgLISH de TBLI	951 pb	33,48 KDa	BamHI&BgIII+HindIII
pNS46	pET41a+	TgLISH de TBLI	1704 pb	62,23 KDa	BamHI&BgIII+HindIII
pNS36	pTOPO	TgKELCH	822 pb		BamHI & Xbal
pNS38	pMAH14	TgKELCH	933 pb	34,88 KDa	BamHI+ClaI
pNS54	pET28a+	TgKELCH	753 pb	27,7 KDa	BamHI+XhoI
pNS19	pTOPO	TgHDAC1trq	1002 pb		BamHI & HindIII
pNS26	pET28a+	TgHDAC1trq	1131 pb	41,7 KDa	BamHI + HindIII
pNS27	pET41a+	TgHDAC1trq	1890 pb	70,7 KDa	BamHI + HindIII
pMAH21	pMAH14	TgHDAC1	1425 pb	53,37 KDa	BamHI + PstI
pNS30	pMAH14	TgHDAC2trq	1176 pb	41,78 KDa	BamHI+EcoRI
pNS13	pMAH14	TgHDAC2	1989 pb	71,95 KDa	BamHI + Pst1 + EcoRI
pNS88	pTOPO	TgHDAC2trq(2)	1435 pb		BamHI & EcoRI
pNS90	pMAH14	TgHDAC2trq(2)	1542 pb	55 KDa	BamHI&BgIII+HindIII
pNS31	pMAH14	TgSIR2	876pb	32,3 KDa	XmaI+ EcoRI
pNS39	pTOPO	TgHDAC4 V#2	957 pb		BamHI & EcoRI
pNS41	pMAH14	TgHDAC4 V#2	1026 pb	37,65 KDa	BamHI+ EcoRI
pNS49	pET28a+	TgHDAC4 V#2	1119 pb	40,67 KDa	BamHI+ EcoRI
pNS51	pET41a+	TgHDAC4 V#2	1899 pb	70 KDa	BamHI+ EcoRI
pNS103	pNS85	TgHDAC4 V#2	1020 pb	37,44 KDa	BamHI+ EcoRI
pNS40	pTOPO	TgHDAC4 V#1	912 pb		BamHI & EcoRI
pNS42	pMAH14	TgHDAC4 V#1	981 pb	36 KDa	BamHI+ EcoRI
pNS50	pET28a+	TgHDAC4 V#1	1074 pb	39 KDa	BamHI+ EcoRI
pNS52	pET41a+	TgHDAC4 V#1	1854 pb	68,55 KDa	BamHI+ EcoRI

## Matériel et méthode

pNS102	pNS85	TgHDAC4 V#1	975 pb	35,8 KDa	BamHI+ EcoRI
pNS70	pTOPO	TgHDAC5 (PRC55)	3814 pb		BgIII & HindIII
pNS74	pMAH14	TgHDAC5 (PRC55)	3882 pb	137 KDa	BgIII/BamHI+HindIII
pNS71	pTOPO	TgHDAC5 (PRC60)	3814 pb		BgIII & HindIII
pNS75	pMAH14	TgHDAC5 (PRC60)	3882 pb	137 KDa	BgIII/BamHI+HindIII
pNS9	pTOPO	TgGCN5	1431 pb		amHI & HindIII
pNS10	pMAH14	TgGCN5	1416 pb	52,5 KDa	BgIII/BamHI +HindIII
pNS11	pET28a+	TgGCN5	1515 pb	55,94 KDa	BamHI +HindIII
pNS12	pET41a+	Tg:GCN5	2268 pb	84,4 KDa	BamHI +HindIII

**Tableau 2 :** Les clonages réalisés ainsi que la taille et le poids moléculaire des inserts avec leur(s) étiquette(s) sont inscrits dans ce tableau.

### 1.6. Souches bactériennes :

- TOP10F' : F' {*lacI<sup>q</sup>*, Tn10(Tet<sup>R</sup>)} *mcrA* Δ(*mrr-hsdRMS-mcrBC*) ϕ80*lacZ*ΔM15 Δ*lacX74 recA1 araD139* Δ(*ara-leu*)7697 *galU galK rpsL* (Str<sup>R</sup>) *endA1 nupG*
- BL21 (DE3) : F' *ompT hsdS<sub>B</sub>(r<sub>B</sub><sup>-</sup>m<sub>B</sub><sup>-</sup>) gal dcm* (DE3)

Les souches bactériennes sont rendues compétentes et transformées par choc thermique selon Sambrook and Russell.

### 1.7. Les anticorps :

<u>Immunofluorescence :</u>	<u>Western blot :</u>	<u>Immunoprécipitation:</u>
<ul style="list-style-type: none"> <li>➤ _-HA monoclonal (<i>Roche</i>)</li> <li>➤ _-49(TgCARM1) polyclonal de lapin (<i>Eurogentec</i>)</li> <li>➤ _-TgHDAC1-polyclonal de lapin (<i>Eurogentec</i>)</li> </ul>	<ul style="list-style-type: none"> <li>➤ _-49(TgCARM1) polyclonal de lapin (<i>Eurogentec</i>)</li> <li>➤ _-TgHDAC1-polyclonal de lapin (<i>Eurogentec</i>)</li> <li>➤ _-Flag monoclonal (<i>SIGMA</i>)</li> </ul>	<ul style="list-style-type: none"> <li>➤ _-Flag monoclonal (<i>SIGMA</i>)</li> <li>➤ _-H3 (K4)<sup>diMe</sup>, _-AcH4 (K5, K8, K12, K16)<sup>Acetyl</sup>, _-H3 (R17)<sup>Me</sup>, _-H4</li> </ul>

## Matériel et méthode

<ul style="list-style-type: none"> <li>➤ <math>\gamma</math>-TgGCN5-A polyclonal de lapin (Bhatti &amp; sullivan, 2005)</li> <li>➤ <math>\gamma</math>-H3 (K4)<sup>diMe</sup>, <math>\gamma</math>-AcH4 (K5, K8, K12, K16)<sup>Acetyl</sup>, <math>\gamma</math>-H3 (R17)<sup>Me</sup>, <math>\gamma</math>-H4 (R3)<sup>Me</sup> (Upstate)</li> </ul>	<ul style="list-style-type: none"> <li>➤ <math>\gamma</math>-HA monoclonal (Roche)</li> <li>➤ <math>\gamma</math>-TgGCN5-A polyclonal de lapin (Bhatti &amp; sullivan, 2005)</li> <li>➤ <math>\gamma</math>-H3 (K4)<sup>diMe</sup>, <math>\gamma</math>-AcH4 (K5, K8, K12, K16)<sup>Acetyl</sup>, <math>\gamma</math>-H3 (R17)<sup>Me</sup>, <math>\gamma</math>-H4 (R3)<sup>Me</sup> (Upstate)</li> </ul>	$\gamma$ -H3 (R17) <sup>Me</sup> , $\gamma$ -H4 (R3) <sup>Me</sup> (Upstate)
---	--	--

## 2. METHODES DE BIOLOGIE MOLECULAIRE

### 2.1. Conditions de culture et de stockage des bactéries :

Les bactéries sont cultivées selon la méthodes de Sambrook *et al.*, 1989 et son stockées dans 20 % de glycérol à -80 °C.

### 2.2. Préparation de E.Coli compétentes :

La méthode utilisée est la méthode classique au chlorure de calcium (cohen, 1972 et Erlich, 1979). Les bactéries de la souche hôte sont mises en culture en milieu LB et sont collectées en phase exponentielle de croissance (absorbance de 0.3 à 600 nm), car dans les cellules en division, les membranes en cours d'allongement sont plus fragiles. Le traitement consiste placer les cellules dans une solution concentrée de Chlorure de Calcium à 4 °C : les ions Ca<sup>2+</sup> altèrent les structures membranaires en créant des microperforations dans la double couche phospholipidique. Pour 40ml de culture, les cellules sont remises en suspension dans 20 ml de CaCl<sub>2</sub> 100 mM glacé, centrifugées à 3000 rpm, à 4 °C pendant 10 minutes et remises en suspension dans 1 ml de CaCl<sub>2</sub> 100 mM.

Pour congeler les cellules compétentes, la dernière remise en suspension se fait en ajoutant 15 % de glycérol au CaCl<sub>2</sub>. Lors de la transformation faut incubé les cellules avec l'ADN à 4°C pendant au moins 30 minutes.

### 2.3. PCR (Polymerase Chain Reaction) :

Les vecteurs plasmidiques utilisés comme matrices pour l'amplification ont été préparés et purifiés par la méthode « Qiagen » (Qiagen<sup>TM</sup> Plasmid Midi Kit). La séquence des amorces oligonucléotidiques (tableau 1) a été choisie de manière à amplifier la région d'intérêt ; des bases additionnelles ont été introduites artificiellement aux extrémités des amorces de façon à créer des sites enzymatiques cohésifs aux extrémités des fragments amplifiés et permettre ainsi un clonage ultérieur. Ces sites enzymatiques sont explicités dans le tableau 2. Quelques bases flottantes ont également été ajoutées derrière chaque site enzymatique de façon à permettre la liaison de l'enzyme de restriction choisie et la libération des extrémités cohésives à partir des fragments amplifiés. Les PCR ont été réalisées à partir d'un thermocycleur T3 (Biometra) à partir de 50 ng de matrice plasmidique, dans un volume de 50 µl, en présence de 0.25 mM d'amorces oligonucléotidiques, d'un mélange de désoxyribonucléotides (dATP, dTTP, dCTP et dGTP, 200 µM, Biolabs), de 10 µl tampon de la TAQ polymérase, de 2 unités de TAQ DNA polymerase Dynazyme EXT (*Finnzyme*). 2 µl de DMSO sont ajoutés si le pourcentage en base GC des amorces excède 50%. Après une dénaturation de 5 minutes à 94 °C, 35 cycles d'amplification ont été réalisés ; ces cycles comportent successivement une étape de dénaturation (30 seconde à 94 °C), une étape d'appariement (30 seconde à 55 °C), puis une étape d'élongation (30 seconde à deux minutes à 72 °C), avant une élongation finale de dix minutes à 72 °C.

La taille des fragments amplifiés a été vérifiée par électrophorèse d'une fraction (tampon de charge : 25 % de bleu de bromophénol et/ ou de xylène cyanol ; 5 mM Tris pH7,6 ; 6% glycérol) chargée sur gel d'agarose 1 % en tampon TBE (tris-HCl 89 mM, pH 7.9, H<sub>3</sub>BO<sub>3</sub> 89 mM, EDTA 0.2 mM). Les produits d'amplification ont été sous-clonés dans le vecteur pTOPO fourni dans le kit TOPO TA cloning (*Invitrogen*).

### 2.4. Clonage dans le vecteur TOPO :

Les produits/ amplicons PCR ont été clonés (ou sous clonés) dans le vecteur pTOPO du Kit TA cloning (*Invitrogen*).

### 2.5. Méthode blanc-bleu :

Les clones bactériens issus d'un sous-clonage dans le vecteur pTOPO et apparus après une nuit de culture sur LB-ampicilline (100 µg/ml), X-gal (5%) et IPTG (4 mM) ont été sélectionnés par la méthode blanc-bleu (Maniatis et al., 1990). Les bactéries contenant les plasmides recombinants, c'est à dire ayant intégré le produit d'amplification, sont blanches alors que les bactéries contenant les plasmides sauvages sont bleues. En effet, le site de clonage du vecteur se trouve dans le gène *LacZ* qui code pour l'enzyme β-Gal (enzyme qui génère un produit bleu à partir des substrats, X-gal et IPTG). L'insertion d'un fragment dans le site perturbe donc l'expression du gène *LacZ* et par conséquent, l'expression de la β-Gal : les colonies restent blanches.

### 2.6. Mini préparation de plasmide par lyse alcaline de bactéries :

Ce type de préparation d'ADN permet de purifier environ 5 µg d'ADN plasmidique à partir d'une culture de la souche désirée de 2 mL (250 rpm, 37°C, toute la nuit). Le matin, centrifuger la culture de 2 mL (Vmax, 5 min) et jeter le surnageant. Resuspendre soigneusement le culot dans 400 µL de tampon de resuspension (Solution 1 : glucose 50 mM, Tris-HCl pH8 25 mM, EDTA 10 mM, RNaseA 10 µg/µL), pour éviter les grumeaux qui empêcheraient une lyse homogène. La présence de RNase est essentielle pour éviter la contamination de l'ADN plasmidique par de l'ARN. Après resuspension, ajouter 400 µL de tampon de lyse (Solution 2 : NaOH 0.2 N, SDS 1%). Mélanger en inversant doucement le tube 4-5 fois jusqu'à ce que la solution prenne un aspect visqueux, signe que l'ADN est libéré. Ne jamais laisser agir plus de 5 min. Pour arrêter la lyse, ajouter 400 µL de tampon de neutralisation (Solution 3 : Acétate de potassium KOAc 3 M, acide acétique glacial 11% volume/volume (v/v)). Mélanger doucement en inversant le tube. La solution prend un aspect trouble, ce sont les protéines qui précipitent. Pour s'en débarrasser, centrifuger (12,000 rpm, 15 min, 4°C). Transférer le surnageant dans un tube en évitant le culot blanchâtre de protéines. Rincer avec de l'éthanol 100 % puis 70 %. Centrifuger pour éliminer toute trace résiduelle d'éthanol (Vmax, 30 s.). Solubiliser l'ADN dans 50 µL d'H<sub>2</sub>O. Quantifier en mesurant la DO.

### 2.7. Midi préparation de plasmides :

Une midi-culture de bactéries transformées (100 mL LB sélectif) estensemencée au 1/100 à partir d'une culture de 1,5 mL sélective, et placée à 37°C sous agitation pendant 12 à 16 h. La purification de plasmides est réalisée par la technique de lyse alcaline des bactéries en suivant le protocole fourni par le kit « NucleoBond PC100» (Machery Nagel). La concentration en ADN est estimée en mesurant la DO à 260 nm (une unité de DO correspond à 50 µg/mL d'ADN double brin).

### 2.8. Séquençage de l'ADN :

Les séquences de tous les inserts ont été vérifiées par séquençage chez GenomExpress (Meylan/France).

### 2.9. Digestion totale de l'ADN plasmidique :

Le vecteur et les fragments d'ADN (produits PCR ou plasmide recombinant) sont digérés par les enzymes adéquates (1 unité d'enzyme permet de digérer 1 µg d'ADN en 1 heure de réaction). La réaction est catalysée à 37°C pendant quelques heures. Les enzymes sont ensuite inactivées par un traitement de 15 min à 65°C, et l'ADN purifié par un traitement au phénol/chloroforme. Les acides nucléiques sont alors précipités par l'addition d'1/10 de volume d'acétate de sodium (3M ; pH5,2) et de 2 volumes d'éthanol absolu froid.

### 2.10. Séparation des fragments d'ADN sur gel :

Les acides nucléiques peuvent être séparés en fonction de leur taille par électrophorèse sur gel d'agarose. Le gel est réalisé en solubilisant 1-2% (m/v) d'agarose dans du tampon TBE 1X (Tris-HCl 100 mM, H<sub>3</sub>BO<sub>3</sub> 100 mM, EDTA 2.5 mM, pH 8). *NB: Pour des fragments de petite taille (≤ 1Kb), un agarose de type Low Melting (NuSieve GTG Low Melting Agarose, Mandel) à 1% est utilisé puisqu'il permet une meilleure résolution de séparation dans les faibles poids moléculaires.* L'ADN à analyser est additionné de tampon de charge 4X (glycérol 50%, bleu de bromophénol 0.25%, TBE 1X) et déposé dans un puit du gel. La migration s'effectue à 120 V dans un appareil à électrophorèse. *NB : Pour les petits fragments (≤ 1Kb), il est conseillé de ne pas prolonger la migration inutilement. En effet, les faibles poids moléculaires ont tendance à migrer au même niveau que le bleu du tampon de charge, si bien que le bruit de fonds du bleu sous UV masque la présence des petits*



## Matériel et méthode

*fragments. Pour limiter le phénomène, le tampon de charge 4X peut être utilisé à des concentrations plus faibles 10X. L'ADN est visualisé à 320 nm, grâce au bromure d'éthidium (BEt, 5 µg/mL), agent intercalant des acides nucléiques.*

### **2.11. Purification d'un fragment d'ADN par élution du gel d'agarose:**

Après migration des acides nucléiques, la bande du gel contenant le fragment d'intérêt est découpée sous UV puis purifiée à l'aide du « Nucleospin Extract II» (Machery Nagel) selon les instructions du fournisseur.

### **2.12. Ligation :**

La ligation se déroule dans un très petit volume (20 µL) et à basse température de façon à faciliter le rapprochement des extrémités cohésives de l'insert et du vecteur et l'action de la T4-ADN ligase. 50-400 ng de vecteur digéré purifié sont mis en présence d'un excès d'insert digéré purifié (0.1-1 µg), dans du tampon de ligation (Tris-HCl 300 mM pH7.8, MgCl<sub>2</sub> 100 mM, DTT 100 mM, ATP 10mM), 1 µL de T4-DNA ligase (Promega), qsp 20 µL. La solution est incubée à 16°C pendant une nuit. *NB : Dans le cas d'un clonage de petits inserts (≤1Kb), il peut être nécessaire d'augmenter le ratio insert/vecteur pour augmenter les chances de succès du clonage. La température de ligation peut également être diminuée pour favoriser le rapprochement de l'insert et du vecteur pour la ligation.*

### **2.13. Transformation de la ligation :**

Les produits de ligation ont ensuite été introduits dans des bactéries *E. Coli* préalablement rendues compétentes pour la transformation par la méthode Hanahan. Après une incubation de 30 minutes sur glace des bactéries avec le produit de ligation suivie d'un choc thermique de 42 secondes à 43°C, les bactéries ont été mises en culture une heure à 37°C en milieu LB, afin de permettre l'expression de la résistance à l'antibiotique portée par les plasmides recircularisés par la ligation. Les bactéries sont étalées sur boîte LB/antibiotique (100µg/ml) et incubées une nuit à 37°C.

### 3. METHODES DE BIOLOGIE CELLULAIRE

#### 3.1. Surexpression de protéines recombinantes chez E.coli :

Les bactéries BL21 (DE3) transformées par le plasmide recombinant sont cultivées à 37°C et induites en phase exponentielle de croissance avec 1 mM d'IPTG pendant 3 heures. Les fractions solubles et insolubles du lysat bactérien contenant la protéine chimérique étiquetée soit par le (His)<sub>6</sub> soit par la GST sont purifiées sur une colonne d'affinité Ni-NTA et sur billes de sépharose respectivement (*Qiagen et Pharmacia/Amersham*). Le protocole détaillé est disponible dans le manuel QIAexpressionist (*Qiagen*).

#### 3.2. Culture in vitro des parasites :

Nous avons utilisé la souche de toxoplasme non kystogène RH pour l'ensemble des extractions d'ARN ou de protéines et la souche RH(HX-), déficiente pour l'enzyme HXGPRT (hypoxanthine-xanthine-guanine phosphoribosyl transferase) pour les manipulations génétiques. Les parasites sont amplifiés par infections de cellules humaines adhérentes HFF (Human Foreskin Fibroblasts) et Hela (Human Epithelial cells cervical carcinoma) et maintenues en culture jusqu'à la lyse complète du tapis cellulaire. Les cellules infectées ou non par les parasites sont cultivées à 37°C, sous 5% CO<sub>2</sub>, en milieu DMEM (Dulbecco Minimum Essential Medium) de *Gibco*, additionné de 10% de sérum de veau fœtal (SVF), 1% de glutamine, 50 U/ml de pénicilline et 50µg/µl de streptomycine.

#### 3.3. Obtention de parasites recombinants stables :

Les parasites sont transfectés par électroporation lorsqu'ils sont extracellulaires (Roos *et al.*, 1994). Les parasites ayant intégré le plasmide d'intérêt sont sélectionnés à l'aide du système HXGPRT. L'acide mycophénolique ajouté dans le milieu de culture bloque la seule voie de synthèse des purines chez les toxoplasmes HXGPRT-. Par contre, les parasites intégrant le plasmide pmini/HXGPRT codant pour l'enzyme HXGPRT, survivent à l'acide mycophénolique en synthétisant leurs purines à partir de xanthine ajoutée au milieu de culture. Les toxoplasmes sont ensuite clonés par dilution limite (1 parasite par puits, dans des plaques de culture de 96 puits contenant des cellules HFF confluentes). Les plaques sont incubées pendant 5 jours puis observées au faible grossissement (x2.5) de façon à cribler les

puits où une seule plage de lyse s'est développée. Les clones sont transférés en plaques de 24 puits après 7 jours de prolifération et testés par immunofluorescence \_HA.

### 3.4. Test des clones par immunofluorescence indirecte (IFI) :

Les tests IFI sont réalisés sur des cellules HFF infectées ou non, et cultivées sur lamelles de verre. Les cellules sont fixées au formaldéhyde 5 %, puis perméabilisées à l'aide d'une solution de triton x100 (0,1 %). Les lamelles sont saturées en sérum de veau fœtal (5%) - sérum de chèvre (5 %). La protéine taguée HA-FLAG exprimée ectopiquement est détectée par incubation des lamelles avec l'anticorps primaire monoclonal anti-HA, dilué au 1/1000<sup>ème</sup>, puis avec un anticorps secondaire dilué au 1/500<sup>ème</sup>. Il s'agit d'un anticorps polyclonal de chèvre anti-IgG de lapin couplé à la fluorescéine (BODIPY, Molecular Probes). Les lamelles sont montées sur lame à l'aide du kit Prolong Antifad Kit (Molecular Probes), observées en épifluorescence au grossissement x100, à l'aide d'un microscope Axioplan 2 (Zeiss) muni d'un contraste de phase et photographiées à l'aide d'une caméra noir et blanc (AxioCam MRm, Zeiss) et du logiciel Axiovision 3.1 (Zeiss).

## 4. Méthodes de biochimie

### 4.1. Analyse d'extraits protéiques :

Les protéines sont séparées sur gels SDS-PAGE. Le tampon de migration est choisi en fonction de la taille de la protéine d'intérêt. Le tampon SDS PAGE Running (500 mL de tampon sont nécessaires pour la migration d'un gel. Le gel pré-coulé (gel gradient 4-12% Bis-Tris NuPAGE, Novex, USA) est rincé dans de l'H<sub>2</sub>O et placé dans l'appareil à électrophorèse (XCell II mini, Novex, USA) rempli de tampon de migration 1X. Les puits sont soigneusement rincés avec du tampon de migration 1X avant de charger les échantillons à un volume maximal de 20  $\mu$ L par puits. Le marqueur SeeBlue (Novex, USA) est utilisé pour visualiser les poids moléculaires. La migration est effectuée à 100V.

Les protéines sont révélées par coloration au nitrate d'argent (Silverexpress Staining Kit, *Invitrogen*) ou coloration au bleu de coomassie (Colloidal Blue Staining Kit, *Invitrogen*).

### 4.2. Transfert par Western blot s :

Après séparation, le transfert des protéines est effectué par électrobuvarinage sur une membrane PVDF de 0.2  $\mu$ m. La membrane est préhumidifiée par immersion 30s. dans le méthanol, rincée 1 min dans de l' ddH<sub>2</sub>O puis 2 min dans le tampon de transfert (Bicine 500 mM, Bis-Tris 500 mM, EDTA 20 mM, tampon de transfert NuPAGE, Novex, USA, additionné de 15% méthanol). Le transfert est effectué à voltage maximum (entre 80-100 V) pendant 60 min, ou à 30 V toute la nuit.

### 4.3. Immunorévélation des protéines :

Après transfert, la membrane est rincée plusieurs fois dans du PBS-T (PBS, 1%(v/v) Tween20), puis saturée 30 min sous agitation pour bloquer les sites aspécifiques dans du PBS-T contenant 5% de lait écrémé (m/v), en renouvelant la solution plusieurs fois. Puis la membrane est incubée (une nuit, à 4°C, sous agitation constante) en présence de l'anticorps primaire dilué dans du PBS-T additionné de lait écrémé à 5%. La membrane est ensuite rincée 15 min dans du PBS-T, en renouvelant la solution plusieurs fois, avant d'ajouter l'anticorps secondaire couplé à la phosphatase alcaline. Après 2h d'incubation sous agitation constante à température pièce, la membrane est rincée 10 min dans du PBS-T, en renouvelant la solution plusieurs fois. La membrane est révélée avec le substrat approprié : le BCIP/NBT (BCIP : 5-Bromo-4-chloro-3-indolyl phosphate ; NBT : Bitro Blue Tetrazolium).

### 4.4. Dosage des protéines par la méthode de Bradford :

Le réactif de Bradford contient du bleu de Coomassie se complexant aux protéines. Il permet, à l'aide d'une gamme étalon de concentration connue en albumine bovine de doser les protéines. La coloration se développe pendant 20 minutes à l'obscurité, puis l'absorbance est lue à 595 nm à l'aide du spectrophotomètre ( *Biophotometer, Eppendorf*).

### 4.5. Production et Purification de la protéine taguée Histidine et GST exprimée dans les bactéries :

#### 4.5.1. Surexpression de la protéine dans les bactéries :

## **Matériel et méthode**

Préparer une pré-culture de la souche d'intérêt en l'incubant pendant une nuit à 37 °C avec agitation dans 3 ml de milieu LB contenant 1/1000. Le 2<sup>ème</sup> jour, incuber 200 ml de milieu LB contenant 1/1000 l'antibiotique et 2 ml de la Pré-culture du Jour 1 à 37 °C sous agitation jusqu'à  $0.4 < DO < 0.8$ . Ajouter 1 mM d'IPTG et incuber pendant 3h à 37 °C sous agitation. Centrifuger à 4500 rpm à température 4 °C pendant 30 minutes pour obtenir le culot bactérien. Resuspendre le culot dans 5 ml de tampon de lyse, puis laisser à 4 °C pendant 30 minutes. Lysér les bactéries par Sonication ou en utilisant la press french. Centrifuger à 12000 rpm pendant 60 minutes (4 °C).

### **4.5.2. Purification de protéine recombinante avec une étiquette 6-His sur colonne Nickel :**

Equilibrer la colonne par 2 lavages avec de l'eau et 2 lavages avec le tampon de lyse (PBS 1X, 0.1% Triton, 5 mM Imidazole). Déposer l'extrait sur la colonne et incuber en batch pendant 1 à 2h. Laver 2 fois la colonne par le tampon de lavage (PBS 1X, 0.1% Triton, 20 mM Imidazole). Eluer la protéine avec 500  $\mu$ l de tampon d'élution (PBS 1X, 0.1% Triton, 0.3 M Imidazole) dans la colonne, répéter cette étape 2 à 3 fois.

### **4.5.3. Purification de protéine recombinante avec une étiquette GST sur colonne GSH-sepharose :**

Equilibrer la colonne par 2 lavages avec de l'eau et 2 lavages avec le tampon de lyse (10% Glycérol, 30mM HEPES, 50mM NaCl, 0.2mM EDTA, 0.5% Triton). Déposer l'extrait sur la colonne et incuber en batch pendant 1 à 2h. Laver 2 fois la colonne par le tampon de lavage (10% Glycérol, 30mM HEPES, 500mM NaCl, 0.2mM EDTA, 0.5% Triton). Eluer la protéine avec 700  $\mu$ l de tampon d'élution (10% Glycérol, 30mM HEPES, 50mM NaCl, 0.2mM EDTA, 0.5% Triton, \*mM NaOH 10mM Glutathione), répéter cette étape 2 à 3 fois.

### **4.5.4. Purification de la protéine taguée HA-Flag exprimée dans T.gondii :**

Après avoir amplifié le clone transgénique stable RH::HAFlag-TgHDAC3 ou RH::TgCARM1-HAFlag, les parasites sont repris dans le tampon A (20 % Glycérol, 20mM Tris HCl pH8, 500 mM KCl, 1.5mM MgCl<sub>2</sub>, 0.2mM EDTA, 0,5% NP-40) et lysés mécaniquement avec un poter. L'extrait cellulaire est placé sur une roue d'agitation (8 rpm) pendant 30 minutes à 4°C, puis centrifugé à 12,000 rpm pendant 30 minutes. Le surnageant contient les protéines solubles. Cet extrait est incubé pendant 1 heure à 4°C en présence de la

## Matériel et méthode

résine  $\mu$ -Flag, afin d'immunoprécipiter la protéine d'intérêt taguée Flag. Après un lavage au tampon A, les protéines sont éluées par compétition avec le peptide Flag (500 $\mu$ g/ml). Les fractions collectées sont analysées par SDS-PAGE et par western blot. Les fractions positives sont passées sur une colonne d'exclusion de taille (Superose 6 ou Superdex 200). Les fractions d'intérêt sont précipitées au TCA et séquencées par MS/MS (Collaboration Jérôme GARIN, INSERM-CEA).

### 4.6. Purification des anticorps sur colonne d'affinité :

#### 4.6.1. Liaison des antigènes sur colonne d'affinité :

L'immunopurification des anticorps est réalisée par colonne d'affinité (Affigel, Biorad). Selon le point isoélectrique (pI) de la protéine à purifier, on choisit la résine *Affigel 10* ( $6 < pI < 12$ ) ou *Affigel 15* ( $4 < pI < 6.5$ ).

La solution d'antigène est préparée dans un tampon dépourvu de groupements amines (sans Tris), par dialyse contre du tampon HEPES si nécessaire. Elle est diluée dans un tampon 100mM HEPES pH8 de manière à obtenir dans la colonne un ratio de 1 à 6 mg d'antigènes pour 1ml de résine *Affigel*, puis incubée 4h à 4°C sous agitation douce (5rpm sur la roue). On récupère ensuite l'excédant d'antigène qui est conservé à -20°C, avant de saturer la résine avec un tampon Tris (1M trispH8) et d'incuber pendant 1h à 2h. La colonne est lavée avec 10 à 20 volumes de colonne (CV) de tampon BC 500 (20 mM Tis-Hcl pH8 ; 0.2 mM EDTA, 10% glycerol et 500 mM Kcl) et équilibrée avec 10CV de tampon PBS.

#### 4.6.2. Liaison des anticorps sur colonne d'affinité :

Une fois la colonne chargée en antigènes et équilibrée, on ajoute le sérum d'anticorps et on incube toute la nuit à 4°C sous agitation douce (5rpm sur la roue). On récupère l'excédant, puis on lave avec 10 à 20 CV de tampon BC 500 et 10 CV de tampon PBS. L'éluat est répété 8 fois avec un tampon 0.1M glycine pH 2.5. L'éluat récupéré est immédiatement complété de 40 $\mu$ l de tris pH 8 à 2M. La colonne est neutralisée avec 10CV de tampon Tris et rééquilibrée avec du tampon PBS (pH 7 à 7.4).

Les anticorps sont dosés en utilisant le réactif de Bradford. Les éluats dont l'absorbance

## Matériel et méthode

est supérieure à 1 sont aliquotés en fraction de 600µl, auxquelles sont ajouté 100µl de glycérol 50%.

Les anticorps ainsi purifiés sont ensuite dilués au 1/500 ou 1/1000 lors des expériences d'immunodétection.

### 4.7. Immunoprécipitation de protéines par des anticorps immunopurifiés :

Le couplage des anticorps immunopurifiés sur protéine A-sepharose est réalisé selon Simanis *et al.*, 1985. Les immunoprécipitations sont réalisées comme décrit précédemment par Hakimi *et al.*, 2002.

Les anticorps immunopurifiés et les billes – protéine A sont mélangés volume à volume (V/V) dans une colonne 10 ml (Biorad), le volume final est ajusté à 10 ml, et incubés à température ambiante 1h sous agitation douce (roue). On récupère l'excédant et la colonne est lavé 2 fois avec 10 volumes de tampon 0,2M borate de sodium pH9 en centrifugeant à 700 rpm 5 minutes. Les billes sont re-suspendues dans 10 volumes de tampon 0,2M borate de sodium pH9 et 20mM de diméthylpimelimidate et incubées 30 minutes sous agitation douce à température ambiante. La réaction est stoppée en lavant les billes avec un tampon 0,2M d'éthanolamine pH8, pendant 2h à température ambiante sur roue. La colonne est lavée 3 fois avec du tampon PBS 1X, puis une fois avec 2 ml de tampon 100mM de glycine pH2,5 et de nouveau avec du tampon PBS 1X pour éliminer la glycine. La résine est alors incubée toute la nuit avec les extraits protéiques (300µl de WCE). Les billes sont lavées 3 fois 5 minutes avec le tampon BC300 avant l'élution des protéines liées aux anticorps. Pour cela, les billes sont incubées avec 30 µl du tampon 0,1M glycine 0,1M pH 2,5 pendant 10 minutes à 4°. On sépare alors les billes du surnageant et on ajoute 10µl de bleu de dépôt 4X au surnageant et 40µl de bleu 2X aux billes. Ces extraits sont déposés sur gel afin de détecter les protéines immunoprécipitées et coélues par western blot.

### 4.8. Extraction des cores histones

Les cores histones sont extraits des noyaux de toxoplasme selon la technique décrite par Workman *et al.*, 1991.

**4.9. Test d'activité des Protéines Histones-méthyltransférase (HMT) :**

Les tests d'activités des protéines arginines méthyltransférases (PRMT1 et CARM1) sont décrits dans l'article ci-dessus (Saksouk *et al.*, 2005).

**4.10. Test d'activité des Histones-acétylltransférase (HAT) :**

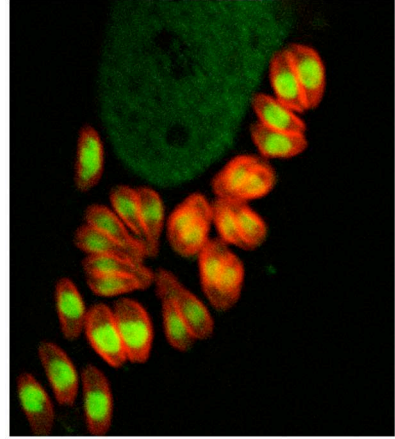
L'activité de l'histone acétyltransférase de la protéine TgGCN5-A a été mesurée comme décrit dans l'article Eberharter *et al.*, 1998.

**4.11. Test d'activité des Histones déacétylase (HDAC)**

Les activités des histones déacétylases ont été mesurées en utilisant le KIT colorimétrique « HDAC activity » (Calbiochem – Novabiochem, Darmstadt, Allemagne).



# BIBLIOGRAPHIE



## BIBLIOGRAPHIE

- Agresti, A. and Bianchi, M. E.  
2003 HMGB proteins and gene expression. *Curr.Opin.Genet.Dev.* 13(2): 170-178.
- Ahringer, J.  
2000 NuRD and SIN3 histone deacetylase complexes in development. *Trends Genet.* 16(8): 351-356.
- Ajzenberg, D., Banuls, A. L., Su, C., Dumetre, A., Demar, M., Carme, B., and Darde, M. L.  
2004 Genetic diversity, clonality and sexuality in *Toxoplasma gondii*. *Int.J.Parasitol.* 34(10): 1185-1196.
- An, W., Kim, J., and Roeder, R. G.  
2004 Ordered cooperative functions of PRMT1, p300, and CARM1 in transcriptional activation by p53. *Cell* 117(6): 735-748.
- An, W. and Roeder, R. G.  
2004 Reconstitution and transcriptional analysis of chromatin in vitro. *Methods Enzymol.* 377:460-474.
- Andres, M. E., Burger, C., Peral-Rubio, M. J., Battaglioli, E., Anderson, M. E., Grimes, J., Dallman, J., Ballas, N., and Mandel, G.  
1999 CoREST: a functional corepressor required for regulation of neural-specific gene expression. *Proc.Natl.Acad.Sci.U.S.A* 96(17): 9873-9878.
- Andrulis, E. D., Guzman, E., Doring, P., Werner, J., and Lis, J. T.  
2000 High-resolution localization of *Drosophila* Spt5 and Spt6 at heat shock genes in vivo: roles in promoter proximal pausing and transcription elongation. *Genes Dev.* 14(20): 2635-2649.
- Aravind, L., Anantharaman, V., Balaji, S., Babu, M. M., and Iyer, L. M.  
2005 The many faces of the helix-turn-helix domain: transcription regulation and beyond. *FEMS Microbiol.Rev.* 29(2): 231-262.
- Aravind, L., Iyer, L. M., Wellems, T. E., and Miller, L. H.  
2003 Plasmodium biology: genomic gleanings. *Cell* 115(7): 771-785.
- Balaji, S., Babu, M. M., Iyer, L. M., and Aravind, L.  
2005 Discovery of the principal specific transcription factors of Apicomplexa and their implication for the evolution of the AP2-integrase DNA binding domains. *Nucleic Acids Res.* 33(13): 3994-4006.
- Bannister, A. J. and Kouzarides, T.  
1996 The CBP co-activator is a histone acetyltransferase. *Nature* 384(6610): 641-643.
- Bauer, U. M., Daujat, S., Nielsen, S. J., Nightingale, K., and Kouzarides, T.  
2002 Methylation at arginine 17 of histone H3 is linked to gene activation. *EMBO Rep.* 3(1): 39-44.
- Berger, S. L.  
2002 Histone modifications in transcriptional regulation. *Curr.Opin.Genet.Dev.* 12(2): 142-148.

## Bibliographie

- Bhatti, M. M. and Sullivan, W. J., Jr.  
2005 Histone acetylase GCN5 enters the nucleus via importin-alpha in protozoan parasite *Toxoplasma gondii*. *J.Biol.Chem.* 280(7): 5902-5908.
- Bird, A. P. and Wolffe, A. P.  
1999 Methylation-induced repression--belts, braces, and chromatin. *Cell* 99(5): 451-454.
- Blander, G. and Guarente, L.  
2004 The Sir2 family of protein deacetylases. *Annu.Rev.Biochem.* 73:417-435.
- Boothroyd, J. C., Black, M., Bonnefoy, S., Hehl, A., Knoll, L. J., Manger, I. D., Ortega-Barria, E., and Tomavo, S.  
1997 Genetic and biochemical analysis of development in *Toxoplasma gondii*. *Philos.Trans.R.Soc.Lond B Biol.Sci.* 352(1359): 1347-1354.
- Braunstein, M., Sobel, R. E., Allis, C. D., Turner, B. M., and Broach, J. R.  
1996 Efficient transcriptional silencing in *Saccharomyces cerevisiae* requires a heterochromatin histone acetylation pattern. *Mol.Cell Biol.* 16(8): 4349-4356.
- Brownell, J. E. and Allis, C. D.  
1996 Special HATs for special occasions: linking histone acetylation to chromatin assembly and gene activation. *Curr.Opin.Genet.Dev.* 6(2): 176-184.
- Burke, L. J. and Baniahmad, A.  
2000 Co-repressors 2000. *FASEB J.* 14(13): 1876-1888.
- Burzio, L. O., Riquelme, P. T., and Koide, S. S.  
1979 ADP ribosylation of rat liver nucleosomal core histones. *J.Biol.Chem.* 254(8): 3029-3037.
- Buxton, D. and Innes, E. A.  
1995 A commercial vaccine for ovine toxoplasmosis. *Parasitology* 110 Suppl:S11-S16.
- Carruthers, V. B.  
2002 Host cell invasion by the opportunistic pathogen *Toxoplasma gondii*. *Acta Trop.* 81(2): 111-122.
- Catez, F., Lim, J. H., Hock, R., Postnikov, Y. V., and Bustin, M.  
2003 HMGN dynamics and chromatin function. *Biochem.Cell Biol.* 81(3): 113-122.
- Chen, C. J., Deng, Z., Kim, A. Y., Blobel, G. A., and Lieberman, P. M.  
2001 Stimulation of CREB binding protein nucleosomal histone acetyltransferase activity by a class of transcriptional activators. *Mol.Cell Biol.* 21(2): 476-487.
- Chen, G., Fernandez, J., Mische, S., and Courey, A. J.  
1999 A functional interaction between the histone deacetylase Rpd3 and the corepressor groucho in *Drosophila* development. *Genes Dev.* 13(17): 2218-2230.
- Chen, H., Lin, R. J., Schiltz, R. L., Chakravarti, D., Nash, A., Nagy, L., Privalsky, M. L., Nakatani, Y., and Evans, R. M.  
1997 Nuclear receptor coactivator ACTR is a novel histone acetyltransferase and forms a multimeric activation complex with P/CAF and CBP/p300. *Cell* 90(3): 569-580.
- Chen, H., Tini, M., and Evans, R. M.  
2001 HATs on and beyond chromatin. *Curr.Opin.Cell Biol.* 13(2): 218-224.

## Bibliographie

- Chen, Z. and Han, M.  
2001 C. elegans Rb, NuRD, and Ras regulate lin-39-mediated cell fusion during vulval fate specification. *Curr.Biol.* 11(23): 1874-1879.
- Cheng, D., Yadav, N., King, R. W., Swanson, M. S., Weinstein, E. J., and Bedford, M. T.  
2004 Small molecule regulators of protein arginine methyltransferases. *J.Biol.Chem.* 279(23): 23892-23899.
- Chervitz, S. A., Aravind, L., Sherlock, G., Ball, C. A., Koonin, E. V., Dwight, S. S., Harris, M. A., Dolinski, K., Mohr, S., Smith, T., Weng, S., Cherry, J. M., and Botstein, D.  
1998 Comparison of the complete protein sets of worm and yeast: orthology and divergence. *Science* 282(5396): 2022-2028.
- Cohen, S. N., Chang, A. C., and Hsu, L.  
1972 Nonchromosomal antibiotic resistance in bacteria: genetic transformation of Escherichia coli by R-factor DNA. *Proc.Natl.Acad.Sci.U.S.A* 69(8): 2110-2114.
- Colletti, S. L., Myers, R. W., Darkin-Rattray, S. J., Gurnett, A. M., Dulski, P. M., Galuska, S., Allocco, J. J., Ayer, M. B., Li, C., Lim, J., Crumley, T. M., Cannova, C., Schmatz, D. M., Wyvratt, M. J., Fisher, M. H., and Meinke, P. T.  
2001 Broad spectrum antiprotozoal agents that inhibit histone deacetylase: structure-activity relationships of apicidin. Part 2. *Bioorg.Med.Chem.Lett.* 11(2): 113-117.
- Darkin-Rattray, S. J., Gurnett, A. M., Myers, R. W., Dulski, P. M., Crumley, T. M., Allocco, J. J., Cannova, C., Meinke, P. T., Colletti, S. L., Bednarek, M. A., Singh, S. B., Goetz, M. A., Dombrowski, A. W., Polishook, J. D., and Schmatz, D. M.  
1996 Apicidin: a novel antiprotozoal agent that inhibits parasite histone deacetylase. *Proc.Natl.Acad.Sci.U.S.A* 93(23): 13143-13147.
- Daujat, S., Bauer, U. M., Shah, V., Turner, B., Berger, S., and Kouzarides, T.  
2002 Crosstalk between CARM1 methylation and CBP acetylation on histone H3. *Curr.Biol.* 12(24): 2090-2097.
- De Nadal, E., Zapater, M., Alepuz, P. M., Sumoy, L., Mas, G., and Posas, F.  
2004 The MAPK Hog1 recruits Rpd3 histone deacetylase to activate osmoresponsive genes. *Nature* 427(6972): 370-374.
- De Rubertis, F., Kadosh, D., Henchoz, S., Pauli, D., Reuter, G., Struhl, K., and Spierer, P.  
1996 The histone deacetylase RPD3 counteracts genomic silencing in Drosophila and yeast. *Nature* 384(6609): 589-591.
- de Ruijter, A. J., van Gennip, A. H., Caron, H. N., Kemp, S., and van Kuilenburg, A. B.  
2003 Histone deacetylases (HDACs): characterization of the classical HDAC family. *Biochem.J.* 370(Pt 3): 737-749.
- Denton, H., Roberts, C. W., Alexander, J., Thong, K. W., and Coombs, G. H.  
1996 Enzymes of energy metabolism in the bradyzoites and tachyzoites of Toxoplasma gondii. *FEMS Microbiol.Lett.* 137(1): 103-108.
- Desmots, G. and Couvreur, J.  
1974a [Isolation of the parasite in congenital toxoplasmosis: its practical and theoretical importance]. *Arch.Fr.Pediatr.* 31(2): 157-166.
- Desmots, G. and Couvreur, J.  
1974b Congenital toxoplasmosis. A prospective study of 378 pregnancies. *N.Engl.J.Med.* 290(20): 1110-1116.

## Bibliographie

- Dhalluin, C., Carlson, J. E., Zeng, L., He, C., Aggarwal, A. K., and Zhou, M. M.  
1999 Structure and ligand of a histone acetyltransferase bromodomain. *Nature* 399(6735): 491-496.
- Drapkin, R., Le Roy, G., Cho, H., Akoulitchev, S., and Reinberg, D.  
1996 Human cyclin-dependent kinase-activating kinase exists in three distinct complexes. *Proc.Natl.Acad.Sci.U.S.A* 93(13): 6488-6493.
- Dunn, A. Y., Melville, M. W., and Frydman, J.  
2001 Review: cellular substrates of the eukaryotic chaperonin TRiC/CCT. *J.Struct.Biol.* 135(2): 176-184.
- Eberharter, A., John, S., Grant, P. A., Utley, R. T., and Workman, J. L.  
1998 Identification and analysis of yeast nucleosomal histone acetyltransferase complexes. *Methods* 15(4): 315-321.
- Eberharter, A., Sterner, D. E., Schieltz, D., Hassan, A., Yates, J. R., III, Berger, S. L., and Workman, J. L.  
1999 The ADA complex is a distinct histone acetyltransferase complex in *Saccharomyces cerevisiae*. *Mol.Cell Biol.* 19(10): 6621-6631.
- Fan, Q., An, L., and Cui, L.  
2004a PfADA2, a Plasmodium falciparum homologue of the transcriptional coactivator ADA2 and its in vivo association with the histone acetyltransferase PfGCN5. *Gene* 336(2): 251-261.
- Fan, Q., An, L., and Cui, L.  
2004b Plasmodium falciparum histone acetyltransferase, a yeast GCN5 homologue involved in chromatin remodeling. *Eukaryot.Cell* 3(2): 264-276.
- Fan, Q., Li, J., Kariuki, M., and Cui, L.  
2004 Characterization of PfPuf2, member of the Puf family RNA-binding proteins from the malaria parasite Plasmodium falciparum. *DNA Cell Biol.* 23(11): 753-760.
- Ferguson, D. J. and Hutchison, W. M.  
1987 An ultrastructural study of the early development and tissue cyst formation of *Toxoplasma gondii* in the brains of mice. *Parasitol.Res.* 73(6): 483-491.
- Finnin, M. S., Donigian, J. R., Cohen, A., Richon, V. M., Rifkind, R. A., Marks, P. A., Breslow, R., and Pavletich, N. P.  
1999 Structures of a histone deacetylase homologue bound to the TSA and SAHA inhibitors. *Nature* 401(6749): 188-193.
- Forsberg, E. C. and Bresnick, E. H.  
2001 Histone acetylation beyond promoters: long-range acetylation patterns in the chromatin world. *Bioessays* 23(9): 820-830.
- Freitas-Junior, L. H., Hernandez-Rivas, R., Ralph, S. A., Montiel-Condado, D., Ruvalcaba-Salazar, O. K., Rojas-Meza, A. P., Mancio-Silva, L., Leal-Silvestre, R. J., Gontijo, A. M., Shorte, S., and Scherf, A.  
2005 Telomeric heterochromatin propagation and histone acetylation control mutually exclusive expression of antigenic variation genes in malaria parasites. *Cell* 121(1): 25-36.
- Grant, P. A. and Berger, S. L.  
1999 Histone acetyltransferase complexes. *Semin.Cell Dev.Biol.* 10(2): 169-177.

## Bibliographie

- Grant, P. A., Berger, S. L., and Workman, J. L.  
1999 Identification and analysis of native nucleosomal histone acetyltransferase complexes. *Methods Mol.Biol.* 119311-317.
- Grant, P. A., Duggan, L., Cote, J., Roberts, S. M., Brownell, J. E., Candau, R., Ohba, R., Owen-Hughes, T., Allis, C. D., Winston, F., Berger, S. L., and Workman, J. L.  
1997 Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex. *Genes Dev.* 11(13): 1640-1650.
- Grant, P. A., Schieltz, D., Pray-Grant, M. G., Yates, J. R., III, and Workman, J. L.  
1998 The ATM-related cofactor Tra1 is a component of the purified SAGA complex. *Mol.Cell* 2(6): 863-867.
- Grant, P. A., Sterner, D. E., Duggan, L. J., Workman, J. L., and Berger, S. L.  
1998 The SAGA unfolds: convergence of transcription regulators in chromatin-modifying complexes. *Trends Cell Biol.* 8(5): 193-197.
- Gray, S. G. and Ekstrom, T. J.  
2001 The human histone deacetylase family. *Exp.Cell Res.* 262(2): 75-83.
- Grunstein, M.  
1997 Histone acetylation in chromatin structure and transcription. *Nature* 389(6649): 349-352.
- Guenther, M. G., Yu, J., Kao, G. D., Yen, T. J., and Lazar, M. A.  
2002 Assembly of the SMRT-histone deacetylase 3 repression complex requires the TCP-1 ring complex. *Genes Dev.* 16(24): 3130-3135.
- Hakimi, M. A., Bochar, D. A., Chenoweth, J., Lane, W. S., Mandel, G., and Shiekhattar, R.  
2002 A core-BRAF35 complex containing histone deacetylase mediates repression of neuronal-specific genes. *Proc.Natl.Acad.Sci.U.S.A* 99(11): 7420-7425.
- Hakimi, M. A., Bochar, D. A., Schmiesing, J. A., Dong, Y., Barak, O. G., Speicher, D. W., Yokomori, K., and Shiekhattar, R.  
2002 A chromatin remodelling complex that loads cohesin onto human chromosomes. *Nature* 418(6901): 994-998.
- Hartl, F. U. and Hayer-Hartl, M.  
2002 Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science* 295(5561): 1852-1858.
- Hassan, A. H., Neely, K. E., and Workman, J. L.  
2001 Histone acetyltransferase complexes stabilize swi/snf binding to promoter nucleosomes. *Cell* 104(6): 817-827.
- Hawkes, N. A., Otero, G., Winkler, G. S., Marshall, N., Dahmus, M. E., Krappmann, D., Scheidereit, C., Thomas, C. L., Schiavo, G., Erdjument-Bromage, H., Tempst, P., and Svejstrup, J. Q.  
2002 Purification and characterization of the human elongator complex. *J.Biol.Chem.* 277(4): 3047-3052.
- Hebbes, T. R., Thorne, A. W., and Crane-Robinson, C.  
1988 A direct link between core histone acetylation and transcriptionally active chromatin. *EMBO J.* 7(5): 1395-1402.
- Hendzel, M. J., Sun, J. M., Chen, H. Y., Rattner, J. B., and Davie, J. R.  
1994 Histone acetyltransferase is associated with the nuclear matrix. *J.Biol.Chem.* 269(36): 22894-22901.

## Bibliographie

- Ho-Yen, D. O., Joss, A. W., Balfour, A. H., Smyth, E. T., Baird, D., and Chatterton, J. M.  
1992 Use of the polymerase chain reaction to detect *Toxoplasma gondii* in human blood samples. *J.Clin.Pathol.* 45(10): 910-913.
- Hudson, B. P., Martinez-Yamout, M. A., Dyson, H. J., and Wright, P. E.  
2000 Solution structure and acetyl-lysine binding activity of the GCN5 bromodomain. *J.Mol.Biol.* 304(3): 355-370.
- Humphrey, G. W., Wang, Y., Russanova, V. R., Hirai, T., Qin, J., Nakatani, Y., and Howard, B. H.  
2001 Stable histone deacetylase complexes distinguished by the presence of SANT domain proteins CoREST/kiaa0071 and Mta-L1. *J.Biol.Chem.* 276(9): 6817-6824.
- Jackson, J. P., Lindroth, A. M., Cao, X., and Jacobsen, S. E.  
2002 Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature* 416(6880): 556-560.
- Jackson, M. H. and Hutchison, W. M.  
1989 The prevalence and source of *Toxoplasma* infection in the environment. *Adv.Parasitol.* 28:55-105.
- Jacobson, R. H., Ladurner, A. G., King, D. S., and Tjian, R.  
2000 Structure and function of a human TAFII250 double bromodomain module. *Science* 288(5470): 1422-1425.
- Jenuwein, T.  
2001 Re-SET-ting heterochromatin by histone methyltransferases. *Trends Cell Biol.* 11(6): 266-273.
- Jenuwein, T. and Allis, C. D.  
2001 Translating the histone code. *Science* 293(5532): 1074-1080.
- Johnson, C. A., Barlow, A. L., and Turner, B. M.  
1998 Molecular cloning of *Drosophila melanogaster* cDNAs that encode a novel histone deacetylase dHDAC3. *Gene* 221(1): 127-134.
- Johnson, C. N., Adkins, N. L., and Georgel, P.  
2005 Chromatin remodeling complexes: ATP-dependent machines in action. *Biochem.Cell Biol.* 83(4): 405-417.
- Joshi, M. B., Lin, D. T., Chiang, P. H., Goldman, N. D., Fujioka, H., Aikawa, M., and Syin, C.  
1999 Molecular cloning and nuclear localization of a histone deacetylase homologue in *Plasmodium falciparum*. *Mol.Biochem.Parasitol.* 99(1): 11-19.
- Kadosh, D. and Struhl, K.  
1997 Repression by Ume6 involves recruitment of a complex containing Sin3 corepressor and Rpd3 histone deacetylase to target promoters. *Cell* 89(3): 365-371.
- Kaplan, C. D., Morris, J. R., Wu, C., and Winston, F.  
2000 Spt5 and spt6 are associated with active transcription and have characteristics of general elongation factors in *D. melanogaster*. *Genes Dev.* 14(20): 2623-2634.
- Khan, A. U. and Hampsey, M.  
2002 Connecting the DOTs: covalent histone modifications and the formation of silent chromatin. *Trends Genet.* 18(8): 387-389.
- Kim, J., Sif, S., Jones, B., Jackson, A., Koipally, J., Heller, E., Winandy, S., Viel, A., Sawyer, A., Ikeda, T., Kingston, R., and Georgopoulos, K.

## Bibliographie

- 1999 Ikaros DNA-binding proteins direct formation of chromatin remodeling complexes in lymphocytes. *Immunity*. 10(3): 345-355.
- Kim, M. S., Son, M. W., Kim, W. B., In, Park Y., and Moon, A.  
2000 Apicidin, an inhibitor of histone deacetylase, prevents H-ras-induced invasive phenotype. *Cancer Lett.* 157(1): 23-30.
- Kingston, R. E. and Narlikar, G. J.  
1999 ATP-dependent remodeling and acetylation as regulators of chromatin fluidity. *Genes Dev.* 13(18): 2339-2352.
- Knoepfler, P. S. and Eisenman, R. N.  
1999 Sin meets NuRD and other tails of repression. *Cell* 99(5): 447-450.
- Koh, S. S., Chen, D., Lee, Y. H., and Stallcup, M. R.  
2001 Synergistic enhancement of nuclear receptor function by p160 coactivators and two coactivators with protein methyltransferase activities. *J.Biol.Chem.* 276(2): 1089-1098.
- Kouzarides, T.  
2002 Histone methylation in transcriptional control. *Curr.Opin.Genet.Dev.* 12(2): 198-209.
- Kreimeyer, A., Wielckens, K., Adamietz, P., and Hilz, H.  
1984 DNA repair-associated ADP-ribosylation in vivo. Modification of histone H1 differs from that of the principal acceptor proteins. *J.Biol.Chem.* 259(2): 890-896.
- Krogan, N. J., Dover, J., Khorrami, S., Greenblatt, J. F., Schneider, J., Johnston, M., and Shilatifard, A.  
2002 COMPASS, a histone H3 (Lysine 4) methyltransferase required for telomeric silencing of gene expression. *J.Biol.Chem.* 277(13): 10753-10755.
- Krogan, N. J. and Greenblatt, J. F.  
2001 Characterization of a six-subunit holo-elongator complex required for the regulated expression of a group of genes in *Saccharomyces cerevisiae*. *Mol.Cell Biol.* 21(23): 8203-8212.
- Kuo, M. H. and Allis, C. D.  
1998 Roles of histone acetyltransferases and deacetylases in gene regulation. *Bioessays* 20(8): 615-626.
- Kuo, M. H., Brownell, J. E., Sobel, R. E., Ranalli, T. A., Cook, R. G., Edmondson, D. G., Roth, S. Y., and Allis, C. D.  
1996 Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines. *Nature* 383(6597): 269-272.
- Kuo, M. H., Zhou, J., Jambeck, P., Churchill, M. E., and Allis, C. D.  
1998 Histone acetyltransferase activity of yeast Gcn5p is required for the activation of target genes in vivo. *Genes Dev.* 12(5): 627-639.
- Kurdistani, S. K. and Grunstein, M.  
2003 In vivo protein-protein and protein-DNA crosslinking for genomewide binding microarray. *Methods* 31(1): 90-95.
- Lachner, M. and Jenuwein, T.  
2002 The many faces of histone lysine methylation. *Curr.Opin.Cell Biol.* 14(3): 286-298.



## Bibliographie

- Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T.  
2001 Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* 410(6824): 116-120.
- Lachner, M., O'Sullivan, R. J., and Jenuwein, T.  
2003 An epigenetic road map for histone lysine methylation. *J.Cell Sci.* 116(Pt 11): 2117-2124.
- Lacoste, N., Utley, R. T., Hunter, J. M., Poirier, G. G., and Cote, J.  
2002 Disruptor of telomeric silencing-1 is a chromatin-specific histone H3 methyltransferase. *J.Biol.Chem.* 277(34): 30421-30424.
- Lee, D. Y., Hayes, J. J., Pruss, D., and Wolffe, A. P.  
1993 A positive role for histone acetylation in transcription factor access to nucleosomal DNA. *Cell* 72(1): 73-84.
- Lee, M. G., Wynder, C., Cooch, N., and Shiekhattar, R.  
2005 An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation. *Nature* .
- Lee, M. S., Kliewer, S. A., Provencal, J., Wright, P. E., and Evans, R. M.  
1993 Structure of the retinoid X receptor alpha DNA binding domain: a helix required for homodimeric DNA binding. *Science* 260(5111): 1117-1121.
- Lespinet, O., Wolf, Y. I., Koonin, E. V., and Aravind, L.  
2002 The role of lineage-specific gene family expansion in the evolution of eukaryotes. *Genome Res.* 12(7): 1048-1059.
- Lo, W. S., Duggan, L., Emre, N. C., Belotserkovskya, R., Lane, W. S., Shiekhattar, R., and Berger, S. L.  
2001 Snf1--a histone kinase that works in concert with the histone acetyltransferase Gcn5 to regulate transcription. *Science* 293(5532): 1142-1146.
- Lorch, Y., Zhang, M., and Kornberg, R. D.  
1999 Histone octamer transfer by a chromatin-remodeling complex. *Cell* 96(3): 389-392.
- Lyons, R. E., McLeod, R., and Roberts, C. W.  
2002 *Toxoplasma gondii* tachyzoite-bradyzoite interconversion. *Trends Parasitol.* 18(5): 198-201.
- Manning, E. T., Ikehara, T., Ito, T., Kadonaga, J. T., and Kraus, W. L.  
2001 p300 forms a stable, template-committed complex with chromatin: role for the bromodomain. *Mol.Cell Biol.* 21(12): 3876-3887.
- Manning, L. R. and Manning, J. M.  
2001 The acetylation state of human fetal hemoglobin modulates the strength of its subunit interactions: long-range effects and implications for histone interactions in the nucleosome. *Biochemistry* 40(6): 1635-1639.
- Marks, P., Rifkind, R. A., Richon, V. M., Breslow, R., Miller, T., and Kelly, W. K.  
2001 Histone deacetylases and cancer: causes and therapies. *Nat.Rev.Cancer* 1(3): 194-202.
- Marks, P. A., Richon, V. M., Breslow, R., and Rifkind, R. A.  
2001 Histone deacetylase inhibitors as new cancer drugs. *Curr.Opin.Oncol.* 13(6): 477-483.

## Bibliographie

- Marks, P. A., Richon, V. M., Miller, T., and Kelly, W. K.  
2004 Histone deacetylase inhibitors. *Adv.Cancer Res.* 91:137-168.
- Marks, P. A., Rifkind, R. A., Richon, V. M., and Breslow, R.  
2001 Inhibitors of histone deacetylase are potentially effective anticancer agents. *Clin.Cancer Res.* 7(4): 759-760.
- Marmorstein, R.  
2001 Structure and function of histone acetyltransferases. *Cell Mol.Life Sci.* 58(5-6): 693-703.
- Marmorstein, R. and Berger, S. L.  
2001 Structure and function of bromodomains in chromatin-regulating complexes. *Gene* 272(1-2): 1-9.
- Meinke, P. T., Colletti, S. L., Doss, G., Myers, R. W., Gurnett, A. M., Dulski, P. M., Darkin-Rattray, S. J., Allocco, J. J., Galuska, S., Schmatz, D. M., Wyvratt, M. J., and Fisher, M. H.  
2000 Synthesis of apicidin-derived quinolone derivatives: parasite-selective histone deacetylase inhibitors and antiproliferative agents. *J.Med.Chem.* 43(25): 4919-4922.
- Meissner, M., Krejany, E., Gilson, P. R., Koning-Ward, T. F., Soldati, D., and Crabb, B. S.  
2005 Tetracycline analogue-regulated transgene expression in Plasmodium falciparum blood stages using Toxoplasma gondii transactivators. *Proc.Natl.Acad.Sci.U.S.A* 102(8): 2980-2985.
- Meissner, M. and Soldati, D.  
2005 The transcription machinery and the molecular toolbox to control gene expression in Toxoplasma gondii and other protozoan parasites. *Microbes.Infect.*
- Mele, A., Paterson, P. J., Prentice, H. G., Leoni, P., and Kibbler, C. C.  
2002 Toxoplasmosis in bone marrow transplantation: a report of two cases and systematic review of the literature. *Bone Marrow Transplant.* 29(8): 691-698.
- Melki, R., Batelier, G., Soulie, S., and Williams, R. C., Jr.  
1997 Cytoplasmic chaperonin containing TCP-1: structural and functional characterization. *Biochemistry* 36(19): 5817-5826.
- Messina, D. N., Glasscock, J., Gish, W., and Lovett, M.  
2004 An ORFeome-based analysis of human transcription factor genes and the construction of a microarray to interrogate their expression. *Genome Res.* 14(10B): 2041-2047.
- Messina, M., Kim, S., and Sibley, L. D.  
1996 A family of dispersed DNA elements that contain GAA repeats in Toxoplasma gondii. *Mol.Biochem.Parasitol.* 81(2): 247-252.
- Messina, M., Niesman, I., Mercier, C., and Sibley, L. D.  
1995 Stable DNA transformation of Toxoplasma gondii using phleomycin selection. *Gene* 165(2): 213-217.
- Metzger, E., Wissmann, M., Yin, N., Muller, J. M., Schneider, R., Peters, A. H., Gunther, T., Buettner, R., and Schule, R.  
2005 LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. *Nature* .
- Miller, T., Krogan, N. J., Dover, J., Erdjument-Bromage, H., Tempst, P., Johnston, M., Greenblatt, J. F., and Shilatifard, A.

## Bibliographie

- 2001 COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. *Proc.Natl.Acad.Sci.U.S.A* 98(23): 12902-12907.
- Mitchell, D. and Bell, A.  
2003 PEST sequences in the malaria parasite Plasmodium falciparum: a genomic study. *Malar.J.* 216.
- Mizzen, C. A., Yang, X. J., Kokubo, T., Brownell, J. E., Bannister, A. J., Owen-Hughes, T., Workman, J., Wang, L., Berger, S. L., Kouzarides, T., Nakatani, Y., and Allis, C. D.  
1996 The TAF(II)250 subunit of TFIID has histone acetyltransferase activity. *Cell* 87(7): 1261-1270.
- Montoya, J. G. and Liesenfeld, O.  
2004 Toxoplasmosis. *Lancet* 363(9425): 1965-1976.
- Morlat, P., Chene, G., Lepout, C., Rousseau, F., Luft, B., Aubertin, J., Hafner, R., Salamon, R., and Vilde, J. L.  
1993 [Primary prevention of cerebral toxoplasmosis in patients with HIV infection: results of a double-blind randomized trial, pyrimethamine versus placebo]. *Rev.Med.Interne* 14(10): 1002.
- Naar, A. M., Lemon, B. D., and Tjian, R.  
2001 Transcriptional coactivator complexes. *Annu.Rev.Biochem.* 70:475-501.
- Nagy, P. L., Griesenbeck, J., Kornberg, R. D., and Cleary, M. L.  
2002 A trithorax-group complex purified from Saccharomyces cerevisiae is required for methylation of histone H3. *Proc.Natl.Acad.Sci.U.S.A* 99(1): 90-94.
- Narlikar, G. J., Fan, H. Y., and Kingston, R. E.  
2002 Cooperation between complexes that regulate chromatin structure and transcription. *Cell* 108(4): 475-487.
- Nishioka, K., Chuikov, S., Sarma, K., Erdjument-Bromage, H., Allis, C. D., Tempst, P., and Reinberg, D.  
2002 Set9, a novel histone H3 methyltransferase that facilitates transcription by precluding histone tail modifications required for heterochromatin formation. *Genes Dev.* 16(4): 479-489.
- Nishioka, K., Rice, J. C., Sarma, K., Erdjument-Bromage, H., Werner, J., Wang, Y., Chuikov, S., Valenzuela, P., Tempst, P., Steward, R., Lis, J. T., Allis, C. D., and Reinberg, D.  
2002 PR-Set7 is a nucleosome-specific methyltransferase that modifies lysine 20 of histone H4 and is associated with silent chromatin. *Mol.Cell* 9(6): 1201-1213.
- Noma, K., Allis, C. D., and Grewal, S. I.  
2001 Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries. *Science* 293(5532): 1150-1155.
- Nowak, S. J. and Corces, V. G.  
2004 Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation. *Trends Genet.* 20(4): 214-220.
- Ogryzko, V. V., Kotani, T., Zhang, X., Schiltz, R. L., Howard, T., Yang, X. J., Howard, B. H., Qin, J., and Nakatani, Y.  
1998 Histone-like TAFs within the PCAF histone acetylase complex. *Cell* 94(1): 35-44.
- Ogryzko, V. V., Schiltz, R. L., Russanova, V., Howard, B. H., and Nakatani, Y.  
1996 The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell* 87(5): 953-959.

## Bibliographie

- Oliva, R., Bazett-Jones, D. P., Locklear, L., and Dixon, G. H.  
1990 Histone hyperacetylation can induce unfolding of the nucleosome core particle. *Nucleic Acids Res.* 18(9): 2739-2747.
- Orlando, V.  
2000 Mapping chromosomal proteins in vivo by formaldehyde-crosslinked-chromatin immunoprecipitation. *Trends Biochem.Sci.* 25(3): 99-104.
- Orphanides, G., LeRoy, G., Chang, C. H., Luse, D. S., and Reinberg, D.  
1998 FACT, a factor that facilitates transcript elongation through nucleosomes. *Cell* 92(1): 105-116.
- Orphanides, G., Wu, W. H., Lane, W. S., Hampsey, M., and Reinberg, D.  
1999 The chromatin-specific transcription elongation factor FACT comprises human SPT16 and SSRP1 proteins. *Nature* 400(6741): 284-288.
- Osley, M. A.  
2004 H2B ubiquitylation: the end is in sight. *Biochim.Biophys.Acta* 1677(1-3): 74-78.
- Otero, G., Fellows, J., Li, Y., de Bizemont, T., Dirac, A. M., Gustafsson, C. M., Erdjument-Bromage, H., Tempst, P., and Svejstrup, J. Q.  
1999 Elongator, a multisubunit component of a novel RNA polymerase II holoenzyme for transcriptional elongation. *Mol.Cell* 3(1): 109-118.
- Peterson, C. L. and Laniel, M. A.  
2004 Histones and histone modifications. *Curr.Biol.* 14(14): R546-R551.
- Petrokovski, S.  
2001 Intein spread and extinction in evolution. *Trends Genet.* 17(8): 465-472.
- Pijnappel, W. W., Schaft, D., Roguev, A., Shevchenko, A., Tekotte, H., Wilm, M., Rigaut, G., Seraphin, B., Aasland, R., and Stewart, A. F.  
2001 The *S. cerevisiae* SET3 complex includes two histone deacetylases, Hos2 and Hst1, and is a meiotic-specific repressor of the sporulation gene program. *Genes Dev.* 15(22): 2991-3004.
- Pray-Grant, M. G., Schieltz, D., McMahon, S. J., Wood, J. M., Kennedy, E. L., Cook, R. G., Workman, J. L., Yates, J. R., III, and Grant, P. A.  
2002 The novel SLIK histone acetyltransferase complex functions in the yeast retrograde response pathway. *Mol.Cell Biol.* 22(24): 8774-8786.
- Rechsteiner, M. and Rogers, S. W.  
1996 PEST sequences and regulation by proteolysis. *Trends Biochem.Sci.* 21(7): 267-271.
- Reeves, R. and Beckerbauer, L.  
2001 HMG1/Y proteins: flexible regulators of transcription and chromatin structure. *Biochim.Biophys.Acta* 1519(1-2): 13-29.
- Rho, J., Choi, S., Seong, Y. R., Cho, W. K., Kim, S. H., and Im, D. S.  
2001 Prmt5, which forms distinct homo-oligomers, is a member of the protein-arginine methyltransferase family. *J.Biol.Chem.* 276(14): 11393-11401.
- Richman, R., Chicoine, L. G., Collini, M. P., Cook, R. G., and Allis, C. D.  
1988 Micronuclei and the cytoplasm of growing *Tetrahymena* contain a histone acetylase activity which is highly specific for free histone H4. *J.Cell Biol.* 106(4): 1017-1026.

## Bibliographie

- Riechmann, J. L., Heard, J., Martin, G., Reuber, L., Jiang, C., Keddie, J., Adam, L., Pineda, O., Ratcliffe, O. J., Samaha, R. R., Creelman, R., Pilgrim, M., Broun, P., Zhang, J. Z., Ghandehari, D., Sherman, B. K., and Yu, G.  
2000 Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. *Science* 290(5499): 2105-2110.
- Roguev, A., Schaft, D., Shevchenko, A., Pijnappel, W. W., Wilm, M., Aasland, R., and Stewart, A. F.  
2001 The *Saccharomyces cerevisiae* Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4. *EMBO J.* 20(24): 7137-7148.
- Roopra, A., Sharling, L., Wood, I. C., Briggs, T., Bachfischer, U., Paquette, A. J., and Buckley, N. J.  
2000 Transcriptional repression by neuron-restrictive silencer factor is mediated via the Sin3-histone deacetylase complex. *Mol. Cell Biol.* 20(6): 2147-2157.
- Santos-Rosa, H., Schneider, R., Bannister, A. J., Sherriff, J., Bernstein, B. E., Emre, N. C., Schreiber, S. L., Mellor, J., and Kouzarides, T.  
2002 Active genes are tri-methylated at K4 of histone H3. *Nature* 419(6905): 407-411.
- Schnitzler, G., Sif, S., and Kingston, R. E.  
1998 Human SWI/SNF interconverts a nucleosome between its base state and a stable remodeled state. *Cell* 94(1): 17-27.
- Schnitzler, G. R., Sif, S., and Kingston, R. E.  
1998 A model for chromatin remodeling by the SWI/SNF family. *Cold Spring Harb. Symp. Quant. Biol.* 63:535-543.
- Schurter, B. T., Koh, S. S., Chen, D., Bunick, G. J., Harp, J. M., Hanson, B. L., Henschen-Edman, A., Mackay, D. R., Stallcup, M. R., and Aswad, D. W.  
2001 Methylation of histone H3 by coactivator-associated arginine methyltransferase 1. *Biochemistry* 40(19): 5747-5756.
- Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J. R., Cole, P. A., Casero, R. A., and Shi, Y.  
2004 Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* 119(7): 941-953.
- Shi, Y. J., Matson, C., Lan, F., Iwase, S., Baba, T., and Shi, Y.  
2005 Regulation of LSD1 Histone Demethylase Activity by Its Associated Factors. *Mol. Cell* .
- Simanis, V. and Lane, D. P.  
1985 An immunoaffinity purification procedure for SV40 large T antigen. *Virology* 144(1): 88-100.
- Singh, U., Brewer, J. L., and Boothroyd, J. C.  
2002 Genetic analysis of tachyzoite to bradyzoite differentiation mutants in *Toxoplasma gondii* reveals a hierarchy of gene induction. *Mol. Microbiol.* 44(3): 721-733.
- Spencer, T. E., Jenster, G., Burcin, M. M., Allis, C. D., Zhou, J., Mizzen, C. A., McKenna, N. J., Onate, S. A., Tsai, S. Y., Tsai, M. J., and O'Malley, B. W.  
1997 Steroid receptor coactivator-1 is a histone acetyltransferase. *Nature* 389(6647): 194-198.
- Stallcup, M. R., Chen, D., Koh, S. S., Ma, H., Lee, Y. H., Li, H., Schurter, B. T., and Aswad, D. W.

## Bibliographie

- 2000 Co-operation between protein-acetylating and protein-methylating co-activators in transcriptional activation. *Biochem.Soc.Trans.* 28(4): 415-418.
- Sterner, D. E., Belotserkovskaya, R., and Berger, S. L.  
2002 SALSA, a variant of yeast SAGA, contains truncated Spt7, which correlates with activated transcription. *Proc.Natl.Acad.Sci.U.S.A* 99(18): 11622-11627.
- Struhl, K.  
1998 Histone acetylation and transcriptional regulatory mechanisms. *Genes Dev.* 12(5): 599-606.
- Struhl, K., Kadosh, D., Keaveney, M., Kuras, L., and Moqtaderi, Z.  
1998 Activation and repression mechanisms in yeast. *Cold Spring Harb.Symp.Quant.Biol.* 63:413-421.
- Struhl, K. and Moqtaderi, Z.  
1998 The TAFs in the HAT. *Cell* 94(1): 1-4.
- Studitsky, V. M., Clark, D. J., and Felsenfeld, G.  
1994 A histone octamer can step around a transcribing polymerase without leaving the template. *Cell* 76(2): 371-382.
- Studitsky, V. M., Kassavetis, G. A., Geiduschek, E. P., and Felsenfeld, G.  
1997 Mechanism of transcription through the nucleosome by eukaryotic RNA polymerase. *Science* 278(5345): 1960-1963.
- Suka, N., Luo, K., and Grunstein, M.  
2002 Sir2p and Sas2p opposingly regulate acetylation of yeast histone H4 lysine16 and spreading of heterochromatin. *Nat.Genet.* 32(3): 378-383.
- Sullivan, W. J., Jr. and Smith, C. K.  
2000 Cloning and characterization of a novel histone acetyltransferase homologue from the protozoan parasite *Toxoplasma gondii* reveals a distinct GCN5 family member. *Gene* 242(1-2): 193-200.
- Syntichaki, P., Topalidou, I., and Thireos, G.  
2000 The Gcn5 bromodomain co-ordinates nucleosome remodelling. *Nature* 404(6776): 414-417.
- Tachibana, M., Sugimoto, K., Nozaki, M., Ueda, J., Ohta, T., Ohki, M., Fukuda, M., Takeda, N., Niida, H., Kato, H., and Shinkai, Y.  
2002 G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes Dev.* 16(14): 1779-1791.
- Tamaru, H. and Selker, E. U.  
2001 A histone H3 methyltransferase controls DNA methylation in *Neurospora crassa*. *Nature* 414(6861): 277-283.
- Tang, J., Gary, J. D., Clarke, S., and Herschman, H. R.  
1998 PRMT 3, a type I protein arginine N-methyltransferase that differs from PRMT1 in its oligomerization, subcellular localization, substrate specificity, and regulation. *J.Biol.Chem.* 273(27): 16935-16945.
- Templeton, T. J., Iyer, L. M., Anantharaman, V., Enomoto, S., Abrahante, J. E., Subramanian, G. M., Hoffman, S. L., Abrahamsen, M. S., and Aravind, L.  
2004 Comparative analysis of apicomplexa and genomic diversity in eukaryotes. *Genome Res.* 14(9): 1686-1695.

## Bibliographie

- Tenant-Flowers, M., Boyle, M. J., Carey, D., Marriott, D. J., Harkness, J. L., Penny, R., and Cooper, D. A.  
1991 Sulphadiazine desensitization in patients with AIDS and cerebral toxoplasmosis. *AIDS* 5(3): 311-315.
- van Holde, K. E., Lohr, D. E., and Robert, C.  
1992 What happens to nucleosomes during transcription? *J.Biol.Chem.* 267(5): 2837-2840.
- Van Lint, C., Emiliani, S., Ott, M., and Verdin, E.  
1996 Transcriptional activation and chromatin remodeling of the HIV-1 promoter in response to histone acetylation. *EMBO J.* 15(5): 1112-1120.
- Van Lint, C., Emiliani, S., and Verdin, E.  
1996 The expression of a small fraction of cellular genes is changed in response to histone hyperacetylation. *Gene Expr.* 5(4-5): 245-253.
- Vannier, D., Balderes, D., and Shore, D.  
1996 Evidence that the transcriptional regulators SIN3 and RPD3, and a novel gene (SDS3) with similar functions, are involved in transcriptional silencing in *S. cerevisiae*. *Genetics* 144(4): 1343-1353.
- Verdin, E., Dequiedt, F., and Kasler, H. G.  
2003 Class II histone deacetylases: versatile regulators. *Trends Genet.* 19(5): 286-293.
- Vermaak, D., Wade, P. A., Jones, P. L., Shi, Y. B., and Wolffe, A. P.  
1999 Functional analysis of the SIN3-histone deacetylase RPD3-RbAp48-histone H4 connection in the *Xenopus* oocyte. *Mol.Cell Biol.* 19(9): 5847-5860.
- Vettese-Dadey, M., Grant, P. A., Hebbes, T. R., Robinson, C., Allis, C. D., and Workman, J. L.  
1996 Acetylation of histone H4 plays a primary role in enhancing transcription factor binding to nucleosomal DNA in vitro. *EMBO J.* 15(10): 2508-2518.
- Vogelauer, M., Wu, J., Suka, N., and Grunstein, M.  
2000 Global histone acetylation and deacetylation in yeast. *Nature* 408(6811): 495-498.
- Wade, P. A., Jones, P. L., Vermaak, D., Veenstra, G. J., Imhof, A., Sera, T., Tse, C., Ge, H., Shi, Y. B., Hansen, J. C., and Wolffe, A. P.  
1998 Histone deacetylase directs the dominant silencing of transcription in chromatin: association with MeCP2 and the Mi-2 chromodomain SWI/SNF ATPase. *Cold Spring Harb.Symp.Quant.Biol.* 63:435-445.
- Wade, P. A., Jones, P. L., Vermaak, D., and Wolffe, A. P.  
1998 A multiple subunit Mi-2 histone deacetylase from *Xenopus laevis* cofractionates with an associated Snf2 superfamily ATPase. *Curr.Biol.* 8(14): 843-846.
- Wade, P. A., Pruss, D., and Wolffe, A. P.  
1997 Histone acetylation: chromatin in action. *Trends Biochem.Sci.* 22(4): 128-132.
- Wade, P. A. and Wolffe, A. P.  
1997 Histone acetyltransferases in control. *Curr.Biol.* 7(2): R82-R84.
- Wang, A., Kurdistani, S. K., and Grunstein, M.  
2002 Requirement of Hos2 histone deacetylase for gene activity in yeast. *Science* 298(5597): 1412-1414.

## Bibliographie

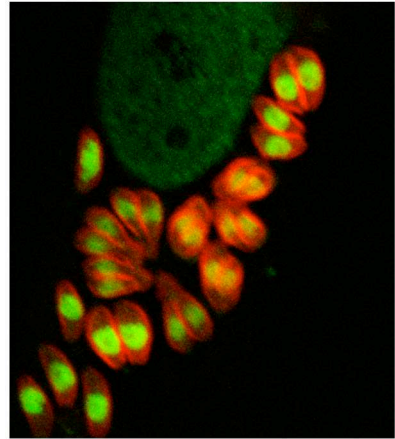
- Wang, L., Liu, L., and Berger, S. L.  
1998 Critical residues for histone acetylation by Gcn5, functioning in Ada and SAGA complexes, are also required for transcriptional function in vivo. *Genes Dev.* 12(5): 640-653.
- Wang, Y., Fischle, W., Cheung, W., Jacobs, S., Khorasanizadeh, S., and Allis, C. D.  
2004 Beyond the double helix: writing and reading the histone code. *Novartis.Found.Symp.* 2593-17.
- Wang, Y., Wysocka, J., Sayegh, J., Lee, Y. H., Perlin, J. R., Leonelli, L., Sonbuchner, L. S., McDonald, C. H., Cook, R. G., Dou, Y., Roeder, R. G., Clarke, S., Stallcup, M. R., Allis, C. D., and Coonrod, S. A.  
2004 Human PAD4 regulates histone arginine methylation levels via demethylimination. *Science* 306(5694): 279-283.
- Weiss, L. M. and Kim, K.  
2000 The development and biology of bradyzoites of *Toxoplasma gondii*. *Front Biosci.* 5D391-D405.
- Weiss, V. H., McBride, A. E., Soriano, M. A., Filman, D. J., Silver, P. A., and Hogle, J. M.  
2000 The structure and oligomerization of the yeast arginine methyltransferase, Hmt1. *Nat.Struct.Biol.* 7(12): 1165-1171.
- Winston, F. and Sudarsanam, P.  
1998 The SAGA of Spt proteins and transcriptional analysis in yeast: past, present, and future. *Cold Spring Harb.Symp.Quant.Biol.* 63553-561.
- Wolffe, A. P. and Hayes, J. J.  
1999 Chromatin disruption and modification. *Nucleic Acids Res.* 27(3): 711-720.
- Workman, J. L. and Kingston, R. E.  
1998 Alteration of nucleosome structure as a mechanism of transcriptional regulation. *Annu.Rev.Biochem.* 67545-579.
- Workman, J. L., Taylor, I. C., and Kingston, R. E.  
1991 Activation domains of stably bound GAL4 derivatives alleviate repression of promoters by nucleosomes. *Cell* 64(3): 533-544.
- Workman, J. L., Taylor, I. C., Kingston, R. E., and Roeder, R. G.  
1991 Control of class II gene transcription during in vitro nucleosome assembly. *Methods Cell Biol.* 35419-447.
- Wysocka, J., Milne, T. A., and Allis, C. D.  
2005 Taking LSD1 to a New High. *Cell* 122(5): 654-658.
- Wysocka, J., Swigut, T., Milne, T. A., Dou, Y., Zhang, X., Burlingame, A. L., Roeder, R. G., Brivanlou, A. H., and Allis, C. D.  
2005 WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell* 121(6): 859-872.
- Xiao, B., Wilson, J. R., and Gambin, S. J.  
2003 SET domains and histone methylation. *Curr.Opin.Struct.Biol.* 13(6): 699-705.
- Xu, W., Chen, H., Du, K., Asahara, H., Tini, M., Emerson, B. M., Montminy, M., and Evans, R. M.  
2001 A transcriptional switch mediated by cofactor methylation. *Science* 294(5551): 2507-2511.



## Bibliographie

- Xu, W., Cho, H., and Evans, R. M.  
2003 Acetylation and methylation in nuclear receptor gene activation. *Methods Enzymol.* 364:205-223.
- Xu, W., Cho, H., Kadam, S., Banayo, E. M., Anderson, S., Yates, J. R., III, Emerson, B. M., and Evans, R. M.  
2004 A methylation-mediator complex in hormone signaling. *Genes Dev.* 18(2): 144-156.
- Yoon, H. G., Chan, D. W., Huang, Z. Q., Li, J., Fondell, J. D., Qin, J., and Wong, J.  
2003 Purification and functional characterization of the human N-CoR complex: the roles of HDAC3, TBL1 and TBLR1. *EMBO J.* 22(6): 1336-1346.
- Yoon, H. G., Chan, D. W., Reynolds, A. B., Qin, J., and Wong, J.  
2003 N-CoR mediates DNA methylation-dependent repression through a methyl CpG binding protein Kaiso. *Mol.Cell* 12(3): 723-734.
- You, A., Tong, J. K., Grozinger, C. M., and Schreiber, S. L.  
2001 CoREST is an integral component of the Co. *Proc.Natl.Acad.Sci.U.S.A* 98(4): 1454-1458.
- Zhang, X. and Cheng, X.  
2003 Structure of the predominant protein arginine methyltransferase PRMT1 and analysis of its binding to substrate peptides. *Structure.(Camb.)* 11(5): 509-520.
- Zhang, X., Zhou, L., and Cheng, X.  
2000 Crystal structure of the conserved core of protein arginine methyltransferase PRMT3. *EMBO J.* 19(14): 3509-3519.
- Zhang, Y., Iratni, R., Erdjument-Bromage, H., Tempst, P., and Reinberg, D.  
1997 Histone deacetylases and SAP18, a novel polypeptide, are components of a human Sin3 complex. *Cell* 89(3): 357-364.
- Zhang, Y., LeRoy, G., Seelig, H. P., Lane, W. S., and Reinberg, D.  
1998 The dermatomyositis-specific autoantigen Mi2 is a component of a complex containing histone deacetylase and nucleosome remodeling activities. *Cell* 95(2): 279-289.
- Zhang, Y., Ng, H. H., Erdjument-Bromage, H., Tempst, P., Bird, A., and Reinberg, D.  
1999 Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. *Genes Dev.* 13(15): 1924-1935.
- Zhang, Y. and Reinberg, D.  
2001 Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes Dev.* 15(18): 2343-2360.
- Zhang, Y., Sun, Z. W., Iratni, R., Erdjument-Bromage, H., Tempst, P., Hampsey, M., and Reinberg, D.  
1998 SAP30, a novel protein conserved between human and yeast, is a component of a histone deacetylase complex. *Mol.Cell* 1(7): 1021-1031.
- Zhang, Y. W., Kim, K., Ma, Y. F., Wittner, M., Tanowitz, H. B., and Weiss, L. M.  
1999 Disruption of the *Toxoplasma gondii* bradyzoite-specific gene BAG1 decreases in vivo cyst formation. *Mol.Microbiol.* 31(2): 691-701.

# ABBREVIATIONS



## ABBREVIATIONS

ACF :	ATP-utilizing <b>C</b> hromatin assembly and remodelling <b>F</b> actor
ADA :	<b>A</b> Daptor <b>P</b> rotein
ADN :	<b>A</b> cide <b>D</b> ésoxyribo <b>N</b> ucléique
ADP:	<b>A</b> denosine <b>D</b> i <b>P</b> hosphate
ARN :	<b>A</b> cide <b>R</b> ibo <b>N</b> ucléique
ARNm :	<b>ARN</b> messenger
ATP :	<b>A</b> dénosine <b>T</b> ri <b>P</b> hosphate
ATPase :	Enzyme catalysant l'ATP
BAF :	<b>B</b> rg1 <b>A</b> ssociated <b>F</b> actor
BRCA1:	<b>B</b> Reast <b>C</b> Ancer 1
BRG1:	<b>B</b> rahma <b>R</b> elated <b>G</b> ene1
CAF1:	<b>C</b> hromatin <b>A</b> ssembly <b>F</b> actor 1
CARM1:	<b>C</b> oactivator- <b>A</b> ssociated arginine <b>M</b> ethyltransferase 1
CBP :	<b>C</b> REB <b>B</b> inding <b>P</b> rotein
CENP-A :	<b>CEN</b> tromeric <b>P</b> rotein <b>A</b>
ChIP :	<b>C</b> hromatin <b>I</b> mmuno <b>P</b> recipitation
CHRAC :	<b>C</b> Hromatin <b>A</b> ccessibility <b>C</b> omplex
CREB :	<b>C</b> yclic AMP <b>R</b> esponse <b>E</b> lement <b>B</b> inding protein
<b>Chromodomain</b> :	chromatin organization modifier
<b>C-terminal</b> :	<b>C</b> arboxy- <i>terminal</i>
DMSO :	<b>D</b> i <b>M</b> éthyl <b>S</b> ulf <b>O</b> xide
DNase :	<b>D</b> ésoxyribo <b>N</b> ucléase
DNMT1:	<b>D</b> N <b>A</b> <b>M</b> ethyl <b>T</b> ransferase 1
DO :	<b>D</b> ensité <b>O</b> ptique
Dot1:	<b>D</b> isruptor <b>of</b> telomeric silencing 1
E(Pc) :	<b>E</b> nhancer of <b>P</b> olycomb
Esa1:	<b>E</b> ssential <b>S</b> AS2-related <b>a</b> cetyltransferase 1
E(Z) :	<b>E</b> nhancer of <b>Z</b> este.
FAT :	<b>F</b> actor <b>A</b> cetyl <b>T</b> ransferase
Gcn5 :	<b>G</b> eneral control <b>n</b> onderepressible 5
GNAT :	<b>G</b> cn5 <b>N</b> <b>A</b> cetyl <b>T</b> ransferase related
GPS2 :	<b>G</b> <b>P</b> rotein pathway <b>S</b> uppressor 2

## Abbreviations

GST:	<i>Glutathione S Transferase</i>
HAT :	<i>Histone AcetylTransfarse</i>
HBO1 :	<i>Histone acetyltransferase Bound to ORC</i>
HDAC :	<i>Histone DéACétylase</i>
HIV :	<i>Human Immunodeficiency Virus,</i>
HKMT:	<i>Histone Lysine MethylTransferase,</i>
HLH	<i>Helix-Loop-Helix</i>
HMG :	<i>High Mobility Group</i>
HMT :	<i>Histone MethylTransferase</i>
HP1 :	<i>Heterochromatin Protein 1</i>
Ino8o :	<i>Inositol 8o</i>
ISWI :	<i>Imitation SWI/SNF</i>
KDa :	<i>KiloDalton</i>
LZ :	<i>Leucine Zipper</i>
MALDI-TOF:	<i>Matrix-Assisted Laser Desorption ionization Time Of Flight</i>
MBD :	<i>Methyl Binding Domain protein</i>
MDa :	<i>Méga Dalton</i>
MLL :	<i>Myeloid-Lymphoid (or mixed-lineage) Leukemia</i>
MOF :	<i>Male On the First</i>
MOZ :	<i>MOncocytic leukemia ZInc finger</i>
MRG15 :	<i>MORF Related Gene on chromosome 15</i>
MORF :	<i>MOncocytic leukemia zinc finger protein-Related Factor</i>
MSL :	<i>Male Specific Lethal</i>
MYST :	<i>MOZ, Ybf2/Sas3, Sas2 et TIP6o</i>
NAD :	<i>Nicotinamide Adenine Dinucleotide</i>
NES :	<i>Nuclear Export Signal</i>
N-CoR	<i>Nuclear hormone receptor Co-Repressor</i>
NF- $\kappa$ B :	<i>Nuclear Factor of kappa light polypeptide gene enhancer in B-cells</i>
NLS:	<i>Nuclearisation Localisation Signal</i>
N-terminal:	<i>amino-terminal</i>
NuA3 :	<i>Nucleosome Acetyltransferase of H3</i>
NuA4 :	<i>Nucleosome Acetyltransferase of H4</i>
NURF:	<i>NUcleosome Remodelling Factor</i>
NuRD :	<i>Nucleosome Remodeling histone Deacetylase complex</i>
ORF :	<i>Open Reading Frame</i>

## Abbreviations

pb :	<i>paire de bases</i>
PCAF :	<i>p300/CBP Associated Factor</i>
PCR :	<i>Polymerase Chain Reaction</i>
PEV :	<i>Position Effect Variegation</i>
PHD :	<i>Plant HomeoDomain</i>
PRMT :	<i>Proteine arginine MethylTransferase</i>
PML :	<i>ProMyelocytic Leukemia protein</i>
POZ :	Domaine caractéristique du <i>POxyvirus</i> et à Doigts de <b>Z</b> inc
RAR :	<i>Retinoic Acid Receptor</i>
RB :	<i>RetinoBlastoma</i>
RNase :	<i>RiboNucléase</i>
rpm :	<i>rotation par minutes</i>
Rpd3 :	<i>Reduced potassium dependency 3</i>
RSC :	<i>Remodel the Structure of Chromatin</i>
RT :	<i>Reverse Transcription</i>
SAGA :	<i>Spt, Ada, Gcn5 Acetyltransferase</i>
SAM :	<i>S Adenosyl Methionine</i>
SANT :	<i>Swi3, Ada2, N-Cor et TFIIIB</i>
Sas :	<i>Something about silencing</i>
SDS-PAGE	<i>SDS PolyAcrylamide Gel Electrophoresis</i>
SET :	<i>Su(var), Enhancer of zeste, Trithorax</i>
Sin3 :	<i>Swi-independent 3</i>
Sir :	<i>Silent information regulator protein</i>
SMRT :	<i>Silencing Mediator for Retinoic acid and Thyroid hormone receptors</i>
SNF :	<i>Sucrose Non Fermenting</i>
Spt :	<i>suppressor of Ty</i>
SRCAP :	<i>Snf2-Related CBP Activator Protein</i>
STAGA :	<i>SPT3-TAF(II)31- Ada2-GCN5L Acetylase</i>
SUV39H1 :	<i>SUppressor of Variegation 3-9 Homologue 1</i>
Su(var) :	<i>Suppressor of variegation</i>
SWI :	mating type <i>Sw</i> itching
SWI/SNF :	mating type <i>SW</i> itching/ <i>S</i> ucrose Non- <i>F</i> ermenting
TAF :	<i>TBP Associated Factor</i>
TAP :	<i>Tandem Affinity Purification</i>
TBP :	<i>TATA Box binding protein</i>

## Abbreviations

TF :	<i>T</i> ranscription <i>F</i> actor
TIP60 :	<i>T</i> at Interacting <i>P</i> rotein ( <i>60</i> kDa)
TSA :	<i>T</i> richo <i>S</i> tatin <i>A</i>
Tup1 :	deoxy- <i>T</i> hymidine monophosphate <i>u</i> ptake <i>p</i> ositive <i>1</i>
UV :	<i>U</i> ltra-violet
WT :	wild <i>T</i> ype

# The capsular polysaccharide biosynthesis of *Streptococcus pneumoniae* serotype 8: functional identification of the glycosyltransferase WciS (Cap8H)

Nehmé SAKSOUK\*, Ludovic PELOSI\*, Pierre COLIN-MOREL†, Manel BOUMEDIENNE\*, Patricia L. ABDIAN‡ and Roberto A. GEREMIA\*†<sup>1</sup>

\*Laboratoire Adaptation et Pathogénie des Micro-organismes, CNRS UMR 5163, Bâtiment Jean Roget, Faculté de Médecine Pharmacie, La Tronche, France, †Laboratoire d'Ecologie Alpine, CNRS UMR 5553, Université Joseph Fourier, 38041 Grenoble cedex 9, France, and ‡Instituto de Investigaciones Bioquímicas, Fundación Leloir, Buenos Aires, Argentina

CPS (capsular polysaccharide) is a major virulence factor in *Streptococcus pneumoniae*. Biosynthesis of CPS RU (repeat unit) proceeds by sequential transfer of sugar residues from the appropriate sugar donor to an activated lipid carrier by committed GTs (glycosyltransferases). While the nucleotide sequence of many *cps* loci is already known, the real substrate specificity of the hypothetical GTs, as well as the sequence of sugar addition is unclear. In the present paper, we report the biochemical characterization of one  $\alpha$ -galactosyltransferase, WciS (Cap8H), a member of GT family 4. This enzyme is implicated in the tetrasaccharide RU biosynthetic pathway of *Strep. pneumoniae* CPS 8 ( $(1 \rightarrow 4)\text{-}\alpha\text{-D-Glcp-(1 \rightarrow 4)\text{-}\alpha\text{-D-Galp-(1 \rightarrow 4)\text{-}\beta\text{-D-GlcAp-(1 \rightarrow 4)\text{-}\beta\text{-D-Glcp-(1 \rightarrow 1)}_n$ ). Expression of WciS–His<sub>6</sub> in *Escherichia coli* BL21 (DE3) strains or BL21 (DE3)/ $\Delta galU$  strain

resulted in synthesis of a 39 kDa membrane-associated protein identified by N-terminal sequencing and recognized by anti-His<sub>6</sub>-tag antibody. This protein was capable of adding a galactose residue cellobiuronic acid [ $\beta\text{-D-GlcAp-(1 \rightarrow 4)\text{-D-Glcp}$ ]-pyrophosphate-polyprenol from UDP-Gal. The newly added galactose residue is removed by  $\alpha$ -galactosidase, indicating that WciS is a retaining GT. Our results suggest that WciS catalyses the addition of the third sugar residue of the CPS 8 RU. The recombinant WciS–His<sub>6</sub> was solubilized and purified as a soluble multimer, opening the way for structural studies.

**Key words:** capsular polysaccharide, galactosyltransferase, glycosyltransferase, *Streptococcus pneumoniae*, virulence factor, WciS.

## INTRODUCTION

*Streptococcus pneumoniae* is an important bacterial pathogen causing invasive diseases such as pneumonia, otitis media and meningitis. Control of this Gram-positive pathogen is based on two completely different ways: the curative action of  $\beta$ -lactam agents, which is becoming less efficient because of the rise in antibiotic multi-resistant strains, and vaccination with a multi-valent vaccine based on CPS (capsular polysaccharide) vaccine.

The extracellular CPS of *Strep. pneumoniae* is the major virulence factor, because it prevents phagocytosis by the host immune system [1]. The CPS also induces T-cell-independent antibodies and this immunoprotection is serotype-specific [1], providing the basis for vaccine development. The current pneumococcal vaccine consists of purified CPS from 23 serotypes out of the 90 different known CPSs [2,3]. Although this vaccine protects healthy adults against pneumococcal infections, its efficiency decreases in high-risk human groups such as young children, elderly people or immunocompromised patients [1], and certain capsular types are poorly immunogenic [4]. A better understanding of CPS biosynthesis pathway may help to elaborate new anti-infectious agents, or even provide biotechnological tools for the synthesis of more efficient conjugated vaccines [5–7].

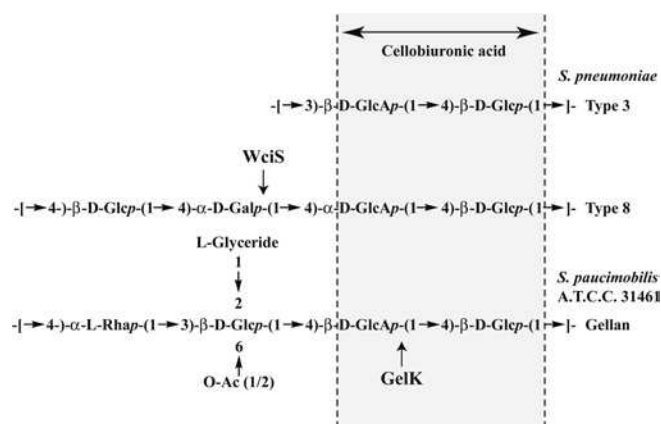
The first committed step for CPS biosynthesis is the activation of monosaccharides by their conversion into nucleotide derivatives. With the exception of two serotypes (3 and 37) that make use of a processive biosynthetic system [8,9], it has been proposed that a membrane-associated GT (glycosyltransferase) complex catalyses the assembly of the lipid-linked RU (repeating

unit) by successive transfer of monosaccharides. This first step is followed by polymerization of the RU, and subsequent export and attachment of the complete CPS to the cell surface [10]. Despite the large number of *cps* loci and chemical structures of CPS available [11], only the biosynthetic pathway of Cps14 has been reported [12–14].

We have studied the serotype 8 CPS, which is included in the 23 CPS-valent vaccine [2]. The chemical structure of the RU of this serotype [15] includes a cellobiuronic acid [ $\beta\text{-D-GlcAp-(1 \rightarrow 4)\text{-D-Glcp}$ ] that is also present in pneumococcal serotype 3 [8] and other bacterial exopolysaccharides, such as gellan (Figure 1) [16,17]. The *cps8* locus was initially identified by Muñoz et al. [18] in a DNA fragment between the *dexB* and *aliA* loci [19], and was capable of transforming the non-capsulated *Strep. pneumoniae* M24 to a capsulated serotype 8 transformant [18]. Analysis of the DNA sequence of the serotype 8 CPS gene cluster [11,18] revealed the presence of 12 hypothetical ORFs (open reading frames). In the present paper, we adopt the nomenclature of Jiang et al. [11] who later renamed the genes according to the Bacterial Polysaccharide Gene Database (BPGD) nomenclature system [20] (see <http://www.microbio.usyd.edu.au/BPGD/default.htm>). The biosynthesis of the cellobiuronic moiety differs with the species. *Strep. pneumoniae* serotype 3 CPS uses a single processive GT [9] (Figure 1). In *Sphingomonas paucimobilis* or *Rhizobium leguminosarum*, the synthesis of cellobiuronic-glycolipid requires two enzymatic activities [21–23], namely a phosphoglucosyltransferase (SpsB, PssA) and a glucuronosyltransferase (GelK, SpsK, PssDE). In the case of CPS 8, which is similar to that of *Sphing. paucimobilis*

Abbreviations used: CPS, capsular polysaccharide; GT, glycosyltransferase; IMAC, immobilized metal-affinity chromatography; IPTG, isopropyl  $\beta$ -D-thiogalactoside; LB, Luria–Bertani; LPS, lipopolysaccharide; ORF, open reading frame; PPL, pyrophosphate-polyprenol; RU, repeating unit; SEC, size-exclusion chromatography.

<sup>1</sup> To whom correspondence should be addressed (email roberto.geremia@ujf-grenoble.fr).



**Figure 1** Structure of the RU of the serotype 8 pneumococcal CPS

The corresponding hypothetical RU of serotype 3 CPS and gellan from *Sphing. paucimobilis* A.T.C.C. 31461 are shown for comparison. The glucuronosyltransferase GelK using  $\alpha$ -D-GlcP-PPL as the expected acceptor is indicated by an arrow. Our results suggest that the  $\alpha$ -galactosyltransferase WciS catalyses the transfer (see the arrow) of a galactose residue to the cellobiuronic acid moiety common to the pneumococcal serotypes and gellan exopolysaccharides.

or *R. leguminosarum*, five ORFs deduced from the *cps8* locus sequences code for hypothetical proteins that share sequence similarity with enzymes involved in RU assembly of Gram-negative bacterial CPS [11, 18]. These proteins are WchA (Cap8E, a glucose-1-phosphate transferase [24]), WciQR {Cap8FG, a putative inverting heterodimeric GT from the Carbohydrate Active enZYme's (CAZY) family 1 (<http://afmb.cnrs-mrs.fr/CAZY/index.html>) [25]}, WciS and WciT (Cap8H and Cap8J, two putative retaining GTs classified in CAZY's family 4 and 32 respectively). Since there are only two sugars with  $\alpha$ -linkages in the serotype 8 CPS RU [15], the putative GTs, WciS and WciT, are predicted to catalyse the addition of the  $\alpha$ -D-GlcP or the  $\alpha$ -D-Galp moieties.

In the present paper, we report the molecular and biochemical characterization of one of the GTs, WciS (Cap8H). Our results show that the purified recombinant GT (WciS-His<sub>6</sub>) is an  $\alpha$ -galactosyltransferase. This work also suggests that WciS adds the third sugar residue of the RU deduced from the serotype 8 CPS structure.

## MATERIALS AND METHODS

### Materials, culture conditions and bacterial strains

*Escherichia coli* strains BL21 (DE3) from Novagen were grown in LB (Luria-Bertani) broth at 37 °C with orbital agitation. The *E. coli* BL21 (DE3)/ $\Delta galU$  strain was constructed by generalized P1 transduction [26] using *E. coli* MC4100 *galU95::Tn10* (Tet<sup>R</sup>) as the donor [27]. This strain was kindly provided by Dr R. Hengge-Aronis and Dr E. Klauk (Institut für Biologie, Pflanzenphysiologie/Mikrobiologie, Berlin, Germany). *Sphing. paucimobilis* A.T.C.C. 31461 (American Type Culture Collection), kindly provided by Professor I. Correia (Instituto Superior Técnico, Lisboa, Portugal), was grown at 30 °C in S medium as described previously [28]. When required, antibiotics were used at the following concentrations: ampicillin (0.1 mg/l), chloramphenicol (0.03 mg/l) and kanamycin (0.1 mg/l). *Strep. pneumoniae* serotype 8 chromosomal DNA was from the A.T.C.C. strain. The pET29b and pET23a<sup>+</sup> expression vectors were from Novagen. The pET29b expression vector carrying the *gelK* gene

(pET-*gelK*) was a gift from Professor I. Correia. Oligonucleotide primers for PCR were purchased from Eurogentec (Seraing, Belgium). UDP-[U-<sup>14</sup>C]Gal (267 mCi/mmol) and UDP-[U-<sup>14</sup>C]-Glc (303 mCi/mmol) were from Amersham Biosciences and UDP-[U-<sup>14</sup>C]GlcA (313 mCi/mmol) was from ICI Biochemicals.

### Cloning of the *wciS* gene

The *wciS* gene was amplified from *Strep. pneumoniae* type 8 chromosomal DNA by PCR using the Expand High Fidelity PCR system (Roche) with the synthetic oligonucleotides *wciS*-for (5'-GGAAATTCATATGACAAAAGTTTTGCAAAATTGG-3') as a 5' primer (the engineered NdeI site is underlined) and *wciS*-rev (5'-TAGCTCGAGTTTTTTTCGTTTCTAATAGAGTATT-3') as a 3' primer (the engineered XhoI site is underlined). The PCR was optimized and carried out in 100  $\mu$ l of a mixture containing 100 ng of *Strep. pneumoniae* type 8 chromosomal DNA, 200  $\mu$ M each dNTP (Promega), 1  $\mu$ M each primer and 5 units Taq DNA polymerase according to the recommendations of the manufacturer (Roche). The predicted 1064 bp PCR product obtained was cloned into the NdeI/XhoI-digested pET23a<sup>+</sup> expression vector (Novagen) and sequenced (Genome Express, Meylan, France). The new plasmid was called pET-*wciS*.

### Expression and purification of recombinant protein WciS-His<sub>6</sub>

*E. coli* BL21(DE3) (Novagen) and BL21 (DE3)/ $\Delta galU$  cells harbouring the plasmid pET-*wciS* or the cloning vector pET23a<sup>+</sup> were grown at 37 °C in 500 ml of LB. At a  $D_{600}$  of 0.6, protein expression was induced by addition of IPTG (isopropyl  $\beta$ -D-thiogalactoside) to 0.4 mM. The bacterial culture was incubated at 37 °C for an additional 2 h. Cells were harvested by centrifugation (5000 g), washed once with 70 mM Tris/HCl, pH 8.2, resuspended in 70 mM Tris/HCl, pH 8.2, 300 mM NaCl, 10 mM EDTA and 1% Tween 20 at 40  $D_{600}$  equivalents/ml. For sub-cellular localization of the recombinant protein in *E. coli*, Tween 20 in the extraction buffer was omitted. The bacterial cells were broken by two passages through a French press (SLM-Amico Instruments, Thermo Spectronic, Rochester, NY, U.S.A.) at 80 MPa and submitted to continuous stirring for 30 min at 4 °C. This last step was performed only with extraction buffer containing detergent. Cell debris and inclusion bodies were then pelleted by centrifugation at 12000 g (12K) for 10 min at 4 °C (Sigma 6K15, rotor Nr 12166) and the supernatant was submitted to a further centrifugation at 50000 g (50K) (Beckman Avanti J-30I, rotor JA-30.50) for 1 h at 4 °C. The recovered insoluble and soluble fractions were stored at -20 °C or were used directly for the GT assay.

For WciS-His<sub>6</sub> purification, the extraction buffer was supplemented with 15 mM imidazole. The 50K supernatant fractionated through IMAC (immobilized metal-affinity chromatography) were prepared in 70 mM Tris/HCl, pH 8.2, 300 mM NaCl and 1% Tween 20. The sample was loaded on to a nickel-nitrilotriacetic acid Superflow agarose column (Amersham Biosciences) pre-equilibrated in the same buffer. After a washing step with 70 mM Tris/HCl, pH 8.2, 300 mM NaCl, 0.05% Tween 20 and 50 mM imidazole, the elution was carried out with an imidazole step gradient (50–500 mM) in the same buffer. The fractions containing the recombinant protein were pooled, desalted (Sephadex G-25M; Amersham Biosciences), concentrated and subjected to SEC (size-exclusion chromatography).

### SEC

Samples of 500  $\mu$ l containing 200  $\mu$ g of IMAC-purified recombinant protein were loaded on to a Superdex 200 HR 10/30 column



(Amersham Biosciences) equilibrated with 70 mM Tris/HCl, pH 8.2, 0.2 mM EDTA, 500 mM NaCl, 10 % glycerol and 0.05 % Tween 20. The chromatographic run was carried out in the same buffer at a flow rate of 0.4 ml/min at 4 °C using the FPLC BioLogic (Bio-Rad). Fractions of 2.5 ml were collected. Fractions containing WciS–His<sub>6</sub> were pooled and concentrated by centrifugation at 5000 g using a Centricon YM-10 filter (Amicon Bioseparations) according to the manufacturer's recommendations. Protein samples (10 µl) containing 0.25–12 mg/ml of WciS–His<sub>6</sub> were analysed directly by native or denaturing PAGE, or were used for the GT assay. The column was calibrated by running a set of gel-filtration markers (Amersham Biosciences).

### Protein quantification, SDS/PAGE, Western blot and amino acid sequencing

The protein concentrations were determined using a Bradford protein assay kit (Bio-Rad) and BSA as a standard. The purification procedure of WciS–His<sub>6</sub> was monitored by SDS/PAGE [12.5 % acrylamide/bisacrylamide (29:1) staining with Coomassie Brilliant Blue R250]. Pre-stained protein standards (Amersham Biosciences) were used for molecular mass estimation. Proteins were electrotransferred from SDS/PAGE gels on to PVDF membranes (Amersham Biosciences) as described previously [29]. Immunoblot analysis was performed with the Promega His<sub>6</sub>-tag Western-blotting kit, using alkaline phosphatase His<sub>6</sub>-protein conjugate, according to the manufacturer's instructions. The blots were developed with NBT (Nitro Blue Tetrazolium)/BCIP (5-bromo-4-chloroindol-3-yl phosphate) (Sigma–Aldrich) systems. N-terminal sequencing was performed with 20 µg of purified recombinant protein as described in [29].

### Electrophoresis of proteins under non-denaturing conditions

Protein samples containing WciS–His<sub>6</sub> were analysed in native PAGE using a gradient gel [4–20 % acrylamide/bisacrylamide (29:1)] in TBE buffer (90 mM Tris/borate, pH 8.4, and 2 mM EDTA). Electrophoresis was at 2 mA/gel using a Bio-Rad Mini-Protean II system at 4 °C, and the gel was stained with Coomassie Brilliant Blue R250. For localization of radiolabelled compounds, the gels were incubated directly for 30 min in ethanol/ethanoic (acetic) acid/water (5:1:4, by vol.) mixture, dried under vacuum and analysed with a phosphorimager (Fujifilm FLA-8000) after 72 h of exposure. The C-terminal His<sub>6</sub>-tag epitope fusion was detected in the native gels as described above.

### Preparation of *Sphing. paucimobilis* and *E. coli* glycolipid acceptors

Cells of *Sphing. paucimobilis* were prepared as described previously [22] from 100 ml cultures of *E. coli* BL21 (DE3) cells carrying either pET-*gelK*, pET23a<sup>+</sup> or pET29b. EDTA-permeabilized bacterial cells were obtained using the method described by Tomalsky et al. [30], modified slightly as described in [22]. Radiolabelled and non-radiolabelled glycolipid acceptors were synthesized *in vitro* by incubating 150 µl of permeabilized cells in the presence of the relevant sugar donor (2 mM UDP-Glc, 2 mM UDP-GlcA, 0.25 µCi of UDP-[U-<sup>14</sup>C]Glc or 0.25 µCi of UDP-[U-<sup>14</sup>C]GlcA), as described previously [22]. The glycolipids were then recovered as described previously [31]. Permeabilized cell suspensions of BL21 (DE3) carrying pET23a<sup>+</sup> (150 µl) were used directly to prepare endogenous glycolipid acceptors. Extracts were finally dried in a SpeedVac to a final volume of 10 µl and were stored at –80 °C until use.

### Galactosyltransferase assay

Galactosyltransferase assays were performed in 100 µl volumes containing 10 µl of the glycolipid acceptor, 70 mM Tris/HCl, pH 8.2, 1 % Triton X-100, 12.5 mM MgCl<sub>2</sub> and 0.10 µCi of UDP-[U-<sup>14</sup>C]Gal (267 mCi/mmol) (or 2 mM UDP-Gal), the reaction being started by addition of WciS–His<sub>6</sub> (approx. 30–300 µg of total protein). Assays were performed in triplicate. After 1 h at 37 °C, the reaction was stopped by the addition of 100 µl of water and 200 µl of chloroform/methanol (1:1, v/v) or analysed directly in native PAGE. Glycolipids were isolated (see below), and the radioactivity was determined by liquid-scintillation counting. The oligosaccharide moiety was released as described previously [22,31] and submitted to analytical TLC [31]. Glucose, galactose, glucuronic acid and oligomaltose (dimer to pentamer) were used as standards. Unlabelled sugar standards on the TLC plate were revealed by charring (100 °C for 10 min) after treatment with 5 % H<sub>2</sub>SO<sub>4</sub> in ethanol. The remaining portion of the plate containing the radiolabelled oligosaccharides was analysed with a phosphorimager Fujifilm FLA-8000.

### Galactosidase digestion

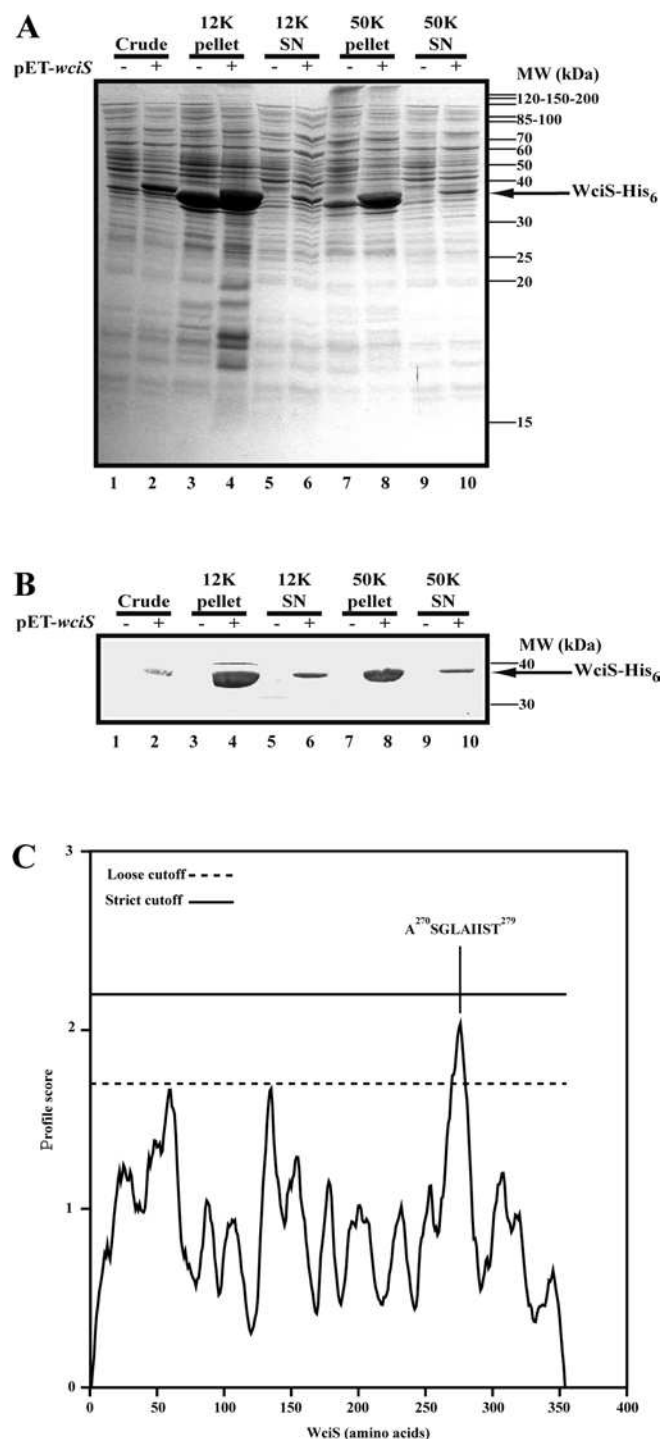
To determine the linkage configuration ( $\alpha$  or  $\beta$ ) of the galactose residue transferred *in vitro*, the oligosaccharides co-migrating with WciS–His<sub>6</sub> in the gel or released directly after mild acid hydrolysis were incubated overnight either in 10 mM Mops, pH 6.5, containing 0.6 units of  $\alpha$ -galactosidase from green coffee beans (G8507, Sigma–Aldrich), or in 100 mM Tris/HCl, pH 7.3, containing 11 units of  $\beta$ -galactosidase (G3153, Sigma–Aldrich) from *E. coli* at 25 and 37 °C respectively. Controls were performed with oligosaccharides incubated in the respective buffers without glycosylhydrolases.

## RESULTS

### WciS–His<sub>6</sub> is expressed mainly as a membrane-associated protein in *E. coli*

The *Strep. pneumoniae wciS* gene was cloned into the expression vector pET23a<sup>+</sup> by PCR amplification based on available nucleotide sequences [10,18]. The resulting recombinant protein, WciS–His<sub>6</sub>, is predicted to have 354 amino acids with a C-terminal His<sub>6</sub>-tag. The production of WciS–His<sub>6</sub> by *E. coli* BL21 (DE3) harbouring pET-*wciS* was confirmed by SDS/PAGE and His<sub>6</sub>-tag Western blotting. A 39 kDa protein was detected after 1 h of IPTG-induction (results not shown). The amount of protein increased for up to 3 h after induction and remained constant after overnight culture (results not shown). In all subsequent experiments, we prepared protein samples from cultures following 2 h of induction. The 39 kDa protein was identified as WciS–His<sub>6</sub> by Western blotting and immunodetection using an antibody against the His<sub>6</sub>-tag epitope (see below) and N-terminal amino acid sequencing.

To investigate the subcellular localization of WciS–His<sub>6</sub> in *E. coli*, crude French press extracts were prepared and submitted to differential centrifugation using the strain harbouring pET23a<sup>+</sup> without the *wciS* insert as a control. The protein content of each fraction was analysed by SDS/PAGE (Figure 2A) and Western blot (Figure 2B). WciS–His<sub>6</sub> polypeptide was the only significant protein expressed and recognized by anti-His<sub>6</sub>-tag antibody. A significant amount of the recombinant protein was recovered in the 12K pellet, suggesting that a portion of WciS–His<sub>6</sub> formed inclusion bodies (Figures 2A and 2B, lanes 4). Recovery of WciS–His<sub>6</sub> in the 50K pellet was almost quantitative (Figures 2A and 2B, lanes 8). Resuspension of the 50K pellet and centrifugation at



**Figure 2** Membrane localization of WciS-His<sub>6</sub>

(A) and (B) *E. coli* BL21 (DE3) strain harbouring pET23a<sup>+</sup> with (+) or without (-) the insert was grown, induced for 2 h and disrupted. The crude fractions obtained (lanes 1 and 2) were subjected to sequential 12 000 g (12K; lanes 3–6) and 50 000 g (50K; lanes 7–10) centrifugation. Electrophoretic profiles of the pellets and the corresponding supernatants (SN) are presented in lanes 3, 4, 7 and 8, and lanes 5, 6, 9 and 10 respectively. (A) Analysis of each aliquot by Coomassie Blue-staining SDS/12.5% PAGE. Approx. 20 µg of proteins were loaded in each lane. Pre-stained protein standards were used for molecular mass (MW) estimation. (B) His<sub>6</sub>-tag Western blotting with alkaline phosphatase His<sub>6</sub>-protein conjugate of each aliquot. The recombinant polypeptide, WciS-His<sub>6</sub>, expressed in the present is indicated by an arrow. (C) Transmembrane segment prediction profile obtained from WciS sequence using the DAS (dense alignment surface) method. The sequence of the main peptide which could interact with the membrane is indicated.

**Table 1** Incorporation of [U-<sup>14</sup>C]sugar mediated by WciS-His<sub>6</sub>

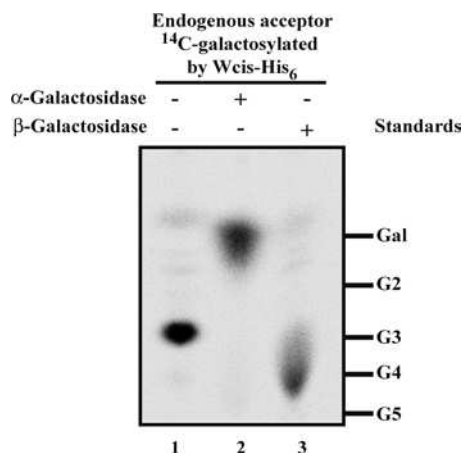
Activity corresponding to 150 µl of EDTA-permeabilized cells of *E. coli* BL21 (DE3)/pET-wciS or BL21 (DE3)/pET23a<sup>+</sup> was expressed as pmol of [U-<sup>14</sup>C]sugar (Glc, GlcA or Gal) incorporated per hour from corresponding UDP-[U-<sup>14</sup>C]sugar onto the *E. coli* BL21 (DE3) endogenous glycolipids.

Plasmid expressed in	Sugar donor			[ <sup>14</sup> C]sugar incorporated into the <i>E. coli</i> glycolipid fraction (pmol/h)
	BL21 (DE3)	UDP-[U- <sup>14</sup> C]Glc	UDP-[U- <sup>14</sup> C]GlcA	
pET23a <sup>+</sup>	+	-	-	< 3
pET-wciS	+	-	-	< 3
pET23a <sup>+</sup>	-	+	-	< 3
pET-wciS	-	+	-	< 3
pET23a <sup>+</sup>	-	-	+	< 3
pET-wciS	-	-	+	458 ± 76

50 000 g did not result in release of WciS-His<sub>6</sub> (results not shown). Furthermore, when membranes were purified by high-speed centrifugation on a 30% (w/v) glycerol cushion, the recombinant protein remained associated to the membrane (L. Pelosi, unpublished work). These data suggest that the recombinant WciS is associated with *E. coli* membrane (Figures 2A and 2B, lanes 8). The presence of a hypothetical transmembrane helix was not predicted by only the DAS (dense alignment surface) method [32] at strict cut-off, a peptide (A<sup>270</sup>SGLAIIST<sup>279</sup>) within the loose cut-off was detected (Figure 2C). Thus it is possible that the association of WciS-His<sub>6</sub> with the membrane is not related to the presence of a clear transmembrane domain, as it is also known for MurG [33] or AceA [34]. Interestingly, a fraction of recombinant WciS-His<sub>6</sub> was also present in the 50K supernatant (Figures 2A and 2B, lanes 10), even after 3 h of centrifugation (results not shown). However, the 'soluble' protein precipitated after 12 h at 4°C. The 12K and 50K supernatants were frozen at -20°C, to preclude precipitation.

#### WciS-His<sub>6</sub> displays α-galactosyltransferase activity *in vitro* on to an endogenous acceptor

In order to understand the function of WciS-His<sub>6</sub> in the biosynthetic pathway of the CPS 8 RU and to rule out the possibility of the presence of a hypothetical endogenous acceptor in *E. coli*, we incubated EDTA-permeabilized BL21 (DE3) cells expressing WciS-His<sub>6</sub> with sugar precursors (UDP-[U-<sup>14</sup>C]Gal, UDP-[U-<sup>14</sup>C]Glc or UDP-[U-<sup>14</sup>C]GlcA) under conditions described previously for other GTs [22,31,34]. The radioactivity incorporated into the glycolipid fraction was measured as described in the Materials and methods section. We did not detect any significant incorporation into the glycolipid fraction when *E. coli* BL21 (DE3) cells harbouring pET23a<sup>+</sup> extracts were incubated with each radiolabelled UDP-sugar, showing the absence of active endogenous GTs in the assay (Table 1). Incubation in the absence of added acceptor resulted in the incorporation of 458 pmol/h of [U-<sup>14</sup>C]Gal moiety into the organic phase (Table 1). This result indicates the presence of an endogenous acceptor. In previous functional characterization of recombinant GT [22,31,34], we have not detected any endogenous acceptor. Neither [U-<sup>14</sup>C]Glc nor [U-<sup>14</sup>C]GlcA residue was incorporated into the glycolipid fraction in a WciS-His<sub>6</sub>-dependent way (Table 1), suggesting that WciS is a galactosyltransferase. This galactosyltransferase activity was also found both in the 12K and 50K supernatants (results



**Figure 3** Characterization of the <sup>14</sup>C-galactosylated endogenous acceptor in *E. coli* BL21 (DE3)

Phosphoimaging of <sup>14</sup>C-galactosylated oligosaccharides moiety. EDTA-permeabilized *E. coli* BL21 (DE3) cells expressing WciS-His<sub>6</sub> were incubated in the presence of UDP-[U-<sup>14</sup>C]Gal. The radiolabelled acceptor was incubated overnight without galactosylhydrolases (lane 1), in the presence of  $\alpha$ -galactosidase (lane 2) or  $\beta$ -galactosidase (lane 3). Galactose (Gal) and oligomaltoses (G2, maltobiose; G3, maltotriose; G4, maltotetraose; G5, maltopentaose) were used as standards.

not shown), suggesting that the endogenous acceptor co-purifies with the recombinant GT during differential centrifugation.

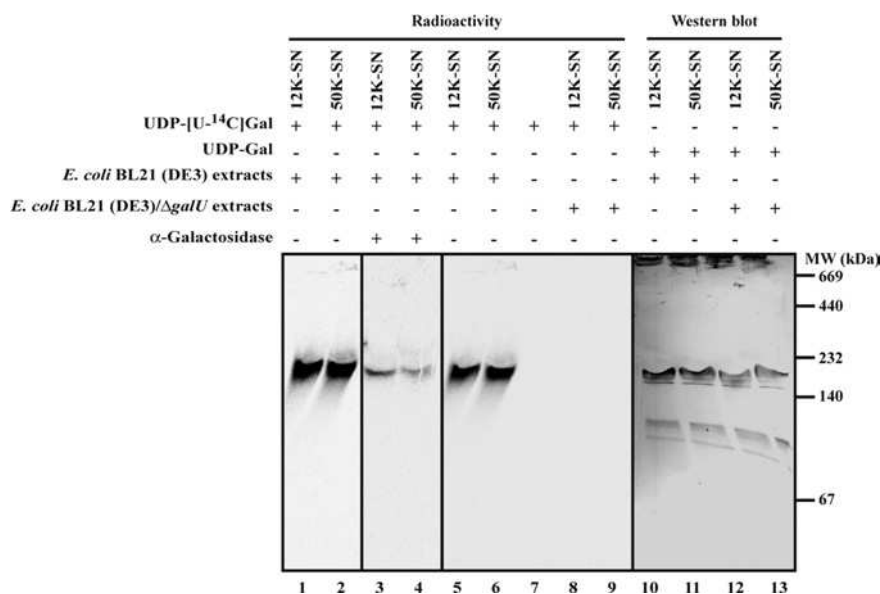
The <sup>14</sup>C-galactosylated moiety of the *E. coli* endogenous acceptor was released from the organic phase by mild acid hydrolysis followed by treatment with bovine alkaline phosphatase, suggesting a pyrophosphate linkage to the lipid moiety (results not shown). Upon delipidation and dephosphorylation, a significant amount of the radiolabelled galactoside was bound to the

ion-exchange resin (results not shown), suggesting the presence of a charged sugar. The size of the [U-<sup>14</sup>C]Gal residue-containing moiety was assessed by TLC (Figure 3, lane 1). A major radioactive spot was detected by phosphoimaging at the same position as the maltotriose marker. The glycosyl moiety of the endogenous acceptor was therefore probably a disaccharide, compatible with the putative disaccharide moiety of the native acceptor of WciS: cellobiuronic acid-PPL (pyrophosphate-polyprenol). The linkage of the added [U-<sup>14</sup>C]Gal was assessed by digestion with  $\alpha$ - and  $\beta$ -galactosidases. The  $\alpha$ -galactosidase treatment resulted in the release of Gal (Figure 3, lane 2). Thus WciS-His<sub>6</sub> is an  $\alpha$ -galactosyltransferase. The streaky migration of the spot on the silica gel corresponding to radiolabelled trisaccharide treated by the  $\beta$ -galactosidase (Figure 3, lane 3) was observed only with this hydrolase, and probably due to the commercial formulation of this enzyme.

The interactions between WciS-His<sub>6</sub> and its glycolipid reaction product were investigated using native PAGE analysis of the 12K and 50K supernatants containing WciS-His<sub>6</sub>. Incubation in the presence of radiolabelled UDP-Gal resulted in the labelling of a protein with an apparent molecular mass of approx. 180 kDa (Figure 4, lanes 1–4), that co-migrated with WciS-His<sub>6</sub> as revealed by Western blotting (Figure 4, lanes 10 and 11). Incubation of this gel in buffer containing  $\alpha$ -galactosidase (Figure 4, lanes 3 and 4) resulted in a significant reduction of the radioactive signals, indicating that the galactosylated glycolipid co-migrates with native WciS-His<sub>6</sub>. These results suggest that the recombinant WciS-His<sub>6</sub> behaves as a multimer, and that the reaction product is associated with this complex.

#### The WciS-His<sub>6</sub>-dependent galactosylation is abolished in the *E. coli* BL21 (DE3)/ $\Delta galU$ strain

The presence of an endogenous acceptor precluded the use of better characterized exogenous acceptors. To overcome this problem,



**Figure 4** Detection of <sup>14</sup>C-galactosylated endogenous acceptor co-migrating with WciS-His<sub>6</sub>

Recombinant proteins (12K and 50K supernatants) were prepared from *E. coli* BL21 (DE3) expressing WciS-His<sub>6</sub> (lanes 1–6, 10 and 11) or *E. coli* BL21 (DE3)/ $\Delta galU$  expressing WciS-His<sub>6</sub> (lanes 8, 9, 12 and 13) and incubated in the presence of radiolabelled (lanes 1–6, 8 and 9) or not (lanes 10–13) UDP-Gal. Native gradient PAGE (4–20% acrylamide/bisacrylamide in Tris/borate/EDTA) was performed for each sample. The migration of the <sup>14</sup>C-galactosylated acceptor and WciS-His<sub>6</sub> expressed in the two strains was monitored in the gel by phosphoimaging and His<sub>6</sub>-tag Western blotting respectively. A control was realized in the same assay conditions with only UDP-[U-<sup>14</sup>C]Gal added to the reaction mixture (lane 7). Identity of radiolabelled product (lanes 1 and 2) was confirmed by  $\alpha$ -galactosidase gel treatment (lanes 3 and 4). SN, supernatant; MW, molecular-mass markers (sizes given in kDa).

**Table 2** Incorporation of [ $^{14}\text{C}$ ]Gal mediated by WciS–His<sub>6</sub> into the GelK product

The 12K supernatants were prepared from *E. coli* BL21 (DE3)/ $\Delta galU$  cells expressing WciS–His<sub>6</sub> (pET-wciS) or not (control vector, pET23a<sup>+</sup>). Activity corresponding to 50  $\mu\text{l}$  of GT preparation (approx. 300  $\mu\text{g}$  of total proteins) was expressed as pmol of [ $^{14}\text{C}$ ]Glc incorporated per h from UDP-[ $^{14}\text{C}$ ]Glc on to the glycolipid fraction prepared from a *E. coli* BL21 (DE3) strain expressing GelK (pET-gelK) or harbouring the corresponding control vector (pET29b, see the Materials and methods section for further details).

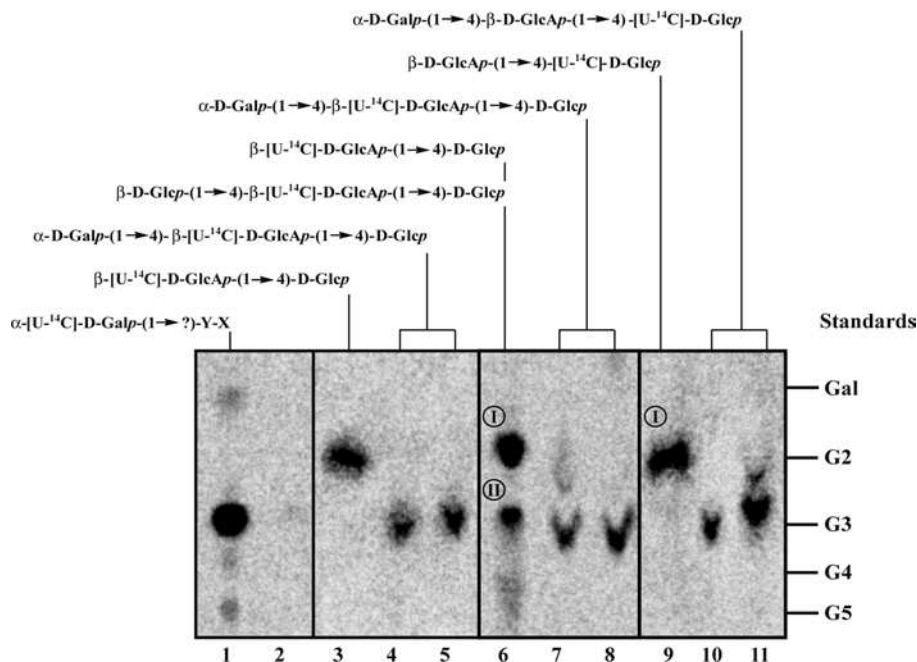
Plasmid expressed	pET29b glycolipid	pET-gelK glycolipid	UDP-[ $^{14}\text{C}$ ]Gal	[ $^{14}\text{C}$ ]Gal incorporated into the acceptor (pmol/h)
pET23a <sup>+</sup>	–	–	+	<3
pET-wciS	–	–	+	<3
pET23a <sup>+</sup>	+	–	+	<3
pET-wciS	+	–	+	154 $\pm$ 31
pET23a <sup>+</sup>	–	+	+	<3
pET-wciS	–	+	+	301 $\pm$ 54

we constructed the *E. coli* BL21 (DE3)/ $\Delta galU$  strain (see the Materials and methods section). The *galU* gene encodes the UTP: $\alpha$ -D-glucose-1-phosphate uridylyltransferase (EC 2.7.7.9) implicated directly in the UDP-Glc biosynthetic pathway [35]. If glucose, galactose or glucuronic acid (derived from UDP-Glc) are present in the endogenous acceptor, the latter will not be synthesized in the  $\Delta galU$  strain. Neither the 12K (Table 2) nor the 50K supernatant (results not shown) prepared from *E. coli* BL21 (DE3)/ $\Delta galU$  extracts expressing WciS–His<sub>6</sub> incorporated any significant radioactivity into the glycolipid fraction (Table 2, com-

pare with the 458 pmol/h in Table 1). Moreover, using the  $\Delta galU$  extracts, we did not detect any radioactive spot corresponding to the [ $^{14}\text{C}$ ]Galactosylated oligosaccharide (Figure 5, lanes 1 and 2), nor any radioactivity co-migrating with WciS–His<sub>6</sub> in native electrophoresis (Figure 4, lanes 8, 9, 12 and 13). However, the WciS–His<sub>6</sub> overexpressed in the  $\Delta galU$  strain was active (154 pmol/h of [ $^{14}\text{C}$ ]Gal) towards the lipid acceptor prepared from *E. coli* BL21 (DE3) harbouring only a control vector (Table 2). Taken together, these results indicate that the mutant strain lacks the endogenous glycolipid acceptor. In addition, these data point to the presence of glucose, galactose, glucuronic acid and/or galacturonic acid in the endogenous acceptor, as expected for the hypothetical physiological substrate of WciS: the cellobiuronic acid-PPL. The *E. coli* BL21 (DE3)/ $\Delta galU$  harbouring pET-wciS was therefore used for the rest of the present study.

### The cellobiuronic acid-PPL is an acceptor of WciS–His<sub>6</sub>

The expected *in vivo* acceptor of WciS is the cellobiuronic acid-PPL [18,24]. Since this product is not commercially available, we decided to synthesize it *in vitro* using two kinds of enzymatic extracts [22]. The first one involved the  $\beta$ -(1  $\rightarrow$  4)-D-glucuronosyltransferase, GelK, from *Sphing. paucimobilis* A.T.C.C. 31461 expressed in *E. coli* BL21 (DE3) strain harbouring pET-gelK [22]. The glycolipidic acceptors were partially purified (see the Materials and methods section) and added to the reaction mixture containing WciS–His<sub>6</sub> (12K supernatant) and UDP-[ $^{14}\text{C}$ ]Gal (Table 2). A control was performed in parallel with the glycolipid fraction prepared from a *E. coli* BL21 (DE3)

**Figure 5** Analytical TLC of oligosaccharides released after WciS–His<sub>6</sub>-mediated transfer of galactose residues on to endogenous and exogenous acceptors

Lanes 1 and 2: oligosaccharides released from the lipid fraction by mild acid hydrolysis after incubation of 12K supernatants prepared from *E. coli* BL21 (DE3) (lane 1) or BL21 (DE3)/ $\Delta galU$  (lane 2) expressing WciS–His<sub>6</sub> with UDP-[ $^{14}\text{C}$ ]Gal. Lanes 3–5: the glycolipid  $\beta$ -[ $^{14}\text{C}$ ]D-GlcAp-(1  $\rightarrow$  4)-D-GlcP-PPL was prepared by incubating *E. coli* BL21 (DE3) expressing GelK extracts with UDP-[ $^{14}\text{C}$ ]GlcA (see the Materials and methods section). Lanes 6–8 and lanes 9–11: the glycolipid  $\beta$ -[ $^{14}\text{C}$ ]D-GlcAp-(1  $\rightarrow$  4)-D-GlcP-PPL and  $\beta$ -D-GlcAp-(1  $\rightarrow$  4)-[ $^{14}\text{C}$ ]D-GlcP-PPL were prepared by incubating *Sphing. paucimobilis* A.T.C.C. 31461 extracts with the relevant UDP-sugar radiolabelled or not (see the Materials and methods section for further details). The oligosaccharides released by mild acid hydrolysis were analysed by TLC before (lanes 3, 6 and 9) or after incubation with WciS–His<sub>6</sub> prepared from *E. coli* BL21 (DE3) 12K supernatants (lanes 4, 7 and 10) or 50K supernatants (lanes 5, 8 and 11). Glucose and oligomaltoses (G2, maltobiose; G3, maltotriose; G4, maltotetraose; G5, maltopentaose) were used as standards. The predicted chemical structure of each oligosaccharide moieties synthesized *in vitro* is indicated. X and Y are unknown sugar residues.

strain harbouring the corresponding control vector (pET29b) to assess the contribution of the endogenous acceptor. Under these conditions, a higher incorporation of [ $^{14}\text{C}$ ]Gal residues (301 pmol/h) in the glycolipid fraction was found in the presence of exogenous GelK-product, WciS–His<sub>6</sub> and UDP-[U- $^{14}\text{C}$ ]Gal (Table 2). Similar results were obtained with the WciS–His<sub>6</sub> 50K supernatant (results not shown). The converse experiment, using non-labelled UDP-Gal and  $\beta$ -[U- $^{14}\text{C}$ ]D-GlcAp-(1  $\rightarrow$  4)-D-Glcp-PPL was also performed (Figure 5, lanes 3–5). The strong decrease in intensity of the radioactive spot, which was the expected  $\beta$ -[U- $^{14}\text{C}$ ]D-GlcAp-(1  $\rightarrow$  4)-D-Glcp-PPL product (Figure 5, lane 3), was accompanied by the appearance of a radio-labelled trisaccharide after incubation with UDP-Gal and WciS–His<sub>6</sub> 12K or WciS–His<sub>6</sub> 50K supernatant (Figure 5, lanes 4 and 5). Incubation with  $\alpha$ -galactosidase released the disaccharide (results not shown). These results, showing that the GelK product (cellobiuronic acid-PPL) is a substrate for WciS *in vitro* are concordant with the proposed biosynthetic RU.

These results were verified by a different approach using radiolabelled lipid-linked oligosaccharides from *Sphing. paucimobilis* A.T.C.C. 31461. To synthesize the appropriated acceptor, EDTA-permeabilized cells from *Sphing. paucimobilis* were incubated in the presence of UDP-Glc/UDP-[U- $^{14}\text{C}$ ]GlcA or UDP-[U- $^{14}\text{C}$ ]Glc/UDP-GlcA. The oligosaccharides synthesized *in vitro* by endogenous GTs were released from the glycolipid fraction and were analysed by TLC. Two radioactive spots (I and II) were detected by phosphoimaging after incubation with UDP-Glc and UDP-[U- $^{14}\text{C}$ ]GlcA. The major oligosaccharide formed (Spot I, Figure 5, lane 6) co-migrated with the maltobiose marker and was assigned to be  $\beta$ -[U- $^{14}\text{C}$ ]D-GlcAp-(1  $\rightarrow$  4)-D-Glcp, based on data of Videira et al. [22]. Spot II represents probably the trisaccharide  $\beta$ -D-Glcp-(1  $\rightarrow$  4)- $\beta$ -[U- $^{14}\text{C}$ ]D-GlcAp-(1  $\rightarrow$  4)-D-Glcp (Figure 5, lane 6), which might be generated by addition of another glucose residue to the previous radiolabelled disaccharide [22]. When permeabilized *Sphing. paucimobilis* cells were incubated with UDP-[U- $^{14}\text{C}$ ]Glc and UDP-GlcA, spot I was mainly obtained. After incubation of the radioactive glycolipids in the presence of WciS–His<sub>6</sub> from the 12K or 50K supernatant (Figure 5, lanes 7, 8, 10 and 11) and UDP-Gal, a new spot appeared that co-migrated with the maltotriose marker. The corresponding radiolabelled trisaccharide was probably the expected  $\alpha$ -D-Galp-(1  $\rightarrow$  4)- $\beta$ -[U- $^{14}\text{C}$ ]D-GlcAp-(1  $\rightarrow$  4)-D-Glcp (Figure 5, lanes 7 and 8) and  $\alpha$ -D-Galp-(1  $\rightarrow$  4)- $\beta$ -D-GlcAp-(1  $\rightarrow$  4)-[U- $^{14}\text{C}$ ]D-Glcp (Figure 5, lanes 10 and 11). The minor spots revealed in Figure 5, lanes 7 and 11 probably represented the unmodified disaccharide. A similar product {radiolabelled cellobiosyl ( $\beta$ -D-Glcp-(1  $\rightarrow$  4)-D-Glcp)-PPL prepared from *Xanthomonas campestris* [34]} was tested in parallel. No WciS–His<sub>6</sub>-dependent incorporation on to this acceptor was found, showing the important role of the GlcA (results not shown). These results support our working hypothesis that WciS catalyses the transfer *in vivo* of galactose residues from UDP-Gal into the cellobiuronic acid moiety of the *Strep. pneumoniae* CPS 8 RU.

### The SEC-purified WciS–His<sub>6</sub> presents an active multimeric structure

We have purified WciS–His<sub>6</sub> from the *E. coli* BL21 (DE3)/ $\Delta galU$  strain to exclude the possibilities that the galactosyltransferase activity depends on other endogenous proteins or small molecules present in the cell extracts. To avoid the precipitation of WciS–His<sub>6</sub> and to increase the solubility of the protein, *E. coli* protein extracts were prepared using 1% Tween 20. Under these conditions, most WciS–His<sub>6</sub> was found in the 50K supernatant (Figures 6A

and 6B, lanes 1–10). Additionally, this treatment prevented the precipitation of the recombinant protein (see above). The WciS–His<sub>6</sub> purification, performed in the presence of Tween 20, was monitored using SDS/PAGE and Western blot (results not shown). The 50K supernatant (Figures 6A and 6B, lane 10) was purified by IMAC on a Ni<sup>2+</sup>-chelate column, as described in the Materials and methods section. A 39 kDa protein that was eluted as a major peak at 50 mM imidazole was identified by Western blot analysis as WciS–His<sub>6</sub> (Figures 6A and 6B, lane 11). Both to purify further and to assess the size of WciS–His<sub>6</sub>, we carried out SEC experiments (Figure 6C). The major protein fraction eluted at 37.5 min (Figure 6D, N-terminus: TKVLQIGPS) with an apparent molecular mass of 130 kDa. After denaturation and SDS/PAGE of this sample, only the monomer of 39 kDa was present (Figures 6A and 6B, lane 12), suggesting that WciS is in multimeric.

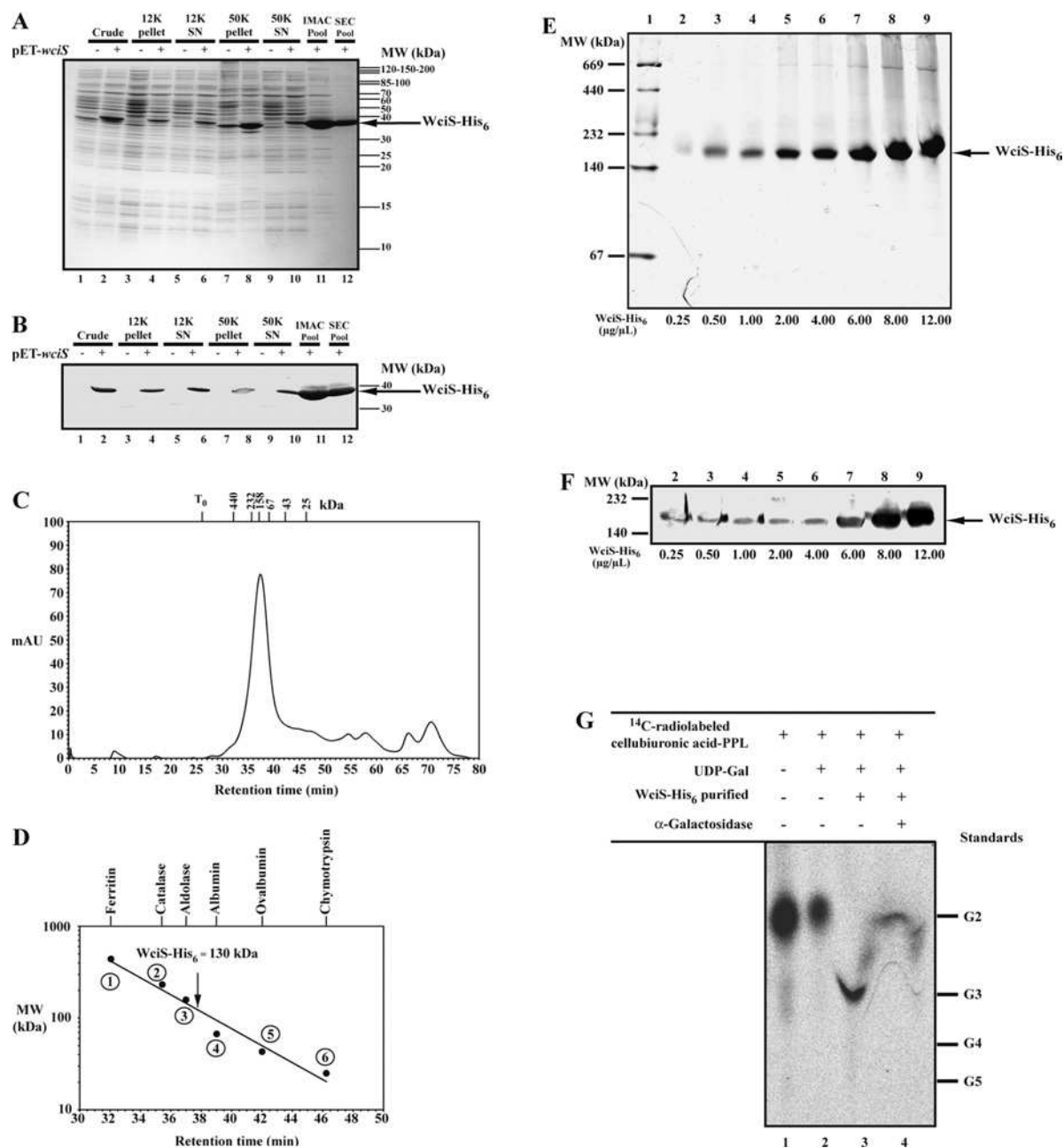
Since protein crystallization requires concentrated protein, we studied the behaviour of WciS–His<sub>6</sub> in concentrated solutions. Native PAGE was performed with concentrations ranging from 0.25 to 12  $\mu\text{g}/\mu\text{l}$  (Figures 6E and 6F). The apparent average molecular mass of WciS–His<sub>6</sub> (approx. 180 kDa) did not change with the concentrations tested. All these results indicate that the multimeric form of WciS–His<sub>6</sub> is stable through all the purification procedures (see results above and Figure 4). The discrepancies observed in the native apparent molecular mass of WciS–His<sub>6</sub> estimated by SEC and native PAGE (130 and 180 kDa respectively) rely on the fact that the two experimental approaches separate proteins on the basis of two different properties.

The  $\alpha$ -galactosyltransferase activity of the SEC-purified WciS–His<sub>6</sub> was assayed in the presence of radiolabelled glycolipid acceptor ( $\beta$ -D-GlcAp-(1  $\rightarrow$  4)-[U- $^{14}\text{C}$ ]D-Glcp-PPL) (Figure 6G, lane 1). The TLC analysis showed that the resulting radiolabelled galactosylated oligosaccharide co-migrated with the maltotriose marker (Figure 6G, lane 3) and was hydrolysed in the presence of  $\alpha$ -galactosidases (Figure 6G, lane 4), indicating that the purified galactosyltransferase is active. It is worthy of mention that in purifications performed using the wild-type *E. coli* strain, the endogenous acceptor co-purified all the way with WciS–His<sub>6</sub>.

## DISCUSSION

In the present study, we report on the functional characterization of WciS as an  $\alpha$ -galactosyltransferase that uses cellobiuronic acid-PPL as an acceptor. This function is in agreement with the sugar sequence of the CPS type 8 and previous proposals based on sequence analyses [11,18], as well as its classification in the CAZY family 4 of retaining GTs [25]. WciS is therefore the first of the three putative GTs, deduced from the *cps8* gene cluster sequence, to be characterized using a biochemical approach. The recombinant WciS co-precipitates on centrifugation with membranes; however, the nature of the association remains unclear. Solubilization of WciS was achieved in the presence of detergents (the present study, and L. Pelosi, unpublished work). We have also found that WciS behaves as a multimeric protein.

As proposed for other GTs that are involved in LPS (lipopolysaccharide) biosynthesis, such as AceA from *Acetobacter xylinum* [34], ExoM from *Sinorhizobium meliloti* [31] or GelK from *Sphing. paucimobilis* [22], features of PPL moiety are required for WciS–His<sub>6</sub> activity (results not shown). We have also found that WciS–His<sub>6</sub> galactosylates an endogenous glycolipid acceptor. The endogenous acceptor co-purifies with WciS. Moreover, the reaction product co-migrates in native electrophoresis with

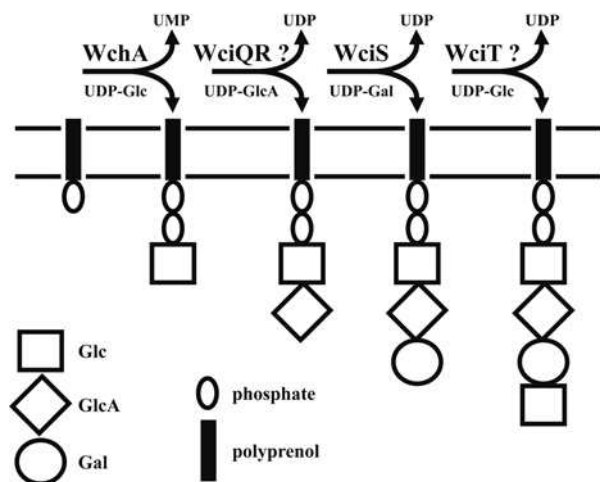
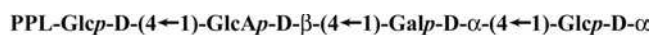


**Figure 6** Purification of active WciS-His<sub>6</sub> in the presence of Tween 20

(A) Coomassie-Blue-stained SDS/PAGE gel of the WciS-His<sub>6</sub> protein samples taken during the purification procedure. (B) His<sub>6</sub>-tag Western blotting of each aliquot. Cells of the BL21 (DE3)/Δ*galU* strain harbouring pET23a<sup>+</sup> with (+) or without (-) the insert were disrupted in the extraction buffer containing 1% Tween 20 and subjected to a differential centrifugation (see legend to Figure 1, lanes 1–10). The recombinant protein from the 50K supernatant was subjected to IMAC purification using a nickel-containing column (lane 11) followed by analytical FPLC gel filtration (SEC) (lane 12). The apparent molecular mass of the native WciS-His<sub>6</sub> was estimated by filtration on a Superose 6 column. (C) Retention profile of the protein purified from the chromatographic column (the position of the retention times of the protein markers used to calibrate the column is indicated by crossbars; T<sub>0</sub>, retention time of the void volume of the column, determined using Dextran Blue). (D) Calibration curve of the gel filtration column; the following protein standards (Amersham Biosciences) were used: (1) ferritin (440 kDa), (2) catalase (232 kDa), (3) aldolase (158 kDa), (4) albumin (67 kDa), (5) ovalbumin (43 kDa) and (6) chymotrypsin (25 kDa). (E) SEC-purified WciS-His<sub>6</sub> was analysed by native gradient PAGE (4–20% acrylamide/bisacrylamide in Tris/borate/EDTA). Lane 1, molecular-mass protein standards (Amersham Biosciences). Lanes 2–9, monitoring of the concentration of WciS-His<sub>6</sub> prepared by centrifugal filtration (from 0.25 to 12.00 μg/μl). The expected WciS-His<sub>6</sub> multimer is indicated by an arrow. Each sample (10 μl) was stained with Coomassie Blue. (F) His<sub>6</sub>-tag Western blotting of the corresponding samples in (E). (G) Analytical TLC of oligosaccharides released after WciS-His<sub>6</sub>-mediated transfer of galactose residue into radiolabelled cellobiuronic acid-PPL acceptor. The glycolipid β-D-GlcAp-(1 → 4)-[U-<sup>14</sup>C]D-Glcp-PPL was prepared as described in Figure 5 (lane 9) and incubated in the presence of UDP-Gal with (lane 3) or without (lane 2) SEC-purified WciS-His<sub>6</sub> (30 μg of total protein). The resulting trisaccharide was then incubated overnight in the presence of α-galactosidase (lane 4). Galactose and oligomaltoses (G2, maltobiose; G3, maltotriose; G4, maltotetraose; G5, maltopentaose) were used as standards. WciS-His<sub>6</sub> polypeptide was indicated by an arrow. SN, supernatant; MW, molecular-mass markers (sizes in kDa).

the multimeric protein. The strong interaction between WciS-His<sub>6</sub> and the reaction product can be understood either as a non-specific hydrophobic interaction or as a specific interaction with

the active site of the protein. Although our data cannot distinguish between these two hypotheses, we can speculate that the specific interaction leads to channelling of RU biosynthesis.



**Figure 7** Model of the assignment of biochemical function to GTs involved in RU assembly of *Strep. pneumoniae* CPS 8

Like the Streptococcal gene products, such as Cps14E, Cps9vE and Cps<sub>III</sub>E characterized previously [12,36,37], we propose that WchA is a UDP-Glc: PPL glucosyl-1-phosphate transferase [24]. WciS catalyses the transfer of galactose residue, from UDP-Gal, on to a cellobiuronic acid-PPL through an  $\alpha$ -linkage. The functions of WciQR and WciT are unknown. However, based on bioinformatic analysis and assuming a sequential expression of the presumed GT genes identified in the CPS 8 cluster, we propose that WciQR and WciT catalyse the addition of the second (GlcA) and fourth (glucose) residues of the CPS 8 RU.

Although a chemical characterization of the endogenous acceptor and the WciS products are beyond the purpose of the present paper, our data supports that the endogenous acceptor is a cellobiuronic glycolipid. Indeed, the properties of the product formed with the endogenous acceptor, as well as the absence of it in the  $\Delta galU$  strain, are compatible with a cellobiuronic acid. Unfortunately, there are no data available on extracellular polysaccharides produced by *E. coli* B strains, precluding any hypothesis on the origin of such a disaccharide. Further structural studies will be necessary to gain further information about the actual chemical structure of the endogenous acceptor. The *E. coli* BL21 (DE3)/ $\Delta galU$  strain may represent a new biological tool for both expression and biochemical characterization of recombinant GTs in decreasing the probability to have endogenous acceptors.

The present study, in addition to the CAZy classifications, allows us to provide a model for the biosynthesis of the RU of serotype 8 CPS (Figure 7). In a previous study, we have shown that WchA primes the biosynthesis of the serotype 8 CPS by transfer of a phosphoglucose residue to a lipid [24]. Given the presence of the cellobiuronic acid in CPS type 8 and the similarity of WciQR with SpsK, GelK and PssDE, it is possible that the second step is the transfer of a glucuronic acid residue catalysed by WciQR. Thus this is the acceptor of WciS that transfers the galactose residue to form the trisaccharide. Finally, WciT may add the remaining glucose. The synthesis *in vitro* of the complete RU may be useful for the study of translocation and the polymerization events.

We thank R. Hengge-Aronis and E. Klauck for providing *E. coli* strains, D. Schneider for help with P1 transduction, P. Videira for her critical reading of the manuscript, and N. Potier and E. Coursange for technical assistance. The N-terminal amino acid sequence of recombinant protein was determined by J.-P. Andrieu, Institut de Biologie Structurale (IBS), Grenoble, France. L.P. and N.S. are supported by a post-doctoral (CNRS) and a doctoral (CNRS) fellowship respectively. This work was supported by CNRS (France) grants.

## REFERENCES

- Alonso de Velasco, E., Verheul, A. F. M., Verhoef, J. and Snippe, H. (1995) *Streptococcus pneumoniae*: virulence factors, pathogenesis, and vaccines. *Microbiol. Rev.* **59**, 591–603
- Robbins, J. B., Austrian, R., Lee, C. J., Rastogi, S. C., Schiffman, G., Henrichsen, J., Mäkelä, P. H., Broome, C. V., Fracklam, R. R., Tiesjema, R. H. and Parke, J. C. J. (1983) Considerations for formulating the second-generation pneumococcal capsular polysaccharide vaccine with emphasis on the cross-reactive types within groups. *J. Infect. Dis.* **148**, 1136–1159
- Henrichsen, J. (1995) Six newly recognized types of *Streptococcus pneumoniae*. *J. Clin. Microbiol.* **33**, 2759–2762
- Butler, J. C., Breiman, R. F., Campbell, J. F., Lipman, H. B., Broome, C. V. and Fracklam, R. R. (1993) Pneumococcal polysaccharide vaccine efficacy: an evaluation of current recommendations. *JAMA, J. Am. Med. Assoc.* **270**, 1826–1831
- Makela, P. H. and Kayhty, H. (2002) Evolution of conjugate vaccines. *Expert Rev. Vaccines* **1**, 399–410
- Alexander, J., Guercio, M. F., Frame, B., Maewal, A., Sette, A., Nahm, M. H. and Newman, M. J. (2004) Development of experimental carbohydrate-conjugate vaccine composed of *Streptococcus pneumoniae* capsular polysaccharides and the universal helper T-lymphocyte epitope (PADRE®). *Vaccine* **22**, 2362–2367
- Lee, C. J., Lee, L. H. and Frasch, C. E. (2003) Protective immunity of pneumococcal glycoconjugates. *Crit. Rev. Microbiol.* **29**, 333–349
- Cartee, R. T., Forsee, W. T., Jensen, J. W. and Yother, J. (2001) Expression of the *Streptococcus pneumoniae* type 3 synthase in *Escherichia coli*: assembly of the type 3 polysaccharide on a lipid carrier. *J. Biol. Chem.* **276**, 48831–48839
- Llull, D., Garcia, E. and Lopez, R. (2001) Tts, a processive  $\beta$ -glucosyltransferase of *Streptococcus pneumoniae*, directs the synthesis of the branched type 37 capsular polysaccharide in *Pneumococcus* and other Gram-positive species. *J. Biol. Chem.* **276**, 21053–21061
- Whitfield, C. (1995) Biosynthesis of lipopolysaccharide O-antigens. *Trends Microbiol.* **3**, 178–185
- Jiang, S. M., Wang, L. and Reeves, P. R. (2001) Molecular characterization of *Streptococcus pneumoniae* Type 4, 6B, 8, and 18C capsular polysaccharide gene clusters. *Infect. Immun.* **69**, 1244–1255
- Kolkman, M. A. B., Morrison, D. A., van der Zeijst, B. A. M. and Nuijten, P. J. M. (1996) The capsule polysaccharide synthesis locus of *Streptococcus pneumoniae* serotype 14: identification of the glycosyltransferase gene *cps14E*. *J. Bacteriol.* **178**, 3736–3741
- Kolkman, M. A., van der Zeijst, B. A. M. and Nuijten, P. J. M. (1997) Functional analysis of glycosyltransferase encoded by the capsular polysaccharide biosynthesis locus of *Streptococcus pneumoniae* serotype 14. *J. Biol. Chem.* **272**, 19502–19508
- Kolkman, M. A. B., Wakarchuk, W., Nuijten, P. J. M. and van der Zeijst, B. A. M. (1997) Capsular polysaccharide synthesis in *Streptococcus pneumoniae* serotype 14: molecular analysis of the complete *cps* locus and identification of genes encoding glycosyltransferases required for biosynthesis of the tetrasaccharide subunit. *Mol. Microbiol.* **26**, 197–208
- Jones, J. K. N. and Perry, M. B. (1957) The structure of the type VIII *Pneumococcus* specific polysaccharide. *J. Am. Chem. Soc.* **79**, 2787–2793
- O'Neil, M. A., Selvendran, R. R. and Morris, V. J. (1993) Structure of the acidic extracellular gelling polysaccharide produced by *Pseudomonas elodea*. *Carbohydr. Res.* **124**, 123–133
- Kuo, M. S., Mort, A. J. and Dell, A. (1986) Identification and location of L-glycerate, an unusual substituent in gellan gum. *Carbohydr. Res.* **56**, 173–183
- Muñoz, R., Mollerach, M., Lopez, R. and Garcia, E. (1999) Characterization of type 8 capsular gene cluster of *Streptococcus pneumoniae*. *J. Bacteriol.* **181**, 6214–6219
- Garcia, E. and Lopez, R. (1997) Molecular biology of the capsular genes of *Streptococcus pneumoniae*. *FEMS Microbiol. Lett.* **149**, 1–10
- Reeves, P. R., Hobbs, M., Valvano, M., Skurnik, M., Whitfield, C., Coplin, D., Kido, N., Klana, J., Maskell, D., Raetz, C. and Rick, P. (1996) Bacterial polysaccharide synthesis and gene nomenclature. *Trends Microbiol.* **4**, 495–503
- Pollock, T. J., van Workum, W. A., Thorne, L., Mikolajczak, M. J., Yamazaki, M., Kijne, J. W. and Armentrout, R. W. (1998) Assignment of biochemical functions to glycosyl transferase genes which are essential for biosynthesis of exopolysaccharides in *Sphingomonas* strain S88 and *Rhizobium leguminosarum*. *J. Bacteriol.* **180**, 586–593
- Videira, P., Fialho, A., Geremia, R. A., Breton, C. and Correia, I. (2001) Biochemical characterization of the (1  $\rightarrow$  4)- $\beta$ -glucuronosyltransferase GelK in the gellan gum-producing strain *Sphingomonas paucimobilis* A.T.C.C. 31461. *Biochem. J.* **358**, 457–464
- Yamazaki, M., Thorne, L., Mikolajczak, M., Armentrout, R. W. and Pollock, T. J. (1996) Linkage of genes essential for synthesis of a polysaccharide capsule in *Sphingomonas* strain S88. *J. Bacteriol.* **178**, 2676–2687

- 24 Pelosi, L., Boumedienne, M., Saksouk, N., Geiselman, J. and Geremia, R. A. (2004) The glucosyl-1-phosphate transferase WchA (Cap8E) primes the capsular polysaccharide repeat unit biosynthesis of *Streptococcus pneumoniae* serotype 8. *Biochem. Biophys. Res. Commun.* **327**, 857–865
- 25 Campbell, J. A., Davies, G. J., Bulone, V. and Henriissat, B. (1997) A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. *Biochem. J.* **326**, 929–939
- 26 Miller, J. H. (1972) *Experiments In Molecular Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor
- 27 Bohringer, J., Fischer, D., Mosler, G. and Hengge-Aronis, R. (1995) UDP-glucose is a potential intracellular signal molecule in the control of expression of  $\sigma^S$  and  $\sigma^S$ -dependent genes in *Escherichia coli*. *J. Bacteriol.* **177**, 413–422
- 28 Videira, P. A., Cortes, L. L., Fialho, A. M. and Correia, I. (2000) Identification of the *pgmG* gene, encoding a bifunctional protein with phosphoglucomutase and phosphomannomutase activities, in the gellan gum-producing strain *Sphingomonas paucimobilis* A.T.C.C. 31461. *Appl. Environ. Microbiol.* **66**, 2252–2258
- 29 Matsudaira, P. (1987) Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. *J. Biol. Chem.* **262**, 10035–10038
- 30 Tolmasky, M. E., Staneloni, R. J. and Leloir, L. F. (1982) Lipid-bound saccharides in *Rhizobium meliloti*. *J. Biol. Chem.* **257**, 6751–6757
- 31 Lellouch, A. C. and Geremia, R. A. (1999) Expression and study of recombinant ExoM, a (1  $\rightarrow$  4)- $\beta$ -D-glucosyltransferase involved in succinoglycan biosynthesis in *Sinorhizobium meliloti*. *J. Bacteriol.* **181**, 1141–1148
- 32 Cserzo, M., Wallin, E., Simon, I., von Heijne, G. and Elofsson, A. (1997) Prediction of transmembrane  $\alpha$ -helices in prokaryotic membrane proteins: the dense alignment surface method. *Protein Eng.* **10**, 673–676
- 33 Mengin-Lecreux, D., Texier, L., Rousseau, M. and van Heijenoort, J. (1991) The murG gene of *Escherichia coli* code for the UDP-N-acetylglucosamine: N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase involved in the membrane steps of peptidoglycan synthesis. *J. Bacteriol.* **173**, 4625–4636
- 34 Geremia, R. A., Roux, M., Ferreiro, D. U., Dauphin-Dubois, R., Lellouch, A. C. and Ielpi, L. (1999) Expression and biochemical characterisation of recombinant AceA, a bacterial  $\alpha$ -mannosyltransferase. *Mol. Gen. Genet.* **261**, 933–940
- 35 Frey, P. A. (1996) The Leloir pathway: a mechanistic imperative for three enzymes to change the stereochemical configuration of a single carbon in galactose. *FASEB J.* **10**, 461–470
- 36 van Selm, S., Kolkman, M. A. B., van der Zeijst, B. A. M., Zwaagstra, K. A., Gastra, W. and van Putten, J. P. M. (2002) Organization and characterization of the capsule biosynthesis locus of *Streptococcus pneumoniae* serotype 9V. *Microbiology.* **148**, 1747–1755
- 37 Chaffin, D. O., Beres, S. B., Yim, H. H. and Rubens, C. E. (2000) The serotype of type Ia and II group B Streptococci is determined by the polymerase gene within the polycistronic capsule operon. *J. Bacteriol.* **182**, 4466–4477

Received 4 February 2005; accepted 15 March 2005

Published as BJ Immediate Publication 15 March 2005, DOI 10.1042/BJ20050217



## The glucosyl-1-phosphate transferase WchA (Cap8E) primes the capsular polysaccharide repeat unit biosynthesis of *Streptococcus pneumoniae* serotype 8

Ludovic Pelosi\*, Manel Boumedienne<sup>1</sup>, Nehmé Saksouk, Johannes Geiselmann, Roberto A. Geremia<sup>2</sup>

Laboratoire Adaptation et Pathogénie des Micro-organismes (LAPM), CNRS UMR 5163, Institut Jean Roget, Faculté de Médecine Pharmacie, Domaine de La Merci, 38700 La Tronche, France

Received 9 December 2004  
Available online 30 December 2004

### Abstract

The gene *wchA* (*cap8E*) belongs to the *cps8* locus that is involved in biosynthesis of the capsular polysaccharide (CPS) repeat unit (RU) of the virulent *Streptococcus pneumoniae* serotype 8. We report here the biochemical characterization of the membrane-associated protein WchA (Cap8E), overexpressed in *Escherichia coli* BL21(DE3)/pLysS. Our results demonstrate that the recombinant enzyme transfers in vitro a glucosyl-1-phosphate from UDP-glucose to an endogenous phosphoryl-polyprenol, thereby priming the RU biosynthetic pathway of *S. pneumoniae* CPS 8. We also show that the C-terminal half of WchA is the glycosyltransferase domain as observed for the galactosyl-1-phosphate transferase WbaP from *Salmonella enterica*, previously described to prime the first step of O-antigen biosynthesis. These results demonstrate that WchA plays a prominent function in the capsule biosynthesis and explain the key role it occupies in the pneumococcal capsule variation.

© 2004 Elsevier Inc. All rights reserved.

**Keywords:** WchA; Cap8E; Glucosyl-1P-transferase; Capsular polysaccharide; Biosynthesis; *Streptococcus pneumoniae*; Serotype 8; Pathogen bacteria

*Streptococcus pneumoniae* (pneumococcus) is an important human pathogen that causes bacterial infections such as pneumonia, bacteremia, meningitis, and otitis media resulting in high mortality rates despite treatment with antimicrobial agents. The capsular polysaccharides (CPS), the major pathogenicity determinant of *S. pneumoniae*, are responsible for pneumococcal virulence [1]. Many of the vaccines that have been developed provide protection by eliciting antibodies against pneumococcal CPS. The efficiency of these vaccines is

however limited [1]. At present, at least 90 different serotypes of *S. pneumoniae* are known, mainly based on antigenic differences in their CPS structures and this number may further increase though genetics exchanges between serotypes [2].

CPS of *S. pneumoniae* are polymers composed of repeating oligosaccharide units (RU) joined by glycosidic linkages. Heterogeneity among capsules is established through variations of the type and linkage of the different monosaccharides constitutive of the RU [3]. CPS synthesis requires a complex pathway, including biosynthesis of the activated monosaccharide nucleotide sugar donors, sequential transfer of each sugar onto a lipid carrier to form the RU and subsequent polymerization of the RU, export, and complete attachment of the CPS to the cell surface [3].

\* Corresponding author. Fax: +33 476635663.

E-mail address: [ludovic.pelosi@ujf-grenoble.fr](mailto:ludovic.pelosi@ujf-grenoble.fr) (L. Pelosi).

<sup>1</sup> Present address: UMR 144-Institut Curie, 12 Rue Lhomond, 75248 Paris Cedex 05, France.

<sup>2</sup> Present address: LECA, CNRS UMR 5553, University Joseph Fourier BP 53, Grenoble Cedex 9, France.

For a number of pneumococcal serotypes, the genes involved in the CPS biosynthesis pathway have been sequenced (serotypes 1, 2, 3, 4, 6B, 8, 9V, 14, 15B, 18C, 19A, 19B, 19C, 19F, 23F, and 33F) and all except one serotype (serotype 37) are clustered within the chromosome between the *dexB* and *aliA* loci [4–21]. A bioinformatics analysis suggested that each *cps* locus groups all of the genes required for CPS biosynthesis and its regulation [21]. Despite the large available number of both sequences and chemical structures of CPS, the functions of only few of these genes have been clearly identified [10,22,23]. To better understand the CPS biosynthetic pathway and the evolution of the serotypes, we need to identify the function of each enzyme encoded by the *cps* loci.

We are currently studying the capsule biosynthesis of the virulent *S. pneumoniae* serotype 8 included in the actual vaccine [24]. CPS 8 consists of a linear backbone of [ $\rightarrow$  4)- $\alpha$ -D-Glcp-(1  $\rightarrow$  4)- $\alpha$ -D-Galp-(1  $\rightarrow$  4)- $\beta$ -D-GlcAp-(1  $\rightarrow$  4)- $\beta$ -D-Glcp-(1  $\rightarrow$  ) based on the cellobiuronic acid unit ( $\beta$ -D-GlcAp-(1  $\rightarrow$  4)-D-Glcp) also present in serotype 3 [25]. The chromosome region of serotype 8 involved in CPS synthesis has been characterized successively by Jian et al. [8] and Muñoz et al. [9]. Among the 12 ORF deduced from the sequence, five, *wchA*, *wciQ*, *wciR*, *wciS*, and *wciT* (according to the Bacterial Polysaccharide Gene Database (BPGD) nomenclature system [26]), code for proteins sharing similarities with glycosyltransferases (GTs) involved in the RU assembly of gram-positive CPS [8,9].

In this work, we provide biochemical evidence that the membrane-associated recombinant protein WchA (Cap8E) (EC 2.7.8.6), coded by the gene *wchA* (*cap8E*), corresponds to an UDP-Glc: phosphoryl-polyprenol (PL) Glc-1-phosphate (Glc-1P)-transferase. The C-terminal half of WchA carries the glucosyltransferase function. Our results: (i) show that WchA is involved in the first step of the RU serotype 8 biosynthetic pathway and

(ii) provide a better understanding of its key role in the pneumococcal capsule phase variation [27].

## Materials and methods

**Material, bacterial strains, plasmids, and media.** The strains and plasmids used in this work are listed in Table 1. *Escherichia coli* strains were grown in Luria–Bertani broth (LB) at 37 °C with orbital agitation. When required, antibiotics were used at the following concentrations: ampicillin (0.1 mg/L), chloramphenicol (0.03 mg/L), and kanamycin (0.1 mg/L). *S. pneumoniae* serotype 8 chromosomal DNA was from the A.T.C.C (American Type Culture Collection) strains. The pET29b expression vector carrying the *gelK* gene (pET-*gelK*) was a gift of Dr. I. Sa-Correia (Instituto Superior Técnico, Lisboa, Portugal). Restriction enzymes and DNA polymerase were from Fermentas. Oligonucleotide primers for PCR were purchased from Eurogentec, Seraing, Belgium. UDP-[U-<sup>14</sup>C]Gal (267 mCi/mmol) and UDP-[U-<sup>14</sup>C]Glc (303 mCi/mmol) were from Amersham/Pharmacia and UDP-[U-<sup>14</sup>C]GlcA (313 mCi/mmol) was from ICI Biochemicals.

**Cloning of the *wchA* and *wchAt* genes.** The *wchA* gene and its truncated form (called *wchAt*) were amplified from *S. pneumoniae* type 8 chromosomal DNA by PCR using the Expand High Fidelity PCR system (Roche), using the oligonucleotide primers *wchA*-for (5'-AGGGATACATATGAATGGAAAAGTAG-TAAATTCTTCATTGGC C-3') and *wchAt*-for (5'-TTATAAGCATATGACCATGATTACG AATTTTATTATTGATATCGTGGGTGC-3'), the engineered *NdeI* site is underlined; *wchA*-rev (5'-ATTCTCGAGCTTCGTCATCCTTCATCAATAACAACCTTTAACT-3') and *wchAt*-rev (5'-ATTCTCGAGCTTCGTCATCCTTCATCAATAACAACCTTTAACTGTCTT CAA-3'), the engineered *XhoI* site is underlined. For each construction, the PCR was optimized and carried out in 100  $\mu$ l of a mixture containing: approximately 100 ng of *S. pneumoniae* type 8 chromosomal DNA, 200  $\mu$ M of each deoxynucleoside triphosphate (Promega), 1  $\mu$ M of each primer, and 5 U *Taq* DNA polymerase according to the recommendations of the manufacturer (Roche). The predicted PCR products were cloned into the *NdeI*–*XhoI* digested pET23a<sup>+</sup> expression vector (Novagen) and sequenced (Genome Express, Meylan, France). The resulting plasmids were called pET-*wchA* and pET-*wchAt*.

**Expression and partial purification of recombinant WchA and WchAt.** *E. coli* BL21(DE3) pLysS (Novagen) cells harboring the plasmids pET-*wchA* and pET-*wchAt* or the cloning vector pET23a<sup>+</sup> were grown at 37 °C in 500 mL LB. At an OD<sub>600 nm</sub> of 0.6, protein expression was induced by the addition of IPTG to 0.6 mM. The bacterial culture was incubated at 37 °C for an additional 2 h.

Table 1  
Bacterial strains and plasmids used in this works

Strains or plasmids	Genotype and relevant characteristics	Sources or references
<i>E. coli</i>		
XL1-Blue	<i>recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac</i> [F <sup>+</sup> <i>proAB locI</i> <sup>q</sup> ZM15::Tn10(Tet <sup>r</sup> )]	Stratagene
BL21(DE3)/pLysS	F <sup>-</sup> <i>ompT hsdSB</i> ( $r_{\beta}^{-}$ m $_{\beta}^{-}$ ) <i>gal dcm</i> (DE3) pLysS(Cm <sup>r</sup> )	Novagen
BL21(DE3)	F <sup>-</sup> <i>ompT hsdSB</i> ( $r_{\beta}^{-}$ m $_{\beta}^{-}$ ) <i>gal dcm</i> (DE3)	Novagen
C43(DE3)	F <sup>-</sup> <i>ompT hsdSB</i> ( $r_{\beta}^{-}$ m $_{\beta}^{-}$ ) <i>gal dcm</i> (DE3) C43	[28]
<i>Plasmids</i>		
pCRII-topo	<i>ColE1</i> , P <sub>T7</sub> , RBS, M13R, M13F, Kan <sup>r</sup> , Amp <sup>r</sup> , P <sub>Lac</sub>	Invitrogen
pET23a <sup>+</sup>	<i>ColE1</i> , P <sub>T7</sub> , RBS, Amp <sup>r</sup> , His <sub>6</sub>	Novagen
pET29a	<i>ColE1</i> , P <sub>T7</sub> , RBS, Kan <sup>r</sup> , S-tag	Novagen
PGelK	pET29a containing <i>Bam</i> HI– <i>Eco</i> RI <i>gelK</i> PCR fragment from <i>S. paucimobilis</i> A.T.C.C. 31461 genomic bank	[29]
PWchA	pET23a <sup>+</sup> containing <i>NdeI</i> – <i>XhoI</i> <i>wchA</i> PCR fragment from <i>S. pneumoniae</i> type 8 chromosomal DNA	This work
pWchAt	pET23a <sup>+</sup> containing <i>NdeI</i> – <i>XhoI</i> <i>wchAt</i> PCR fragment from <i>S. pneumoniae</i> type 8 chromosomal DNA	This work

Cells were harvested by centrifugation (5000g), washed once with 70 mM Tris–HCl, pH 8.2, and resuspended in 70 mM Tris–HCl, pH 8.2, 300 mM NaCl, 10 mM EDTA, and 10% saccharose (w/v) at 40 OD<sub>600</sub> equivalent/mL. The resulting suspension was lysed by two passages through a French press (SLM-Amico Instruments, Thermo Spectronic) at 80 MPa. Cell debris and inclusion bodies were pelleted by centrifugation for 10 min at 12,000g (12 K) at 4 °C (Sigma 6K15, Rotor No. 12166) and the supernatant was further centrifuged at 50,000g (50 K) (Beckman Avanti J-30I, rotor JA-30.50) for 1 h at 4 °C. The recovered insoluble and soluble fractions were stored at –20 °C or directly used for GT assays.

**Protein quantification, SDS–PAGE, Western blot, and amino acid sequencing.** The protein concentrations were determined using a Bradford protein assay kit (Bio-Rad) and BSA as a standard. The expression and differential centrifugation of WchA and WchAt were monitored by SDS–PAGE (12.5% acrylamide-bis [29:1], stained with Coomassie brilliant blue R250). Amersham/Pharmacia prestained protein standards were used for molecular weight estimation. Proteins were electro-transferred from SDS–PAGE gels onto PVDF membranes (Amersham/Pharmacia) as described previously [30]. Immunoblot analysis was performed with the Promega His-Tag Western-blotting kit, using alkaline phosphatase His-protein conjugate according to the manufacturer's instructions. The blots were developed with the NBT/BCIP (Sigma/Aldrich) system. N-terminal sequencing of recombinant proteins was directly performed from the PVDF Western blot [30].

**Preparation of EDTA-permeabilized bacterial cells.** *E. coli* BL21 (DE3) cells carrying either pET23a<sup>+</sup>, pET-*wchA* or pET-*wchAt* were grown in 100 mL LB at 37 °C, induced with 0.6 mM IPTG, and harvested by centrifugation after 2 h. EDTA-permeabilized bacterial cells were obtained using the method described by Tomalsky et al. [31], slightly modified as described below. Bacterial cells were washed in 70 mM Tris–HCl, pH 8.2, resuspended in the same buffer containing 10 mM EDTA, and submitted to three freeze–thaw cycles in liquid nitrogen. Permeabilized cells were collected by centrifugation, resuspended in the same buffer at 200 OD<sub>600 nm</sub> equivalents/mL and finally stored as 300 µL aliquots at –80 °C.

**Glucosyltransferase assay.** Glucosyltransferase assays were performed in 200 µL containing 70 mM Tris–HCl, pH 8.2, 1% Triton X-100, 12.5 mM MgCl<sub>2</sub>, and 0.10 µCi UDP-[U-<sup>14</sup>C]Glc, the reaction being started by the addition of recombinant proteins (approximately 30–300 µg of total proteins). Assays were done in triplicate and each experiment was reproduced three times. After 1 h at 37 °C, the reaction was stopped by the addition of 100 µL water and 200 µL chloroform/methanol (1:1, v/v). Glycolipids were isolated by solvent extraction as described previously [29]. The radioactivity incorporated into an aliquot of the samples was measured in 4 mL liquid scintillation cocktail, using a Packard Tri-Carb 1600 Tr counter. The oligosaccharide moiety synthesized in vitro was released by mild acid hydrolysis. The glycolipid fraction was dried in a Speed-Vac at 65 °C and resuspended in 50 µL *n*-butanol. To this solution, we added 50 µL of 0.05 M trifluoroacetic acid and the solution was heated for 20 min at 90 °C. After desalting over a mixed bed resin TMD-8 (Sigma–Aldrich), the pellet was dried (in a Speed-Vac at 65 °C), resuspended in 40% 2-propanol in water, and submitted to analytical TLC (silica gel 60, 0.25 mm; Aldrich). Glc, Gal, GlcA, maltose, and cellobiose were used as standards. The TLC plate was developed in 1-butanol/ethanol/water (5:3:2 v/v/v). Unlabeled sugar standards on the TLC plate were revealed by charring (100 °C for 10 min) after treatment with 5% H<sub>2</sub>SO<sub>4</sub> in ethanol. The remaining portion of the plate containing the radiolabeled oligosaccharides was analyzed with a phosphorimager Fujifilm FLA-8000 after two days of exposure.

**Glucuronosyltransferase assay.** The radiolabeled glycolipids ([U-<sup>14</sup>C]D-GlcP-PPL) synthesized in vitro by WchA and WchAt from EDTA-permeabilized bacterial cells were recovered by extraction with a chloroform/methanol/water (1:2:0.3, v/v/v) mixture. Extracts were dried in a Speed-Vac to a final volume of 10 µL (5000–10000 cpm) and stored at –20 °C until use. Glucuronosyltransferase (GelK) assays

were performed in a final volume of 200 µL containing: 10 µL of radiolabeled glycolipid acceptor, 70 mM Tris–HCl, pH 8.2, 8.5 mM MgCl<sub>2</sub>, 1 mM ATP, and 0.12 µCi UDP-[U-<sup>14</sup>C]GlcA [30]. The reaction was started by the addition of French press cell extracts (approximately 300 µg) prepared from IPTG-induction *E. coli* BL21(DE3) culture carrying pET-*gelK* as previously described by Videira et al. [29]. The radiolabeled cellobiuronic acid moiety (D-GlcAp-(1 → 4)-β-[U-<sup>14</sup>C]D-GlcP) synthesized in vitro in a GelK-dependent pathway was then analyzed by TLC as explained above.

## Results and discussion

### Computational analysis of the sequence and cloning of *WchA*

WchA-like proteins are separated into two classes, both of which possess monosaccharide transferase activity [32]. The sequence of the *wchA* gene product has a high level of identity (between 70% and 95%) with the pneumococcal enzymes Cps14E (class I) and Cps9v (Class II) proposed to be UDP-Glc: PL Glc-1P-transferases [10,22]. We therefore suggest the same function for WchA. All of these enzymes are similar to WbaP, an UDP-Gal: PL Gal-1P-transferase from *Salmonella enterica*, priming the first step of O-antigen synthesis [33]. WchA-like proteins differ from most other GTs by a unique hydrophobic profile with, respectively, four and one transmembrane segments in their N- and C-terminal domains as predicted by the DAS method (Fig. 1) [34]. It has been suggested that WbaP is a bifunctional protein with a GT function in the C-terminal half [33]. Furthermore, the work of Wang et al. [33] has shown that the basal GT activity of WbaP was dependent on the membrane association of the enzyme. On the basis of all these results, the *S. pneumoniae wchA* gene was cloned into the expression vector pET23a<sup>+</sup> by PCR amplification guided by previously published nucleotide sequences [8,9]. The resulting recombinant protein, WchA, is predicted to possess 463 amino acids (53.8 kDa) including a C-terminal His-tag epitope. To complete the functional characterization of WchA, and from the results obtained with *S. enterica* enzyme, we have constructed a truncated version of WchA, called WchAt, limited to the 203 C-terminal amino acids (23.3 kDa) and comprising the putative transmembrane helix numbered 5, as shown in Fig. 1.

### Expression and subcellular localization of *WchA* and *WchAt* in *E. coli*

Overexpression in *E. coli* BL21(DE3)/pLysS of both recombinant proteins WchA and WchAt was monitored by SDS–PAGE and Western blot (Fig. 2). After 2 h of IPTG-induction, new bands of approximately 55 and 25 kDa were detected by protein staining (Fig. 2A, lanes 5 and 6) and immunodetection using an antibody against the His-Tag epitope (Fig. 2B, lanes 5 and 6) in

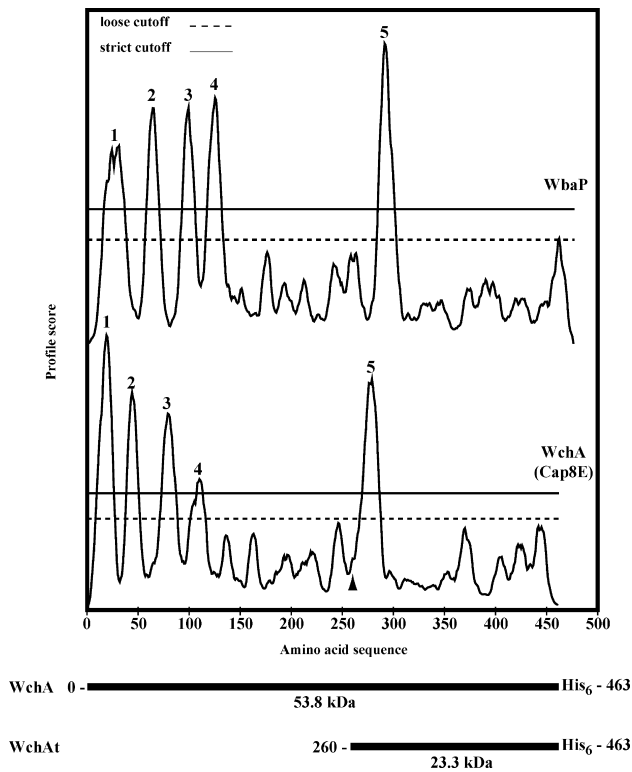


Fig. 1. Transmembrane segment profile alignment obtained from sequences of WchA/Cap8E (with the C-terminal His-Tag epitope) and the UDP-Gal: PL Gal-1P-transferase WbaP from *S. enterica* [33]. The predicted profiles were generated using the DAS method [34]. Two cutoffs are indicated on the plot: the strict cutoff (represented by the solid line) informs about the number of matching segments. The loose cutoff (represented by the dashed line) gives the actual location of the transmembrane segments numbered from 1 to 5. Horizontal lines below the profiles illustrate the amino acid sequences of both recombinant proteins: WchA (coded by the gene *wchA*) and WchAt (coded by the 3'-half of *wchA*, see the arrowhead).

cells harboring pET-*wchA* and pET-*wchAt*. These bands were absent from cells transformed with the cloning vector pET23a<sup>+</sup> (without insert) used as a control in this work (Fig. 2, lane 4). The induced polypeptides have approximately the predicted size and their identity was confirmed by N-terminal amino-acid sequencing (data not shown). The amount of protein appeared to increase up to 3 h and then decreased significantly in overnight cultures (data not shown). The signal observed in Fig. 2B, lane 3 by immunodetection and before the IPTG-induction corresponds to the basal level of WchAt expression in *E. coli* BL21(DE3)/pLysS. We have observed that WchAt was produced in higher amount than WchA, probably due to the lack of the 4 predicted transmembrane helices in the N-terminal domain (Fig. 1). Even with other *E. coli* strains such as C43 (DE3), which is optimized for expression of membrane-associated proteins [28], we were unable to produce larger amount of WchA (data not shown). In all subsequent experiments, the recombinant protein samples were prepared from *E. coli* BL21 cultures after 2 h of IPTG induction.

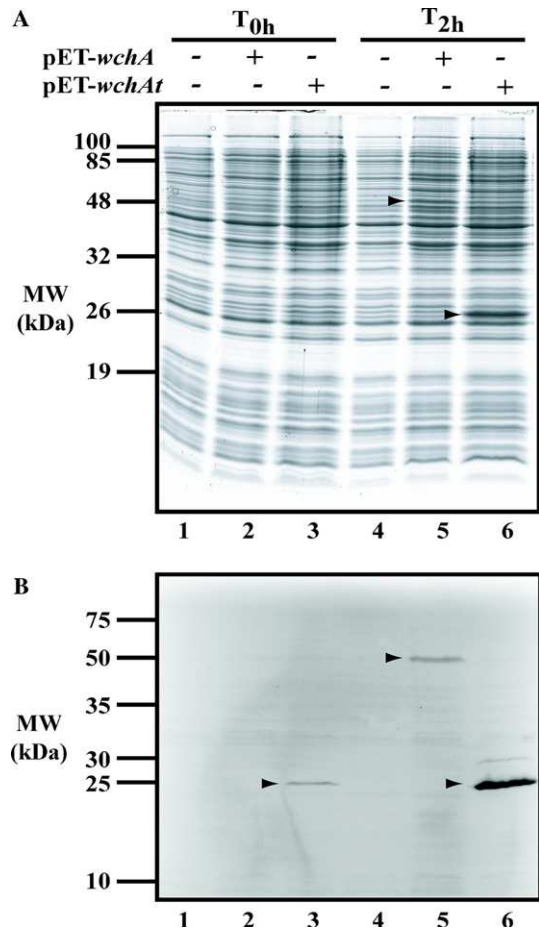


Fig. 2. Expression of WchA and WchAt in *E. coli* BL21(DE3)/pLysS harboring pET23a<sup>+</sup> with (+) or without (-) the insert *wchA* and *wchAt* were cultured as described in Materials and methods section. Aliquots were removed before (lanes 1–3) or after 2 h (lanes 4–6) of IPTG induction. (A) Analysis of each aliquot by Coomassie blue-staining SDS-12.5% PAGE. Twenty micrograms of proteins was loaded. (B) His-tag Western blot with alkaline phosphatase His-protein conjugate of each aliquot. The overexpressed WchA and WchAt polypeptides are indicated by arrowheads. The anti-His-Tag signal detected in lane 3 is due to the leaky expression of T7 polymerase, which leads to a basal expression of WchAt in the absence of IPTG. MW, molecular weight.

We assessed the subcellular localization of WchA and WchAt in *E. coli* from crude cell extracts obtained with a French press and submitted to differential centrifugation. The strain harboring pET23a<sup>+</sup> was used as a control. The content of each fraction was analyzed by SDS-PAGE (Fig. 3A) and the identity of each polypeptide was confirmed by an anti-His-Tag Western blot (Fig. 3B). A significant amount of recombinant proteins was recovered in the 12 K-pellets, suggesting that a fraction of WchA and WchAt may form inclusion bodies (Figs. 3A and B, lanes 7–9), as had been observed for other bacterial GTs overexpressed in *E. coli* [29,35]. The proteins recovered in the 12 K-supernatant were centrifuged at 50 K. Both recombinant proteins were found in the pellet, indicating their membrane localization in *E. coli* (Figs. 3A and B, lanes 13–15). Interestingly, the single



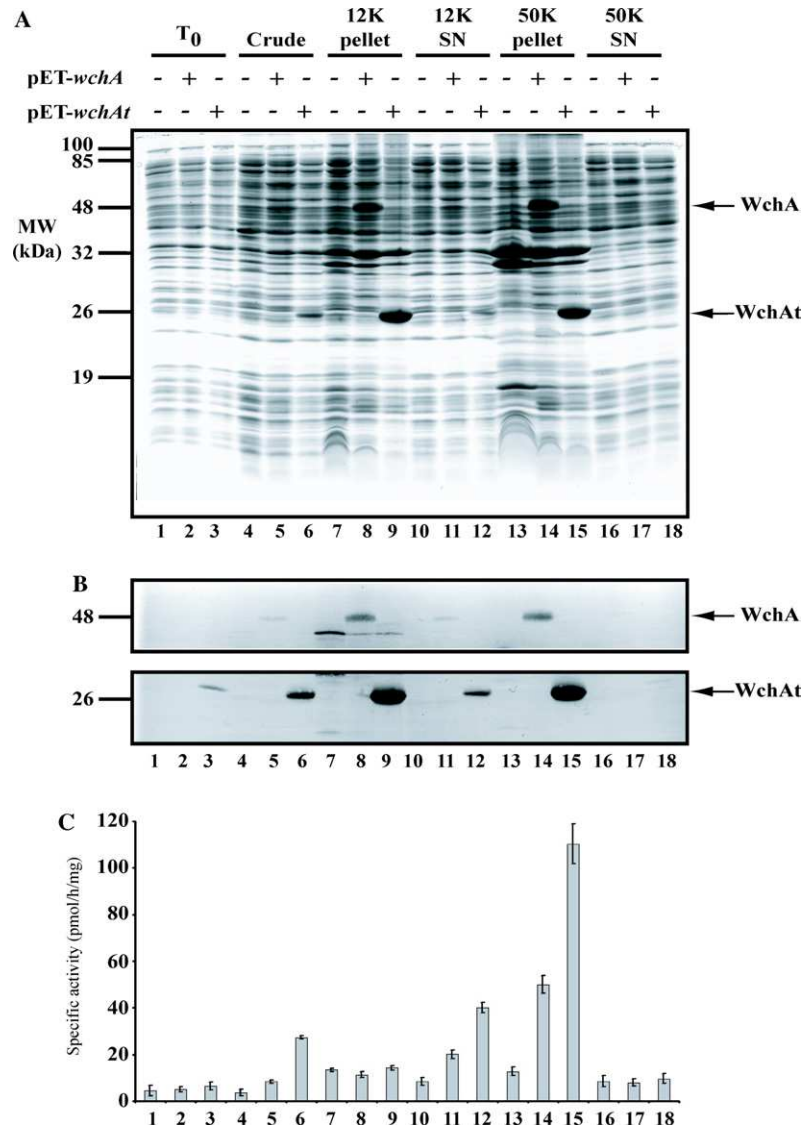


Fig. 3. Membrane localization of WchA and WchAt. *E. coli* BL21(DE3)/pLysS harboring pET23a<sup>+</sup> with (+) or without (-) the insert *wchA* and *wchAt* were grown and cells were disrupted as described in Materials and methods. Aliquots were removed before (lanes 1–3) or after 2 h (lanes 4–18) of IPTG induction. The crude fractions obtained (lanes 4 and 6) were subjected to sequential 12,000g (12 K) (lanes 7–12) and 50,000g (50 K) (lanes 13–18) centrifugations. Electrophoretic profiles of the pellets (lanes 7–9 and 13–15) and the corresponding supernatants (SN) (lanes 10–12 and 16–18) are shown. (A) Analysis of each aliquot by a Coomassie blue-stained SDS-PAGE, 12.5% acrylamide-bis. Twenty micrograms of proteins was loaded in each lane. (B) His-tag Western blot with alkaline phosphatase His-protein conjugate of each aliquot. The recombinant polypeptides are indicated by arrowheads. (C) Glucosyltransferase enzymatic assays of each aliquot. The specific activity corresponding to 100  $\mu$ L GT preparation (approximately 30  $\mu$ g of total proteins) was expressed as pmol of [U-<sup>14</sup>C]Glc incorporated per hour and per mg of proteins from UDP-[U-<sup>14</sup>C]Glc into the *E. coli* glycolipid fraction. SN, supernatant; MW, molecular weight.

predicted transmembrane helix seems sufficient for localizing WchAt in the membrane. This result was in agreement with the previous data obtained with WbaP [33] and confirmed the experimental possibility to express Gram positive GTs in *E. coli* [36,37].

#### *wchA* And *wchAt* encode an UDP-Glc: PL Glc-1P-transferase

The *wchA* gene of CPS type 8 encodes a protein with high sequence and hydrophobicity profile similarities

with the *cps14E* gene product from *S. pneumoniae* serotype 14. Genetics approaches indicated that this enzyme corresponds to an UDP-Glc: PL Glc-1P-transferase [22]. The RU of serotype 8 CPS contains two Glc residues [25], but based on structural homology between serotype 8 and 14 (which contains only one Glc residue in the RU) we postulate that WchA may possess the same enzymatic activity as Cps14E.

To confirm the function of *wchA* as a Glc-1P-transferase priming the capsule RU biosynthesis of serotype 8, i.e., the transfer of Glc residue to the lipid carrier,

we measured the enzymatic activity of the recombinant proteins. EDTA-permeabilized cells expressing WchA or WchAt were incubated in the presence of UDP-[U-<sup>14</sup>C]Glc. As a control, the substrates, UDP-[U-<sup>14</sup>C]Gal and UDP-[U-<sup>14</sup>C]GlcA, were also tested in the same conditions. The synthesis of labeled intermediates was assessed after extraction of the glycolipid fraction. The results showed that a basal level of radioactivity is detected when *E. coli* BL21 (DE3)/pLysS harboring pET23a<sup>+</sup> (vector control) was incubated with each radiolabeled UDP-sugar (Table 2). This weak activity is most likely due to endogenous GTs. As observed in Table 2, neither UDP-[U-<sup>14</sup>C]Gal nor UDP-[U-<sup>14</sup>C]GlcA was significantly incorporated as a substrate. This means that these monosaccharides are probably not the first monosaccharides to be added in the RU. As expected, the [U-<sup>14</sup>C]Glc moiety was incorporated into the organic fraction specifically in the presence of EDTA-permeabilized cells harboring pET-*wchA* or pET-*wchAt* (Table 2). WchAt presented a higher specific activity than WchA, probably due to a better expression level observed for the truncated protein (see above). Similar results were obtained from membrane extracts prepared from *E. coli* cultures expressing the recombinant enzymes. Indeed, we have observed that the glucosyltransferase activity copurified with WchA and WchAt polypeptides submitted to differential centrifugation (Fig. 3C). In these conditions, WchAt presented an approximately twofold higher specific activity than WchA in the membrane fractions (Fig. 3C, lanes 14 and 15). Interestingly, a low activity found in the 12 K-pellet corresponding to the inclusion body fraction indicated that part of the overexpressed enzymes was inactive (Fig. 3C, lanes 8 and 9). Despite numerous attempts at using mild detergents, such as CHAPS, *n*-octyl-β-D-glucoside, Mega-10 or Zwittergent ZW 3-12, tested at or above their critical micellar concentration, we were unable to extract and solubilize from the membrane an active form of either enzyme (data not shown).

These results suggest that the lipid environment is vital for the enzymatic activity of WchA, as had been observed for other intrinsic GTs such as the well-studied glycan synthases [38,39].

Analysis by thin layer chromatography (TLC) of the in vitro synthesized intermediates after mild acid hydrolysis (Fig. 4) demonstrated that both recombinant enzymes catalyzed the transfer of [U-<sup>14</sup>C]Glc residue (see spot I, Fig. 4) to an *E. coli* lipid carrier, characterized previously in *Salmonella* as a PL compound [40]. The product synthesized by our Glc-1P-transferases was

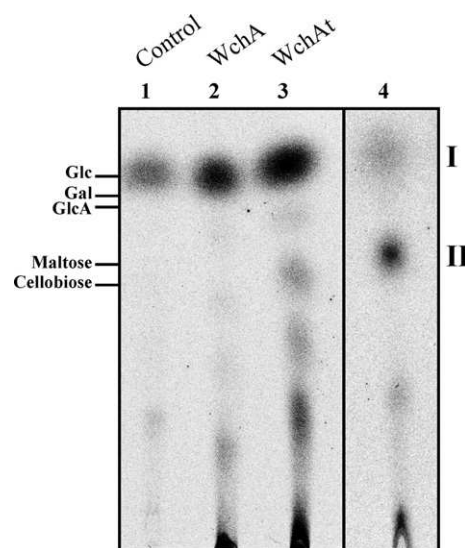


Fig. 4. TLC analysis of radiolabeled oligo (I) and monosaccharides (II) released by mild acid hydrolysis of the glycolipid fractions. Lane 1–3, prepared from the EDTA-permeabilized cells harboring pET23a<sup>+</sup> (used as a control vector), pET-*wchA* and pET-*wchAt* after WchA and WchAt-mediated transfer of UDP-[U-<sup>14</sup>C]Glc to unlabeled endogenous PL acceptor. Lane 4, after GelK-mediated transfer of UDP-GlcA to the radiolabeled glycolipid acceptor synthesized in vitro in a WchA-dependent pathway and expected to be [U-<sup>14</sup>C]D-GlcP-PPL. Glc, Gal, GlcA, maltose, and cellobiose were used as standards.

Table 2

Incorporation of [U-<sup>14</sup>C]sugar mediated by the recombinant proteins, WchA and WchAt, into permeabilized *E. coli* BL21 (DE3)/pLyS cells harboring pET23a<sup>+</sup>, pET-*wchA* or pET-*wchAt*

Plasmid surexpressed	Sugar donor			[U- <sup>14</sup> C]sugar incorporated into the <i>E. coli</i> glycolipid fraction (pmol/h/mg)
	UDP-[U- <sup>14</sup> C]Glc	UDP-[U- <sup>14</sup> C]GlcA	UDP-[U- <sup>14</sup> C]Gal	
pET23a <sup>+</sup>	+	–	–	27 ± 3
pET- <i>wchA</i>	+	–	–	64 ± 4
pET- <i>wchAt</i>	+	–	–	277 ± 8
pET23a <sup>+</sup>	–	+	–	6 ± 4
pET- <i>wchA</i>	–	+	–	9 ± 3
pET- <i>wchAt</i>	–	+	–	7 ± 4
pET23a <sup>+</sup>	–	–	+	30 ± 2
pET- <i>wchA</i>	–	–	+	38 ± 5
pET- <i>wchAt</i>	–	–	+	27 ± 4

Note. The activity corresponding to 100 μL GT preparation (approximately 300 μg total protein) is expressed as pmol of [U-<sup>14</sup>C]sugar incorporated per hour from UDP-[U-<sup>14</sup>C]sugar into the *E. coli* glycolipid fractions.

therefore expected to be a D-Glcp-PPL. Some other minor radioactive spots were also visible on the TLC pattern (Fig. 4). These minor products are probably due to the presence of active endogenous GTs in the preparation. To confirm our results, we used the ability of the (1 → 4)-β-D-glucuronosyltransferase, GelK, to use as an acceptor the product synthesized in vitro by WchA and WchAt. The corresponding recombinant GT (see Table 1) is able to synthesize the cellobiuronic acid-PLL (β-D-GlcAp-(1 → 4)-D-Glcp-pyrophosphoryl-poly-prenol) from UDP-GlcA and its acceptor, thought to be D-Glcp-PPL [29]. The radiolabeled glycolipidic product synthesized in vitro in a WchA-dependent manner was incubated in the presence of the GelK protein extract prepared as previously described [29] and UDP-GlcA. The corresponding radiolabeled oligosaccharide synthesized in vitro by GelK was released from the glycolipid fraction by mild acid hydrolysis and analyzed by TLC (Fig. 4, lane 4). Two radioactive spots were detected by phosphoimaging. The major oligosaccharide formed (spot II) co-migrated with the maltobiose marker and was thus assigned to β-D-GlcAp-(1 → 4)-[U-<sup>14</sup>C]D-Glcp-PPL [29]. As described previously by Videira et al. [29], no GlcA transfer was observed in the absence of the donor (UDP-GlcA) or in the presence of an enzymatic extract prepared by French press lysis of an *E. coli* BL21 (DE3) strain harboring the control vector pET29a (data not shown). The remaining Spot I represents the unmodified monosaccharide used as a GelK acceptor (Fig. 4, lane 4). A similar pattern was also obtained with the glycolipid synthesized in vitro by WchAt (data not shown). Taken together, these results confirm the structure of the product synthesized in vitro by WchA and WchAt as a D-Glcp-PPL.

Taken together, these results clearly show that WchA is an UDP-Glc: PL Glc-1P-transferase. We also have demonstrated that the C-terminal part of WchA displays the GT function. Indeed, as observed in a diverse range of GT families, both prokaryotic and eukaryotic, the C-terminal region of WchA contains a conserved catalytic motif DXD (residues 342 and 344) preceded by three hydrophobic amino acids (MFXA) [41]. Another putative catalytic region, characterized by a pair of acid residues, Glu and Asp (residues 385 and 437), previously reported to be part of the catalytic site in inverting GTs, is also present in the WchA sequence [42]. Furthermore, the fact that the 25 kDa recombinant polypeptide encoded by the 3' half of *wchA* possesses GT activity suggests that the primary interaction with the PL occurs in the C-terminal part of this protein. As observed for the Gal-1P-transferase WbaP, a single transmembrane segment in this domain (helix numbered 5, see Fig. 1), is enough to preserve enzymatic activity and the membrane location of WchA or WchAt was vital for both enzymes most likely due to the lipidic nature of their acceptor. All these data agree with previous

results obtained from WchA-like proteins, such as Cps9vE, Cps14E, CpsIII E, CpsE, and WbaP, that have been characterized in several bacterial species using molecular and/or biochemical approaches [10,22,33,37,43]. It had been proposed that all these enzymes catalyze, in their respective organism, the first step of exopolysaccharide biosynthesis by means of their Glc-1P-transferase or Gal-1P-transferase activities. The functional characterization of WchA further illustrates the key role that this family of enzymes plays in pneumococcal infections. In fact, it has been recently demonstrated that the tandem duplication of the 5' part of the pneumococcal gene *wchA* confers to serotype 8 the ability to “flip” between capsular and acapsular phenotypes, thereby profiting from the advantages of each phenotype, as is the case also for *Neisseria meningitidis* [44]. We can corroborate and confirm that the introduction of a premature stop codon downstream of the duplicated sequence of *wchA* leads to the synthesis of a C-terminally truncated protein without the GT domain, thus inhibiting the biosynthesis of the CPS 8. As proposed by Waite et al. [27], this phenomenon may help pneumococci to progress to an invasive disease beyond simple carriage.

Finally, the *Cps8* locus contains five ORFs (WchA, WciQ, WciR, WciS, and WciT) coding for hypothetical proteins sharing sequence similarity with nonprocessive GTs involved in RU assembly of Gram-positive bacterial CPS. On the basis of their putative functions and current knowledge about capsule biosynthesis, the following sequence of reactions can be proposed for the biosynthesis of CPS serotype 8. First, as suggested for other pneumococcal serotypes, such as 9V and 14, a Glc residue is attached to a lipid carrier, probably a PL [10,22]. This reaction is catalyzed by WchA as we show in this work. After formation of the D-Glc-PPL intermediate, the four additional GTs, WciQ, WciR, WciS, and WciT sequentially add monosaccharide to the growing RU. Recently, we have shown that WciS catalyzes in vitro the transfer of a Gal residue to the cellobiuronic acid-PPL (work in progress), suggesting that WciS is the best candidate, of the five GTs of *cps8* locus, to add the third sugar residue. The RU of serotype 8 contains two Glc residues. Based on our results about the biological functions of WchA and WciS, we propose also that the reaction catalyzed by WchA probably involves the Glc residue that will be linked to the GlcA residue. On the basis of sequence similarities and the observation that genes in *cps* loci of *S. pneumoniae* seem to be organized in the order of assembly of the RU, we suggest that the putative heterodimer WciQ/WciR and WciT exhibit, respectively (1 → 4)-β-D-glucuronosyltransferase and (1 → 4)-α-D-glucosyltransferase activities. Using a biochemical approach, our data confirm the predicted function of the GTs from *cps8* locus [8,9]. A detailed molecular understanding of the different

pneumococcal GTs will eventually permit the synthesis in vitro of the entire glycolipid RU serotype 8. Such an approach will allow us to tackle the elucidation of the still unknown polymerization step and the biochemical mechanism of the processive GT.

## Acknowledgments

We thank Dr. Jean Gagon for his critical reading of the manuscript and his useful suggestions. The N-terminal amino acid sequence of recombinant proteins was determined by J.-P. Andrieu, Institut de Biologie Structurale (IBS), Grenoble, France. L. Pelosi and N. Saksouk are supported by post-doctoral (CNRS) and doctoral (CNRS) fellowships, respectively. This work was supported by CNRS (France) grants.

## References

- [1] E. Alonsodevelasco, A.F.M. Verheul, J. Verhoef, H. Snippe, *Streptococcus pneumoniae*: virulence factors, pathogenesis, and vaccines, *Microbiol. Rev.* 59 (1995) 591–603.
- [2] T.J. Coffey, M.C. Enright, M. Daniels, J.K. Morona, R. Morona, W. Hryniewicz, J.C. Paton, B.G. Spratt, Recombinational exchanges at the capsular polysaccharide biosynthetic locus lead to frequent serotype changes among natural isolates of *Streptococcus pneumoniae*, *Mol. Microbiol.* 27 (1998) 73–83.
- [3] C. Whitfield, Biosynthesis of lipopolysaccharide O-antigens, *Trends Microbiol.* 3 (1995) 178–185.
- [4] R.M. Munoz, R. Mollerach, R. Lopez, E. Garcia, Molecular organization of the genes required for the synthesis of type 1 capsular polysaccharide of *Streptococcus pneumoniae*: formation of binary encapsulated pneumococci and identification of cryptic dTDP-rhamnose biosynthesis genes, *Mol. Microbiol.* 25 (1997) 79–92.
- [5] F.B. Iannelli, J. Pearce, G. Pozzi, The type 2 capsule locus of *Streptococcus pneumoniae*, *J. Bacteriol.* 181 (1999) 2652–2654.
- [6] J.P. Dillard, M.W. Vandersea, J. Yother, Characterization of the cassette containing genes for type 3 capsular polysaccharide biosynthesis in *Streptococcus pneumoniae*, *J. Exp. Med.* 181 (1995) 973–983.
- [7] C.E. Arrecubieta, E. Garcia, R. Lopez, Sequence and transcriptional analysis of a DNA region involved in the production of capsular polysaccharide in *Streptococcus pneumoniae* type 3, *Gene* 167 (1995) 1–7.
- [8] S.M. Jiang, L. Wang, P.R. Reeves, Molecular characterization of *Streptococcus pneumoniae* Type 4, 6B, 8, and 18C capsular polysaccharide gene clusters, *Infect. Immun.* 69 (2001) 1244–1255.
- [9] R. Muñoz, M. Mollerach, R. Lopez, E. Garcia, Characterization of type 8 capsular gene cluster of *Streptococcus pneumoniae*, *J. Bacteriol.* 181 (1999) 6214–6219.
- [10] S. van Selm, M.A.B. Kolkman, B.A.M. van der Zeijst, K.A. Zwaagstra, W. Gaastra, J.P.M. van Putten, Organization and characterization of the capsule biosynthesis locus of *Streptococcus pneumoniae* serotype 9V, *Microbiology* 148 (2002) 1747–1755.
- [11] M.A. Kolkman, B.A.M. van der Zeijst, P.J.M. Nuijten, Functional analysis of glycosyltransferase encoded by the capsular polysaccharide biosynthesis locus of *Streptococcus pneumoniae* serotype 14, *J. Biol. Chem.* 272 (1997) 19502–19508.
- [12] M.A.B. Kolkman, W. Wakarchuk, P.J.M. Nuijten, B.A.M. van der Zeijst, Capsular polysaccharide synthesis in *Streptococcus pneumoniae* serotype 14: molecular analysis of the complete *cps* locus and identification of genes encoding glycosyltransferases required for biosynthesis of the tetrasaccharide subunit, *Mol. Microbiol.* 26 (1997) 197–208.
- [13] S. van Selm, L.M. van Cann, M.A.B. Kolkman, B.A.M. van der Zeijst, J.P.M. van Putten, Genetic basis for the structural difference between *Streptococcus pneumoniae* serotype 15B and 15C capsular polysaccharide, *Infect. Immun.* 71 (2003) 6192–6198.
- [14] J.K. Morona, R. Morona, J.C. Paton, Molecular and genetic characterization of the capsule biosynthesis locus of *Streptococcus pneumoniae* type 19B, *J. Bacteriol.* 179 (1997) 4953–4958.
- [15] J.K. Morona, R. Morona, J.C. Paton, Characterization of the locus encoding the *Streptococcus pneumoniae* type 19F capsular polysaccharide biosynthetic pathway, *Mol. Microbiol.* 23 (1997) 751–763.
- [16] J.K. Morona, R. Morona, J.C. Paton, Comparative genetics of capsular polysaccharide biosynthesis in *Streptococcus pneumoniae* types belonging to serogroup 19, *J. Bacteriol.* 181 (1999) 5355–5364.
- [17] M. Ramirez, A. Tomasz, Molecular characterization of the complete 23F capsular polysaccharide locus of *Streptococcus pneumoniae*, *J. Bacteriol.* 180 (1998) 5273–5278.
- [18] J.K. Morona, D.C. Miller, T.J. Coffey, C.J. Vindurampulle, B.G. Spratt, R. Morona, J.C. Paton, Molecular and genetic characterization of the capsule biosynthesis locus of *Streptococcus pneumoniae* type 23F, *Microbiology* 145 (1999) 781–789.
- [19] D. Llull, R. Lopez, E. Garcia, R. Munoz, Molecular structure of the gene cluster responsible for the synthesis of the polysaccharide capsule of *Streptococcus pneumoniae* type 33F, *Biochim. Biophys. Acta* 1443 (1998) 217–224.
- [20] D. Llull, R. Lopez, E. Garcia, R. Munoz, A single gene (*tts*) located outside the *cap* locus directs the formation of *Streptococcus pneumoniae* type 37 capsular polysaccharide. Type 37 pneumococci are natural, genetically binary strains, *J. Exp. Med.* 190 (1999) 241–251.
- [21] E. Garcia, R. Lopez, Molecular biology of the capsular genes of *Streptococcus pneumoniae*, *FEMS Microbiol. Lett.* 149 (1997) 1–10.
- [22] M.A.B. Kolkman, D.A. Morrison, B.A.M. van der Zeijst, P.J.M. Nuijten, The capsule polysaccharide synthesis locus of *Streptococcus pneumoniae* serotype 14: identification of the glycosyltransferase gene *cps14E*, *J. Bacteriol.* 178 (1996) 3736–3741.
- [23] M.A. Kolkman, B.A. van der Zeijst, P.J. Nuijten, Diversity of capsular polysaccharide synthesis gene clusters in *Streptococcus pneumoniae*, *J. Biochem. (Tokyo)* 123 (1998) 937–945.
- [24] J.B. Robbins, R. Austrian, C.J. Lee, S.C. Rastogi, G. Schiffman, J. Henrichsen, P.H. Mäkelä, C.V. Broome, R.R. Fracklam, R.H. Tiesjema, J.C.J. Parke, Considerations for formulating the second-generation pneumococcal capsular polysaccharide vaccine with emphasis on the cross-reactive types within groups, *J. Infect. Dis.* 148 (1983) 1136–1159.
- [25] J.K.N. Jones, M.B. Perry, The structure of the type VIII pneumococcus specific polysaccharide, *J. Am. Chem. Soc.* 79 (1957) 2787–2793.
- [26] P.R. Reeves, M. Hobbs, M. Valvano, M. Skurnik, C. Whitfield, D. Coplin, N. Kido, J. Klana, D. Maskell, C. Raetz, P. Rick, Bacterial polysaccharide synthesis and gene nomenclature, *Trends Microbiol.* 4 (1996) 495–503.
- [27] R.D. Waite, D.W. Penfold, J.K. Struthers, C.G. Dowson, Spontaneous sequence duplications within genes *cap8E* and *tts* control phase variation in *Streptococcus pneumoniae* serotypes 8 and 37, *Microbiology* 149 (2003) 497–504.
- [28] I. Arechaga, B. Miroux, S. Karrasch, R. Huijbregts, B. de Kruijff, M.J. Runswick, J.E. Walker, Characterization of new intracellular membranes in *Escherichia coli* accompanying large scale overproduction of the b subunit of F<sub>1</sub>F<sub>0</sub> ATP synthase, *FEBS Lett.* 482 (2000) 215–219.



- [29] P. Videira, A. Fialho, R.A. Geremia, C. Breton, I. Correia, Biochemical characterization of the (1 → 4)- $\beta$ -glucuronosyltransferase GelK in the gellan gum-producing strain *Sphingomonas paucimobilis* A.T.C.C. 31461, *Biochem. J.* 358 (2001) 457–464.
- [30] P. Matsudaira, Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes, *J. Biol. Chem.* 262 (1987) 10035–10038.
- [31] M.E. Tolmasky, R.J. Staneloni, L.F. Leloir, Lipid-bound saccharides in *Rhizobium meliloti*, *J. Biol. Chem.* 257 (1982) 6751–6757.
- [32] J.K. Morona, R. Morona, J.C. Paton, Analysis of the 5' portion of the type 19A capsule locus identifies two classes of *cpsC*, *cpsD* and *cpsE* genes in *Streptococcus pneumoniae*, *J. Bacteriol.* 181 (1999) 3599–3605.
- [33] L. Wang, D. Liu, P.R. Reeves, C-terminal half of *Salmonella enterica* WbaP (RfbP) is the galactosyl-1-phosphate transferase domain catalyzing the first step of O-antigen synthesis, *J. Bacteriol.* 178 (1996) 2598–2604.
- [34] M. Cserzo, E. Wallin, I. Simon, G. von Heijne, A. Elofsson, Prediction of transmembrane  $\alpha$ -helices in prokaryotic membrane proteins: the Dense Alignment Surface method, *Protein Eng.* 10 (1997) 673–676.
- [35] A.C. Lellouch, R.A. Geremia, Expression and study of recombinant ExoM, a (1 → 4)- $\beta$ -D-glucosyltransferase involved in succinoglycan biosynthesis in *Sinorhizobium meliloti*, *J. Bacteriol.* 181 (1999) 1141–1148.
- [36] R.T. Cartee, W.T. Forsee, J.W. Jensen, J. Yother, Expression of the *Streptococcus pneumoniae* type 3 synthase in *Escherichia coli*. Assembly of the type 3 polysaccharide on a lipid carrier, *J. Biol. Chem.* 276 (2001) 48831–48839.
- [37] E. Almiron-Roig, F. Mulholland, M.J. Gasson, A.M. Griffin, The complete *cps* cluster from *Streptococcus thermophilus* NCFB 2393 involved in the biosynthesis of a new exopolysaccharide, *Microbiology* 146 (2000) 2793–2802.
- [38] J.L. Him, L. Pelosi, H. Chanzy, J.L. Putaux, V. Bulone, Biosynthesis of (1 → 3)- $\beta$ -D-glucan (callose) by detergent extracts of a microsomal fraction from *Arabidopsis thaliana*, *Eur. J. Biochem.* 268 (2001) 4628–4638.
- [39] J. Lai-Kee-Him, H. Chanzy, M. Muller, J.L. Putaux, T. Imai, V. Bulone, In vitro versus in vivo cellulose microfibrils from plant primary wall synthases: structural differences, *J. Biol. Chem.* 277 (2002) 36931–36939.
- [40] A. Wright, M. Dankert, P. Fennessey, P.W. Robbins, Characterization of a polyisoprenoid compound functional in O-antigen biosynthesis, *Proc. Natl. Acad. Sci. USA* 57 (1967) 1798–1803.
- [41] C. Wiggins, S. Munro, Activity of the yeast *MNN1*  $\alpha$ -1,3-mannosyltransferase requires a motif conserved in many other families of glycosyltransferases, *Proc. Natl. Acad. Sci. USA* 95 (1998) 7945–7950.
- [42] I. Saxena, M. Brown, M. Fevre, R. Geremia, B. Henrissat, Multidomain architecture of  $\beta$ -glycosyl transferases: implications for mechanism of action, *J. Bacteriol.* 177 (1995) 1419–1424.
- [43] D.O. Chaffin, S.B. Beres, H.H. Yim, C.E. Rubens, The serotype of type Ia and II group B *Streptococci* is determined by the polymerase gene within the polycistronic capsule operon, *J. Bacteriol.* 182 (2000) 4466–4477.
- [44] S. Hammerschmidt, A. Muller, H. Sillmann, M. Muhlenhoff, R. Borrow, A. Fox, J. van Putten, W.D. Zollinger, R. Gerardy-Schahn, M. Frosch, Capsule phase variation in *Neisseria meningitidis* serogroup B by slipped-strand mispairing in the polysialyltransferase gene (*siaD*): correlation with bacterial invasion and the outbreak of meningococcal disease, *Mol. Microbiol.* 20 (1996) 1211–1220.